



Universidad
de Alcalá

ESCUELA POLITÉCNICA SUPERIOR

Tesis Doctoral

**DISEÑO DE DETECTORES ROBUSTOS
EN APLICACIONES RADAR**

Autor: David Anastasio de la Mata Moya

Directora: María del Pilar Jarabo Amores

Año 2012

Agradecimientos

Sentir que aún queda tiempo para intentarlo, para cambiar tu destino. Necesito saber que queda más tiempo para seguir aprendiendo cosas, viviendo experiencias y sentir que no hay nada escrito. De hecho, he pensado en tirar la toalla muchas veces y al final ha llegado el momento de dar las gracias a mucha gente. Ellos saben quiénes son y, aunque se lo diré personalmente a cada uno de ellos, quiero dejarles por escrito mi gratitud.

Lo primero es gracias a la Tesis, a lo que significa. Para mí ha sido un proceso complicado que se podría resumir en una prueba de superación que me ha permitido continuar mi desarrollo profesional y aumentar la confianza en mí mismo. Aún así, queda mucho por hacer porque esto es un punto y seguido.

Principalmente se trata de un hito profesional así que en ese orden seguiré con los agradecimientos,

Gracias a mis mentores. Gracias por enseñarme, por valorarme y por confiar en mí. Me siento muy orgulloso de trabajar con vosotros y sólo espero no defraudaros.

Gracias a la gente del fondo de pasillo. Gracias por hacerme más llevadero el día a día, por nuestros desayunos y, principalmente, porque todos habéis participado en esta Tesis.

Gracias a los compañeros. Gracias porque todos me han ayudado a hacer mi trabajo cada día un poco mejor.

Y ahora la parte personal,

Gracias a mi familia. Gracias por todo. Sé que todos me habéis apoyado en este camino muy de cerca aunque a veces la distancia haya sido grande. Y seguramente el tiempo que he dedicado a la Tesis os lo he robado a vosotros. El tiempo no se devuelve, se vive. Espero pasar más tiempo juntos, pero, de momento, quiero compartir toda la alegría de este momento con vosotros y dedicaros esta Tesis.

Gracias a mis amigas (son mayoría y algunas muy especiales). Gracias por no hablarme de trabajo, por tener vida social, por compartir momentos inolvidables,... Como ya sabéis el éxito de uno, es un éxito compartido (ya lo celebraremos).

Resumen

El problema de la detección automática de blancos radar puede ser formulado como un test de hipótesis binaria, en el que el sistema tiene que decir a favor de la hipótesis alternativa H_1 (blanco presente) o de la hipótesis nula H_0 (blanco ausente). El criterio de Neyman-Pearson, NP, es el más extendido en aplicaciones radar. Este detector trata de maximizar la probabilidad de detección, P_D , manteniendo la probabilidad de falsa alarma, P_{FA} , igual o inferior a un valor determinado. Cuando las funciones de verosimilitud son conocidas, una posible implementación del detector NP consiste en comparar el cociente de verosimilitud con un umbral fijado por los requisitos de P_{FA} (LRT). Se trata de un detector paramétrico que puede presentar grandes pérdidas de detección cuando las características estadísticas del blanco y/o interferencia asumidas en el diseño difieren de las reales.

En situaciones prácticas, los parámetros de la interferencia pueden estimarse a partir de medidas obtenidas en el entorno del radar, pero las propiedades del blanco pueden ser difíciles de estimar. Por lo que, para el diseño de detectores, se asume diferentes modelos de blanco cuyos parámetros, como su coeficiente de correlación o su frecuencia Doppler, son variables aleatorias con funciones de densidad de probabilidad conocidas. En estos casos, el problema de la detección se plantea como un test de hipótesis compuesto y, una regla de decisión basada en el cociente de verosimilitud promediado (ALR) es una posible implementación del detector NP. Esta realización requiere la resolución de integrales muy complejas que pueden hasta no tener una solución cerrada y se proponen soluciones sub-óptimas basadas en técnicas de integración numérica y otras aproximaciones numéricas.

En esta Tesis Doctoral, se aborda el diseño de detectores basados en inteligencia artificial como solución alternativa para la detección de blancos con parámetros desconocidos en diferentes entornos de clutter. En [Jarabo2009], se demuestra la capacidad de aproximar el detector NP utilizando sistemas adaptativos entrenados de manera supervisada para minimizar la función de coste adecuada. En [Jarabo2005a, Jarabo2009], se calcula la función aproximada por agentes inteligentes, como los perceptrones multicapa (MLP), redes neuronales con funciones de base radial (RBFNN) y redes neuronales de segundo orden (SONN), entrenados con el error cuadrático medio o la entropía. En esta Tesis, el estudio teórico presentado en [Jarabo2009] ha sido extendido para tests de hipótesis compuestos, confirmando que la condición suficiente puede ser aplicada para probar si un sistema adaptativo entrenado de manera supervisada con una función de error adecuada es capaz de aproximar el detector NP para cualquier par de funciones de verosimilitud. Otra contribución importante de la Tesis, es el estudio teórico de la función aproximada por una Máquina de Vectores Soporte (SVM) cuando en el entrenamiento se utiliza la función de error definida en [Shawe2004]. Se trata de una contribución importante en este campo, porque

aporta claves importantes para explicar, desde el punto de vista teórico, las limitaciones de las prestaciones de las C-SVM y 2C-SVM en diferentes aplicaciones de detección presentados en la literatura.

Como esta Tesis se enmarca en proyectos financiados por el Ministerio de Ciencia e Innovación, la Comunidad de Madrid, la Universidad de Alcalá y la empresa AMPER SISTEMAS, S.A. centrados en aplicaciones de radares marinos, se han estudiado distintos modelos de clutter marino. Estos modelos se han utilizado para generar datos sintéticos para entrenar, validar y probar las soluciones basadas en inteligencia artificial y simular un escenario radar. Se han considerado tres casos de estudio: Detección de blancos fluctuantes Gaussianos con coeficientes de correlación o pulsación Doppler desconocida en ruido blanco Gaussiano aditivo; detección de blancos fluctuantes Gaussianos con coeficientes de correlación o pulsación Doppler desconocida en clutter Gaussiano correlado más ruido blanco Gaussiano aditivo; y detección de blancos no fluctuantes con pulsación Doppler desconocida en clutter K-distribuido impulsivo. Se ha realizado un estudio de la sensibilidad de los detectores LRT para blancos con parámetros desconocidos para todos los casos y se han diseñado aproximaciones basadas en el cociente de verosimilitud generalizado restringido (CGLR) para ser utilizadas como detectores de referencia para analizar las capacidades de detección y el coste computacional de las soluciones basadas en inteligencia artificial.

Para cada uno de los casos de estudio, se han diseñado y evaluado detectores basados en MLPs, RBFNNs, SONNs y SVMs que presenten un buen compromiso entre capacidad de detección y coste computacional. La propuesta de soluciones basadas en SONNs es otra contribución importante de esta Tesis. Los detectores SONN, con una única unidad neuronal cuadrática, presentan una gran robustez frente al coeficiente de correlación o frecuencia Doppler del blanco en interferencia Gaussiana. También se proponen soluciones basadas en mezclas de expertos para mejorar las capacidades de detección y/o reducir el coste computacional. Se han propuesto diferentes técnicas novedosas de combinación de las salidas de los expertos. Los detectores propuestos han sido, finalmente, evaluados en un escenario radar simulado, cuyos resultados han sido comparados con los obtenidos con técnicas CA-CFAR.

Summary

The automatic radar detection problem can be formulated as a binary hypothesis test, where the system must decide in favour of hypothesis H_1 (target present) or H_0 (target absent). The Neyman-Pearson (NP) criterion is the most extended for this task. This type of detector maximizes the probability of detection (P_D), while maintaining the probability of false alarm (P_{FA}) lower than or equal to a given value. When the likelihood functions are known, a decision rule based on comparing the Likelihood Ratio with a detection threshold fixed attending to P_{FA} requirements (LRT), is an implementation of the NP detector. Due to its parametric nature, if the interference and/or target statistical features assumed in the design differ from real ones, the detection losses can be very important.

In practical situations, the interference parameters can be estimated and tracked from the operating environment in some degree, but target ones are really difficult to estimate. Because of that, different target models are assumed, considering that parameters such as the one-lag correlation coefficient or the Doppler shift are random variables with known probability density functions (PDFs). In these cases, the detection problem must be formulated as a composite hypothesis test, for which the decision rule based on the Average Likelihood Ratio (ALR) is an implementation of the NP detector. This approach usually leads to intractable integrals without a closed-form solution, and sub-optimum solutions based on numerical approximations of the ALR, or the Generalized Likelihood Ratio (GLR), can be proposed.

This PhD. Thesis tackles the study of the suitability of Artificial Intelligence (AI) based radar detectors as alternative solutions for the problem of detecting targets with unknown parameters in different radar clutter environments. Learning machines trained in a supervised manner using a suitable error function, have been previously proved to be able to approximate the optimum NP detector in simple hypothesis tests [Jarabo2009]. The calculus of the function approximated by the learning machine after training is a key element for the analysis of the suitability of an error function for training a learning machine to approximate the NP detector. In [Jarabo2005a, Jarabo2009] the function approximated by a learning system such as a MultiLayer Perceptron (MLP), a Radial Basis Function Neural Network (RBFNN) or a Second Order Neural Network (SONN) when trained using the mean-squared error and the cross-entropy, was calculated. In this PhD Thesis, the theoretical study presented in [Jarabo2009] has been extended to composite hypothesis tests, confirming that the proposed sufficient condition can be applied for testing if an error function is suitable for training learning machines in a supervised manner, for approximating the ALR based detector for any pair of likelihood functions. Another important contribution of this PhD. Thesis, is the theoretical study of the function approximated by a Super Vector Machine (SVM), when the error used for training is the classification error proposed by [Shawe2004].

This is an important contribution in the field, because provides important keys to justify, from theoretical foundations, the observed performances and limitations of C-SVMs and 2C-SVMs in different detection applications presented in the literature.

This PhD. Thesis has been carried out in the scope of research projects such as the Spanish Ministry, the Comunidad de Madrid, the University of Alcalá, and the company AMPER SISTEMAS, S.A., which focused on marine radar applications, radar sea clutter models have been studied and used for generating synthetic data sets for training, validating and testing the proposed AI solutions, and for simulating a radar scenario. Three cases of study have been considered taking into account different clutter models and target unknown parameters: detection of Gaussian fluctuating targets with unknown correlation coefficient or unknown Doppler shift, in presence of Additive White Gaussian Noise (AWGN); detection of Gaussian fluctuating targets with unknown correlation coefficient or unknown Doppler shift, in presence of correlated Gaussian clutter and AWGN; and detection of non fluctuating targets with unknown Doppler shift, in presence of spiky K-distributed clutter. For all these cases, an analysis of the sensitivity of the LRT detectors for targets of known parameters has been carried out in a first step, in order to study the detection loss associated to a mismatch between the assumed design target parameters and the actual ones. Constrained approximations of the GLR test have been designed, to be used as references for the design and performance analysis of the AI proposed solutions, in terms of detection capabilities and computational complexity.

For each case of study, detectors based on MLPs, RBFNNs, SONNs and SVMs have been designed and tested. The proposal of SONN based solutions is another important contribution of the PhD Thesis. SONNs with only one quadratic neural unit present a high robustness against target correlation coefficient and Doppler shift, in Gaussian interference. Mixture of experts are also designed for improving the detection capabilities and/or reducing the computational cost. Different combination techniques have been considered. Some of them, as far as the author knows, are novel contributions. A final AI solution has been proposed as a compromise between detection performance and computational cost for each case of study. These detectors have been finally evaluated in a simulated radar scenario, comparing their performances with those provided with CA-CFARs.

Índice general

I	Estudio Preliminar	1
1.	Antecedentes y Organización de la Tesis	3
1.1.	Introducción	3
1.2.	Detección robusta de blancos con parámetros desconocidos	4
1.2.1.	Aplicación de métodos estadísticos	5
1.2.2.	Aplicación de inteligencia artificial	6
1.3.	Problemas pendientes de resolver	11
1.4.	Estructura de la tesis	12
2.	Modelos de señal radar	15
2.1.	Introducción	15
2.2.	Radar primario de impulsos	15
2.3.	Caracterización de la señal recibida	19
2.3.1.	Modelos de blancos radar	19
2.3.2.	Modelo de ruido en el receptor	23
2.3.3.	Modelos de clutter en el receptor	24
2.3.3.1.	Clutter marino	26
2.3.3.2.	Clutter terrestre	29
2.3.3.3.	Clutter atmosférico	31
2.4.	Casos de estudio	31
3.	Detectores óptimos	33
3.1.	Introducción	33
3.2.	Planteamiento como test de hipótesis binario	33
3.3.	Caracterización del espacio de observaciones	35
3.3.1.	Hipótesis H_0 : blanco ausente	36
3.3.2.	Hipótesis H_1 : blanco presente	41
3.3.3.	Relación señal a interferencia	42
3.4.	Detectores óptimos	43
3.4.1.	Caso de estudio 1: Detector NP de parámetros conocidos	44
3.4.2.	Caso de estudio 2: Detector NP de parámetros conocidos	46
3.4.3.	Caso de estudio 3: Detector NP de parámetros conocidos	50

4. Inteligencia artificial y el detector de Neyman-Pearson	53
4.1. Introducción	53
4.2. Esquema de detección	55
4.3. Condición suficiente para que una función discriminante pueda aproximar el detector de Neyman-Pearson	56
4.4. Función de error: el cuadrático medio	58
4.5. Función de error: la entropía cruzada	59
4.6. Función de error: el error de clasificación	60
4.7. Agentes inteligentes	61
4.7.1. Perceptrones multicapa, MLP	62
4.7.2. Redes neuronales con funciones de base cuadrática (SONN)	65
4.7.3. Redes neuronales con funciones de base radial (RBFNN)	66
4.7.4. Máquinas de vectores soporte	70
II Investigación y Resultados	73
5. Detección de blancos con parámetros desconocidos	75
5.1. Introducción	75
5.2. Estudio de sensibilidad de los detectores óptimos de blancos con parámetros conocidos	76
5.2.1. Estudio de sensibilidad del detector LR para el caso de estudio 1	76
5.2.2. Estudio de sensibilidad del detector LR para el caso de estudio 2	77
5.2.3. Estudio de sensibilidad del detector LR para el caso de estudio 3	80
5.3. Detectores subóptimos para aproximar el ALR	82
5.3.1. Detectores CGLR para el caso de estudio 1	83
5.3.2. Detectores CGLR para el caso de estudio 2	86
5.3.3. Detectores CGLR para el caso de estudio 3	89
5.4. Conclusiones	91
6. Soluciones basadas en agentes inteligentes	95
6.1. Introducción	95
6.2. Análisis y diseño de detectores basados en redes neuronales	96
6.2.1. Perceptrones multicapa, MLPs	96
6.2.2. Redes neuronales con funciones de base de segundo orden, SONN	98
6.2.3. Redes neuronales con funciones de base radial, RBFNN	100
6.2.4. Comités de redes neuronales	104
6.3. Análisis y diseño de detectores basados en SVM	111
6.3.1. Función discriminante aproximada al minimizar el error de clasificación	112
6.3.2. Detectores basados en C-SVMs para blancos Swerling I y II en AWGN	114
6.3.3. SVMs con coste asimétrico	120
6.3.4. Detectores basados en 2C-SVMs para blancos Swerling I y II en AWGN	121
6.4. Conclusiones	123

7. Resultados	127
7.1. Introducción	127
7.2. Bases de datos sintéticos de entrenamiento y simulación	128
7.2.1. Conjuntos de datos sintéticos	128
7.2.2. Salidas deseadas	129
7.2.3. Estimación de las curvas ROC	130
7.3. Detección de blancos gaussianos con coeficiente de correlación desconocido en AWGN	130
7.3.1. Soluciones basadas en redes neuronales	131
7.3.2. Soluciones basadas en SVMs	133
7.4. Detección de blancos gaussianos con pulsación Doppler desconocida en AWGN .	135
7.4.1. Soluciones basadas en redes neuronales	135
7.4.2. Soluciones basadas en SVMs	137
7.5. Detección de blancos gaussianos con coeficiente de correlación desconocido en interferencia gaussiana	138
7.5.1. Soluciones basadas en redes neuronales	139
7.5.2. Soluciones basadas en SVMs	142
7.6. Detección de blancos gaussianos con pulsación Doppler desconocida en interferencia gaussiana	144
7.6.1. Soluciones basadas en redes neuronales	145
7.6.2. Soluciones basadas en SVMs	148
7.7. Detectores de blancos Swerling V con pulsación Doppler desconocida en clutter K-distribuido	151
7.7.1. Fronteras de decisión del detector CGLR	152
7.7.2. Soluciones basadas en redes neuronales	154
7.7.3. Soluciones basadas en SVMs	155
7.8. Detectores robustos aplicados a un escenario radar	158
7.8.1. Escenario radar simulado	159
7.8.2. Detectores CA-CFAR	166
7.8.3. Resultados de detección en el escenario simulado para el caso de estudio 1	168
7.8.4. Resultados de detección en el escenario simulado para el caso de estudio 2	169
7.8.5. Resultados de detección en el escenario simulado para el caso de estudio 3	170
7.9. Conclusiones	174

III Conclusions and Future Research Lines **179**

8. Conclusions and Future Research Lines	181
8.1. Conclusions	181
8.1.1. Regarding the Capability of Intelligent Agents for Approximating the Neyman-Pearson Detector in Composite Hypothesis Test	183
8.1.2. Regarding the Detection of Targets with Unknown Parameters	184
8.1.3. Regarding the Design of Robust AI Detectors	185
8.1.4. Regarding the Application of the Proposed Robust NN Detectors to a Simulated Radar Scenario	188

8.2. Future Research Lines 188

IV Bibliografía **191**
Bibliografía

Índice de figuras

1.1. Solución de una SVM binaria no lineal.	10
2.1. Esquema básico de un radar primario monoestático	17
2.2. Ejemplo de envolvente de la señal recibida	20
2.3. Ejemplo ilustrativo de la recepción de varios ecos de un mismo blanco durante una exploración	21
2.4. Geometría del área iluminada de clutter superficial	25
2.5. Imagen recibida por el radar IPIX situado en Dartmouth, Canadá.	27
2.6. Comparativa de las FDPs de la amplitud de los datos adquiridos con el radar IPIX y diferentes modelos estadísticos [Greco2004].	28
2.7. Características y condiciones de adquisición de datos del radar IPIX [Greco2004].	28
2.8. Funciones de densidad de probabilidad de datos radar correspondientes a diferentes trozos de terreno de una imagen adquirida por el TerraSAR-X [Mata2011b].	29
2.9. Comparativa de las FDPs de la amplitud de los datos adquiridos con el radar Phase One y diferentes modelos estadísticos [Billingsley1999].	30
3.1. Ejemplo de la matriz obtenida a la salida del detector síncrono (se han marcado con una "x" las posiciones en las que hay un blanco)	36
3.2. Espectros de clutter gaussiano para diferentes valores de coeficientes de correlación)	38
3.3. Modelo SIRP para una secuencia compleja correlada no gaussiana	39
3.4. Histogramas de la envolvente de patrones de clutter K entre pulsos de una exploración, azul, y entre pulsos de distintas exploraciones, rojo.	41
3.5. Espectro de un blanco gaussiano para distintas pulsaciones Doppler.	42
3.6. Curvas ROC del LR para el caso de estudio 1. $\Omega = 0$	46
3.7. Curvas ROC del LR para el caso de estudio 1. $SNR = 7$ dB.	47
3.8. Curvas ROC del LR del caso de estudio 2. $CNR = 20$ dB y $\Omega = \pi/2$	48
3.9. Curvas ROC del LR para el caso de estudio 2. $CNR = 20$ dB y $\rho_s = 0,9$	49
3.10. Curvas ROC del LR para el caso de estudio 2. $CNR = 40$ dB y $\rho_c = 0,995$	49
3.11. Curvas ROC del LR para el caso de estudio 3. $\nu = 0,5$ y $\theta_R = \pi/4$	51
4.1. Esquema del detector neuronal propuesto	56
4.2. Detector neuronal coherente.	62
4.3. Esquema de una red multicapa de propagación directa	63
4.4. Modelo de una neurona artificial con función de base lineal	64
4.5. Unidad neuronal cuadrática (QNU)	65

4.6. Modelo de red neuronal con funciones de base radial	66
5.1. Robustez del LR para el caso de estudio 1 con respecto a $DSNR$. $\Omega^d = \Omega^s = 0$ y $\rho_s^d = \rho_s^s$	77
5.2. Robustez del LR para el caso de estudio 1 con respecto a $DSNR$ y Ω^d . $SSNR = 7$ dB.	78
5.3. Robustez del LR para el caso de estudio 1 con respecto a ρ_s^d . $DSNR = SSNR = 7$ dB y $\Omega^d = \Omega^s = 0$	78
5.4. Robustez del LR para el caso de estudio 2 con respecto a $DSIR$. $CNR = 20$ dB y $\Omega^d = \Omega^s = \pi/2$	78
5.5. Robustez del LR para el caso de estudio 2 con respecto a ρ_s^d . $CNR = 20$ dB y $\Omega^d = \Omega^s = \pi/2$. Las líneas continuas se corresponden con $\rho_s^s = 0, 1$, mientras que la discontinuas se corresponden con $\rho_s^s = 0, 9$	79
5.6. Robustez del LR para el caso de estudio 2 con respecto a Ω^d . $CNR = 20$ dB, $\rho_s^d = \rho_s^s = 0, 9$ y $\Omega^s = \pi/2$	79
5.7. Robustez del LR para el caso de estudio 2 asumiendo $CNR = 40$ dB y $\rho_c = 0, 995$	80
5.8. Robustez del LR para el caso de estudio 3 con respecto a $DSCR$. $\theta_R^d = \theta_R^s = \pi/4$ y $\Omega^d = \Omega^s = \pi/2$	81
5.9. Robustez del LR para el caso de estudio 3 con respecto a θ_R^d . $\theta_R^s = \pi/4$ y $\Omega^d = \Omega^s = \pi/2$	81
5.10. Robustez del LR para el caso de estudio 3 con respecto a Ω^d . $\theta_R^d = \theta_R^s = \pi/4$ y $\Omega^s = \pi/2$	82
5.11. Curvas ROC del CGLR para el caso de estudio 1 con $\rho_s^d = \rho_s^s$ variable uniforme en el intervalo $[0, 1]$	84
5.12. Curvas ROC del CGLR para el caso de estudio 1 con $\Omega^d = \Omega^s$ variable uniforme en el intervalo $[0, 2\pi)$	85
5.13. Curvas ROC del CGLR para el caso de estudio 2 con $\rho_s^d = \rho_s^s$ variable uniforme en el intervalo $[0, 1]$. $\Omega^d = \Omega^s = \pi/2$	87
5.14. Curvas ROC del CGLR para el caso de estudio 2 con $\Omega^d = \Omega^s$ variable uniforme en el intervalo $[0, 2\pi)$. $\rho_s^d = \rho_s^s = 1$	88
5.15. Curvas ROC del detector $CGLR_{16}$ para el caso de estudio 3 con $\Omega^d = \Omega^s$ variable uniforme en el intervalo $[0; 2\pi)$. $\theta_R = \pi/4$	90
5.16. Curvas ROC del detector $CGLR_{16}$ para el caso de estudio 3 con $\Omega^d = \Omega^s$ variable uniforme en un intervalo centrado en $\Omega_c = \pi/2$. $DSCR = SSCR$ y $\Omega_R = \pi/4$	90
5.17. Curvas ROC del detector $CGLR_K$ para el caso de estudio 3 con $\Omega^d = \Omega^s$ variable uniforme en un intervalo $[\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\theta_R = \pi/4$	91
6.1. Red neuronal con una capa oculta de M unidades neuronales cuadráticas.	99
6.2. Esquema basado en un perceptrón multicapa equivalente a una red neuronal con una capa oculta de M QNU's.	99
6.3. Esquema de un algoritmo evolutivo.	102
6.4. Comité de expertos de estructura estática con combinación lineal de sus salidas.	105
6.5. Esquema de entrenamiento de un comité de expertos con <i>boosting</i> por filtrado.	106
6.6. Comité de NNs con estructura dinámica: Mezcla de expertos.	107

6.7. Mezcla de expertos basada en combinación por selección de máxima salida.	107
6.8. Mezcla de dos MLPs con selección por máximo: MLP_1 ha sido entrenado con blancos radar con $\rho_s^t \in [0; 0, 5]$ y MLP_2 con blancos radar con $\rho_s^t \in (0, 5; 1]$	108
6.9. Curvas ROC de los detectores CGLR, MLP y mezcla de dos MLPs para blancos gaussianos con $\rho_s \in [0; 1]$ en AWGN. $SNR = 7$ dB.	108
6.10. Mezcla de dos NNs con combinación basada en una OR lógica.	109
6.11. Curvas ROC de expertos (MLPs, RBFNNs y SONNs) pertenecientes de un comité simulados con un blanco con ρ_s conocidos perteneciente al margen de entrenamiento correspondiente. $SNR = 7$ dB.	110
6.12. Estudio de la capacidad de mezclas de expertos basados en MLPs, RBFNNs y SONNs de aproximar el detector óptimo de blancos gaussianos con $\rho_s \in [0; 1]$ en AWGN. $SNR = 7$ dB.	111
6.13. Resultados obtenidos para la regla de decisión en $F_0(\mathbf{z})$, (6.18), para blancos Swerling II en AWGN para tres valores de k . $SNR = 3$ dB.	115
6.14. Resultados obtenidos para la regla de decisión en $F_0(\mathbf{z})$, (6.18), para blancos Swerling II en AWGN para tres valores de k . $SNR = 7$ dB.	116
6.15. Curvas ROC de detectores C-SVM de blancos Swerling II en presencia de AWGN.	117
6.16. P_{FA} vs η_0 de detectores C-SVM de blancos Swerling II en presencia de AWGN.	117
6.17. Resultados obtenidos para la $F_0(\mathbf{z})$ aproximada para blancos Swerling I en AWGN. $SNR = 3$ dB.	118
6.18. Resultados obtenidos para la $F_0(\mathbf{z})$ aproximada para blancos Swerling I en AWGN. $SNR = 7$ dB.	118
6.19. Curvas ROC de detectores C-SVM de blancos Swerling I en presencia de AWGN.	119
6.20. P_{FA} vs η_0 de detectores C-SVM de blancos Swerling I en presencia de AWGN.	119
7.1. Curvas ROC de detectores basados en un único MLP 16/M/1 para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.	131
7.2. Estudio del tamaño de la capa oculta de los MLPs que forman la mezcla de expertos para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.	132
7.3. Curvas ROC de detectores basados en RBFNNs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.	133
7.4. Resumen de soluciones basadas en NNs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.	134
7.5. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.	134
7.6. Curvas ROC de detectores basados en un MLP o en una RBFNN para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.	136
7.7. Curvas ROC de detectores basados en una SONN o en un comité de NNs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.	136
7.8. Resumen de soluciones basadas en NNs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.	137
7.9. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.	138

7.10. Curvas ROC de detectores basados en un MLP y una RBFNN para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$	140
7.11. Curvas ROC de detectores basados en un MLP y una RBFNN para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = DSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$	140
7.12. Curvas ROC de detectores basados en un MLP y una RBFNN para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = DSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$	140
7.13. Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$	141
7.14. Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$	141
7.15. Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$	142
7.16. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$	142
7.17. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$	143
7.18. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$	143
7.19. Curvas ROC de detectores basados en MLPs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$	146
7.20. Curvas ROC de detectores basados en MLPs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$	146
7.21. Curvas ROC de detectores basados en MLPs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$	146
7.22. Curvas ROC de detectores basados en una única RBF 16/M/1 para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $\rho_s^t = \rho_s^s = 1$	147
7.23. Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$	148
7.24. Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$	148
7.25. Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$	149
7.26. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$	149
7.27. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$	150
7.28. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$	150
7.29. Fronteras de decisión formadas por un detector CGLR para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$	153

7.30. Corte bidimensional de las fronteras de decisión formadas por un detector CGLR con $P8$ para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$, $\rho_c = 0$, $DSCR = 9$ dB y $\varphi = \pi/4$	153
7.31. Curvas ROC de detectores RBFNNs 16/M/1 para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$	154
7.32. Curvas ROC de detectores MLPs 16/M/1 para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$	155
7.33. Resumen de soluciones basadas en NNs para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0$, $TSCR = SSCR = 9$ dB y $\varphi = \pi/4$	155
7.34. Resumen de soluciones basadas en NNs para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0,9$, $TSCR = SSCR = -3$ dB y $\varphi = \pi/4$	156
7.35. Fronteras de decisión formadas por un detector SONN 16/152/1 para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$	156
7.36. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0$, $TSCR = 9$ dB y $\varphi = \pi/4$	157
7.37. Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0,9$, $TSCR = -3$ dB y $\varphi = \pi/4$	157
7.38. Emplazamiento del radar y sector de cobertura simulado.	160
7.39. Trayectoria de ida: movimiento rectilíneo uniforme con velocidad constante de 10,7 nudos.	161
7.40. Trayectoria de vuelta: movimiento rectilíneo uniformemente acelerado con velocidad inicial de 25 nudos y aceleración de 0,058 nudos/s.	162
7.41. Modelo escalado a las dimensiones del yate simulado en la trayectoria de vuelta	163
7.42. RCS en dBsm simulada para el yate simulado en la trayectoria de vuelta	163
7.43. Retorno de clutter marino en dBs para clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$	164
7.44. Retorno de clutter marino en dBs para clutter K-distribuido con $\rho_c = 0$ y $\sigma^{(0)} = -48$ dBsm.	165
7.45. Retorno de clutter marino en dBs para clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$ después de corregir los efectos de la propagación y la variación del área de clutter iluminada con la distancia.	165
7.46. Retorno de clutter marino en dBs para clutter K-distribuido con $\rho_c = 0$ y $\sigma^{(0)} = -48$ dBsm después de corregir los efectos de la propagación y la variación del área de clutter iluminada con la distancia.	165
7.47. Intensidad de las muestras generadas para la exploración 20 (en dB) en la que se muestra el detalle de las zonas en las que se encuentran los dos blancos en unidades naturales.	166
7.48. Esquema básico de un detector CA-CFAR.	167
7.49. Aplicación del CA-CFAR en rango sobre la matriz radar.	167
7.50. Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 1.	169
7.51. Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 2 con $CNR = 20$ y $\rho_c = 0,7$	170

- 7.52. Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 2 con $CNR = 20$ y $\rho_c = 0,995$ 171
- 7.53. Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 2 con $CNR = 40$ y $\rho_c = 0,995$ 171
- 7.54. Superposición de pulsos detectados de las 905 exploraciones sin etapa de reducción de falsas alarmas para el caso de estudio 3 con $\rho_c = 0$ 173
- 7.55. Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 3 con $\rho_c = 0,9$. . 173

Parte I

Estudio Preliminar

Capítulo 1

Antecedentes y Organización de la Tesis

1.1. Introducción

Según [IEEE-Std1997] un sistema radar se define como sigue:

Un sistema electromagnético para la detección y localización de objetos que funciona transmitiendo señales electromagnéticas, recibiendo ecos de objetos (blancos) que estén dentro de su volumen de cobertura y extrayendo información de la señal del eco.

Notas: 1. Radar es un acrónimo de “radio detection and ranging”. 2. El equipamiento del radar puede funcionar con el transmisor apagado, como un rastreador pasivo para detectar fuentes que radien en la banda del sistema receptor.

Las dos funciones principales de un radar son, por tanto, la detección de señales reflejadas por objetos deseados (blancos) y la extracción de información relativa al blanco, como su posición, su velocidad o forma.

En sentido amplio, el término *blanco* se aplica a cualquier objeto que interfiere con la onda transmitida y dispersa parte de su energía hacia el sistema radar [IEEE-Std1997]. Más específicamente, se distingue entre *blanco*, objeto que se desea detectar, y “*clutter*”, objetos no deseados que también interceptan la energía del radar y la dispersan. Típicamente se trata de reflexiones del mar, la tierra, la lluvia y otros tipos de precipitación, los pájaros, los insectos, los meteoros, la aurora o las producidas por *chaff*¹. En esta tesis se considera la segunda definición en la que se distingue claramente entre objetos deseados o blancos y objetos no deseados o clutter.

El problema de la detección de blancos radar es un *test de hipótesis binario* en el que H_1 , la hipótesis alternativa, representa la hipótesis de *blanco presente* y H_0 , la hipótesis nula, representa la hipótesis de *blanco ausente*. En ambos casos, la señal recibida es ruidosa debido a fuentes radiantes externas captadas por la antena y al ruido térmico generado por componentes de carácter óhmico de la antena, del receptor y de los distintos medios de transmisión empleados. Esto, unido al hecho de que es muy probable la aparición de blancos no deseados (clutter), hace que el problema de diseño de detectores radar sea muy complicado. Aunque se pueden

¹Segn [IEEE-Std1997], se trata de tiras de metal o un material metalizado que se lanzan en grandes cantidades para simular un blanco o para crear una señal de clutter muy potente que enmascare la detección de un blanco deseado

utilizar varios criterios para resolver este problema, el más apropiado en radar es el criterio de Neyman-Pearson, donde se trata de maximizar la probabilidad de detección, P_D , manteniendo la probabilidad de falsa alarma, P_{FA} , igual o inferior a un valor determinado [Neyman1933]. Se trata de un detector paramétrico que para poderse aplicar deben conocerse las funciones de verosimilitud bajo las hipótesis nula y alternativa [Srinath1996][Aloisio1994] y, por tanto, requiere una caracterización total del blanco y de la interferencia. En muchas situaciones prácticas, las propiedades estadísticas de la interferencia pueden estimarse a partir de medidas obtenidas en el entorno del radar, pero la estimación de las propiedades del blanco no es posible, por lo que para diseñar un detector basado en el criterio de Neyman-Pearson ha de asumirse un determinado modelo de blanco. Dado que es muy probable que los blancos reales tengan propiedades estadísticas diferentes a las del modelo asumido en el diseño, es necesario realizar un estudio de la robustez del detector diseñado ante variaciones de distintos parámetros del blanco, como su sección radar o su frecuencia doppler.

En esta tesis se aborda el problema de la detección de blancos con parámetros desconocidos en presencia de interferencia aditiva con soluciones basadas en inteligencia artificial. Un problema fundamental hasta ahora no resuelto, es la posibilidad de implementar aproximaciones válidas del detector óptimo según el criterio de Neyman-Pearson, donde el problema de la detección se plantea como un test de hipótesis compuesto en el que la función de densidad de probabilidad del blanco es desconocida o no está completamente especificada. Además, en muchas ocasiones el detector óptimo requiere la resolución de integrales muy complejas que pueden hasta no tener una solución cerrada y, entonces, se presentarían soluciones subóptimas mediante técnicas de integración numérica u otras aproximaciones [Nayebi1996], que aunque pueden conllevar costes computacionales muy elevados, nos permiten evaluar las soluciones propuestas.

En este capítulo, se presenta el estado del arte en la detección robusta en aplicaciones radar en presencia de blancos con parámetros desconocidos en el apartado 1.2. De esta forma, se pondrán de manifiesto las tareas susceptibles de ser investigadas en el apartado 1.3. Finalmente, en el apartado 1.4 se presenta la estructura de la tesis.

1.2. Detección robusta de blancos con parámetros desconocidos

Las dos funciones principales de un radar son la detección de señales reflejadas por blancos y la extracción de información relativa a los mismos, como su posición, velocidad o forma. La señal recibida es ruidosa debido a fuentes radiantes externas captadas por la antena y al ruido térmico generado por componentes de carácter óhmico de la antena, del receptor y de los distintos medios de transmisión empleados. Esto, unido al hecho de que es muy probable la aparición de blancos no deseados (clutter), hace que el problema de diseño de detectores radar sea muy complicado.

Uno de los criterios de detección más utilizados en aplicaciones radar es el de Neyman-Pearson. Este detector maximiza la probabilidad de detección, manteniendo la probabilidad de falsa alarma menor o igual a un valor fijado de antemano [Neyman1933]. Una posible realización de este detector consiste en evaluar el cociente de verosimilitud (LR, *Likelihood Ratio*) y compararlo con un umbral fijado atendiendo a los requisitos de probabilidad de falsa alarma. Dado que requiere del conocimiento de las funciones de verosimilitud del problema y, por tanto, de una caracterización total del blanco y la interferencia, se trata de un detector paramétrico. En

muchas situaciones prácticas, las propiedades estadísticas de la interferencia pueden estimarse a partir de medidas obtenidas en el entorno del radar, pero la estimación de las propiedades del blanco no es posible, por lo que para poder construir un detector paramétrico ha de asumirse un determinado modelo de blanco. Dado que es muy probable que los blancos reales tengan propiedades estadísticas diferentes a las del modelo asumido en el diseño, es necesario plantear el estudio de detectores robustos ante variaciones de distintos parámetros del blanco como su sección radar o su frecuencia Doppler.

En el caso de asumir variaciones en parámetros del blanco, el problema de detección de blancos radar se plantea como un test de hipótesis compuesto en el que la función de densidad de probabilidad del blanco es desconocida o no está completamente especificada. Una de las realizaciones del detector de Neyman-Pearson consiste en comparar el cociente de verosimilitud promediado (ALR, *Average Likelihood Ratio*) con un umbral especificado para cumplir los requisitos de probabilidad de falsa alarma [Van1968]. En muchas ocasiones este cociente de verosimilitud requiere la resolución de integrales muy complejas que pueden hasta no tener una solución cerrada, por lo que se plantean aproximaciones basadas en técnicas de integración numérica, estimadores de máxima probabilidad u otras soluciones subóptimas.

1.2.1. Aplicación de métodos estadísticos

L. Brennan, I. Reed, y W. Sollfrey estudian la detección de blancos no fluctuantes con frecuencia Doppler desconocida en presencia de blanco gaussiano aditivo y comparan las características de detección de un detector basado en el ALR con otro basado en un estimador de máxima probabilidad, el cociente de verosimilitud generalizado (GLR, *Generalized Likelihood Ratio*), obteniendo resultados similares en ambos [Brennan1968]. En [Nayebi1996], se continua la línea de detección de blancos no fluctuantes con frecuencia Doppler desconocida en presencia, en este caso, de ruido coloreado. Al plantear el ALR se obtiene una integral que no se puede resolver analíticamente y se proponen diferentes aproximaciones numéricas: el GLR, el GLR considerando un número finito de valores discretos del parámetro desconocido (CGLR, *Constrained Generalized Likelihood Ratio*), desarrollos de Taylor del integrando para valores bajos de relación señal a ruido o sustituir la integral por una suma de los valores obtenidos del integrando para un número finito de valores discretos del parámetro desconocido (CALR, *Constrained Average Likelihood Ratio*). Las aproximaciones basadas en el CGLR y el CALR se pueden implementar de forma práctica porque suponen un número finito de filtros que generalmente depende del número de ecos de cada exploración. En este trabajo se concluye que los detectores basados en el CGLR y el CALR con un número de valores discretos de la frecuencia Doppler igual al doble del número de ecos por exploración aproximan al óptimo, basado en el criterio de NP, para ruido blanco y coloreado; además se deduce que la complejidad computacional es proporcional al cuadrado del número de ecos por exploración. E. Lovli formula un ALR para el caso de detección de blancos Swerling I con frecuencia Doppler desconocida en presencia de ruido blanco gaussiano. Resuelve la integral del ALR y especifica que la estructura del receptor resultante no tiene porqué ser óptima para ruido coloreado. Propone una etapa de preprocesado para blanquear el ruido y poder utilizar el esquema resultante de resolver la integral. La etapa de blanqueamiento es equivalente a un filtro MTI adaptativo que atenúa y decorrela los ecos del clutter a nivel de ruido de manera que resulta que el espectro de la interferencia es blanco.

A. di Vito y M. Naldi estudian, en [diVito1999], la robustez de un detector basado en el cociente de verosimilitud, LR, diseñado para blancos gaussianos en interferencia gaussiana con matrices de autocorrelación conocidas. Los resultados presentan unas pérdidas de detección cuando el coeficiente de correlación y la frecuencia Doppler del blanco no coinciden con las consideradas respectivamente para el diseño, reflejando la sensibilidad de los detectores óptimos a los parámetros de diseño.

Las capacidades de detección de detectores diseñados bajo la suposición de interferencia gaussiana quedan muy mermadas cuando el modelo de señal no deseada, clutter, es no gaussiano. En [Conte1991], [Pentini1992] y [Conte1994], se propone el diseño de detectores basados en el test GLR para la detección de blancos no fluctuantes de amplitud desconocida y fase desconocida en presencia de clutter K distribuido. En [Conte1998] se estudia la sensibilidad de detectores basados en el test GLR para la detección coherente en presencia de clutter Weibull dominante. Contempla el caso de detección de blancos fluctuantes y no fluctuantes con amplitudes y fases desconocidas y se obtienen resultados con las estadísticas del clutter Weibull completamente o parcialmente conocidas. Los resultados muestran que los detectores propuestos son adecuados para eludir la incertidumbre del escenario planteado.

En estos trabajos la caracterización estadística del clutter se basa en la teoría de procesos aleatorios invariantes esféricamente (SIRP, *Spherically Invariant Random Process*) que permite definir de forma independiente la función densidad de probabilidad de la amplitud y la función de covarianza del clutter. Esta teoría asume un modelo compuesto por el producto de un proceso gaussiano, llamado *speckle*, y una variable no negativa que representa la potencia local, llamada *textura*. Una estrategia posible de diseño de detectores subóptimos se puede basar en considerar la *textura* como una variable determinista desconocida en lugar de una variable aleatoria. Este método se utiliza para la detección coherente de blancos en presencia de clutter gaussiano compuesto con una amplitud K distribuida en [Gini1997] y el resultado se aproxima a las curvas ROC del detector óptimo. Continuando con la detección de blancos con parámetros desconocidos en presencia de clutter gaussiano compuesto, en [Conte2000] se propone un detector canónico basado en el GLR cuya estructura es independiente de la función de densidad de probabilidad de la amplitud que aproxima el ALR y mejora los resultados obtenidos con un receptor convencional de ley cuadrática.

1.2.2. Aplicación de inteligencia artificial

En esta tesis se proponen soluciones basadas en inteligencia artificial para el diseño de detectores no paramétricos capaces de aproximar el detector óptimo de Neyman-Pearson en el test de hipótesis compuesto. Se han considerado distintas técnicas de inteligencia artificial con un aprendizaje supervisado para el diseño de detectores robustos que consideran variaciones en los parámetros del blanco y/o en los parámetros de la interferencia.

En [Ruck1990] y [Wan1990], se demuestra la capacidad de las redes neuronales artificiales (NN) de aproximar el clasificador bayesiano óptimo cuando se entrena con el algoritmo de retropropagación del error, utilizando como función de error el error cuadrático medio. En [Jarabo2005a], se demuestra la posibilidad de aproximar el criterio de Neyman-Pearson utilizando sistemas adaptativos, como redes neuronales artificiales (NNs), que son entrenados de manera supervisada para minimizar la función de coste adecuada. En este trabajo se plantea un caso

de detección de blancos Swerling I y II en presencia de ruido aditivo gaussiano y se proponen detectores basados en perceptrones multicapa (MLP) y redes neuronales con funciones de base radial (RBFNN).

Soluciones basadas en perceptrones multicapa con funciones de base de primer orden

Las redes neuronales facilitan la adquisición automática de información, la adaptación a las condiciones cambiantes del blanco y del entorno y la posibilidad de un refinamiento continuo. En cuanto a los requerimientos hardware, la utilización de ordenadores de propósito general para el entrenamiento de redes neuronales requiere de gran capacidad y rapidez de cálculo. En cambio, la simplicidad computacional de la neurona artificial permite que una red neuronal pueda realizarse con hardware específico de procesamiento masivo paralelo.

Son destacables los trabajos de J.L. Sanz y D. Andina, que han realizado un estudio profundo sobre la aplicación de MLPs a la detección radar. Han considerado la utilización de funciones objetivo distintas al error cuadrático medio [Andina1995a], comprobando que la función objetivo propuesta por El-Jaroudi y J. Makhoul [El-Jaroudi1990] permite obtener buenos resultados para probabilidades de falsa alarma más bajas [Andina1995b][Andina1995c]. Estos mismos autores han comparado el detector propuesto con el óptimo en el sentido de Neyman-Pearson [Andina1996a] utilizando blancos no fluctuantes, demostrando que con una red neuronal puede aproximarse muy bien dicho detector óptimo. Además, han realizado el estudio de la dependencia del funcionamiento con el número de pulsos integrados [Andina1996b], poniendo de manifiesto una gran dependencia, no sólo con el número de pulsos integrados, sino también con la relación señal a ruido utilizada durante el entrenamiento. En la misma línea de los investigadores anteriores, P.P. Gandhi y V. Ramamurti [Gandhi1997] estudiaron el funcionamiento de perceptrones multicapa para la detección de señales conocidas contaminadas con ruido de media nula y distintas funciones de densidad de probabilidad: gaussiana, laplaciana, gaussiana contaminada y Cauchy. Además, a diferencia de J.L. Sanz y D. Andina, que siempre consideran un ruido gaussiano de varianza unidad, Gandhi y Ramamurti estudian la influencia de la varianza de los distintos tipos de ruido. En [Burian1999], se estudia la robustez de detectores basados en MLP diseñados para los casos propuestos en [Gandhi1997] frente a variaciones de la relación señal a ruido.

Teniendo en cuenta la posibilidad de expresar el eco radar recibido como una señal compleja, T. Bucciarelli y otros [Bucciarelli1993a][Bucciarelli1993b] consideraron la aplicación de MLP con aritmética compleja, demostrando una mejora en las prestaciones con respecto a los que utilizaban aritmética real. Sólo consideraron blancos no fluctuantes en presencia de ruido blanco, aditivo, gaussiano, de media nula.

Las soluciones más extendidas hasta el momento para simular distintas probabilidades de falsa alarma con detectores neuronales se basan en utilizar un umbral variable en la fase de test, [Andina1995d], o añadir una neurona extra en la capa de salida que fija el umbral de detección, [Burian1999].

Una estrategia completamente diferente, pero también basada en el uso de redes neuronales, es la propuesta por G.A. Hanson y R.A. Iltis [Hanson1998]. En este trabajo se propone un detector casi-óptimo de señales complejas con parámetros desconocidos basado en una aproximación numérica de la razón de verosimilitud. En general, el cálculo de la razón de verosimilitud, que para este tipo de problemas da lugar a una integral que no tiene una solución cerrada, y se

demuestra que un detector construido con una red neuronal entrenada para aproximar al detector óptimo consigue resultados casi idénticos a los proporcionados con un detector construido siguiendo el criterio minimax.

En [Blacknell1994], se utilizan MLPs para clasificar datos no gaussianos, típicos de imágenes SAR, y se comprueba que los porcentajes de error obtenidos con las herramientas basadas en MLPs aproximan los correspondientes al clasificador óptimo de mínima probabilidad de error. Las funciones de densidad de probabilidad no gaussianas consideradas son la función de distribución K y Weibull.

También se han estudiado las redes neuronales en la aproximación de detectores CFAR. Amoozegar y otros [Amoozegar1994] compararon el funcionamiento de detectores basados en redes multicapa de propagación directa con el de detectores CFAR convencionales, demostrando las mejores prestaciones de los primeros cuando la longitud de la ventana de referencia es pequeña y su robustez frente a la presencia de blancos interferentes o discontinuidades de clutter. Recientemente, [Cheikh2004] ha confirmado que se obtienen mejores resultados con las técnicas neuronales que con los detectores CA-CFAR (*Cell-Averaging Constat False Alarm*) convencionales. En este último trabajo, en el caso de estudio se ha considerado interferencia no gaussiana modelada con la función de distribución K y se demuestra que las soluciones basadas en redes neuronales mantienen razonablemente constante la probabilidad de falsa alarma.

Soluciones basadas redes de neuronales con funciones de base de mayor orden

En principio, las NNs no son capaces de aprender aquello que no son capaces de hacer y, cuando son capaces, el aprendizaje depende de su estructura y las características de los elementos que la componen. Por lo tanto, en el caso de que el objetivo de aprendizaje no se haya conseguido, hay que modificar la estructura de la red o el espacio de entrada de manera que se plantee el problema de una nueva manera [Bishop1995]. Una de las limitaciones de los MLP con funciones de base de primer orden es que las fronteras de decisión que se forman se basan en hiperplanos.

Las redes de alto orden (HONN, *High Order Neural Network*) se caracterizan por definir las entradas a las neuronas de la red como una suma ponderada de productos de sus entradas (dando lugar a conexiones tanto de primer orden como de órdenes superiores). Las HONNs son capaces de aproximar cualquier asignación funcional de precisión arbitraria y por consiguiente constituyen una clase universal de asignación no lineal parametrizada [Villalobos1993].

El conjunto de entrenamiento de las redes de primer orden se diseña para conseguir que las redes sean capaces de generalizar y distinguir entre diferentes clases aunque existan ciertas variaciones en los parámetros de entrada. Las capas ocultas deben ser capaces de extraer las características invariantes de los patrones de entrenamientos para conseguir la generalización en el aprendizaje. Sin embargo, la arquitectura de la HONNs implica de forma inherente robustez frente a posibles variaciones, como la posición, la escala o la rotación y, por lo tanto, el entrenamiento de la red se centra en distinguir diferentes clases y no en la generalización, reduciendo considerablemente el tiempo para aprender de una forma más generalizada [Spirkovska1993].

C.L. Giles y T. Maxwell describieron en [Giles1987] el funcionamiento de las HONNs con la ecuación (1.1), donde $\omega_j^{(1)}$ son los pesos de las conexiones de primer orden, $\omega_{ji}^{(q)}$ son los pesos de las conexiones de orden q y S es la función sigmoide de activación de las neuronas.

$$y = S \left(\sum_j \omega_j^{(1)} z_j + \sum_j \sum_i \omega_{ji}^{(2)} z_j z_i + \dots \right) \quad (1.1)$$

En [Karayiannis1995] se estudia de forma detallada la arquitectura, entrenamiento y propiedades de las HONNs confirmando las prometedoras capacidades de este tipo de redes. Una de las desventajas que se presenta en este trabajo es el número elevado de pesos necesarios para el diseño. Aún así, las HONNs han sido utilizadas para el reconocimiento de patrones en diferentes aplicaciones [Wan1996], [Redlapalli2003], [Zhang2004] y [Artyomov2004]. Por ejemplo, en [Park2000] se aplican las HONNs junto a bancos de filtros direccionales para la detección automática de blancos en imágenes SAR (*Synthetic Aperture Radar*).

Soluciones basadas en redes neuronales de funciones de base radial

Las redes neuronales de funciones de base radial (RBFNN) presentan la atractiva característica de tener un entrenamiento rápido y sencillo. Las redes de funciones de base radial, también se han propuesto para resolver problemas de detección de señales en presencia de ruido. Por ejemplo, D.M. Hummels, W. Ahmed y M.T. Musavi [Hummels1995] [Ahmed1994] han propuesto el uso de este tipo de redes para la detección localmente óptima.

Al igual que con los perceptrones multicapa, la manera de aproximar el criterio de NP más extendida es variando el umbral de decisión a la salida de la red. Para mejorar estos métodos tradicionales, en [Casasent2003a] y [Casasent2003b] se propone una nueva red *RBFNN confidence-clustering*. El entrenamiento de este tipo de redes conlleva técnicas de *clustering de clases modificadas* para seleccionar los centros de las neuronas ocultas y elección de los pesos que unen las neuronas para asegurar la minimización del error cuadrático medio.

Teniendo en cuenta la capacidad de este tipo de redes de aproximar el detector de NP óptimo, se estudia en [Khairnar2008] el funcionamiento de RBFNN para la detección de señales conocidas contaminadas con ruido de media nula y distintas funciones de densidad de probabilidad: gaussiana, laplaciana, gaussiana contaminada y Cauchy.

Cheikh y Faozi presentan en [Cheikh2006] un detector radar que utiliza técnicas CFAR basadas en RBFNNs. Los resultados se comparan con técnicas CA-CFAR, OS-CFAR, MLPs entrenados con el algoritmo de retropropagación del error y MLPs entrenados con algoritmos genéticos. Los resultados demuestran que las técnicas basadas en redes neuronales mejoran los resultados de técnicas más convencionales y además mantienen razonablemente constante la probabilidad de falsa alarma ante cambios temporales en la distribución del clutter. El resultado de aplicar técnicas neuronales CFAR a datos reales también ha sido estudiado. En [Leung2002], se utiliza RBFNNs entrenadas con algoritmos genéticos para reconstruir la variación temporal no gaussiana del clutter marítimo y, además, utiliza estas redes para mejorar la detectabilidad de objetos pequeños manteniendo fija la probabilidad de falsa alarma. Los resultados han sido obtenidos tras aplicar datos adquiridos por dos tipos de radar, un radar coherente en banda X (situado en Osborne Head Gunnery Range, Nova Scotia, Canada) y por el sistema SAR RADARSAT, a las RBFNNs.

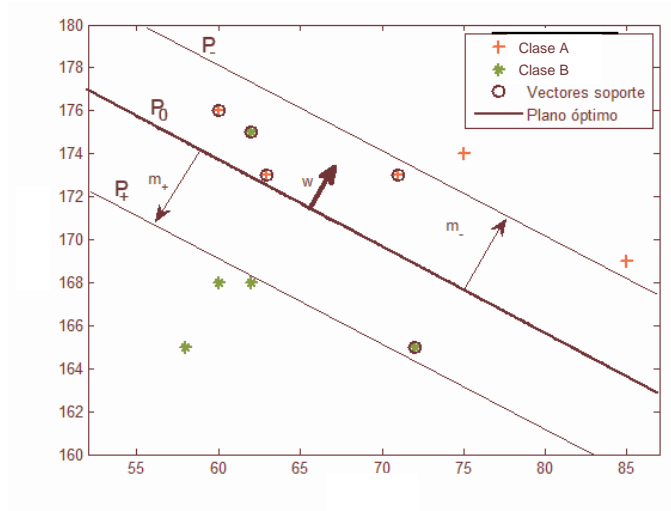


Figura 1.1: Solución de una SVM binaria no lineal.

Soluciones basadas en máquinas de vectores soporte

La teoría de las máquinas de vectores soporte (SVM, *Support Vector Machine*) es una técnica de clasificación desarrollada inicialmente por Vladimir Vapnik en [Vapnik1995]. La teoría de las SVM consiste en minimizar el riesgo estructural, con una cantidad finita de datos, que permita llevar a cabo una buena generalización. En muchas aplicaciones, las SVMs han mostrado tener gran desempeño, más que las máquinas de aprendizaje tradicional como las redes neuronales [Burges1998] y han sido introducidas como herramientas poderosas para resolver problemas de clasificación.

En la bibliografía se pueden encontrar muchos trabajos en los que se aplican las SVMs a la clasificación y reconocimiento de objetos incluso en escenarios radar, como por ejemplo en [Eryildirim2011] y [Fukuda2001]. Esta tesis se centra en la aproximación del criterio de Neyman-Pearson, NP, más extendido en aplicaciones de detección radar. Para entender la capacidad de las SVMs de aproximar este tipo de detectores es necesario formular el problema.

Una SVM primero mapea los puntos de entrada a un espacio de características de una dimensión mayor y encuentra un hiperplano que los separe y maximice el margen entre las clases en este espacio, tal y como se aprecia en la Figura 1.1.

El entrenamiento de una SVM requiere establecer una función de coste que penalice los vectores de entrenamiento que queden fuera de los márgenes y que proporcionan generalización en las regiones de decisión. En [Cortes1995] se formula la C-SVM según la expresión (1.2), donde C representa el parámetro de penalización de error o parámetro de regularización.

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \quad (1.2)$$

Sujeto a

$$\begin{aligned} t_i(G(\mathbf{w}, \mathbf{x}_i) + b) &\geq 1 - \xi_i & i = 1, \dots, n \\ \xi_i &\geq 0 & i = 1, \dots, n \end{aligned} \quad (1.3)$$

$G(\mathbf{w}, \mathbf{x}_i)$ representa la función kernel no lineal que relaciona el vector de entrenamiento \mathbf{x}_i con los vectores soporte \mathbf{w} , ξ_i representa el error asociado a \mathbf{x}_i y t_i la salida deseada para el vector \mathbf{x}_i .

En la formulación (1.2) se propone la misma penalización al error asociado a las dos clases. Surge entonces, una extensión de la C-SVM de penalización asimétrica, conocida como 2C-SVM y utilizada en [Osuna1997] y [Veropoulos1999] para entrenar y controlar la sensibilidad de las SVMs.

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C\gamma \sum_{i \in I_+} \xi_i + C(1 - \gamma) \sum_{i \in I_-} \xi_i \quad (1.4)$$

Sujeto a

$$\begin{aligned} t_i(G(\mathbf{w}, \mathbf{x}_i) + b) &\geq 1 - \xi_i & i = 1, \dots, n \\ \xi_i &\geq 0 & i = 1, \dots, n \end{aligned} \quad (1.5)$$

Donde γ es el parámetro de asimetría que penaliza de forma distinta el error cometido por los patrones pertenecientes a la clase cuya salida deseada es 1, (I_+), de los pertenecientes a la clase cuya salida deseada $t_i = -1$, (I_-).

Con esta nueva formulación, se plantean dos métodos principales para estudiar la capacidad de las SVMs en detección radar. El detector óptimo se caracteriza por su curva característica (ROC, *Receiver Operating Characteristic*) donde se relaciona la probabilidad de detección (P_D) para cada probabilidad de falsa alarma (P_{FA}). En [Bach2006] se presenta una forma para estudiar el comportamiento de las SVM para distintas P_{FA} basada en la obtención de \mathbf{w} con la formulación (1.2) y variar la bias o b en función de la P_{FA} requerida. En este mismo trabajo de F.R. Bach, D. Heckerman y E. Horvitz proponen una alternativa basada en la utilización de funciones de coste asimétricas, es decir la formulación (1.4) correspondiente a las 2C-SVM, que mejoran las curvas obtenidas con el método convencional. Utilizando este entrenamiento desbalanceado, en [Davenport2006] y [Davenport2010] se estudia la aproximación de un clasificador basado en una SVM a uno basado en el criterio Minimax y el de NP. En este trabajo, para aproximar el clasificador de NP, se propone la búsqueda de los parámetros libres de entrenamiento de la SVM para maximizar la P_D para un valor fijado de P_{FA} .

1.3. Problemas pendientes de resolver

Del análisis de la revisión realizada sobre la detección de blancos con parámetros desconocidos se pueden extraer las siguientes conclusiones:

- En la literatura, se han propuesto soluciones subóptimas basadas en métodos estadísticos para aproximar el cociente de verosimilitud promediado y se ha demostrado que estas soluciones pueden conllevar un alto coste computacional.
- Aunque se ha demostrado que los modelos teóricos de blanco y de interferencia se ajustan a los datos reales adquiridos por sistemas radar en diferentes escenarios, hay que tener en cuenta que se tratan de modelos paramétricos. Puesto que en la práctica dichos parámetros pueden ser variables es necesario plantear un estudio exhaustivo de la sensibilidad de los detectores óptimos en cada escenario, para plantear el diseño de detectores radar robustos.

- Se ha demostrado que detectores basados en redes neuronales entrenados de forma supervisada son capaces de aproximar el detector óptimo si se elige la función de error adecuada. Esta aproximación depende de la estructura y del tipo de red utilizado. Se deberá estudiar la posibilidad de proponer diferentes tipos de redes neuronales para aproximar el detector óptimo en el caso de test de hipótesis compuesto.
- Las soluciones basadas en inteligencia artificial aplicadas a la detección radar se centran principalmente en redes de primer orden, perceptrones multicapa y redes de funciones de base radial, quedando un estudio más detallado por hacer para redes de órdenes superiores o máquinas de vectores soporte.
- En los trabajos revisados relacionados con la aplicación de las máquinas de vectores soporte para aproximar el criterio de Neyman-Pearson, normalmente sólo se plantean los métodos para obtener mejores resultados sin presentar la capacidad matemática de aproximar dicho estadístico para cualquier valor de probabilidad de falsa alarma.

Todos estos puntos susceptibles de investigación justifican la realización de esta tesis doctoral. En ella se ha tratado de abordar la mayor parte de los temas planteados, ofreciendo soluciones razonables en muchos casos.

1.4. Estructura de la tesis

En la presente tesis se propone presentar casos de estudio de detección de blancos con parámetros desconocidos en presencia de diferentes modelos de clutter presentados en la literatura, que nos permitan aplicar las soluciones basadas en inteligencia artificial y poder comparar para extraer tendencias y conclusiones lo más generalizadas posible. Como el objetivo de las soluciones propuestas son las aplicaciones radar, los parámetros de los casos de estudio se determinan para ajustarse a los casos de interés en este tipo de aplicaciones.

Teniendo todo esto en cuenta, la tesis se estructura en los siguientes capítulos.

- En el capítulo 2 se presenta la estructura del detector considerado en esta tesis. Se caracteriza la señal transmitida y la señal recibida. Para ello, se consideran los distintos modelos de blanco utilizados durante el desarrollo de la tesis, así como los modelos de interferencia en el receptor.

Para los modelos de blanco se consideran blancos no fluctuantes, de fluctuación lenta, moderada y rápida basándose en los modelos estadísticos presentados en [Swerling1954]. Los modelos de interferencia utilizados se basan en los diferentes modelos estadísticos de clutter que se ajustan a datos adquiridos por diferentes sistemas radar reales. Para sistemas de baja resolución o con ángulos de incidencia elevados se propone un modelo de clutter gaussiano, mientras que para los sistemas radar de alta resolución o con ángulos de incidencia pequeños se proponen modelos de clutter gaussiano compuesto con textura K-distribuida.

- En el capítulo 3 se caracterizan los detectores óptimos de parámetros conocidos basados en el criterio de Neyman-Pearson para los casos de estudio considerados.

El primer caso plantea la detección de blancos fluctuantes (lentos, moderados y rápidos) en presencia de ruido blanco gaussiano aditivo. En [Jarabo2003] y [Jarabo2005a] ya se presenta este caso, por lo que en éste se resumen las características del detector de NP basado en el test LR, como punto de partida para la investigación posterior.

El siguiente caso considera el mismo tipo de blancos pero en presencia de interferencia gaussiana, presentado previamente en [Aloisio1994] y [diVito1999]. El objetivo de este apartado es presentar los parámetros de blanco y clutter que se han utilizado en estos trabajos y que se utilizarán a lo largo de la tesis para comparar los resultados en la curva ROC debidos a la presencia del clutter gaussiano.

Por último, se considerará la detección de blancos no fluctuantes en presencia de clutter K distribuido. En [Conte1991], [Pentini1992] y [Conte1994] se plantea el LR cuando la amplitud y la fase del blanco son conocidas. Teniendo en cuenta que en estos trabajos el clutter de distribución se modela como un proceso aleatorio esféricamente invariante (SIRP), en este apartado se presentan distintas curvas ROC obtenidas para los parámetros de blanco y clutter considerados en estos trabajos y que se mantendrán para los estudios posteriores.

- En el capítulo 4 se presenta la condición suficiente para que un detector basado en un agente inteligente entrenado para minimizar una determinada función de error sea capaz de aproximar al detector óptimo de Neyman-Pearson. Además, se presentan los agentes inteligentes que se van a utilizar para diseñar detectores robustos en la investigación de la tesis. Se han considerado redes neuronales con funciones de base de primer orden, de segundo orden, de base radial y máquinas de vectores soporte. El objetivo de considerar diferentes agentes se debe a que, aunque cumplan la condición suficiente para ser capaces de aproximar el detector NP, la realización de las fronteras de decisión está limitada por la arquitectura del agente inteligente.
- En el capítulo 5 se estudia la robustez de los detectores basados en el test LR cuando las propiedades estadísticas del blanco difieren de las asumidas en el modelo asumido en el diseño del detector óptimo. Los parámetros que se han considerado como variables son el coeficiente de correlación y la frecuencia Doppler del blanco, es decir, se plantea la detección de blancos, independientemente del tipo de blanco Swerling presente, y la detección de blancos en movimiento con diferentes velocidades.

El estudio de sensibilidad facilitará la comprensión de los resultados obtenidos para el detector basado en el test de hipótesis compuesto. Dado que en la mayoría de los casos presentados, el detector NP basado en el ALR conduce a una integral sin solución cerrada, se propone obtener aproximaciones basadas en estimadores de máxima probabilidad para valores discretos del parámetro variable. Estas aproximaciones basadas en el CGLR, se pueden implementar de forma práctica, por lo tanto, este estudio nos permitirá comparar tanto las capacidades de detección de las soluciones basadas en inteligencia artificial, como el coste computacional asociado a cada detector.

- En el capítulo 6 se presentarán las estructuras consideradas para cada agente inteligente, el algoritmo de entrenamiento y la función de error considerada para cumplir la condición

suficiente de aproximar el detector de NP descrita en el capítulo 4.

En todos los agentes inteligentes se considerará aritmética real, por lo que la entrada estará formada por un vector con una dimensión igual al doble de los ecos considerados en cada exploración, correspondientes a las partes reales e imaginarias de dichos ecos. Por otro lado, la salida de los agentes inteligentes será un único valor que se comparará con un umbral fijado por la probabilidad de falsa alarma deseado.

En este capítulo se hará especial hincapié en la estructura de las NNs de alto orden, en el algoritmo de entrenamiento completamente supervisado de las RBFNNs y en la capacidad de las SVMs de aproximar un único punto de la curva ROC del detector NP para un conjunto de parámetros de entrenamiento.

- Los resultados obtenidos para las soluciones basadas en agentes inteligentes para cada uno de los casos de estudio propuestos se presentan en el capítulo 7. Los resultados se encuentran estructurados por caso de estudio y, dentro de caso de estudio, por el parámetro del blanco desconocido considerado. Las mejores soluciones se aplicarán a bases de datos sintéticos generados a partir de un escenario radar real. El escenario se centra en la detección de embarcaciones con diferentes velocidades en el estrecho de Gibraltar con un radar cuyas características coincidan con las del radar IPIX [Drosopoulos1994]. Los resultados de detección se compararán con detectores utilizados en aplicaciones radar basados en CA-CFAR. Al final de dicho capítulo se presenta un resumen con las soluciones que presentan el mejor compromiso entre capacidad de detección y coste computacional.
- Finalmente, se resumen las principales conclusiones de la tesis y se presentan futuras líneas de investigación.

Este trabajo ha sido financiado por diferentes proyectos gracias a los cuales ha sido posible dotar al grupo de los medios técnicos necesarios y del apoyo económico adicional para la difusión de resultados:

- *Análisis de secuencias de imágenes de la superficie marina obtenidas con radares de navegación para aplicaciones de ingeniería costera*, Referencia: TEC2009-14217, financiado por la Secretaría de Estado de Investigación del Ministerio de Ciencia e Innovación.
- *Investigación en Tecnologías para la Gestión de la Inmigración (INTEGRA)*, Referencia UAH: 116/2010, convenio de colaboración en la modalidad de subcontratación por AMPER SISTEMAS, S.A. dentro del programa CENIT 2008-2011.
- *Estudio de campos de oleaje costero mediante el análisis de imágenes obtenidas con radares de navegación en banda X*, Referencia: CCG08-UAH/TIC-4242, financiado por la Comunidad de Madrid a través del programa de ayudas para apoyar las líneas de I+D en el programa de creación y consolidación de Grupos de Investigación de la Universidad de Alcalá.
- *Filtrado del clutter marino en series temporales de imágenes radar*, Referencia CCG06-UAH/TIC-0726, financiado por la Comunidad de Madrid a través del programa de ayudas para apoyar las líneas de I+D en el programa de creación y consolidación de Grupos de Investigación de la Universidad de Alcalá.

Capítulo 2

Modelos de señal radar

2.1. Introducción

En esta tesis se aborda el problema de detección utilizando un radar de exploración, primario de impulsos. El criterio más utilizado es el de Neyman-Pearson, NP, donde se trata de maximizar la P_D , manteniendo la P_{FA} igual o inferior a un valor determinado. Para poder aplicarlo deben conocerse las funciones de verosimilitud bajo las hipótesis nula y alternativa [Aloisio1994].

Se han utilizado modelos estadísticos de las señales radar para generar datos, de manera que son conocidas las funciones de verosimilitud y es posible formular el detector NP. Los blancos se modelan atendiendo a los conocidos como *blancos Swerling* [Swerling1954][Swerling1997], donde las observaciones se modelan como procesos con función de densidad de probabilidad gaussiana con distintas matrices de covarianza en función del coeficiente de correlación, la frecuencia Doppler normalizada y la potencia de la señal recibida. Los casos más utilizados I y II, se corresponden a los casos extremos en términos de coeficiente de correlación (el caso I corresponde a una correlación total de los pulsos recibidos en una exploración, mientras que el caso II modela la otra situación extrema en la que los pulsos recibidos están incorrelados).

Para la hipótesis nula se consideran diferentes modelos de clutter en función del tipo de clutter y el sistema de adquisición radar. El clutter gaussiano se suele asociar al clutter atmosférico, chaff y clutter marino y terrestre observados con ángulos de incidencia elevados o baja resolución. Cuando se trabaja con radares de alta resolución y ángulos de incidencia pequeños, los modelos de clutter no se ajustan a envolventes con distribución Rayleigh y se proponen modelos gaussianos compuestos cuya amplitud sigue, por ejemplo, una distribución K.

En el apartado 2.2 se explican las características principales de los radares primarios de exploración y de impulsos y la estructura de los sistemas de detección convencionales. En el apartado 2.3 se estudian los modelos estadísticos utilizados en la literatura para caracterizar la señal recibida del blanco y de la interferencia. Finalmente, en el apartado 4.2 se presenta el esquema de detección que se considera en esta tesis doctoral.

2.2. Radar primario de impulsos

La finalidad de este apartado es describir de una forma muy general la estructura de los detectores radar más comunes.

Un radar primario es aquel en el que la señal recibida es el resultado de la reflexión de la onda transmitida por el propio sistema al incidir sobre un objeto [IEEE-Std1997]. Al ser éste el modo de funcionamiento más común, normalmente se omite la palabra *primario*, a no ser que sea necesario establecer una clara distinción con los radares *secundarios*. Estos últimos son sistemas de identificación de blancos en los que se transmite una señal codificada, esperando respuesta del blanco.

Los sistemas radar se clasifican atendiendo a diversos criterios dentro de las siguientes categorías [IEEE-Std1997]:

- Según la posición relativa del transmisor y el receptor, los radares se clasifican en monoestáticos y multiestáticos. Los primeros transmiten y reciben a través de una antena común mientras que los segundos tienen dos o más antenas transmisoras o receptoras separadas distancias mayores que su tamaño.
- Atendiendo a su finalidad, los radares se clasifican en tres grupos: de vigilancia o exploración, de seguimiento y multifunción.
- Según su resolución los radares pueden ser convencionales (blancos puntuales), de alta resolución y de apertura sintética.
- Según el tipo de señal, pueden ser de impulsos o de onda continua.

El esquema básico de un radar primario monoestático es el que se presenta en la figura 2.1. En esta tesis se considerarán los radares primarios, monoestáticos, de exploración, convencionales y de impulsos.

En un radar de impulsos la señal transmitida es un tren de pulsos de radiofrecuencia. El duplexor permite utilizar la misma antena para transmisión y para recepción. La señal recibida es enviada al receptor, que es casi siempre heterodino. En frecuencias de microondas, el ruido presente es el generado por el propio receptor. En aplicaciones civiles suele emplearse una primera etapa amplificadora de bajo nivel de ruido. En cambio, en aplicaciones militares la salida del duplexor se aplica directamente a la entrada del mezclador. Esto es así porque en un ambiente con un alto nivel de interferencias, intencionadas o no, un radar con un receptor de bajo nivel de ruido es más susceptible. Además, el rango dinámico del receptor suele ser menor, característica no deseable ante contramedidas electrónicas hostiles (*ECM, Electronic Counter Measures*¹) o cuando hay que detectar blancos por efecto doppler en entornos con altos niveles de clutter. La mayor figura de ruido del mezclador puede compensarse por un incremento equivalente en la potencia transmitida.

En la inmensa mayoría de los sistemas radar, el amplificador de frecuencia intermedia se diseña como un filtro adaptado [Skolnik2008][Skolnik2002]. En la práctica, el filtro adaptado no puede obtenerse siempre de forma exacta. Con la finalidad de simplificar su implementación o eliminar otros tipos de interferencia, se emplean filtros que lo aproximan.

Casi todos los sistemas de detección radar están basados en el resultado de la comparación de la salida del receptor con un umbral. En el receptor radar se genera una señal que es función

¹Cualquier técnica electrónica cuya finalidad es evitar que el radar detecte los blancos deseados o adquiera información precisa [IEEE-Std1997].

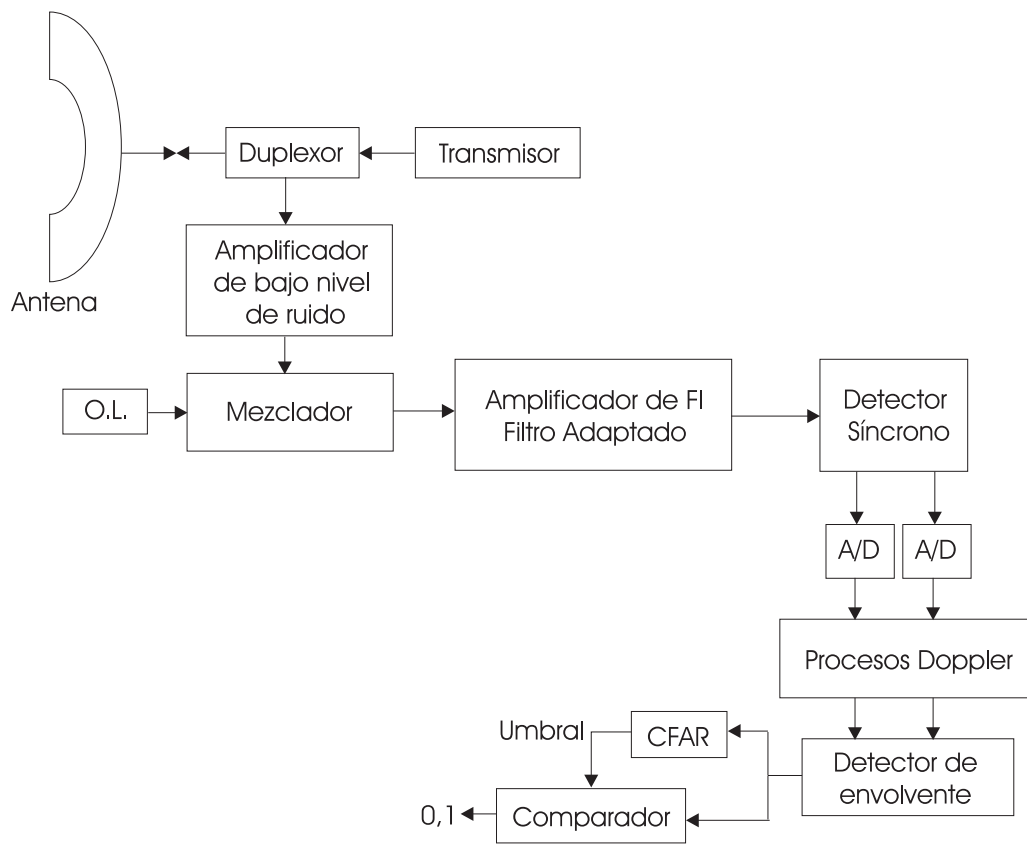


Figura 2.1: Esquema básico de un radar primario monoestático

de la envolvente de la señal de frecuencia intermedia. En sistemas que utilizan la información doppler para detectar blancos móviles en presencia de clutter (*sistemas MTI -“Moving Target Indicator”-*, *Radars Doppler Pulsados* o *Sistemas MTD -“Moving Target Detector”-*) es necesario mantener la información de fase. En estos casos, existe una primera etapa de detección, que suele implementarse con un detector de fase, un detector de pasos por cero o un detector síncrono, tras la cual se utilizan diversos procesos de filtrado cuya finalidad es eliminar las señales de clutter.

Debe indicarse que también es posible reconocer la componente doppler de un blanco móvil a partir de la fluctuación de la envolvente de la señal: *sistemas MTI incoherentes*. En estos casos los procesos de filtrado se realizan sobre la señal a la salida del detector del envolvente. El principal inconveniente de esta estrategia de detección es el hecho de que un blanco móvil sólo puede detectarse en presencia de clutter intenso y es posible que no se reciban señales de clutter en toda la zona de cobertura del sistema. Como solución, se puede desconectar el procesado MTI incoherente cuando las señales de clutter son poco intensas.

La última etapa del proceso de detección está formada por un detector binario que compara la salida del detector de envolvente con un umbral y decide la presencia o ausencia de blanco (si la salida del receptor es mayor o menor que el umbral, respectivamente). Si el clutter se elimina totalmente en las etapas previas de filtrado, puede emplearse un umbral fijo para conseguir una determinada P_{FA} . En la práctica la señal aplicada al detector binario tiene residuos de clutter, por lo que un umbral fijo dará lugar a un elevado número de falsas alarmas a no ser que el entorno sea conocido y homogéneo o las señales reflejadas por los blancos sean más intensas que las que provienen del clutter. Un receptor CFAR (*“Constant False Alarm Rate”*) mantiene la P_{FA} constante, a pesar de la naturaleza cambiante del nivel de ruido del receptor, el clutter o las contramedidas electrónicas [IEEE-Std1997]. Las tres técnicas principales que se han empleado para resolver este problema son los umbrales adaptativos, los procesadores CFAR no paramétricos y los mapas de clutter [Skolnik2008][Skolnik2002][Eaves1987]:

- En las técnicas de umbral adaptativo, se supone conocida la función densidad de probabilidad del ruido excepto alguno de sus parámetros, cuyos valores se estiman a partir de las observaciones del entorno. La técnica CFAR por promediado de celdas (*CA-CFAR*, *“Cell Averaging Constant False Alarm Rate”*) es la más sencilla.
- Los procesadores CFAR no paramétricos consiguen una P_{FA} constante sin utilizar ninguna información sobre las funciones densidad de probabilidad de la señal o la interferencia. El objetivo de este tipo de procesadores es ser insensibles a las variaciones de la función densidad de probabilidad, más que adaptarse a ellas. El detector no paramétrico más común es el detector de rango [Gonzalez2005].
- Los mapas de clutter están formados por las medidas de clutter realizadas con el radar para cada celda de resolución del sistema o para una determinada región, que son almacenadas con el fin de establecer los umbrales para cada celda de resolución en un sistema de detección CFAR, o bien para ajustar otros parámetros de procesado [IEEE-Std1997].

2.3. Caracterización de la señal recibida

En el caso más general, la señal transmitida por un sistema radar puede representarse como una señal de banda estrecha modulada tanto en amplitud como en fase:

$$s_T(t) = \sqrt{2E_T}a(t) \cos(\omega_p t + \theta(t)), \quad 0 \leq t \leq \tau \quad (2.1)$$

Donde ω_p es la pulsación de la portadora transmitida, $a(t)$ representa la modulación en amplitud, $\theta(t)$ la modulación de fase y E_T la energía de la señal.

En el caso de los radares pulsados, la forma de onda más común de $s_T(t)$ consiste en un tren de pulsos rectangulares de duración τ y periodo de repetición T_0 , que modula una portadora senoidal. La señal moduladora se expresa en (2.2).

$$a(t) = \frac{1}{\sqrt{\tau}}, \quad 0 \leq t \leq \tau \quad (2.2)$$

De esta forma, la expresión (2.1) puede escribirse como:

$$s_T(t) = \sqrt{\frac{2E_T}{\tau}} \cos(\omega_p t + \theta(t)), \quad 0 \leq t \leq \tau \quad (2.3)$$

Cuando esta onda incide sobre un determinado blanco, las propiedades dispersoras del mismo determinan las características de la señal recibida. Al estudio de estas características y de las interferencias que acompañan a la señal deseada se dedican los siguientes apartados.

2.3.1. Modelos de blancos radar

Sea una esfera conductora perfecta de radio r , cuyas coordenadas respecto al sistema radar son (R, ϕ, ϑ) , donde R es la distancia, ϕ el acimut y ϑ la elevación. Si en el instante $t = 0$ el radar emite un pulso de radiofrecuencia, la distancia a la que se encuentra el blanco se determina midiendo el tiempo t_0 que tarda la señal dispersada por el blanco en llegar al receptor. Así, la señal recibida por el sistema con retardo t_0 viene dada por:

$$s_R(t) = \alpha \sqrt{2\sigma_e E_T} a(t - t_0) \cos(\omega_p(t - t_0) + \theta(t - t_0)), \quad 0 \leq t - t_0 \leq \tau \quad (2.4)$$

donde:

- σ_e es la sección radar de la esfera. En el caso monoestático (las direcciones de incidencia y dispersión son iguales), la sección radar de un objeto se define como 4π veces el cociente entre la potencia dispersada por el objeto en la dirección del radar por unidad de ángulo sólido y la potencia por unidad de área de la onda plana que incide sobre dicho objeto [IEEE-Std1997]. En general, la sección radar es función de la geometría y del material del que está fabricado el objeto, de la posición relativa del objeto respecto al radar (dirección de incidencia), de la frecuencia de la señal transmitida y de la polarización de la onda incidente. Para la esfera conductora perfecta de radio r , si $r \gg \lambda$, siendo λ la longitud de onda de la portadora ($\lambda = \frac{2\pi c}{\omega_p}$), σ_e se aproxima a πr^2 [Skolnik2002][Skolnik2008].

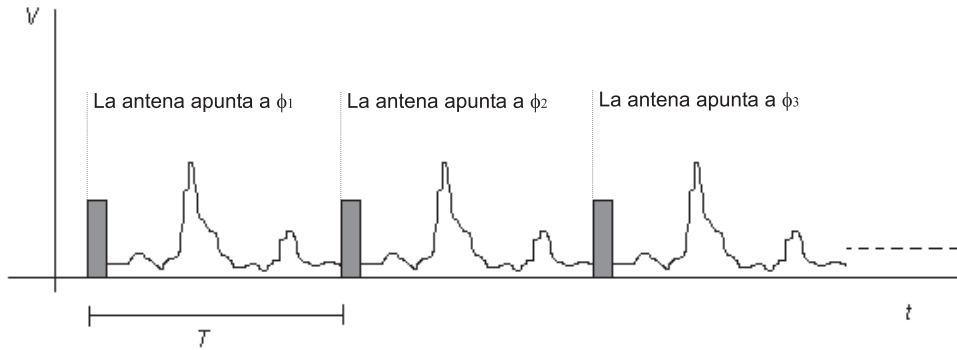


Figura 2.2: Ejemplo de envolvente de la señal recibida

- En el factor α se han agrupado las pérdidas por propagación en el espacio libre, cuyo valor depende de la frecuencia de la señal y de la distancia a la que se encuentra el blanco.

Si el receptor y el transmisor están sincronizados, es decir, si se puede asumir que el retardo t_0 es conocido por el receptor, podría eliminarse la dependencia de t_0 en la envolvente de la señal recibida:

$$s_R(t) = \alpha \sqrt{2\sigma_e E_T} a(t) \cos(\omega_p t - \theta), \quad 0 \leq t \leq \tau \quad (2.5)$$

Un ejemplo de la envolvente de la señal recibida es el que se presenta en la Figura 2.2, en la que las barras verticales gruesas representan el tiempo en que se transmiten los pulsos de radiofrecuencia. En esa figura se ha supuesto la presencia de dos blancos de las mismas características, uno más próximo, cuyo eco se recibe antes y tiene más energía, y otro más alejado. Cuando la antena apunta a una dirección caracterizada por un valor de acimut $\phi = \phi_1$, el transmisor lanza un pulso de radiofrecuencia. Pasado un tiempo T_0 el sistema apunta a $\phi = \phi_2$ y de nuevo emite un pulso de radiofrecuencia. Si hay un blanco presente, en el receptor se obtiene el eco del pulso con el retardo correspondiente. Esto se repite sucesivamente hasta completar una exploración completa. Resulta evidente que la diferencia de retardos mínima entre las señales dispersadas por dos blancos para que el sistema radar sea capaz de diferenciarlos es τ .

Las resoluciones en acimut y elevación vienen determinadas por los anchos del haz de la antena transmisora. Aunque la antena no apunte hacia la dirección en la que se encuentra un blanco, lo iluminará si queda dentro de su ancho de haz y recibirá la energía reflejada por el mismo. En la Figura 2.3 se muestra un ejemplo en el que aunque la antena apunte a tres direcciones distintas y sólo una coincide con la de los blancos, se recibirán tres pulsos por blanco.

Con estas consideraciones, es necesario conocer el número de muestras con información del blanco en una exploración, es decir, el número de pulsos recibidos en una exploración, P . Considerando únicamente un barrido en acimut, P puede calcularse con la expresión (2.6),

$$P = \frac{BW}{6\omega_g T_0} \quad (2.6)$$

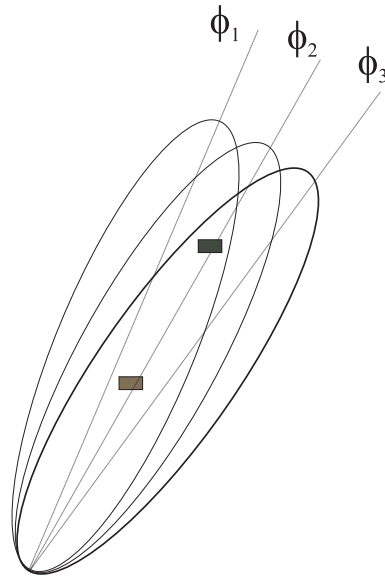


Figura 2.3: Ejemplo ilustrativo de la recepción de varios ecos de un mismo blanco durante una exploración

donde BW es el ancho de haz en acimut de la antena, ω_g es la velocidad angular a la que gira la antena expresada en revoluciones por minuto (rpm) y T_0 es el periodo de repetición de pulsos en segundos. De esta manera se construye una matriz de datos resultado de muestrear toda la región de cobertura del sistema en acimut y distancia.

Sin considerar la modulación en amplitud provocada por la forma del haz de la antena, en el caso de una esfera conductora perfecta, la amplitud y la fase de los P pulsos recibidos del blanco en una exploración y en exploraciones sucesivas son constantes. Este modelo corresponde al blanco puntual estacionario y fue el supuesto por Marcum [Marcum1948], con la salvedad de que realizaba una detección incoherente de la señal recibida, por lo que perdía la información de fase.

En el caso de considerar un blanco más complejo, fijada la frecuencia de la portadora y la polarización de la antena transmisora, la sección radar será función de la dirección de incidencia, por lo que su valor fluctuará como resultado de vibraciones, rotaciones y flexiones del objeto. Es necesario, por tanto, caracterizar el blanco mediante medidas, solución que es muy costosa. Otra posibilidad consiste en descomponer el blanco en numerosos objetos dispersores y calcular teóricamente la sección radar de la agrupación de dichos objetos haciendo uso de la Teoría Electromagnética. También se puede obtener un modelo matemático, con el que se supone un blanco genérico compuesto por numerosas superficies reflectantes de orientación arbitraria y se le modela como un blanco puntual cuya sección radar es aleatoria. Esta última opción es la que se considera en el desarrollo de la presente tesis.

Un método muy extendido para representar la fluctuación de los blancos son los cuatro modelos estadísticos propuestos por Peter Swerling en 1954 [Swerling1954]. Según Swerling, si el blanco es de fluctuación lenta, la amplitud de los P pulsos recibidos durante una exploración es constante, pero dicha amplitud varía de forma independiente de exploración a exploración (blancos Swerling I y III). En cambio, si la amplitud recibida varía de forma independiente pulso

a pulso, se trata de un blanco de fluctuación rápida (blancos Swerling II y IV). En función de las características del objeto y de los parámetros del sistema, las fluctuaciones pueden ser más o menos rápidas. En cualquier caso, se supone que la sección radar se mantiene constante durante la duración del pulso. La función de densidad de probabilidad de la sección radar de los blancos Swerling I y II es exponencial, mientras que para los blancos Swerling III y IV es χ^2 con cuatro grados de libertad.

Swerling sólo consideró el caso de detección incoherente de blancos radar empleando detectores de ley cuadrática, por lo que le bastaba con caracterizar las fluctuaciones de la sección radar. En el caso de detección coherente, es necesario modelar las funciones de densidad de probabilidad tanto de la amplitud como de la fase de la señal recibida.

Los casos I y II modelan blancos formados por varios reflectores (basta con que sean 4 o 5) con la misma sección radar, σ' , que fluctúan de forma independiente. La señal recibida puede calcularse como la suma de las ondas dispersadas por cada uno de ellos, tal y como se expresa en (2.7).

$$s_R(t)|_{n \text{ dispersores}} = \sum_{i=1}^n \alpha \sqrt{2\sigma' E_T} a(t) \cos(\omega_p t - \theta_i), \quad 0 \leq t \leq \tau \quad (2.7)$$

Como cada objeto dispersor fluctúa aleatoriamente, el desfase θ_i , $i = 1, \dots, n$, puede modelarse como una variable aleatoria uniforme en el intervalo $[0, 2\pi)$. La señal recibida puede expresarse como en (2.8), donde σ es la sección radar del blanco complejo, que se modela como una variable aleatoria de tipo exponencial. La amplitud de $s_R(t)|_{n \text{ dispersores}}$ se modela como una variable aleatoria de tipo Rayleigh [Papoulis1991] y la fase θ_R se modela como una variable aleatoria uniforme en el intervalo $[0, 2\pi)$.

$$s_R(t)|_{n \text{ dispersores}} = \alpha \sqrt{2\sigma E_T} a(t) \cos(\omega_p t + \theta_R), \quad 0 \leq t \leq \tau \quad (2.8)$$

Los casos III y IV modelan blancos compuestos por un gran reflector especular y un determinado número de reflectores pequeños que fluctúan aleatoriamente, o un gran reflector sometido a cambios de orientación muy pequeños. La amplitud de la señal reflejada tendrá una función de densidad de probabilidad Rayleigh generalizada con 4 grados de libertad (raíz cuadrada de una χ^2 con 4 grados de libertad) y la fase será uniforme en el intervalo $[0, 2\pi)$.

En general, un blanco queda completamente caracterizado por las funciones de densidad de probabilidad de su sección radar σ (o su raíz cuadrada, $\sqrt{\sigma}$) y de la fase de la señal recibida, θ_R , así como las correspondientes densidades espectrales de potencia. Sin embargo, en la práctica, estos parámetros pueden ser variables aleatorias con una función de densidad de probabilidad desconocida, lo que supone plantear el diseño de detectores radar robustos frente a dichos parámetros.

Si denotamos $E_{RS} = E_T \alpha^2 \sigma$ y expresamos (2.9) en términos de su envolvente compleja:

$$s_R(t) = \text{Re} \left\{ \sqrt{2E_{RS}} a(t) \exp(j(\omega_p t + \theta_R)) \right\}, \quad 0 \leq t \leq \tau \quad (2.9)$$

La envolvente compleja de la señal en frecuencia intermedia:

$$\tilde{s}_R(t) = \sqrt{2E_{RS}}a(t) \exp(j\theta_R), \quad 0 \leq t \leq \tau \quad (2.10)$$

Esta señal se aplica a la entrada de un filtro adaptado al pulso transmitido. Como $a(t)$ tiene energía unidad, la salida de este filtro será

$$\tilde{s} = \sqrt{2E_{RS}} \exp(j\theta_R) \exp(j\Omega) \quad (2.11)$$

Se añade un término que depende de la pulsación Doppler (Ω) que describe blancos en movimiento, es decir, con desplazamiento Doppler.

2.3.2. Modelo de ruido en el receptor

El ruido que aparece en los receptores radar se debe en parte al ruido térmico generado por los propios componentes del sistema receptor, pero también intervienen otros procesos. La mayoría de estos procesos producen ruidos que, en los anchos de banda típicos de los receptores radar, tienen las mismas características que el ruido térmico [Skolnik2002][Skolnik2008].

Dado que el nivel de ruido a la entrada del detector es el que se debe tener en cuenta para el cálculo de la relación señal a ruido, sólo se consideran las etapas de la cadena receptora desde la antena hasta la entrada del detector. Así, la potencia de ruido disponible a la entrada del detector, N , se puede expresar en términos de la *temperatura de ruido del sistema*, T_s , tal y como se indica en (2.12) [Skolnik2002][Skolnik2008][Hovanessian1984]:

$$N = kT_s B_n G_0 \quad (2.12)$$

donde k es la *constante de Boltzmann*, B_n es el ancho de banda de ruido de los circuitos de pre-detección definido en la expresión (2.13). En dicha expresión, $H(f)$ es la función de transferencia de los circuitos de pre-detección, f_0 es la frecuencia nominal del sistema y $G_0 = |H(f_0)|^2$.

$$B_n = \frac{\int_0^\infty |H(f)|^2 df}{|H(f_0)|^2} \quad (2.13)$$

Con todo esto, el ruido a la entrada del filtro de frecuencia intermedia se asume que es un proceso aleatorio blanco, gaussiano, de media nula y una densidad espectral de potencia $N_0/2 = kT_s/2$. Evidentemente, este modelo es una aproximación y tendrán que considerarse modelos más complejos cuando $\frac{f}{T} > 10^8$, donde f es la frecuencia expresada en (Hz) y T es la temperatura expresada en grados Kelvin. A 30 GHz la suposición es adecuada si la temperatura es al menos de 300K, por lo que es perfectamente válida para los sistemas considerados [Skolnik2008].

En realidad, este ruido ha sido filtrado en etapas previas de la cadena receptora, pero el ancho de banda de estos circuitos es mucho mayor que el de la señal y el del filtro de frecuencia intermedia (adaptado al pulso de radiofrecuencia transmitido cuya duración es τ). Por último, debe indicarse que en el caso de una antena sin pérdidas, su temperatura equivalente es función de la frecuencia y, en el margen de frecuencias de microondas, del ángulo de elevación. Aunque el ruido asociado a la antena no es blanco, puede considerarse como tal en los anchos de banda de los receptores típicos.

La función de autocorrelación de la envolvente compleja del ruido se puede aproximar por (2.14)[Kailath1966].

$$R_{\tilde{n}\tilde{n}}(\tau) = 2N_0\delta(\tau) \quad (2.14)$$

La muestra obtenida a la salida del filtro $y = \int_0^\tau \tilde{n}(t)\tilde{h}^*(t)$, supuesto que $\int_0^\tau |\tilde{h}(t)|^2 = 1$, es una variable aleatoria compleja gaussiana de media cero y varianza $2N_0$, por lo que sus componentes en fase y cuadratura también son gaussianas de media cero y varianza N_0 .

2.3.3. Modelos de clutter en el receptor

Clutter es un término utilizado para referirse a los ecos no deseados procedentes del entorno [IEEE-Std1997]. Típicamente se trata de reflexiones del mar, la tierra, la lluvia y otros tipos de precipitación, los pájaros, los insectos, los meteoros, la aurora o las producidas por *chaff*. Los ecos del clutter pueden enmascarar a los ecos de los blancos deseados limitando la capacidad de detección. Los objetos que dan origen al clutter, generalmente, se encuentran distribuidos espacialmente, de manera que su tamaño es mucho más grande que el correspondiente a la celda de resolución del radar. Se puede utilizar un modelo matemático que supone que la señal recibida como clutter está compuesta por la suma de numerosos elementos reflectantes de orientación arbitraria. Un clutter de superficie es aquel que considera que los elementos reflectantes se distribuyen aleatoriamente en una superficie, mientras que un clutter volumétrico considera la distribución aleatoria en un volumen.

La magnitud de un eco de clutter superficial (como por ejemplo el clutter marino o terrestre) es proporcional al área iluminada. Se define la sección radar normalizada, $\sigma^{(0)}$, con la sección radar del clutter por unidad de superficie:

$$\sigma^{(0)} = \frac{\sigma_c}{A_c} \quad (2.15)$$

Donde σ_c es la sección radar total para el clutter de superficie A_c .

De forma similar, se puede calcular la sección radar por unidad de volumen, en caso de clutter volumétrico (como por ejemplo el clutter atmosférico o *chaff*), en función de la sección radar total, σ_c , de un clutter de volumen V_c :

$$\eta = \frac{\sigma_c}{V_c} \quad (2.16)$$

Por lo tanto, el nivel del eco de un clutter volumétrico o de superficie depende tanto de η o $\sigma^{(0)}$ como del volumen o área iluminada, respectivamente.

La potencia de la señal recibida debida a un clutter superficial se describe en (2.17) donde se observa que la potencia recibida depende de la sección radar del clutter σ_c .

$$C = \frac{P_t G A_e \sigma_c}{(4\pi)^2 R^4} \quad (2.17)$$

donde P_t es la potencia transmitida, G y A_e son la ganancia y el área efectiva de la antena respectivamente y R es la distancia. Atendiendo a la geometría de los sistemas radar, Figura 2.4,

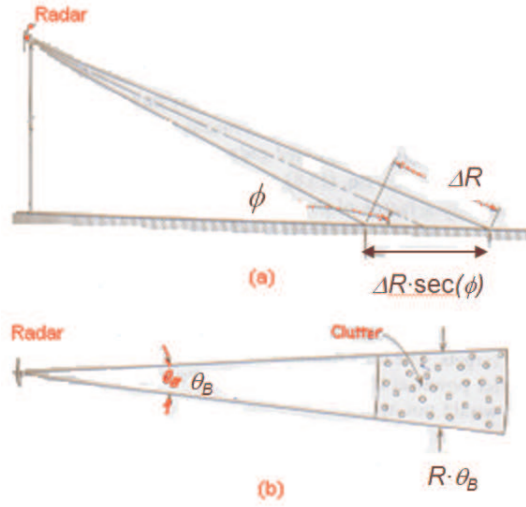


Figura 2.4: Geometría del área iluminada de clutter superficial

del ángulo de incidencia, ϕ , de la resolución del radar, ΔR , y del ancho de haz de la antena, θ_B , se puede reescribir (2.17) como:

$$C = \frac{P_t G A_e \sigma^{(0)} B W \Delta R \sec \phi}{(4\pi)^2 R^3} \quad (2.18)$$

En un escenario donde la potencia de clutter es tal que hace despreciable el efecto del ruido, la ecuación (2.19) expresa la relación entre la potencia recibida por blanco con una sección radar σ y la del clutter.

$$\frac{S}{C} = \frac{P_t G A_e \sigma_s}{(4\pi)^2 R^4} \frac{(4\pi)^2 R^3}{P_t G A_e \sigma^{(0)} B W \Delta R \sec \phi} = \frac{\sigma_s}{\sigma^{(0)} B W \Delta R \sec \phi R} \quad (2.19)$$

La envolvente compleja de la señal en frecuencia intermedia a la salida del filtro adaptado:

$$\tilde{c}_R = \sqrt{2E_{RC}} \exp(j\psi_c) = c_p + jc_q, \quad 0 \leq t \leq \tau \quad (2.20)$$

Las características de los ecos del clutter, c_p , c_q , v y ψ_c , se suelen modelar con distribuciones estadísticas en función del tipo de clutter, del ángulo de incidencia y de la resolución del sistema radar. Si consideramos, en el modelo matemático, que no predomina ninguno de los elementos reflectantes de orientación arbitraria, el clutter puede ser modelado como un proceso gaussiano. El clutter gaussiano se suele asociar al clutter atmosférico, chaff y clutter marítimo y terrestre observado con baja resolución o con alta resolución pero con ángulos de incidencia rasante altos. Cuando se trabaja con radares de alta resolución, los modelos de clutter no se ajustan a envolventes con distribución Rayleigh y los modelos más utilizados son modelos gaussianos compuestos cuya amplitud sigue una distribución K o Weibull [Skolnik2008].

Si comparamos las varianzas de las variables aleatorias de los modelos asumidos para cada tipo de señal podemos definir la relación señal a interferencia. Se distinguen tres tipos de escenarios

en función de la relación existente entre la potencia de clutter y la potencia de ruido calculada con la expresión 2.12:

- Escenarios limitados por clutter, en los que la potencia relativa a los ecos procedentes del clutter son mucho más grandes que los generados por ruido. Si llamamos C a la potencia del clutter, N a la de ruido y S a la de la señal, en este tipo de escenarios se puede aproximar $S/(N + C) \approx S/C = SCR$ (Relación señal a clutter).
- Escenarios limitados por ruido, en los que se puede despreciar los ecos procedentes del clutter por ser de baja intensidad y, por lo tanto, se puede asumir $S/(N+C) \approx S/N = SNR$ (Relación señal a ruido).
- Escenarios combinados, en los que ninguno de los efectos del ruido o el clutter son dominantes $S/(N + C) = SIR$ (Relación señal a interferencia).

Donde los valores de S y C dependerán de $2E[E_{RC}]$ y $2E[E_{RS}]$ respectivamente.

2.3.3.1. Clutter marino

La naturaleza de un eco generado por el mar depende de la forma de la superficie del mar y, por lo tanto, de las condiciones instantáneas del mar. El estado del mar es un término relacionado con la medida de la altura de las olas. Los cambios continuos que experimenta el estado del mar pueden deberse a la velocidad del viento, la dirección del viento, la polarización, a la existencia de manchas contaminantes e, incluso, a la temperatura del mar. En la Figura 2.5 se presenta la imagen del radar IPIX de alta resolución situado en Dartmouth, Canadá, en un acantilado frente al Océano Atlántico, a una altura de 100 metros sobre el nivel medio del mar y una vista al mar abierto de aproximadamente 130° .

En función del ángulo de incidencia y del tamaño de la celda de resolución se utilizan modelos teóricos del clutter marino distintos. En el caso de tener un tamaño de celda de resolución grande (en comparación con la longitud de onda del oleaje) y un ángulo de incidencia rasante superior a 10° , la variación de la amplitud del clutter en función de la distancia puede ser expresada en términos medios y se puede modelar mediante una distribución Rayleigh. En radares de baja resolución, el clutter marítimo es el resultado de la suma vectorial de un número elevado de contribuciones con fase aleatoria. La aplicación del teorema del límite central al problema da lugar a una estadística gaussiana de la señal recibida. Además, los ecos recibidos del clutter presentan tiempos de decorrelación muy pequeños, en torno a los 5-10 mseg, y son prácticamente incorrelados de pulso a pulso debido a la agilidad en frecuencia [Skolnik2008].

A medida que la resolución aumenta, y para ángulos de incidencia muy pequeños, los modelos de clutter pasan a ser correlados y de cola larga: se producen retornos radar potentes con una frecuencia mayor de lo esperado para un proceso gaussiano de la misma potencia media [Ward2006]. Diferentes modelos no gaussianos basados en, por ejemplo, distribuciones log-normal, distribución K y Weibull se han descrito en la literatura y se ha demostrado que aproximan las propiedades estadísticas de datos reales adquiridos por sistemas radar marinos convencionales. Generalmente estos modelos de clutter se basan en un modelo gaussiano compuesto que se puede interpretar como la composición de dos componentes: una modela las variaciones lentas del valor medio de la amplitud, llamada *textura*, mientras que la otra componente modela una variación rápida según

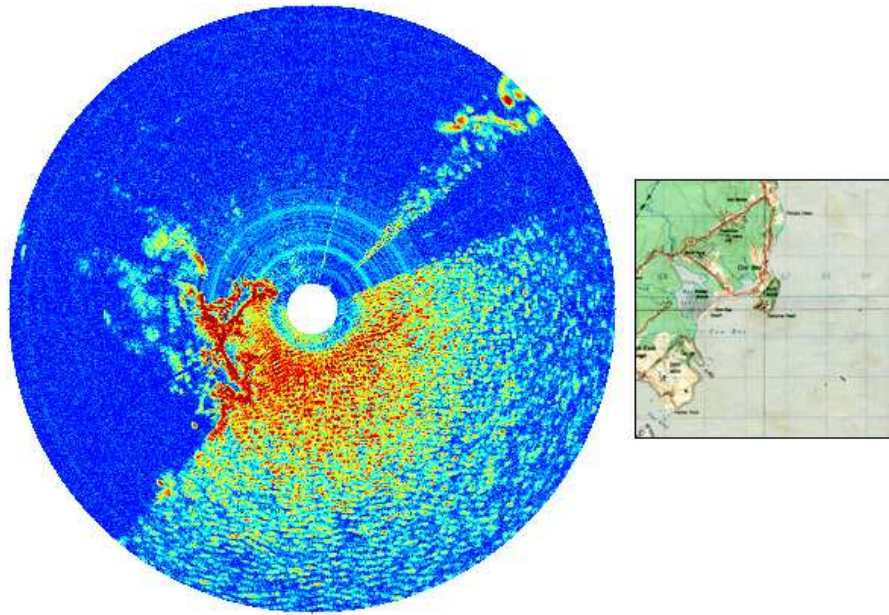


Figura 2.5: Imagen recibida por el radar IPIX situado en Dartmouth, Canadá.

una distribución Rayleigh, denominada como *speckle*. Por lo tanto, se puede interpretar como la composición de un valor medio debido a la altura del mar del área iluminada y un valor aleatorio causado por ondas capilares que están continuamente cambiando por los vientos locales.

En [Trunk1970] y [Trunk1972], se realiza un estudio sobre medidas de ecos procedentes del mar proporcionados en [Findlay1968], y se propone modelos en los que la textura sigue una función de distribución log-normal para aproximar las propiedades estadísticas de los datos reales, principalmente cuando se trata de ecos generados cuando el estado del mar es alto. Jakeman y Pusey introdujeron, en [Jakeman1976], la distribución K para modelar la sección radar del clutter marino, considerando correlaciones espaciales y temporales. Siguiendo con esta línea aparecen las publicaciones de S. Watts y su grupo de trabajo en [Watts1985], [Watts1987a], [Watts1987b], [Watts1990], [Ward1990], [Baker1991] and [Watts2000]. A lo largo de estos trabajos se presenta la capacidad de modelar la textura que modela la variación gaussiana de las ondas capilares con una función de distribución K, tanto para las variaciones temporales como espaciales, y se presentan curvas de detección para diferentes detectores basados en este modelo de clutter teórico.

En [Farina1997], se hace un estudio para comparar los modelos de clutter teóricos con datos reales, mostrando que cuando utilizamos polarización vertical el modelo de clutter de distribución K se ajusta correctamente a los datos reales, mientras que cuando utilizamos polarización horizontal presentan mejor aproximación los modelos de clutter basados en distribuciones log-normal. Los modelos presentados en ese trabajo degeneran en procesos aleatorios invariantes esféricamente (SIRP, *Spherically invariant random process*) cuando el clutter es observado en intervalos pequeños de tiempo. La teoría de los SIRP se describe en [Conte1987], [Conte1991] y [Rangaswamy2003] y también se ha comprobado experimentalmente que se ajusta a las propiedades estadísticas de datos reales de clutter marino adquiridos por radares de alta resolución en [Conte2004a], [Conte2004b], [Greco2004] y [Carretero2010]. En concreto, M. Greco, F. Bordoni y F. Gini en [Greco2004] comparan este modelo basado en la teoría SIRP considerando diferentes

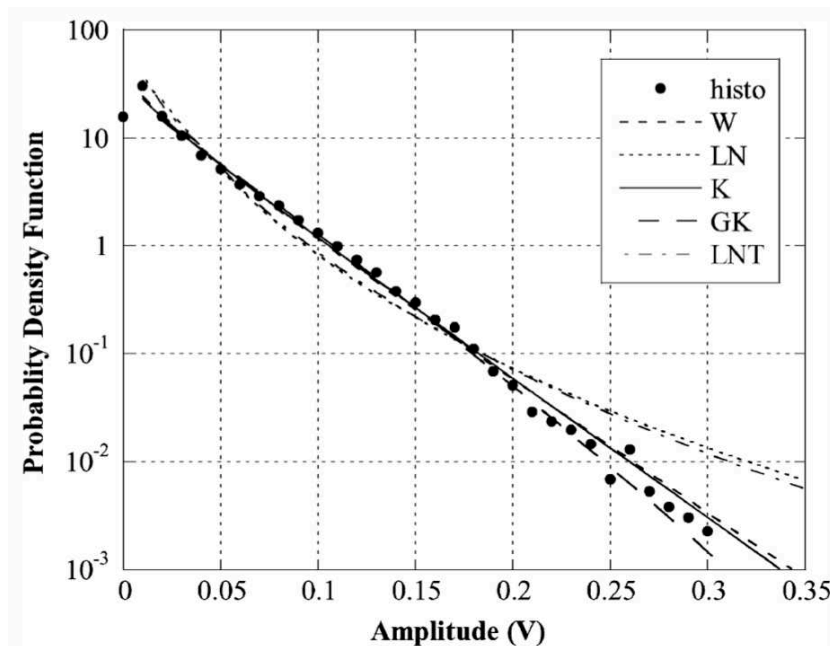


Figura 2.6: Comparativa de las FDPs de la amplitud de los datos adquiridos con el radar IPIX y diferentes modelos estadísticos [Greco2004].

funciones de densidad de probabilidad para la componente textura (función de distribución K, Weibull,...) con una base de datos prácticos adquiridos por un sistema radar de alta resolución (radar IPIX de exploración experimental en banda X con posibilidad de operar con agilidad en frecuencia). En la Figura 2.6, se presentan modelos de clutter que se ajustan a las estadísticas de los datos reales, comparando la función densidad de probabilidad de los datos adquiridos en las condiciones detalladas en la Figura 2.7 con las funciones de densidad Weibull (W), Log-Normal (LN), K distribuida (K), K distribuida generalizada con textura gamma (GK) y K distribuida generalizada con textura log-normal (LNT).

Los modelos de clutter Weibull también son propuestos en [Sekine1990] y [Ward2006] para modelar el clutter marino. Haykin, Bakker y Currier en [Haykin2002] comparan un modelo de

Número de celdas	7	Resolución en rango	30 m
PRF (agilidad en polarización)	2 kHz	Frecuencia RF	9,39 GHz
Número de barridos	262,144	Polarización	VV
Ángulo de incidencia	0,305°	Azimut (desde N)	134°
Dirección del viento (desde N)	280°	Velocidad del viento	7 km/h
Altura de la olas significativa	2,23 m	Periodo del campo de oleaje significativo	8,3 s

Figura 2.7: Características y condiciones de adquisición de datos del radar IPIX [Greco2004].

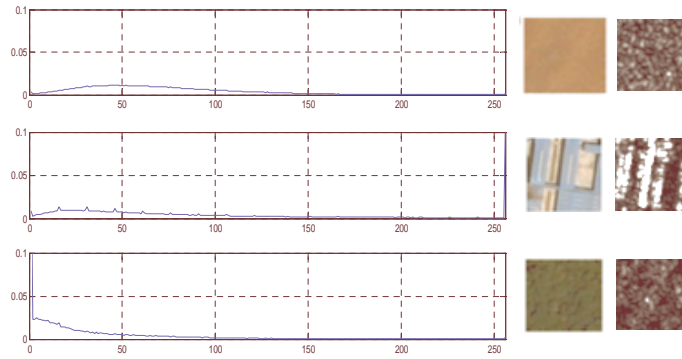


Figura 2.8: Funciones de densidad de probabilidad de datos radar correspondientes a diferentes trozos de terreno de una imagen adquirida por el TerraSAR-X [Mata2011b].

clutter de distribución K con la aplicación de la teoría del caos al clutter marino. Los resultados experimentales que se presentan en este trabajo muestran la posibilidad de utilizar el caos como un mecanismo dinámico no lineal para generar clutter marino.

Esta tesis se centra en los modelos de clutter K-distribuidos para modelar los ecos procedentes del clutter marino adquiridos por radares de alta resolución con ángulos de incidencia pequeños. En este escenario se han presentado diferentes estudios de detección de blancos radar [Conte1991], [Pentini1992], [Gini1997], [Conte1994] y [Gonzalez2005].

Un caso especial de clutter marino, son los ecos procedentes de los bloques de hielo flotantes. Cuando la temperatura baja, se pueden formar cristales de hielo en el mar que pueden crecer hasta convertirse en laminas de hielo con un cierto grosor. Estos bloques pueden ser estacionarios o moverse libremente en el mar. Los ecos procedentes de los bloques de hielo dependen de la antigüedad de la formación, del grosor, de la rugosidad de la superficie, de la temperatura del hielo e, incluso, de sus propiedades dieléctricas. En [Sayama2002], [Sayama2004] and [Hwang2006] se presentan modelos de clutter log-Weibull y Weibull, para modelar la variación estadística de la amplitud de dichos ecos.

2.3.3.2. Clutter terrestre

El clutter terrestre depende principalmente del tipo de terreno que se encuentre iluminado por el radar. Los estudios para modelar la amplitud de los ecos procedentes del clutter terrestre han considerado la dependencia con la longitud de onda, la resolución y la altura de la antena. Los valores medios de amplitud procedentes de ecos de distintos terrenos pueden ser iguales, sin embargo, la ley de variación no tiene por qué ser la misma. Por lo tanto, los modelos de clutter son muy importantes para el diseño de detectores radar que operan adecuadamente en entornos terrestres. En la Figura 2.8, se representa las distribuciones de datos SAR procedentes de diferentes terrenos.

Para estudiar los modelos estadísticos que se ajustan a la sección radar del clutter terrestre se han utilizado medidas radar $\sigma^{(0)} F^4$ que tienen en cuenta tanto la sección radar por unidad de superficie, $\sigma^{(0)}$, como los efectos de la propagación de dos rayos y la geometría de los mismos, F^4 . Una vez corregido este factor, la variación de la señal de clutter depende de la función de densidad de probabilidad de $\sigma^{(0)}$.

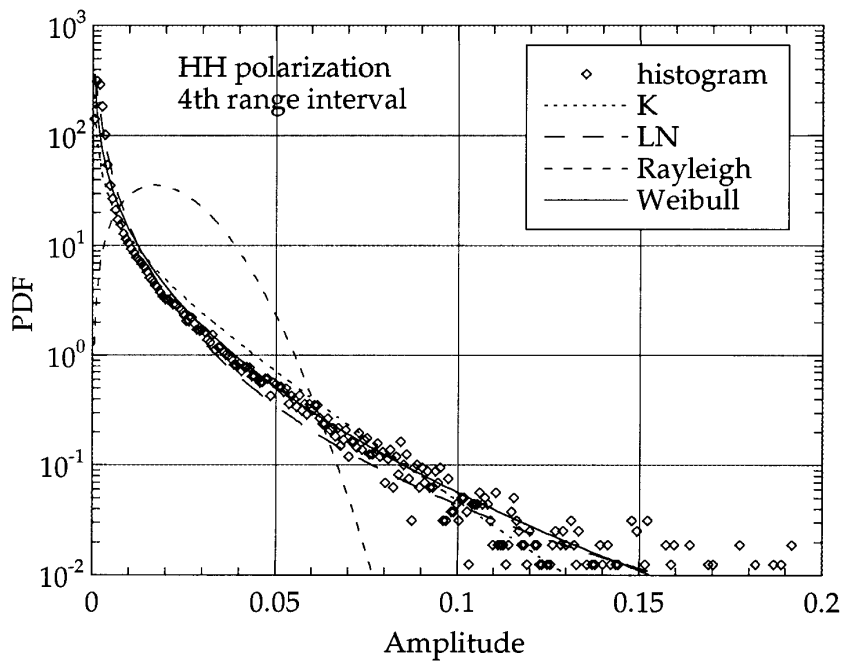


Figura 2.9: Comparativa de las FDPs de la amplitud de los datos adquiridos con el radar Phase One y diferentes modelos estadísticos [Billingsley1999].

Para ángulos de incidencia rasante pequeños, el clutter terrestre presenta un carácter impulsivo debido a valores elevados de amplitud causados por ecos procedentes de la presencia de obstáculos (árboles, edificios, montañas,...), así como valores muy bajos generados por las sombras de los mismos obstáculos. Cambios de elevación del terreno de 1 o 2 metros se corresponden a variaciones de amplitud de unos 10 dB de diferencia. Esta variación impulsiva nos conduce a tener en cuenta modelos teóricos de clutter no gaussianos [Barton1985]. Sin embargo, el efecto de los obstáculos se ve reducido según aumenta el ángulo de incidencia, y el clutter terrestre se puede modelar con un clutter Rayleigh [Skolnik2008].

En [Billingsley1999] y [Billingsley2002] se utilizan datos de clutter terrestre procedentes del *MIT Lincoln Laboratory Phase One program* para hacer estudios estadísticos de la variación espacial de los datos. Las medidas fueron adquiridas en banda X en Wolseley, Saskatchewan, localizado en Canadá. El área iluminada estaba compuesta por 83% de cultivos agrícolas, 11% de árboles de hoja caduca, 4% de lagos y un 2% de edificios. El terreno presenta elevaciones de menos de 2° de inclinación. En la Figura 2.9 se representa el histograma de la amplitud de los datos adquiridos en polarización HH cuando la antena cubre en ángulo desde los 360° a los 270° y en distancia desde los 15,2 a los 20 km. Las características del sistema radar (PRF=500 Hz y duración del pulso=100 nseg) permiten una resolución en rango de 15 m. Se proponen distribuciones Rayleigh, log-normal, distribución K y Weibull como modelos teóricos para aproximar el comportamiento de los datos prácticos. Después de ajustar los parámetros libres de cada distribución con el Método de los Momentos, se observa en la Figura 2.9 que el modelo de clutter Weibull se ajusta bastante al histograma de los datos reales. Similares resultados se obtienen para otras tomas de adquisición de datos y, finalmente, se concluye en [Billingsley1999] que el modelo de clutter Weibull es adecuado para generar datos simulados de clutter terrestre.

Un estudio similar se presenta en [Conte2004b] al comparar modelos teóricos con datos adquiridos con el radar IPIX McMaster con diferentes resoluciones. Se consideran resoluciones de 30, 15 y 3 metros. Independientemente de los datos adquiridos para las diferentes resoluciones, E. Conte, A. de Maio y C. Galdi demostraron que la distribución de la amplitud de los datos reales dista de ser Rayleigh y estudiaron los modelos teóricos basados en procesos invariantes esféricamente (SIRP). Los SIRP se describen como procesos gaussianos compuestos cuya amplitud puede seguir una función densidad de probabilidad K distribuida o Weibull [Conte1987]. Los resultados determinan que las FDP K distribuida se ajusta más al histograma de la amplitud de los datos adquiridos para cualquier resolución si la polarización es VV y cualquier polarización si la resolución es muy pequeña.

En [Sekine1990] y [Sayama2001], aparte de considerar modelos de clutter Weibull y K distribuidos, proponen modelos basados en funciones de distribución log-Weibull. En [Jao1984] y [Yamaguchi2002], se centran en modelos de clutter de distribución K para el clutter terrestre, ajustando el valor del parámetro de forma de la función de distribución K dependiendo del tipo de terreno considerado y la resolución del sistema radar. En [Greco2000], se utiliza un modelo de interferencia compuesto por clutter gaussiano más ruido aditivo blanco gaussiano para proponer un cancelador de clutter terrestre en un radar que opera en banda L.

2.3.3.3. Clutter atmosférico

Los radares, en general, se ven menos afectados por el clima que los sensores ópticos o de infrarrojos. Sin embargo, si trabajamos con radares navales en banda X, los ecos procedentes de la lluvia pueden ocupar una gran parte del espectro Doppler y enmascarar la potencia recibida de ecos procedentes de blancos que tienen una sección radar pequeña.

Se puede considerar que el número de gotas de lluvia iluminadas en cada celda de resolución de un radar es un valor muy elevado [Skolnik2008]. Cada gota presenta su sección radar, σ_i . La sección radar total de la lluvia, σ_c , es el resultado de la suma de los ecos recibidos de cada uno de los elementos independientes. En [Miller1988] se propone un modelo de clutter gaussiano definido por su función de autocorrelación gaussiana, porque en el caso de que la duración de un pulso sea mucho más grande que el periodo de repetición de pulsos y que las gotas estén distribuidas aleatoriamente en el espacio, se puede considerar que la fase y la amplitud de los ecos que se reciben en una exploración presentan una distribución uniforme y Rayleigh respectivamente. El mismo modelo se propone en [Thomson2001], para simular clutter atmosférico y compararlo con datos reales obtenidos con tres tipos de radar que trabajan en bandas X, C y S.

En varios artículos donde M. Sekine aparece como primer autor, [Sekine1979], [Sekine1980] y [Sekine1981], tras analizar datos reales de clutter atmosférico de la ciudad de Tokio, se demuestra que se mejoran las capacidades de detección cuando se utilizan modelos de clutter Weibull, tanto para las técnicas CFAR, como para sistemas MTI.

2.4. Casos de estudio

El objetivo de esta tesis es el diseño de detectores radar robustos en diferentes escenarios. Se han considerado los siguientes modelos de interferencia para abarcar el mayor número de posibles escenarios radar:

- Ruido blanco gaussiano aditivo (AWGN): se utiliza como caso de referencia.
- Clutter Gaussiano: Modelo que se ajusta a la variación estadística del clutter superficial en sistemas de baja resolución o con ángulos de incidencia elevados.
- Clutter de distribución K: Modelo que se utiliza en sistemas radar de alta resolución o con ángulos de incidencia pequeños para simular la señal de clutter terrestre o marítimo.

Para los modelos de blanco se consideran blancos no fluctuantes, de fluctuación lenta, moderada y rápida basándose en los modelos estadísticos presentados en [Swerling1954].

La combinación de estos modelos de interferencia y de blanco, definen los casos de estudio que se van a tratar en la tesis: en el capítulo 3 se estiman las curvas ROC para los detectores NP con parámetros conocidos, en el capítulo 5 se presentan aproximaciones al óptimo basadas en el CGLR cuando alguno de los parámetros del blanco es desconocido y en el capítulo 7 se estudian las soluciones basadas en inteligencia artificial.

Los casos de estudio considerados se basan en estudios previos que han sido publicados por diferentes autores:

- **Caso de estudio 1:** Detección de blancos fluctuantes (lentos, moderados y rápidos) en presencia de AWGN. En [Jarabo2003] y [Jarabo2005a] se estudia la detección de blancos de fluctuación lenta (Swerling I) y rápida (Swerling II) en AWGN. En esta tesis, se incluye la detección de blancos moderadamente fluctuantes con un coeficiente de correlación $\rho \in [0, 1]$.
- **Caso de estudio 2:** Detección de blancos fluctuantes (lentos, moderados y rápidos) gaussianos en presencia de interferencia gaussiana. En [Aloisio1994] y [diVito1999] se hace un estudio del detector óptimo para diferentes tipos de clutter gaussiano definidos por su potencia y coeficiente de correlación: clutter poco correlado y no dominante frente a AWGN, clutter muy correlado y no dominante frente AWGN y clutter muy correlado y dominante frente a AWGN.
- **Caso de estudio 3:** Detección de blancos no fluctuantes en presencia de clutter K distribuido. En [Conte1991], [Pentini1992] y [Conte1994] se plantea el detector óptimo cuando la amplitud y la fase del blanco son conocidas. El modelo considerado de clutter K-distribuido se basa en un proceso aleatorio esféricamente invariante (SIRP). El estudio contempla la posibilidad de la presencia de un clutter correlado o incorrelado.

Como este trabajo ha sido financiado por proyectos del Ministerio de Ciencia e Innovación (proyecto TEC2009-14217), de la Comunidad de Madrid y la Universidad de Alcalá (proyectos CCG08-UAH/TIC-4242 y CCG06-UAH/TIC-0726) y de la empresa AMPER SISTEMAS, S.A. (proyecto UAH: 116/2010) centrados en el área de clutter marino, los resultados están más orientados a este tipo de clutter superficial. En concreto, los mejores detectores propuestos basados en agentes inteligentes han sido aplicados a un escenario sintético generado a partir de un sistema radar real situado en un entorno costero. A partir de los datos proporcionados por la herramienta [Localizotodo], se han simulado las trayectorias de dos barcos dentro del área de cobertura de un radar situado en el Estrecho de Gibraltar. En función de la resolución del radar, de los ángulos de incidencia y del estado del mar se pueden plantear los casos de estudio descritos y estimar las capacidades de detección de los detectores propuestos en un escenario realista.

Capítulo 3

Detectores óptimos

3.1. Introducción

La detección de blancos radar utilizando sistemas inteligentes de aprendizaje supervisado requiere de conjuntos de entrenamiento construidos a partir de modelos de blancos y de interferencia para los cuales es posible obtener las características de operación del detector óptimo o una aproximación. De este modo, con las características de detección de los detectores óptimos se puede evaluar los resultados obtenidos con los esquemas de detección propuestos en la investigación desarrollada.

Para poder obtener las características de operación del detector, es necesario conocer las funciones de verosimilitud que aparecen en el cociente o razón de verosimilitud. A su vez, para poder obtener estas funciones, es necesario caracterizar estadísticamente el espacio de observaciones del problema de detección planteado.

En el apartado 3.2 se presenta la detección radar como un test de hipótesis binario, planteamiento válido cuando las funciones de distribución bajo ambas hipótesis son conocidas. En el apartado 3.3 se caracterizan los modelos estadísticos que se utilizarán para el espacio de observaciones bajo hipótesis H_0 de blanco ausente, donde se detallará el modelo de clutter gaussiano y de distribución-K, y bajo hipótesis H_1 de blanco, estacionario o fluctuante, presente. En este mismo apartado, se define la relación señal a interferencia y sus derivados en función de la relación clutter a ruido. Finalmente, en el apartado 3.4 se detallan los casos de estudio considerados en esta tesis y se obtendrán las reglas de decisión de los detectores óptimos siguiendo el criterio de Neyman-Pearson cuando los parámetros del blanco son conocidos. Finalmente, en el apartado 5.3, se estudiarán las reglas de decisión de los detectores óptimos en el caso de parámetros desconocidos del blanco. El problema de la detección se plantea como un test de hipótesis compuesto que conlleva la resolución de integrales complejas que puede que no presenten una solución cerrada y, como alternativa, se presentan aproximaciones sub-óptimas basadas en el cociente de verosimilitud generalizado (GLR).

3.2. Planteamiento como test de hipótesis binario

El problema de detección a resolver, puede plantearse como un test de hipótesis binario en el que el sistema tiene que decidir entre dos hipótesis, H_1 (blanco presente) y H_0 (blanco ausente),

a partir de un conjunto de observaciones representado por un vector de variables complejas de dimensión $P \times 1$: $\tilde{\mathbf{z}} = [\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_P]^T$.

El vector $\tilde{\mathbf{z}}$ estará compuesto por las siguientes contribuciones:

- Bajo la hipótesis H_0 de blanco ausente: $\tilde{\mathbf{z}} = \tilde{\mathbf{c}} + \tilde{\mathbf{n}}$.
- Bajo la hipótesis H_1 de blanco presente: $\tilde{\mathbf{z}} = \tilde{\mathbf{s}} + \tilde{\mathbf{c}} + \tilde{\mathbf{n}}$.

Donde $\tilde{\mathbf{s}}$, $\tilde{\mathbf{c}}$ y $\tilde{\mathbf{n}}$ son vectores de dimensión $P \times 1$ que se corresponden con las muestras complejas recibidas en banda base debidas al blanco, clutter y ruido respectivamente. El ruido, con las consideraciones realizadas en el apartado 2.3.2, se modela como un proceso aleatorio gaussiano de media nula y con una matriz de covarianza $\Sigma_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}$ que viene determinada por las características del sistema.

Como se vio en el capítulo 2, los modelos de clutter se pueden dividir principalmente en gaussianos, cuando los ángulos de incidencia rasante son elevados o cuando la resolución del radar es baja o, en caso contrario, no gaussianos. En esta tesis se considera dos modelos de clutter:

- Clutter gaussiano definido por la matriz de covarianza $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ de la variable compleja $\tilde{\mathbf{c}}$.
- Clutter compuesto por el producto de un proceso gaussiano complejo de variación rápida y un proceso modulador, independiente y de variación lenta, modelada por una distribución K.

Partiendo de la ecuación de la señal recibida en banda base por el clutter expresada en (2.20), una caracterización sólo de la envolvente $E_{RC}(t)$ puede no ser suficiente, y es necesario caracterizar la envolvente compleja $\tilde{c}(t)$, definiendo funciones de distribución de amplitud y funciones de correlación que se ajusten a modelos obtenidos con datos experimentales. En esta tesis se utiliza la teoría de procesos invariantes esféricamente (SIRP), definida en [Conte1987], [Conte1991b] y [Rangaswamy2003], por la cual, las funciones de densidad de probabilidad de cualquier orden pueden ser expresadas en función de formas cuadráticas definidas no negativas. En [Conte2004b] y [Greco2004], se comprueba que estos modelos se ajustan a datos prácticos adquiridos por radares de alta resolución en presencia de clutter terrestre y marino, respectivamente.

Análogamente, la señal recibida de la reflejada o dispersada por el blanco se puede expresar como una variable compleja \tilde{s} según la ecuación (3.1) obtenida a partir de (2.11). El blanco puede presentar un movimiento relativo con respecto a la interferencia, por lo que para modelar la fase de la variable compleja en banda base de un detector coherente, es necesario añadir un factor que dependa de la diferencia entre la frecuencia Doppler del blanco con respecto a la del clutter, $\delta(t)$.

$$\tilde{s}(t) = \sqrt{E_R(t)} \cdot \exp(j\theta_R(t)) \cdot \exp(j\delta(t)) \quad (3.1)$$

donde

$$\delta = [0 \quad \Omega \quad 2\Omega \quad \dots \quad (P-1)\Omega]^T \quad (3.2)$$

$\Omega = \varpi_D/PRF$ es la frecuencia Doppler angular normalizada. Debido a la ambigüedad en Ω , en radares con PRF bajas o medias, la frecuencia Doppler normalizada puede ser considerada

como una variable aleatoria uniforme entre $[0, 2\pi)$. Teniendo en cuenta la amplitud y la fase de la ecuación (3.1), pueden darse las siguientes situaciones:

- $\tilde{s}(t)$ es determinista. Tanto su envolvente como su fase son constantes conocidas: blanco no fluctuante de fase conocida.
- $a(t)$ es determinista pero su fase, $\theta_R(t)$, es aleatoria, asumiéndose que tiene una función de densidad de probabilidad uniforme en el intervalo $[0, 2\pi)$.
- Tanto la amplitud como la fase son parámetros aleatorios. En esta tesis se contempla el modelado de la variable compleja debida al blanco como un proceso gaussiano definido por su función de autocorrelación gaussiana [Scholtz1976].

En caso de conocer todos los parámetros para definir las FDP condicionadas a cada hipótesis, puede plantearse el diseño del detector óptimo basado en el cálculo de relación de verosimilitudes (LR, *Likelihood Ratio*) expresado en la ecuación (3.3).

$$\Lambda(\tilde{\mathbf{z}}) = \frac{f(\tilde{\mathbf{z}}|H_1)}{f(\tilde{\mathbf{z}}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{lr}(P_{FA}) \quad (3.3)$$

3.3. Caracterización del espacio de observaciones

Como ya se indicó en el apartado 2.2, el filtro de frecuencia intermedia de la Figura 2.1 se suele diseñar como un filtro adaptado al pulso transmitido. Si en el instante t_1 el transmisor transmite un pulso de radiofrecuencia y la antena apunta a $\phi = \phi_1$, al muestrear las salidas del detector síncrono se obtiene una secuencia de muestras separadas $\tau(s)$ (una para cada celda de resolución en distancia del sistema). Pasados $T_0(s)$, la antena apunta a $\phi = \phi_2$ y el transmisor lanza un nuevo pulso de radiofrecuencia. A la salida del detector síncrono de nuevo se obtiene una secuencia de muestras correspondientes a todas las celdas de distancia. Este proceso se repite hasta completar un barrido completo y así se construyen las filas de una tabla como la indicada en la Figura 3.1, resultado de muestrear en distancia y acimut para toda la zona de cobertura del radar.

El vector de observaciones $\tilde{\mathbf{z}} = [\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_P]^T$ está compuesto por muestras obtenidas a la salida del detector síncrono, para el mismo valor de distancia. Si el sistema recibe P pulsos de un blanco en una exploración, a la salida del detector síncrono se dispone de un total de $2P$ muestras (P componentes en fase y P componentes en cuadratura), correspondientes a la celda de distancia en la que se encuentra el blanco, y para distintas celdas de acimut. Estas muestras se han obtenido con una separación de $T_0(s)$, el periodo de repetición de pulsos del sistema.

Entonces el vector de observaciones se puede definir como un vector de número reales de dimensión $2P \times 1$ correspondiente las P primeras muestras a las componentes en fase y las P siguientes a las componentes en cuadratura:

$$\begin{aligned} \mathbf{z} &= [Re\{\tilde{z}_1\}, Re\{\tilde{z}_2\}, \dots, Re\{\tilde{z}_P\}, Im\{\tilde{z}_1\}, Im\{\tilde{z}_2\}, \dots, Im\{\tilde{z}_P\}]^T \\ \mathbf{z} &= [z_1, z_2, \dots, z_{2P}]^T \\ \mathbf{n} &= [n_1, n_2, \dots, n_{2P}]^T \\ \mathbf{c} &= [c_1, c_2, \dots, c_{2P}]^T \\ \mathbf{s} &= [s_1, s_2, \dots, s_{2P}]^T \end{aligned} \quad (3.4)$$

		Distancia (celdas de distancia)									
		d_1	...	d_{i-1}	d_i	d_{i+1}	...	d_{j-1}	d_j	d_{j+1}	...
Ángulo de apuntamiento de la antena (celdas de acimut)

	ϕ_1	0	...	0	x	0	...	0	x	0	...
	ϕ_2	0	...	0	x	0	...	0	x	0	...
	ϕ_3	0	...	0	x	0	...	0	x	0	...

Figura 3.1: Ejemplo de la matriz obtenida a la salida del detector síncrono (se han marcado con una "x" las posiciones en las que hay un blanco)

3.3.1. Hipótesis H_0 : blanco ausente

Ruido blanco gaussiano aditivo

Bajo la hipótesis H_0 la señal está compuesta por los ecos recibidos del clutter más ruido. Tal y como se ha explicado en el apartado 2.3.2, el ruido a la salida del filtro adaptado se obtiene ruido gaussiano coloreado de media nula y varianza N_0 (suponiendo que la respuesta al impulso del filtro se ha normalizado para que tenga energía unidad). Sus componentes en fase y cuadratura también son gaussianas de media cero y varianza $\sigma_n^2 = N_0 = P_n$ [Papoulis1991]. Como estas salidas se muestrean cada $\tau(s)$ y el ancho de banda del filtro de frecuencia intermedia es de $\frac{1}{\tau}(Hz)$, las muestras obtenidas son variables aleatorias gaussianas de media cero y varianza $\sigma_n^2 = N_0 = P_n$, independientes entre sí. Además, el vector de observaciones está formado por muestras de la misma celda de distancia pero distinto valor de acimut, por lo que también estará formado por variables aleatorias gaussianas independientes. Así, la función de densidad de probabilidad del vector de observaciones bajo hipótesis H_0 puede calcularse como el producto de $2P$ funciones de densidad de probabilidad de variables aleatorias gaussianas de media nula y varianza P_n :

$$f_{2P}(\mathbf{n}) = \frac{1}{(2\pi)^P P_n^P} \exp\left(-\frac{1}{2P_n} \mathbf{n}^T \cdot \mathbf{n}\right) \quad (3.5)$$

La FDP de la correspondiente variable compleja $\tilde{\mathbf{n}}$:

$$f_P(\tilde{\mathbf{n}}) = \frac{1}{(\pi)^P |\Sigma_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}|} \exp(-\tilde{\mathbf{n}}^T \Sigma_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}^{-1} \tilde{\mathbf{n}}^*) \quad (3.6)$$

Donde la matriz de covarianza, $\Sigma_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}$, es una matriz diagonal, $2P_n \mathbf{I}$ (\mathbf{I} es una matriz de $P \times P$ con unos en la diagonal y ceros en el resto), cuyos valores se corresponden a la suma de las varianzas de las componentes en fase y cuadratura.

Clutter

Para generar y simular la envolvente compleja del clutter, se ha utilizado la teoría SIRP. Un proceso aleatorio invariante esféricamente es aquel en el que cada vector de observación es un vector aleatorio invariante esféricamente (SIRV). Un vector $\tilde{\mathbf{c}}$ con media μ y matriz de covarianza $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ es un SIRV, si su FDP tiene la forma expresada en la ecuación (3.7) y la FDP de cualquier orden se puede obtener en función de una forma cuadrática definida positiva de la forma (3.8).

$$f_P(\tilde{\mathbf{c}}) = \frac{1}{(\pi)^P |\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}|} h_P[q(\tilde{\mathbf{c}})] \quad (3.7)$$

$$q(\tilde{\mathbf{c}}) = (\tilde{\mathbf{c}} - \mu)^T \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}^{-1} (\tilde{\mathbf{c}} - \mu)^* \quad (3.8)$$

h_P tiene que ser una función que asegure no sólo que f_P sea una FDP, sino que también lo sean f_{P-1}, f_{P-2}, \dots , obtenidas por integraciones sucesivas de componentes de $\tilde{\mathbf{c}}$. El tipo de funciones que cumplen esta condición tienen la forma descrita en (3.9).

$$h_P(q) = \int_0^\infty s^{-2P} \exp\left(-\frac{q(\tilde{\mathbf{c}})}{s^2}\right) f(s^2) ds^2 \quad (3.9)$$

Sustituyendo la ecuación (3.9) en (3.7):

$$f_P(\tilde{\mathbf{c}}) = \int_0^\infty \frac{1}{(\pi)^P |s^2 \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}|} \exp\left(-(\tilde{\mathbf{c}} - \mu)^T (s^2 \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}})^{-1} (\tilde{\mathbf{c}} - \mu)^*\right) f(s^2) ds^2 \quad (3.10)$$

Atendiendo a la ecuación (3.10), podemos decir que $f_P(\tilde{\mathbf{c}})$ representa una FDP gaussiana generalizada de orden P con $E[s^2] \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ como matriz de covarianza, siendo s una variable aleatoria definida positiva. Es decir, se puede obtener un SIRV como el producto de un vector gaussiano por un factor de escala s aleatorio, independiente y definido positivo. El vector gaussiano considerado es de media cero y matriz de covarianza $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ y el valor cuadrático medio de s es la unidad. Los SIRPs se ajustan a situaciones estacionarias y no estacionarias, y permite controlar de forma independiente las funciones de correlación y la FDP de la amplitud debido al modelo compuesto en el que se basa. Por lo tanto, los SIRPs pueden ser utilizados para generar clutter gaussiano y no gaussiano.

Clutter Gaussiano

En un clutter gaussiano se cumple la expresión (3.11).

$$h_P(q) = \exp(-q(\tilde{\mathbf{c}})) \quad (3.11)$$

Las propiedades de correlación, y por tanto las características espectrales, de un clutter gaussiano vienen definidas por la matriz de covarianza de $\tilde{\mathbf{c}}$ que se puede calcular con la expresión (3.12).

$$\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}(h, k) = 2p_c \cdot \rho_c^{|h-k|} \quad (3.12)$$

Donde

- $h, k = 1, 2, \dots, P$

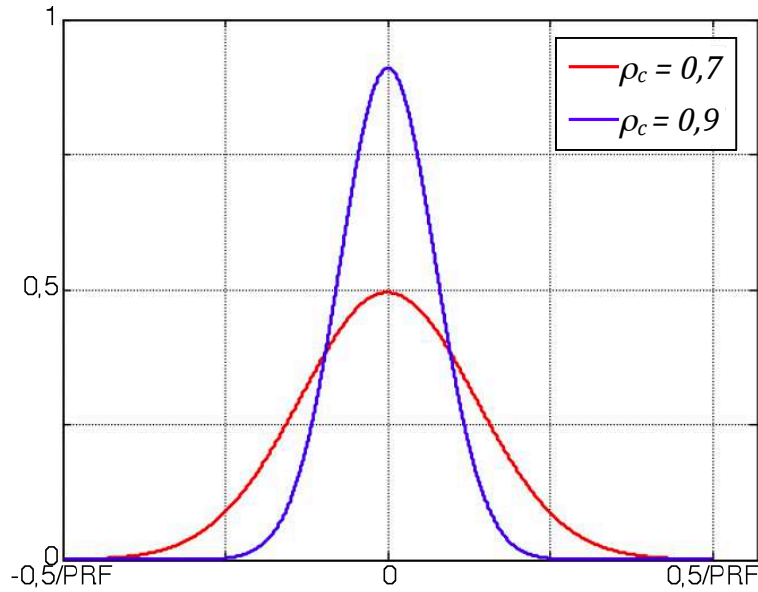


Figura 3.2: Espectros de clutter gaussiano para diferentes valores de coeficientes de correlación)

- p_c es la varianza de la partes real e imaginaria de la señal compleja del clutter que según el modelo expresado en (2.20) es igual a $E[E_{RS}]$.
- ρ_c es el coeficiente de correlación del clutter

La ecuación 3.13 define la relación que existe entre el coeficiente de correlación y la desviación estándar del espectro del clutter σ_c [diVito1999].

$$\rho_c = \exp\left(-2\pi^2 \left(\frac{\sigma_c}{PRF}\right)^2\right) \quad (3.13)$$

En la Figura 3.2, se representa el espectro de diferentes clutter gaussianos generados con la misma potencia y distintos coeficientes de correlación. Los valores de correlación $\rho_c = 0,7$ y $\rho_c = 0,9$ corresponden a valores tomados en [Aloisio1994] y [diVito1999] para el estudio de capacidades de detección de blancos en presencia de interferencia gaussiana. Según el valor de ρ_c se hace más pequeño, el espectro se va ensanchando tendiendo a un espectro plano cuando $\rho_c = 0$ correspondiente a un ruido blanco.

Para simular un clutter gaussiano de propiedades $\mu_{\tilde{\mathbf{c}}}$ y $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ se utiliza una transformación lineal a partir de un proceso de media nula, incorrelado y varianza unidad. Esta transformación se puede explicar a partir de una de las propiedades de las SIRV.

Si tenemos un SIRV, $\tilde{\mathbf{x}}$, de dimensión P con parámetros $\mu_{\tilde{\mathbf{x}}}$ y $\Sigma_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}$, se puede obtener otro SIRV, $\tilde{\mathbf{y}} = \mathbf{A}\tilde{\mathbf{x}} + \mathbf{b}$, también de dimensión P con parámetros $\mu_{\tilde{\mathbf{y}}} = \mathbf{A}\mu_{\tilde{\mathbf{x}}} + \mathbf{b}$ y $\Sigma_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}} = \mathbf{A}\Sigma_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}\mathbf{A}^T$, siempre que \mathbf{A} sea una matriz de $P \times P$ tal que $\mathbf{A}\mathbf{A}^T$ sea no singular y \mathbf{b} sea un vector de dimensión P .

Por ejemplo, los valores de \mathbf{A} y \mathbf{b} que permiten la transformación de una variable $\tilde{\mathbf{w}}$ con $\mu_{\tilde{\mathbf{w}}} = 0$ y $\Sigma_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}} = \mathbf{I}$ (siendo \mathbf{I} la matriz unitaria) en una variable $\tilde{\mathbf{c}}$ ($\tilde{\mathbf{c}} = \mathbf{A}\tilde{\mathbf{w}} + \mathbf{b}$) con $\mu_{\tilde{\mathbf{c}}} = 0$ y $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ se describen en (3.14) y (3.15) respectivamente.



Figura 3.3: Modelo SIRP para una secuencia compleja correlada no gaussiana

$$\mathbf{A} = \mathbf{E}^* \mathbf{D}^{\frac{1}{2}} \quad (3.14)$$

$$\mathbf{b} = \mathbf{0} \quad (3.15)$$

Donde \mathbf{D} es una matriz diagonal con los autovalores de $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$ y \mathbf{E} es la matriz de los autovectores normalizados de $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$.

La FDP de la variable $\tilde{\mathbf{c}}$ generada de esta forma se puede obtener a partir de las expresiones (3.10) y (3.11):

$$f_P(\tilde{\mathbf{c}}) = \frac{1}{(\pi)^P |\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}|} \exp(-\tilde{\mathbf{c}}^T (\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}})^{-1} \tilde{\mathbf{c}}^*) \quad (3.16)$$

Esta expresión nos permite diseñar el detector óptimo de Neyman-Pearson descrito en la ecuación (3.3), ya que forma parte de la función de verosimilitud condicionada a la hipótesis de blanco ausente H_0 cuando la interferencia es considerada gaussiana.

Clutter de distribución K

Hay situaciones en las que el modelo gaussiano no es aplicable porque la señal recibida no está compuesta por un número de dispersores elementales independientes e idénticamente distribuidos, sino que sólo un número limitado de los dispersores son los que realmente contribuyen en la formación del eco recibido. Este es el caso de trabajar con radares de alta resolución y ángulos de incidencia pequeños, donde la FDP de las amplitudes medidas dista mucho de ser Rayleigh [Watts1990] [Ward1990] y se ajusta más a modelos Weibull o de distribución K.

En esta tesis se ha utilizado el modelo de clutter de distribución K para este tipo escenarios radar. Según la ecuación (3.10), un SIRV tiene una matriz de covarianza resultado del producto $E[s^2] \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$. Para un modelo de clutter K distribuido la interpretación física es el producto de una intensidad distribuida en las celdas de distancia según una distribución gamma multiplicada por una componente compleja gaussiana independiente e idénticamente distribuida recibida en cada instante.

En la Figura 3.3, se presenta el esquema del modelo compuesto que genera una variable aleatoria $\tilde{\mathbf{c}}$ descrita en la ecuación (3.17).

$$\tilde{\mathbf{c}}_k = s_k \tilde{\mathbf{y}}_k \quad (3.17)$$

donde $\tilde{\mathbf{y}}$ representa el proceso estocástico complejo gaussiano que modela la reflexión local de la señal radar en la superficie del mar (*speckle*) y s el proceso aleatorio de variación lenta (*textura*), consecuencia de la modulación en potencia de la reflexión local debida al paso de los

frentes de ola por la celda de resolución del radar. La diferente velocidad de fluctuación de ambas componentes permite aproximar la secuencia de clutter recibida por la ecuación:

$$\tilde{\mathbf{c}}_k = s_m \tilde{\mathbf{y}}_k \quad m = k - L_c + 1, \dots, k + L_c \quad (3.18)$$

L_c es la longitud de coherencia de la textura, es decir, el número de pulsos durante las cuales es posible considerar que la textura es constante [Conte1997]. En cada intervalo de L_c pulsos, $\tilde{\mathbf{c}}_k$ es el producto de una constante aleatoria y un proceso estocástico gaussiano. La longitud de coherencia de la textura es función de la resolución en distancia, puesto que para una velocidad aproximadamente constante de los frentes de ola, el tiempo de tránsito de dichos frentes por la celda de resolución disminuye según se reduce su dimensión. Es interesante destacar, que dicho periodos L_c son del orden de duración de los intervalos de procesado de muchos sistemas radar de vigilancia [Carretero2010].

El proceso gaussiano $\tilde{\mathbf{y}}$ se caracteriza por una media nula y matriz de covarianza $\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$, generado a partir de una variable gaussiana $\tilde{\mathbf{w}}$ de media cero, incorrelada y potencia unidad, mientras que s^2 se caracteriza por un tener una función de distribución gamma:

$$f(s^2) = \frac{(\nu/\eta)^\nu}{\Gamma(\nu)} s^{2(\nu-1)} \exp(-\nu s^2/\eta) \quad (s^2 \geq 0) \quad (3.19)$$

donde $\Gamma(\cdot)$ es la función gamma, ν es el parámetro de forma (el clutter es más impulsivo cuanto menor es el valor de ν) y η es el valor medio de la distribución, que a partir de ahora y sin pérdida de generalidad supondremos igual a la unidad.

Sustituyendo e integrando en (3.9) la ecuación (3.19) de la FDP de una gamma, llegamos a la siguiente expresión:

$$h_P(q) = \frac{\nu^P 2^{1-\nu+P}}{\Gamma(\nu)} \left(\sqrt{4q\nu}\right)^{\nu-P} K_{\nu-P} \left(\sqrt{4q\nu}\right) \quad (3.20)$$

donde $q = \tilde{\mathbf{c}}^T \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}^{-1} \tilde{\mathbf{c}}^*$. La FDP de $\tilde{\mathbf{c}}$ cuya envolvente sigue una función de distribución K se define con la siguiente expresión:

$$f_P(\tilde{\mathbf{c}}) = \frac{\nu^P 2^{1-\nu+P}}{(\pi)^P |\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}| \Gamma(\nu)} \left(\sqrt{4q\nu}\right)^{\nu-P} K_{\nu-P} \left(\sqrt{4q\nu}\right) \quad (3.21)$$

La FDP (3.21) describe un vector $\tilde{\mathbf{c}}$ gaussiano cuya varianza está modulada según una distribución gamma. El espectro de la variable es gaussiano y su forma dependerá principalmente del factor de correlación de la parte gaussiana. Como ejemplo, se han generado patrones de clutter de distribución K con $\rho_c = 0,995$ y $\nu = 0,5$. En la Figura 3.4, se presenta el histograma de la envolvente de los ecos de un mismo patrón y del mismo eco de diferentes patrones, mostrando como en el primer caso la curva representada aproxima una distribución Rayleigh mientras que la segunda aproxima una gamma.

La Figura 3.4 también refleja la facilidad que proporciona trabajar con modelos SIRP ya que permite la caracterización de manera independiente la FDP de la amplitud y las funciones de correlación. Funciones que serán utilizadas para diseñar el detector de Neyman-Pearson cuando la interferencia sea debida a clutter de distribución K.

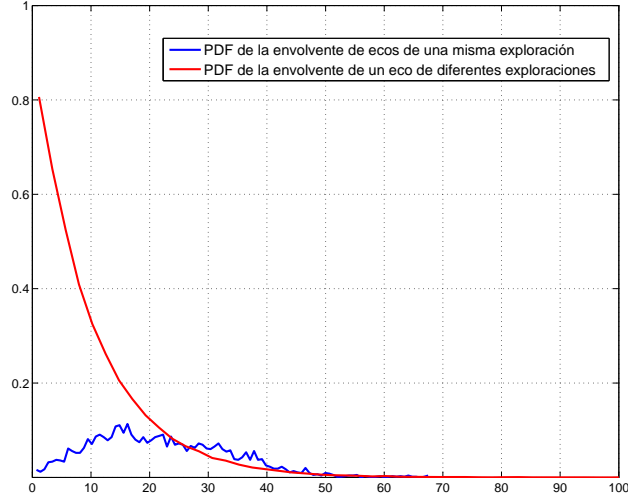


Figura 3.4: Histogramas de la envolvente de patrones de clutter K entre pulsos de una exploración, azul, y entre pulsos de distintas exploraciones, rojo.

3.3.2. Hipótesis H_1 : blanco presente

Según los modelos estadísticos de la variable aleatoria compleja debida a la señal recibida por los ecos del blanco, $\tilde{\mathbf{s}}$, definida en la expresión 3.1, se distingue entre blancos fluctuantes y no fluctuantes. Para blancos no fluctuantes se considera que la amplitud $a(t) = a$ es determinista. Sin embargo, la fase ϕ puede ser un parámetro desconocido y asumir una distribución uniforme entre $[0, 2\pi)$. Además, debido a la ambigüedad en la frecuencia Doppler normalizada del blanco con respecto a la interferencia, Ω puede ser considerada, también, como una variable aleatoria uniforme entre $[0, 2\pi)$. La FDP de $\tilde{\mathbf{s}}$ asumiendo un blanco no fluctuante de fase y frecuencia Doppler normalizada desconocidas se describe a continuación:

$$f_P(\tilde{\mathbf{s}}) = \frac{1}{2\pi} \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} \sqrt{2E_{RS}} \cdot \exp(j\theta_R) \exp(j\Omega) d\theta_R d\Omega \quad (3.22)$$

Por otro lado, los blancos fluctuantes son considerados procesos gaussianos con una función de autocorrelación, $\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}$, gaussiana. Se trata de blancos que dispersan un tren de pulsos correlados cuya envolvente sigue una variación Rayleigh. El coeficiente de correlación puede valer desde 0, modelo de blanco Swerling II, hasta 1, modelo de blanco Swerling I [Swerling1954][Scholtz1976].

Los ecos recibidos de un blanco pueden considerarse como el resultado de la suma de la contribución de dispersores elementales como, por ejemplo, los *puntos calientes* de un avión. Para este tipo de ecos, dentro de un mismo tiempo de barrido, se puede considerar que las componentes en fase y cuadratura de un detector coherente son procesos gaussianos estacionarios en sentido amplio con la correspondiente función de autocorrelación [Aloisio1994].

La generación de muestras de blancos fluctuantes sigue la misma transformación lineal que depende de las expresiones (3.14) y (3.15), donde \mathbf{D} y \mathbf{E} son los autovalores y autovectores normalizados de la matriz de covarianza del blanco, $\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}$. La matriz de covarianza $\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}$ se define:

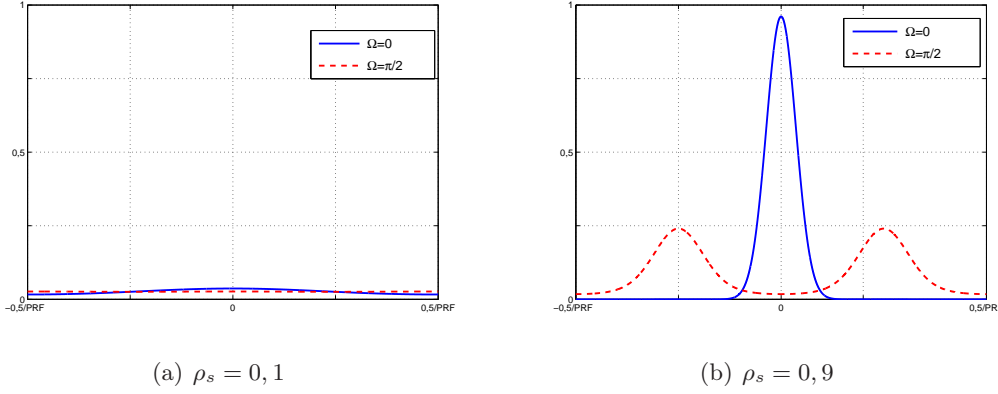


Figura 3.5: Espectro de un blanco gaussiano para distintas pulsaciones Doppler.

$$(\mathbf{M}_s)_{h,k} = 2p_s \cdot \rho_s^{|h-k|} \cdot \exp(j|h-k|\Omega) \quad (3.23)$$

Donde

- $h, k = 1, 2, \dots, P$
- p_s es la varianza de las componentes en fase y cuadratura de los ecos recibidos por el blanco que según la expresión (2.11) es igual a $E[E_{RS}]$.
- ρ_s es el coeficiente de correlación del blanco
- Ω frecuencia Doppler normalizada

Para calcular ρ_s se puede utilizar la ecuación 3.13 sustituyendo la desviación estándar del espectro del clutter, σ_c , por el del blanco, σ_s .

Suponiendo que todos los parámetros que caracterizan el blanco son conocidos podemos expresar la FDP de la variable compleja $\tilde{\mathbf{s}}$ con la siguiente ecuación:

$$f_P(\tilde{\mathbf{s}}) = \frac{1}{(\pi)^P |\boldsymbol{\Sigma}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}|} \exp(-\tilde{\mathbf{s}}^T (\boldsymbol{\Sigma}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}})^{-1} \tilde{\mathbf{s}}^*) \quad (3.24)$$

En la Figura 3.5 se representan espectros para diferentes parámetros del blanco. Se observa como blancos con ρ_s próximos a 0, es decir, a Swerling II, presentan un espectro más plano que aquellos que tienen un ρ_s próximo a 1 y, por lo tanto, al Swerling I. También se observa que la dependencia de la frecuencia Doppler normalizada en el espectro es mucho mayor cuanto mayor sea el coeficiente de correlación.

3.3.3. Relación señal a interferencia

Teniendo en cuenta la caracterización del vector de observaciones bajo las hipótesis H_0 y H_1 estudiada en los apartados 3.3.1 y 3.3.2, el criterio que se sigue para la generación de datos sintéticos de exploraciones radar considera que las muestras son independientes de exploración a exploración y que las muestras en fase son independientes de las componentes en cuadratura.

Se define la relación clutter a ruido a la entrada del detector como se indica en (3.25), donde CNR representa esta magnitud expresada en unidades logarítmicas (dB), mientras que cnr representa la misma magnitud en unidades naturales.

$$CNR(dB) = 10 \log_{10}(cnr) = 10 \log_{10} \left(\frac{2p_c}{2p_n} \right) \quad (3.25)$$

Como se explicó en el apartado 2.3.2, la varianza del ruido a la salida del filtro adaptado es función de la temperatura de ruido del sistema, T_s . Dicha temperatura a su vez depende de los distintos elementos de la cadena receptora hasta la entrada del detector, por lo que, con las consideraciones hechas en 2.3.2, en una determinada situación es una cantidad constante. Como criterio de normalización y sin pérdida de generalidad, se asume que las componentes en fase y cuadratura tienen potencia unidad y, por lo tanto, $P_n = 1$, con lo que la relación clutter a ruido queda:

$$CNR(dB) = 10 \log_{10}(cnr) = 10 \log_{10}(p_c) \quad (3.26)$$

Se define la relación señal a interferencia a la entrada del detector como se indica en (3.27), donde SIR representa esta magnitud expresada en unidades logarítmicas (dB), mientras que sir representa la misma magnitud en unidades naturales.

$$SIR(dB) = 10 \log_{10}(sir) = 10 \log_{10} \left(\frac{2p_s}{2p_c + 2p_n} \right) = 10 \log_{10} \left(\frac{p_s}{p_c + 1} \right) \quad (3.27)$$

En función de la relación entre la potencia del clutter y la del ruido definimos dos tipos de escenarios en los que se simplifica la expresión (3.27):

- La potencia del clutter es despreciable y sólo se tiene en cuenta la presencia de ruido, definiendo la relación señal a ruido (3.28), donde SNR representa esta magnitud expresada en unidades logarítmicas (dB), mientras que snr representa la misma magnitud en unidades naturales.

$$SNR(dB) = 10 \log_{10}(snr) = 10 \log_{10} \left(\frac{2p_s}{2p_n} \right) = 10 \log_{10}(p_s) \quad (3.28)$$

- La potencia del clutter es tan elevada que hace despreciable el efecto de ruido blanco gaussiano, definiendo la relación señal a clutter (3.29), donde SCR representa esta magnitud expresada en unidades logarítmicas (dB), mientras que scr representa la misma magnitud en unidades naturales.

$$SCR(dB) = 10 \log_{10}(scr) = 10 \log_{10} \left(\frac{2p_s}{2p_c} \right) \quad (3.29)$$

3.4. Detectores óptimos

En este apartado se estudian los detectores óptimos de blancos estacionarios (no fluctuantes) y de blancos fluctuantes de parámetros conocidos en interferencia gaussiana o de distribución K, atendiendo al criterio de Neyman-Pearson.

El objetivo de este apartado es presentar, a partir de estudio previos sobre los casos de estudio presentados en el apartado 2.4, el punto de referencia que se utilizará para definir los diferentes parámetros de los casos de estudio en los que se va a llevar a cabo la investigación de la tesis y para poder comparar los resultados que se vayan obteniendo. A continuación, se realiza una breve descripción de las publicaciones asociadas a cada caso de estudio:

- **Caso de estudio 1:** Detección de blancos fluctuantes gaussianos en presencia de ruido blanco gaussiano aditivo. Algunos de los artículos en los que se utiliza este escenario son [Jarabo2003] y [Jarabo2005a]. En estos artículos se considera la detección de blancos de fluctuación lenta (Swerling I) y rápida (Swerling II). En el apartado 3.4.1, se consideran, además, blancos moderadamente fluctuantes con un coeficiente de correlación $\rho \in [0, 1]$.
- **Caso de estudio 2:** Detección de blancos fluctuantes gaussianos en presencia de interferencia gaussiana. En [Aloisio1994] y [diVito1999] se hace un estudio del detector óptimo para diferentes parámetros de blanco y de clutter que se presentarán en el apartado 3.4.2.
- **Caso de estudio 3:** Detección de blancos no fluctuantes en presencia de clutter K distribuido. En el apartado 3.4.3 se presenta las curvas ROC del detector NP formulado en [Conte1991], [Pentini1992] y [Conte1994] para la detección de blancos estacionarios con amplitud y fase conocidas en presencia de clutter de distribución K generado como un SIRP. En este apartado se estudiará, además, el efecto del movimiento de los blancos en las curvas ROC del detector óptimo.

Para representar las curvas ROC de los detectores óptimos presentados en este apartado se ha considerado, sin pérdida de generalidad, que el vector de observación está formado por $P = 8$ pulsos.

3.4.1. Caso de estudio 1: Detector NP de parámetros conocidos

En [Jarabo2005a] se hace un estudio detallado de la detección de blancos fluctuantes en ruido blanco y gaussiano (AWGN). El estudio para blancos fluctuantes se centró en los modelos Swerling I y Swerling II. En esta tesis, se extiende este estudio para cualquier blanco fluctuante gaussiano con un valor de coeficiente de correlación $\rho_s = [0, 1]$.

Para resolver el problema de la detección, es necesario caracterizar el vector de observaciones bajo cada una de las hipótesis:

$$H_0 : \tilde{\mathbf{z}} = \tilde{\mathbf{n}} \quad (3.30)$$

$$H_1 : \tilde{\mathbf{z}} = \tilde{\mathbf{s}} + \tilde{\mathbf{n}} \quad (3.31)$$

Como ya se indicó en los apartados 3.3.1 y 3.3.3, $\tilde{\mathbf{n}}$ es una variable aleatoria compleja gaussiana independiente de media nula y varianza $\sigma_n^2 = N_0 = 2$. Si se aplica este criterio de normalización, la función de densidad de probabilidad bajo hipótesis H_0 será:

$$f(\tilde{\mathbf{z}}/H_0) = \frac{1}{(2\pi)^P} \exp(-\frac{1}{2}\tilde{\mathbf{z}}^T \cdot \tilde{\mathbf{z}}^*) = \frac{1}{(2\pi)^P} \exp(-\frac{1}{2}|\tilde{\mathbf{z}}|^2) \quad (3.32)$$

El vector de observaciones $\tilde{\mathbf{z}}$ bajo hipótesis H_1 está compuesto por la suma de un vector gaussiano, $\tilde{\mathbf{s}}$, de media cero y matriz de covarianza $\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}$ y un vector blanco y gaussiano debido al ruido. Por lo tanto, $\tilde{\mathbf{z}}$ también será un vector gaussiano de media cero y matriz de covarianza:

$$\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + \Sigma_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} = \Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2\mathbf{I} \quad (3.33)$$

donde \mathbf{I} representa la matriz unitaria. La función de densidad de probabilidad de $\tilde{\mathbf{z}}$ bajo hipótesis H_1 tiene la siguiente expresión:

$$f(\tilde{\mathbf{z}}/H_1) = \frac{1}{\pi^P |\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2\mathbf{I}|} \exp(-\tilde{\mathbf{z}}^T (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2\mathbf{I})^{-1} \tilde{\mathbf{z}}^*) \quad (3.34)$$

Planteando la regla de decisión a partir del cociente de verosimilitud, aplicando logaritmos y despejando se obtiene la siguiente expresión:

$$\tilde{\mathbf{z}}^T [(2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2 * \mathbf{I})^{-1}] \tilde{\mathbf{z}}^* \underset{H_0}{\overset{H_1}{\gtrless}} \text{Ln} \left(\eta_{cv} (P_{FA}) \frac{|\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2 * \mathbf{I}|}{2^P} \right) = \eta_s \quad (3.35)$$

donde η_{cv} es el umbral a utilizar con el cociente de verosimilitud que, al igual que η_s , se determina en función del criterio de detección elegido. Si se desea diseñar un detector de Neyman-Pearson, los valores de η_{cv} y η_s dependen de la P_{FA} elegida. La regla (3.35) depende de la matriz de covarianza del blanco, por lo que el valor de η_s para un valor de P_{FA} también dependerá.

En este apartado se asumen todos los parámetros de diseño conocidos, p_s , ρ_s y Ω , y se estudian las características del detector óptimo basado en el logaritmo del LR.

Si llamamos $\mathbf{Q} = (2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2 * \mathbf{I})^{-1}$, podemos decir, que según (3.35), la forma cuadrática $q = \tilde{\mathbf{z}}^T \mathbf{Q} \tilde{\mathbf{z}}^*$ es un estadístico suficiente. Bajo ambas hipótesis, el vector de observación es un proceso normal complejo de media cero. Una expresión sencilla para la transformada de Fourier de la forma cuadrática q [Turin1960][Barrett1987]:

$$\Phi_q(j\omega) = \frac{1}{|1 - j\omega \Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}^* \mathbf{Q}^*|} = \prod_{k=1}^P \frac{1}{1 - j\omega \lambda(k)} \quad (3.36)$$

donde $\lambda(k)$ son los autovalores de $\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}^* \mathbf{Q}^*$.

La FDP de la forma cuadrática q , se puede obtener a partir de la transformada inversa de Fourier utilizando el Teorema de los Residuos para, finalmente, obtener la probabilidad de superar un umbral T [Schwartz1956]:

$$\int_T^\infty f_q(q) dq = \sum_{k=1}^P \left[\exp\left(-\frac{T}{\lambda(k)}\right) \prod_{l=1; l \neq k}^P \frac{\lambda(k)}{\lambda(k) - \lambda(l)} U[\lambda(k)] \right] \quad (3.37)$$

donde $U[\cdot]$ es la función escalón, por lo que el sumatorio sólo se extiende a los autovalores, $\lambda(k)$, positivos. La ecuación (3.37) propocionará valores de P_{FA} , si $\lambda(k)$ son los autovalores de $\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}^* \mathbf{Q}^*$ bajo hipótesis H_0 , es decir, los autovalores de $(2 * \mathbf{I})^* \mathbf{Q}^*$. Mientras que si consideramos la matriz de covarianza del vector de observación bajo hipótesis H_1 , $\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2\mathbf{I}$, calculamos los autovalores $\lambda(k)$ y sustituimos en la ecuación (3.37), obtenemos valores de P_D .

En las Figuras 3.6 y 3.7 se representan curvas ROC del detector óptimo asumiendo que los parámetros de los vectores de observación bajo H_1 son conocidos e iguales a los considerados

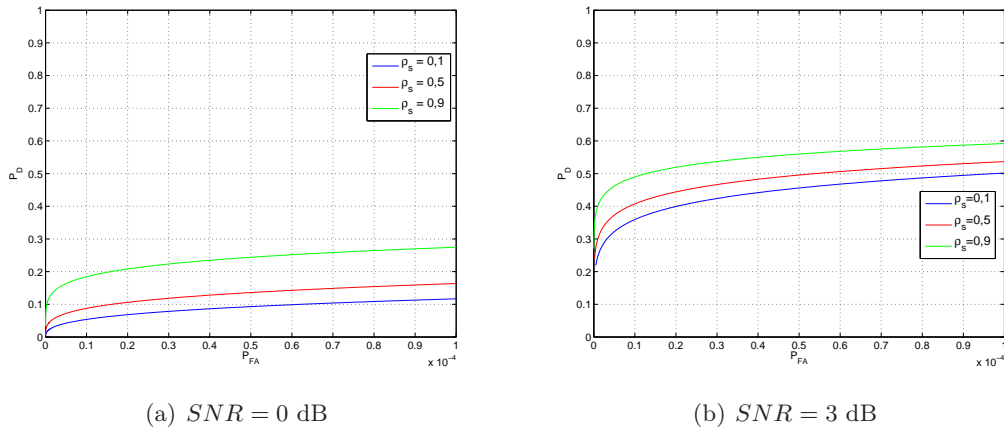


Figura 3.6: Curvas ROC del LR para el caso de estudio 1. $\Omega = 0$.

para el diseño de \mathbf{Q} . Debido a la ausencia de clutter, se utiliza la relación señal a ruido, definida en (3.28), para relacionar las potencias de blanco e interferencia.

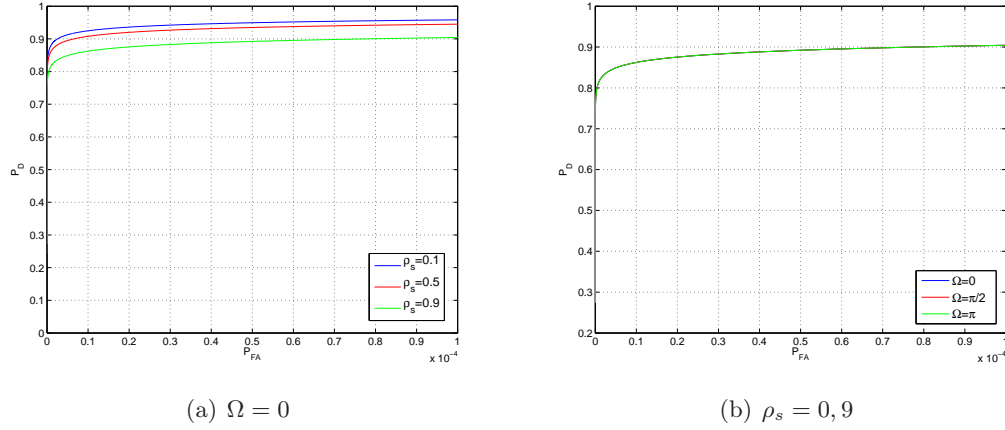
Los resultados de las Figuras 3.6(a), 3.6(b) y 3.7(a) asumen una frecuencia Doppler normalizada igual a cero. Cada una de las gráficas corresponde a una $SNR = 0, 3$ y 7 dB respectivamente. En cada gráfica se comparan los resultados obtenidos para coeficientes de correlación $\rho_s = 0, 1; 0, 5$ y $0, 9$. Según la SNR aumenta la P_D para una misma P_{FA} también aumenta, independientemente del valor de ρ_s . Para valores de SNR que proporcionan baja P_D , las capacidades de detección son mejores para valores altos de ρ_s , pero cuando el valor de SNR elegido proporciona alta P_D , los mejores resultados se obtienen para valores bajos de ρ_s . Esta inversión en las capacidades de detección, también se produce en las curvas ROC del detector de envolvente [Eaves1987].

Como el ruido es blanco y gaussiano, el valor de Ω seleccionado no afecta a las capacidades de detección. Este resultado se comprueba en la Figura 3.7(b), donde se representa la curva ROC para una $SNR = 7$ dB, un $\rho_s = 0, 9$ y diferentes valores de Ω . Se ha elegido un blanco bastante correlado, porque las propiedades del espectro son muy dependientes del valor de Ω , como se presenta en la Figura 3.5(b).

3.4.2. Caso de estudio 2: Detector NP de parámetros conocidos

En los sistemas radar, se recibe, a parte de ruido, ecos no deseados debidos al clutter. En radares coherentes de baja resolución o con un ángulo de incidencia rasante elevado, la variabilidad de la amplitud de estos ecos no deseados se modela con una distribución Rayleigh y una variabilidad de fase uniforme entre $[0, 2\pi)$. En esta sección vamos a plantear el detector de Neyman-Pearson para detectar blancos gaussianos con función de autocorrelación gaussiana en presencia de ruido blanco y gaussiano más clutter gaussiano con función de autocorrelación gaussiana. Este caso de estudio, para parámetros conocidos del blanco, es abordado en [Aloisio1994] y [diVito1999].

El vector de observaciones bajo cada una de las hipótesis es el siguiente:


 Figura 3.7: Curvas ROC del LR para el caso de estudio 1. $SNR = 7$ dB.

$$H_0 : \tilde{\mathbf{z}} = \tilde{\mathbf{n}} + \tilde{\mathbf{c}} \quad (3.38)$$

$$H_1 : \tilde{\mathbf{z}} = \tilde{\mathbf{s}} + \tilde{\mathbf{n}} + \tilde{\mathbf{c}} \quad (3.39)$$

Bajo hipótesis H_0 , el vector de observación es la suma de dos procesos gaussianos, por lo tanto, $\tilde{\mathbf{z}}$ será un proceso gaussiano de media cero y matriz de covarianza $\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2\mathbf{I}$, cuya FDP es:

$$f(\tilde{\mathbf{z}}/H_0) = \frac{1}{\pi^P |\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2\mathbf{I}|} \exp(-\tilde{\mathbf{z}}^T (\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2\mathbf{I})^{-1} \tilde{\mathbf{z}}^*) \quad (3.40)$$

Análogamente, el vector de observación bajo hipótesis H_1 , resultado de la suma de tres procesos gaussianos, será un proceso gaussiano de media cero y matriz de covarianza $\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2\mathbf{I}$, cuya FDP es:

$$f(\tilde{\mathbf{z}}/H_1) = \frac{1}{\pi^P |\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2\mathbf{I}|} \exp(-\tilde{\mathbf{z}}^T (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2\mathbf{I})^{-1} \tilde{\mathbf{z}}^*) \quad (3.41)$$

El detector del Neyman-Pearson se puede implementar comparando el cociente de verosimilitud, LR, a un umbral atendiendo a los requisitos de P_{FA} . A partir del LR, aplicando logaritmos y despejando:

$$\tilde{\mathbf{z}}^T \mathbf{Q} \tilde{\mathbf{z}}^* \underset{H_0}{\overset{H_1}{\gtrless}} \text{Ln} \left(\eta_{cv}(P_{FA}) \frac{|\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I}|}{|\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I}|} \right) = \eta_s \quad (3.42)$$

donde $\mathbf{Q} = (\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I})^{-1}$, y los umbrales η_{cv} y η_s se determinan en función de la P_{FA} elegida. La regla (3.42) representa un detector que sólo será óptimo cuando la matriz de covarianza del vector de observación coincida con la asumida en el diseño del detector basado en el logaritmo del LR.

La FDP de q se puede obtener con la expresión (3.43) a partir de la transformada de Fourier de q descrita en la ecuación (3.36).

$$f_q(q) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_q(j\omega) \exp(-j\omega q) d\omega \quad (3.43)$$

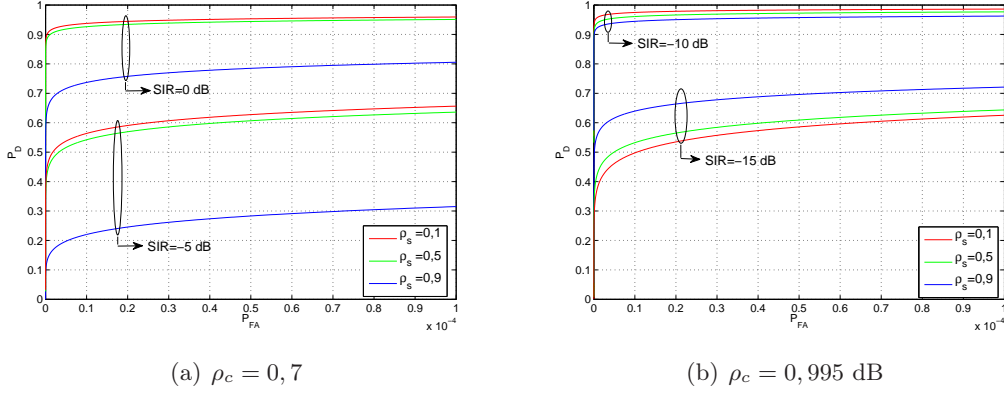


Figura 3.8: Curvas ROC del LR del caso de estudio 2. $CNR = 20$ dB y $\Omega = \pi/2$.

Entonces, la probabilidad de superar un umbral T se expresa en la ecuación (3.37), donde $\lambda(k)$ son los autovalores de $\Sigma_{\tilde{z}\tilde{z}}^* \mathbf{Q}^*$:

- Si $\Sigma_{\tilde{z}\tilde{z}} = \Sigma_{\tilde{c}\tilde{c}} + 2\mathbf{I}$, es decir, se calcula bajo H_0 , la ecuación (3.37) proporciona valores de P_{FA} .
- Si $\Sigma_{\tilde{z}\tilde{z}} = \Sigma_{\tilde{s}\tilde{s}} + \Sigma_{\tilde{c}\tilde{c}} + 2\mathbf{I}$, bajo H_1 , la ecuación (3.37) proporciona valores de P_D .

En [Aloisio1994] y [diVito1999], se proponen valores para caracterizar los vectores de observación bajo hipótesis H_0 y H_1 que utilizaremos a continuación para representar las curvas ROC del detector óptimo.

Primero estudiaremos un escenario en el que el clutter no predomine claramente sobre el ruido ($CNR = 20$ dB). En la Figura 3.8 se representan curvas ROC del detector óptimo asumiendo que los parámetros son conocidos y se observa la dependencia de las características de detección con el coeficiente de correlación del clutter, la relación SIR y el coeficiente de correlación del blanco. Se ha considerado que $\Omega = \pi/2$, lo que explica que a medida que el clutter es más correlado mejora considerablemente las probabilidades de detección. Para conseguir P_D en torno a 0,9 para P_{FA} de 10^{-4} es necesario una $SIR = -10$ dB para un $\rho_c = 0,995$ mientras que para $\rho_c = 0,7$ hay que aumentar ese valor 10 dB, $SIR = 0$ dB.

La dependencia de P_D con ρ_s es menor en presencia de un clutter más correlado. En la Figura 3.8(a), se observa una probabilidad de detección muy baja para blancos más correlados independientemente de la SIR considerada. Si $\rho_c = 0,995$, hay una inversión en la detección en función de ρ_s (Figura 3.8(b)), de manera que para valores de SIR que proporcionan P_D altas se detecta mejor blancos poco correlados, y para valores de SIR que proporcionan P_D bajas se detecta mejor blancos muy correlados, al igual que pasaba en la detección de blancos fluctuantes en presencia de AWGN.

Los resultados de la Figura 3.9, demuestran que cuando el espectro del blanco y del clutter están centrados a la misma frecuencia la P_D decrece considerablemente. Si el clutter está muy correlado, la energía está muy concentrada en pocas frecuencias, por lo que la dependencia del detector con Ω es de tipo escalón, la P_D crece significativamente para los primeros valores de $\Omega > 0$, pero, a partir de $\Omega = \pi/2$, el crecimiento disminuye (Figura 3.9(b)). En un clutter menos

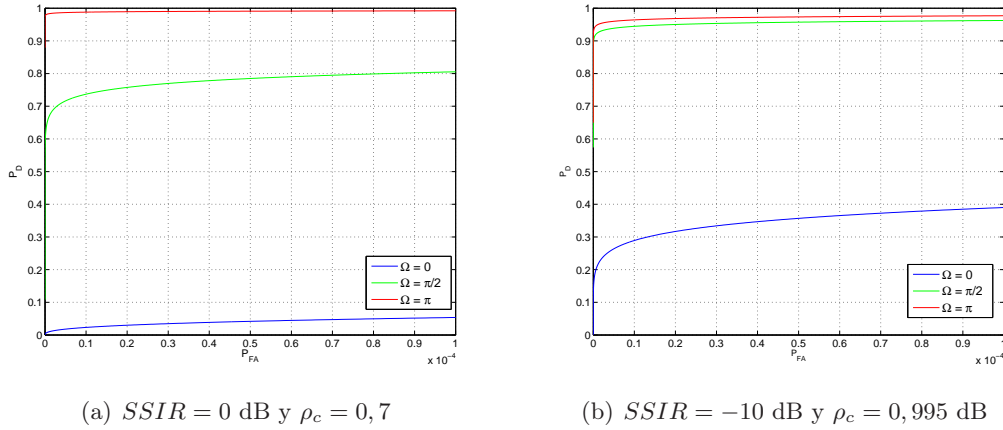


Figura 3.9: Curvas ROC del LR para el caso de estudio 2. $CNR = 20$ dB y $\rho_s = 0,9$.

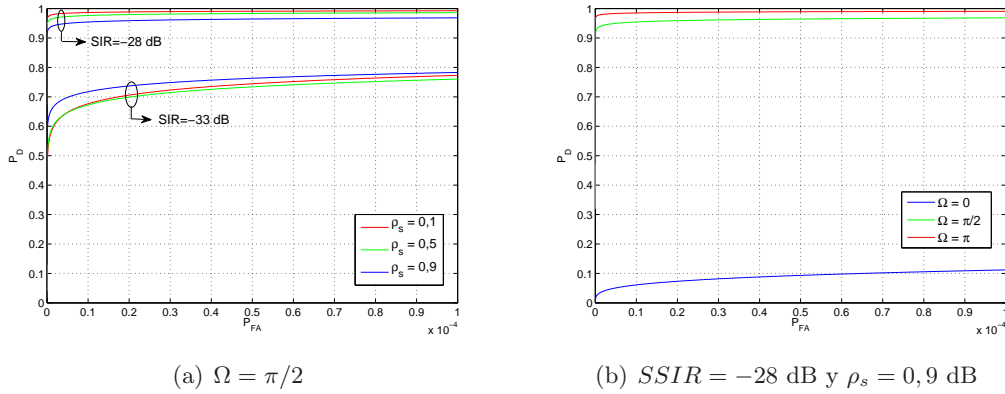


Figura 3.10: Curvas ROC del LR para el caso de estudio 2. $CNR = 40$ dB y $\rho_c = 0,995$.

correlado, el crecimiento es más lineal porque el espectro del clutter se encuentra repartido en más frecuencias, presentando los mejores resultados para $\Omega = \pi$ (Figura 3.9(a)).

En un escenario donde la potencia del clutter hace despreciable el efecto del ruido blanco gaussiano, se utilizan valores de coeficientes de correlación muy altos para el clutter [Aloisio1994] [diVito1999]. En la Figura 3.10, se representan las curvas ROC para detectores LR diseñados con $CNR = 40$ dB y $\rho_c = 0,995$. Las características de detección son análogas a las obtenidas para un detector diseñado con $CNR = 20$ dB y $\rho_c = 0,995$ (Figuras 5.4(b) y 5.5(b)): se produce una inversión en el tipo de blancos Swerling en función de la $SSIR$ y el comportamiento de la detección de blancos muy correlados presenta un efecto escalón en función del valor de Ω .

Los resultados mostrados para los diferentes modelos de clutter gaussiano, permiten deducir que las capacidades de detección de un detector óptimo diseñado para detectar blancos fluctuantes en interferencia gaussiana depende más del coeficiente de correlación del clutter considerado en el diseño que de CNR .

3.4.3. Caso de estudio 3: Detector NP de parámetros conocidos

En esta sección se estudiará el detector coherente óptimo basado en el criterio de Neyman-Pearson para blancos no fluctuantes en presencia de clutter de distribución K. En [Conte1991], [Pentini1992] y [Conte1994], se presenta la expresión del detector basado en el cociente de verosimilitud, LR, cuando la amplitud y la fase del blanco son conocidas y diferentes aproximaciones subóptimas del cociente de verosimilitud promediado, ALR, cuando la fase del blanco es desconocida. En esta tesis, también se tendrá en cuenta el movimiento del blanco con respecto al clutter. El escenario presentado en estas referencias consideran un clutter dominante donde el efecto del ruido es despreciable y definen la relación de potencias del blanco y la interferencia a través de la relación señal a clutter (3.29).

$$H_0 : \tilde{\mathbf{z}} = \tilde{\mathbf{c}} \quad (3.44)$$

$$H_1 : \tilde{\mathbf{z}} = \tilde{\mathbf{s}} + \tilde{\mathbf{c}} \quad (3.45)$$

Los valores que se suelen tomar para el parámetro de forma del clutter, ν , son valores muy pequeños, $\nu \leq 0,5$, para estudiar las capacidades de detección en presencia de un clutter muy impulsivo, y un valor $\nu = \infty$, para el que se puede aproximar el clutter de distribución K a un modelo gaussiano. Como en el apartado 3.4.2 se estudia el clutter gaussiano, en este caso de estudio se asumirá el modelo de clutter impulsivo con un valor de ν igual a 0,5.

En la sección 3.3.1, se presenta la teoría SIRP en la que se basa el modelo de clutter no gaussiano. El clutter de distribución K presenta una matriz de correlación gaussiana, $E[s^2]\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}$, que describe la correlación entre ecos de una misma exploración, considerando que de exploración a exploración, las variables son independientes y están idénticamente distribuidas (IID) [Conte1991]. Las exploraciones se ven *moduladas* por una variable aleatoria, s^2 , definida positiva y que sigue una distribución gamma. El valor de $E[s^2]$ se asume igual a la unidad sin pérdida de generalidad porque el estudio se basa en la relación de potencia entre el blanco y el clutter según la expresión (3.29). La FDP de $\tilde{\mathbf{z}}$ bajo hipótesis H_0 , a partir de (3.21):

$$f(\tilde{\mathbf{z}}|H_0) = \frac{\nu^P 2^{1-\nu+P}}{(\pi)^P |\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}|\Gamma(\nu)} b^{\nu-P} K_{\nu-P}(b) \quad (3.46)$$

donde $b = \sqrt{4\nu(\tilde{\mathbf{z}}^T \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}^{-1} \tilde{\mathbf{z}}^*)}$.

Suponiendo todos los parámetros del blanco conocidos, la FDP bajo hipótesis H_1 :

$$f(\tilde{\mathbf{z}}|H_1) = f((\tilde{\mathbf{z}} - \tilde{\mathbf{s}})|H_0) = \frac{\nu^P 2^{1-\nu+P}}{(\pi)^P |\Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}|\Gamma(\nu)} a^{\nu-P} K_{\nu-P}(a) \quad (3.47)$$

donde $a = \sqrt{4\nu((\tilde{\mathbf{z}} - \tilde{\mathbf{s}})^T \Sigma_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}}^{-1} (\tilde{\mathbf{z}} - \tilde{\mathbf{s}})^*)}$.

Al implementar el detector óptimo a partir del cociente de verosimilitud, se obtiene la regla de decisión (3.48), la cual depende de parámetros del blanco, SCR , φ y Ω , y del clutter, ρ_c y ν . En la Figura 3.11, se estudia el comportamiento del detector óptimo en función de distintas combinaciones del conjunto de parámetros que definen ambas hipótesis. Sólo se ha considerado el caso de un clutter impulsivo con $\nu = 0,5$. En la Figura 3.11(a) se asume un clutter incorrelado, mientras que los resultados para un clutter muy correlado, $\rho_c = 0,9$, se representan en la Figura

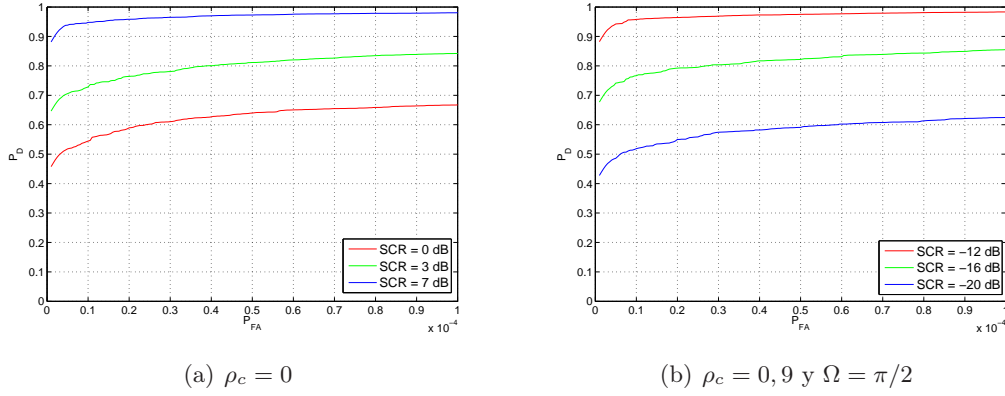


Figura 3.11: Curvas ROC del LR para el caso de estudio 3. $\nu = 0,5$ y $\theta_R = \pi/4$.

3.11(b). Efectivamente, cuando el clutter está muy correlado y la información del blanco se encuentra en una zona del espectro diferente a la del clutter, $\Omega \neq 0$, la *SCR* necesaria para obtener buenas características de detección son mucho menores.

$$\frac{a^{\nu-P} K_{\nu-P}(a)}{b^{\nu-P} K_{\nu-P}(b)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{cv} \quad (3.48)$$

Capítulo 4

Inteligencia artificial y el detector de Neyman-Pearson

4.1. Introducción

Una manera concisa de definir la inteligencia artificial es el estudio de cómo hacer que los procesadores hagan cosas que, hasta el momento, hace mejor el ser humano [Rich1983]. Para ello se utiliza el diseño de agentes inteligentes, entendiéndose como tales a los sistemas capaces de percibir su entorno (recibir entradas) y realizar acciones (proporcionar salidas) con el objetivo de maximizar sus probabilidades de éxito [Ertel2011]. Los agentes establecen una asignación directa entre percepciones y acciones a tomar, que generalmente depende de una gran cantidad de información aprendida. Por lo tanto, un sistema basado en inteligencia artificial tiene que ser capaz de adquirir información del entorno, aprender de la información adquirida e inferir las acciones consecuentes. Dentro del aprendizaje se incluye la adquisición de habilidades matemáticas que permitan la generalización de la información utilizada durante el entrenamiento de los agentes inteligentes para conseguir una determinada función objetivo. Un agente inteligente es capaz de aprender de forma automática a través de la experiencia. Entre los agentes inteligentes, se encuentran las redes neuronales artificiales, cuyo objetivo es el de emular el funcionamiento físico del cerebro de animales y humanos. Las redes neuronales pueden definirse del modo siguiente [Haykin1999]:

Procesador masivo paralelo distribuido, formado por unidades simples, con una propensión natural a almacenar conocimiento experimental y hacerlo disponible para ser usado. Se asemeja al cerebro en dos aspectos:

- *La red adquiere el conocimiento del entorno a través de un proceso de aprendizaje.*
- *La intensidad de las conexiones interneuronales, conocidas como pesos sinápticos, se emplean para almacenar el conocimiento adquirido.*

Análisis estadísticos de su comportamiento han revelado su versatilidad y potencial en tareas de detección, clasificación y, en general, de aproximación de funciones. En las redes neuronales, las neuronas se organizan formando capas. Atendiendo a cómo se conectan entre sí las neuronas,

las redes pueden ser de propagación directa o recurrentes. En las redes de propagación directa, la información fluye unidireccionalmente de una capa a otra (desde la capa de entrada a las capas ocultas y de éstas a la capa de salida) y, además, no se admiten conexiones intracapa. En las redes recurrentes, la información puede formar bucles y pueden establecerse conexiones intracapa, incluso de una unidad consigo misma. Las conexiones entre una capa y otra pueden ser totales, es decir, que cada unidad se conecta con todas las unidades de la capa siguiente, o parciales, en las cuales una unidad se conecta con sólo algunas de las unidades de la capa siguiente.

En esta tesis doctoral nos hemos centrado en redes neuronales de propagación directa con neuronas artificiales con funciones de base de primer orden, cuadráticas y radiales. El aprendizaje se produce mediante la variación de los parámetros asociados a sus unidades, de acuerdo a una estrategia o algoritmo de aprendizaje, con la finalidad de conseguir un objetivo de diseño. La arquitectura, el tipo y la efectividad de las conexiones representan, en cierto modo, la memoria o el conocimiento adquirido por la red.

Otro tipo de agentes inteligentes de aprendizaje supervisado son las máquinas de vectores soporte (SVM), que se suelen aplicar a tareas de clasificación y regresión y cuya teoría se basa en la minimización del riesgo estructural, R_{est} , expresado en la ecuación (4.1), donde $F(z)$ es la función que aproxima la SVM, z es el patrón de entrada, R_{emp} representa la función de error o precisión evaluada sobre un conjunto particular de datos de entrenamiento y $\lambda Q(F(z))$ es el parámetro de regularización, que permite la capacidad de generalizar para otros conjuntos de patrones de entrada.

$$R_{est}(F(z)) = R_{emp}(F(z)) + \lambda Q(F(z)) \quad (4.1)$$

El estudio del entrenamiento de las SVM puede ser abordado a través de la teoría clásica de la regularización, al igual que las redes neuronales con funciones de base radial [Burges1998]. Una SVM mapea los datos de entrada a un espacio Hilbert de características, generalmente, de dimensión superior. La transformación al nuevo espacio de características se hace a través de *funciones kernel*. La SVM encuentra un hiperplano lineal en el espacio de dimensión mayor que clasifique los datos de entrada maximizando el margen entre las clases. La solución del hiperplano óptimo puede ser escrita como la combinación de unos pocos datos de entrada que son llamados *vectores soporte*.

De acuerdo a la Teoría de la Detección, la formulación de la regla de decisión óptima para la realización de un detector de Neyman-PEarson o un clasificador de mínima probabilidad de error requiere de una caracterización estadística de las clases consideradas. Supóngase un problema de clasificación en el que se cumple que las probabilidades a priori de las clases son iguales y los costes asociados al acierto valen cero, mientras que los asociados a la equivocación valen uno. Si, dado un patrón, el objetivo es minimizar la probabilidad de error, debe ser asignado a la clase cuya probabilidad a posteriori sea mayor (criterio *MAP*, del inglés: “*Maximum a Posteriori Probability*”). Ante este problema, son varias las posibles estrategias a seguir:

- Estimar las distintas funciones de densidad de probabilidad que permitan caracterizar probabilísticamente el problema y diseñar el clasificador óptimo.
- Para evitar la necesidad de estimar funciones de densidad de probabilidad, pueden utilizarse

funciones discriminantes parametrizadas y determinar el valor de los parámetros que las caracterizan mediante un proceso de aprendizaje a partir de un conjunto de datos pre-clasificados.

El principio de funcionamiento de las soluciones basadas en inteligencia artificial está íntimamente ligado a la segunda opción. El proceso de aprendizaje o entrenamiento pretende aproximar las funciones discriminantes a partir de la minimización del riesgo estructural y, por lo tanto, del riesgo empírico. La minimización del riesgo empírico depende de la función de coste asociada al error o, de forma abreviada, función de error.

A parte de la función de error, la función que aproxima los agentes inteligentes dependen de los conjuntos de entrenamiento y de test diseñados, de la estructura del sistema y del algoritmo de aprendizaje. El conjunto de entrenamiento debe ser un subconjunto representativo del espacio de entrada para poder utilizarlo en el algoritmo de aprendizaje que trata de encontrar los parámetros adaptativos del agente inteligente que minimizan la función de error. En este sentido, en [Saerens2002], los autores afirman que un sistema de aprendizaje adaptativo es capaz de llegar a una aproximación de la función discriminante si se consigue un *entrenamiento perfecto*, entendido como un entrenamiento capaz de conseguir el mínimo de la función de error considerada, y el agente inteligente es un *modelo suficientemente potente* que es capaz de aproximar la función discriminante para cualquier grado de exactitud.

Dada una regla de decisión, llamaremos \mathcal{Z}_i al conjunto de vectores de entrada que serán asignados a la hipótesis H_i , y \mathcal{Z} al conjunto de todos los vectores de entrada o espacio de observaciones ($\mathcal{Z} = \mathcal{Z}_0 \cup \mathcal{Z}_1$). $F(\mathbf{z})$ será la salida del sistema adaptativo, y t_{H_i} la salida deseada. Por último, el conjunto de entrenamiento será denominado $\mathcal{Z} = \mathcal{Z}_0 \cup \mathcal{Z}_1$, donde \mathcal{Z}_1 está compuesto por N_1 patrones asociados a la hipótesis H_1 , y \mathcal{Z}_0 está compuesto por N_0 patrones pertenecientes a la hipótesis H_0 ($N = N_0 + N_1$).

En la sección 4.2 se presenta el esquema de detección basado en inteligencia artificial que se considera en esta tesis doctoral. Posteriormente, en la sección 4.3, se define la condición suficiente que tiene que cumplir la función discriminante aproximada por el agente inteligente para aproximar el detector NP. En los apartados 4.4, 4.5 y 4.6 se estudian distintas funciones de error: el error cuadrático medio, la entropía cruzada y el error de clasificación. Por, último, en el apartado 4.7, se presentan los agentes inteligentes utilizados en esta Tesis.

4.2. Esquema de detección

El primer paso a la hora de diseñar un detector es la caracterización de las señales recibidas. Si se trata de un problema de detección binaria, han de determinarse las funciones de densidad de probabilidad de las observaciones bajo las dos hipótesis de trabajo: H_0 y H_1 . En el caso de la detección de blancos radar y bajo hipótesis H_0 , el origen de las observaciones es el ruido del receptor o el clutter, mientras que bajo hipótesis H_1 las observaciones son el resultado de reflexiones producidas por blancos deseados que vienen contaminadas con interferencia aditiva de las mismas características que las indicadas para la hipótesis H_0 .

En los casos en los que se desconocen las funciones de verosimilitud del problema, dos son los posibles caminos a seguir:

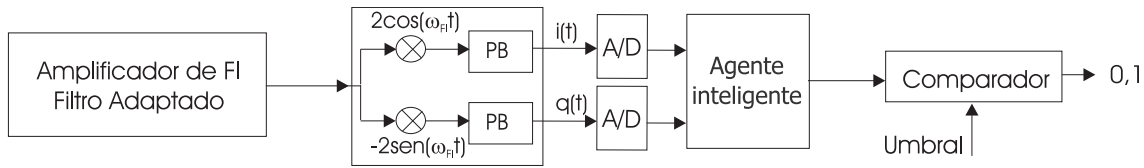


Figura 4.1: Esquema del detector neuronal propuesto

- Pueden estimarse las funciones de verosimilitud y diseñar el detector en base a estas estimaciones.
- Puede utilizarse una técnica no paramétrica, por ejemplo, utilizando soluciones basadas en inteligencia artificial.

En esta tesis se contempla el caso de la detección de blancos de parámetros desconocidos. La estimación de las funciones de verosimilitud conlleva la resolución de integrales que pueden tener una solución no cerrada. Por lo tanto, en esta tesis, se plantea tanto el estudio de soluciones subóptimas que aproximen el cociente de verosimilitud promediado, ALR, como soluciones basadas en inteligencia artificial que mantengan las mismas características de detección con menor complejidad computacional.

El detector que se considera en esta tesis es el que se presenta en la Figura 4.1. Las etapas previas al filtro de frecuencia intermedia se han obviado en el esquema al ser idénticas a las de la Figura 2.1. Aunque la etapa de digitalización en la Figura 4.1 se propone en banda base, se podría haber contemplado realizar la conversión analógico/digital a la salida del filtro de frecuencia intermedia sin alterar los resultados de los estudios realizados en la tesis.

Como ya se ha explicado en el apartado 2.3, a la salida del filtro de frecuencia intermedia, la señal bajo cualquiera de las dos hipótesis es de banda estrecha, por lo que sus componentes en fase y cuadratura pueden obtenerse con un detector síncrono [Papoulis1991].

Los detectores basados en redes neuronales o máquinas de vectores soporte se implementarán con aritmética real y las entradas serán muestras de las componentes en fase y cuadratura de la señal obtenida a la salida del filtro de frecuencia intermedia. Teniendo en cuenta que el sistema recibe P pulsos por blanco, para utilizar toda esta información en el proceso de detección, los detectores propuestos deberán tener $2P$ entradas (P componentes en fase y P componentes en cuadratura).

4.3. Condición suficiente para que una función discriminante pueda aproximar el detector de Neyman-Pearson

El detector de Neyman-Pearson se basa en comparar el cociente de verosimilitud o cualquier otra función discriminante equivalente con un umbral fijado en función de los requisitos de P_{FA} . En [Jarabo2005a][Jarabo2009], se propone un método para determinar si la función discriminante aproximada por el agente inteligente, cuando es comparada con un umbral de detección, η_0 , fijado por la P_{FA} deseada, permite implementar el detector de NP. El método consiste en los siguientes tres pasos:

- Reescribir la regla de decisión en función del cociente de verosimilitud.
- Despejar el cociente de verosimilitud para obtener una regla de decisión equivalente y determinar la relación entre el nuevo umbral de detección, η_{lr} , y el anterior, η_0 .
- Comprobar si la relación entre η_{lr} y η_0 depende o no del vector de entrada.

Si η_0 puede ser determinado a partir de cada pareja (η_{lr}, P_{FA}) y no depende del vector de entrada, la regla de decisión resultante de comparar la función discriminante con η_0 es un implementación del detector de NP.

En [Jarabo2009], este método es aplicado a tres expresiones equivalentes de la función discriminante óptima de Bayes:

$$P(H_1|\mathbf{z}) - P(H_0|\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \eta_0 \quad (4.2)$$

$$P(H_1|\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \eta_0 \quad (4.3)$$

$$P(H_1)f(\mathbf{z}|H_1) - P(H_0)f(\mathbf{z}|H_0) \underset{H_0}{\overset{H_1}{\geq}} \eta_0 \quad (4.4)$$

Para despejar el cociente de verosimilitud de estas expresiones, debe tenerse en cuenta lo siguiente:

$$P(H_i|\mathbf{z}) = \frac{f(\mathbf{z}|H_i)P(H_i)}{f(\mathbf{z})} \quad i = 0, 1 \quad (4.5)$$

donde $f(\mathbf{z}) = f(\mathbf{z}|H_1)P(H_1) + f(\mathbf{z}|H_0)P(H_0)$ y $P(H_i|\mathbf{z})$ es la probabilidad a posteriori de las clases. Operando en las expresiones (4.2), (4.3) y (4.4):

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)(1 + \eta_0)}{P(H_1)(1 - \eta_0)} = \eta_{lr} \quad (4.6)$$

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)\eta_0}{P(H_1)(1 - \eta_0)} = \eta_{lr} \quad (4.7)$$

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0) + \eta_0 f^{-1}(\mathbf{z}|H_0)}{P(H_1)} = \eta_{lr} \quad (4.8)$$

En las expresiones (4.6) y (4.7) se puede obtener un η_0 para cada η_{lr} fijado para cada P_{FA} , que no depende del vector de entrada, por lo tanto, estas reglas de decisión son implementaciones del detector NP. Sin embargo, en la regla expresada en (4.8), la relación entre η_0 y η_{lr} depende del vector de entrada. Si para cada valor de P_{FA} es posible encontrar, en función de las características del cociente de verosimilitud y la función de verosimilitud bajo hipótesis H_0 , un valor de η_0 independiente del vector de entrada, la regla de decisión (4.8) será una realización del detector de NP. Por ejemplo, en el caso de la detección de blancos Swerling II en ruido blanco y gaussiano, es posible encontrar un umbral independiente del patrón de entrada [Jarabo2009]. Si no es posible encontrar un valor de η_0 para cada valor de P_{FA} independiente del vector de entrada, podemos

variar la P_{FA} ajustando el valor de η_0 , pero la correspondiente P_D no tiene porque ser máxima a diferencia de la P_D que obtendríamos con el detector NP. Por lo tanto, la condición que se ha comprobado en una condición suficiente pero no necesaria.

4.4. Función de error: el cuadrático medio

Si E es el error estimado sobre los N datos pre-clasificados que forman el conjunto de entrenamiento entre la salida de la red, $F(\mathbf{z})$, y las salidas deseadas, el error cuadrático medio o superficie de error del problema de clasificación a resolver puede estimarse como se indica en (4.9).

$$E = \frac{1}{N} \sum_{i=0}^1 \sum_{\forall \mathbf{z} \in \mathcal{Z}_i} |F(\mathbf{z}) - t_{H_i}|^2 \quad (4.9)$$

Si el número de patrones tiende a infinito ($N \rightarrow \infty$), el error se expresa en (4.10) y, aplicando la Ley de los Grandes Números, se llega a la expresión (4.11).

$$E_m = \lim_{N \rightarrow \infty} \sum_{i=0}^1 \frac{N_i}{N} \frac{1}{N_i} \sum_{\forall \mathbf{z} \in \mathcal{Z}_i} |F(\mathbf{z}) - t_{H_i}|^2 \quad (4.10)$$

$$E_m = \sum_{i=0}^1 P(H_i) \int_{\mathcal{Z}} |F(\mathbf{z}) - t_{H_i}|^2 f(\mathbf{z}|H_i) d\mathbf{z} \quad (4.11)$$

La función $F(\mathbf{z})$ que minimiza E_m , que llamaremos $F_0(\mathbf{z})$, se puede obtener aplicando *cálculo de variaciones*, en particular, la ecuación diferencial de Euler-Lagrange [Gelfand2000] [Weinstock1952]. El cálculo de variaciones puede ser utilizado para encontrar la función $F(\mathbf{z})$ que minimiza el funcional $J(F)$:

$$J(F) = \int_{\mathcal{Z}} I \left(\mathbf{z}, F(\mathbf{z}), \frac{\partial F(\mathbf{z})}{\partial z_1}, \frac{\partial F(\mathbf{z})}{\partial z_2}, \dots, \frac{\partial F(\mathbf{z})}{\partial z_L} \right) d\mathbf{z} \quad (4.12)$$

donde $\mathbf{z} = [z_1, z_2, \dots, z_L]^T$, I es dos veces diferenciable con respecto a los argumentos indicados y F es una función que pertenece a $\mathcal{C}^2(\mathcal{Z})$ que asume los valores prescritos en todos los puntos de la frontera $\partial\mathcal{Z}$ en el dominio \mathcal{Z} . La función F que minimiza $J(F)$ se puede obtener resolviendo la ecuación de Euler-Lagrange (4.13), donde $F'_k = \partial F / \partial z_k$.

$$\frac{\partial I}{\partial F} - \sum_{k=1}^L \frac{\partial}{\partial z_k} \left(\frac{\partial I}{\partial F'_k} \right) = 0 \quad (4.13)$$

En nuestro problema:

$$J(F) = E_m(F) \quad (4.14)$$

$$I(\mathbf{z}, F(\mathbf{z})) = \sum_{i=0}^{\infty} P(H_i) |F(\mathbf{z}) - t_{H_i}|^2 f(\mathbf{z}|H_i) \quad (4.15)$$

Como $I(\mathbf{z}, F(\mathbf{z}))$ no depende de las primeras derivadas de F , sólo necesitamos definir F en $\mathcal{C}^0(\mathcal{Z})$ y la ecuación de Euler-Lagrange se reduce a (4.16) y la función $F_0(\mathbf{z})$ que minimiza E_m es (4.17).

$$\frac{\partial}{\partial F} \left(\sum_{i=0}^1 P(H_i) |F(\mathbf{z}) - t_{H_i}|^2 f(\mathbf{z}|H_i) \right) = 0 \quad (4.16)$$

$$F_0(\mathbf{z}) = \frac{P(H_1)f(\mathbf{z}|H_1)t_{H_1} + P(H_0)f(\mathbf{z}|H_0)t_{H_0}}{P(H_1)f(\mathbf{z}|H_1) + P(H_0)f(\mathbf{z}|H_0)} \quad (4.17)$$

Si η_0 es el umbral de detección correspondiente a la P_{FA} deseada, la regla decisión aproximada viene dada por (4.18).

$$F_0(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_1)f(\mathbf{z}|H_1)t_{H_1} + P(H_0)f(\mathbf{z}|H_0)t_{H_0}}{P(H_1)f(\mathbf{z}|H_1) + P(H_0)f(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \eta_0 \quad (4.18)$$

La regla (4.18) puede escribirse en función del cociente de verosimilitud:

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)(\eta_0 - t_{H_0})}{P(H_1)(t_{H_1} - \eta_0)} = \eta_r \quad (4.19)$$

La relación que existe entre el umbral correspondiente al cociente de verosimilitud, η_r , y el umbral requerido por la función $F_0(\mathbf{z})$ viene dada por (4.20).

$$\eta_0 = \frac{P(H_1)\eta_r t_{H_1} + P(H_0)t_{H_0}}{P(H_1)\eta_r + P(H_0)} \quad (4.20)$$

La regla anterior pone de manifiesto que la relación que existe entre η_r y η_0 no depende del vector de entrada, por lo tanto cumple la condición suficiente para que un agente inteligente entrenado para minimizar el error cuadrático medio pueda implementar el detector de NP.

D. W. Ruck y otros [Ruck1990] demostraron que una red neuronal del tipo perceptrón multicapa entrenado con el objetivo de minimizar el error cuadrático medio con salidas deseadas $t_{H_1} = 1$ y $t_{H_0} = -1$, utilizando la función tangente hiperbólica como función de activación de la neurona de salida, converge a una aproximación de una de las expresiones equivalentes de la función discriminante óptima de Bayes. Hasta la publicación de [Jarabo2005a], las redes se forzaban para trabajar en un punto subóptimo en el que al cambiar el umbral o la bias de la neurona de salida se podía ajustar la P_{FA} pero no la P_D .

4.5. Función de error: la entropía cruzada

A. El-Jaroudi y J. Makhoul, en [El-Jaroudi1990], proponen una función de error basada en la entropía cruzada entre las salidas deseadas y las proporcionadas por el agente inteligente, con el objetivo de maximizar la salida correspondiente a la clase a la que pertenece el patrón de entrada y minimizar las otras. Para el caso de clasificación entre dos clases basta con una única salida y la función de error se calcularía como:

$$E = -\frac{1}{N} \sum_{i=0}^1 \sum_{\mathbf{z} \in Z_i} \frac{1}{2} \log((F(\mathbf{z}) - \overline{t_{H_i}})^2) \quad (4.21)$$

donde $\overline{t_{H_i}}$ representa el complementario de t_{H_i} , es decir, si la salida deseada de la hipótesis H_i es $t_{H_i} = 1$, entonces $\overline{t_{H_i}} = 0$. Se ha demostrado que para que las salidas $F(\mathbf{z})$ aproximen las probabilidades a posteriori de ambas hipótesis utilizando como criterio de error la entropía cruzada es necesario que las salidas asociadas a cada clase sumen la unidad ($\sum_{i=0}^1 F_{H_i}(\mathbf{z}) = 1$) [Jarabo2005b].

Si el número de patrones de entrenamiento tiende a infinito y se aplica la Ley de los Grandes Números (4.22), donde se ha despreciado el factor $\frac{1}{2}$ que no es relevante en el proceso de optimización.

$$E_m = -\sum_{i=0}^1 P(H_i) \int_{\mathbf{z}} \log(F(\mathbf{z}) - \overline{t_{H_i}}) f(\mathbf{z}|H_i) d\mathbf{z} \quad (4.22)$$

Para encontrar la función $F_0(\mathbf{z})$ que minimiza E_m se resolverá la ecuación de Euler-Lagrange, definida en (4.13), aplicando cálculo de variaciones. Para ello, se iguala $E_m(F)$ al funcional $J(F)$ definido en (4.12) y se define, para este caso, $I(\mathbf{z}, F(\mathbf{z}))$ en (4.23), el cuál no depende de las primeras derivadas de F . La función $F_0(\mathbf{z})$ que resuelve la ecuación de Euler-Lagrange viene definida en (4.24):

$$I(\mathbf{z}, F(\mathbf{z})) = -\sum_{i=0}^1 P(H_i) \log(F(\mathbf{z}) - \overline{t_{H_i}}) f(\mathbf{z}|H_i) \quad (4.23)$$

$$F_0(\mathbf{z}) = \frac{P(H_1) f(\mathbf{z}|H_1) \overline{t_{H_0}} + P(H_0) f(\mathbf{z}|H_0) \overline{t_{H_1}}}{P(H_1) f(\mathbf{z}|H_1) + P(H_0) f(\mathbf{z}|H_0)} \quad (4.24)$$

Si comparamos la función discriminante aproximada con un umbral η_0 correspondiente a la P_{FA} deseada, la regla decisión resultante es (4.25) y al despejar el cociente de verosimilitud queda la regla de decisión equivalente expresada en (4.26).

$$F_0(\mathbf{z}) = \frac{P(H_1) f(\mathbf{z}|H_1) \overline{t_{H_0}} + P(H_0) f(\mathbf{z}|H_0) \overline{t_{H_1}}}{P(H_1) f(\mathbf{z}|H_1) + P(H_0) f(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \eta_0 \quad (4.25)$$

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)(\eta_0 - \overline{t_{H_1}})}{P(H_1)(\overline{t_{H_0}} - \eta_0)} = \eta_{lr} \quad (4.26)$$

La regla (4.26) demuestra que la relación que existe entre η_{lr} y η_0 no depende del vector de entrada y, por lo tanto, la función discriminante aproximada por un agente inteligente entrenado para minimizar la función de error (4.21) basada en la entropía cruzada, cumple la condición suficiente para ser una posible realización del detector de NP.

4.6. Función de error: el error de clasificación

Para problemas de clasificación con dos clases, J. Shawe-Taylor y N. Cristianini [Cristianini2000] propusieron como función de error para el aprendizaje de máquinas de vectores soporte una estimación de error de clasificación sobre los patrones del conjunto de entrenamiento:

$$E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 - t_{H_0} F(\mathbf{z}))) + \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - t_{H_1} F(\mathbf{z}))) \quad (4.27)$$

donde $u(\cdot)$ es la función escalón. Las etiquetas correspondientes a las dos clases se consideran igual a $t_{H_0} = -1$ y $t_{H_1} = 1$, de manera que los casos que pueden surgir son:

- $F(\mathbf{z}) < -1$, entonces si \mathbf{z} pertenece a la hipótesis H_0 estará correctamente clasificado y la salida de la función escalón será cero, mientras que si pertenece a la hipótesis H_1 consistirá en un error que hay que minimizar, quedando $E(\mathbf{z}) = \frac{1}{N} \sum (u(1 - F(\mathbf{z})))$.
- $-1 < F(\mathbf{z}) < 1$, corresponden a patrones de clasificación dudosa de manera que se penaliza el error en función de la clase a la que corresponda para que la salida de $F(\mathbf{z})$ sea -1 ó 1 :

$$E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 + F(\mathbf{z}))) + \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - F(\mathbf{z}))).$$
- $F(\mathbf{z}) > 1$, se tratará de un error si \mathbf{z} pertenece a la hipótesis H_0 , de manera que la función de error se simplifica a $E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 + F(\mathbf{z})))$.

Esta función de error se utiliza en el entrenamiento de SVMs aplicadas a tareas de clasificación. En esta tesis se pretende comprobar si función discriminante aproximada por una SVM o cualquier otro agente inteligente, tras un *entrenamiento perfecto* supervisado para minimizar este error de clasificación, cumple la condición suficiente de aproximar el detector NP.

4.7. Agentes inteligentes

En este apartado se presentan los diferentes tipos de agentes inteligentes considerados en esta tesis: redes neuronales artificiales con funciones base lineal (*Multilayer Perceptron*, MLP), redes neuronales con funciones base cuadráticas (*Second Order Neural Network*, SONN), redes neuronales con funciones base radial (*Radial Basis Function Neural Network*, RBFNN) y máquinas de vectores soporte (*Support Vector Machine*, SVM).

Las redes neuronales artificiales, cuyo objetivo es el de emular el funcionamiento físico del cerebro de animales y humanos, son capaces de aproximar el clasificador bayesiano óptimo [Ruck1990] (el que minimiza la probabilidad media de error). En 2009, también se demostró teóricamente la posibilidad de aproximar el detector óptimo de Neyman-Pearson [Jarabo2009].

Las redes que se presentan se caracterizan por ser de propagación directa y tener las neuronas artificiales organizadas en capas. Dado que el objetivo es decidir ante la presencia o ausencia de blanco, los agentes inteligentes tendrán una única salida que se comparará con un umbral elegido para cumplir los requisitos de P_{FA} . En la Figura 4.2 se presenta la estructura de un detector basado en redes neuronales, donde el detector neuronal se puede sustituir por otro agente inteligente.

En el caso de las máquinas de vectores soporte, se han presentado estrategias de aprendizaje y simulación para aproximar el criterio de detección de NP para distintas P_{FA} [Bach2006] y [Davenport2010]. Sin embargo, no hay ningún desarrollo matemático que demuestre la capacidad de un detector basado en una única SVM de aproximar el detector NP. En esta tesis, se realiza un estudio teórico sobre la idoneidad de este tipo agentes inteligentes en aplicaciones de detección radar.

Todos los agentes inteligentes trabajan con aritmética real, de manera que si se asume que el radar es capaz de recibir P pulsos por exploración de un blanco, se requiere $2P$ entradas correspondientes a las componentes en fase y cuadratura de cada uno de los P pulsos.

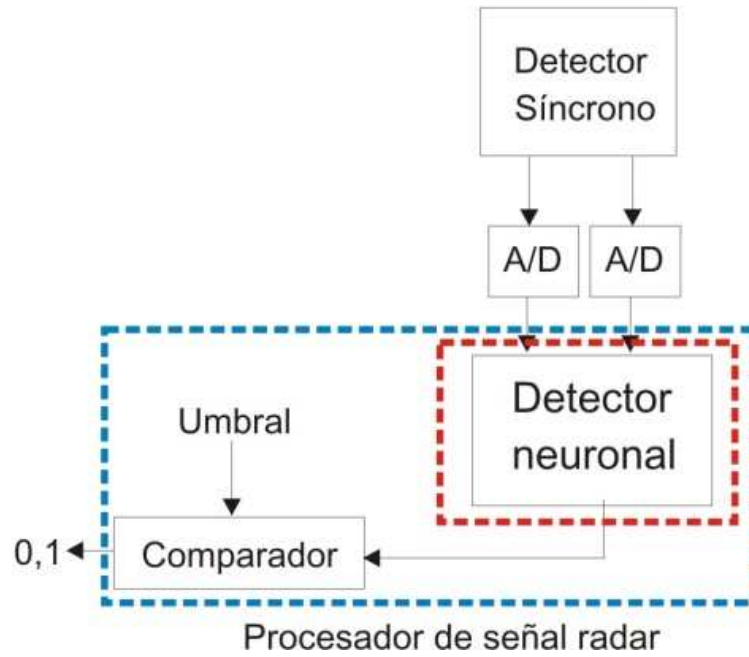


Figura 4.2: Detector neuronal coherente.

4.7.1. Perceptrones multicapa, MLP

El *perceptron* fue propuesto por Rosenblatt [Rosenblatt1959] como el primer modelo de sistema capaz de aprender mediante un entrenamiento supervisado. Los perceptrones multicapa son redes neuronales formadas por más de una capa de neuronas con funciones de base lineal y funciones de activación de tipo umbral o sigmoide. Las neuronas están agrupadas en capas conectadas de forma directa. Al tratarse de una red de propagación directa, sus salidas pueden calcularse como funciones explícitas de las entradas y los pesos. Además, si las funciones de activación son diferenciables, puede aplicarse el método de retropropagación del error [Rumelhart1986] para ajustar los valores de los pesos.

Matemáticamente, una red de $2P$ entradas, L capas con M_k , $k = 1, \dots, L$ neuronas en la capa k -ésima y, por tanto, M_L salidas, representa una función $F(\mathbf{z}) : \mathbb{R}^M \rightarrow \mathbb{R}^{M_L}$, donde \mathbf{z} es el vector de entrada dado por: $\mathbf{z} = [z_1, z_2, \dots, z_{2P}]^T$.

El esquema básico de una red multicapa es el que se representa en la Figura 4.3, en el que puede observarse que se trata de una red de propagación directa totalmente interconectada:

- $g_k(\cdot)$, $k = 1, \dots, L$ es la función de activación de las neuronas de la capa k -ésima.
- En la capa k -ésima, $\omega_{ij}^{(k)}$, $i = 1, \dots, M_k$, $j = 1, \dots, M_{k-1}$, es el peso de la rama que conecta la salida de la j -ésima neurona de la capa anterior, $k - 1$, con la neurona i -ésima de la capa considerada, k . Por tanto puede definirse el vector de pesos de la neurona i -ésima de la capa k como:

$$\mathbf{w}_i^{(k)} = [\omega_{i1}^{(k)}, \omega_{i2}^{(k)}, \dots, \omega_{iM_{k-1}}^{(k)}]^T \quad (4.28)$$

Nótese que si $k = 1$, $M_{k-1} = 2P$, siendo $2P$ el número de entradas de la red.

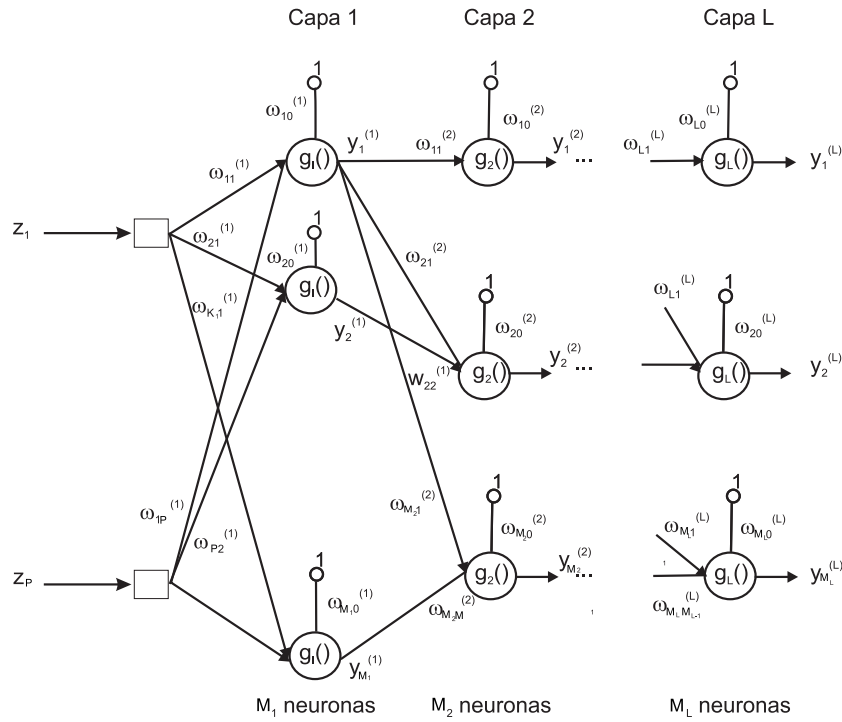


Figura 4.3: Esquema de una red multicapa de propagación directa

- $y_i^{(k)}$, $i = 1, \dots, M_k$, es la salida de la i -ésima neurona de la capa k . Así, el vector de las salidas de la capa k , $\mathbf{y}^{(k)}$ se define como:

$$\mathbf{y}^{(k)} = [y_1^{(k)}, y_2^{(k)}, \dots, y_{M_k}^{(k)}]^T \quad (4.29)$$

En el caso de una red con una única capa oculta, $L = 2$, la salida de la i -ésima neurona de la segunda capa puede expresarse como:

$$y_i^{(2)} = g_2 \left(\sum_{j=0}^{M_1} \omega_{ij}^{(2)} g_1 \left(\sum_{m=0}^{2P} z_m \omega_{jm}^{(1)} \right) \right) \quad (4.30)$$

El teorema de la aproximación universal establece que un perceptrón de una sola capa oculta puede calcular una aproximación uniforme a un conjunto de entrenamiento representado por un conjunto de entradas y las correspondientes salidas deseadas. Las funciones de activación deben ser no constantes, continuas, monótonas crecientes y estar acotadas, condiciones que cumplen las funciones tipo sigmoide. G. Cybenko [Cybenko1989] fue el primero que demostró rigurosamente que con una única capa oculta puede aproximarse uniformemente cualquier función continua con soporte en el hipercubo unidad. Otros autores han estudiado esta propiedad, por ejemplo [Hecht-Nielsen1989]. El corolario que puede extraerse en el contexto de un problema de clasificación es que redes con dos capas con funciones de activación de tipo sigmoide pueden aproximar cualquier frontera de decisión con una precisión arbitraria, siempre que el número de neuronas en la capa oculta sea suficientemente grande. Por tanto, pueden proporcionar funciones discriminante no lineales universales y pueden modelar las probabilidades a posteriori de las clases.

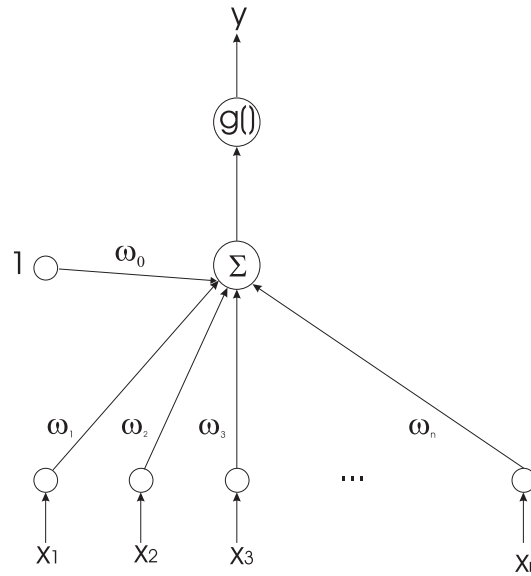


Figura 4.4: Modelo de una neurona artificial con función de base lineal

En [Jarabo2009] se demostró que estas funciones discriminantes son equivalentes al cociente de verosimilitud si la función de error objetivo es el error cuadrático medio o el error de entropía.

De entre las funciones tipo sigmoide, las más utilizadas son la logística y la tangente hiperbólica, expresiones (4.31) y (4.32) respectivamente. La utilización de funciones de activación logísticas permite interpretar las salidas de la red neuronal como probabilidades a posteriori de las clases.

$$g_k(x) = \frac{1}{1 + \exp(-ax)} \quad (4.31)$$

$$g_k(x) = \tanh(x) \quad (4.32)$$

En la Figura 4.4 se representa el esquema básico de una neurona artificial con una función de base lineal que consiste en una suma ponderada de las entradas tal y como se indica en (4.33):

$$I(\mathbf{x}) = \omega_0 + \sum_1^n \omega_i x_i \quad (4.33)$$

donde ω_0 es un umbral, también llamado “*bias*” (sesgo), que hace que ante una entrada nula la salida no sea necesariamente nula. Los pesos sinápticos (ω_i , $i = 1, 2, \dots, n$) ponderan el efecto de cada entrada sobre la activación de la neurona. Si $\mathbf{w} = [\omega_0, \omega_1, \dots, \omega_n]^T$ es el vector de pesos de la neurona, $I(\mathbf{x})$ puede expresarse de forma más compacta como $I(\mathbf{x}) = \mathbf{x}^T \mathbf{w}$. Por tanto, una función de base lineal define un hiperplano en el espacio de entrada y la activación de la neurona ($I(\mathbf{x})$) es proporcional a la distancia de la entrada considerada al plano definido por el conjunto de pesos sinápticos que definen la función de base lineal.

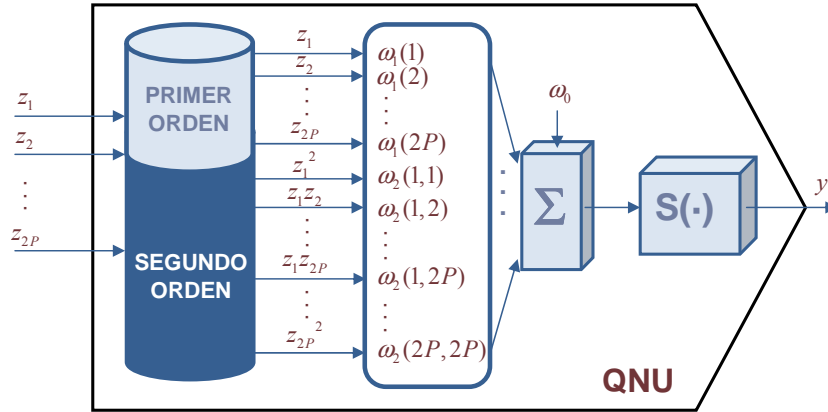


Figura 4.5: Unidad neuronal cuadrática (QNU)

4.7.2. Redes neuronales con funciones de base cuadrática (SONN)

Una de las propiedades de las redes neuronales artificiales es su capacidad para extraer parámetros esenciales a partir de datos complejos del espacio de entrada. En el caso de la detección, la red se convierte en un conjunto de conexiones complejas de un espacio de dimensión $2P$ a un espacio de salida de dimensión unidad. En la sección 4.7.1, se utilizan redes con neuronas con funciones base lineales que limitan la búsqueda de parámetros invariantes a las correlaciones de primer orden en los vectores de entrenamiento.

En este apartado se propone utilizar unidades de mayor orden para poder completar la información adquirida por la red a través de las correlaciones de órdenes superiores. La salida de una neurona de orden N es:

$$I(\mathbf{z}) = \omega_0 + \sum_j^{2P} \omega_1(j) \mathbf{z}_j + \sum_j^{2P} \sum_i^{2P} \omega_2(j, i) \mathbf{z}_j \mathbf{z}_i + \dots \quad (4.34)$$

donde $\omega_N(j, i, \dots)$ son los pesos ajustables especializados en las correlaciones de orden N entre los vectores de entrenamiento y la salida. El término ω_0 se corresponde con la bias de la neurona. El número de pesos total asociados a una unidad depende del orden N de la neurona. Una de las desventajas de este tipo de arquitectura es que el número de parámetros libres de una red con neuronas con funciones de base de orden N a determinar en el proceso de aprendizaje puede ser excesivo.

Esta tesis se centra en soluciones basadas en redes neuronales con funciones de base de segundo orden, SONN (*Second Order Neural Network*), para controlar el coste computacional de las soluciones aportadas. En la Figura 4.5 se representa la unidad neuronal cuadrática (QNU, *Quadratic Neural Unit*) cuya salida depende de las relaciones de primer y segundo orden entre los vectores de entrada.

La matriz de pesos asociados a una QNU es:

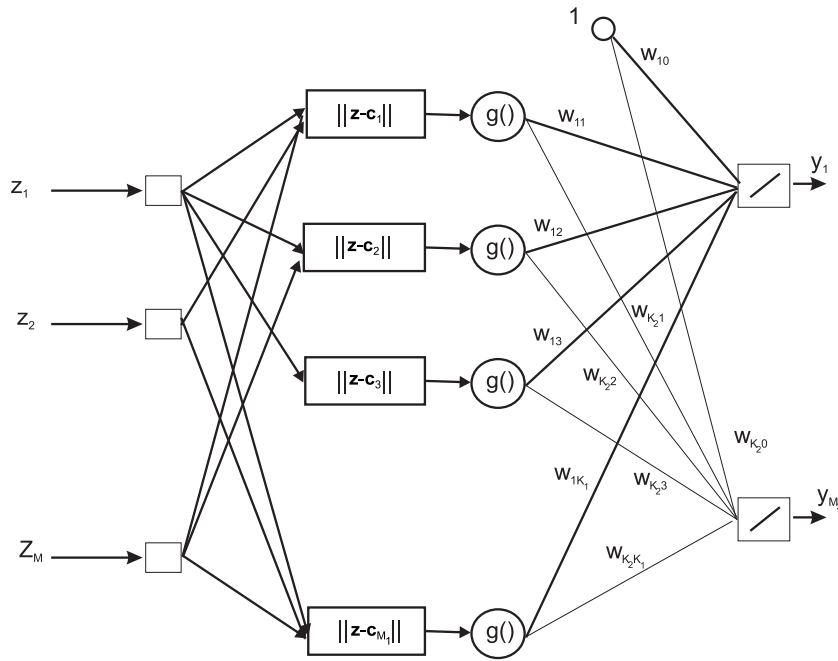


Figura 4.6: Modelo de red neuronal con funciones de base radial

$$\mathbf{w} = \begin{bmatrix} \omega_0 & \omega_1(1) & \omega_1(2) & \cdots & \omega_1(2P) \\ 0 & \omega_2(1,1) & \omega_2(1,2) & \cdots & \omega_2(1,2P) \\ 0 & 0 & \omega_2(2,2) & \cdots & \omega_2(2,2P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega_2(2P,2P) \end{bmatrix} \quad (4.35)$$

La matriz de pesos \mathbf{w} tiene elementos nulos en la parte inferior de la diagonal porque los pesos simétricos con respecto a la diagonal $\omega_2(i,j)$ y $\omega_2(j,i)$, se refieren al mismo término cuadrático $z_i z_j = z_j z_i$. Por lo tanto, el número de elementos no nulos de la matriz es $2P^2 + 3P + 1$ y la expresión matemática de la salida de una QNU puede simplificarse de la siguiente manera:

$$I(\mathbf{z}) = \sum_{j=0}^{2P} \sum_{i=j}^{2P} \omega(j,i) \mathbf{z}_j \mathbf{z}_i = [1, \mathbf{z}]^T \mathbf{w} [1, \mathbf{z}] \quad (4.36)$$

Las neuronas de función base lineal son un caso particular de una QNU, puesto que la primera fila de la matriz \mathbf{w} , expresión (4.35), son los pesos de una neurona con función base lineal. Una de las principales características de la morfología de una unidad neuronal cuadrática es su mayor generalidad que puede ayudar a mejorar el rendimiento de la red de forma significativa.

4.7.3. Redes neuronales con funciones de base radial (RBFNN)

Estas redes suelen estar formadas por dos capas: una primera compuesta por neuronas con funciones de base radial y una de salida de neuronas lineales (Figura 4.6).

Las funciones de base radial fueron utilizadas por primera vez para la solución de problemas de interpolación. Este es el enfoque que suele seguirse para la explicación de su principio de

funcionamiento, pero se puede explicar desde otros puntos de vista, como son la *teoría de la regularización*, la *interpolación de datos ruidosos*, *técnicas de estimación de funciones de regresión basadas en métodos kernel para la estimación de funciones de densidad de probabilidad*, la *teoría de la clasificación y las funciones potencial* [Haykin1999][Bishop1995].

Para explicarlas a partir de la teoría de la regularización hay que acudir a Tikhonov, que propuso el método de la regularización para resolver problemas mal formulados de aprendizaje a partir de ejemplos [Tikhonov1977]. Aprender de los ejemplos puede ser considerado como un problema de regresión para aproximar una función de múltiples variables a partir de pocos datos [Haykin1999]. La teoría clásica de la regularización pretende estabilizar la solución por medio de un funcional no-negativo auxiliar que incorpora información previa acerca de la solución.

$$\min_{\mathbf{F} \in \mathcal{H}} (E + \lambda \|\mathbf{D}\mathbf{F}\|_K^2) \quad (4.37)$$

En la ecuación (4.37) se describe el funcional que hay que minimizar, donde:

- $\lambda \|\mathbf{D}\mathbf{F}\|_K^2$ es el término de regularización.
- \mathbf{D} es un operador diferencial lineal que contiene información previa sobre la forma de la solución.
- λ es un número real positivo, llamado *parámetro de regularización*, que controla la influencia de las restricciones impuestas por el término de regularización en la solución final. Un valor pequeño de λ implica que la aproximación sea muy precisa, sin que eso genere una penalización muy grave. Si el valor de λ es muy grande, la aproximación debe ser sacrificada en detrimento de una función más simple.
- $\|\cdot\|_K^2$ denota la norma impuesta en el espacio de Hilbert \mathcal{H} al que pertenecen el conjunto de funciones K .
- E es el error estimado para los patrones de entrenamiento entre la salida de $F(\mathbf{z})$ y la salida esperada t . En las redes neuronales regularizadas, se suele utilizar el error cuadrático medio como función de error $\left(E = \frac{1}{N} \sum_{i=0}^1 \sum_{\mathbf{z} \in Z_i} |F(\mathbf{z}) - t_{H_i}|^2 \right)$.

Si dividimos el funcional (4.37) en dos términos asociados a la función de error y a la regularización, nos queda la expresión (4.38). El diferencial de Fréchet aplicado a $\xi(F)$, se define formalmente en (4.39)[Debnath1990].

$$\xi(F) = \xi_e(F) + \xi_r(F) = [F(\mathbf{z}) - t]^2 + \lambda \|\mathbf{D}\mathbf{F}\|_K^2 \quad (4.38)$$

$$d\xi(F, h) = \left[\frac{d}{d\beta} \xi(F + \beta h) \right]_{\beta=0} \quad (4.39)$$

Evaluando el diferencial de Fréchet, obtenemos la ecuación (4.40), que se puede expresar en función de productos escalares, $(\cdot, \cdot)_{\mathcal{H}}$, de dos funciones en el espacio de Hilbert \mathcal{H} .

$$\begin{aligned}
d\xi(F) &= -\sum_{i=1}^N [F(\mathbf{z}_i) - t_i] h(\mathbf{z}_i) + \lambda \int_{\mathcal{Z}} \mathbf{D}F \mathbf{D}h d\mathbf{z} \\
d\xi(F) &= -\left(h, \sum_{i=1}^N (F(\mathbf{z}_i) - t_i) \delta(\mathbf{z} - \mathbf{z}_i) \right)_{\mathcal{H}} + \lambda (\mathbf{D}h, \mathbf{D}F)_{\mathcal{H}}
\end{aligned} \tag{4.40}$$

Según la identidad de Green [Lanczos1961], se puede definir de manera única un operador $\tilde{\mathbf{D}}$, tal que:

$$\int_{\mathcal{Z}} \mathbf{D}F \mathbf{D}h d\mathbf{z} = \int_{\mathcal{Z}} h \tilde{\mathbf{D}}\mathbf{D}F d\mathbf{z} = (h, \tilde{\mathbf{D}}\mathbf{D}F)_{\mathcal{H}} \tag{4.41}$$

Aplicando (4.41) en el diferencial de Fréchet nos queda el producto escalar (4.42). Como el parámetro de regularización λ se le asigna de forma ordinaria un valor en el intervalo abierto $(0, \infty)$, el diferencial de Fréchet es cero para cualquier $h(\mathbf{z})$ en el espacio \mathcal{H} si, y sólo si, se cumple la condición (4.43), que también se puede expresar como la ecuación Euler-Lagrange de $\xi(f)$ definida en (4.44) [Debnath1990].

$$d\xi(F) = \left(h, \tilde{\mathbf{D}}\mathbf{D}F - \frac{1}{\lambda} \sum_{i=1}^N (F(\mathbf{z}_i) - t_i) \delta(\mathbf{z} - \mathbf{z}_i) \right)_{\mathcal{H}} \tag{4.42}$$

$$\tilde{\mathbf{D}}\mathbf{D}F - \frac{1}{\lambda} \sum_{i=1}^N (F(\mathbf{z}_i) - t_i) \delta(\mathbf{z} - \mathbf{z}_i) = 0 \tag{4.43}$$

$$\tilde{\mathbf{D}}\mathbf{D}F(\mathbf{z}) = \frac{1}{\lambda} \sum_{i=1}^N (F(\mathbf{z}_i) - t_i) \delta(\mathbf{z} - \mathbf{z}_i) \tag{4.44}$$

La transformación de un problema diferencial en una ecuación integral se puede realizar por medio de una función de Green o su equivalente. Si consideramos la ecuación diferencial ordinaria (4.45), en la que L es un operador diferencial lineal ordinario, $f(x)$ es una función conocida y $u(x)$ la función que se ha de determinar. Un método para resolver la ecuación consiste en hallar el operador inverso L^{-1} , tal que $L^{-1}L$ es el operador identidad. L^{-1} es un operador integral, cuyo núcleo es la llamada *función de Green*, $G(x, y)$, del operador diferencial L (4.46).

$$Lu(x) = f(x) \tag{4.45}$$

$$L^{-1}Lu(x) = u(x) = L^{-1}f(x) = \int G(x, y) f(y) dy \tag{4.46}$$

Para hallar la función de Green, considerada como una función de x , se sabe que se cumple la ecuación diferencial (4.47) [Haykin1999].

$$LG(x, y) = \delta(x - y) \tag{4.47}$$

Considerando $L = \tilde{\mathbf{D}}\mathbf{D}$ y sustituyendo la propiedad de las funciones de Green, expresada en (4.47), en la ecuación de Euler-Lagrange (4.44) se obtiene que la función a determinar F_N , para N patrones de entrenamiento, es una combinación lineal de funciones de Green del operador diferencial $L = \tilde{\mathbf{D}}\mathbf{D}$.

$$F_N(\mathbf{z}) = \frac{1}{\lambda} \sum_{i=1}^N (F_N(\mathbf{z}_i) - t_i) G(\mathbf{z}, \mathbf{z}_i) \tag{4.48}$$

Si el estabilizador \mathbf{D} es traslacionalmente invariante, las funciones de Green centradas en \mathbf{z}_i dependerán únicamente de la diferencia entre \mathbf{z} y \mathbf{z}_i ($G(\mathbf{z}, \mathbf{z}_i) = G(\mathbf{z} - \mathbf{z}_i)$). Si \mathbf{D} es, además, rotacionalmente invariante, las funciones de Green dependerán de la distancia euclídea entre \mathbf{z} y \mathbf{z}_i , por lo que las funciones de Green serán funciones de base radial ($G(\mathbf{z}, \mathbf{z}_i) = G(\|\mathbf{z} - \mathbf{z}_i\|)$).

Se define:

- $\mathbf{F}_N = [F_N(\mathbf{z}_1), F_N(\mathbf{z}_2), \dots, F_N(\mathbf{z}_N)]^T$
- $\mathbf{G} = G(\mathbf{z}_i, \mathbf{z}_j)$ es una matriz de $N \times N$ donde $i, j = 1, \dots, N$
- $\mathbf{w} = [\omega_1, \omega_2, \dots, \omega_N]^T$ de manera que $\omega_i = \frac{1}{\lambda}(F_0(\mathbf{z}_i) - t_i)$ $i = 1, \dots, N$
- $\mathbf{t} = [t_1, t_2, \dots, t_N]^T$

Sustituyendo estas definiciones en (4.48):

$$\begin{aligned} F_N(\mathbf{z}) &= \sum_{i=1}^N \omega_i G(\mathbf{z}, \mathbf{z}_i) \\ \mathbf{F}_N &= \mathbf{G}\mathbf{w} \end{aligned} \quad (4.49)$$

La ecuación (4.48) se puede interpretar como la salida de una red neuronal con una capa oculta con neuronas cuya función de activación es una función de dos variables $G(\mathbf{z}, \mathbf{z}_i)$ del espacio de Hilbert \mathcal{H} y las conexiones entre las neuronas, que dependen de información a priori de la solución según el operador \mathbf{D} , cumplen la ecuación (4.50).

$$\begin{aligned} \mathbf{w} &= \frac{1}{\lambda}(\mathbf{F}_N - \mathbf{t}) = \frac{1}{\lambda}(\mathbf{G}\mathbf{w} - \mathbf{t}) \\ (\mathbf{G} - \lambda\mathbf{I})\mathbf{w} &= \mathbf{t} \end{aligned} \quad (4.50)$$

donde \mathbf{I} representa la matriz identidad de $N \times N$.

De entre las funciones de Green que tienen más interés para el diseño de redes con funciones de base radial se encuentran [Haykin1999]:

$$G(\mathbf{z}, \mathbf{x}) = (\|\mathbf{z} - \mathbf{x}\|^2 + \alpha^2)^{\frac{1}{2}} \quad (4.51)$$

$$G(\mathbf{z}, \mathbf{x}) = (\|\mathbf{z} - \mathbf{x}\|^2 + \alpha^2)^{-\frac{1}{2}} \quad (4.52)$$

$$G(\mathbf{z}, \mathbf{x}) = \exp\left(\frac{-\|\mathbf{z} - \mathbf{x}\|^2}{2\alpha^2}\right) \quad (4.53)$$

En todos los casos, $\alpha > 0$.

Según la ecuación (4.49), la función F_0 aproximada es el resultado de combinar linealmente N funciones de Green correspondientes a los N patrones utilizados en el proceso de aprendizaje. La estrategia de utilizar tantas funciones (neuronas de la capa oculta) como datos de entrada, centradas en dichos datos, puede no ser adecuada para el diseño de redes con funciones de base radial. Cuando el número de datos disponibles es mucho mayor que el número de grados de libertad del problema a resolver, esta estrategia disminuirá de forma drástica las capacidades generalizadoras de la red. Para resolver este problema, se reduce la complejidad de la red para obtener una aproximación a la solución regularizada.

La aproximación parte de la base de una búsqueda de una solución en un espacio de dimensión menor, es decir, la solución dependerá de un número de funciones base, linealmente independientes, menor que los datos de entrada.

$$F_N^*(\mathbf{z}) = \sum_{i=1}^M \omega_i G(\|\mathbf{z} - \mathbf{c}_i\|) \quad (4.54)$$

donde \mathbf{c}_i son los centros de las M neuronas de la capa oculta que tienen que ser determinados en el proceso de entrenamiento de la red, al igual que el vector de pesos. La expresión (4.54) formula matemáticamente el modelo de RBFNN representado en la Figura 4.6.

4.7.4. Máquinas de vectores soporte

Las máquinas de vectores soporte (SVM) tienen la capacidad de ser aproximadores universales de funciones de variables múltiples para cualquier grado deseado de exactitud [Wang2005]. Al igual que las redes neuronales artificiales, las SVMs son agentes inteligentes que necesitan ser entrenados bajo supervisión.

El objetivo del proceso de aprendizaje es la minimización del riesgo estructural que puede ser explicado a partir de la teoría de la regularización clásica [Tikhonov1977]:

$$\min_{F \in \mathcal{H}} \left(E + \frac{1}{2C} \|\mathbf{D}F\|_K^2 \right) \quad (4.55)$$

El parámetro C se ha introducido para respetar la notaciones estándar de las C-SVM [Vapnik1995] [Borges1998] y está relacionado con el parámetro de regularización, $\lambda = \frac{1}{2C}$.

Las funciones de error que se utilizan en el entrenamiento de las SVMs son:

- SVMs para clasificación se utiliza (4.56), donde $u[\cdot]$ representa la función escalón y t_i toma dos valores posibles para cada una de las clases, -1 y 1 [Shawe2004].

$$E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 - t_{H_0} F(\mathbf{z}))) + \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - t_{H_1} F(\mathbf{z}))) \quad (4.56)$$

- SVMs para regresión utilizan (4.57), conocida como función de error ϵ -insensitiva de Vapnik [Vapnik1995], donde t_{H_0} y t_{H_1} son números reales.

$$E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(|t_{H_0} - F(\mathbf{z})| - \epsilon)) + \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(|t_{H_1} - F(\mathbf{z})| - \epsilon)) \quad (4.57)$$

Según [Evgeniou2000], para cualquiera de las funciones de error consideradas, la función que minimiza el funcional expresado en (4.55) tiene la siguiente forma:

$$F_N(\mathbf{z}) = \sum_{i=1}^N \omega_i G(\mathbf{z}, \mathbf{z}_i) \quad (4.58)$$

Dado que esta tesis se centra en la aplicación de las SVMs a tareas de detección, en el estudio sólo se considera la función de error de clasificación (4.56). El objetivo del entrenamiento con

un número N de ejemplos finitos consiste en encontrar el vector de pesos, ω_i , que definen $F_N(\mathbf{z})$ en (4.58) que minimiza el funcional (4.55). Debido a la dificultad de trabajar directamente con $e(n)$, se plantea un problema equivalente, es decir, la solución a este problema también minimiza el siguiente funcional:

$$\min_{F \in \mathcal{H}, \xi} C \sum_{i=1}^N \xi_i + \frac{1}{2} \|DF\|^2 \quad (4.59)$$

Sujeto a

$$\begin{aligned} t_i F_N(\mathbf{z}_i) &\geq 1 - \xi_i & i = 1, \dots, N \\ \xi_i &\geq 0 & i = 1, \dots, N \end{aligned} \quad (4.60)$$

En la ecuación (4.59), $\xi_i \geq 1$ si $t_i \neq F(\mathbf{x}_i)$, de manera que $\sum_{i=1}^N \xi_i$ es una cota superior del número de errores que se cometen dentro del conjunto de entrenamiento. La equivalencia de los dos problemas se basa en que los dos tratan de minimizar el mismo error. Además el parámetro C mantiene el significado del parámetro de regularización, ya que si C es muy grande se asigna más importancia a minimizar el error que al término de regularización o de suavizado $\|DF\|^2$ y viceversa. Desde el punto de vista de interpretación geométrica, C representa el compromiso entre la obtención de un hiperplano de separación con la menor inclinación posible y una desviación mínima de los valores objetivo.

Teniendo en cuenta la ecuación diferencial (4.47) que se cumple para las funciones de Green y considerando $\mathbf{L} = \tilde{\mathbf{D}}\mathbf{D}$, la norma en el espacio de Hilbert al que pertenece la función $F(\mathbf{z})$ se puede calcular como:

$$\|DF(\mathbf{z})\|^2 = (F, \tilde{\mathbf{D}}\mathbf{D}F)_{\mathcal{H}} = \left(\sum_{i=1}^N \omega_i G(\mathbf{z}, \mathbf{z}_i), \sum_{j=1}^N \omega_j \delta(\mathbf{z} - \mathbf{z}_j) \right)_{\mathcal{H}} = \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j G(\mathbf{z}_i, \mathbf{z}_j) \quad (4.61)$$

Las funciones de Green de interés para las SVMs son definidas positivas con media cero y tienen la propiedad de que ninguna de ellas tiene la norma $\|DF\|_K^2$ igual a cero, por lo tanto el espacio nulo del funcional $\|DF\|_K^2$ está compuesto de funciones constantes. La función $F(\mathbf{z})$ se puede reescribir como (4.62) cumpliendo la condición (4.63) [Evgeniou2000].

$$F_N(\mathbf{z}) = \sum_{i=1}^N \omega_i G(\mathbf{z}, \mathbf{z}_i) + b \quad (4.62)$$

$$\sum_{i=1}^N \omega_i = 0 \quad (4.63)$$

Utilizando la técnica de los multiplicadores de Lagrange se llega a la función objetivo dual:

$$L(\alpha, \beta) = C \sum_{i=1}^N \xi_i + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j G(\mathbf{z}_i, \mathbf{z}_j) - \sum_{i=1}^N \alpha_i \left[t_i \left(\sum_{j=1}^N \omega_j G(\mathbf{z}_j, \mathbf{z}_i) + b \right) - 1 + \xi_i \right] - \sum_{i=1}^N \beta_i \xi_i \quad (4.64)$$

La solución a este problema de minimización con restricciones, donde $\alpha_i, \beta_i \geq 0$, viene dado por las condiciones de punto de silla, resultado de la derivada con respecto a las variables primales ω_j, b y ξ_i , y las condiciones de Karush-Kuhn-Tucker (KKT):

$$\begin{aligned}\frac{dL(\alpha, \beta)}{d\omega_j} &= \sum_{i=1}^N \omega_i G(\mathbf{z}_i, \mathbf{z}_j) - \sum_{i=1}^N \alpha_i t_i G(\mathbf{z}_i, \mathbf{z}_j) = 0 \\ \frac{dL(\alpha, \beta)}{db} &= - \sum_{i=1}^N \alpha_i t_i = 0 \\ \frac{dL(\alpha, \beta)}{d\xi_i} &= C - \alpha_i - \beta_i = 0\end{aligned}\tag{4.65}$$

De donde se deduce (4.66) y (4.67).

$$\omega_i = \alpha_i t_i\tag{4.66}$$

$$C = \alpha_i + \beta_i\tag{4.67}$$

Aplicando en (4.64), nos queda (4.68):

$$\text{mín } \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j t_i t_j G(\mathbf{z}_i, \mathbf{z}_j) - \sum_{i=1}^N \alpha_i\tag{4.68}$$

Sujeto a:

$$\begin{aligned}0 &\leq \alpha_i \leq C & i = 1, \dots, N \\ \sum_{i=1}^N \alpha_i t_i &= \sum_{i=1}^N \omega_i = 0\end{aligned}\tag{4.69}$$

Una de las condiciones coincide con (4.63) impuesta por considerar funciones de Green definidas positivas.

Se trata de un problema de programación convexo cuadrático que obtiene los valores de los multiplicadores de Lagrange α_i , que nos permiten reconstruir la función que minimiza el funcional requerido:

$$F_N(\mathbf{z}) = \sum_{i=1}^{N_{sv}} \alpha_i t_i G(\mathbf{s}_i, \mathbf{z}) + b\tag{4.70}$$

donde N_{sv} denota el número de vectores soporte y \mathbf{s}_i son los vectores soporte del conjunto $\mathbf{z}_1, \dots, \mathbf{z}_N$. Algunos de los coeficientes de Lagrange, α_i son nulos, en concreto aquellos que se corresponden a vectores \mathbf{z}_i que no son vectores soporte, por lo tanto, se cumple que $N_{sv} \leq N$.

Como conclusión, podemos establecer que los parámetros de diseño de las SVMs son las funciones de Green consideradas y el valor de C que determina el efecto de la regularización en forma de límite superior a los coeficientes de Lagrange. La SVM óptima se caracteriza por tener las funciones de Green y el valor C que minimizan el error de clasificación.

Parte II

Investigación y Resultados

Capítulo 5

Detección de blancos con parámetros desconocidos

5.1. Introducción

En el capítulo 3, se han presentado los casos de estudio planteando el problema de la detección como un test de hipótesis simple donde se asumen conocidos todos los parámetros necesarios para definir las funciones de densidad de probabilidad bajo ambas hipótesis.

En muchas situaciones prácticas, las propiedades estadísticas de la interferencia pueden estimarse a partir de medidas obtenidas en el entorno del radar, pero la estimación de las propiedades del blanco es muy difícil, por lo que para poder construir un cociente de verosimilitud ha de asumirse un determinado modelo de blanco.

En el apartado 5.2 se presenta un estudio de sensibilidad del detector NP de blancos con parámetros conocidos en distintos entornos de clutter. Para ello, se evalúan las capacidades detectoras de detectores NP diseñados para unos parámetros de blancos ante blancos con parámetros diferentes a los asumidos para el diseño. En concreto, en los casos de estudio considerados, se ha analizado la sensibilidad respecto al coeficiente de correlación y la frecuencia Doppler del blanco. Los resultados obtenidos revelan una alta sensibilidad de este tipo de detectores respecto de los parámetros considerados (para una P_{FA} , las pérdidas de P_D asociadas son muy elevadas cuando los parámetros del blanco son distintos a los supuestos en el diseño).

Como consecuencia, se propone plantear el problema de detección como un test de hipótesis compuesto. Si los parámetros desconocidos son variables aleatorias con FDP desconocida, una de las realizaciones del detector de Neyman-Pearson consiste en comparar el cociente de verosimilitud promediado, ALR, con un umbral especificado para cumplir los requisitos de probabilidad de falsa alarma. Si se desconoce la función densidad de probabilidad de los parámetros, suele asumirse uniforme en el intervalo de variación esperable [Aref1994]. En muchas ocasiones, este cociente de verosimilitud requiere la resolución de integrales muy complejas que pueden hasta no tener una solución cerrada y que obligan a plantear su resolución a través de aproximaciones, dando lugar a detectores subóptimos.

En el apartado 5.3 se lleva a cabo un estudio de las capacidades de detección de un detector basado en el cociente de verosimilitud generalizado limitado, CGLR, como aproximación subóptima del ALR que puede ser implementada. En general, estos detectores suelen conllevar costes

computacionales muy elevados, pero resultan útiles para comprobar la aproximación realizada por las soluciones basadas en inteligencia artificial que se presentarán en el capítulo 7.

Los casos de estudio que se contemplan en este capítulo están relacionados con los presentados en el capítulo 3, a los que se ha añadido la variación de algún parámetro característico del modelo de blanco:

- **Caso de estudio 1:** Detección de blancos fluctuantes gaussianos en presencia de ruido blanco gaussiano aditivo. Se estudia la sensibilidad del LR y las características de detección de aproximaciones basadas en el CGLR, cuando el coeficiente de correlación o la frecuencia Doppler del blanco son desconocidos.
- **Caso de estudio 2:** Detección de blancos fluctuantes gaussianos en presencia de interferencia gaussiana. Se estudia la sensibilidad del LR y las características de detección de aproximaciones basadas en el CGLR, cuando el coeficiente de correlación o la frecuencia Doppler del blanco son desconocidos.
- **Caso de estudio 3:** Detección de blancos no fluctuantes en presencia de clutter K distribuido. Se estudia la sensibilidad del LR y las características de detección de aproximaciones basadas en el CGLR cuando la frecuencia Doppler del blanco es desconocida.

Algunas de las aportaciones resultado de la investigación desarrollada en este tema han sido publicadas en [Jarabo2006], [Mata2006] y [Mata2008].

5.2. Estudio de sensibilidad de los detectores óptimos de blancos con parámetros conocidos

Los detectores de NP son detectores paramétricos que requieren una caracterización completa del blanco y de la interferencia. Dado que una posible implementación del detector de NP se basa en el cociente de verosimilitud, es necesario asumir en el diseño las FDPs condicionadas a ambas hipótesis. En el capítulo 3, se analizaron las funciones de verosimilitud y los detectores óptimos para tres casos de estudio en los que los parámetros de los modelos de blanco eran conocidos. En este apartado se extiende el estudio de estos detectores cuando los blancos tienen propiedades estadísticas diferentes a las del modelo asumido en el diseño. A partir de ahora, se denotará a los parámetros de diseño como $DSNR$, $DSIR$, $DSCR$, ρ_s^d , θ_R^d y Ω^d , y a los de simulación, utilizados para sintetizar los vectores de observación bajo hipótesis H_1 , $SSNR$, $SSIR$, $SSCR$, ρ_s^s , θ_R^s y Ω^s . Los resultados se han obtenido asumiendo un vector de observación formado por $P = 8$ pulsos.

5.2.1. Estudio de sensibilidad del detector LR para el caso de estudio 1

En el apartado 3.4.1 se obtuvo la expresión simplificada del detector LR:

$$\tilde{\mathbf{z}}^T [(2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{ss}}} + 2 * \mathbf{I})^{-1}] \mathbf{z}^* \underset{H_0}{\overset{H_1}{\gtrless}} Ln \left(\eta_{cv}(P_{FA}) \frac{|\Sigma_{\tilde{\mathbf{ss}}} + 2 * \mathbf{I}|}{2^P} \right) = \eta_s \quad (5.1)$$

Esta regla de decisión depende de $\Sigma_{\tilde{\mathbf{ss}}}$ y, por lo tanto, depende de los valores asumidos en el diseño para $DSNR$, ρ_s^d y Ω^d .

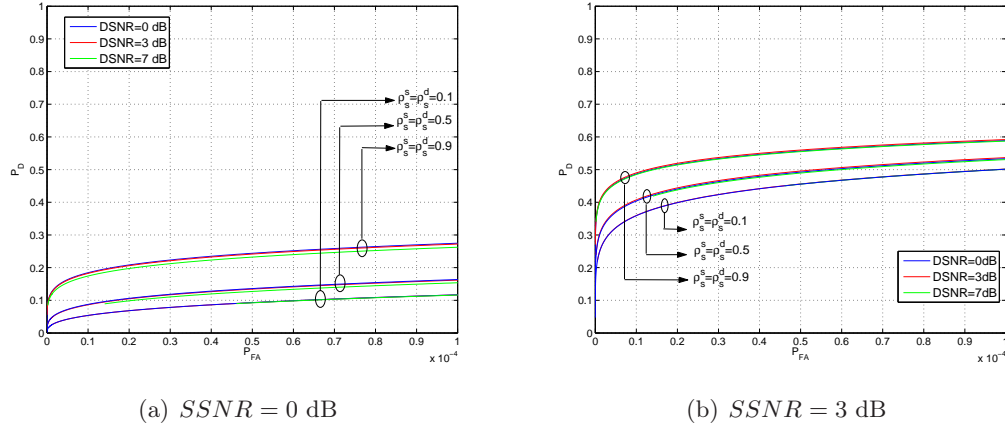


Figura 5.1: Robustez del LR para el caso de estudio 1 con respecto a $DSNR$. $\Omega^d = \Omega^s = 0$ y $\rho_s^d = \rho_s^s$.

En las Figuras 5.1(a), 5.1(b) y 5.2(a) se representan las curvas ROC cuando $SSNR = 0, 3$ y 7 dB respectivamente. Los resultados permiten estudiar la robustez del detector óptimo frente a la $DSNR$ para diferentes valores de ρ_s . En general, podemos decir que la dependencia con respecto a $DSNR$ es prácticamente despreciable. Para una P_{FA} de 10^{-4} , la mayor pérdida de detección es menor del 1%, excepto para $SSNR = 0$ dB que es algo mayor. La variación de P_D es mucho menor para valores del coeficiente de correlación próximos a cero.

El estudio se centra en $SSNR = 7$ dB, ya que proporciona valores de P_D razonables en aplicaciones radar. Cuando el blanco se encuentra correlado, la pérdida de detección del detector de Neyman-Pearson cuando $\Omega^s \neq \Omega^d$ puede ser muy elevada. La Figura 5.2(b) representa cómo las capacidades de detección se ven mermadas según Ω^d se aleja de Ω^s cuando $\rho_s^s = \rho_s^d = 0, 9$.

Los resultados de la Figura 5.3 reflejan la dependencia del detector basado en el LR con ρ_s^d . Las mayores pérdidas de detección se observan cuando ρ_s^d es próximo a uno y el vector de observación tiene un ρ_s^s próximo a cero. Por lo tanto, un detector de blancos fluctuantes en AWGN basado en el criterio de Neyman-Pearson presenta una dependencia con el coeficiente de correlación y la frecuencia Doppler normalizada considerados en el diseño. Para minimizar el efecto de los parámetros desconocidos del blanco en las capacidades de detección, es necesario plantear el diseño de detectores robustos frente a variaciones de ρ_s y Ω .

5.2.2. Estudio de sensibilidad del detector LR para el caso de estudio 2

En este apartado se realiza un estudio de la robustez del detector óptimo cuando el vector de observación tiene unos parámetros de simulación ($SSIR$, ρ_s^s y Ω^s) distintos a los valores de $DSIR$, ρ_s^d y Ω^d asumidos en el diseño del detector óptimo definido en la expresión (3.42). En las Figuras 5.4, 5.5 y 5.6, se representa la dependencia del detector frente a $DSIR$, ρ_s^d y Ω^d respectivamente, ante un clutter no dominante, $CNR = 20$ dB, con distintos coeficientes de correlación.

Para todos los valores de ρ_c , ρ_s y de $SSIR$ considerados, se observa que la pérdida de detección cuando $DSIR \neq SSIR$ es tan pequeña que se puede considerar despreciable (Figura 5.4). Para el resto de los estudios se fijará, para cada valor de ρ_c , una SIR que permita obtener P_D de

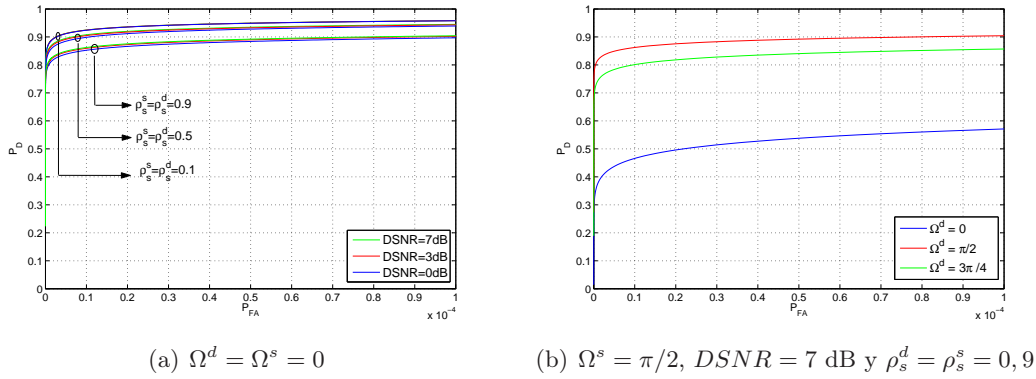


Figura 5.2: Robustez del LR para el caso de estudio 1 con respecto a $DSNR$ y Ω^d . $SSNR = 7$ dB.

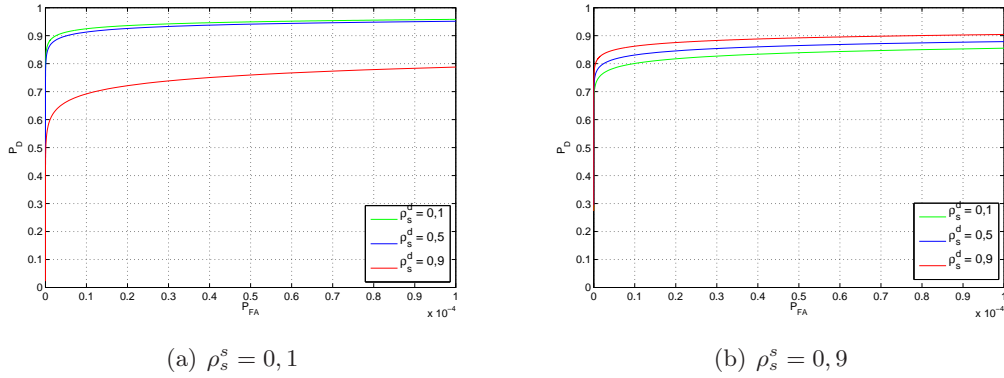


Figura 5.3: Robustez del LR para el caso de estudio 1 con respecto a ρ_s^d . $DSNR = SSNR = 7$ dB y $\Omega^d = \Omega^s = 0$.

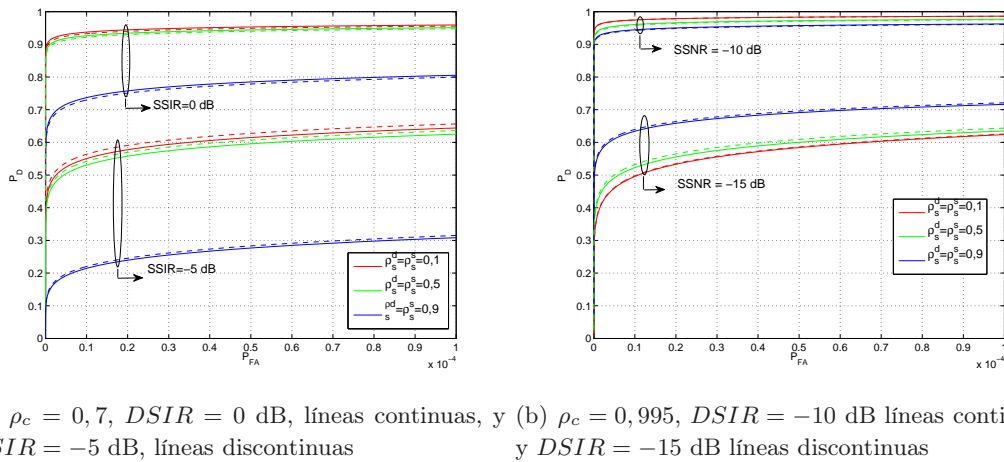


Figura 5.4: Robustez del LR para el caso de estudio 2 con respecto a $DSIR$. $CNR = 20$ dB y $\Omega^d = \Omega^s = \pi/2$.

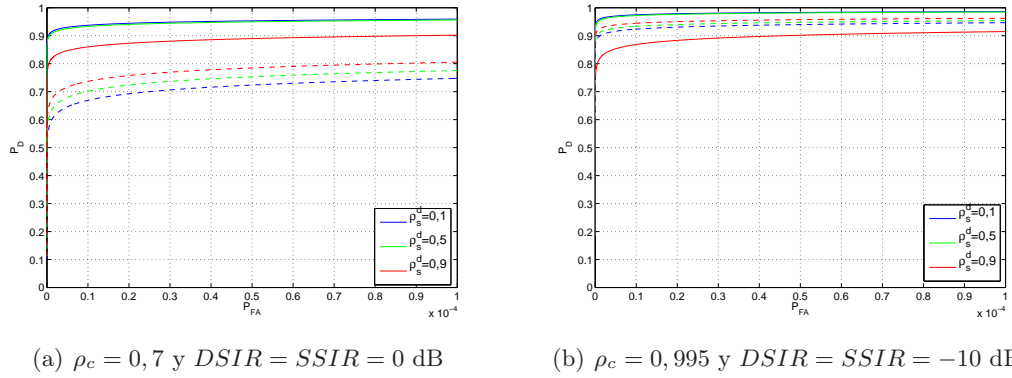


Figura 5.5: Robustez del LR para el caso de estudio 2 con respecto a ρ_s^d . $CNR = 20$ dB y $\Omega^d = \Omega^s = \pi/2$. Las líneas continuas se corresponden con $\rho_s^s = 0,1$, mientras que la discontinuas se corresponden con $\rho_s^s = 0,9$.

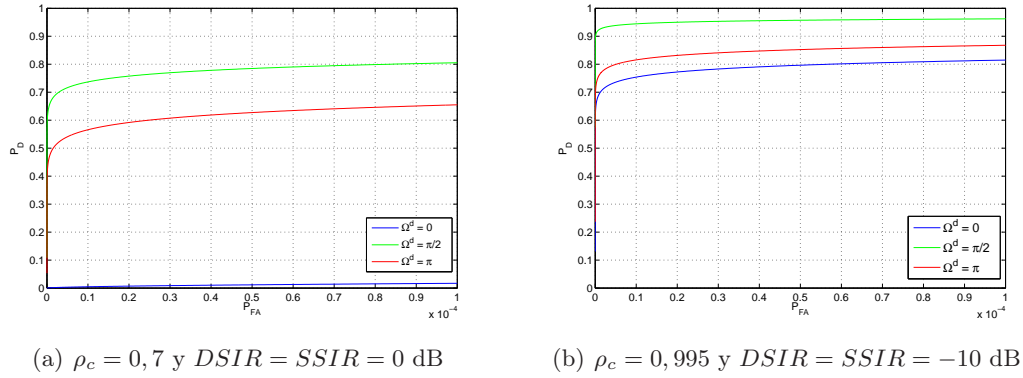
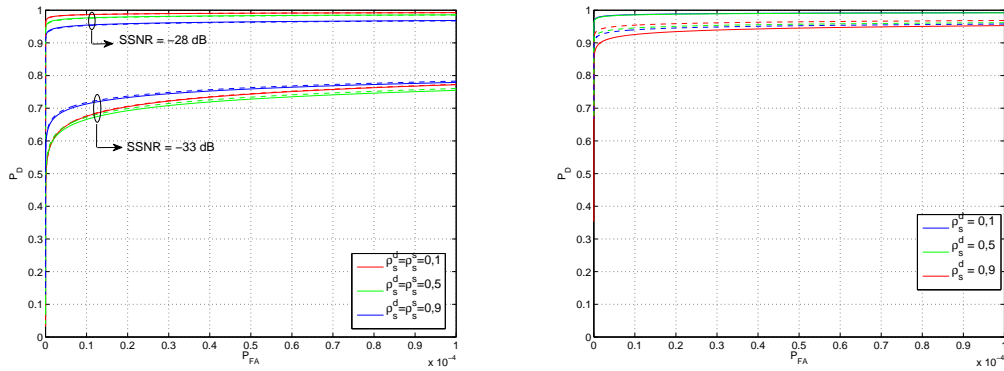


Figura 5.6: Robustez del LR para el caso de estudio 2 con respecto a Ω^d . $CNR = 20$ dB, $\rho_s^d = \rho_s^s = 0,9$ y $\Omega^s = \pi/2$.

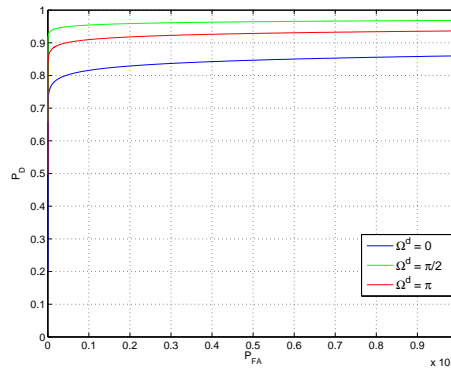
interés en aplicaciones radar.

En la Figura 5.5, se representan las curvas ROC obtenidas cuando se simulan detectores diseñados con diferentes valores de ρ_s^d con vectores de observación generados con una $\rho_s^s = 0,1$, líneas continuas, y $\rho_s^s = 0,9$, líneas discontinuas. Se ha asumido $\Omega^d = \Omega^s = \pi/2$ como un punto intermedio, ya que $\Omega = 0$ y $\Omega = \pi$ representan los casos extremos correspondientes a espectros del blanco lo más cerca y lo más lejos posible del espectro del clutter respectivamente. Cuando se simula con $\rho_s^s = 0,9$, la pérdida de detección es casi proporcional a la diferencia $\rho_s^s - \rho_s^d$ para los valores de ρ_c considerados.

La sensibilidad del detector óptimo frente a Ω^d cuando $\Omega^s = \pi/2$ se presenta en la Figura 5.6. Las capacidades de detección son peores cuando se diseña para valores de Ω^d próximos a cero que para valores cercanos a π . Este comportamiento asimétrico con respecto a la Ω^d considerada se puede explicar teniendo en cuenta que $\Omega = 0$ representa que el espectro del blanco y del clutter coinciden, dificultando mucho más la detección de blancos sin movimiento. Al comparar los resultados obtenidos para $\rho_s = 0,9$ en las Figuras 5.6(a) y 5.6(b), para valores $\rho_c = 0,7$ y $\rho_c = 0,995$ respectivamente, se observa que cuando el $\rho_c > \rho_s$ la sensibilidad del detector con



(a) Robustez frente a $DSIR$. $\Omega^d = \Omega^s = \pi/2$, (b) Robustez frente a ρ_s^d . $\Omega^d = \Omega^s = \pi/2$, $DSIR = DSIR = -28$ y -33 dB para las líneas continuas $SSIR = -28$ dB y $\rho_s^s = 0, 1$ y $0, 9$ para las líneas y discontinuas respectivamente



(c) Robustez frente a Ω^d . $DSIR = SSIR = -28$ dB, $\rho_s^d = \rho_s^s = 0, 9$ y $\Omega^s = \pi/2$.

Figura 5.7: Robustez del LR para el caso de estudio 2 asumiendo $CNR = 40$ dB y $\rho_c = 0, 995$.

$\Omega^d = 0$ es mucho menor.

En la Figura 5.7 se estudia la robustez del LR ante la presencia de un clutter dominante muy correlado, $CNR = 40$ dB y $\rho_c = 0, 995$. Al igual que ocurre con los detectores LR diseñados para un clutter no dominante, $CNR = 20$ dB, la sensibilidad frente $DSIR$ es despreciable (Figura 5.7(a)), la sensibilidad con respecto a ρ_s^d es peor para detectores diseñados con ρ_s^d próximos a la unidad (Figura 5.7(b)) y la sensibilidad con respecto a Ω^d es asimétrica presentando mayores pérdidas cuando Ω^s es próxima a cero (Figura 5.7(c)).

5.2.3. Estudio de sensibilidad del detector LR para el caso de estudio 3

Una vez formulado el detector óptimo de blancos estacionarios en presencia de clutter K distribuido en (3.48), debe estudiarse cuál es la pérdida de detección si los parámetros del blanco simulado, $SSCR$, θ_R^s y Ω^s , no coinciden con los considerados en el diseño del detector, $DSCR$, θ_R^d y Ω^d . En la Figura 5.8, se presentan los resultados de la robustez del detector basado en el criterio de Neyman-Pearson en función de la SCR considerada en el diseño, $DSCR$. Comparando estos resultados con los obtenidos en un escenario dominado por clutter gaussiano (Figura 5.7(a)), se

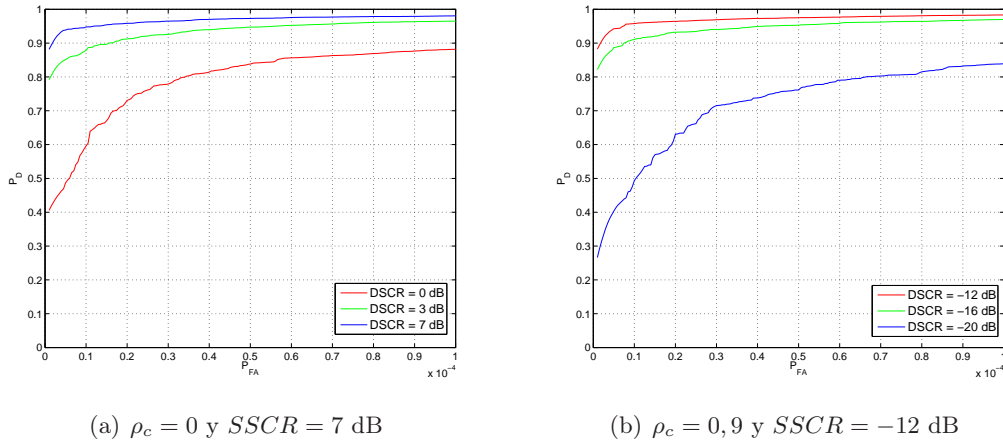


Figura 5.8: Robustez del LR para el caso de estudio 3 con respecto a $DSCR$. $\theta_R^d = \theta_R^s = \pi/4$ y $\Omega^d = \Omega^s = \pi/2$.

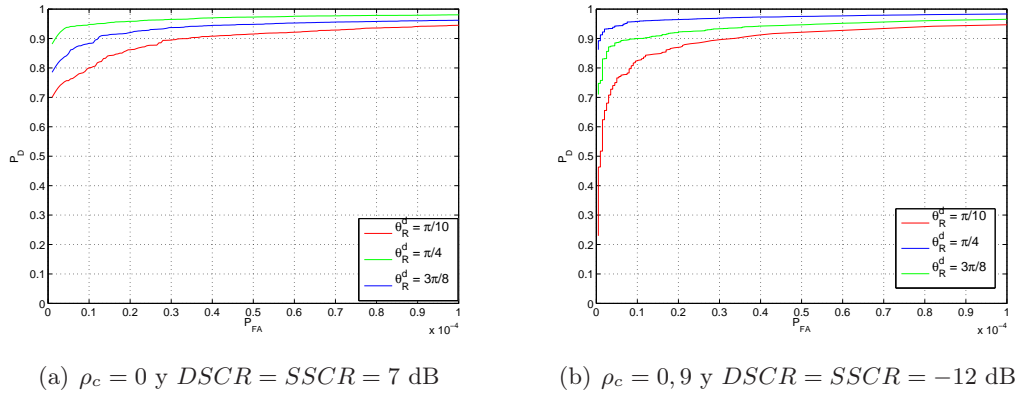


Figura 5.9: Robustez del LR para el caso de estudio 3 con respecto a θ_R^d . $\theta_R^s = \pi/4$ y $\Omega^d = \Omega^s = \pi/2$.

observa que la dependencia con respecto a la $DSCR$ es mucho mayor. Esta dependencia, que se produce independientemente del coeficiente de correlación del clutter, confirma que el diseño de detectores robustos en presencia de clutter hankel será mucho más crítico.

Aunque en los detectores incoherentes se pierde la información de la fase, en la Figura 5.9 se demuestra que cuando la fase asumida en el diseño no coincide con la de simulación hay pérdidas de detección. El orden de la pérdida de detección es similar tanto para un clutter incorrelado como para un clutter muy correlado. Esta especie de independencia con respecto a ρ_c , se explica porque en el caso del clutter muy correlado, el espectro del blanco no fluctuante no coincide con el del clutter.

Al tratarse de un blanco no fluctuante, su espectro es una delta a la frecuencia correspondiente a su velocidad radial. En el caso de detección en presencia de clutter K distribuido, cuando el detector se ha diseñado asumiendo una frecuencia Doppler del blanco determinada, cualquier mínima variación de la velocidad del blanco simulado con respecto a la asumida en el diseño supondrá una gran pérdida de detección. En la Figura 5.10 se comprueba la sensibilidad del

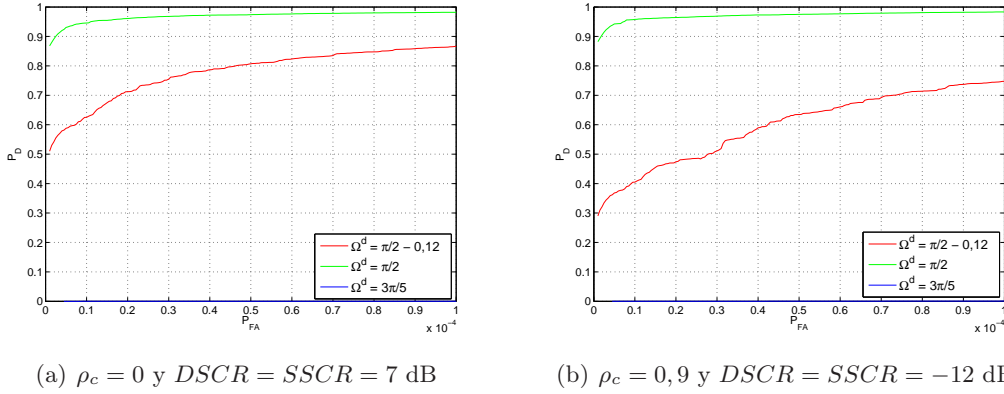


Figura 5.10: Robustez del LR para el caso de estudio 3 con respecto a Ω^d . $\theta_R^d = \theta_R^s = \pi/4$ y $\Omega^s = \pi/2$.

detector LR diseñado con parámetros conocidos con respecto a la Ω^d . En estos resultados, se observa una considerable pérdida en P_D para una misma P_{FA} , cuando la frecuencia Doppler de simulación es $\Omega^d - 0,12$, tanto para $\rho_c = 0$ como para $\rho_c = 0,9$, y se confirma que cualquier diferencia mayor supone una pérdida total de detección. El espectro del clutter también influye en esta sensibilidad. Al comparar las Figuras 5.10(a) y 5.10(b), el comportamiento del detector empeora cuando el clutter K distribuido está correlado.

5.3. Detectores subóptimos para aproximar el ALR

Una vez analizada la sensibilidad de los detectores óptimos para blancos de parámetros conocidos con respecto a dichos parámetros para los tres casos de estudio, se propone el estudio del ALR como solución óptima para aquellos problemas en los que los parámetros desconocidos son variables aleatorias con FDPs desconocidas. En este caso, el problema de detección puede plantearse como un test de hipótesis compuesto en el que la función de densidad de probabilidad del vector de observación bajo hipótesis H_1 depende de una serie de parámetros, φ , en un espacio de parámetros, χ , siendo las hipótesis subespacios de χ . $f(\mathbf{z}|\varphi)$ es la FDP que describe el mapeo de χ a \mathcal{Z} . Si se conoce $f(\mathbf{z}|\varphi)$ para todos los valores de φ en χ , una de las realizaciones del detector de Neyman-Pearson consiste en comparar el cociente de verosimilitud promediado (ALR) con un umbral especificado para cumplir los requisitos de probabilidad de falsa alarma, descrita en (5.2), donde se han calculado las verosimilitudes bajo cada hipótesis por separado [Van1968]. Si se desconoce la función densidad de probabilidad de los parámetros suele asumirse uniforme en el intervalo de variación esperable [Aref1994].

$$\Lambda(\tilde{\mathbf{z}}) = \frac{f(\tilde{\mathbf{z}}|H_1)}{f(\tilde{\mathbf{z}}|H_0)} = \frac{\int_{\chi} f(\tilde{\mathbf{z}}|\varphi, H_1) f(\varphi|H_1) d\varphi}{f(\tilde{\mathbf{z}}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{ALR}(P_{FA}) \quad (5.2)$$

En muchas ocasiones este cociente de verosimilitud requiere la resolución de integrales muy complejas que pueden hasta no tener una solución cerrada y que obligan a plantear su resolución a través de aproximaciones subóptimas. Un esquema basado en el cociente de verosimilitud generalizado (GLR, *Generalized Likelihood Ratio*) permite el diseño de detectores subóptimos

para el caso de test de hipótesis compuesto [Van1968].

$$\max_{\varphi} \Lambda(\varphi) \underset{H_0}{\overset{H_1}{\geq}} \eta_{GLR}(P_{FA}) \quad (5.3)$$

El test GLR se basa en sustituir φ por su estimación de máxima probabilidad (máxima verosimilitud). Sin embargo, en muchas situaciones la regla de decisión (5.3) no puede ser implementada de forma práctica, ya que necesita un número infinito de filtros para cubrir todos los posibles valores del set de parámetros φ . Se define entonces el cociente de verosimilitud generalizado limitado (CGLR, *Constrained Generalized Likelihood Ratio*):

$$\max_{\varphi_k} \Lambda(\varphi_k) \underset{H_0}{\overset{H_1}{\geq}} \eta_{CGLR}(P_{FA}) \quad k = 1, \dots, K \quad (5.4)$$

donde K representa el número finito de detectores LR diseñados para valores discretos que cubren el rango de valores que puede tomar φ . En este apartado, a los detectores basados en un CGLR con K valores discretos, se les denotará como $CGLR_K$. Según [Brennan1968], el valor de K que se utiliza usualmente es igual a la dimensión P del vector de entrada, pero en [Nayebi1996] se comprueba que en función del parámetro desconocido es necesario un $CGLR_{2P}$ para conseguir una buena aproximación del detector óptimo.

En general, los detectores basados en el CGLR suelen conllevar costes computacionales muy elevados. En esta tesis, se utilizan estos detectores subóptimos para estudiar la aproximación realizada por las soluciones basadas en inteligencia artificial, así como la relación entre complejidad computacional y capacidades de detección de las soluciones propuestas.

Al igual que en el apartado anterior, se asume que los vectores de observación están formados por $P = 8$ pulsos para obtener las curvas ROC de los detectores CGLR para los diferentes casos de estudio.

5.3.1. Detectores CGLR para el caso de estudio 1

De los resultados de robustez del LR para el caso de detección de blancos fluctuantes gaussianos en AWGN presentados en las Figuras 5.1, 5.2 y 5.3, se concluye que el detector presenta mucha robustez frente a la $DSNR$, mientras que la pérdida de detección puede ser grande cuando la frecuencia Doppler o el coeficiente de correlación considerado en el diseño es distinto del correspondiente al vector de observación.

Si se llama φ al parámetro desconocido, ρ_s o Ω , y se asume una variación uniforme en el intervalo $[\varphi_1; \varphi_2]$, se puede formular el ALR en presencia de AWGN como:

$$\Lambda(\tilde{\mathbf{z}}) = \frac{\int_{\varphi_1}^{\varphi_2} \exp(\tilde{\mathbf{z}}^T (\boldsymbol{\Sigma}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}(\varphi) + 2 * \mathbf{I})^{-1} \tilde{\mathbf{z}}^*) d\varphi}{\exp(-\frac{1}{2} |\tilde{\mathbf{z}}|^2)} \underset{H_0}{\overset{H_1}{\geq}} \eta_{ALR}(P_{FA}) \quad (5.5)$$

Debido a que la regla de decisión (5.5) conlleva una integral compleja, se propone utilizar una aproximación al ALR basada en el diseño de un detector $CGLR_K$, el cuál utiliza una estimación de máxima probabilidad dentro de valores discretos del margen de variación de φ :

$$\max_{\varphi_k} \frac{\exp(\tilde{\mathbf{z}}^T (\boldsymbol{\Sigma}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}(\varphi_k) + 2 * \mathbf{I})^{-1} \tilde{\mathbf{z}}^*)}{\exp(-\frac{1}{2} |\tilde{\mathbf{z}}|^2)} \underset{H_0}{\overset{H_1}{\geq}} \eta_{CGLR}(P_{FA}) \quad k = 1, \dots, K \quad (5.6)$$

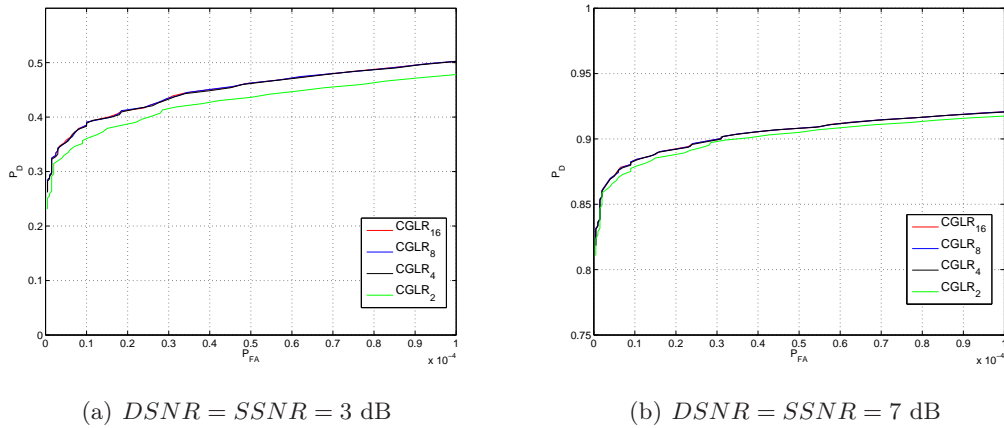


Figura 5.11: Curvas ROC del CGLR para el caso de estudio 1 con $\rho_s^d = \rho_s^s$ variable uniforme en el intervalo $[0, 1]$.

Simplificando la ecuación nos queda la regla de decisión (5.7) basada en la forma cuadrática $q = \tilde{\mathbf{z}}^T \mathbf{Q}(\varphi_k) \tilde{\mathbf{z}}^*$, donde $\mathbf{Q} = (2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}(\varphi_k) + 2 * \mathbf{I})^{-1}$.

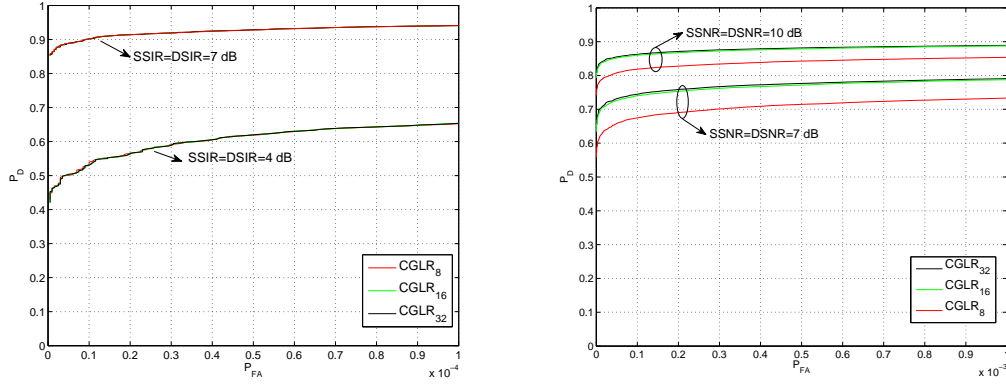
$$\underset{\varphi_k}{\text{máx}} q(\varphi_k) \underset{H_0}{\overset{H_1}{\geq}} \eta_1 \quad k = 1, \dots, K \quad (5.7)$$

La expresión (5.7) se puede interpretar como una comparación de la salida máxima de un banco de detectores LR diseñados para K valores discretos de θ con un umbral fijado por la P_{FA} deseada. Si cada suma, producto y función exponencial es considerada como una operación simple en el procesador, la forma cuadrática q requiere de $8P^2 + 6P - 2$ operaciones. Y un detector $CGLR_K$ necesita $(K) \times (8P^2 + 6P - 2) + K - 1$ operaciones. Como la dimensión del espacio de entrada considerada es 8, el número de operaciones requerido por un detector $CGLR_K$ es $559K - 1$.

Coefficiente de correlación desconocido

Los blancos gaussianos con una función de autocorrelación gaussiana pueden presentar desde una fluctuación lenta, blanco Swerling I con un coeficiente de correlación $\rho_s = 1$, hasta una fluctuación rápida, blanco Swerling II con un coeficiente de correlación $\rho_s = 0$. En caso de ser desconocida la fluctuación del blanco, el detector óptimo se puede aproximar con un detector $CGLR_K$ compuesto por K detectores LR diseñados con K valores de ρ_s equiespaciados en el margen de variación de $[0; 1]$. Las capacidades de detección de este esquema subóptimo cuando se simula con un blanco con un ρ_s uniformemente distribuido en $[0, 1]$ y $SNR = 3$ y 7 dB se representan en las Figuras 5.11(a) y 5.11(b) respectivamente.

Independientemente del valor de SNR , con un detector $CGLR_{P/2}$ se consigue una buena aproximación al detector óptimo, ya que al aumentar el valor K la mejora presentada es despreciable. En el caso de $SNR = 3$ dB, Figura 5.11(a), para un valor de K menor se produce una pérdida de P_D considerable para una misma P_{FA} . A medida que la SNR aumenta, esta caída se reduce y para $SNR = 7$ dB, Figura 5.11(b), el detector $CGLR_2$ se aproxima mucho más a las curvas obtenidas con detectores $CGLR$ con un K mayor.



(a) $\rho_s^d = \rho_s^s = 0, 5$; $DSNR = SSNR = 4$ y 7 dB (b) $\rho_s^d = \rho_s^s = 1$; $DSNR = SSNR = 7$ y 10 dB

Figura 5.12: Curvas ROC del CGLR para el caso de estudio 1 con $\Omega^d = \Omega^s$ variable uniforme en el intervalo $[0, 2\pi)$.

El estudio de las soluciones basadas en inteligencia artificial se centra en los resultados obtenidos con $SNR = 7$ dB, para los que se obtienen P_D de interés en aplicaciones radar, y un detector $CGLR_4$ como aproximación al ALR. El detector $CGLR_4$ requiere 2.235 operaciones para tomar la decisión de presencia o ausencia de blanco.

Frecuencia Doppler normalizada desconocida

Ahora se asume la detección de un blanco con un coeficiente de correlación conocido que puede estar parado o en movimiento con diferentes velocidades. La velocidad radial del movimiento determina la pulsación Doppler y el valor de Ω :

$$\Omega = \frac{\varpi_D}{PRF} = \frac{2\omega_p v_R}{c \cdot PRF} \quad (5.8)$$

Para la detección de blancos con Ω desconocida se propone un esquema basado en $CGLR_K$ implementado por un banco de detectores LR diseñados para K valores discretos de Ω en el intervalo $[0; 2\pi)$. Se han considerado dos coeficientes de correlación del blanco, $\rho_s = 0, 5$ y 1 , y los resultados se presentan en las Figuras 5.12(a) y 5.12(b) respectivamente.

Atendiendo a los resultados, se comprueba que las características de detección dependen de las características espectrales del blanco, puesto que bajo hipótesis H_0 se supone un ruido blanco gaussiano. De la Figura 5.12(a) puede concluirse que cuando el blanco está menos correlado y reparte su información en un espectro más ancho, el detector es prácticamente independiente de Ω^d . El peor caso es la detección de un blanco muy correlado, es decir, un Swerling I (Figura 5.12(b)). Tal y como se afirma en [Nayebi1996], es necesario un detector basado en un $CGLR$ de al menos $2P$ filtros para aproximar el ALR, mostrándose una gran diferencia entre $CGLR_P$ y $CGLR_{2P}$.

Para estudios comparativos posteriores nos centraremos en la detección de blancos Swerling I con $SNR = 7$ dB. El detector $CGLR_{16}$, capaz de aproximar el ALR, requiere 8.943 operaciones para su implementación, un número elevado de operaciones para sistemas de tiempo real.

5.3.2. Detectores CGLR para el caso de estudio 2

Tal y como se estudia en el apartado 5.2.2 para blancos fluctuantes gaussianos en interferencia gaussiana, el detector NP depende de las matrices de covarianza bajo ambas hipótesis. En las Figuras 5.4 y 5.7(a), se demuestra que la pérdida de detección, cuando la *SIR* asumida en el diseño es distinta de la del vector de observación, es mínima independientemente de \mathbf{M}_c . Sin embargo, los detectores LR presentan más sensibilidad con respecto a ρ_s^d y Ω^d . En este apartado, se estudia las capacidades de detección de detectores subóptimos basados en CGLR para aproximar el ALR en el caso de la detección de blancos gaussianos con ρ_s^d o Ω_d desconocidos en presencia de interferencia gaussiana.

El ALR se formula a partir de las FDPs condicionadas bajo las hipótesis H_0 y H_1 . El vector de observación generado bajo H_1 depende, a su vez, de un parámetro φ desconocido, ρ_s o Ω . Si consideramos que $f(\varphi|H_1)$ es una distribución uniforme en el intervalo $[\varphi_1; \varphi_2]$, el ALR se expresa como:

$$\Lambda(\tilde{\mathbf{z}}) = \frac{\int_{\varphi_1}^{\varphi_2} \exp(\tilde{\mathbf{z}}^T (\boldsymbol{\Sigma}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}(\varphi) + \boldsymbol{\Sigma}_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I})^{-1} \tilde{\mathbf{z}}^*) d\varphi}{\exp(\tilde{\mathbf{z}}^T (\boldsymbol{\Sigma}_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I})^{-1} \tilde{\mathbf{z}}^*)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{ALR}(P_{FA}) \quad (5.9)$$

De nuevo, el ALR presenta una integral de difícil solución que se puede aproximar con un detector $CGLR_K$, que, una vez simplificado, se puede expresar en función de la forma cuadrática $q = \tilde{\mathbf{z}}^T \mathbf{Q}(\varphi_k) \tilde{\mathbf{z}}^*$, donde $\mathbf{Q} = (\boldsymbol{\Sigma}_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I})^{-1} - (\boldsymbol{\Sigma}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}(\varphi_k) + \boldsymbol{\Sigma}_{\tilde{\mathbf{c}}\tilde{\mathbf{c}}} + 2 * \mathbf{I})^{-1}$.

$$\underset{\varphi_k}{\text{máx}} q(\varphi_k) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_1 \quad k = 1, \dots, K \quad (5.10)$$

Al igual que en el caso de estudio 1, tanto la expresión (5.7) como (5.10) dependen de una forma cuadrática que requiere de $8P^2 + 6P - 2$ operaciones. Por lo tanto, para $P = 8$, un detector $CGLR_K$ necesita $559K - 1$ operaciones para generar una decisión.

Manteniendo los valores utilizados en [Aloisio1994] y [diVito1999], las curvas ROC de los detectores $CGLR_K$ se han obtenido para los siguientes casos:

- Clutter no dominante poco correlado: $CNR = 20$ dB y $\rho_c = 0,7$.
- Clutter no dominante muy correlado: $CNR = 20$ dB y $\rho_c = 0,995$.
- Clutter dominante muy correlado: $CNR = 40$ dB y $\rho_c = 0,995$.

Coefficiente de correlación desconocido

Los resultados se han obtenido simulando los diferentes detectores con blancos con un ρ_s que varía uniformemente en el intervalo $[0, 1]$ (en [Aloisio1994] y [diVito1999] se realizaron estudios de sensibilidad para parámetros de blanco fijos distintos a los asumidos en el diseño). Partiendo del estudio de sensibilidad del LR, sabemos que los detectores diseñados con ρ_s^d alto son poco robustos, justo al contrario que los diseñados con un blanco poco correlado. Es de esperar que con dos valores discretos de ρ_s utilizados en la estimación de máxima probabilidad, $CGLR_2$, se consiga una buena aproximación del ALR, Figura 5.13. Solamente se advierten pequeñas pérdidas de detección en comparación con cualquier $CGLR_K$ de $K > 2$, para valores de *SIR* cuya probabilidad de detección es pequeña. Como en aplicaciones radar estos valores de *SIR* que

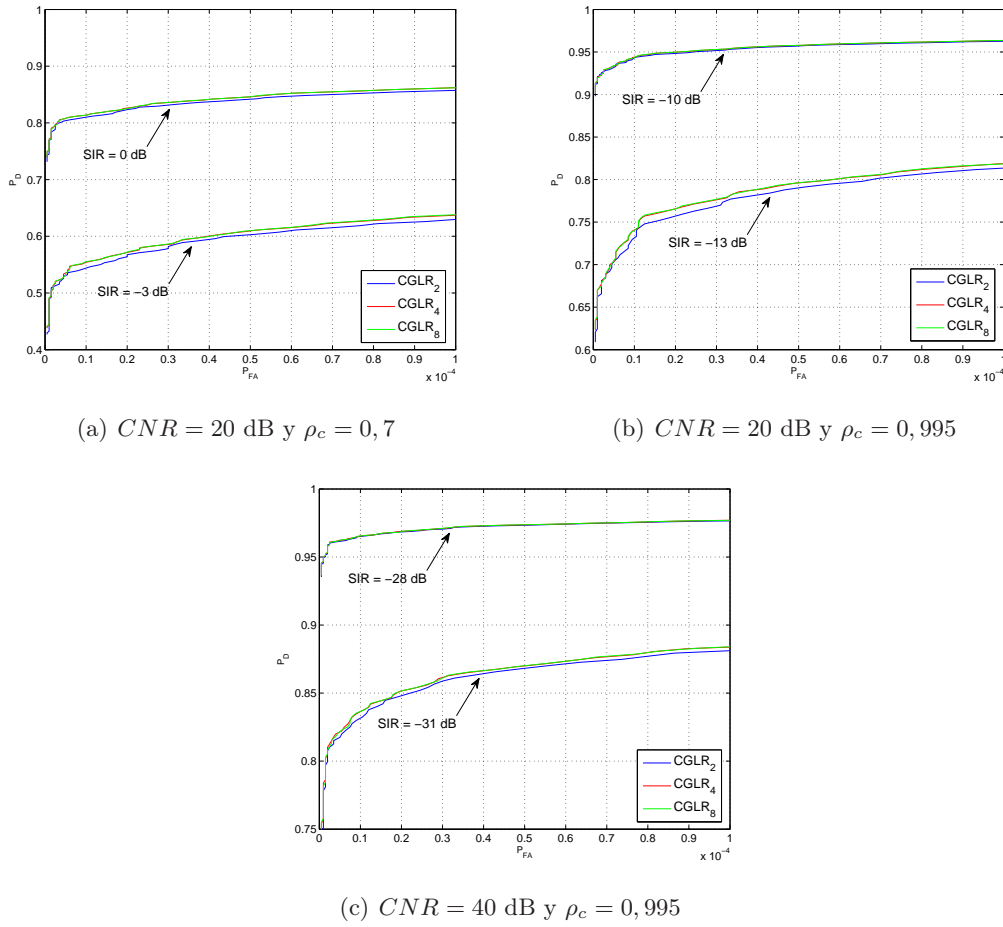


Figura 5.13: Curvas ROC del CGLR para el caso de estudio 2 con $\rho_s^d = \rho_s^s$ variable uniforme en el intervalo $[0, 1]$. $\Omega^d = \Omega^s = \pi/2$.

producen valores de P_D pequeños no son de interés, se elegirá el detector $CGLR_2$, el cuál requiere 1.117 operaciones para su implementación, como buena aproximación del ALR. Se asumió $\Omega^d = \Omega^s = \pi/2$. como caso intermedio entre $\Omega = 0$, blanco y clutter centrados en el origen, y $\Omega = \pi$, máxima separación espectral entre el blanco y el clutter.

Frecuencia Doppler normalizada desconocida

La mayor sensibilidad respecto a la frecuencia Doppler del blanco la mostraban los detectores diseñados para blancos Swerling I. La densidad espectral de estos blancos está compuesta por dos deltas en $-\Omega$ y Ω . Precisamente esta propiedad justifica las pérdidas de detección que se presentaban en las Figuras 5.6 y 5.7(c) cuando $\Omega^d \neq \Omega^s$. Las curvas ROC estimadas para un esquema basado en el $CGLR_K$, para un blanco con una Ω que varía uniformemente en el intervalo de $[0, 2\pi)$ en presencia de ruido y clutter gaussianos se presentan en la Figura 5.14.

Las potencias del blanco necesarias para conseguir probabilidades de detección de interés para aplicaciones radar, en torno al 90 %, son mucho más altas que las necesarias cuando la frecuencia Doppler es conocida. Este efecto ratifica la sensibilidad del detector LR con respecto a la Ω^d .

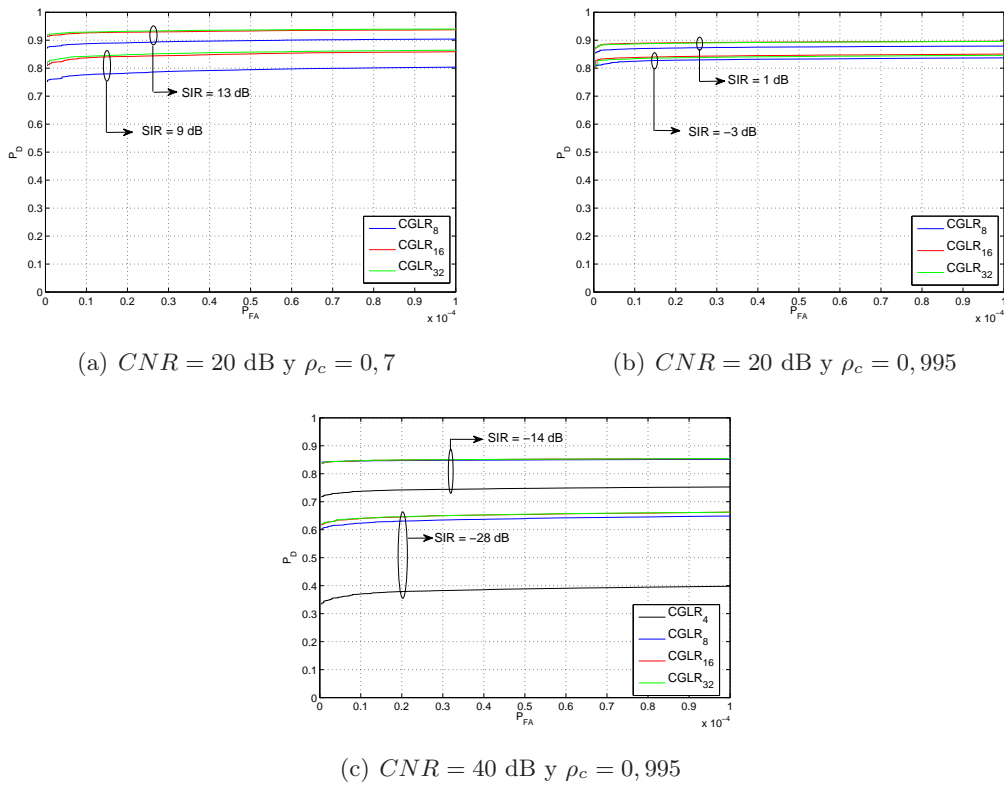


Figura 5.14: Curvas ROC del CGLR para el caso de estudio 2 con $\Omega^d = \Omega^s$ variable uniforme en el intervalo $[0, 2\pi)$. $\rho_s^d = \rho_s^s = 1$.

Para conseguir una buena aproximación al ALR se necesita un detector basado en un $CGLR_{2P}$, aunque se observa que la diferencia entre las curvas pertenecientes a $CGLR_P$ y $CGLR_{2P}$ es cada vez menor según el clutter está más correlado y su potencia con respecto al ruido, CNR, es mayor, es decir, a medida que la densidad espectral del clutter es de banda cada vez más estrecha. Esta diferencia entre las características de detección entre $CGLR_P$ y $CGLR_{2P}$, también depende de la potencia del blanco, habiendo menor diferencia entre las probabilidades de detección de cada una según la SIR es mayor. En la Figura 5.14(c), donde la CNR y ρ_c corresponden a los valores más altos utilizados, se observa que para una $SIR = -14$ dB se consigue una buena aproximación con un detector $CGLR_P$ pero se demuestra que para $K < P$ la pérdida de detección es considerable.

El detector $CGLR_{2P}=CGLR_{16}$ necesita 8.943 operaciones para obtener una buena aproximación del ALR independientemente de las características del clutter considerado. Sólo en el caso de $CNR = 40$ dB y $\rho_c = 0,995$ se podría reducir el número de operaciones sin mermar significativamente las capacidades de detección utilizando un esquema basado en el $CGLR_8$ que requiere 4.471 operaciones.

5.3.3. Detectores CGLR para el caso de estudio 3

La expresión del detector basado en el criterio de Neyman-Pearson para blancos no fluctuantes en presencia de clutter K distribuido para parámetros conocidos es una expresión compleja que conlleva el cálculo de funciones de Bessel de segunda especie. La regla de decisión (3.48) depende de la señal recibida debida al blanco modelada como $\tilde{\mathbf{s}} = a \cdot \exp(j\theta_R) \cdot \exp(j\Omega)$.

Puesto que en casos prácticos, la velocidad del blanco suele ser desconocida y se demuestra una alta sensibilidad del detector LR diseñado para una Ω^d ante $\Omega^s \neq \Omega^d$, en este apartado el estudio se centra en el diseño y la evaluación de detectores robustos de blancos con Ω desconocida. El ALR planteado para la detección de blancos no fluctuantes con una Ω variable uniformemente de $[0, 2\pi)$ en presencia de clutter K -distribuido es:

$$\frac{\int_0^{2\pi} a(\Omega)^{\nu-P} K_{\nu-P}(a(\Omega)) d\Omega}{b^{\nu-P} K_{\nu-P}(b)} \underset{H_0}{\overset{H_1}{\geq}} \eta_{ALR}(P_{FA}) \quad (5.11)$$

donde $a = \sqrt{4\nu((\tilde{\mathbf{z}} - \tilde{\mathbf{s}}(\Omega))^T \Sigma_{\tilde{\mathbf{c}\tilde{\mathbf{c}}}^{-1}} (\tilde{\mathbf{z}} - \tilde{\mathbf{s}}(\Omega))^*)}$ y $b = \sqrt{4\nu(\tilde{\mathbf{z}}^T \Sigma_{\tilde{\mathbf{c}\tilde{\mathbf{c}}}^{-1}} \tilde{\mathbf{z}}^*)}$.

El cálculo del ALR conlleva la resolución de una integral, que se puede aproximar con un estimador de máxima probabilidad entre valores discretos dentro del margen de variación de Ω según la expresión (5.12).

$$\max_{\Omega_k} \frac{a(\Omega_k)^{\nu-P} K_{\nu-P}(a(\Omega_k))}{b^{\nu-P} K_{\nu-P}(b)} \underset{H_0}{\overset{H_1}{\geq}} \eta_{CGLR}(P_{FA}) \quad (5.12)$$

La implementación del detector $CGLR_K$ asumiendo $P = 8$ y $\nu = 0,5$ requiere de $1180K - 1$ operaciones. Para este cálculo se ha supuesto que la raíz cuadrada y el acceso a un resultado tabulado de la función de Bessel de segunda especie es una única operación al igual que las sumas, productos y funciones exponenciales.

En la Figura 5.15, se representan las curvas ROC de un detector basado en un $CGLR_{2P}$ para detectar un blanco con una Ω que varía uniformemente entre $[0, 2\pi)$. Se ha elegido un $CGLR_{2P}$ porque en los apartados 5.3.1 y 5.3.2, los detectores basados en este esquema obtienen buenas aproximaciones al ALR cuando Ω es desconocida.

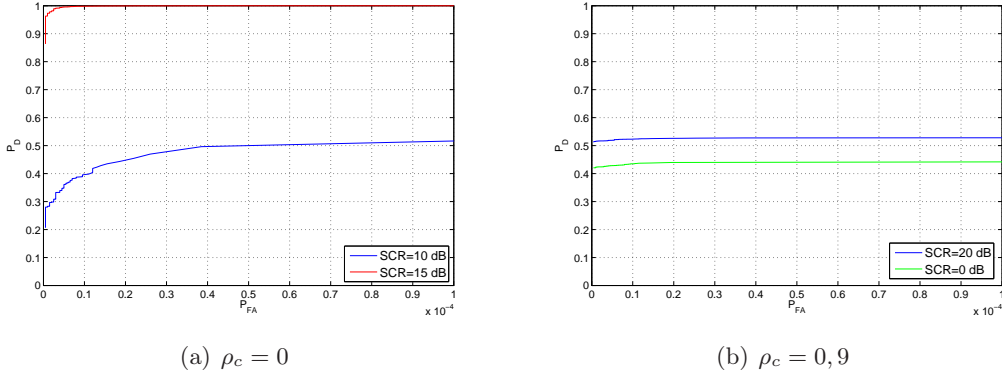


Figura 5.15: Curvas ROC del detector $CGLR_{16}$ para el caso de estudio 3 con $\Omega^d = \Omega^s$ variable uniforme en el intervalo $[0; 2\pi)$. $\theta_R = \pi/4$.

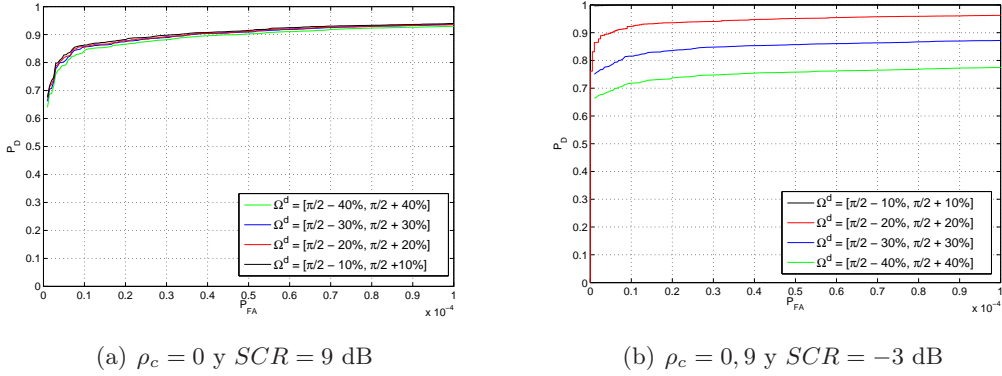


Figura 5.16: Curvas ROC del detector $CGLR_{16}$ para el caso de estudio 3 con $\Omega^d = \Omega^s$ variable uniforme en un intervalo centrado en $\Omega_c = \pi/2$. $DSCR = SSCR$ y $\Omega_R = \pi/4$.

Se ha asumido un blanco de amplitud y fase conocida, $\theta_R = \pi/4$, para centrar el estudio en los efectos de la variación de Ω en las capacidades de detección. En el caso de clutter incorrelado, Figura 5.15(a), hay que aumentar la relación SCR en unos 8 dB para obtener probabilidades de interés en aplicaciones radar, en comparación a un detector donde $\Omega^d = \Omega^s = \pi/2$ (Figura 5.10(a)). Sin embargo, cuando el clutter presenta $\rho_c = 0,9$, Figura 5.15(b), aunque se aumente en más de 30 dB la relación SCR con respecto al detector donde $\Omega^d = \Omega^s = \pi/2$ (Figura 5.10(b)), sólo se consigue superar una P_D del 50 %.

Se plantea reducir el margen de variación de la frecuencia Doppler, para diseñar detectores que puedan ser útiles en aplicaciones radar. Consideraremos una frecuencia central, Ω_0 , y un margen de variación uniforme proporcional al valor central, $[\Omega_0 - p\%, \Omega_0 + p\%]$. En la Figura 5.16, se ha ido aumentando el porcentaje $p\%$ del 10 % al 40 % y se ha obtenido la curva ROC de un detector $CGLR_{2P}$ dentro del margen de variación considerado.

Cuando el clutter, al igual que el blanco, está muy correlado, Figura 5.16(b), la respuesta del detector depende mucho del margen de variación considerado, empeorando la probabilidad de detección según aumenta el margen de variación. Si el clutter es incorrelado, esa dependencia es mucho menor y se mantiene mucho más la P_D según aumenta el margen de variación con

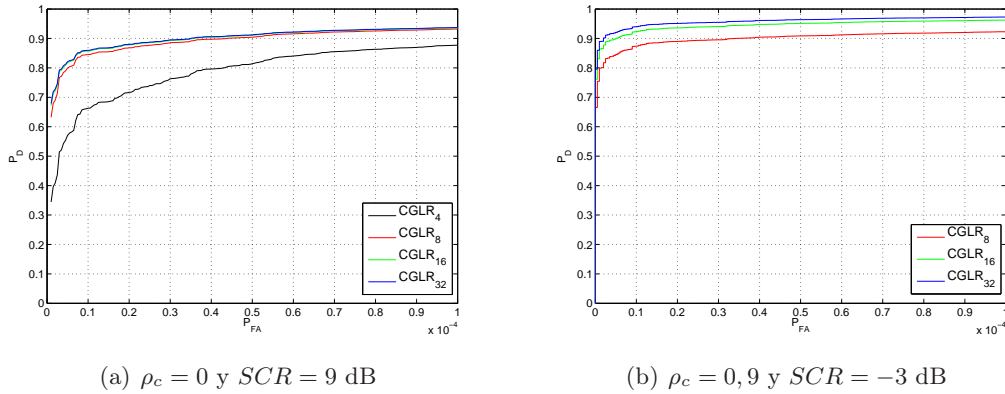


Figura 5.17: Curvas ROC del detector $CGLR_K$ para el caso de estudio 3 con $\Omega^d = \Omega^s$ variable uniforme en un intervalo $[\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\theta_R = \pi/4$.

respecto a Ω_0 .

Para hacer un estudio del mínimo número de detectores LR que es capaz de aproximar razonablemente el ALR, nos centramos en el margen de variación correspondiente al 20 %, $[\pi/2 - \pi/10; \pi/2 + \pi/10]$, y compararemos las curvas ROC obtenidas con un detector $CGLR_K$ con un K variable, Figura 5.17.

Al reducir el margen de variación de Ω , las características de detección en presencia de clutter de distribución K incorrelado permiten que un $CGLR_P$ obtenga una buena aproximación del ALR, Figura 5.17(a). Cuando la información se distribuye en espectros estrechos concentrados en dos núcleos separados por una distancia variable, blanco y clutter correlados, es necesario un detector $CGLR_{2P}$ para considerar que para $K > 2P$ la mejora introducida es muy pequeña.

Resumiendo, para comparar resultados posteriores de detectores de blancos estacionarios con una Ω variable uniformemente entre $[\pi/2 - \pi/10, \pi/2 + \pi/10]$ en presencia de clutter K distribuido incorrelado con una $SCR = 9$ dB (correspondiente a probabilidades de detección de interés en aplicaciones radar), utilizaremos la curva ROC del detector $CGLR_8$, para la que se ha necesitado llevar a cabo 9.439 operaciones. En caso de tener un clutter correlado con una $SCR = -3$ dB, la aproximación del ALR que se utilizará en los siguientes apartados es el detector $CGLR_{16}$ que necesita 18.879 operaciones para obtener resultados.

5.4. Conclusiones

En este apartado se incluyen, a modo de resumen, las principales conclusiones y aportaciones realizadas en este capítulo:

- En presencia de interferencia gaussiana, ya sea sólo ruido blanco o ruido más clutter gaussiano correlado, los detectores óptimos diseñados para una relación señal a interferencia determinada presentan una gran robustez frente a vectores de observación con un SIR diferente a la asumida en el diseño.

Si el parámetro bajo estudio es el coeficiente de correlación del blanco, se observa que el detector LR presenta pérdidas de detección principalmente cuando se diseña con un ρ_s

próximo a uno y se recibe un eco de un blanco con un ρ_s próximo a cero. El estudio de sensibilidad nos permite concluir que los diseños basados en $\rho_s \leq 0,5$ presentan una gran robustez y podrían proponerse, en función de las características de la interferencia, como una aproximación al ALR.

En cualquier caso, el parámetro más crítico en el diseño de los detectores LR en presencia de interferencia gaussiana es la pulsación Doppler. Un detector diseñado con una Ω^d presenta un comportamiento asimétrico, salvo en $\Omega^d = \pi$. Para $\Omega^d < \pi$, se produce más pérdida de detección cuando el $\Omega^d < \Omega^s$ que cuando $\Omega^d > \Omega^s$, al contrario de lo que ocurre cuando $\Omega^d > \pi$. Esta asimetría depende principalmente de las características del clutter, presentando una mayor sensibilidad los detectores diseñados con valores de Ω^d pequeños en presencia de clutter gaussiano poco correlado.

- Los detectores diseñados para el caso de blancos estacionarios con amplitud, fase y frecuencia Doppler conocidos en presencia de clutter de distribución K impulsivo presentan poca robustez frente a los parámetros asumidos en el diseño.

La sensibilidad con respecto a las diferencias entre $DSCR \neq SSCR$ y $\theta_R^d \neq \theta_R^s$ puede llegar a ser asumible si se compara con las obtenidas para $\Omega^d \neq \Omega^s$, principalmente si se considera un clutter K distribuido correlado. Puesto que en aplicaciones reales los blancos suelen estar en movimiento con distintas velocidades radiales con respecto al radar, se ha decidido centrar los posteriores estudios en esta sensibilidad crítica para proponer soluciones robustas de interés en aplicaciones radar.

- Como primer paso en la búsqueda de soluciones robustas ante blancos de parámetros desconocidos, se ha formulado el detector ALR, asumiendo que dichos parámetros eran variables aleatorias uniformes en su intervalo de variación. Como aproximación al ALR, se han estudiado detectores CGLR, con el objetivo de obtener detectores de referencia para la evaluación de las soluciones basadas en agentes inteligentes.
- El detector $CGLR_4$ realiza una buena aproximación al ALR para la detección de blancos gaussianos con coeficiente de correlación variable uniformemente en el intervalo $[0; 1]$. Aunque el detector $CGLR_2$ presenta una pérdida de detección tan leve que también se puede considerar como una buena aproximación subóptima que mantiene mejor compromiso con el coste computacional, expresado a través del número de operaciones requeridas, 2.235 y 1.117 para los detectores $CGLR_4$ y $CGLR_2$ respectivamente, para tomar una decisión.
- La mayoría de los detectores $CGLR_K$ en caso de detección de blancos con Ω variable uniforme en el intervalo $[0; 2\pi)$ requieren al menos $K = 2P = 16$ filtros LR para aproximar el ALR independientemente del modelo de clutter asumido. Estos esquemas de detección requieren un coste computacional elevado que puede resultar de difícil implementación en sistemas de tiempo real. El número de operaciones que requiere un detector $CGLR_{16}$ en presencia de interferencia gaussiana son 8.943 mientras que en presencia de clutter de distribución K son 18.879 operaciones. La diferencia entre estos dos valores pone de manifiesto el problema de diseño de soluciones robustas para la detección de blancos con Ω desconocida, principalmente en clutter K distribuido impulsivo. Tanto es así que para poder

obtener P_D de interés en aplicaciones radar se ha optado limitar el margen de variación de Ω en función del escenario radar considerado.

Capítulo 6

Soluciones basadas en agentes inteligentes

6.1. Introducción

La estrategia de entrenamiento de un agente inteligente define la interacción con el entorno y cómo deben modificarse sus parámetros libres para conseguir el funcionamiento deseado. Suelen emplearse técnicas supervisadas con el objetivo de minimizar una función error definida entre la salida de la red y la salida deseada para un conjunto de datos pre-clasificados.

En este capítulo se describen las estrategias de entrenamiento que se van a considerar para que los agentes inteligentes aprendan la información necesaria para ser capaces de aproximar un función discriminante que al ser comparada con un umbral ajustado para cumplir los requisitos de P_{FA} permitan implementar el detector NP. En la sección 4.3, se formula la condición suficiente para que un entrenamiento supervisado sea capaz de aproximar el detector de NP. Esta condición reside en la función discriminante aproximada al minimizar la función de error considerada en el aprendizaje.

Si $f(\mathbf{z}|H_i)$ es conocida, una posible implementación del NP es la regla de decisión resultante de comparar el LR con un umbral calculado para cumplir los requisitos de P_{FA} . En los casos en los que una o ambas funciones de verosimilitud dependen de uno o varios parámetros, el problema de detección se plantea como un test de hipótesis compuesto.

Sin embargo, una estrategia de entrenamiento no viene sólo determinada por la función de error, sino que también depende, por ejemplo, de la arquitectura del agente inteligente considerado. Hay arquitecturas que permiten entrenamientos basados en diferentes etapas que no tienen porqué ser todas supervisadas. Para conseguir que el entrenamiento maximice la capacidad de aproximar el detector óptimo, hay que garantizar que el entrenamiento es completamente supervisado para minimizar la función de error adecuada.

Aún así, no se puede olvidar que un agente inteligente entrenado de manera eficiente para aproximar el detector NP puede presentar malos resultados. Este inconveniente se debe a las limitaciones propias de la estructura del agente inteligente y su capacidad de aproximar diferentes formas de fronteras de decisión. Esta es la razón de porqué en esta tesis se consideran distintos tipos de agentes inteligentes:

- Perceptrones multicapa con funciones de base de primer orden (MLP).
- Redes neuronales con funciones de base de segundo orden (SONN).
- Redes neuronales con funciones de base radial (RBFNN).
- Máquinas de vectores soporte (SVM).

En el apartado 6.2 se estudiarán las estrategias de entrenamiento y las arquitecturas propuestas con el coste computacional asociado para cada uno de los tipos de NNs considerados. Mientras que para los MLPs el entrenamiento considera la minimización del error de entropía, las SONNs y RBFNNs minimizan el error cuadrático medio. Además, para las RBFNNs se describe un entrenamiento completamente supervisado basado en la utilización de algoritmos genéticos. Por último, se describen estrategias de combinación de NNs para mejorar las capacidades de detección basándose en el principio de *divide y vencerás*.

En el apartado 6.3 se estudia la capacidad matemática de aproximar el detector NP con un esquema basado en C-SVMs entenada para minimizar el error de clasificación y se demuestra experimentalmente en los casos de detección de blancos Swerling I y II en AWGN donde el detector óptimo se puede formular analíticamente.

6.2. Análisis y diseño de detectores basados en redes neuronales

6.2.1. Perceptrones multicapa, MLPs

Para garantizar que los detectores basados en MLPs con capaces de aproximar el detector óptimo, van a considerarse entrenamientos supervisados que minimicen el error de entropía. Si se eligen como funciones de activación logísticas, las salidas deseadas son $t_{H_0} = 0$ y $t_{H_1} = 1$ bajo hipótesis H_0 y H_1 respectivamente y la función discriminante aproximada se obtiene de (4.25).

$$F_0(\mathbf{z}) = \frac{P(H_1)f(\mathbf{z}|H_1)}{P(H_1)f(\mathbf{z}|H_1) + P(H_0)f(\mathbf{z}|H_0)} = P(H_1|\mathbf{z}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_0 \quad (6.1)$$

El MLP tiende a aproximar la probabilidad a posteriori de la clase H_1 , una de las funciones equivalentes de la función discriminante óptima de Bayes en problemas de clasificación binarios que al reescribir en función del cociente de verosimilitud se comprueba que la relación entre η_r y η_0 no depende del vector de entrada.

Para corregir este error, los algoritmos supervisados más extendidos para el entrenamiento de las redes neuronales están basados en estimaciones del gradiente [Haykin1999]. El más sencillo es el método del gradiente. En un perceptrón multicapa, si w_{ji} es el peso que conecta la salida de la neurona i con la neurona j de la capa siguiente, la corrección a aplicar a este peso en la iteración n -ésima se calculará a partir del error estimado $e(n)$ del modo siguiente:

$$\nabla w_{ji}(n) = -\gamma \frac{\delta e(n)}{\delta w_{ji}(n)} \quad (6.2)$$

donde γ es la constante de aprendizaje del algoritmo. Cuanto menor sea la constante de aprendizaje, menores serán las fluctuaciones del algoritmo en su trayectoria sobre la superficie de error hacia el mínimo, pero esta mejora es a costa de una menor velocidad de aprendizaje.

Al aumentar su valor, aumenta la velocidad de aprendizaje pero el algoritmo puede hacerse inestable.

Para incrementar la velocidad de aprendizaje, evitando el peligro de inestabilidad, se propuso modificar la regla anterior incluyendo un nuevo parámetro, el momento, y dando lugar a la regla generalizada:

$$\nabla \mathbf{w}_{ji}(n) = \mu \nabla w_{ji}(n-1) - \gamma \frac{\delta e(n)}{\delta w_{ji}(n)} \quad (6.3)$$

Otros algoritmos están basados en el método de Newton. Este algoritmo se basa en aproximar la función de error por una función cuadrática y minimizar esa función de forma exacta. Si \mathbf{w} es el vector que contiene todos los pesos del perceptrón multicapa, el nuevo vector de pesos en la iteración $n+1$ se obtiene a partir del vector de pesos en la iteración anterior y del error medido mediante la expresión:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mathbf{H}^{-1}(n) \nabla (e(\mathbf{w}(n))) \quad (6.4)$$

Donde \mathbf{H} es la matriz Hessiana que ha de estimarse e invertirse en cada iteración del algoritmo. Por esta razón se han desarrollado los algoritmos quasi-Newton, los cuales calculan estimaciones de la inversa de la matriz Hessiana. Uno de los métodos más aplicados es el de Davidon-Fletcher-Powell.

Como alternativa, los métodos basados en las direcciones conjugadas tienen propiedades intermedias entre el método básico del gradiente y el de Newton. Tienen como objetivos principales acelerar la lenta convergencia que suele caracterizar al algoritmo del gradiente y evitar la necesidad de calcular, almacenar e invertir, en su caso, la matriz Hessiana. El más extendido es el método del gradiente conjugado.

En esta tesis, se ha utilizado en todos los casos un entrenamiento en modo *batch* u *off-line* (por lotes) [Haykin1999], donde los pesos se ajustan utilizando el algoritmo descrito en [El-Jaroudi1990] atendiendo al error estimado con el conjunto completo de entrenamiento.

Coste computacional asociado a esquemas *MLP 2P/M/1*

Teniendo en cuenta el esquema de los perceptrones multicapa, Figura 4.3, y el esquema de detección propuesto en esta tesis, Figura 4.1, las soluciones basadas en MLPs utilizarán una arquitectura definida por:

- Una capa de entrada formada por $2P$ dimensiones correspondientes a las P partes reales e imaginarias de los vectores de observación.
- Una capa oculta con M neuronas.
- Una capa de salida formada por una única neurona cuya salida se comparará con un umbral fijado por los requisitos de P_{FA} .

Como el número de neuronas ocultas puede ser variable, utilizaremos la nomenclatura *MLP 2P/M/1* para identificar a un MLP con M unidades ocultas.

Si cada suma, producto y función exponencial es considerada como una operación simple en el procesador, los esquemas de detección basados en comparar la salida de un *MLP 2P/M/1*

con un umbral requieren de $(4P + 3)M + 2$ operaciones. Para un sistema radar que recibe $P = 8$ ecos de un blanco, el número de operaciones dependen del tamaño de la capa oculta: $35M + 2$.

6.2.2. Redes neuronales con funciones de base de segundo orden, SONN

Las estrategia de entrenamiento que se ha utilizado para modificar los pesos de la red para conseguir el funcionamiento deseado, se basa las mismas técnicas supervisadas que utilizan los MLPs. Para este tipo de arquitectura se ha optado por el error cuadrático medio como función de error y como función de activación la tangente hiperbólica. Se ha demostrado heurísticamente que la red aprende mejor con la función tangente sigmoide [Haykin1999] y en [Jarabo2005a] se demostró que la sensibilidad ante errores de entrenamiento era menor, por lo que finalmente se han entrenado las SONNs con tangentes sigmoides y, $t_{H_0} = -1$ y $t_{H_1} = 1$. Sustituyendo en (4.18), la función discriminante aproximada es

$$F_0(\mathbf{z}) = \frac{P(H_1)f(\mathbf{z}|H_1)t_{H_1} - P(H_0)f(\mathbf{z}|H_0)t_{H_0}}{f(\mathbf{z})} = P(H_1|\mathbf{z}) - P(H_0|\mathbf{z}) \stackrel{H_1}{\geq} \eta_0 \quad (6.5)$$

Si reescribimos (6.5) en función del LR nos queda la regla de decisión (4.6), en la que se demuestra que la relación entre η_0 y η_r no depende del vector de entrada y, por lo tanto, si ajustamos η_0 en función de los requisitos de P_{FA} obtendremos la máxima P_D posible.

Para entender la estrategia de entrenamiento de las SONNs, se presenta en la Figura 6.1 un esquema de una SONN con una capa compuesta por M unidades neuronales cuadráticas (QNUs). El esquema de la Figura 6.1 se puede interpretar como un MLP con una entrada compuesta por la transformación del vector de entrada en sus relaciones de primer y segundo orden, una capa oculta con M neuronas y una única neurona de salida. En la Figura 6.2, se presenta un esquema equivalente basado en un MLP con una etapa de preprocesado que convierte el espacio de entrada original de dimensión $2P$, en uno de dimensión mayor que contiene tanto las relaciones lineales como las cuadráticas de dimensión $2P^2 + 3P$. El número de pesos sinápticos de las Figuras 6.1 y 6.2 es el mismo y, puesto que son equivalentes, la función discriminante a aproximar también es la misma.

El algoritmo de aprendizaje utilizado para corregir el error cuadrático medio es el de Levenberg-Marquardt [Hagan1994], basado en el método de Gauss-Newton que estima la matriz hessiana a partir de una estimación del jacobiano.

Coste computacional asociado a esquemas SONN $2P/2P^2 + 3P/1$

Las soluciones propuestas en esta tesis, se basan en SONNs de estructura sencilla, compuestas por una QNU, para compensar la inherente complejidad computacional de este tipo de arquitectura:

- Una capa de entrada formada por $2P$ dimensiones correspondientes a las P partes reales e imaginarias de los vectores de observación.
- Una transformación del espacio de entrada a uno de dimensión superior con las relaciones de primer y segundo orden de las $2P$ componentes de entrada.

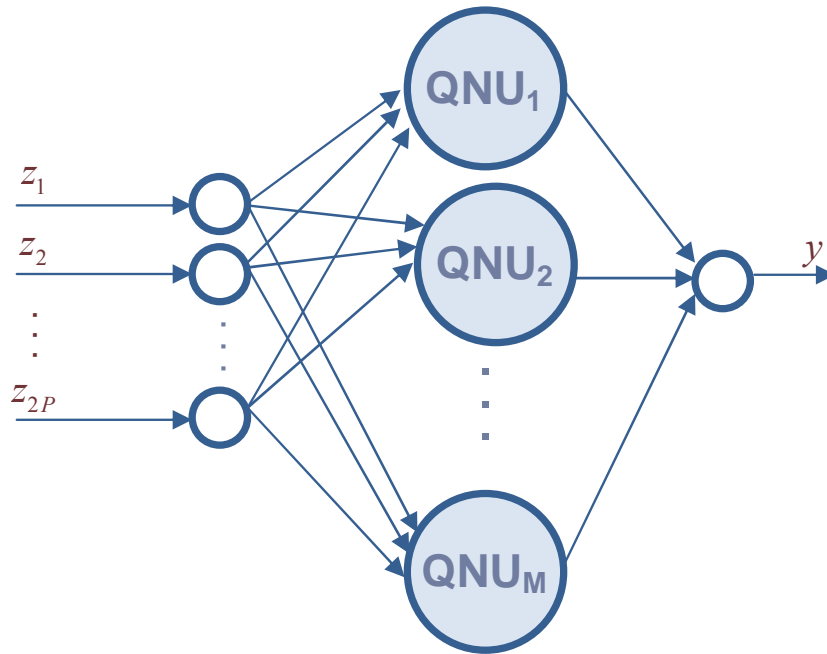


Figura 6.1: Red neuronal con una capa oculta de M unidades neuronales cuadráticas.

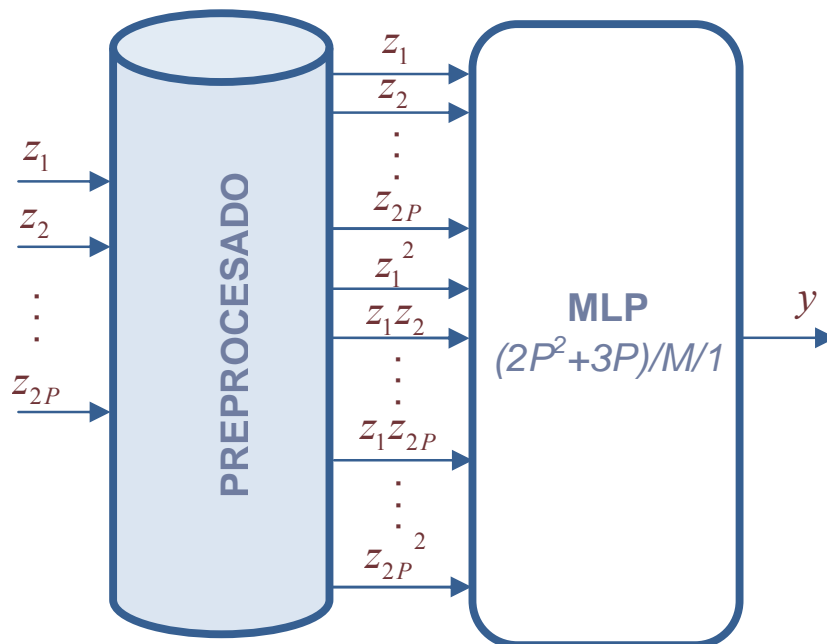


Figura 6.2: Esquema basado en un perceptrón multicapa equivalente a una red neuronal con una capa oculta de M QNU's.

- Como el aumento de información debido a las relaciones de segundo orden facilita la adquisición de conocimiento, se considerará SONNs compuestas por una QNU.
- La salida se comparará con un umbral fijado por los requisitos de P_{FA} .
- La nomenclatura que se ha utilizado para estas redes es $SONN\ 2P/2P^2 + 3P/1$.

Las soluciones basadas en tomar una decisión tras comparar la salida de una $SONN\ 2P/2P^2 + 3P/1$ con un umbral fijado por la P_{FA} requieren de $6P^2 + 7P + 1$ operaciones. Al utilizar una única unidad cuadrática, el número de operaciones sólo depende de la dimensión del espacio de entrada. Para el caso de $P = 8$, considerado en la investigación de esta tesis, el número de operaciones resultante es 441.

6.2.3. Redes neuronales con funciones de base radial, RBFNN

En muchas aplicaciones, la primera capa, compuesta por las funciones de base radial, y la capa de salida se entrenan de forma separada [Musavi1992]. En el entrenamiento por fases, los centros de las funciones base y las matrices que caracterizan las métricas a emplear se determinan en una primera etapa mediante técnicas no supervisadas utilizando los vectores de entrada \mathbf{z}_i , $i = 1, \dots, N$. Consiste en aplicar algoritmos de clustering que dividan el conjunto de vectores de entrada basándose en la similitud entre los mismos. El objetivo de esta fase es revelar la estructura interna del conjunto de vectores de entrada identificando clusters e insertar una RBF en el centro geométrico de cada cluster y con un radio relacionado con el espacio ocupado por el propio cluster. En [Moody1989], uno de los primeros trabajos sobre este tema, se utilizó el algoritmo de las C medias. Posteriormente se han propuesto otros algoritmos de clustering más potentes como el algoritmo de las C medias difuso [Karayiannis1997].

Una vez determinados sus parámetros, las funciones de base se consideran fijas, por lo que el problema de determinar los pesos de salida es equivalente al de la obtención de la función discriminante lineal generalizada. En el caso de considerar el error cuadrático medio, la solución se obtiene a partir de la matriz pseudoinversa de los datos de entrada, si esta es no singular, o a partir de su descomposición en valores singulares, si es singular. Para comprobarlo consideraremos las siguientes definiciones:

- $\mathbf{t} = [t_1, t_2, \dots, t_N]^T$: las salidas deseadas de los N patrones de entrenamiento.
- $\mathbf{w} = [\omega_1, \omega_2, \dots, \omega_M]^T$: pesos sinápticos que representan la conexión entre las M RBFs de la capa oculta con la salida.
- $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_M]^T$: centros de las M RBFs de las unidades ocultas.
- $\mathbf{G} = G(\mathbf{z}_i, \mathbf{c}_j)$ es una matriz de $N \times M$ donde $i = 1, \dots, N$ y $j = 1, \dots, M$
- $\mathbf{G}_0 = G_0(\mathbf{c}_i, \mathbf{c}_j)$ es una matriz de $M \times M$ donde $i, j = 1, \dots, M$

donde $G(\mathbf{x}, \mathbf{y}) = G(\|\mathbf{x} - \mathbf{y}\|)$ representa la función de Green de base radial. Según la teoría de la regularización, la solución viene determinada por la expresión (4.54). Para determinar el vector de pesos, \mathbf{w} , se plantea el diferencial de Tikhonov:

$$d\xi = \|\mathbf{G}\mathbf{w} - \mathbf{t}\|^2 + \lambda\|\mathbf{D}F_N^*\|^2 = (\mathbf{G}\mathbf{w} - \mathbf{t}, \mathbf{G}\mathbf{w} - \mathbf{t})_{\mathcal{H}} + \lambda(\mathbf{D}F_N^*, \mathbf{D}F_N^*)_{\mathcal{H}} \quad (6.6)$$

Al aplicar la identidad de Green nos queda $(\mathbf{D}F_N^*, \mathbf{D}F_N^*)_{\mathcal{H}} = (F_N^*, \tilde{\mathbf{D}}\mathbf{D}F_N^*)_{\mathcal{H}}$, donde el operador diferencial $L = \tilde{\mathbf{D}}\mathbf{D}$ de la función de Green $G(x, y)$ cumple la igualdad $LG(x, y) = \delta(x - y)$. Sustituyendo en el término de regularización del diferencial (6.6) queda:

$$\begin{aligned} (F_N^*, \tilde{\mathbf{D}}\mathbf{D}F_N^*)_{\mathcal{H}} &= \left(\sum_{i=1}^M \omega_i G(\mathbf{x}, \mathbf{c}_i), \tilde{\mathbf{D}}\mathbf{D} \sum_{i=1}^M \omega_i G(\mathbf{x}, \mathbf{c}_i) \right)_{\mathcal{H}} \\ &= \left(\sum_{i=1}^M \omega_i G(\mathbf{x}, \mathbf{c}_i), \sum_{i=1}^M \omega_i \delta(\mathbf{x} - \mathbf{c}_i) \right)_{\mathcal{H}} \\ &= \sum_{i=1}^M \sum_{j=1}^M \omega_j \omega_i G(\mathbf{c}_j, \mathbf{c}_i) = \mathbf{w}^T \mathbf{G}_0 \mathbf{w} \end{aligned} \quad (6.7)$$

Por lo tanto la minimización de (6.6) con respecto al vector de pesos se basa en resolver la siguiente ecuación:

$$(\mathbf{G}^T \mathbf{G} + \lambda \mathbf{G}_0) \mathbf{w} = \mathbf{G}^T \mathbf{t} \quad (6.8)$$

Según el parámetro de regularización tiende a cero, \mathbf{w} converge a la solución pseudoinversa (6.9), donde el operador $(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$ es la pseudoinversa de G .

$$\mathbf{w} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{t} \quad (6.9)$$

Estos pesos se han calculado para minimizar el error cuadrático medio y, en ese caso, F_N^* aproxima una función discriminante equivalente al cociente de verosimilitud de los vectores transformados por la primera capa y, por lo tanto, al detector de NP si la salida se compara con un umbral fijado por los requisitos de P_{FA} .

Sin embargo, las técnicas de clustering que se utilizan para obtener los parámetros libres de las RBFs no son supervisadas. El problema que surge aquí es que la parte supervisada sólo puede actuar sobre los pesos de la capa de salida, por lo que si tras la transformación de la primera capa el problema no es linealmente separable, el error será muy elevado.

Las nuevas estrategias de entrenamiento tienden a incorporar información referente a la salida esperada para determinar los parámetros de las RBFs. Esta información permite realizar una inicialización supervisada de los centros, de forma que incrementa su densidad en aquellas zonas del espacio de entrada en las que la salida de la función objetivo es más variable, en vez de en aquellas en las que haya más vectores de entrenamiento. Estos algoritmos de entrenamiento revelan la estructura subyacente de los datos de entrada preservando la homogeneidad de las salidas de los ejemplos pertenecientes a un mismo cluster. En [Gonzalez2002] se describe el algoritmo CFA (*Clustering for Function Approximation*) que permite un entrenamiento supervisado para las dos capas de la red neuronal de funciones de base radial.

Otra forma de tener en cuenta la información de la salida esperada consiste en métodos de entrenamiento basados en algoritmos evolutivos. En [Michalewicz1996], se introduce el concepto de individuo para referirse a una solución codificada con una estructura natural y cercana al problema correspondiente. También, se define el concepto de algoritmo evolutivo como un algoritmo genético en el que se habrá cambiado la representación de las soluciones por la estructura de datos más convenientes para su procesamiento y evaluación, y en el que se aplicarán operadores evolutivos que tratarán de explotar al máximo la información almacenada en cada individuo mediante el uso de conocimiento específico sobre el problema, manteniendo el resto de características

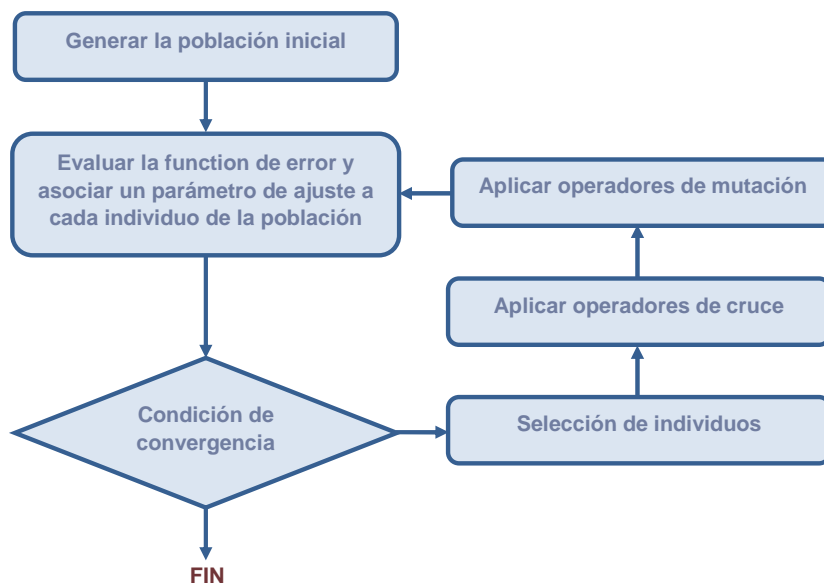


Figura 6.3: Esquema de un algoritmo evolutivo.

de los algoritmos genéticos clásicos. En la Figura 6.3 se presenta un esquema de un algoritmo evolutivo. Estos algoritmos disfrutan de todas las ventajas de los algoritmos genéticos, resuelven sus problemas de precisión, control del efecto de las mutaciones, ajuste local e incorporación de restricciones y además añaden heurísticas y conocimiento específico para ayudar en la búsqueda del óptimo global.

Los algoritmos evolutivos han sido aplicados al diseño de redes neuronales de muchas formas [Yao1999]. Harpham et al. [Harpham2004] han revisado algunos de los métodos más conocidos de aplicación de algoritmos evolutivos al diseño de RBFNN. Como característica general, concluyen que los métodos se concentran, normalmente, en una única característica, cómo los centros y/o los radios de las RBF, y determinan el resto de parámetros que conforman la red con métodos no evolutivos. Inicialmente los algoritmos usaban codificación binaria y estaban restringidos por el número de neuronas ocultas, que tenían que ser fijadas a priori. En [Whitehead1996] se trabaja con una población de neuronas que competían pero también cooperaban para encontrar la RBFNN óptima.

Los algoritmos basados en representación numérica presentan algunas ventajas como operadores rápidos, optimización dinámica del tamaño de la capa oculta, fácil cómputo de la función de evaluación, y rápida convergencia a una solución válida. En [Gonzalez2003], se describe un algoritmo evolutivo para diseñar RBFNN para aproximación de funciones. El algoritmo a través de un entrenamiento supervisado determina el tamaño de la red así como todos los parámetros de la misma. Gonzalez et al. describe operadores de cruce y mutación propuestos para evolucionar específicamente las redes que aproximan funciones [Gonzalez2003].

Aunque en esta tesis se hace especial hincapié en los aprendizajes completamente supervisados, se han implementado dos estrategias de entrenamiento basadas en la minimización del error cuadrático medio que cumplen la condición suficiente de aproximar el detector óptimo:

- La primera combina un aprendizaje no supervisado en la primera etapa con un aprendizaje

supervisado en la segunda etapa.

Los centros de las funciones base de la primera capa se determinan ajustando un modelo de mezclas gaussianas utilizando el algoritmo EM, mientras que la varianza de las gaussianas se calcula como el máximo del cuadrado de las distancias medidas entre todos los centros calculados [Lazaro2003].

Para determinar los pesos que conectan la primera capa con la de salida, se ha utilizado un método basado en el cálculo de la matriz pseudoinversa (6.9).

- La segunda estrategia implementa un aprendizaje completamente supervisado para las dos capas, asegurando que todos los parámetros de la red se han ajustado para minimizar el error cuadrático medio facilitando la capacidad teórica de aproximar el detector NP.

Los centros y radios de las funciones base de la primera capa se determinan a partir de un algoritmo genético basado en la metodología de entrenamiento descrita en [Awad2010]. Cada cromosoma representa los centros y radios de las RBFs según se define en (6.10).

$$cromosoma = [\mathbf{c}_1, r_1, \mathbf{c}_2, r_2, \dots, \mathbf{c}_M, r_M] \quad (6.10)$$

La población inicial estará compuesta por los centros obtenidos al aplicar técnicas de clustering no supervisadas basadas en el algoritmo de EM al conjunto de entrenamiento, y los radios en función del máximo del cuadrado de las distancias medias entre todos los centros calculados.

El algoritmo genético evoluciona de generación en generación hasta alcanzar los criterios de parada. Cada generación está formada por 100 individuos y se consideran dos criterios de parada: un número máximo de generaciones, o cuando se obtiene un valor de la función de evaluación menor o igual a un umbral prefijado. La función de evaluación definida es el error cuadrático medio obtenido para un set de patrones de entrenamiento y las salidas deseadas asociadas, tras haber ajustado los pesos que conectan las neuronas de la capa oculta con la salida basándose en el cálculo de la matriz pseudoinversa.

Para el paso de una generación a la siguiente se aplican operadores genéticos. En este trabajo se han considerado operadores de selección por Ranking, basado en calcular las probabilidades de reproducción atendiendo a la ordenación de la población por el valor de adaptación en vez de atender simplemente al valor de la función de evaluación, operadores de cruce aritméticos, que generan dos individuos nuevos resultado de la combinación lineal de dos padres con una proporción de cruce de 0,5, y operadores de mutación uniformes que intercambian parámetros entre los cromosomas de los individuos padre con una proporción de mutación de 0,05.

Coste computacional asociado a esquemas *RBFNN 2P/M/1*

Las soluciones basadas en RBFNNs presentadas en esta investigación se caracterizan por:

- Una capa de entrada formada por $2P$ dimensiones correspondientes a las P partes reales e imaginarias de los vectores de observación.

- Una capa oculta con M neuronas con funciones de base radial.
- La función de Green que implementan las neuronas ocultas es $G(\mathbf{z}, \mathbf{c}) = \exp\left(\frac{-\|\mathbf{z}-\mathbf{c}\|^2}{2\alpha^2}\right)$
- Una capa de salida formada por una única neurona lineal cuya salida se comparará con un umbral fijado por los requisitos de P_{FA} .

Este tipo de arquitecturas serán denominadas *RBFNN 2P/M/1*, etiqueta en la que se indica el variabilidad del tamaño de la capa oculta con el parámetro M . Los detectores basados en *RBFNN 2P/M/1* necesitan realizar $(6P + 3)M + 1$ operaciones para tomar la decisión de presencia o ausencia de blanco. Si $P = 8$, el número de operaciones sólo depende del número de neuronas en la capa oculta: $51M + 1$.

6.2.4. Comités de redes neuronales

De acuerdo con el principio de *divide y vencerás*, un problema de resolución compleja puede dividirse un problemas más sencillos cuyos resultados combinados adecuadamente también resuelvan el problema complejo. En un aprendizaje supervisado, la simplificación del problema se basa en dividir la tarea del aprendizaje en un grupo de expertos o agentes inteligentes especializados cuya salidas se combinan para formar un comité. El objetivo del comité es la fusión del conocimiento adquirido por cada experto para poder tomar decisiones supuestamente superiores a las que se hubieran tomado cada elemento por separado. Para conseguir reducir los errores de clasificación de las salidas de los componentes del comité es necesario asegurar que estas salidas sean lo suficientemente diferentes o incorreladas [Haykin1999].

Los comités son aproximadores universales que se pueden clasificar en dos categorías principalmente:

- Estructuras estáticas: Comités en los que las respuestas de los expertos se combinan según un mecanismo en el que no influye el valor de la entrada para dar lugar a la salida final de la red.

La estrategia más básica en la combinación de las salidas consiste en un combinación lineal de éstas, para aplicaciones de aproximación de funciones, o en la selección de la salida más votada, para aplicaciones de clasificación.

En la Figura 6.4 se presenta una arquitectura compuesta por una serie de agentes inteligentes (expertos) cuyas salidas se combinan linealmente para dar lugar a la respuesta del comité. Durante el proceso de aprendizaje, todos los expertos son entrenados con los mismos datos, pero pueden diferir entre sí en las condiciones iniciales del proceso de entrenamiento. Al entrenar en paralelo los agentes expertos, el tiempo de entrenamiento es mucho menor que el que se emplearía en ajustar todos los parámetros de un agente con mayor complejidad estructural. Además, dado que el número de parámetros a ajustar es menor, disminuye el riesgo de sobre-especialización sobre los datos de entrenamiento.

El hecho de utilizar el mismo conjunto de entrenamiento para todos los agentes aumenta la correlación entre los diferentes estimadores, reduciendo la posibilidad de mejorar la generalización del conocimiento con la combinación lineal de las salidas. Breiman propuso el *bagging (bootstrap aggregation)* cuyo objetivo es tender a una covarianza nula entre

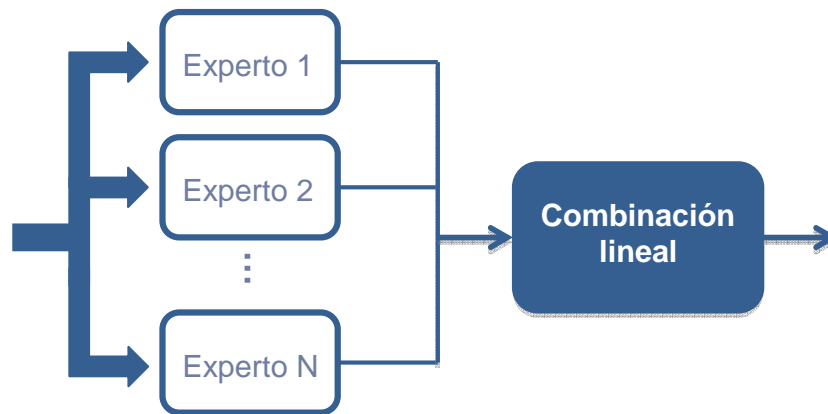


Figura 6.4: Comité de expertos de estructura estática con combinación lineal de sus salidas.

los agentes individuales generando conjuntos de entrenamiento distintos para cada uno [Breiman1994]. En caso de no tener la capacidad de generar o disponer de tantos patrones de entrenamiento, Breiman propone generar conjuntos de entrenamiento compuesto por patrones seleccionados aleatoriamente para cada elemento del conjunto original de patrones disponibles para el entrenamiento. Como consecuencia, algunos patrones del conjunto original se repetirán más de una vez en los conjuntos respectivos de los componentes del comité y otros pueden no aparecer en ninguno. Experimentalmente se ha demostrado que esta estrategia mejora a la basada en la reutilización del mismo conjunto de entrenamiento para todos los elementos.

En [Schapire1990] se describe el *boosting* por filtrado para un comité de redes neuronales. Esta estrategia de entrenamiento del comité es secuencial y, por lo tanto, el entrenamiento de cada miembro depende del resultado del anterior como se refleja en la Figura 6.5. La primera red es entrenada con un conjunto de entrenamiento de un tamaño determinado. La segunda red también es entrenada con el mismo número de patrones resultantes de simular la primera red y obtener la mitad de patrones bien clasificados y otra mitad mal clasificados. Por último, se entrena una tercera red con los patrones cuya salida de clasificación es distinta en función de si se aplican a la primera o a la segunda red.

Esta estrategia requiere disponer de un número elevado de ejemplos para el entrenamiento. *AdaBoost* (*Adaptive Boosting*) es una combinación entre *boosting* y *bagging* reduciendo el tamaño del conjunto de entrada [Freund1996]. Existen diferentes variantes de AdaBoost implementadas, por ejemplo, basadas en el submuestreo del conjunto de entrenamiento [Breiman1994] o la ponderación del mismo [Friedman1998].

- Estructuras dinámicas: Comités donde el patrón de entrada o espacio de entrada influye en la manera en que se integran las respuestas parciales de los expertos para dar lugar a la salida global de la red.

En la Figura 6.6 se presenta un esquema de lo que se denomina como mezcla de expertos. La motivación inicial de la mezcla de expertos se basa en el diseño de sistemas formados por redes neuronales responsables de modelar diferentes regiones del espacio de entrada [Sharkey1999]. Esta modularidad conduce a una mayor capacidad de adquisición de cono-

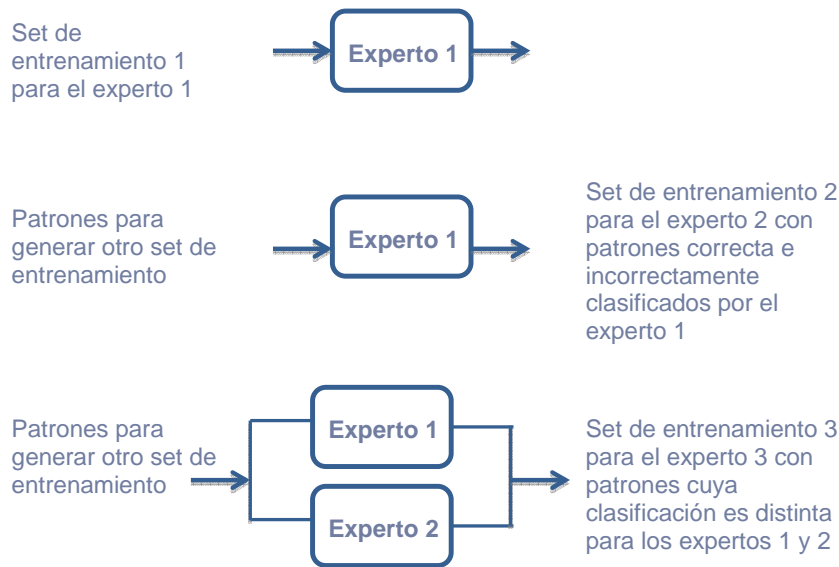


Figura 6.5: Esquema de entrenamiento de un comité de expertos con *boosting* por filtrado.

cimiento segmentado cuya combinación permite dar mayor generalización global al proceso de aprendizaje.

La salida de sistemas basados en la mezcla de expertos se obtiene a través de una estrategia de combinación que determina cómo relacionar las diferentes regiones del espacio de entrada a la salida o qué experto es el responsable de la región del espacio del patrón de entrada correspondiente en cada caso.

En esta tesis, se propone utilizar detectores basados en comités de expertos con estructuras dinámicas para mejorar la aproximación al detector de NP en tests de hipótesis compuestos. Como se describe en el capítulo 5, el espacio de entrada se caracteriza por patrones bajo hipótesis H_1 que dependen de un parámetro desconocido, φ , cuya FDP asumimos uniforme en el intervalo $[\varphi_1; \varphi_2]$. La mezcla de expertos, que se propone para simplificar el problema de entrada, se basa en especializar o entrenar cada experto en un rango de variación más pequeño de manera que entre todos los expertos se cubra todo el margen de variación, Figura 6.7. En el esquema de la Figura 6.7, no se incluye un control de la salida en función del vector de entrada, en concreto de una estimación del parámetro desconocido φ para determinar que experto se debe seleccionar como salida del comité, tal y como se proponen en el esquema de la Figura 6.6. Sin embargo, como en el entrenamiento se ha dividido el espacio de entrada, es de esperar que cada experto proporcione la salida máxima cuando el vector de entrada pertenezca a su subespacio, y por lo tanto se puede utilizar como estrategia de combinación la selección de la máxima salida de los expertos que forman la mezcla.

Para entender mejor el diseño de detectores basados en la mezcla de expertos para aproximar el detector de NP, vamos a proponer un caso de estudio de detección de blancos gaussianos con coeficiente de correlación desconocido en AWGN. El escenario considerado se definirá por los siguientes parámetros: $SNR = 7$ dB y $\rho_s^s \in [0; 1]$. Los resultados obtenidos se compararán con los estimados para el detector CGLR, utilizado como detector de referencia que aproxima el ALR.

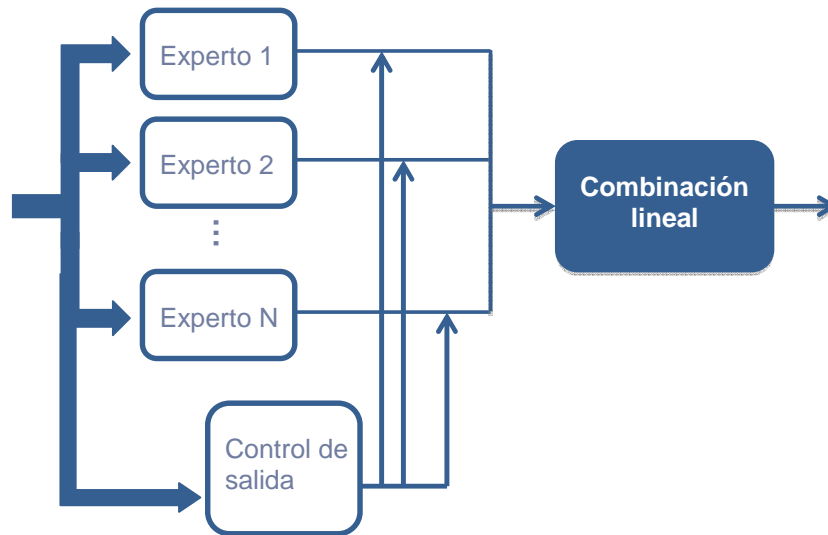


Figura 6.6: Comité de NNs con estructura dinámica: Mezcla de expertos.

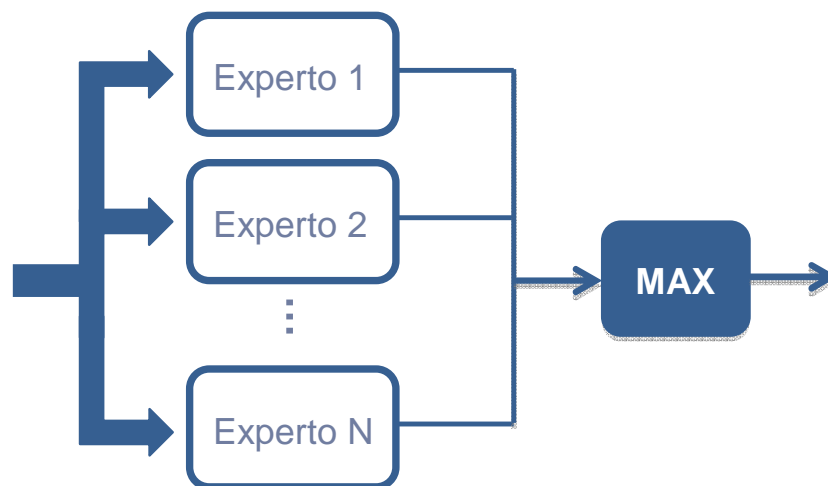


Figura 6.7: Mezcla de expertos basada en combinación por selección de máxima salida.

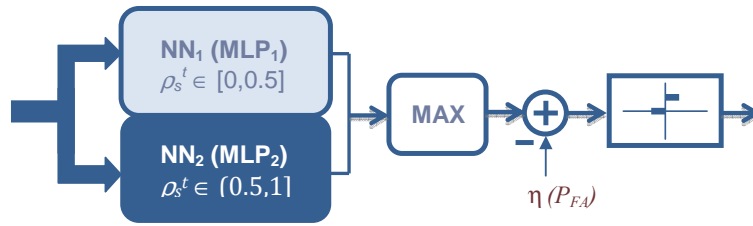


Figura 6.8: Mezcla de dos MLPs con selección por máximo: MLP_1 ha sido entrenado con blancos radar con $\rho_s^t \in [0; 0,5]$ y MLP_2 con blancos radar con $\rho_s^t \in (0,5; 1]$

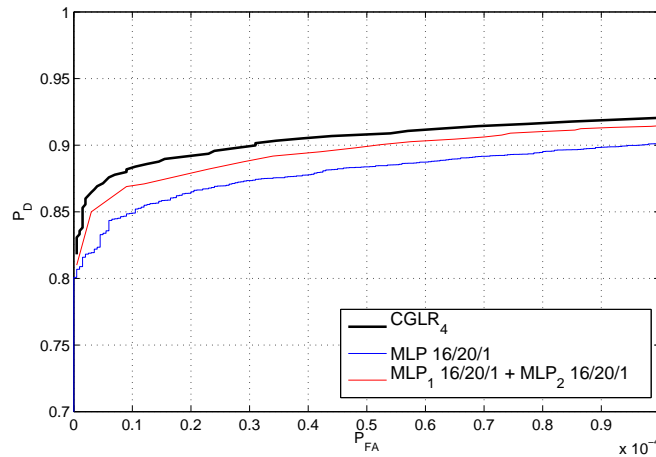


Figura 6.9: Curvas ROC de los detectores CGLR, MLP y mezcla de dos MLPs para blancos gaussianos con $\rho_s \in [0; 1]$ en AWGN. $SNR = 7$ dB.

Como punto de partida, se propone entrenar un único MLP, cuyos patrones bajo H_1 se han generado suponiendo una variación uniforme en $[0, 1]$, de forma supervisada para minimizar la entropía cruzada para cumplir la condición suficiente de aproximar el detector óptimo. Tras un estudio del tamaño de la capa oculta, que se detallará en el capítulo 7, se observa que las mejores capacidades de detección se obtienen para un MLP con 20 neuronas ocultas para $P = 8$ (16 entradas). En la Figura 6.9 se comprueba que el detector propuesto presenta peor P_D que la curva ROC del detector CGLR para una misma P_{FA} .

Para mejorar la aproximación al detector óptimo, se propone diseñar una mezcla de expertos basada en la combinación de las salidas de dos MLPs especializados en diferentes rangos de variación de ρ_s , $[0; 0,5]$ y $(0,5; 1]$. En la Figura 6.8 se presenta el esquema de la mezcla de expertos cuyas salidas se combinan por selección máxima para, finalmente, tomar la decisión en función de un umbral fijado por los requisitos de P_{FA} . Los patrones bajo hipótesis H_1 de los conjuntos de entrenamiento de los MLPs del esquema, MLP_1 y MLP_2 , se han generado con valores de ρ_s que varían uniformemente entre $[0; 0,5]$ y $(0,5; 1]$ respectivamente. Estos valores se denotan como ρ_s^t (*training* ρ_s).

El resultado de combinar dos redes MLP_1 16/20/1 + MLP_2 16/20/1 mejora ligeramente la aproximación basada en un único MLP, pero sigue existiendo un pérdida de detección con respecto al detector $CGLR_4$, Figura 6.9. Una estrategia válida para mejorar la aproximación

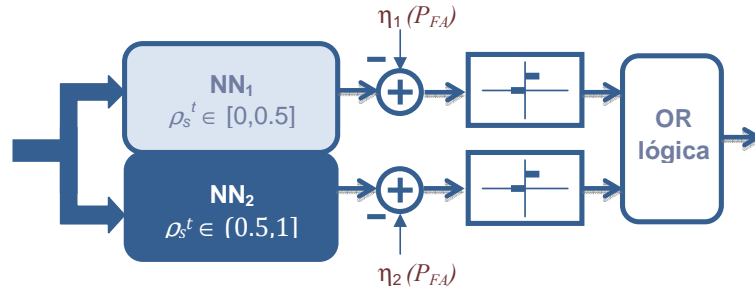


Figura 6.10: Mezcla de dos NNs con combinación basada en una OR lógica.

puede basarse en dividir el espacio de entrada de entrenamiento en subespacios más pequeños, es decir, aumentar el número de redes que se combinan en la mezcla de expertos, puesto que se ha comprobado que el comité mejora los resultados obtenidos por una única red neuronal. Otra posibilidad, se centra en combinar las salidas de distintos tipos de redes neuronales, por ejemplo, una mezcla de expertos entre una RBFNN y un MLP. La combinación de estos tipos de redes se propone a partir del estudio realizado en [Jarabo2003] que demuestra que las RBFNNs son más adecuadas para detectar blancos Swerling II ($\rho_s = 0$) en AWGN y los MLPs para blancos Swerling I ($\rho_s = 1$) en AWGN.

El problema del diseño de comités formados por distintos tipos de redes neuronales se encuentra en la forma de combinar sus salidas, puesto que como se ha visto en los apartados 6.2.1, 6.2.3 y 6.2.2, aunque todas las redes se entrenan de forma supervisada para cumplir la condición suficiente para aproximar el criterio de NP, las funciones discriminantes que aproximan las redes son equivalentes pero no idénticas. Mientras que los MLPs aproximan $P(H_1|\mathbf{z})$, las RBFNNs y las SONNs aproximan $P(H_1|\mathbf{z}_m)$, donde \mathbf{z}_m es el vector de entrada modificado por las RBFs o resultado del conjunto de las relaciones de primer y segundo orden respectivamente. En la Figura 6.10 se presenta una mezcla de expertos de diferentes tipos de redes en la que se sigue la siguiente estrategia de combinación de sus salidas:

- Para cada valor de P_{FA} , los umbrales de detección para NN_1 y NN_2 se determinarán por separado.
- Las salidas de NN_1 y NN_2 se compararán con los umbrales correspondientes para obtener una salida $y \in \{0, 1\}$.
- Se aplica una función OR lógica a las salidas tras haber umbralizado.

La idea de aplicar dos tipos de redes distintas se justifica si la arquitectura de cada una de las redes del comité es especialmente idónea para el subespacio para el que va a ser entrenada. En uno de los límites del parámetro del blanco desconocido, $\rho_s = 1$, blanco Swerling I, las regiones de decisión para la detección en presencia de AWGN son hipercilindros en el espacio de entrada que requieren de al menos 3 hiperplanos para encerrarlos. En el otro extremo, $\rho_s = 0$, las regiones de decisión para la detección de blanco Swerling II en presencia de AWGN son hiperesferas en el espacio de entrada que requieren de al menos $2P + 1$ planos para encerrarlas según se justifica en el apartado 6.3.4.

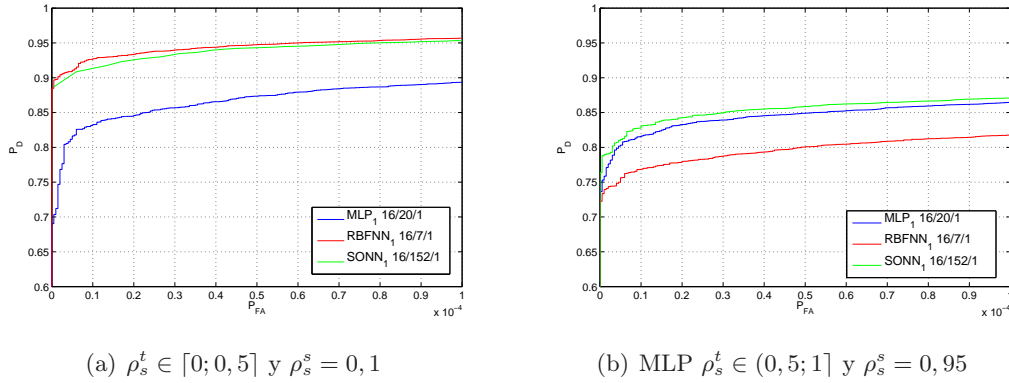


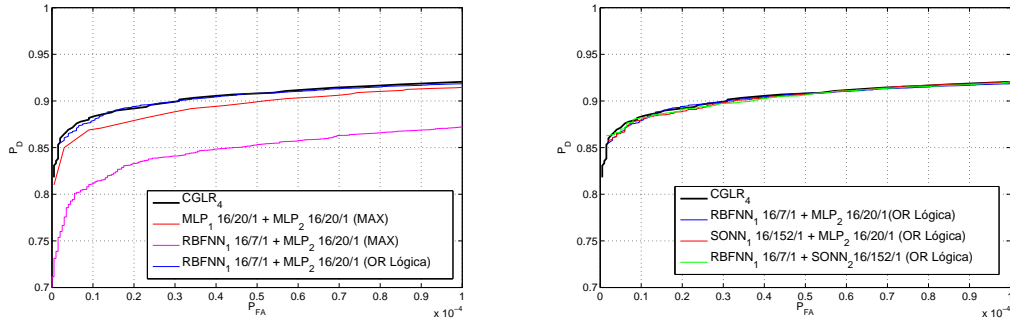
Figura 6.11: Curvas ROC de expertos (MLPs, RBFNNs y SONNs) pertenecientes de un comité simulados con un blanco con ρ_s conocidos perteneciente al margen de entrenamiento correspondiente. $SNR = 7$ dB.

Para estudiar la idoneidad de distintos tipos de red para formar las regiones definidas por hipercilindros o hiperesferas se han entrenado MLPs, SONNs y RBFNNs con patrones bajo H_1 con $\rho_s^t \in [0; 0, 5]$ y se han simulado las redes con un blanco con un coeficiente de correlación fijo $\rho_s^s = 0, 1$, Figura 6.11(a). Las mejores capacidades de detección se presentan para las SONNs y las RBFNNs que se propondrán para sustituir la red NN_1 de la Figura 6.10. En la Figura 6.11(b), se observa que los mejores tipos de red para NN_2 son los MLPs o las SONNs. En este caso las distintas redes se han entrenado con patrones bajo H_1 con $\rho_s^t \in (0, 5; 1]$ y se han obtenido las curvas ROC con un blanco simulado con $\rho_s^s = 0, 95$. En este apartado se ha considerado que las NNs tienen un arquitectura fija, dejando el estudio del tamaño de la capa oculta para el capítulo 7, para centrarse en el beneficio de la combinación de las salidas de los expertos. Las redes que se han utilizado en este apartado son MLP 16/20/1, $RBFNN$ 16/7/1 y $SONN$ 16/152/1 puesto que, como se observa en el apartado 7.3, presentan las mejores capacidades de detección para los casos de estudio. Como consecuencia las mezclas de expertos que se proponen son: $RBFNN_1 + MLP_2$, $SONN_1 + MLP_2$ y $RBFNN_1 + SONN_2$.

En la Figura 6.12(a) se compara las curvas ROC del detector basado en $MLP_1 + MLP_2$ con detectores basados en $RBFNN_1 + MLP_2$ combinados por selección máxima y basados en una OR lógica respectivamente. Como se ha demostrado en la Figura 6.11(a), la $RBFNN_1$ presenta mejores capacidades de detección que el MLP_1 para ρ_s^s pequeños, sin embargo no siempre la combinación de las salidas de $RBFNN_1$ y MLP_2 proporcionan una mejor aproximación al detector de NP. Como se ha comentado, las funciones discriminantes aproximadas por las dos redes no son idénticas y se necesita obtener el umbral asociado a cada P_{FA} de forma independiente para cada red antes de combinar sus salidas. La curva ROC basada en la mezcla de expertos $RBFNN_1 + MLP_2$ con una combinación basada en la OR lógica no sólo mejora la correspondiente al comité $MLP_1 + MLP_2$ sino que también es capaz de aproximar el detector CGLR.

En la Figura 6.12(b), se observa que la capacidad de aproximar el detector óptimo se mantiene para los detectores propuestos basados en la mezcla de expertos de dos tipos de redes distintas, con arquitecturas especialmente idóneas para el subespacio de entrada para el cuál está especializado y cuya salida se combina según el esquema de la Figura 6.10.

Resumiendo, los esquemas basados en comités de redes neuronales se plantearán en los casos



(a) Comparativa de las curvas ROC basadas en mezclas de expertos según los esquemas de las Figuras 6.8 y 6.10 (b) Curvas ROC de detectores basados en el esquema de la Figura 6.10 considerando distintas combinaciones de redes neuronales

Figura 6.12: Estudio de la capacidad de mezclas de expertos basados en MLPs, RBFNNs y SONNS de aproximar el detector óptimo de blancos gaussianos con $\rho_s \in [0; 1]$ en AWGN. $SNR = 7$ dB.

de estudio en los que una única red no sea capaz de aproximar el detector de NP. En ese caso, primero se planteará dividir el espacio de entrada para especializar los expertos y obtener una mejora en la detección a través de una combinación adecuada de las salidas de éstos. Si la mezcla de expertos supone el mismo tipo de red neuronal, la estrategia de combinación se basará en la selección de la máxima salida proporcionada por las redes. Si la mezcla es entre distintas arquitecturas de red, primero se compararán las salidas de cada red con su umbral fijado atendiendo a la P_{FA} , para posteriormente aplicar una OR lógica y obtener la salida del comité. Aunque se recomienda, siempre que sea posible, utilizar el mismo tipo de redes para la mezcla de expertos, pues sólo requieren del ajuste de un umbral.

Desde el punto de vista de coste computacional, la mejora en la aproximación al ALR basada en la combinación de dos redes supone un aumento en el número de operaciones requerido. En la Tabla 6.1, se presenta el número de operaciones necesario por los comités propuestos y se comprueba que es mayor que la solución basada en un único MLP. Este estudio no es definitivo. En realidad, al reducir el espacio de entrada en el que la red tiene que especializarse, sería necesario realizar un estudio de la arquitectura para saber si esa disminución también permite un tamaño más pequeño en la capa oculta de la red sin provocar una pérdida de detección. Por lo tanto, aunque estas soluciones suponen a priori un aumento del coste computacional, hay que estudiar con detalle la arquitectura de las redes que componen la mezcla y posteriormente calibrar si el número de operaciones resultante es competitivo en términos de mejora de la aproximación al detector óptimo.

6.3. Análisis y diseño de detectores basados en SVM

Según la teoría de las SVMs, su objetivo es crear un hiperplano de separación que minimice el riesgo estructural, $R_{est}(F) = E + \frac{1}{2C} \|\mathbf{D}F\|_K^2$, que depende de una función de error y un parámetro de regularización que define la capacidad de generalización. La función de error considerada es el error de clasificación definido en [Shawe2004], que se repite aquí por conveniencia (6.11).

Soluciones basadas en inteligencia artificial	Número de operaciones
$MLP_{16/20/1}$	702
$MLP_{16/20/1} + MLP_{216/20/1}$ (MAX)	1.404
$RBFNN_{16/7/1} + MLP_{216/20/1}$ (OR Lógica)	1.061
$SONN_{16/152/1} + MLP_{216/20/1}$ (OR Lógica)	1.144
$RBFNN_{16/7/1} + SONN_{216/152/1}$ (OR Lógica)	800

Tabla 6.1: Número de operaciones de las aproximaciones basadas en mezcla de expertos para la detección de blancos gaussianos con ρ_s desconocida en AWGN.

$$E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 - t_{H_0} F(\mathbf{z}))) + \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - t_{H_1} F(\mathbf{z}))) \quad (6.11)$$

Para estudiar la capacidad de una SVM para aproximar el detector de NP, como primer paso se propone la determinación de la función discriminante aproximada por la red durante su entrenamiento de forma análoga a como se hizo en [Jarabo2009]. Si el número de patrones de entrenamiento tiende a infinito, el parámetro de regularización tiende a anularse y, por lo tanto, el funcional a minimizar sólo depende del error de clasificación [Tikhonov1977].

En la práctica, el conjunto de entrenamiento será finito, lo más reducido posible para controlar el coste computacional, por lo que se incluirá el término de regularización con el objetivo de mejorar las capacidades generalizadoras de la red. Como resultado, la función que implementa una SVM presenta un error de aproximación a la función discriminante teórica que depende del conjunto de entrenamiento y del espacio transformado definido por la función de Green asumida en su diseño.

6.3.1. Función discriminante aproximada al minimizar el error de clasificación

A partir de la función de error objetivo a minimizar en el entrenamiento de las SVMs, (6.11), y aplicando la Ley de los Grandes Números cuando el número de patrones de entrenamiento tiende a infinito ($N \rightarrow \infty$), se llega a la expresión (6.12).

$$E_m = \int_{\mathbf{z}} (P(H_0)u(1 + F(\mathbf{z}))f(\mathbf{z}|H_0) + P(H_1)u(1 - F(\mathbf{z}))f(\mathbf{z}|H_1)) dz \quad (6.12)$$

Para encontrar la función $F_0(\mathbf{z})$ que minimiza E_m , se plantea la ecuación de Euler-Lagrange y se obtiene que $F_0(\mathbf{z})$ es la $F(\mathbf{z})$ que cumple la siguiente condición:

$$\frac{\partial I}{\partial F} = \frac{\partial}{\partial F} (P(H_0)u(1 + F(\mathbf{z}))f(\mathbf{z}|H_0) + P(H_1)u(1 - F(\mathbf{z}))f(\mathbf{z}|H_1)) = 0 \quad (6.13)$$

Para aplicar el cálculo de variaciones, I tiene que ser dos veces diferenciable. Para cumplir el requisito se ha aproximado la función escalón por la función logística, la cual se ha elegido la función logística, la cual está caracterizada por el parámetro $k > 0$, que regula la pendiente de la función en el tramo en que puede aproximarse como una función lineal (6.13). Aplicando límites cuando $k \rightarrow \infty$, la función logística tiende a la función escalón (6.14).

$$g(x) = \frac{1}{1 + \exp(-2kx)} \quad (6.14)$$

$$u(x) = \lim_{k \rightarrow \infty} \frac{1}{1 + \exp(-2kx)} \quad (6.15)$$

Sustituyendo esta aproximación en (6.13), para un valor k , se obtiene la expresión (6.16), en la que se ha considerado $t_{H_0} = -1$ y $t_{H_1} = 1$. Tras derivar respecto a $F(\mathbf{z})$, (6.17), despejar $F_0(\mathbf{z})$ y aplicar logaritmos se obtiene (6.18).

$$\frac{\partial}{\partial F} \left(\frac{P(H_0)f(\mathbf{z}|H_0)}{1 + \exp(-2k(F(\mathbf{z}) + 1))} + P(H_1) \left(1 - \frac{1}{1 + \exp(-2k(F(\mathbf{z}) - 1))} \right) f(\mathbf{z}|H_1) \right) = 0 \quad (6.16)$$

$$\frac{P(H_0)f(\mathbf{z}|H_0) \cdot 2k \cdot \exp(-2k(x + 1))}{(1 + \exp(-2k(x + 1)))^2} - \frac{P(H_1)f(\mathbf{z}|H_1) \cdot 2k \cdot \exp(-2k(x - 1))}{(1 + \exp(-2k(x - 1)))^2} = 0 \quad (6.17)$$

$$F_0(\mathbf{z}) = \frac{1}{2k} \ln \left(\frac{\sqrt{P(H_0)f(\mathbf{z}|H_0)} - \exp(-2k)\sqrt{P(H_1)f(\mathbf{z}|H_1)}}{\sqrt{P(H_1)f(\mathbf{z}|H_1)} - \exp(-2k)\sqrt{P(H_0)f(\mathbf{z}|H_0)}} \right) \quad (6.18)$$

Para comprobar si cumple la condición suficiente para aproximar el detector de NP, se parte, como hipótesis inicial, de la regla de decisión (6.19), que compara $F_0(\mathbf{z})$ con un umbral η_0 fijado para cumplir los requisitos de P_{FA} , y se estudia la relación de η_0 con el umbral asociado al cociente de verosimilitud η_{lr} . Para ello, (6.19) se reescribe en función de dicho cociente en (6.20) y, finalmente, se despeja el cociente de verosimilitud en (6.21).

$$\frac{1}{2k} \ln \left(\frac{\sqrt{P(H_0)f(\mathbf{z}|H_0)} - \exp(-2k)\sqrt{P(H_1)f(\mathbf{z}|H_1)}}{\sqrt{P(H_1)f(\mathbf{z}|H_1)} - \exp(-2k)\sqrt{P(H_0)f(\mathbf{z}|H_0)}} \right) \underset{H_0}{\overset{H_1}{\geq}} \eta_0 \quad (6.19)$$

$$\frac{1 - \exp(-2k)\sqrt{\frac{P(H_1)}{P(H_0)}\Lambda(\mathbf{z})}}{\sqrt{\frac{P(H_1)}{P(H_0)}\Lambda(\mathbf{z})} - \exp(-2k)} \underset{H_0}{\overset{H_1}{\geq}} \exp(2k\eta_0) \quad (6.20)$$

$$\Lambda(\mathbf{z}) \underset{H_1}{\overset{H_0}{\geq}} \frac{P(H_0)}{P(H_1)} \left(\frac{1 + \exp(-2k)\exp(2k\eta_0)}{\exp(-2k) + \exp(2k\eta_0)} \right)^2 = \eta_s \quad (6.21)$$

Sin embargo, en la regla de decisión del detector de NP, $\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \eta_{lr}$, los signos de comparación asociados a cada hipótesis están intercambiados, resultando que no es posible igualar η_s a η_{lr} . Lo cierto es que no se aplicó ningún criterio a la hora de escribir la regla de decisión (6.19) y simplemente se optó por la tendencia general de decidir a favor de la presencia de blanco ante salidas mayores. Para poder llegar a una solución viable, hay que cambiar los sentidos de los signos de comparación en la hipótesis de partida al comparar $F_0(\mathbf{z})$ con un umbral η_0 fijado por los requisitos de P_{FA} , (6.22), y al despejar $\Lambda(\mathbf{z})$ resulta la regla (6.23).

$$\frac{1}{2k} \ln \left(\frac{\sqrt{P(H_0)f(\mathbf{z}|H_0)} - \exp(-2k)\sqrt{P(H_1)f(\mathbf{z}|H_1)}}{\sqrt{P(H_1)f(\mathbf{z}|H_1)} - \exp(-2k)\sqrt{P(H_0)f(\mathbf{z}|H_0)}} \right) \underset{H_1}{\overset{H_0}{\geq}} \eta_0 \quad (6.22)$$

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)}{P(H_1)} \left(\frac{1 + \exp(-2k) \exp(2k\eta_0)}{\exp(-2k) + \exp(2k\eta_0)} \right)^2 = \eta_s = \eta_{lr} \quad (6.23)$$

Aplicando el límite cuando $k \rightarrow \infty$, la relación entre η_{lr} y η_0 viene expresada en (6.24). El resultado muestra que la SVM sólo es capaz de aproximar tres puntos de la curva ROC del detector de NP, correspondientes a los umbrales $\eta_{lr} = \{0, \frac{P(H_0)}{P(H_1)}, \infty\}$. La función P_{FA} versus η_0 es, por tanto, una función escalón con valores de $P_{FA} = 0$ y $P_{FA} = 1$ para $\eta_0 < 0$ y $\eta_0 > 0$, respectivamente. Cuando $\eta_0 = 0$, la P_{FA} es la asociada al criterio de mínima probabilidad de error.

$$\eta_{lr} = \begin{cases} 0 & \eta_0 > 0 \\ \frac{P(H_0)}{P(H_1)} & \eta_0 = 0 \\ \infty & \eta_0 < 0 \end{cases} \quad (6.24)$$

En la práctica, el tamaño del conjunto de entrenamiento es finito. Teniendo en cuenta el principio de funcionamiento de las SVMs, dado que el número de vectores soporte, N_{vs} , depende del tamaño del conjunto de entrenamiento, es de esperar que para una buena aproximación al detector óptimo sea necesario un coste computacional muy elevado relacionado con el valor del tamaño del conjunto de entrenamiento, N .

La relación entre el número de vectores soporte y el coste computacional viene definido por la función que implementa una SVM (6.25). Para calcular el número de operaciones requerido, hay que determinar el tipo de funciones de Green que se van a utilizar en el desarrollo de esta tesis. Las funciones de Green consideradas son de tipo cuadrático y su expresión viene definida en (6.26). Estas funciones no añaden parámetros libres para ajustar en el entrenamiento y han dado buenos resultados en tareas de detección en otros ámbitos [Paisitkriangkrai2008], [Mert2011]. El número de operaciones requerido para tomar una decisión para cada vector de entrada es igual a $N_{vs}(8P + 3)$. Para reducir este número puede utilizarse un conjunto de entrenamiento más pequeño, pero entonces es necesario tener mayor capacidad de generalización de la información adquirida en el entrenamiento. La capacidad de generalización de las SVMs depende del parámetro de regularización, un C más pequeño permite una mayor generalización a costa de permitir un mayor error en la aproximación. Diseñar un detector SVM con un coste computacional controlado y competitivo frente a los esquemas basados en NNs, requiere utilizar un tamaño del conjunto de entrenamiento lo más pequeño posible asumiendo un posible mayor error de aproximación controlado por el parámetro de regularización.

$$F_N(\mathbf{z}) = \sum_{i=1}^{N_{sv}} \alpha_i t_i G(\mathbf{s}_i, \mathbf{z}) + b \quad (6.25)$$

$$G(\mathbf{z}_i, \mathbf{z}_j) = \langle \mathbf{z}_i, \mathbf{z}_j \rangle \cdot (1 + \langle \mathbf{z}_i, \mathbf{z}_j \rangle) \quad (6.26)$$

6.3.2. Detectores basados en C-SVMs para blancos Swerling I y II en AWGN

En este apartado se estudia la función discriminante que aproxima la C-SVM y se propone un esquema de detección basado en una 2C-SVM para blancos Swerling I y II en presencia de AWGN, escenarios en los que se puede formular analíticamente el detector de Neyman-Pearson.

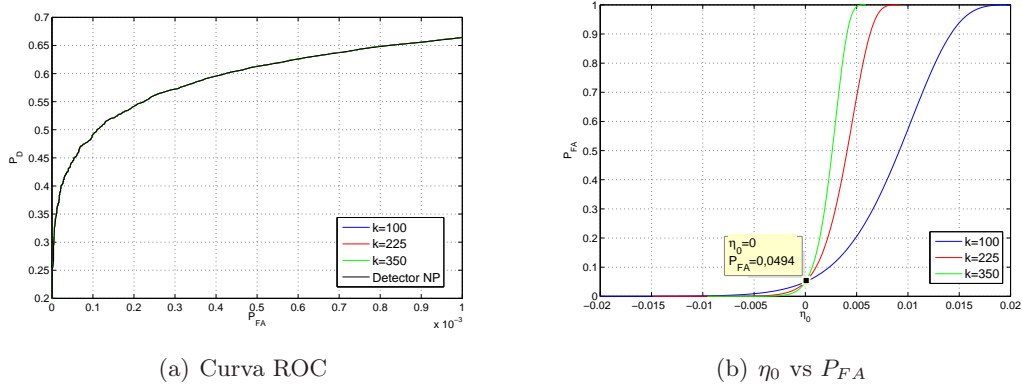


Figura 6.13: Resultados obtenidos para la regla de decisión en $F_0(\mathbf{z})$, (6.18), para blancos Swerling II en AWGN para tres valores de k . $SNR = 3$ dB.

Un blanco tipo Swerling II es un proceso $\tilde{\mathbf{s}}$ gaussiano de media cero y matriz de covarianza $\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} = 2p_s\mathbf{I} = 2snr\mathbf{I}$, siendo \mathbf{I} la matriz identidad de dimensiones $P \times P$. La regla de decisión que se obtiene después de simplificar la basada en el cociente de verosimilitud es la que se indica en (6.27).

$$\tilde{\mathbf{z}}^T [(2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2 * \mathbf{I})^{-1}] \tilde{\mathbf{z}}^* \underset{H_0}{\overset{H_1}{\gtrless}} \text{Ln} (\eta_r (snr + 1)^P) = \eta_s \quad (6.27)$$

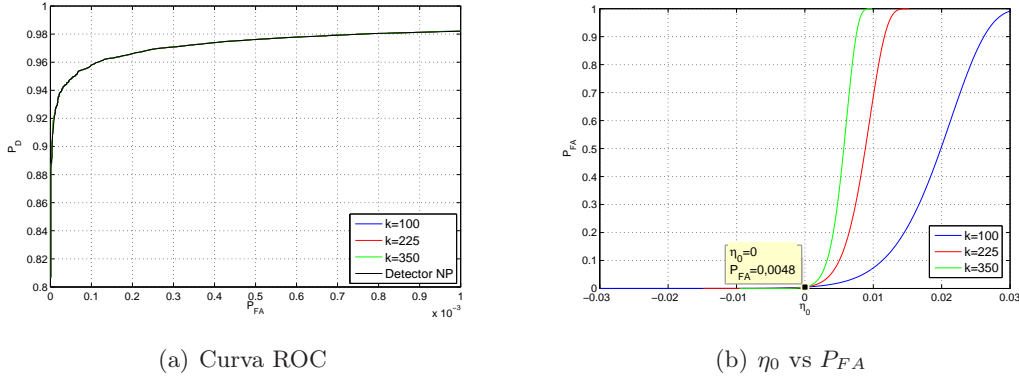
$[(2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} + 2 * \mathbf{I})^{-1}]$ es una matriz diagonal que tiene todos sus autovalores iguales, por lo que las fronteras de decisión que se forman son hipersferas en el espacio de entrada de $2P$ dimensiones. El estadístico suficiente es el módulo al cuadrado del vector de entrada y el umbral de detección, η'_s , es el cuadrado del radio de la hipersfera (6.28).

$$|\tilde{\mathbf{z}}^2| \underset{H_0}{\overset{H_1}{\gtrless}} \eta'_s \quad (6.28)$$

En las Figuras 6.13(a) y 6.14(a) se demuestra cómo para un $k > 0$ suficientemente elevado, la función $F_0(\mathbf{z})$, expresada en (6.18), permite implementar el detector NP cuando se compara con un umbral fijado para cumplir los requisitos de P_{FA} (6.22). Las curvas estimadas en estas Figuras para los distintos valores de k son indistinguibles de la obtenida para el detector NP. Para estimar la P_{FA} se han utilizado técnicas de muestreo enfatizado y la simulación de Montecarlo para estimar la P_D con un error asociado de menos del 10 %.

La expresión de la función discriminante aproximada por la SVM entrenada con $N \rightarrow \infty$, se obtiene aplicando el límite cuando $k \rightarrow \infty$, por esta razón, se han estimado las curvas ROC para valores de k cada vez más grandes hasta llegar a la limitación impuesta por la precisión de MatLab. Los resultados de este estudio se reflejan mejor en las Figuras 6.13(b) y 6.14(b), donde se representan los umbrales versus P_{FA} y se observa que según k aumenta, el intervalo de variación de los umbrales se hace más pequeño tendiendo a un único valor $\eta_0 = 0$.

El hecho de que la pendiente de la curva P_{FA} versus η_0 tienda a infinito en $\eta_0 = 0$, hace que se necesite una precisión muy elevada para obtener los umbrales dentro del margen de variación de la salida de $F_0(\mathbf{z})$. En cualquier caso, para cualquier k finito mayor que cero, la relación entre



(a) Curva ROC

(b) η_0 vs P_{FA}

Figura 6.14: Resultados obtenidos para la regla de decisión en $F_0(\mathbf{z})$, (6.18), para blancos Swerling II en AWGN para tres valores de k . $SNR = 7$ dB.

η_0 y η_{lr} no depende del vector de entrada, cumpliendo la condición suficiente para aproximar el detector NP.

Hasta este punto se ha asumido que el conjunto de entrenamiento $N \rightarrow \infty$ y $C \rightarrow \infty$, la siguiente cuestión es comprobar si estas tendencias se mantienen cuando la C-SVM es entrenada de manera supervisada con un conjunto de entrenamiento y un parámetro C finitos. Se han entrenado C-SVMs con conjuntos de entrenamiento con $P(H_0) = P(H_1)$ de diferentes tamaños: $N = 200$, $N = 1000$ y $N = 2000$ patrones. Las funciones de Green consideradas son del tipo (6.26). En la Figura 6.15 se representan las curvas ROC obtenidas tras simular C-SVMs con distintos valores de C entrenadas con los tres conjuntos de entrenamiento diseñados para detectar Swerling II en AWGN. La Figura 6.15(a) se corresponde a una $TSNR = SSNR = 3$ dB y la Figura 6.15(b) a $TSNR = SSNR = 7$ dB. Según C y N crecen, el error de aproximación se hace más pequeño. De hecho, utilizar un número de patrones de entrenamiento mayor permite obtener mejores resultados para valores de C cada vez más altos. Para los dos valores de SNR considerados, se obtiene una buena aproximación al detector NP con SVMs entrenadas para $C = 10^6$ y $N = 2000$. Si se representa la P_{FA} asociada a cada η_0 , Figuras 6.16(a) y 6.16(b), se confirma que según el error de aproximación es menor el margen de variación de los umbrales es cada vez más pequeño, al igual que ocurría en el estudio teórico de $F_0(\mathbf{z})$. En la Tabla 6.2 se resume un estudio del coste computacional en función del número de vectores soporte resultante del entrenamiento de las C-SVMs. Según aumenta N , también lo hace el número de vectores soporte y el número de operaciones requeridas. Resumiendo, se demuestra que la C-SVM es capaz de aproximar el detector NP para blancos Swerling II en AWGN aunque con dos grandes inconvenientes para aplicaciones reales: se necesita una gran precisión para establecer los umbrales asociados a cada P_{FA} y el número de operaciones requerido para obtener una buena aproximación resulta muy elevado y poco competitivo frente a otras soluciones basadas en redes neuronales.

De forma análoga, se ha realizado un estudio para blancos Swerling I, procesos $\tilde{\mathbf{s}}$ gaussianos de media cero y matriz de covarianza $\Sigma_{\tilde{\mathbf{s}}} = 2p_s \mathbf{U}$, siendo \mathbf{U} una matriz de dimensiones $P \times P$ con todos los elementos iguales a la unidad. La regla de decisión simplificada resultante es:

$$\tilde{\mathbf{z}}^T [(2 * \mathbf{I})^{-1} - (\Sigma_{\tilde{\mathbf{s}}} + 2 * \mathbf{I})^{-1}] \mathbf{z}^* \underset{H_0}{\overset{H_1}{\gtrless}} Ln(\eta_{lr} 2^P (P \cdot snr + 1)) = \eta_s \quad (6.29)$$

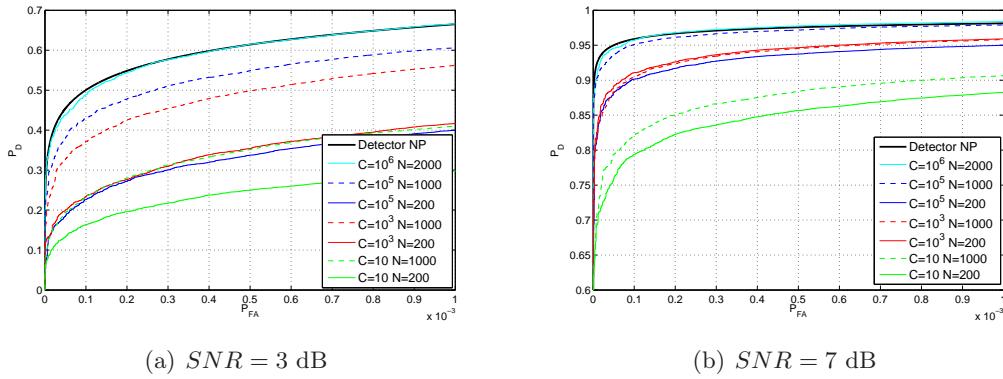


Figura 6.15: Curvas ROC de detectores C-SVM de blancos Swerling II en presencia de AWGN.

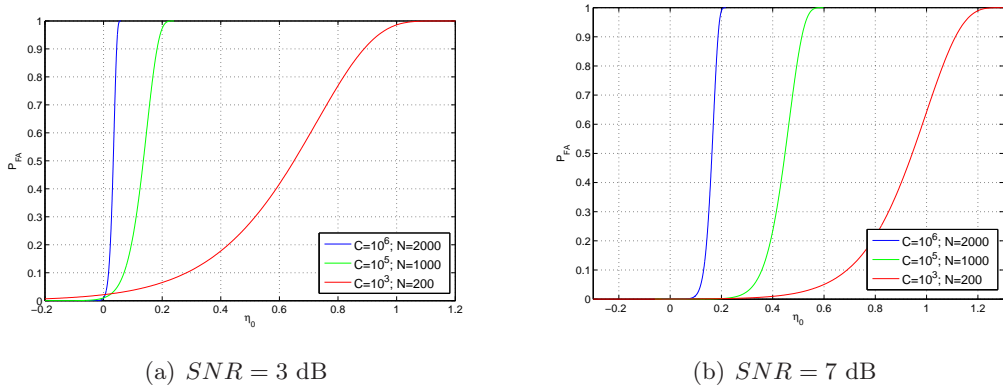


Figura 6.16: P_{FA} vs η_0 de detectores C-SVM de blancos Swerling II en presencia de AWGN.

	$TSNR = SSNR = 3$ dB		$TSNR = SSNR = 7$ dB	
	Nvs	Nop	Nvs	Nop
$C = 10^6$ y $N = 2000$	1.995	133.665	1.989	133.263
$C = 10^5$ y $N = 1000$	991	66.397	964	64.588
$C = 10^3$ y $N = 200$	175	11.725	129	8.643

Tabla 6.2: Coste computacional de los detectores C-SVM de blancos Swerling II en presencia de AWGN: N° de vectores soporte (Nvs) y N° de operaciones requerido ($Nop = Nvs(8P + 3)$).

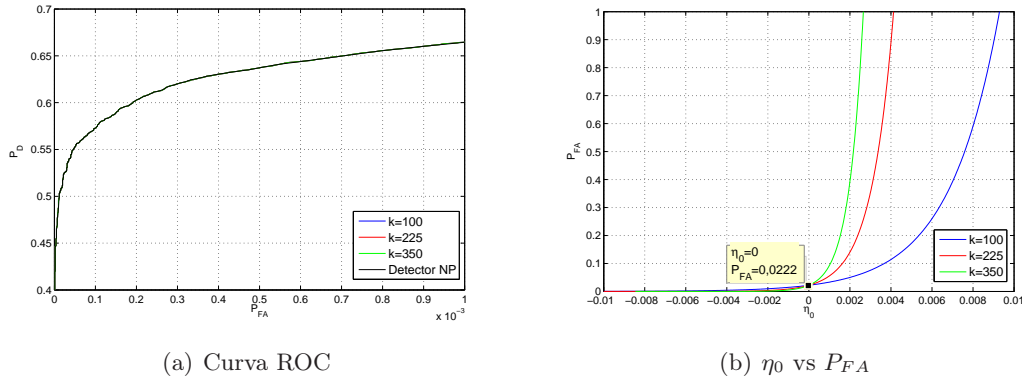


Figura 6.17: Resultados obtenidos para la $F_0(\mathbf{z})$ aproximada para blancos Swerling I en AWGN. $SNR = 3$ dB.

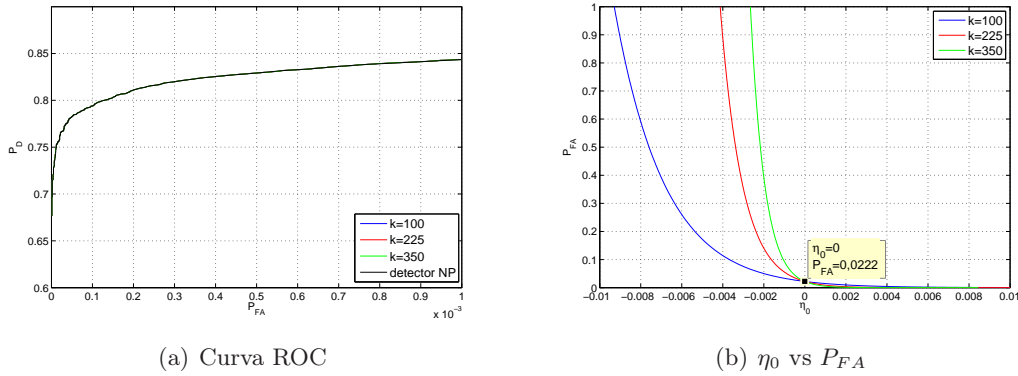


Figura 6.18: Resultados obtenidos para la $F_0(\mathbf{z})$ aproximada para blancos Swerling I en AWGN. $SNR = 7$ dB.

Sólo dos de los autovalores de $[(2 * \mathbf{I})^{-1} - (\mathbf{\Sigma}_{ss} + 2 * \mathbf{I})^{-1}]$, que pueden calcularse a partir de los autovalores de $\mathbf{\Sigma}_{zz}$, son no nulos de valor $\frac{P \cdot snr}{1 + P \cdot snr}$. En una forma cuadrática caracterizada por una matriz definida positiva, las fronteras de decisión son hiperelipsoides cuyos ejes principales son proporcionales a las raíces cuadradas de los inversos de sus autovalores. En este caso, $2P - 2$ autovalores son nulos y el hiperelipsoide degenera en un hipercilindro en el espacio de entrada de $2P$ dimensiones.

Las Figuras 6.17(a) y 6.18(a) representan las curvas ROC obtenidas por el cociente de verosimilitud y las funciones $F_0(\mathbf{z})$ obtenidas al aproximar la función escalón por una sigmoide con distintos valores de k , (6.18).

De nuevo, el margen de variación de los umbrales es cada vez menor según aumenta el valor de k . Todas las curvas que relacionan η_0 con P_{FA} , Figuras 6.17(b) y 6.18(b), se cruzan en $\eta_0 = 0$ y la P_{FA} asociada a ese punto se corresponde con el criterio de mínima probabilidad de error ($\eta_{lr} = \frac{P(H_0)}{P(H_1)}$).

Para evaluar la capacidad de aproximar el detector NP en situaciones en las que se dispone de un número finito de patrones de entrenamiento, se han diseñado conjuntos de entrenamiento para detectar blancos Swerling I en AWGN formados por $N = 200$ y $N = 1000$ patrones

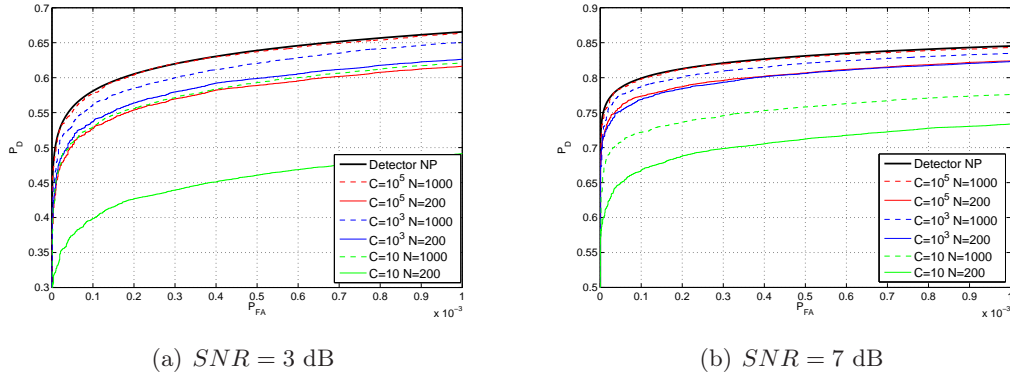


Figura 6.19: Curvas ROC de detectores C-SVM de blancos Swerling I en presencia de AWGN.

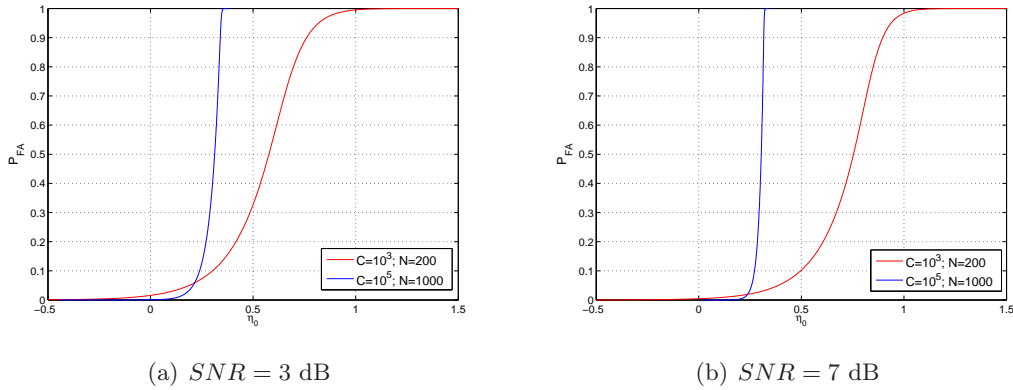


Figura 6.20: P_{FA} vs η_0 de detectores C-SVM de blancos Swerling I en presencia de AWGN.

con $P(H_0) = P(H_1)$. Las Figuras 6.19(a) y 6.19(b) muestran las curvas ROC estimadas para los esquemas basados en C-SVMs entrenadas con los conjuntos de entrenamiento mencionados y distintos valores del parámetro C que controla el parámetro de regularización. Un mayor valor de C reduce la capacidad de generalización, por lo que se necesitan más patrones de entrenamiento para mejorar los resultados. Este comportamiento de las SVMs conduce a que el error de aproximación se reducirá según N aumenta y C aumenta, tal y como se comprueba para las dos SNR estudiadas en las Figuras 6.19(a) y 6.19(b). Por otro lado, una mejor aproximación al detector de NP también conlleva una reducción del margen de variación de los umbrales η_0 , Figura 6.20, y una mayor complejidad computacional asociada al número de vectores soporte resultante del enternamiento de las C-SVM con valores de N y C grandes, Tabla 6.3.

	$TSNR = SSNR = 3 \text{ dB}$		$TSNR = SSNR = 7 \text{ dB}$	
	Nvs	Nop	Nvs	Nop
$C = 10^5$ y $N = 1000$	924	61.908	793	53.131
$C = 10^3$ y $N = 200$	150	10.050	132	8.844

Tabla 6.3: Coste computacional de los detectores C-SVMs de blancos Swerling I en presencia de AWGN: N° de vectores soporte (Nvs) y N° de operaciones requerido ($Nop = Nvs(8P + 3)$).

6.3.3. SVMs con coste asimétrico

Tal y como se ha demostrado en la sección 6.3.1, si se entrena una SVM con $N \rightarrow \infty$ y $C \rightarrow \infty$, la función discriminante aproximada permite implementar el detector de mínima probabilidad de error cuando se fija un umbral de detección $\eta_0 = 0$. En la práctica, se demuestra, que pueden obtenerse buenas aproximaciones al detector NP entrenando con un número patrones de entrenamiento y un valor de C finitos, tanto mejores cuanto mayores son N y C , aunque a consta de una elevada carga computacional.

A continuación se propone una modificación de la función de error a minimizar en el proceso de entrenamiento, cuyo objetivo es reducir el error en la aproximación del detector NP en un punto de la curva ROC diferente al de mínima probabilidad de error. La nueva función de error es la que se define en (6.30), en la que:

- C_{01} es el coste asociado a las pérdidas de detección (haber decidido por H_0 cuando H_1 es cierta).
- C_{10} es el coste asociado a las falsas alarmas (haber decidido por H_1 cuando H_0 es cierta).
- Se asume que los costes asociados a los aciertos son iguales a cero, $C_{00} = C_{11} = 0$.

$$E = C_{10} \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 - t_{H_0} F(\mathbf{z}))) + C_{01} \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - t_{H_1} F(\mathbf{z}))) \quad (6.30)$$

De forma análoga a como se hizo en el apartado 6.3.1, si se aproxima la función escalón de (6.30) por sigmoides cuya pendiente en la zona lineal está controlada por el parámetro k , la función discriminante que minimiza este error es la que se expresa en (6.31) y la regla de decisión resultante de comparar dicha función discriminante con un umbral fijado para cumplir los requisitos de P_{FA} , η_0 , es la que se indica en (6.32) (donde ya se han cambiado los signos de la desigualdad como en (6.22)).

$$F_0(\mathbf{z}) = \frac{1}{2k} \ln \left(\frac{\sqrt{C_{10}P(H_0)f(\mathbf{z}|H_0)} - \exp(-2k)\sqrt{C_{01}P(H_1)f(\mathbf{z}|H_1)}}{\sqrt{C_{01}P(H_1)f(\mathbf{z}|H_1)} - \exp(-2k)\sqrt{C_{10}P(H_0)f(\mathbf{z}|H_0)}} \right) \quad (6.31)$$

$$\frac{1}{2k} \ln \left(\frac{\sqrt{C_{10}P(H_0)f(\mathbf{z}|H_0)} - \exp(-2k)\sqrt{C_{01}P(H_1)f(\mathbf{z}|H_1)}}{\sqrt{C_{01}P(H_1)f(\mathbf{z}|H_1)} - \exp(-2k)\sqrt{C_{10}P(H_0)f(\mathbf{z}|H_0)}} \right) \underset{H_1}{\overset{H_0}{\geq}} \eta_0 \quad (6.32)$$

Despejando el cociente de verosimilitud en (6.32) se obtiene (6.33) y la relación entre η_{lr} y η_0 para cualquier valor de P_{FA} asumiendo k igual a un número finito mayor que cero.

$$\Lambda(\mathbf{z}) \underset{H_0}{\overset{H_1}{\geq}} \frac{C_{10}P(H_0)}{C_{01}P(H_1)} \left(\frac{1 + \exp(-2k)\exp(2k\eta_0)}{\exp(-2k) + \exp(2k\eta_0)} \right)^2 = \eta_{lr} \quad (6.33)$$

Si se calcula el límite cuando $k \rightarrow \infty$, puede concluirse que la SVM entrenada con $N \rightarrow \infty$ y $C \rightarrow \infty$ sólo es capaz de aproximar el punto de la curva ROC correspondiente al coste medio mínimo definido por los costes C_{01} y C_{10} . Para este punto, $\eta_0 = 0$, y de (6.33) se concluye que $\eta_{lr} = \frac{C_{10}P(H_0)}{C_{01}P(H_1)}$. De esta forma, para una P_{FA} deseada, C_{10} y C_{01} se elegirán de modo que η_{lr} sea el requerido por el LR para esa P_{FA} y si se ha utilizado la función de error (6.30) para esos costes y $k \rightarrow \infty$, η_0 será igual a cero siempre.

	$TSNR = SSNR = 3 \text{ dB}$		$TSNR = SSNR = 7 \text{ dB}$	
	η_r	P_D	η_r	P_D
$P_{FA} = 10^{-3}$	70,4286	0,6651	7,1037	0,9813
$P_{FA} = 10^{-4}$	665,6667	0,5004	122,4568	0,9589

Tabla 6.4: η_r y P_D asociados a los requisitos de P_{FA} para el detector óptimo de blancos Swerling II en presencia de AWGN.

Existe una formulación de las SVMs, conocida como 2C-SVM, que depende del parámetro C (coste asociado al error total) y del parámetro $\gamma \in [0, 1]$ que permite desbalancear el coste asociado al error total [Davenport2006]. Es posible relacionar γ con los costes utilizados en (6.30), para obtener: $\gamma = C_{10}$ y $(1 - \gamma) = C_{01}$. Si se fija $\gamma = 0,5$, $C_{10} = C_{01}$ y la 2C-SVM pasa a ser una C-SVM.

La formulación primal de 2C-SVM es la siguiente:

$$\min_{F \in \mathcal{H}, \xi, \gamma} C\gamma \sum_{i \in H_0} \xi_i + C(1 - \gamma) \sum_{i \in H_1} \xi_i + \frac{1}{2} \|DF\|^2 \quad (6.34)$$

Sujeto a

$$\begin{aligned} t_i F(\mathbf{z}_i) &\geq 1 - \xi_i & i = 1, \dots, N \\ \xi_i &\geq 0 & i = 1, \dots, N \end{aligned} \quad (6.35)$$

Y la formulación dual simplificada aplicando los multiplicadores de Lagrange es la que se indica a continuación:

$$\min \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j t_i t_j G(\mathbf{z}_i, \mathbf{z}_j) - \sum_{i=1}^N \alpha_i \quad (6.36)$$

Sujeto a:

$$\begin{aligned} 0 &\leq \alpha_i \leq C\gamma & i \in H_0 \\ 0 &\leq \alpha_i \leq C(1 - \gamma) & i \in H_1 \\ \sum_{i=1}^N \alpha_i t_i &= \sum_{i=1}^N \omega_i = 0 \end{aligned} \quad (6.37)$$

Según la formulación dual, para cada valor del parámetro γ se obtiene una solución distinta. Si se desea diseñar un detector basado en una SVM con un umbral $\eta_0 = 0$, que se ajuste a los requisitos de P_{FA} , es necesaria una búsqueda en rejilla de la pareja de parámetros C y γ que permita obtener la máxima P_D posible correspondiente al detector óptimo.

6.3.4. Detectores basados en 2C-SVMs para blancos Swerling I y II en AWGN

En el apartado 6.3.3 se demuestra que para $\eta_0 = 0$, el punto aproximado de la curva ROC del detector NP se puede ajustar con el parámetro γ de las 2C-SVMs. En la Tabla 6.4 se adjuntan valores de η_r asociados a distintas P_{FA} y sus correspondientes P_D del detector NP.

A continuación se presentan los resultados de detectores 2C-SVMs entrenados con los mismos conjuntos de entrenamiento ya diseñados con $P(H_0) = P(H_1)$, $N = 200$ y $N = 1000$, y cuya

	$N = 200$			$N = 1000$		
$[\gamma; C]$	P_{FA}	P_D	Nvs	P_{FA}	P_D	Nvs
$[0, 9860; 10^5]$	0,0014	0,4203	197	0,0011	0,6519	993
$[0, 9860; 10^3]$	0,0013	0,4257	177	0,0014	0,6061	664
$[0, 9985; 10^5]$	0,00014	0,2362	199	0,00012	0,4966	996
$[0, 9985; 10^3]$	0,00008	0,2348	172	0,00014	0,5210	645

Tabla 6.5: Capacidades de detección para detectores 2C-SVMs de blancos Swerling II en presencia de AWGN para un umbral fijo $\eta_0 = 0$. $TSNR = SSNR = 3$ dB

	$N = 200$			$N = 1000$		
$[\gamma; C]$	P_{FA}	P_D	Nvs	P_{FA}	P_D	Nvs
$[0, 8766; 10^5]$	0,0024	0,9605	198	0,0011	0,9815	999
$[0, 8766; 10^3]$	0,0018	0,9673	132	0,0015	0,9740	273
$[0, 9919; 10^5]$	0,00017	0,9196	197	0,00012	0,9585	998
$[0, 9919; 10^3]$	0,00019	0,9224	148	0,00014	0,9522	516

Tabla 6.6: Capacidades de detección para detectores 2C-SVMs de blancos Swerling II en presencia de AWGN para un umbral fijo $\eta_0 = 0$. $TSNR = SSNR = 7$ dB

salida es comparada con un umbral fijo $\eta_0 = 0$ para tomar la decisión de presencia o ausencia de blanco. Aunque el umbral es fijo, la P_{FA} y P_D se puede aproximar a un punto cualquiera de la curva ROC del detector NP debido a la relación existente entre γ y η_{tr} :

$$\eta_{tr} = \frac{C_{10}P(H_0)}{C_{01}P(H_1)} = \frac{C_{10}}{C_{01}} = \frac{\gamma}{(1-\gamma)} \quad (6.38)$$

Para detectar blancos Swerling II en AWGN, se han entrenado 2C-SVMs para los valores γ asociados a $P_{FA} = 10^{-3}$ y 10^{-4} y distintos valores de C . En la Tablas 6.5 y 6.6 se presentan las probabilidades estimadas para $\eta_0 = 0$ y $SNR = 3$ y 7 dB respectivamente. Los resultados demuestran que un entrenamiento de una 2C-SVM basado en un valor concreto de γ permite aproximar un punto cualquiera de la curva ROC del detector NP aunque la salida de la 2C-SVM se compare siempre con el mismo umbral fijo $\eta_0 = 0$. El error de la aproximación se reduce, al igual que ocurría con los detectores C-SVM, según N y C con cada vez más grandes independientemente del valor de SNR . Además, una vez fijada la SNR y el parámetro C , el número de vectores soporte es aproximadamente el mismo para los valores de γ considerados en el entrenamiento. Si no se conoce a priori el η_{tr} asociado a una P_{FA} , la forma de maximizar la P_D obtenida con una solución basada en una 2C-SVM es realizando una búsqueda en rejilla entre los parámetros γ y C que permitan aproximar el comportamiento del detector óptimo para ese punto con el menor coste computacional asociado.

Se ha repetido el ejercicio para el caso de detección de blancos Swerling I en AWGN. En la Tabla 6.7 se presentan los valores de η_{tr} y P_D asociadas a las P_{FA} consideradas en este estudio.

Los resultados para los valores de γ correspondientes a los umbrales η_{tr} de la Tabla 6.7 se presentan en la Tabla 6.8 para una $SNR = 3$ dB y en la Tabla 6.9 para una $SNR = 7$ dB. Las Tablas 6.8 y 6.9 reflejan cómo la P_{FA} viene fijada por el valor de γ con el que se ha entrenado

	$TSNR = SSNR = 3$ dB		$TSNR = SSNR = 7$ dB	
	η_r	P_D	η_r	P_D
$P_{FA} = 10^{-3}$	37,6100	0,6679	16,1527	0,8459
$P_{FA} = 10^{-4}$	343,8276	0,5819	139,8451	0,7701

Tabla 6.7: η_r y P_D asociados a los requisitos de P_{FA} para el detector óptimo de blancos Swerling I en presencia de AWGN.

$[\gamma; C]$	$N = 200$			$N = 1000$		
	P_{FA}	P_D	Nvs	P_{FA}	P_D	Nvs
$[0, 9741; 10^5]$	0,0012	0,6594	194	0,0010	0,6674	996
$[0, 9741; 10^3]$	0,0012	0,6548	162	0,0011	0,6600	478
$[0, 9971; 10^5]$	0,00012	0,5708	198	0,00009	0,5810	997
$[0, 9971; 10^3]$	0,00013	0,5722	139	0,00008	0,5686	534

Tabla 6.8: Capacidades de detección para detectores 2C-SVMs de blancos Swerling I en presencia de AWGN para un umbral fijo $\eta_0 = 0$. $TSNR = SSNR = 3$ dB

la 2C-SVM, aunque el umbral siempre sea el mismo $\eta_0 = 0$, mientras que el valor de P_D se aproxima más al óptimo según N y C se incrementan.

A partir de los resultados obtenidos, las soluciones basadas en SVMs planteadas para aproximar el detector NP en los casos de estudio planteados en esta Tesis se centrarán en detectores 2C-SVMs que maximicen la P_D para una $P_{FA} = 10^{-4}$. En cada caso, se llevará a cabo una búsqueda en rejilla de los parámetros C y γ que cumplen los requisitos de P_{FA} . Puesto que la complejidad computacional depende del número de patrones de entrenamiento, se ha propuesto utilizar a partir de ahora $N = 1000$ patrones como compromiso entre el error de aproximación y el número de operaciones requerido.

6.4. Conclusiones

En este capítulo se ha llevado a cabo un estudio de las capacidades de aproximar el detector óptimo y la complejidad computacional asociada a los esquemas de detección basados en inteligencia artificial. Se han considerado soluciones basadas en redes neuronales (MLPs, RBFNNs y

$[\gamma; C]$	$N = 200$			$N = 1000$		
	P_{FA}	P_D	Nvs	P_{FA}	P_D	Nvs
$[0, 9417; 10^5]$	0,0015	0,8465	199	0,0010	0,8456	925
$[0, 9417; 10^3]$	0,0015	0,8494	125	0,0013	0,8465	332
$[0, 9929; 10^5]$	0,00011	0,7983	198	0,00010	0,7992	996
$[0, 9929; 10^3]$	0,00014	0,7951	114	0,00013	0,8004	349

Tabla 6.9: Capacidades de detección para detectores 2C-SVMs de blancos Swerling I en presencia de AWGN para un umbral fijo $\eta_0 = 0$. $TSNR = SSNR = 7$ dB.

NN	Función de error	Función aproximada
MLP	Entropía cruzada	$F_0(\mathbf{z}) = P(H_1 \mathbf{z})$
SONN	Error cuadrático medio	$F_0(\mathbf{z}) = P(H_1 \mathbf{z}) - P(H_0 \mathbf{z})$
RBFNN		

Tabla 6.10: Funciones de error utilizadas en los entrenamientos supervisados de las soluciones basadas en redes neuronales y las funciones aproximadas resultantes.

SONNs) y SVMs. A continuación se detallan las conclusiones y aportaciones más relevantes:

- En la Tabla 6.10 se resumen las funciones discriminantes aproximadas por cada tipo de red, para la función de error considerada en su entrenamiento supervisado. Como las reglas de decisión resultantes de comparar las funciones aproximadas con un umbral fijado bajo las condiciones de P_{FA} cumplen la condición suficiente definida en [Jarabo2009], en los tres casos se concluye que un detector basado en las NNs propuestas entrenadas para minimizar los errores asociados puede aproximar el detector NP.
- Aunque la función de error cumpla la condición suficiente para aproximar el detector de NP, el error de aproximación de los esquemas basados en NNs dependerá de la idoneidad de la arquitectura de la NN considerada, del conjunto de patrones diseñado para el entrenamiento y del algoritmo de aprendizaje.
- En el caso de los MLPs y las SONNs, el ajuste de los pesos de conexión se puede hacer completamente supervisado basado en estimaciones del gradiente. Sin embargo, el entrenamiento por fases de las RBFNNs no garantiza que todos los parámetros libres se ajusten para minimizar la función de error considerada. En esta tesis, se utiliza un entrenamiento basado en un algoritmo genético que ajustar los centros y radios de las RBFs de forma supervisada, en lugar de utilizar técnicas de segmentación tradicionales. Con esta estrategia se fomentará la máxima eficiencia para aproximar la función discriminante que resulta de minimizar el error cuadrático medio.
- Los esquemas de detección basados en redes neuronales se caracterizan por una capa de entrada con $2P$ dimensiones correspondientes a las P componentes en fase y P componentes en cuadratura del vector de observación $\tilde{\mathbf{z}}$. En las SONNs este espacio de entrada se transforma en uno de dimensión $2P^2 + 3P$ correspondiente al conjunto de las relaciones de primer y segundo orden. Mientras que en los MLPs y las RBFNNs se ha considerado una capa oculta de dimensión M variable, en las SONNs se ha considerado $M = 1$, para minimizar el número de unidades neuronales cuadráticas que llevan asociado un coste computacional muy elevado. La capa de salida es una única neurona, cuya salida se compara con un umbral fijado atendiendo a los requisitos de P_{FA} . En la Tabla 6.11 se hace un estudio de la complejidad computacional de las arquitecturas consideradas asumiendo que cada suma, producto y función exponencial es considerada como una operación simple en el procesador.
- En los test de hipótesis compuestos en los que los errores de aproximación obtenidos con esquemas basados en una única red son considerables, se propone aplicar soluciones basadas

NN	Número de operaciones	Número de operaciones con $P = 8$
MLP	$(4P + 3)M + 2$	$35M + 2$
SONN	$6P^2 + 7P + 1$	441
RBFNN	$(6P + 3)M + 1$	$51M + 1$

Tabla 6.11: Complejidad computacional asociada a soluciones basadas en redes neuronales.

en comités de NNs. Para mejorar las capacidades de aprendizaje, las NNS son entrenadas para especializarse en un subespacio del espacio de entrada y luego sus salidas son combinadas. El resultado es una mezcla de expertos compuesta por agentes del mismo tipo de NN o de distinto tipo:

- En caso de diseñar mezclas de expertos del mismo tipo de NN, se propone como estrategia de combinación de las salidas la selección por máximo, de forma análoga a como el CGLR. En este caso, en lugar de diseñar un LR para cada valor del parámetro desconocido (CGLR), se diseñan redes neuronales capaces de aproximar el ALR para subintervalos de dicho parámetro.
- En el caso del comité de distintos tipos de redes, se ha implementado la misma estrategia de combinación por máximo, pero se ha probado que puede mejorarse las capacidades de detección estableciendo, para una P_{FA} , diferentes umbrales para cada experto y aplicando una OR lógica a las salida binarias de cada detector basado en una NN para obtener la salida de la mezcla de expertos.

- Se ha realizado un estudio teórico sobre las capacidades de las SVMs para aproximar el detector de NP sin más que comparar su salida con un umbral establecido para cumplir los requisitos de P_{FA} . Las C-SVMs se han asumido entrenadas supervisadamente para minimizar $E = \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 - t_{H_0} F(\mathbf{z}))) + \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - t_{H_1} F(\mathbf{z})))$ [Shawe2004].

Debido a la función escalón, no es posible calcular la función aproximada por la SVM a través del cálculo de variaciones. Para realizar el estudio, se ha aproximado la función escalón por una función logística y se ha aplicado el límite cuando $k \rightarrow \infty$ a la solución obtenida, siendo k el parámetro que regula la pendiente de la función sigmoide.

- Se ha demostrado que cuando el conjunto de entrenamiento tiende a infinito y $C \rightarrow \infty$, la función aproximada cuando se minimiza la aproximación de E basada en la función sigmoide de parámetro k permite implementar el detector NP, cuando se compara con un umbral fijado atendiendo a los requisitos de $P_{FA} \left(F_0(\mathbf{z}) \underset{H_1}{\overset{H_0}{\gtrless}} \eta_0(P_{FA}) \right)$.
- Si se aplica el límite cuando $k \rightarrow \infty$, las funciones logísticas se convierten en las funciones escalón de E . En este caso se demuestra que la C-SVM es capaz de aproximar sólo tres puntos de la curva ROC del detector NP, correspondientes a $P_{FA} = 0$, $P_{FA} = 1$ y la P_{FA} de mínima probabilidad de error. En este último caso, el umbral η_0 siempre es nulo.
- En situaciones prácticas, el tamaño del conjunto de entrenamiento y el parámetro de regularización de C son finitos. Se ha demostrado experimentalmente que para la detección

de blancos Swerling I y II en AWGN pueden lograrse buenas aproximaciones del detector NP para tamaños del conjunto y valores de C suficientemente grandes a costa de elevados costes computacionales. En los casos de estudio, se requieren de una precisión muy elevada para ajustar el umbral requerido para cada P_{FA} .

- Se ha propuesto una modificación de la función de error a minimizar en el proceso de entrenamiento de una SVM, con el objetivo de reducir el error en la aproximación del detector NP en un punto de la curva ROC del detector NP distinto del asociado al de mínima probabilidad de error: $E = C_{10} \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_0} (u(1 - t_{H_0} F(\mathbf{z}))) + C_{01} \frac{1}{N} \sum_{\forall \mathbf{z} \in Z_1} (u(1 - t_{H_1} F(\mathbf{z})))$.

Se han aproximado las funciones escalón por funciones logísticas para después aplicar el límite cuando k , el parámetro que controla la pendiente de la función logística en su zona lineal, tiende a infinito. Se ha demostrado que para $\eta_0 = 0$, la SVM entrenada con $N \rightarrow \infty$ y $C \rightarrow \infty$ es capaz de aproximar el punto de la curva ROC del detector NP correspondiente a $\eta_r = \frac{C_{10} P(H_0)}{C_{01} P(H_1)}$. Por tanto, eligiendo C_{10} y C_{01} apropiadamente se puede aproximar cualquier punto de la curva ROC.

- Para los casos en los que N y C son finitos, se han diseñado detectores basados en 2C-SVMs para la detección de blancos Swerling I y II en AWGN. Como en estos casos el detector NP está completamente caracterizado, se han obtenido los valores de γ asociados a las P_{FA} deseadas y se ha realizado un estudio de N . Los resultados demuestran que el error de aproximación se reduce al aumentar N y C . Además, una vez se ha fijado la SNR , C , el número de vectores soporte es aproximadamente igual para los diferentes valores de γ .
- Si no se conoce a priori el η_r asociado a una P_{FA} , la forma de maximizar la P_D para una P_{FA} , con una solución basada en una 2C-SVM, es realizando una búsqueda en rejilla de entre los parámetros γ y C que permitan aproximar el punto del detector NP correspondiente con un coste computacional (N_{vs}) controlado.

Capítulo 7

Resultados

7.1. Introducción

En los capítulos precedentes se ha desarrollado un estudio de soluciones subóptimas basadas en el CGLR para la detección de blancos con parámetros desconocidos, y se han propuesto soluciones basadas en inteligencia artificial capaces de aproximar el ALR. En este capítulo, se presenta un estudio comparativo de capacidades de detección y complejidad computacional entre las soluciones basadas en el CGLR y las basadas en redes neuronales o máquinas de vectores soporte.

En el capítulo 6 se describieron los algoritmos de entrenamiento que se iban a utilizar en los esquemas de detección basados en inteligencia artificial, para cumplir la condición suficiente de aproximar la curva ROC del detector NP o, al menos, un punto concreto de la curva. Sin embargo, esta aproximación depende, a parte del conjunto de entrenamiento, de la arquitectura de cada tipo de agente inteligente considerado. En este capítulo, también se estudiará la idoneidad de cada arquitectura para aproximar la fronteras de decisión de los casos de estudio que se han considerado para desarrollar esta tesis.

Para presentar los resultados se ha diseñado un escenario lo más real posible donde se puede evaluar las prestaciones de los detectores propuestos. Este escenario asume un hipotético radar situado en el estrecho de Gibraltar, con unas características similares a las del radar IPIX [Drosopoulos1994], sistema desarrollado en la Universidad McMaster de Canadá para la adquisición de datos reales con el fin de caracterizar el clutter marino y desarrollar estrategias de detección de embarcaciones pequeñas. En este entorno marino, se simularán los diferentes tipos de interferencia, en función del ángulo de incidencia y del estado del mar, junto con la trayectoria de los barcos con una frecuencia Doppler variable en función de la velocidad del barco y de su posición.

En una primera fase se realiza un estudio de las diversas soluciones basadas en inteligencia artificial consideradas en la tesis, utilizando bases de datos sintéticas generadas asumiendo diversos modelos de clutter. En una segunda fase, los detectores diseñados se evaluarán sobre el escenario radar simulado el cual incluye los efectos de la propagación de la onda electromagnética, la variación de la celda del clutter iluminada por el radar con la distancia y trayectorias reales de blancos importadas de [Localizado].

Para cada escenario, se detallará el razonamiento seguido en la investigación para presentar

los resultados y obtener la solución capaz de aproximar el detector óptimo con el menor coste computacional. En los casos de los detectores basados en una única red cuyos errores de aproximación son muy elevados, se plantearán esquemas basados en mezcla de expertos. En cada estrategia se detallará el razonamiento seguido en la investigación para obtener una solución de compromiso entre capacidades detectoras y coste computacional asociado. El punto de la curva ROC elegido para las SVMs corresponde a una $P_{FA} = 10^{-4}$. Para estimar la P_D correspondiente se hará una búsqueda en rejilla de los parámetros C y γ de las $2C$ -SVMs que maximicen la detección para un umbral fijado por los requisitos de falsa alarma.

En el apartado 7.2, se describirán el escenario radar simulado, el diseño de los conjuntos de entrenamiento, validación y test y los algoritmos de estimación de P_D y P_{FA} . Y en los apartados 7.3, 7.4, 7.5, 7.6, y 7.7 se presentan los resultados obtenidos por soluciones basadas en inteligencia artificial sobre los conjuntos de datos sintéticos para cada caso de estudio considerado en esta tesis. Primero se estudia el mejor esquema basado en una NN y, posteriormente, se presentan los resultados obtenidos para un detector $2C$ -SVM. Estos resultados se compararán con las aproximaciones basadas en el CGLR presentadas en el apartado 5.3.1. En el apartado 7.8, se presentan los resultados de aplicar los mejores detectores obtenidos a una secuencia temporal de exploraciones en el entorno presentado en el apartado 7.2.

Algunas de las aportaciones relacionadas con la propuesta de detectores robustos basados en inteligencia artificial han sido publicadas en [Mata2006], [Mata2007], [Mata2008],[Mata2008b], [Mata2009] y [Mata2011a].

7.2. Bases de datos sintéticos de entrenamiento y simulación

En este apartado se describen los conjuntos de datos sintéticos a partir de los cuales se han generado los conjuntos de entrenamiento, validación y test, y los algoritmos utilizados para estimar las curvas ROC de los detectores propuestos basados en agentes inteligentes.

Además, se presenta el escenario radar simulado el cual permitirá comprobar las capacidades de detección de los detectores propuestos para blancos con frecuencia Doppler desconocida debido a la variación de la posición relativa del blanco respecto al radar, según se desplaza siguiendo una trayectoria determinada. La descripción incluye la localización del radar y su geometría, así como las trayectorias de dos barcos reales que se encuentran dentro del margen de cobertura del radar en un intervalo de tiempo determinado. Estas trayectorias se han importado de la herramienta *Localizatodo* [Localizatodo].

7.2.1. Conjuntos de datos sintéticos

Para entrenar los MLPs, las SONNs, las RBFNNs y las SVMs se han construido conjuntos de entrenamiento, validación y test formados por patrones pertenecientes a las hipótesis H_0 y H_1 uniformemente distribuidos con la misma probabilidad a priori ($P(H_0) = P(H_1) = 0,5$). En los vectores de entrada bajo hipótesis H_0 , todas las muestras son de interferencia, clutter y/o ruido, y bajo H_1 , todas las muestras son de blanco más clutter y/o ruido.

Para elegir el tamaño de los conjuntos de entrenamiento se han tenido en cuenta los siguientes criterios:

- En el caso de los MLPs y las SONNs, al aumentar el tamaño del conjunto de entrenamiento se minimiza la posibilidad de que la red se sobre-especialice, es decir, se *aprenda* el conjunto de entrenamiento, y se aumenta la precisión en la estimación de la entropía o el error cuadrático medio utilizados como función de error en cada caso respectivamente. La característica negativa es el aumento en el tiempo requerido para realizar el entrenamiento de la red.

Para presentar un estudio suficientemente detallado que aporte conclusiones útiles, y teniendo en cuenta la capacidad de cálculo de los ordenadores disponibles, se han construido conjuntos de 50,000 patrones de entrenamiento.

En todos los casos se ha empleado una estrategia de validación cruzada, incluso en el de las redes más pequeñas para las que este número de patrones de entrenamiento garantiza prácticamente la no sobre-especialización, por lo que ha sido necesario construir una doble batería de conjuntos de datos: una de entrenamiento y otra de validación.

- Para el entrenamiento de las RBFNNs se han construido conjuntos de entrenamiento análogos a los descritos para el entrenamiento de los MLPs. Debido a que la red a entrenar mediante técnicas supervisadas es muy pequeña (compuesta por la única neurona de salida), no se ha utilizado ninguna estrategia de validación.

La característica diferenciadora respecto a los conjuntos generados para el entrenamiento de los MLPs es el número de patrones elegido. Debido al funcionamiento de estas redes (explicado a partir de la teoría de la regularización en la sección 4.7.3), se ha tomado la decisión de utilizar 10,000 patrones, elegidos aleatoriamente de entre los 50,000 que componen el conjunto generado para los MLPs.

- En las SVMs, al igual que en las RBFNNs, el conjunto de entrenamiento no necesita ser tan extenso como en el caso de los MLPs. Además, el entrenamiento de las SVM conlleva interacción entre todos los patrones que forman parte del conjunto de entrenamiento, y, por lo tanto, el coste computacional crece considerablemente con su tamaño. Se ha tomado la decisión de utilizar 500 patrones, elegidos aleatoriamente de entre los 5,000 que componen el conjunto considerado para las redes con funciones de base radial.

7.2.2. Salidas deseadas

En el apartado 6.2 se estudiaron las estrategias de entrenamiento y las arquitecturas propuestas para cada uno de los tipos de NNs considerados.

- En los MLPs se considera neuronas con funciones de activación de tipo logístico y las salidas deseadas correspondientes son $\{0, 1\}$.
- En las SONNs se utilizan neuronas con funciones de activación de tipo tangente hiperbólica y las salidas deseadas asociadas son $\{-1, 1\}$.
- En las RBFNNs se ha optado por utilizar una neurona de salida con función de activación lineal. En estas redes, si se tienen en cuenta las características estadísticas de los problemas a resolver y la estrategia de entrenamiento elegida, la primera capa adquiere mucha más

importancia que la segunda. En este caso, sólo se ha considerado el conjunto de salidas deseadas $\{0, 1\}$.

Por otro lado, en el apartado 6.3 se estudia la arquitectura de las C-SVMs y 2C-SVMs, y se definen las etiquetas $\{-1, 1\}$ correspondientes a las clases de salida que utiliza la función de error de clasificación definida en [Shawe2004].

7.2.3. Estimación de las curvas ROC

Para estimar las curvas ROC de las redes entrenadas es necesario estimar las probabilidades de detección y falsa alarma para un amplio conjunto de umbrales de detección. En el diseño de detectores radar sólo tiene interés práctico el margen de valores del umbral de detección más elevados, que es el que se corresponde con las probabilidades de falsa alarma muy bajas. El número de patrones necesarios para estimar una determinada probabilidad con un error relativo asociado, ϵ , utilizando la simulación de montecarlo viene definido por la ecuación (7.1).

$$N = \frac{1 - P}{\epsilon^2 \cdot P} \quad (7.1)$$

Es decir, se han utilizado tantos patrones bajo la hipótesis nula para garantizar P_{FA} en torno a 10^{-6} con un error relativo menor o igual al 10% ($\epsilon \leq 10\%$). En los casos de estudio basados en interferencia gaussiana, se ha utilizado la técnica de muestreo enfatizado para reducir el número de patrones generados bajo hipótesis H_0 sin reducir el error relativo [Sanz-Gonzalez1999][Orsak1993][Grajaj1999]. Para estimar las probabilidades de detección se ha utilizado siempre la simulación de montecarlo clásica con $M = 80,000$ patrones que garantiza la estimación de una $P_D = 0,01$ con un $\epsilon < 10\%$.

7.3. Detección de blancos gaussianos con coeficiente de correlación desconocido en AWGN

En este apartado se proponen distintos esquemas de detección basados en agentes inteligentes para aproximar el ALR en el caso de blancos gaussianos con ρ_s desconocido en presencia de AWGN. La evolución de las soluciones propuestas se basa en el conocimiento de los detectores óptimos para blancos gaussianos con parámetros conocidos. En el apartado 6.3.4 se demostró como el detector óptimo para Swerling II en AWGN formaba hiperesferas de $2P$ dimensiones como fronteras de decisión, mientras que para Swerling I formaba hipercilindros. En [Jarabo2003] se demostró que para aproximar hiperesferas son más adecuadas las RBFNNs, mientras que para aproximar hipercilindros son más eficientes los MLPs. Con este punto de partida, se pretende diseñar un detector capaz de aproximar el detector NP para blancos con ρ_s o Ω desconocidos, con el menor coste computacional posible.

Según el apartado 5.3.1 relativo a los detectores CGLR para este mismo caso de estudio, los resultados que se van a estudiar se corresponden a un blanco con una $SNR = 7$ dB.

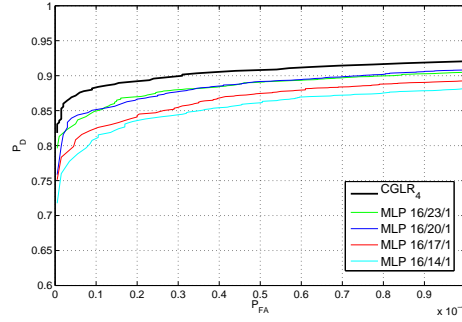


Figura 7.1: Curvas ROC de detectores basados en un único MLP 16/ M /1 para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.

7.3.1. Soluciones basadas en redes neuronales

Como primera aproximación, se propone un detector basado en un **único MLP** entrenado con un conjunto de entrenamiento con patrones bajo hipótesis H_1 con una variación uniforme en $\rho_s \in [0; 1]$. En la Figura 7.1 se presentan las curvas ROC estimadas para distintas MLPs junto a la aproximación basada en el $CGLR_4$. Los peores resultados se obtienen para la red MLP 16/14/1 y se van produciendo mejoras significativas hasta la red MLP 16/20/1. Sin embargo esta ROC sigue estando por debajo de la correspondiente al $CGLR_4$, por lo que se demuestra que este tipo de arquitectura no consigue una buena aproximación del ALR a pesar de cumplir con la condición suficiente para poder aproximar el detector NP.

Como la función de error es adecuada, para mejorar la aproximación se propone cambiar la estrategia de entrenamiento. La idea se basa en simplificar el problema de entrada generando patrones de entrenamiento bajo H_1 con un margen de variación más pequeño. En este caso, para cubrir todo el intervalo de variación será necesario más de una red, cada una especializada en un margen distinto, tal y como se explicó en la sección 6.2.4. El esquema de detección resultante se basa en una **mezcla de expertos MLPs**, $MLP_1 + MLP_2$ entrenados con patrones con $\rho_s \in [0; 0, 5]$ y $\rho_s \in (0, 5; 1]$, respectivamente. La combinación de las salidas de los MLPs se basa en la selección por máximo según el esquema de la Figura 6.8 del apartado 6.2.4.

Si se analiza cada uno de los intervalos de variación por separado:

- MLP_2 está especializado en valores de ρ_s próximos a uno. En el límite, $\rho_s = 1$, blanco Swerling I, las fronteras de decisión son hipercilindros en el espacio de entrada que requieren de al menos 3 hiperplanos. Por lo tanto, se debería poder reducir el tamaño de la capa oculta de MLP_2 sin empeorar la aproximación del ALR y disminuyendo el coste computacional. En la Figura 7.2(a), se presentan los resultados al variar el número de neuronas ocultas del MLP_2 de 20 a 5 neuronas. Se observa que con una MLP_2 16/8/1 se mantienen los resultados obtenidos con un comité MLP_1 16/20/1 + MLP_2 16/20/1. Con menos unidades ocultas, las capacidades de detección se ven mermadas y con más unidades la curva ROC no mejora significativamente.
- Una vez reducida la complejidad de MLP_2 , se pasa a estudiar la estructura del MLP_1 , especializado en valores de ρ_s próximos a cero. Las regiones de decisión para la detección de

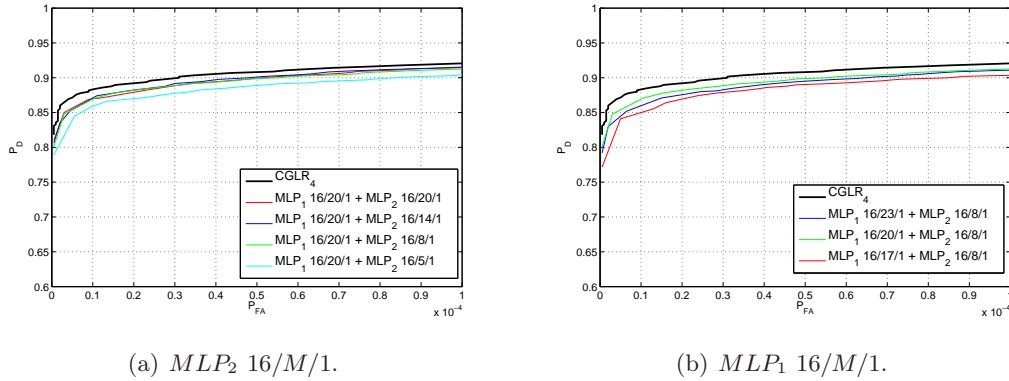
(a) MLP_2 16/M/1.(b) MLP_1 16/M/1.

Figura 7.2: Estudio del tamaño de la capa oculta de los MLPs que forman la mezcla de expertos para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.

blanco Swerling II, $\rho_s = 0$, son hiperesferas que requieren de al menos $2P + 1$ hiperplanos. En este caso, la reducción del número de neuronas no parece viable, pero un aumento podría disminuir el error de aproximación. En la Figura 7.2(b) se comprueba que esta mejora no tiene lugar al variar el tamaño de la capa oculta del MLP_1 y el resultado más próximo a la curva ROC del $CGLR_4$ es el obtenido por un detector basado en una combinación de MLP_1 16/20/1 + MLP_2 16/8/1.

Para mejorar los resultados de la Figura 7.2 se propone aprovechar la estrategia de dividir el intervalo de variación de ρ_s para elegir un tipo de red que sea más adecuado para cada intervalo restringido. Es decir, se plantea una **mezcla de expertos entre una RBF y un MLP**, $RBFNN_1 + MLP_2$ especializados en los intervalos $\rho_s \in [0; 0, 5]$ y $\rho_s \in (0, 5; 1]$, respectivamente. Se ha sustituido el MLP_1 por la $RBFNN_1$, porque puede mejorar la aproximación a las hiperesferas de decisión a las que se tiende cuando $\rho_s \rightarrow 0$. El MLP se entrena para minimizar la entropía cruzada y la RBFNN el error cuadrático medio, por lo tanto tienden a aproximar funciones discriminantes equivalentes pero no idénticas. En este caso, la estrategia para combinar las salidas no se puede simplificar a una selección por máximo, y se utiliza un esquema basado en una OR lógica descrito en la Figura 6.10.

La adecuación de las funciones de base radial para la detección de blancos con valores pequeños de ρ_s se confirma en la Figura 7.3(a), donde se observa una mejora significativa de la capacidad de detección cuando se utilizan esquemas basados en comités $RBFNN_1 + MLP_2$. La solución que presenta el mejor compromiso entre capacidad de aproximar el detector de NP y el coste computacional, es la basada en la combinación $RBFNN_1$ 16/3/1 + MLP_2 16/8/1. A partir de ahora, la investigación se centrará en buscar nuevas soluciones que sean capaces de reducir el coste computacional sin mermar las características de detección. Atendiendo a los resultados en la Figura 7.3(a), el siguiente paso es proponer detectores basados en **una única RBFNN**. El problema que presentan las RBFNN para aproximar el cociente de verosimilitud promediado es la estrategia de entrenamiento elegida. Las redes se entrenan para minimizar el error cuadrático medio, que cumple la condición suficiente para aproximar el detector NP. Sin embargo, los centros y varianzas de las funciones de base radial suelen ajustarse en base a técnicas de clustering de forma no supervisada, de modo que la parte supervisada del entrenamiento sólo tiene

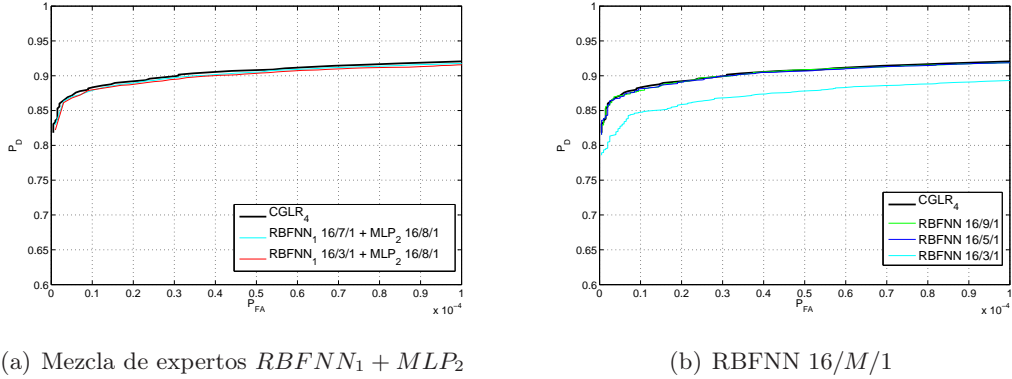


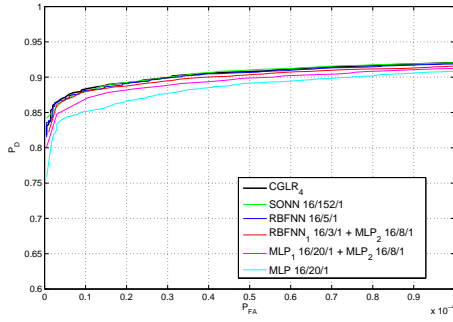
Figura 7.3: Curvas ROC de detectores basados en RBFNNs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.

la posibilidad de ajustar los parámetros asociados a la neurona de salida a partir de los patrones transformados por las funciones base y las salidas deseadas. Para evitar este problema, se han entrenado las RBFNNs de manera completamente supervisada utilizando algoritmos genéticos según se ha descrito en el apartado 6.2.3. En la Figura 7.3(b) se demuestra que una única RBFNN también es capaz de aproximar el ALR. El número de neuronas ocultas mínimo necesario es 5, de manera que una RBFNN 16/9/1 no presenta mejoras significativas. Las soluciones basadas en RBFNNs suelen obtener buenos resultados con un tamaño de la capa oculta tal que permite reducir considerablemente el número de operaciones requerido.

Para completar el estudio de soluciones basadas en NNs, falta por contemplar detectores basados en **una única SONN**. Las SONNs compuestas por una QNU que transforma el espacio del vector de entrada en uno de dimensión mayor dependiente de las combinaciones lineales y de segundo orden. Este espacio de entrada transformado permitirá aproximar fronteras de decisión con hiperplanos lineales y de segundo orden, es decir, combina las propiedades de aproximación de los MLPs y las RBFNNs. En la Figura 7.4(a) se presenta la curva ROC obtenida por una SONN entrenada con patrones bajo H_1 con una variación uniforme en $[0; 1]$ y se demuestra que este tipo de detectores también son capaces de aproximar el detector óptimo. En esta misma Figura, se presenta un resumen con los mejores resultados obtenidos en cada paso de la investigación, junto con el coste computacional asociado a cada solución. Los resultados reflejan que los detectores basados en la mezcla de expertos $RBFNN_1$ 16/3/1 + MLP_2 16/8/1, en una única RBFNN 16/5/1 y en una única SONN 16/152/1 son capaces de aproximar el detector óptimo. Estos tres detectores presentan un número de operaciones mucho menor que los necesarios para implementar el $CGLR_4$. Entre las soluciones propuestas, la RBFNN 16/5/1 es la que mejor responde al compromiso entre capacidades detectoras y coste computacional.

7.3.2. Soluciones basadas en SVMs

El estudio se completa con el análisis de soluciones basadas en 2C-SVMs. Siguiendo la metodología propuesta en el apartado 6.3.3, se entrenarán 2C-SVMs con distintas parejas de valores C y γ . A partir de los detectores 2C-SVM obtenidos, se estimarán la P_{FA} y la P_D cuando su salida se compara con un umbral igual a cero. El objetivo es encontrar, entre los valores considerados,

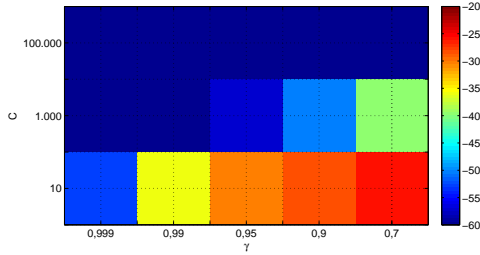
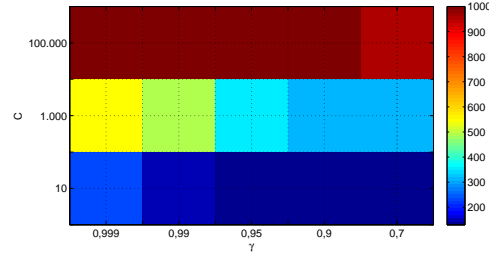


(a) Curvas ROC

Detector	Número de operaciones
CGLR ₄	2.235
SONN 16/152/1	441
RBFNN 16/5/1	256
RBFNN ₁ 16/3/1+ MLP ₂ 16/8/1	437
MLP ₁ 16/20/1+ MLP ₂ 16/8/1	984
MLP 16/20/1	702

(b) Número de operaciones necesarias

Figura 7.4: Resumen de soluciones basadas en NNs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.

(a) $10 \log(P_{FA})$ 

(b) Número de vectores soporte

Figura 7.5: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.

la pareja de parámetros C y γ que nos permite obtener una $P_{FA} \leq 10^{-4}$ y proporcionan la mayor P_D .

En la Figura 7.5 se presentan las P_{FAS} y el número de vectores soporte obtenidos de cada entrenamiento. El conjunto de entrenamiento se ha diseñado con un total de 1000 patrones con $P(H_0) = P(H_1)$. Partiendo de las conclusiones extraídas en el apartado 6.3, se sabe que el error de aproximación se reduce según C aumenta. En la Figura 7.5(a) se observa que para $C = 10^5$, las 2C-SVMs entrenadas no son capaces de aprender dado que para $\eta_0 = 0$ tanto P_{FA} como P_D son nulas, por lo que se necesitarían más patrones de entrenamiento para mejorar la aproximación al punto correspondiente de la curva ROC del detector óptimo deseado. Los entrenamientos con $C = 10^3$ y $C = 10$ muestran una P_{FA} descendente según γ aumenta. Por otro lado, en la Figura 7.5(b) se observa que según C aumenta el número de vectores soporte aumenta y que para un mismo C este número decrece según γ aumenta.

En la Tabla 7.1 se presentan los datos numéricos de las probabilidades estimadas para los detectores 2C-SVM entrenados. El mejor resultado corresponde a la combinación de un $C = 10^3$ y un $\gamma = 0,7$. La P_D obtenida es muy parecida a la resultante de la mejor solución basada en redes neuronales (RBFNN 16/5/1), igual a 0,9197, sin embargo el coste computacional asociado a

C	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	$2,00 \cdot 10^{-6}$	0,8652	$9,00 \cdot 10^{-6}$	0,9070	$9,96 \cdot 10^{-5}$	0,9203
10	$2,39 \cdot 10^{-4}$	0,9108	$9,72 \cdot 10^{-4}$	0,9249	$1,50 \cdot 10^{-3}$	0,9544	$2,50 \cdot 10^{-3}$	0,9594

Tabla 7.1: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 1 con $\rho_s^s \in [0, 1]$. $TSNR = SSNR = 7$ dB.

la 2C-SVM es mucho más grande que el correspondiente a la RBFNN. El número de operaciones requerido por una 2C-SVM entrenada con $C = 10^3$ y $\gamma = 0,7$ es 21.172 (correspondiente a los 316 vectores soporte necesarios), más de 80 veces las operaciones requeridas por la RBFNN 16/5/1.

7.4. Detección de blancos gaussianos con pulsación Doppler desconocida en AWGN

Según el apartado 5.3.1 relativo a los detectores CGLR para este mismo caso de estudio, los resultados que se van a estudiar se corresponden a un blanco Swerling I con una $SNR = 10$ dB.

7.4.1. Soluciones basadas en redes neuronales

Las fronteras de decisión formadas por la forma cuadrática del detector NP para blancos Swerling I con frecuencia Doppler desconocida en ruido AWGN son hipercilindros en el espacio de entrada de 2P dimensiones, puesto que sólo dos de sus autovalores son no nulos. Ante el desconocimiento de cómo los hipercilindros se deformarán debido a la variación uniforme Ω en $[0; 2\pi)$, la primera solución que se propone se basa en **un único MLP** capaz de encerrar un hipercilindro con tres hiperplanos. El MLP utilizará un conjunto de entrenamiento con patrones bajo hipótesis H_1 con una frecuencia Doppler normalizada variable en el intervalo $[0; 2\pi)$. Se utilizará la entropía cruzada como función de error de manera que el entrenamiento supervisado cumpla la condición suficiente de poder aproximar una función que al ser comparada con un umbral definido para una P_{FA} permita implementar el detector NP.

En la Figura 7.6(a) se presentan las curvas ROC estimadas para MLPs para distintos valores de neuronas de la capa intermedia, M . Se propone el esquema basado en un MLP 16/17/1 como solución de compromiso entre las capacidades detectoras y la carga computacional. Esta red requiere 597 operaciones. Aunque el número de operaciones es mucho menor que el requerido por el detector $CGLR_{16}$ (8.043 operaciones), las P_D para una misma P_{FA} presentan una pérdida en torno al 2%, más o menos constante para todos los valores de P_{FA} considerados. Es necesario proponer otras alternativas que sean capaces de aproximar la deformación de los hipercilindros de separación debida a la variación uniforme de Ω .

Siguiendo un proceso análogo al desarrollado en la sección 7.3, también se ha analizado la posibilidad de diseñar detectores basados en **una única RBFNN**. La estrategia de entrenamiento aplicada, con patrones generados bajo H_1 asumiendo una variación uniforme $\Omega^t \in [0; 2\pi)$, es completamente supervisada para ajustar todos los parámetros libres de las RBFNNs. Los mejores

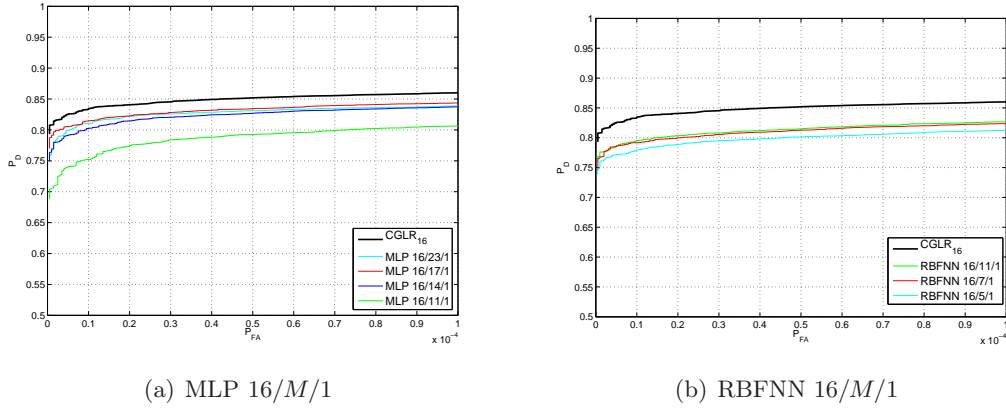


Figura 7.6: Curvas ROC de detectores basados en un MLP o en una RBFNN para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.

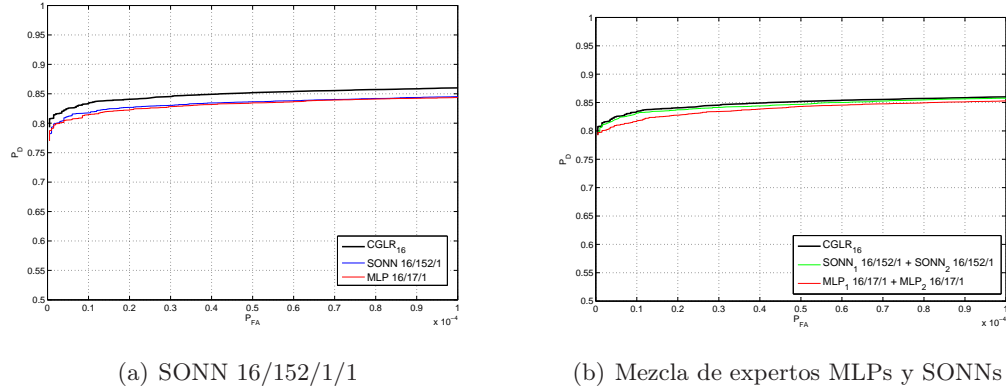


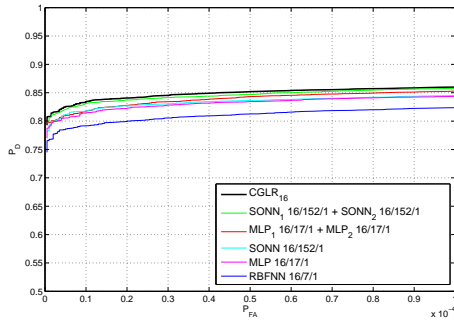
Figura 7.7: Curvas ROC de detectores basados en una SONN o en un comité de NNs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.

resultados se han obtenido para una RBFNN 16/7/1, Figura 7.6(b), puesto que presenta mejores P_D para una misma P_{FA} que otra RBFNN con un $M < 7$ y las mejoras no son significativas para RBFNNs con $M > 7$. A pesar de haber utilizado un entrenamiento completamente supervisado que cumple la condición suficiente de aproximar el detector óptimo, las soluciones basadas en RBFNNs presentan más error en la aproximación que los detectores MLP.

Para mejorar los resultados obtenidos con MLPs se propone tener en cuenta la información contenida en las relaciones de segundo orden de los elementos de los vectores de entrada a través de esquemas basados en **una única SONN**. Los resultados de la Figura 7.7(a) permiten concluir, que la capacidad de detección obtenida con una SONN no obtiene una mejora significativa con respecto al MLP 16/17/1.

Con el objetivo de mejorar las aproximaciones obtenidas se propone dividir el problema en varios más sencillos, planteando un **mezcla de expertos NNs** especializadas en distintos márgenes de variación de Ω más pequeños y cuya combinación de salidas mejore las capacidades de detección global.

Teniendo en cuenta las propiedades de periodicidad del espectro, se propone entrenar una



(a) Curvas ROC

Detector	Número de operaciones
$CGLR_{16}$	8.943
$SONN_1$ 16/152/1 + $SONN_2$ 16/152/1	882
MLP_1 16/17/1 + MLP_2 16/17/1	1.994
$SONN$ 16/152/1	441
MLP 16/17/1	597
$RBFNN$ 16/7/1	358

(b) Número de operaciones necesarias

 Figura 7.8: Resumen de soluciones basadas en NNs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.

red asumiendo un intervalo $\Omega^t \in [0; \pi/4) \cup [3\pi/4; 2\pi)$ y otra red con $\Omega^t \in [\pi/4; 3\pi/4)$. Las redes entrenadas en el intervalo de bajas frecuencias se denotará con el subíndice 1, y las de altas frecuencias con el subíndice 2. Las soluciones basadas en comités consideran la combinación de dos MLPs o dos SONNs, es decir, la mezcla entre dos expertos iguales cuyas salidas se pueden combinar por selección máxima para obtener una salida del comité que se comparará con un umbral fijado por los requisitos de P_{FA} para decidir por la presencia o no de blanco (Figura 6.8).

El número de operaciones de los detectores basados en la combinación de dos redes es mayor que los basados en una única red, sin embargo en la Figura 7.7(b) se observa cómo una mezcla de expertos SONNs consigue una buena aproximación al $CGLR_{16}$. La mezcla de expertos MLPs también mejora la aproximación basada en un único MLP, pero no lo suficiente para para aproximar las P_D del $CGLR_{16}$.

En la Figura 7.8, se presenta un resumen de los resultados obtenidos desde el punto de vista de detección y de carga computacional asociada. Al comparar las curvas ROC en la Figura 7.8(a) se observa que la solución basada en un comité $SONN_1$ 16/152/1 + $SONN_2$ 16/152/1 es la única que presenta capacidades detectoras comparables al $CGLR_{16}$. La gran ventaja de la mezcla de expertos SONNs es que requieren 882 operaciones frente a las 8.943 necesarias para el $CGLR_{16}$.

7.4.2. Soluciones basadas en SVMs

En la Figura 7.5(a) se representa la P_{FA} estimada para cada combinación de $[C; \gamma]$ con la que se ha entrenado la 2C-SVM. El mejor resultado generalmente se obtendrá para la 2C-SVM que es capaz de converger a una solución con el conjunto de entrenamiento diseñado para el valor de C más grande, ya que, para valores superiores de C , la P_{FA} y P_D obtenidas son iguales a cero. Con 1000 patrones de entrenamiento, el comportamiento de la 2C-SVM no es capaz de converger para $C = 10^5$ en este escenario (Figura 7.5(a) y Tabla 7.2). Las P_{FA} obtenidas con las combinaciones $[C = 10^3; \gamma = 0, 7]$ y $[C = 10; \gamma = 0, 99]$ son las más próximas a 10^{-4} . Una de las diferencias entre las dos aproximaciones se basa en el número de vectores soporte resultante, Figura 7.9(b). Los resultados asociados a una C menor se corresponden a un número menor de vectores soporte y un mayor error en la aproximación al óptimo. Los valores estimados de P_D se

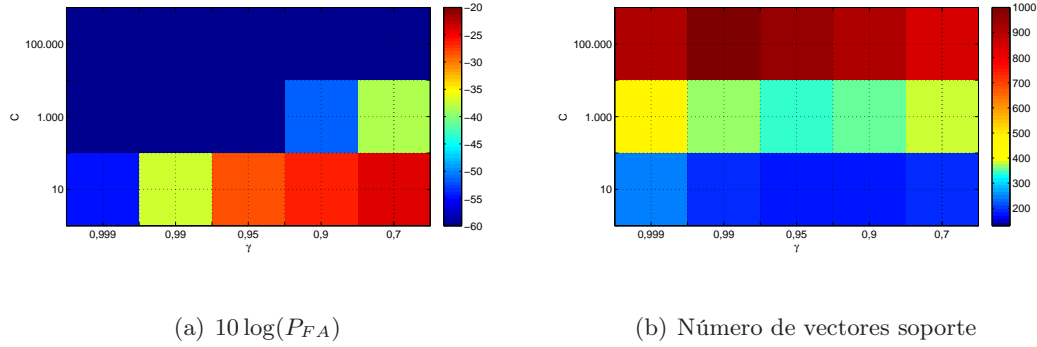


Figura 7.9: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.

C	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	-	-	$6,56 \cdot 10^{-6}$	0,8194	$1,01 \cdot 10^{-4}$	0,8601
10	$1,95 \cdot 10^{-4}$	0,8018	$1,50 \cdot 10^{-3}$	0,8763	$2,30 \cdot 10^{-3}$	0,8786	$3,70 \cdot 10^{-3}$	0,8820

Tabla 7.2: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 1 con $\Omega^s \in [0, 2\pi)$. $TSNR = SSNR = 10$ dB.

presentan en la Tabla 7.2. Esta Tabla confirma que la mejor aproximación se consigue para la combinación $[C = 10^3; \gamma = 0,7]$, que necesita 388 vectores soporte para obtener una $P_D = 0,8601$ manteniendo prácticamente los requisitos de falsa alarma $P_{FA} = 1,01 \cdot 10^{-4}$.

La mejor solución basada en redes neuronales utiliza una mezcla de expertos $SONN_1$ 16/152/1 + $SONN_2$ 16/152/1. La P_D de este detector para una $P_{FA} = 10^{-4}$ es 0,8585, resultados similares a los obtenidos por la 2C-SVM. La gran diferencia de los esquemas de detección propuestos se basa en el número de operaciones requerido: el detector $SONN_1$ 16/152/1 + $SONN_2$ 16/152/1 necesita 882 operaciones frente a las 25.996 que necesita el detector 2C-SVM. Además, el detector 2C-SVM está adaptado a $P_{FA} = 10^{-4}$ y no se ha visto que sucede para otros valores de P_{FA} . El detector $SONN_1$ 16/152/1 + $SONN_2$ 16/152/1 proporciona buenos resultados para todas las P_{FA} consideradas.

7.5. Detección de blancos gaussianos con coeficiente de correlación desconocido en interferencia gaussiana

Para ángulos de incidencia rasante superiores a los 10° , el modelo de clutter gaussiano se ajusta al clutter marítimo [Skolnik2008]. Aunque se pueden hacer estimaciones del entorno para distinguir entre distintos tipos de interferencia gaussiana, es conveniente diseñar detectores robustos basados en inteligencia artificial que sean capaces de detectar blancos gaussianos con parámetros desconocidos con un coste computacional mucho menor que el asociado a aproximaciones al óptimo basadas en el CGLR.

Los tipos de clutter considerados en el desarrollo de la tesis y que se van a utilizar para estudiar las aproximaciones propuestas son: Clutter no dominante poco correlado ($CNR = 20$ dB y $\rho_c = 0,7$), Clutter no dominante muy correlado ($CNR = 20$ dB y $\rho_c = 0,995$) y Clutter dominante muy correlado ($CNR = 40$ dB y $\rho_c = 0,995$).

Las soluciones basadas en los agentes inteligentes dependerán de las características estadísticas y espectrales de la interferencia gaussiana, compuesta por clutter y AWGN, que se asumirán constantes en el tiempo.

Según el apartado 5.3.2 relativo a los detectores CGLR para este mismo caso de estudio, los datos asumidos para obtener los resultados dependen del tipo de clutter:

- Clutter no dominante poco correlado: $CNR = 20$ dB, $\rho_c = 0,7$ y $SIR = 0$ dB.
- Clutter no dominante muy correlado: $CNR = 20$ dB, $\rho_c = 0,995$ y $SIR = -10$ dB.
- Clutter dominante muy correlado: $CNR = 40$ dB, $\rho_c = 0,995$ y $SIR = -28$ dB.

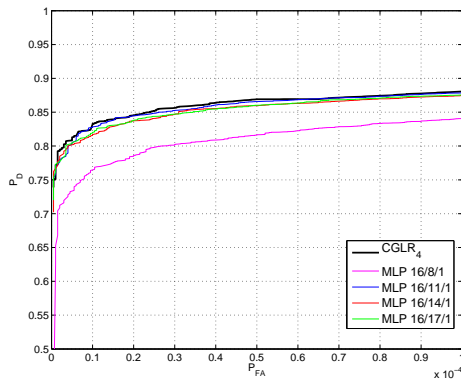
7.5.1. Soluciones basadas en redes neuronales

Según los resultados presentados en 5.2.2, los detectores diseñados con ρ_s^d próximos a cero son bastante robustos, mientras que los diseñados para $\rho_s^d \simeq 1$ presentan mucha sensibilidad a la ρ_s^d asumida en el diseño. Tendencia que también se ha demostrado en la detección de blancos gaussianos con ρ_s desconocido en AWGN, apartado 7.3, donde la mejor solución se basaba en un detector RBFNN con una arquitectura especialmente idónea para aproximar hiperesferas en el espacio de entrada.

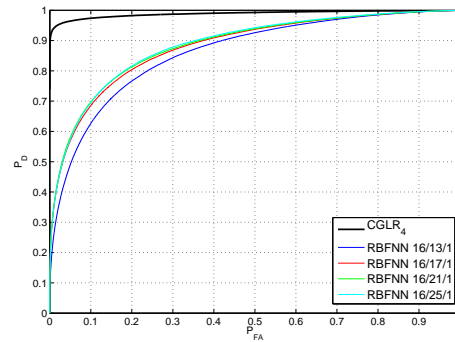
Un análisis de la detección de blanco Swerling II en presencia de AWGN y clutter gaussiano muy correlado permite inferir que las fronteras de decisión tienden a ser hipercilindros en un espacio de $2P$ dimensiones. Según ρ_c tiende a la unidad, el espectro del clutter se hace más estrecho y sólo dos autovalores de la forma cuadrática del detector óptimo son no nulos, confirmando que las fronteras de decisión tienden a formar hipercilindros en el espacio de entrada. Según [Jarabo2003], las soluciones basadas en **un único MLP** son más adecuadas para aproximar este tipo de fronteras de decisión. Para aproximar el ALR se ha diseñado con un conjunto de entrenamiento con patrones bajo H_1 con un ρ_s^t variable uniformemente en el intervalo $[0; 1]$.

En las Figuras 7.10(a), 7.11(a) y 7.12(a) se representa el estudio del tamaño de la capa oculta de los MLPs para cada tipo de clutter. A medida que el clutter se va haciendo más correlado y más dominante, el valor de SIR necesario para obtener una P_D parecida es menor. En la Figura 7.10(a) para una $CNR = 20$ dB, $\rho_c = 0,7$ y $SIR = 0$ dB, un esquema basado en un $MLP_{16/11/1}$ es capaz de aproximar el detector óptimo, sin embargo, cualquier aumento en la CNR o en el ρ_c , Figuras 7.11(a) y 7.12(a), supone que un detector basado en un único MLP ($MLP_{16/17/1}$) presenta elevadas pérdidas de detección aunque el número de neuronas ocultas sea mayor. De hecho, la diferencia con las capacidades de detección obtenidas con el $CGLR_4$ son mayores cuando $CNR = 40$ dB, $\rho_c = 0,995$ y $SIR = -28$ dB, Figura 7.11(a), que cuando $CNR = 40$ dB, $\rho_c = 0,995$ y $SIR = -10$ dB, Figura 7.12(a).

Por otro lado, las Figuras 7.10(b), 7.11(b) y 7.12(b) reflejan que los detectores basados en RBFNNs no presentan una arquitectura idónea para este caso de estudio. El error de aproximación es mayor según el clutter es más correlado y más dominante, es decir, cuando las fronteras de decisión tienden a ser más abiertas.

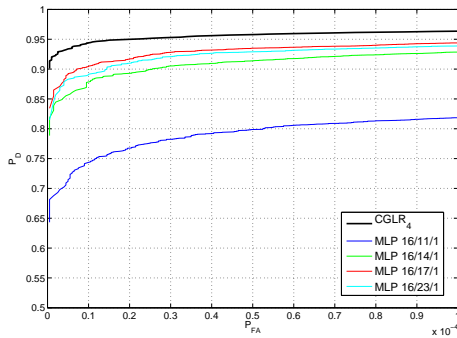


(a) MLP 16/M/1

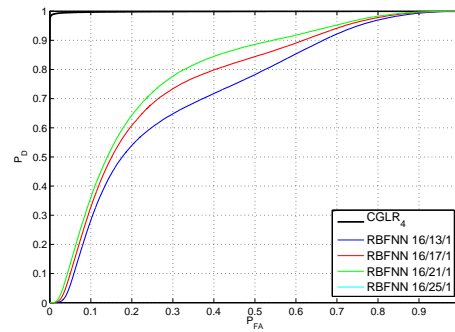


(b) RBFNN 16/M/1

Figura 7.10: Curvas ROC de detectores basados en un MLP y una RBFNN para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$.

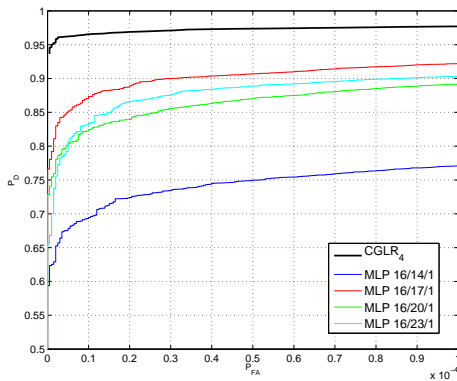


(a) MLP 16/M/1

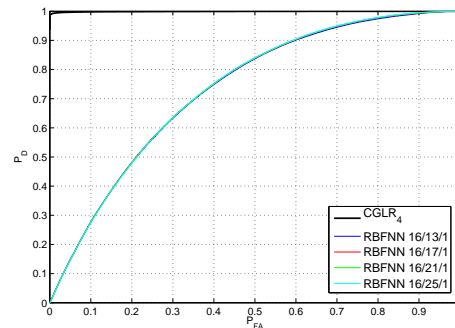


(b) RBFNN 16/M/1

Figura 7.11: Curvas ROC de detectores basados en un MLP y una RBFNN para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0, 995$, $TSIR = DSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$.

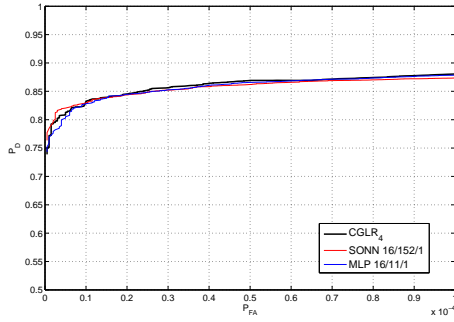


(a) MLP 16/M/1



(b) RBFNN 16/M/1

Figura 7.12: Curvas ROC de detectores basados en un MLP y una RBFNN para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 40$ dB, $\rho_c = 0, 995$, $TSIR = DSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$.

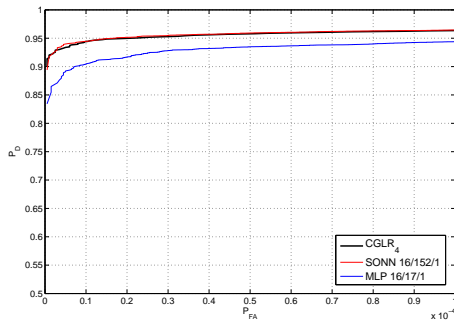


(a) Curvas ROC

Detector	Número de operaciones
CGLR ₄	2.235
SONN 16/152/1	441
MLP 16/11/1	387

(b) Número de operaciones necesarias

Figura 7.13: Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$.



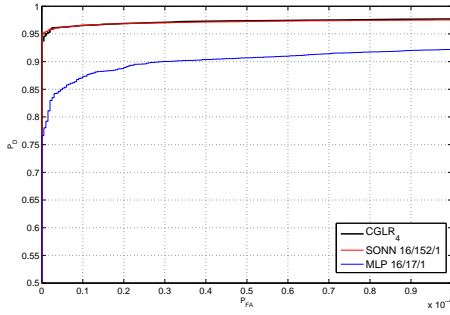
(a) Curvas ROC

Detector	Número de operaciones
CGLR ₄	2.235
SONN 16/152/1	441
MLP 16/17/1	597

(b) Número de operaciones necesarias

Figura 7.14: Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$.

Para los casos en los que $\rho_c = 0,995$ se propone entrenar **una única SONN** con los mismos patrones de entrenamiento y comprobar que el error de aproximación se reduce gracias este tipo de arquitectura que facilita la formación de fronteras de decisión no lineales. Las Figuras 7.13, 7.14 y 7.15, correspondientes a los distintos tipos de clutter gaussianos considerados, demuestran que los detectores SONN son capaces de aproximar el CGLR₄ en todos los casos. Las soluciones basadas en SONNs no sólo consiguen mejores aproximaciones que las basadas en MLPs, sino que lo hacen con un complejidad computacional menor, excepto para $CNR = 20$ dB, $\rho_c = 0,7$ y $SIR = 0$ dB para el que detector MLP 16/11/1 sólo requiere 387 operaciones. Por lo tanto, independientemente del escenario bajo estudio de detección de blancos gaussianos con coeficiente de correlación desconocido en presencia de interferencia gaussiana, se puede aproximar el comportamiento del detector óptimo con una SONN con un coste computacional controlado, que en función de las características del clutter y de la potencia de blanco puede reducirse con un MLP.

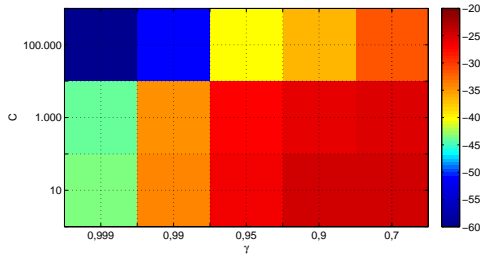
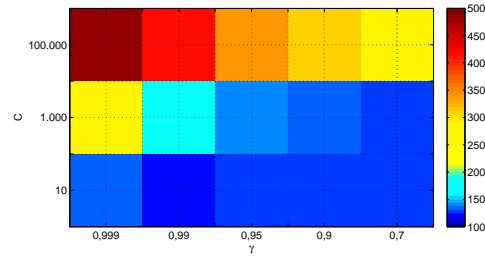


(a) Curvas ROC

Detector	Número de operaciones
CGLR ₄	2.235
SONN 16/152/1	441
MLP 16/17/1	597

(b) Número de operaciones necesarias

Figura 7.15: Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$.

(a) $10 \log(P_{FA})$ 

(b) Número de vectores soporte

Figura 7.16: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$.

7.5.2. Soluciones basadas en SVMs

El punto de partida de las soluciones basadas en SVMs es un búsqueda en rejilla de los parámetros C y γ de la formulación de las 2C-SVM que permiten aproximar el punto de la curva ROC del detector NP para una $P_{FA} = 10^{-4}$ con el menor coste computacional. Las Figuras 7.16, 7.17 y 7.18 muestran los resultados de esta búsqueda para cada uno de los tipos de clutter gaussiano estudiados en términos de P_{FA} estimada, cuando la salida se compara con un umbral fijo $\eta_0 = 0$ y el número de vectores soporte. Sólo para un clutter no dominante y poco correlado, la 2C-SVM es capaz de converger a una solución para $C = 10^5$, Figura 7.16(a), por lo que es de esperar que la aproximación al óptimo sea mejor para este tipo de clutter a pesar de que el número de vectores soporte sea mayor, Figura 7.16(b). La mejor combinación de parámetros de C y γ para este tipo de clutter es 10^5 y $0,95$ respectivamente. Los resultados para un tipo de clutter muy correlado independientemente del valor de CNR , Figuras 7.17 y 7.18, son muy parecidos en términos de P_{FA} , ya que la mejor aproximación se obtiene para $C = 10^3$ y $\gamma = 0,7$ en ambos casos. Para $C = 10^5$, las P_{FA} y P_D obtenidas con un clutter muy correlado son nulas. Sin embargo, el número de vectores soporte se reduce según el clutter se hace más dominante.

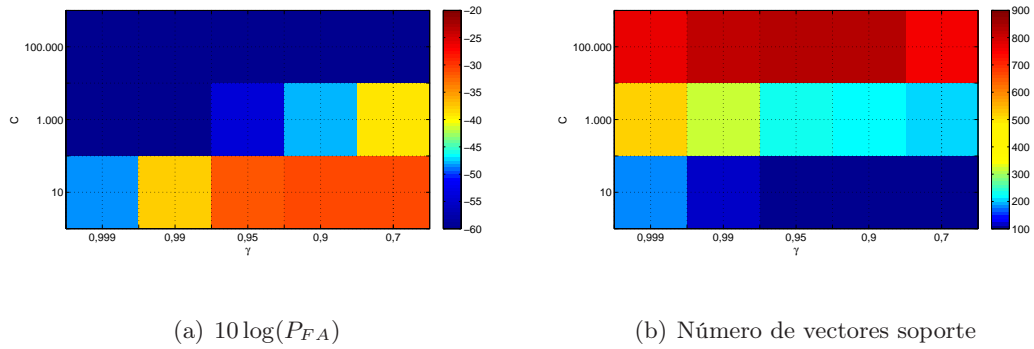


Figura 7.17: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$.

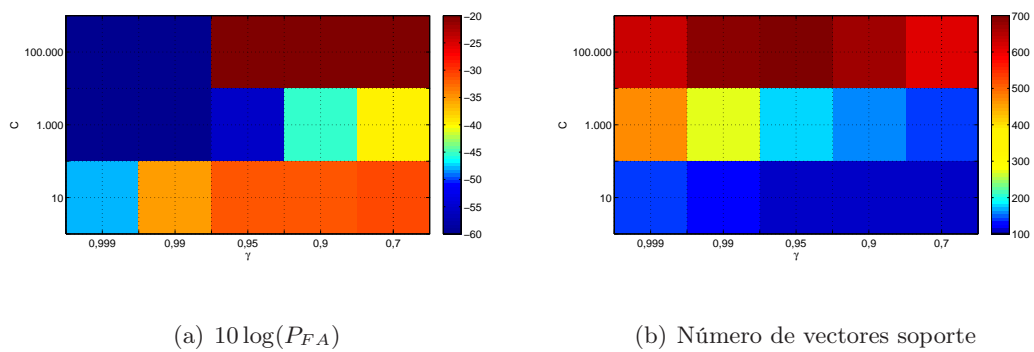


Figura 7.18: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0, 1]$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$.

	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
C	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	$8,50 \cdot 10^{-6}$	0,6654	$9,85 \cdot 10^{-5}$	0,8792	$2,14 \cdot 10^{-4}$	0,9061	$7,22 \cdot 10^{-4}$	0,9468
10^3	$3,34 \cdot 10^{-4}$	0,8525	$1,90 \cdot 10^{-3}$	0,9011	$2,50 \cdot 10^{-3}$	0,9041	$3,90 \cdot 10^{-3}$	0,9230
10	$4,03 \cdot 10^{-4}$	0,8167	$2,20 \cdot 10^{-3}$	0,8824	$3,20 \cdot 10^{-3}$	0,8974	$5,20 \cdot 10^{-3}$	0,9175

Tabla 7.3: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0,1]$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 0$ dB y $\Omega^t = \Omega^s = \pi/2$.

	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
C	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	$4,60 \cdot 10^{-6}$	0,9155	$1,65 \cdot 10^{-5}$	0,9321	$1,02 \cdot 10^{-4}$	0,9449
10	$1,57 \cdot 10^{-4}$	0,9330	$7,21 \cdot 10^{-4}$	0,9509	$7,82 \cdot 10^{-4}$	0,9517	$9,44 \cdot 10^{-4}$	0,9581

Tabla 7.4: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0,1]$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -10$ dB y $\Omega^t = \Omega^s = \pi/2$.

Las Tablas 7.3, 7.4 y 7.5 resumen los valores numéricos de las probabilidades estimadas. También se adjunta la Tabla 7.6 que compara las soluciones basadas en 2C-SVMs con los mejores detectores robustos basados en NNs. Estos resultados muestran que el error de aproximación de las 2C-SVMs es mayor para un clutter más correlado, correspondientes a soluciones con un C más pequeño. A nivel de coste computacional, queda evidente que los detectores 2C-SVMs no son competitivos frente a los esquemas basados en NNs.

7.6. Detección de blancos gaussianos con pulsación Doppler desconocida en interferencia gaussiana

Según el apartado 5.3.2 relativo a los detectores CGLR para este mismo caso de estudio, los datos asumidos para obtener los resultados dependen del tipo de clutter:

	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
C	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	$3,00 \cdot 10^{-6}$	0,9233	$2,80 \cdot 10^{-5}$	0,9300	$9,89 \cdot 10^{-5}$	0,9591
10	$2,95 \cdot 10^{-4}$	0,9054	$7,07 \cdot 10^{-4}$	0,9169	$7,43 \cdot 10^{-4}$	0,9170	$8,74 \cdot 10^{-4}$	0,9182

Tabla 7.5: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 2 con $\rho_s^s \in [0,1]$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -28$ dB y $\Omega^t = \Omega^s = \pi/2$.

	$CNR = 20$ dB $\rho_c = 0,7$	$CNR = 20$ dB $\rho_c = 0,995$	$CNR = 40$ dB $\rho_c = 0,995$
NNs	<i>SONN</i> 16/152/1/1	<i>SONN</i> 16/152/1/1	<i>SONN</i> 16/152/1/1
P_D	0,8796	0,9636	0,9753
Nº de operaciones	441	441	441
2C-SVM	$C = 10^5; \gamma = 0,95$	$C = 10^3; \gamma = 0,7$	$C = 10^3; \gamma = 0,7$
P_D	0,8792	0,9449	0,9591
Nº de operaciones	27,671	13,601	9,648

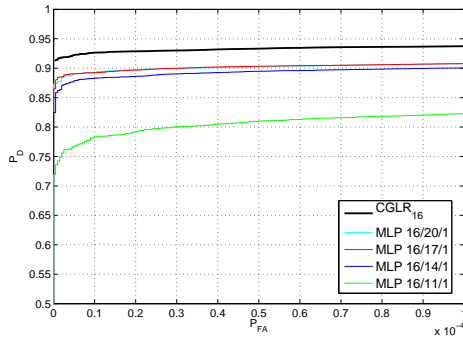
Tabla 7.6: Comparativa de aproximaciones basadas en NNs y SVMs para $P_{FA} = 10^{-4}$ para el caso de estudio 2 con $\rho_s^s \in [0, 1]$.

- Clutter no dominante poco correlado: $CNR = 20$ dB, $\rho_c = 0,7$ y $SIR = 13$ dB.
- Clutter no dominante poco correlado: $CNR = 20$ dB, $\rho_c = 0,995$ y $SIR = 1$ dB.
- Clutter no dominante poco correlado: $CNR = 40$ dB, $\rho_c = 0,995$ y $SIR = -14$ dB.

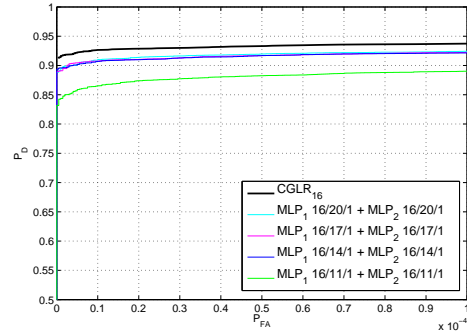
7.6.1. Soluciones basadas en redes neuronales

Los detectores diseñados con parámetros conocidos en los que $\Omega^d > \Omega^s$ presentan menor pérdida de detección que los diseñados con $\Omega^d < \Omega^s$ debido al comportamiento asimétrico observado en 3.4.2. Entonces, para plantear detectores robustos frente a Ω^s desconocida, se propone dar prioridad a Ω^d elevadas, es decir, aquella región del espectro donde las componentes espectrales del clutter son reducidas o casi nulas. En esa región del espectro, habrá blanco Swerling I más AWGN. La primera propuesta se centra en **un único MLP** entrenado con patrones bajo H_1 asumiendo una variación uniforme $\Omega \in [0; 2\pi)$. Los resultados para cada clutter gaussiano se han presentado en las Figuras 7.19(a), 7.20(a) y 7.21(a). El número de neuronas a partir del cuál deja de haber una mejora significativa es $M = 17$ (*MLP*16/17/1) para un $\rho_c = 0,7$ (Figura 7.19(a)) y $M = 14$ (*MLP*16/14/1) para un $\rho_c = 0,995$ (Figuras 7.20(a) y 7.21(a)). Los resultados obtenidos demuestran la limitación de este tipo de arquitecturas basadas en un único MLP. El error de las aproximaciones, independientemente de los parámetros del clutter, es aproximadamente un 3% con respecto al detector CGLR. Debido a que la pérdida de la detección es considerable, se propone diseñar **mezclas de expertos MLPs** donde el espacio de entrada se divide en subespacios más pequeños en los que se especializarán cada una de las redes. El comité $MLP_1 + MLP_2$ está formado por dos redes entrenadas con patrones bajo H_1 asumiendo una variación uniforme $\Omega^t \in [0; \pi/4) \cup [3\pi/4; 2\pi)$ y $\Omega^t \in [\pi/4; 3\pi/4)$ respectivamente. Los resultados de combinar las salidas de los expertos MLPs por selección máxima se representan en las Figuras 7.19(b), 7.20(b) y 7.21(b) para los distintos escenarios.

La capacidad de detección, como era de esperar, ha mejorado. La pérdida de detección para una misma P_{FA} en comparación con el *CGLR*₁₆ se ve reducida aproximadamente a la mitad, un 1,5%. Además, se ha estudiado la posibilidad de reducir la complejidad de los MLPs que forman el comité al especializarse en un problema en principio más sencillo. Sólo en el caso de un clutter poco correlado, $\rho_c = 0,7$ (Figura 7.19(b)), es viable reducir el tamaño de las redes MLPs (*MLP*₁ 16/14/1 + *MLP*₂ 16/14/1) con respecto al número de neuronas ocultas elegido para la soluciones

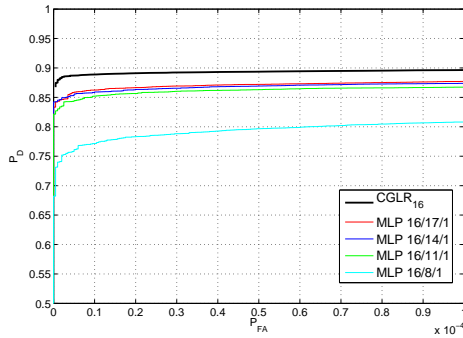


(a) MLP 16/M/1

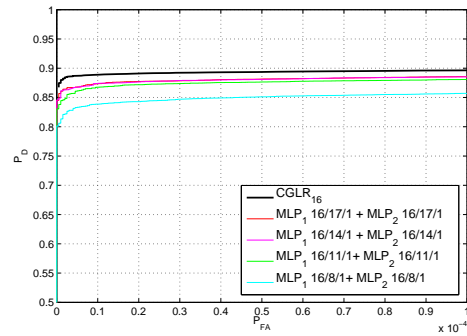


(b) MLP₁ 16/M/1 + MLP₂ 16/M/1

Figura 7.19: Curvas ROC de detectores basados en MLPs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$.

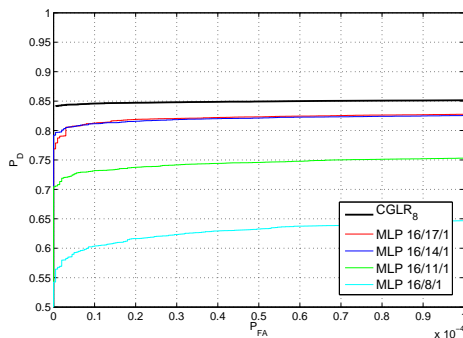


(a) MLP 16/M/1

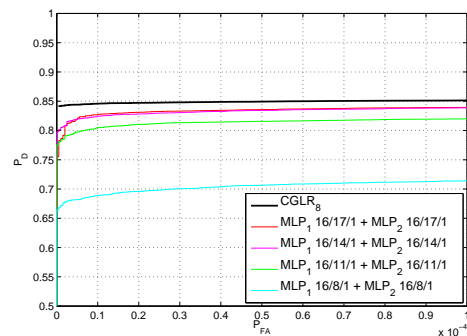


(b) MLP₁ 16/M/1 + MLP₂ 16/M/1

Figura 7.20: Curvas ROC de detectores basados en MLPs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$.

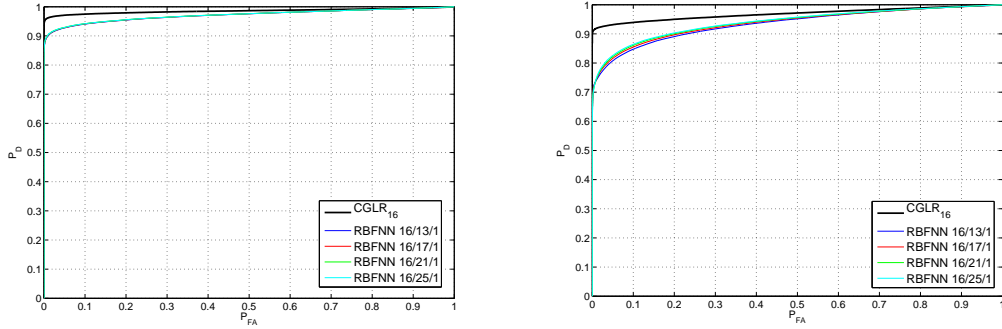


(a) MLP 16/M/1

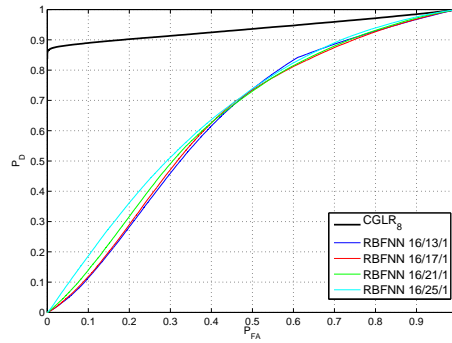


(b) MLP₁ 16/M/1 + MLP₂ 16/M/1

Figura 7.21: Curvas ROC de detectores basados en MLPs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$.



(a) $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 13$ dB (b) $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = 1$ dB



(c) $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -14$ dB

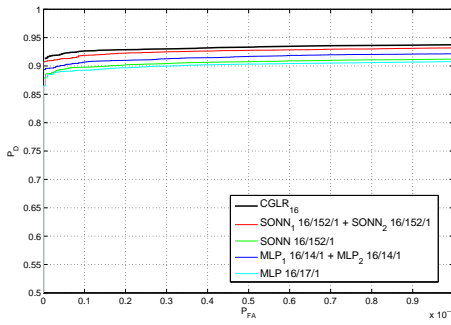
Figura 7.22: Curvas ROC de detectores basados en una única RBF 16/M/1 para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $\rho_s^t = \rho_s^s = 1$.

basadas en una única MLP (MLP 16/17/1) sin sufrir una pérdida en P_D . Probablemente, la aproximación se podría mejorar, para todos los escenarios considerados, dividiendo los márgenes de variación de Ω en intervalos más pequeños. La desventaja de esta estrategia es el aumento en el número de operaciones requerido.

Al igual que en el apartado 7.5, la Figura 7.22 confirma que las soluciones basadas en RBFNNs no presentan la capacidad de aproximar el CGLR. Las peores capacidades de detección se presentan para un clutter gaussiano de $CNR = 40$ dB y $\rho_c = 0,995$, Figura 7.22(c).

Como alternativa se proponen soluciones basadas en **una única SONN** y en una **mezcla de expertos SONNs**. Las unidades neuronales cuadráticas presentan una arquitectura que ya se ha demostrado en la sección 7.4 que son capaces de aproximar el $CGLR_{16}$ para blancos Swerling I con Ω desconocida en AWGN.

En las Figuras 7.23(a), 7.24(a) y 7.25(a) se representan las curvas ROC de las soluciones más destacadas para todos los tipos de clutter gaussiano considerado. El error de aproximación basado en una única SONN disminuye a medida que el clutter es más correlado y más dominante, hasta tal punto que para un clutter con $CNR = 40$ dB y $\rho_c = 0,995$, el detector $SONN$ 16/152/1 es capaz de aproximar el detector $CGLR_8$, Figura 7.25(a). Sin embargo para una $CNR = 20$ dB, tanto para $\rho_c = 0,7$ como para $\rho_c = 0,995$, es necesario un esquema basado en un comité

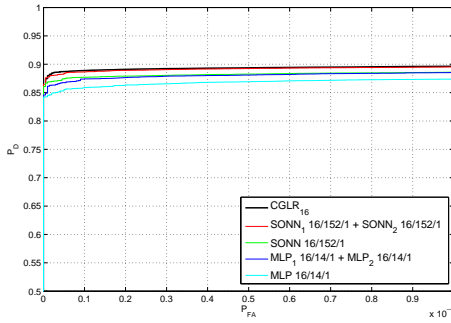


(a) Curvas ROC

Detector	Número de operaciones
CGLR ₁₆	8.943
SONN ₁ 16/152/1 + SONN ₂ 16/152/1	882
SONN 16/152/1	441
MLP ₁ 16/14/1 + MLP ₂ 16/14/1	984
MLP 16/17/1	597

(b) Número de operaciones requerido

Figura 7.23: Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$.



(a) Curvas ROC

Detector	Número de operaciones
CGLR ₁₆	8.943
SONN ₁ 16/152/1 + SONN ₂ 16/152/1	882
SONN 16/152/1	441
MLP ₁ 16/14/1 + MLP ₂ 16/14/1	984
MLP 16/14/1	492

(b) Número de operaciones requerido

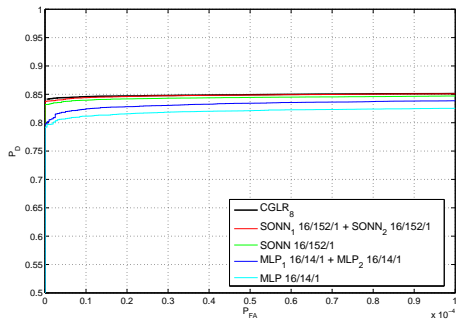
Figura 7.24: Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$.

$SONN_1$ 16/152/1 + $SONN_2$ 16/152/1 para obtener una curva ROC similar a la aproximación basada en el $CGLR_{16}$.

Los detectores basados en mezclas de expertos SONNs necesitan 882 operaciones para obtener un esquema robusto frente a Ω aproximando el comportamiento del CGLR en todos los casos planteados, un número mucho más pequeño que el número de operaciones requerido por los detectores CGLR. Para un clutter dominante muy correlado ($CNR = 40$ dB y $\rho_c = 0,995$), la mejora introducida por la combinación de redes es tan pequeña que se podría proponer el detector $SONN$ 16/152/1/1, que requiere 441 operaciones, como aproximación válida al detector óptimo.

7.6.2. Soluciones basadas en SVMs

En las Figuras 7.26(a), 7.27(a) y 7.28(a) se representa la P_{FA} estimada al comparar la salida de cada 2C-SVM con una umbral fijo $\eta_0 = 0$ para los tipos de clutter considerados. Al igual que ocurría cuando el parámetro desconocido era el coeficiente de correlación del blanco, el tamaño del conjunto de entrenamiento permite a la 2C-SVM converger a una solución para $C = 10^5$ cuando el clutter es poco correlado, mientras que para un clutter correlado, el mayor valor de C

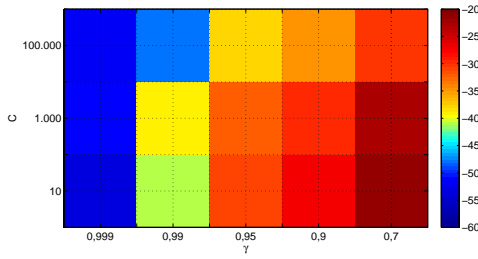


(a) Curvas ROC

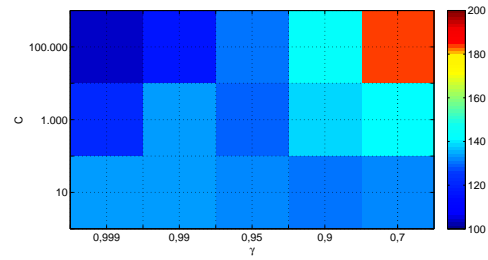
Detector	Número de operaciones
CGLR ₈	4.467
SONN ₁ 16/152/1 + SONN ₂ 16/152/1	882
SONN 16/152/1	441
MLP ₁ 16/14/1 + MLP ₂ 16/14/1	984
MLP 16/14/1	492

(b) Número de operaciones requerido

Figura 7.25: Resumen de soluciones basadas en NNs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$.



(a) $10 \log(P_{FA})$



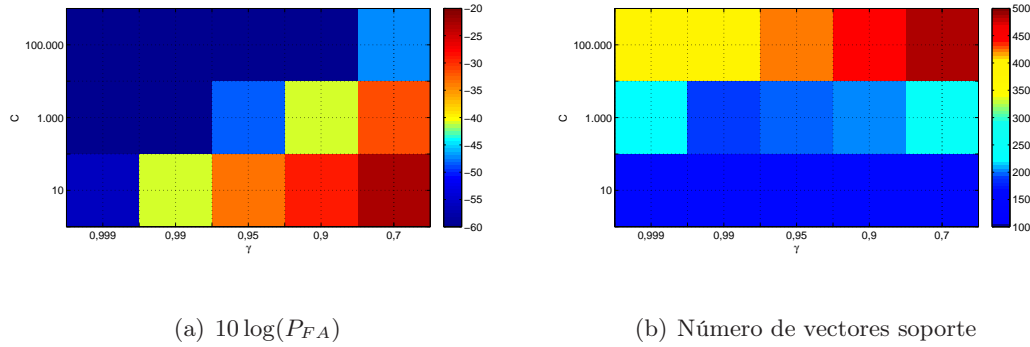
(b) Número de vectores soporte

Figura 7.26: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$.

que permite llegar a una solución en las que P_{FA} y P_D son no nulas es 10^3 .

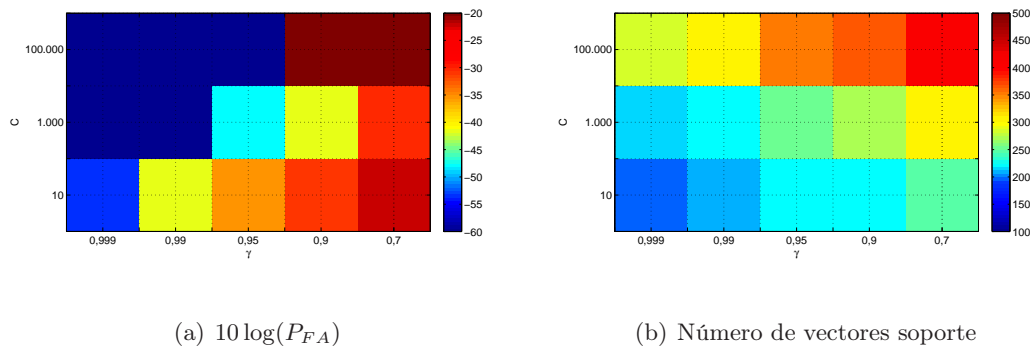
El número de vectores soporte se representa en las Figuras 7.26(b), 7.27(b) y 7.28(b). En general, el número de vectores soporte se ha reducido considerablemente, especialmente para un clutter poco correlado donde la solución con mayor número de vectores soporte ($C = 10^5$ y $\gamma = 0,7$) sólo presenta 183 vectores. Este número tan pequeño, cuando el valor de C es tan grande que casi no hay generalización, invita a pensar que los resultados no van a proporcionar buenas aproximaciones al detector de NP.

La probabilidades de detección se reflejan en las Tablas 7.7, 7.8 y 7.9, donde se ratifica que los detectores 2C-SVM no consiguen una buena aproximación del detector óptimo para $P_{FA} = 10^{-4}$. Una posible forma de mejorar estos resultados es aumentar el tamaño del conjunto de entrenamiento. En otros escenarios, ya había quedado patente que las soluciones basadas en 2C-SVM presentan mayor complejidad computacional que las basadas en NNs, pero además en este caso los detectores neuronales basados en mezclas de expertos SONNs son capaces de obtener una buena aproximación de toda la curva ROC del detector NP.

(a) $10 \log(P_{FA})$

(b) Número de vectores soporte

Figura 7.27: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$.

(a) $10 \log(P_{FA})$

(b) Número de vectores soporte

Figura 7.28: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$.

C	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	$1,55 \cdot 10^{-5}$	0,5364	$1,67 \cdot 10^{-4}$	0,5559	$3,18 \cdot 10^{-4}$	0,5439	$8,84 \cdot 10^{-4}$	0,5339
10^3	$1,31 \cdot 10^{-4}$	0,5423	$6,02 \cdot 10^{-4}$	0,5245	$1,10 \cdot 10^{-3}$	0,5173	$4,90 \cdot 10^{-3}$	0,5304
10	$8,55 \cdot 10^{-4}$	0,4919	$1,90 \cdot 10^{-3}$	0,5045	$6,90 \cdot 10^{-3}$	0,5270	$8,72 \cdot 10^{-3}$	0,5386

Tabla 7.7: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,7$, $TSIR = SSIR = 13$ dB y $\rho_s^t = \rho_s^s = 1$.

	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
C	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	$1,45 \cdot 10^{-5}$	0,6027	$8,1 \cdot 10^{-5}$	0,6687	$6,72 \cdot 10^{-4}$	0,6793
10	$8,40 \cdot 10^{-5}$	0,6220	$4,63 \cdot 10^{-4}$	0,6280	$1,3 \cdot 10^{-3}$	0,6641	$5,40 \cdot 10^{-3}$	0,6694

Tabla 7.8: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 20$ dB, $\rho_c = 0,995$, $TSIR = SSIR = 1$ dB y $\rho_s^t = \rho_s^s = 1$.

	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
C	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	$5,00 \cdot 10^{-7}$	0,5791	$1,70 \cdot 10^{-5}$	0,6226	$8,85 \cdot 10^{-5}$	0,6302
10	$4,50 \cdot 10^{-6}$	0,6220	$7,30 \cdot 10^{-5}$	0,6548	$2,81 \cdot 10^{-4}$	0,6720	$8,56 \cdot 10^{-4}$	0,6782

Tabla 7.9: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 2 con $\Omega^s \in [0, 2\pi)$. $CNR = 40$ dB, $\rho_c = 0,995$, $TSIR = SSIR = -14$ dB y $\rho_s^t = \rho_s^s = 1$.

7.7. Detectores de blancos Swerling V con pulsación Doppler desconocida en clutter K-distribuido

Para ángulos de incidencia rasante pequeños y sistemas radar de alta resolución, la amplitud de los ecos recibidos por un clutter marino siguen una PDF que dista de ser Rayleigh y que se puede modelar con un clutter compuesto K-distribuido. Para estados del mar elevados, el parámetro de forma de la distribución K suele tomar valores en torno $0,1 - 1$. Por ejemplo, en [Greco2004], los datos adquiridos por el radar IPIX el 7 de Noviembre de 1993 con una altura significativa de las olas de 2,23 m siguen una PDF de amplitud que se ajusta a una distribución K con unos parámetros de forma que van desde $0,28 - 0,45$ en función de la celda de rango observada.

Continuando los estudios de [Conte1991], [Pentini1992] y [Conte1994], los tipos de clutter K distribuidos considerados se caracterizan por $\nu = 0,5$ para modelar un clutter impulsivo incorrelado ($\rho_c = 0$) y correlado ($\rho_c = 0,9$). La aproximación basada en el $CGLR_K$ conlleva el cálculo de una función de Bessel modificada de segunda especie por cada filtro K considerado, cuyo coste computacional puede ser muy elevado, por lo que las soluciones basadas en inteligencia artificial, que como se ha comprobado en los apartados 7.4 y 7.6 reducen considerablemente el número de operaciones requerido, pueden resultar de especial interés en este caso de estudio.

En los apartados anteriores, se partía de un conocimiento previo de las fronteras de decisión formadas por el detector NP, que permitía intuir a priori el tipo de red más adecuado para cada situación. Para poder determinar la base experimental de las soluciones basadas en redes neuronales, en este caso, se propone hacer un estudio de las superficies de decisión de la salida del detector $CGLR_K$ para blancos Swerling V con una Ω variable uniformemente en el intervalo $[\pi/2 - \pi/10; \pi/2 + \pi/10]$ en presencia de clutter K distribuido.

Según el apartado 5.3.3 relativo a los detectores CGLR para este mismo caso de estudio, los datos asumidos para obtener los resultados dependen del tipo de clutter:

- Clutter incorrelado: $\rho_c = 0$ y $SCR = 9$ dB.
- Clutter correlado: $\rho_c = 0,9$ y $SCR = -3$ dB..

7.7.1. Fronteras de decisión del detector CGLR

Para poder representar las regiones de decisión, se ha supuesto que los vectores de entrada son de dimensión 2D, correspondientes a la parte real e imaginaria de un pulso. El objetivo es representar las salidas obtenidas de un detector basado en un CGLR, ecuación (7.2), que asume un margen de variación de Ω restringido ($[\pi/2 - \pi/10; \pi/2 + \pi/10]$).

$$F(\tilde{\mathbf{z}}) = \max_{k=[\pi/2-\pi/10:\pi/2+\pi/10]} \Lambda(\Omega_k) \quad (7.2)$$

El valor de K depende del tipo de clutter considerado. Según los resultados presentados en 5.3.3, para un clutter incorrelado $K = 8$ valores discretos de Ω son suficientes para obtener una aproximación que casi no se ve mejorada al aumentar con un valor mayor de K , mientras que para un clutter incorrelado se han considerado como comportamiento de referencia el obtenido con un $CGLR_{32}$.

Los resultados obtenidos se presentan en la Figura 7.29. En ambos casos, las fronteras de decisión corresponden a un segmento de una corona circular cuyo radio depende de la potencia del blanco, siendo mayor para $SCR = 9$ dB (Figura 7.29(a)) que para $SCR = -3$ dB (Figura 7.29(b)). El segmento circular se corresponde con las fases $[3\pi/4 - \pi/10; 3\pi/4 + \pi/10]$, resultado de sumar la fase asumida en el diseño, $\varphi = \pi/4$, y la frecuencia Doppler uniformemente distribuida en $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. Si Ω variase uniformemente en $[0; 2\pi]$, en el resultado se observaría una corona circular completa de radio definido por el valor de SCR .

Para extender este resultado a $2P$ dimensiones, se presenta en la Figura 7.30 el corte de un plano de 2D, cuyos ejes son la parte real e imaginaria de uno de los pulsos del vector de entrada, de las fronteras de decisión formadas un CGLR de 8 pulsos para un clutter K incorrelado. Las fronteras de decisión formadas dependen de las formas cuadráticas $\tilde{\mathbf{z}}^T \Sigma_{\tilde{\mathbf{c}\tilde{\mathbf{c}}}}^{-1} \tilde{\mathbf{z}}^*$ y $(\tilde{\mathbf{z}} - \tilde{\mathbf{s}})^T \Sigma_{\tilde{\mathbf{c}\tilde{\mathbf{c}}}^{-1}} (\tilde{\mathbf{z}} - \tilde{\mathbf{s}})^*$. Como en la Figura 7.30 se observa, que para un corte cualquiera, se sigue teniendo el segmento circular. Por lo que, se podría deducir que $\Sigma_{\tilde{\mathbf{c}\tilde{\mathbf{c}}}}$ y su correspondiente $\Sigma_{\mathbf{c}\mathbf{c}}$ son matrices diagonales que tiene todos sus ejes de igual longitud. Entonces, las superficies de decisión que se formarán en un espacio de $2P$ dimensiones posiblemente se definan con dos hiperesferas, de manera que la región interna delimitadas por las dos hiperesferas se corresponderá con una decisión de presencia de blanco y el resto con la ausencia de blanco. Conviene recordar que como el margen de variación de Ω es restringido, las fronteras de decisión serán segmentos de las fronteras de decisión definidas por las hiperesferas.

Para el caso de clutter correlado, no se puede inferir a priori la forma de las fronteras de decisión que se formarán en un espacio de entrada de $2P$ dimensiones. De todas formas, como las RBFNN tienen facilidad para encerrar hiperesferas, la evolución de los diseños basados en redes neuronales empezará por este tipo de redes para ir comparando las mejoras que se introducen con otras arquitecturas.

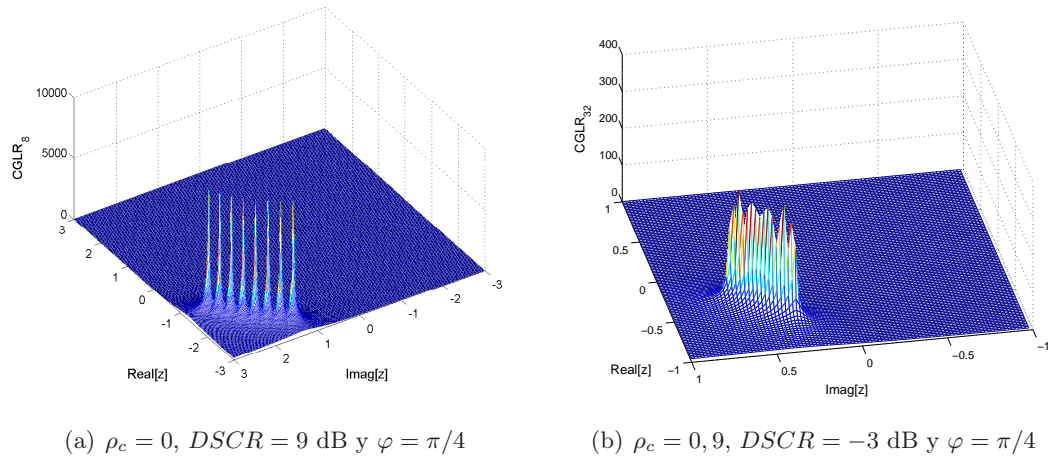


Figura 7.29: Fronteras de decisión formadas por un detector CGLR para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$.

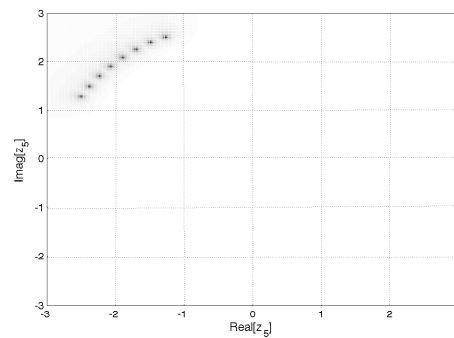
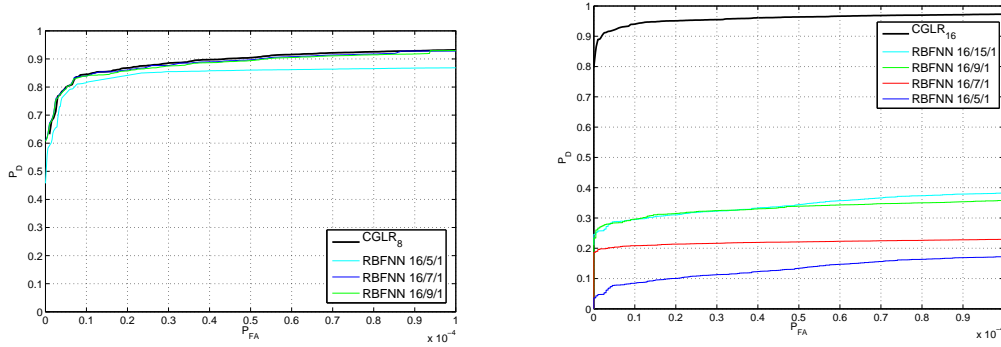


Figura 7.30: Corte bidimensional de las fronteras de decisión formadas por un detector CGLR con $P8$ para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$, $\rho_c = 0$, $DSCR = 9$ dB y $\varphi = \pi/4$.



(a) $\rho_c = 0$, $TSCR = SSCR = 9$ dB y $\varphi = \pi/4$ (b) $\rho_c = 0, 9$, $TSCR = SSCR = -3$ dB y $\varphi = \pi/4$

Figura 7.31: Curvas ROC de detectores RBFNNs 16/ M /1 para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$.

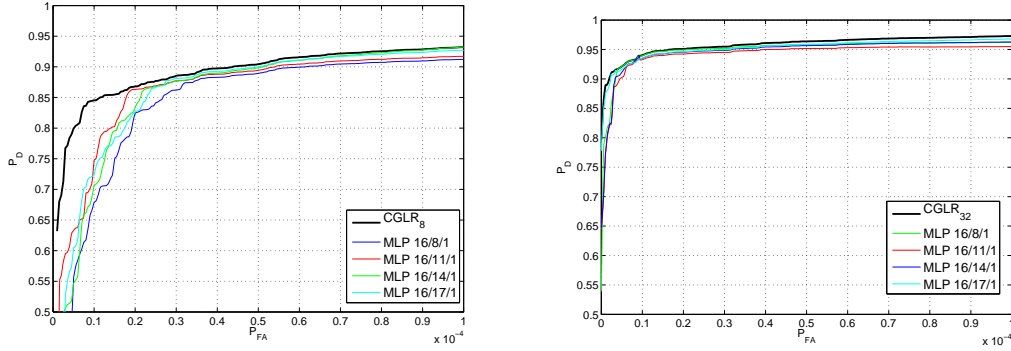
7.7.2. Soluciones basadas en redes neuronales

El estudio de resultados de las soluciones basadas en redes neuronales para el caso de estudio 3 se realiza en paralelo para interferencias impulsivas K-distribuidas con coeficientes de correlación $\rho_c = 0$ y $\rho_c = 0, 9$. Las primeras propuestas se basarán en **una única RBFNN**. El conjunto de entrenamiento se ha diseñado con patrones bajo H_1 asumiendo una variación uniforme $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$.

Los resultados obtenidos en 7.7.1 para $\rho_c = 0$, invitan a pensar que las RBFNNs son más adecuadas para este tipo de clutter. En la Figura 7.31(a), se muestra como una RBFNN con 7 funciones base puede implementar una aproximación muy buena al $CGLR_8$. En cambio, para una $\rho_c = 0, 9$ (Figura 7.31(b)), todas las RBFNNs consideradas presentan una pérdida de detección muy importante. Además la mejora observada al incrementar el número de neuronas de 9 a 15 es insignificante.

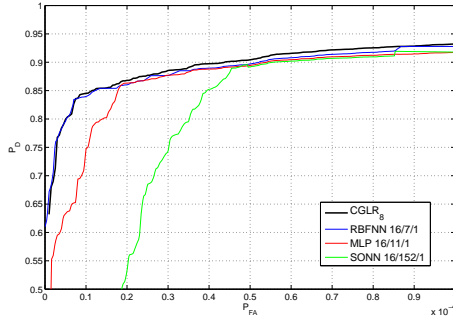
Estos resultados inducen a esperar que las soluciones basadas en MLPs, las cuales suelen ser complementarias a las basadas en RBFNNs, puedan responder mejor ante ρ_c más elevados. En la Figura 7.32(b), se observa cómo un MLP 16/17/1 es capaz de aproximar el comportamiento del $CGLR_{16}$ para $\rho_c = 0, 9$, mejorando de forma clara los resultados obtenidos con RBFNNs (Figura 7.31(b)). En cambio, para $\rho_c = 0$ (Figura 7.32(a)), las soluciones basadas en MLPs son claramente inferiores a las basadas en RBFNNs.

Por último, queda comprobar si una solución basada en **una única SONN** es capaz de obtener una buena aproximación. En las Figuras 7.33(a) y 7.34(a) se observa, tanto para $\rho_c = 0$ como para $\rho_c = 0, 9$, una fuerte pérdida de detección a partir de un cierto valor de P_{FA} (en torno a 0,00005 para $\rho_c = 0$ y 0,00001 para $\rho_c = 0, 9$). La arquitectura de las SONNs combina las relaciones de primer y segundo orden, es decir, comparte propiedades de los MLPs y las RBFNNs. Para entender porqué el error de aproximación es tan grande se han estudiado las fronteras de decisión formadas por los detectores SONNs para vectores de entrada compuestos por las componentes en fase y cuadratura de un pulso. Si se comparan las Figuras 7.35 y 7.29, se observa cómo es capaz de definir la región interna del segmento circular deseado de forma correcta, pero no ocurre lo mismo con la región externa, dando lugar a una P_{FA} mayor. Posiblemente,



(a) $\rho_c = 0$, $TSCR = SSCR = 9$ dB y $\varphi = \pi/4$ (b) $\rho_c = 0, 9$, $TSCR = SSCR = -3$ dB y $\varphi = \pi/4$

Figura 7.32: Curvas ROC de detectores MLPs 16/M/1 para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$.



(a) Curvas ROC

Detector	Número de operaciones
CGLR ₈	4.471
RBFNN 16/7/1	358
MLP 16/11/1	387
SONN 16/152/1	441

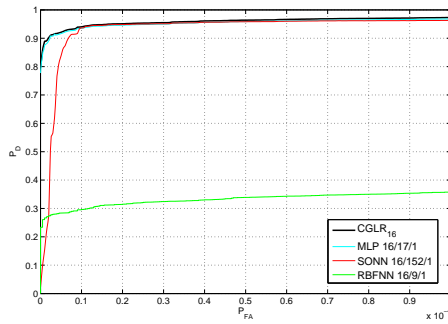
(b) Número de operaciones requerido

Figura 7.33: Resumen de soluciones basadas en NNs para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0$, $TSCR = SSCR = 9$ dB y $\varphi = \pi/4$.

este comportamiento radique en que los detectores SONNs propuestos en esta Tesis, basados en una única unidad neuronal cuadrática, son demasiado sencillas para resolver este problema tan complejo. Los resultados podrían mejorar planteando una mezcla de expertos SONNs o una única SONN con una capa oculta con M neuronas cuadráticas. Sin embargo, estas propuestas no son de interés puesto que el agente inteligente resultante tendría mayor coste computacional que las aproximaciones basadas en un MLP 16/17/1, que requiere 593 operaciones, y una RBFNN 16/7/1, que necesita 358 operaciones, para un clutter de distribución K correlado e incorrelado, respectivamente.

7.7.3. Soluciones basadas en SVMs

En este apartado se presentan las aproximaciones basadas en 2C-SVMs para blancos Swerling V con Ω variable uniformemente en el intervalo $[\pi/2 - \pi/10; \pi/2 + \pi/10]$, en presencia de clutter impulsivo K-distribuido. Aunque la salida de la 2C-SVM se compara con un umbral fijo, este se puede asociar a unos requisitos de P_{FA} determinados entrenando la 2C-SVM con el parámetro γ adecuado. Este valor de γ guarda una relación con η_{lr} ($\eta_{lr} = \frac{\gamma}{1-\gamma}$) que es independiente del

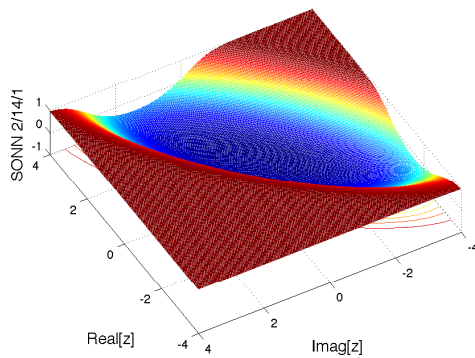


(a) Curvas ROC

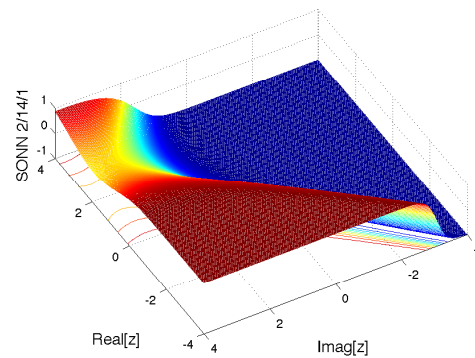
Detector	Número de operaciones
CGLR ₁₆	18.879
MLP 16/17/1	597
RBFNN 16/9/1	460
SONN 16/152/1	441

(b) Número de operaciones requerido

Figura 7.34: Resumen de soluciones basadas en NNs para el caso de estudio 3 con $\Omega^s \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0,9$, $TSCR = SSCR = -3$ dB y $\varphi = \pi/4$.



(a) $\rho_c = 0$, $TSCR = 9$ dB y $\varphi = \pi/4$



(b) $\rho_c = 0,9$, $TSCR = -3$ dB y $\varphi = \pi/4$

Figura 7.35: Fronteras de decisión formadas por un detector SONN 16/152/1 para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$.

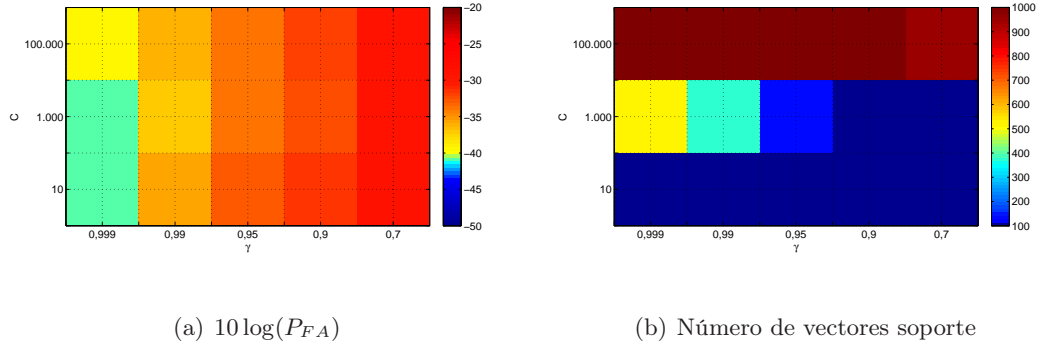


Figura 7.36: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0$, $TSCR = 9$ dB y $\varphi = \pi/4$.

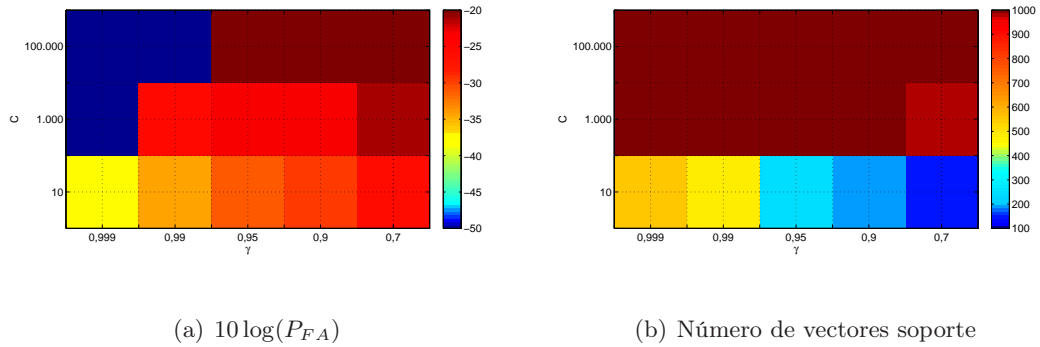


Figura 7.37: Búsqueda en rejilla $[C; \gamma]$ de las 2C-SVMs para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. $\rho_c = 0,9$, $TSCR = -3$ dB y $\varphi = \pi/4$.

vector de entrada. En este caso de estudio, el ALR presenta una integral muy compleja que hace muy complejo el cálculo de η_{lr} asociado a $P_{FA} = 10^{-4}$. Se propone un búsqueda en rejilla de los parámetros C y γ que nos proporcionen las mejores capacidades de detección.

La P_{FA} y el número de vectores soporte resultante de cada entrenamiento se presentan en la Figura 7.36 para un clutter incorrelado y en 7.37 para un clutter correlado. Mientras que para un clutter incorrelado se pueden encontrar soluciones para $C = 10^5$, en un clutter correlado sólo se tienen resultados para $C = 10$ (ya que para $C > 10$ resulta $P_{FA} = P_D = 0$). Antes de estudiar la P_D correspondientes y el número de vectores soporte, ya se puede intuir que la aproximación será mejor y más costosa computacionalmente para un clutter incorrelado, ya que según C es más grande el error, el parámetro de regularización tiende a anularse y se reduce el error de aproximación. A partir de las Figuras 7.36 y 7.37, y de las Tablas 7.10 y 7.11 donde se indican la P_{FA} y la P_D para $\eta_0 = 0$, se puede determinar como mejor pareja de parámetros $C = 10^5$ y $\gamma = 0,999$ para un clutter con $\rho_c = 0$ y $C = 10$ y $\gamma = 0,999$ para $\rho_c = 0,9$.

La Tabla 7.12 resume las capacidades de detección y el coste computacional asociado a las mejores soluciones basadas en NNs y en 2C-SVMs. De la comparación podemos concluir que todas las soluciones presentan P_D parecidas y que los detectores 2C-SVM necesitan muchas

C	$\gamma = 0,999$		$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$	
	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	$1,13 \cdot 10^{-4}$	0,9307	$2,54 \cdot 10^{-4}$	0,9435	$4,29 \cdot 10^{-4}$	0,9520	$6,35 \cdot 10^{-4}$	0,9553
10^3	$8,10 \cdot 10^{-5}$	0,9261	$1,87 \cdot 10^{-4}$	0,9424	$4,13 \cdot 10^{-4}$	0,9520	$5,43 \cdot 10^{-4}$	0,9546
10	$8,20 \cdot 10^{-5}$	0,9249	$2,72 \cdot 10^{-4}$	0,9337	$5,38 \cdot 10^{-4}$	0,9429	$7,01 \cdot 10^{-4}$	0,9547

Tabla 7.10: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$, $\rho_c = 0$, $TSCR = 9$ dB y $\varphi = \pi/4$.

C	$\gamma = 0,99$		$\gamma = 0,95$		$\gamma = 0,9$		$\gamma = 0,7$	
	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
10^5	-	-	-	-	-	-	-	-
10^3	-	-	-	-	-	-	-	-
10	$1,92 \cdot 10^{-4}$	0,9217	$3,53 \cdot 10^{-4}$	0,9368	$7,74 \cdot 10^{-4}$	0,9428	$1,1 \cdot 10^{-3}$	0,9685

Tabla 7.11: Probabilidades estimadas con un umbral fijo $\eta_0 = 0$ para detectores 2C-SVMs para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$, $\rho_c = 0,9$, $TSCR = -3$ dB y $\varphi = \pi/4$.

más operaciones que las soluciones basadas en NNs. En concreto, para el caso de clutter de distribución K, se confirma que la aproximación es mejor para $\rho_c = 0$ que para $\rho_c = 0,9$ y que el número de operaciones requeridas también es mucho mayor que para un clutter correlado.

7.8. Detectores robustos aplicados a un escenario radar

En esta sección se presentan los resultados obtenidos tras aplicar los detectores neuronales propuestos a los casos de estudio en un escenario radar simulado. Se han aplicado técnicas CA-CFAR incoherentes como detectores de referencia para comparar los resultados.

Debido a las características del sistema radar, un blanco podría detectarse en P celdas de acimut. El extractor de datos radar tiene por misión estimar el centroide de los pulsos detectados para determinar la posición puntual del blanco (plot) en rango y acimut. Además, esta etapa de procesamiento puede utilizarse para reducir falsas alarmas. Las fases del algoritmo implementado se resumen en los siguientes pasos:

	$\rho_c = 0$	$\rho_c = 0,9$
NNs	<i>RBFNN</i> 16/7/1	<i>MLP</i> 16/17/1
P_D	0,9296	0,9728
Nº de operaciones	358	597
2C-SVM	$C = 10^5; \gamma = 0,999$	$C = 10; \gamma = 0,999$
P_D	0,9307	0,9217
Nº de operaciones	66,665	37,386

Tabla 7.12: Comparativa de aproximaciones basadas en NNs y SVMs para $P_{FA} = 10^{-4}$ para el caso de estudio 3 con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$.

Potencia de pico	8 kW	Ancho de banda	5,5 MHz
Ganancia de antena	44 dB	Resolución	30 m
Ancho de haz en acimut	0,9°	PRF	1,600 kHz
Polarización	HH	Velocidad de giro de la antena	30 rpm
Figura de ruido	1,2 dB	Frecuencia portadora	9,39 GHz

Tabla 7.13: Parámetros considerados del radar IPIX para el escenario simulado.

- Etapa de clustering: detección de bloques de *hints* (unos de la matriz) conectados.
- Estimación del número de *hints* de cada cluster.
- Eliminación de aquellos clusters que no tengan, al menos, una celda en rango y $P/2$ en acimut.
- Estimación de los centroides de los clusters confirmados.

7.8.1. Escenario radar simulado

Los casos de estudio se han utilizado para la simulación de un escenario radar, en el que se considera un hipotético radar situado en el sur de la península Ibérica. Se asume que este radar es similar al radar IPIX de la Universidad McMaster de Canadá. El IPIX es un radar coherente de exploración experimental en banda X. En la Tabla 7.13, se describen los parámetros considerados en la simulación del escenario, los cuales se encuentran dentro de los márgenes de funcionamiento descritos en [Drosopoulos1994],

Con los parámetros de la Tabla 7.13, el tiempo durante el cual un blanco es iluminado en una exploración es igual a 5s, que, para una $PRF = 1.600$ Hz, da lugar a un total de $P = 8$ pulsos recibidos de un blanco en una exploración. Se considera un sector de cobertura de $23,9^\circ$ centrado en el punto de coordenadas 36,00116 N, 5,60966 O, con una cobertura de 18 MN (33,34 km). El modelo teórico de clutter marino depende del ángulo de incidencia, por lo que se asumió un ángulo de 13° para un escenario con clutter gaussiano y de $0,1^\circ$ para clutter K-distribuido.

Se ha utilizado la herramienta *Localizatodo* para importar trayectorias de barcos reales [Localizatodo]. La herramienta web *Localizatodo* usa y combina datos, presentaciones y funcionalidad procedentes de Google Maps y Google Earth para crear un nuevo servicio que facilita la localización y geolocalización en tiempo real de personas, barcos y aviones. Para aplicaciones marítimas, incluye información de puertos, detalles de los barcos y sus rutas, faros, radares, pasillos marítimos y helicópteros SAR. En la Figura 7.38 se muestra una imagen en la que aparece el emplazamiento del radar, su sector de cobertura y los barcos que en el momento de captura de la imagen se encontraban allí.

En dicho entorno, se simulará la trayectoria de dos barcos basándose en datos reales adquiridos con la herramienta Localizatodo:

A. Trayectoria de ida (Figura 7.39):

- Buque de carga AITANA de 27 m de eslora y 7 m de manga.
- Velocidades máxima y media registradas: 10,7 y 9,7 nudos, respectivamente.

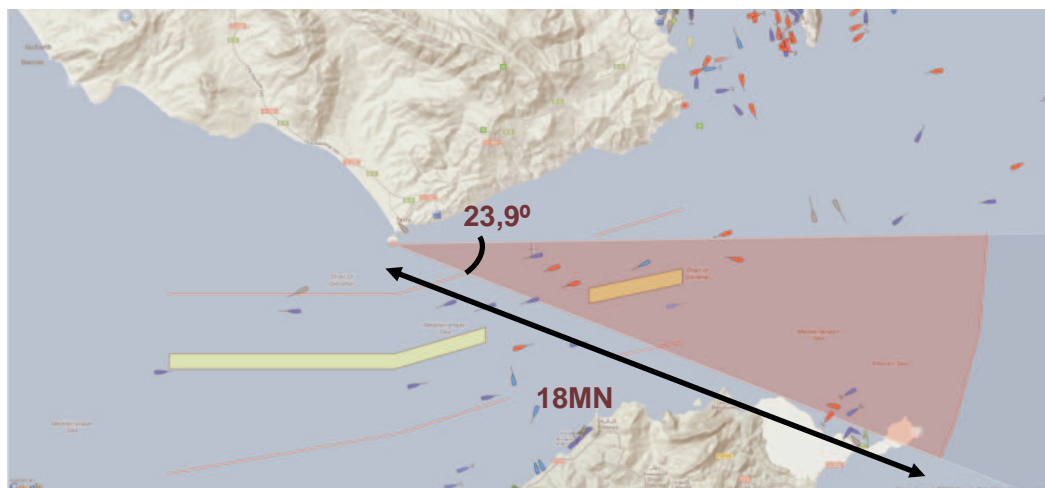


Figura 7.38: Emplazamiento del radar y sector de cobertura simulado.

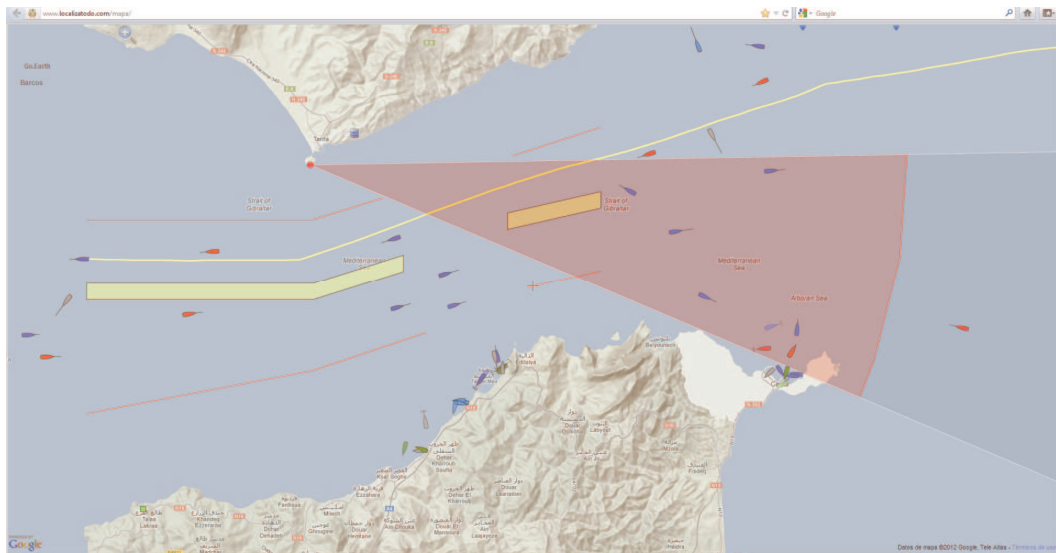
- Movimiento rectilíneo uniforme (MRU) con velocidad constante de 10,7 nudos.
- Intervalo de variación de la frecuencia Doppler de 1,0178 rad a 1,3176 rad.

B. Trayectoria de vuelta (Figura 7.40):

- Yate AZIMUT 53 de 16,6 m de eslora y 4 m de manga.
- Velocidades máxima y media registradas: 33 y 28 nudos, respectivamente.
- Movimiento rectilíneo uniformemente acelerado (MRUA) caracterizado por una velocidad inicial de 25 nudos y una aceleración de 0,058 nudos/s.
- Intervalo de variación de la frecuencia Doppler de 2,3593 rad a 3,54309 rad.

Dado que esta tesis se centra en la detección de blancos puntuales, los barcos considerados no superan en longitud los 30 metros de resolución del radar. Para la simulación de la sección radar de los blancos se ha utilizado el programa POFACETS, desarrollado por *The Naval Postgraduate School* en California. Este software está basado en la aproximación de óptica física e incluye el modelo de un yate escalable. Este modelo escalado a las dimensiones del yate de la trayectoria de vuelta se muestra en la Figura 7.41. Los resultados obtenidos de la simulación de la RCS del yate para los ángulos de incidencia considerados en función del caso de estudio se muestran en la Figura 7.42, para los siguientes parámetros:

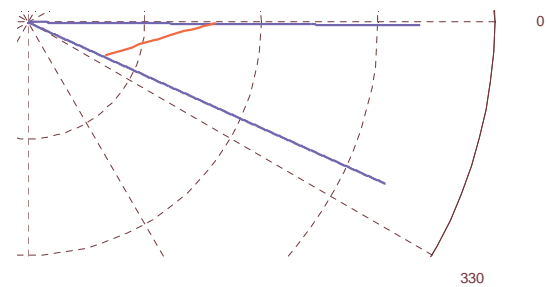
- Frecuencia: 9,39 GHz
- Polarización: horizontal
- Ángulo de elevación: $90-13 = 77^\circ$ para escenarios con clutter gaussiano y $90-0.1^\circ=89,9^\circ$ para escenarios con clutter con distribución K.
- Ángulo de acimut: Variable de 0 a 360°
- Material: VTR/GRP. Glass reinforced polystyrene.



(a) Imagen adquirida de [Localizotodo]

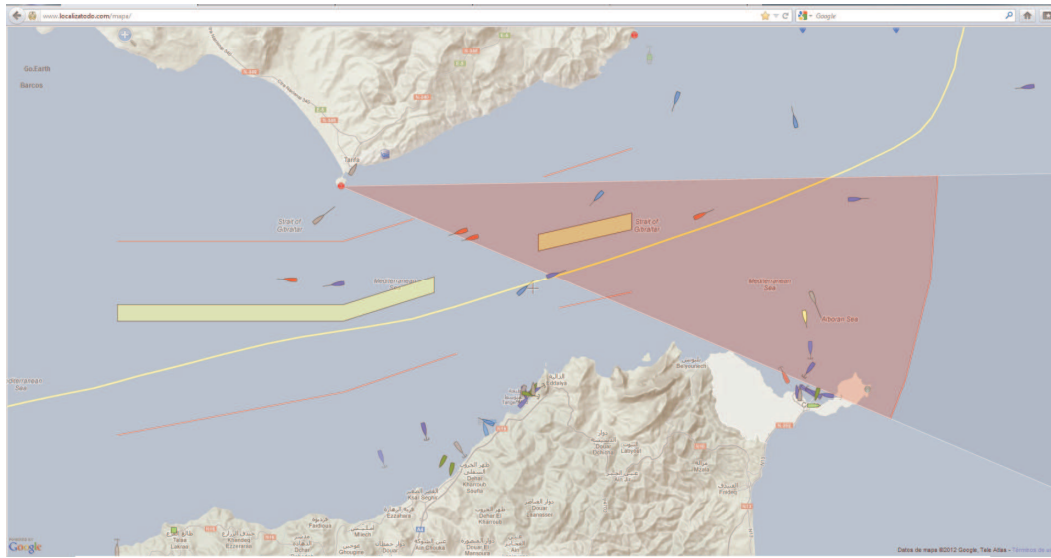


(b) Buque de carga AITANA



(c) Datos importados para la simulación

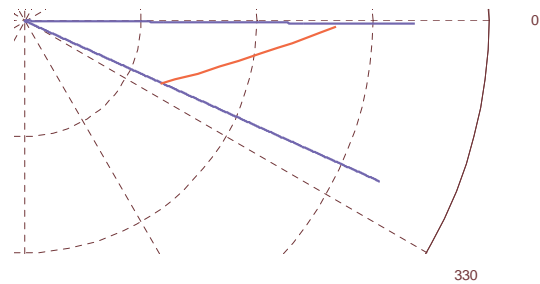
Figura 7.39: Trayectoria de ida: movimiento rectilíneo uniforme con velocidad constante de 10,7 nudos.



(a) Imagen adquirida de [Localizotodo]



(b) Buque de carga AITANA



(c) Datos importados para la simulación

Figura 7.40: Trayectoria de vuelta: movimiento rectilíneo uniformemente acelerado con velocidad inicial de 25 nudos y aceleración de 0,058 nudos/s.

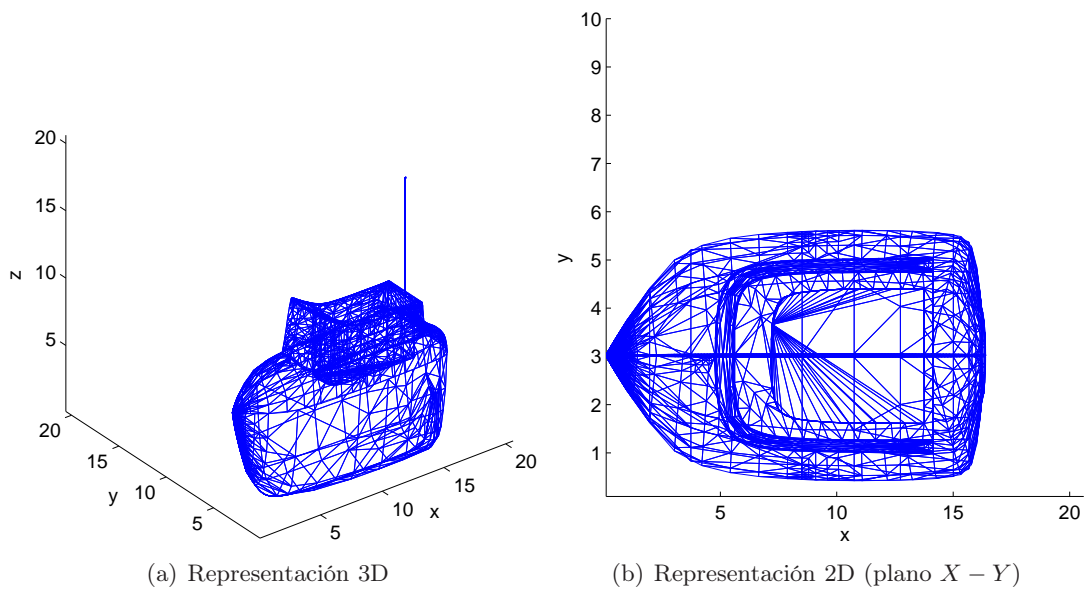


Figura 7.41: Modelo escalado a las dimensiones del yate simulado en la trayectoria de vuelta

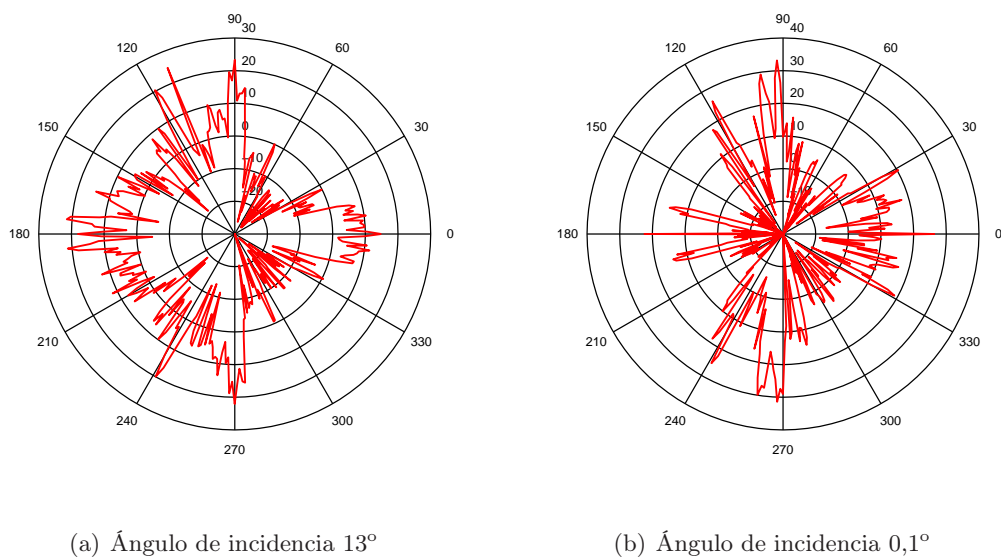
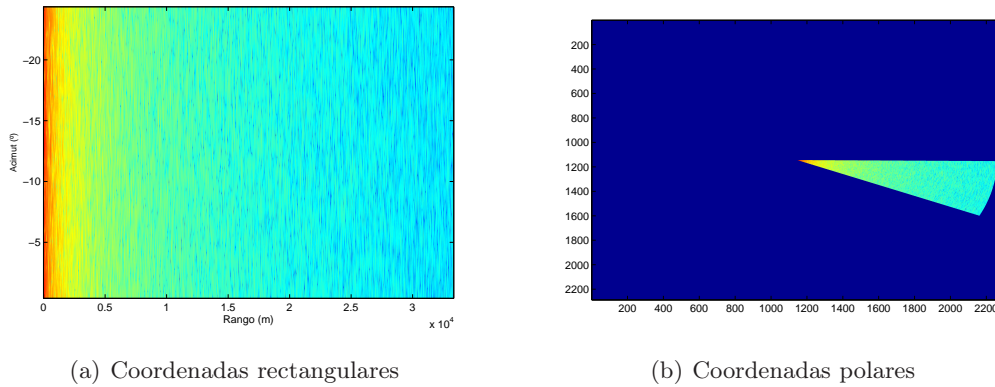


Figura 7.42: RCS en dBsm simulada para el yate simulado en la trayectoria de vuelta



(a) Coordenadas rectangulares

(b) Coordenadas polares

Figura 7.43: Retorno de clutter marino en dBs para clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$.

Tras promediar los valores de RCS obtenidos para todos los valores de acimut considerados, se ha obtenido una RCS media del blanco de $-2,305$ dBsm para un ángulo de incidencia de 13° y $0,226$ dBsm para $0,1^\circ$.

Se han calculado los valores de la potencia recibida de cada uno de los blancos a lo largo de las trayectorias simuladas y las potencias de clutter asociadas a la celda de resolución en la que se encuentran en cada momento. Según el apartado 2.3.3, el eco de clutter superficial recibido depende de la distancia, puesto que el área iluminada es mayor según nos alejamos del radar (2.19). En las Figuras 7.43 y 7.44 se representa, en coordenadas rectangulares (rango, acimut) y polares, el retorno de clutter marino en dBs para los dos ángulo de incidencia considerados. En la Figura 7.43 se representa un clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$, modelo teórico que se ajusta para un ángulo de incidencia de 13° [Skolnik2008]. Para un ángulo de incidencia de $0,1^\circ$ y 30 metros de resolución, el modelo de clutter que se ha considerado es K-distribuido con $\nu = 0,5$ para estados del mar alterados. En la Figura 7.44 se representa un clutter K-distribuido con $\rho_c = 0$ y $\sigma(0) = -48$ dBsm, obtenido en campañas de medidas para un ángulo de incidencia de $0,1^\circ$ descritas [Nathanson1999].

Para conseguir que la potencia de clutter sea constante a lo largo de toda la cobertura del radar, se ha corregido el efecto de la propagación de la onda electromagnética y de la variación de la celda de clutter iluminada con la distancia. El resultado final se presenta en las Figuras 7.45 y 7.46 para cada uno de los tipos de clutter, respectivamente. Esta corrección, favorecerá el rendimiento de los detectores CA-CFAR y neuronales y mantendrá la relación entre la potencia del blanco y la interferencia, por lo que no incluirá artefactos indeseados en las estimaciones de P_{FA} y P_D .

En la Figura 7.47 se presenta la exploración número 20 sumando los retornos de los blancos considerados. Se muestran detalles de las zonas en las que se encuentran los blancos. En estos detalles se ha representado la intensidad en unidades naturales para poder observar los distintos niveles que se reciben de los blancos.

Se ha simulado un intervalo de adquisición de datos radar de 1.810 segundos (905 exploraciones) correspondiente al tiempo en el que al menos una trayectoria es visible en la cobertura de azimut de del radar. Los resultados se presentan en una única gráfica en la que se superponen

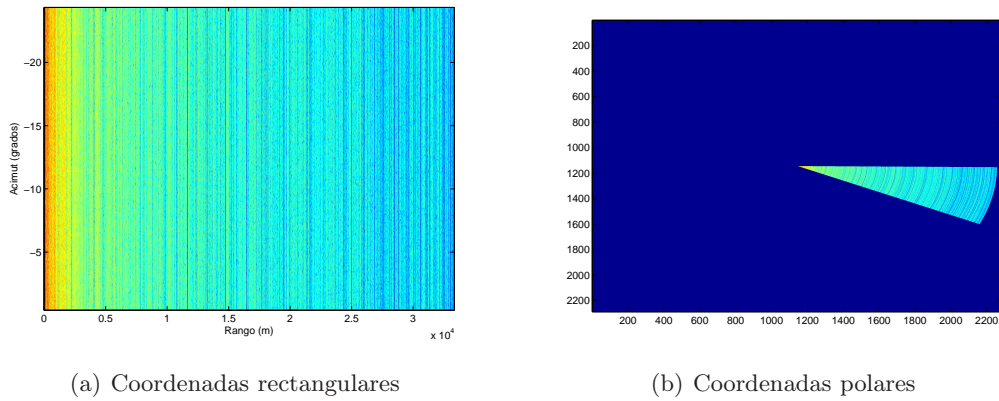


Figura 7.44: Retorno de clutter marino en dBs para clutter K-distribuido con $\rho_c = 0$ y $\sigma^{(0)} = -48$ dBsm.

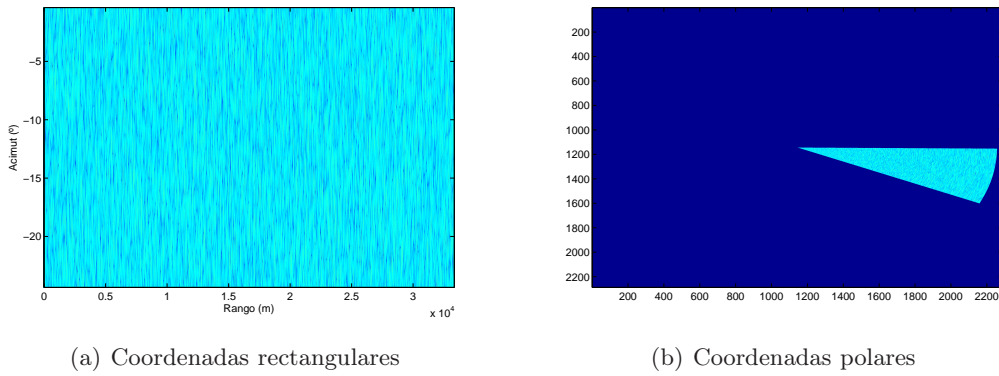


Figura 7.45: Retorno de clutter marino en dBs para clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$ después de corregir los efectos de la propagación y la variación del área de clutter iluminada con la distancia.

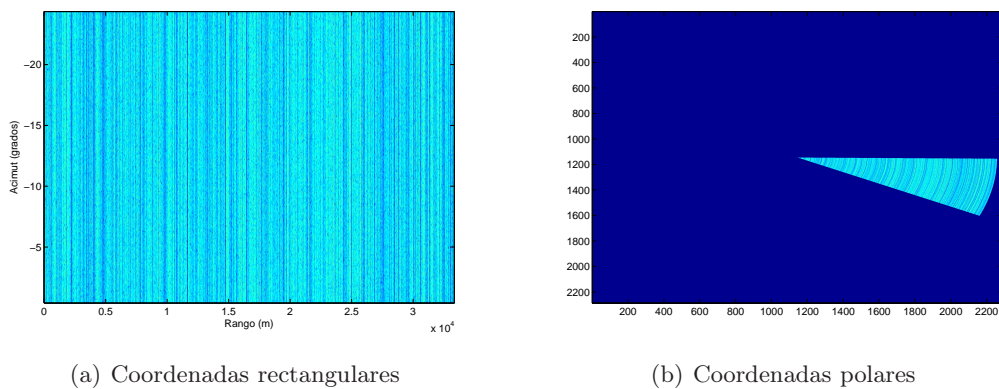
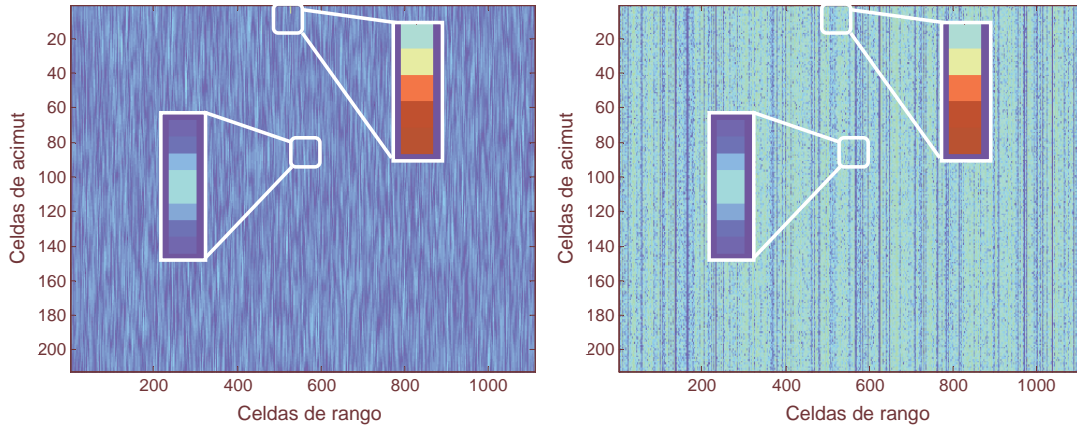


Figura 7.46: Retorno de clutter marino en dBs para clutter K-distribuido con $\rho_c = 0$ y $\sigma^{(0)} = -48$ dBsm después de corregir los efectos de la propagación y la variación del área de clutter iluminada con la distancia.



(a) Blanco más clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$ (b) Blanco más clutter K-distribuido con $\rho_c = 0$

Figura 7.47: Intensidad de las muestras generadas para la exploración 20 (en dB) en la que se muestra el detalle de las zonas en las que se encuentran los dos blancos en unidades naturales.

todos los pulsos detectados y centroides estimados de las 905 exploraciones.

7.8.2. Detectores CA-CFAR

Las técnicas CA-CFAR de umbral adaptativo espaciales, suponen conocida la FDP de la interferencia excepto alguno de sus parámetros y estiman esos parámetros a partir de un promediado de las observaciones del entorno adquiridas en la exploración actual, con el fin de cumplir el objetivo de mantener la P_{FA} constante maximizando la P_D . Al tratarse de técnicas paramétricas, sus capacidades detectoras se deterioran de forma significativa cuando los estadísticos del clutter son diferentes a los asumidos en su diseño.

El esquema básico más extendido de estos detectores, Figura 7.48, utiliza ventanas de referencia que se extienden a lo largo de la dimensión rango de la matriz de entrada, a ambos lados de la CUT (*Cell Under Test*), dejando unas celdas de guarda (Figura 7.49). La salida del detector para la CUT se compara con un umbral que se calcula según (7.3), donde el parámetro T se determina en función del modelo de clutter y la P_{FA} deseada.

$$T_q = T \sum_{i=1}^N q_i \quad (7.3)$$

En un clutter Gaussiano incorrelado, las L celdas de la ventana de referencia del CA-CFAR, a la salida del detector de ley cuadrática, son variables aleatorias exponenciales independientes entre sí y la relación que existe entre T , el parámetro multiplicativo a utilizar en (7.3), y la P_{FA} es la que se indica en (7.4)[Ward2006].

$$T = (P_{FA})^{-\frac{1}{L}} - 1 \quad (7.4)$$

En el caso de estudio 2 se considera un clutter gaussiano que presenta una correlación a lo largo de la dimensión acimut de la matriz de datos de entrada y está incorrelado a lo largo de la

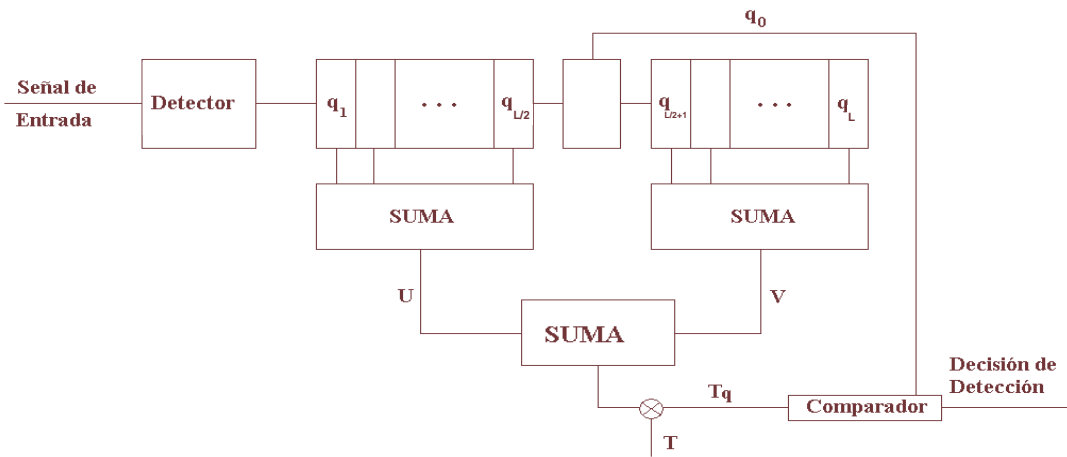


Figura 7.48: Esquema básico de un detector CA-CFAR.

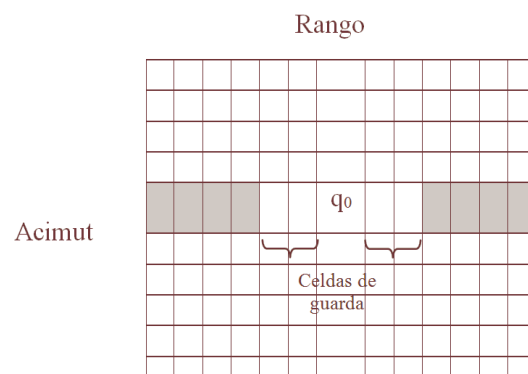


Figura 7.49: Aplicación del CA-CFAR en rango sobre la matriz radar.

dimensión rango. No se ha aplicado integración de pulsos porque, sólo en este caso, se cumplen las hipótesis para las que es válida la expresión (7.4).

La FDP de las muestras de un clutter de distribución K a la salida del detector de envolvente de ley lineal sin integración de pulsos se indica en (7.5)[Ward2006].

$$f(q) = \frac{4q^\nu}{\Gamma(\nu)} \left(\frac{\eta}{\nu}\right)^{\frac{\nu+1}{2}} K_{\nu-1} \left(2q\sqrt{\frac{\eta}{\nu}}\right) \quad (7.5)$$

La FDP del umbral T_q es la suma de N muestras independientes K-distribuidas, por lo que, la P_{FA} se puede calcular a partir de la expresión (7.6). Finalmente, la expresión (7.7) relaciona el valor de T con la P_{FA} deseada.

$$P_{FA} = \int_0^\infty \left(\int_t^\infty f(q) dq \right) P(T_q) dT_q \quad (7.6)$$

$$P_{FA} = \frac{2}{\Gamma(\nu)} (TL)^\nu \left(\frac{\sqrt{\pi}\Gamma(\nu+0,5)}{2\Gamma(\nu)} \right)^\nu K_\nu \left(TL \frac{\sqrt{\pi}\Gamma(\nu+0,5)}{\Gamma(\nu)} \right) \quad (7.7)$$

Para las dos hipótesis se han considerado 2 celdas de rango de guarda y 32 celdas de referencia a cada lado de la CUT ($L = 64$).

Si bien las técnicas CA-CFAR en rango procesan de forma independiente cada fila de la matriz radar, deslizando la ventana de referencia a lo largo de la dimensión rango, Figura 7.49, las soluciones basadas en redes neuronales procesan de forma independiente cada columna de la matriz, deslizando una ventana a lo largo de la dimensión acimut con el fin de integrar la información procedente de P pulsos para tomar la decisión relativa a la celda central de la ventana. Debido a esto, las 8 celdas que en cada momento forman el vector de entrada del detector neuronal no tienen por qué pertenecer todas a la misma hipótesis. Sin embargo, en los conjuntos de entrenamiento se han diseñado de manera que los $P = 8$ pulsos de cada vector de entrada pertenecen a H_0 o a H_1 para las salidas deseadas 0 y 1, respectivamente. Los resultados reflejarán la robustez de la estrategia de diseño de los conjuntos de entrenamiento.

7.8.3. Resultados de detección en el escenario simulado para el caso de estudio 1

Las 905 exploraciones adquiridas por el radar han sido aplicadas a un detector CA-CFAR incoherente sin integración de pulsos y al detector $SONN_1$ 16/152/1 + $SONN_2$ 16/152/1 diseñado en el apartado 7.4, donde se mostraba que este detector neuronal, basado en la mezcla de dos expertos SONNs, es el que mejores resultados proporcionaba para la detección de blancos Swerling I con una Ω variable uniformemente en $[0; 2\pi)$ en AWGN.

Para mantener las P_D de interés en aplicaciones radar, se ha calculado la σ_{RCS} que deberían presentar los blancos para que la SNR a la distancia máxima de cobertura (18 MN) sea 10 dB, resultado del estudio de las curvas ROC del CGLR en el apartado 5.3.1 (Figura 5.12(b)). El valor de σ_{RCS} calculado es -9,01 dBsm, un valor mucho más pequeño que los valores medios de RCS obtenidos con el programa POFACETS para cualquiera de los ángulos de incidencia considerados ($\sigma_{RCS} = -2,305$ dBsm para un ángulo de 13° y $\sigma_{RCS} = 0,226$ dBsm para $0,1^\circ$), por lo que independientemente del ángulo de incidencia, los blancos serán detectados sin problemas en el área de cobertura.

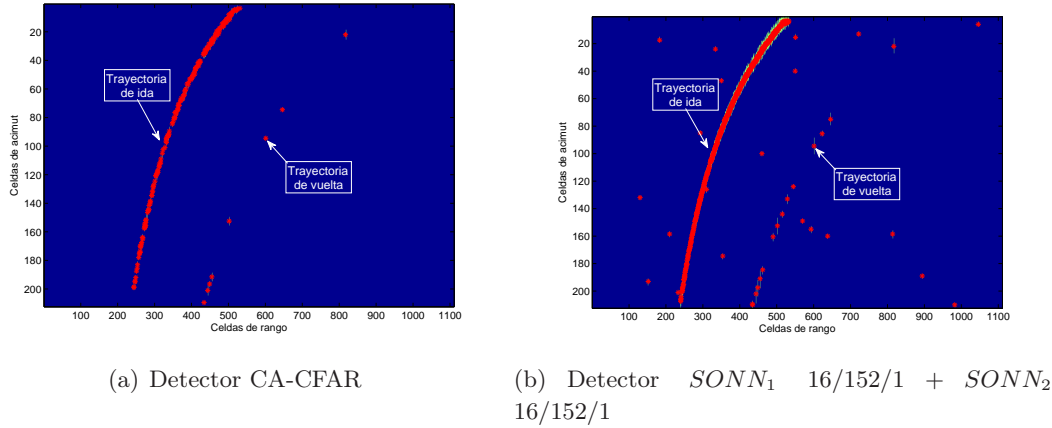


Figura 7.50: Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 1.

	Detector neuronal
$CNR = 20$ dB, $\rho_c = 0,7$ y $SIR = 13$ dB	$SONN_1$ 16/152/1 + $SONN_2$ 16/152/1
$CNR = 20$ dB, $\rho_c = 0,995$ y $SIR = 1$ dB	$SONN_1$ 16/152/1 + $SONN_2$ 16/152/1
$CNR = 40$ dB, $\rho_c = 0,995$ y $SIR = -14$ dB	$SONN$ 16/152/1

Tabla 7.14: Detectores neuronales propuestos para el caso de estudio 2 en función del tipo de clutter gaussiano.

En las Figuras 7.50(a) y 7.50(b) se representa la superposición de los pulsos detectados y centroides estimados tras aplicar detectores CA-CFAR y $SONN_1$ 16/152/1 + $SONN_2$ 16/152/1, respectivamente, y la etapa de reducción de falsas alarmas. El esquema basado en el detector neuronal presenta menos pérdidas de detección en la trayectoria de vuelta aunque el número de falsas alarmas finales es mayor. Debe matizarse que el carácter disperso de estas falsas alarmas a lo largo de exploraciones sucesivas, facilitará su eliminación en la etapa de seguimiento de los blancos.

7.8.4. Resultados de detección en el escenario simulado para el caso de estudio 2

En este caso se plantea la detección de blancos Swerling I con una Ω variable uniformemente en $[0, 2\pi)$ en interferencia gaussiana, compuesta por AWGN y clutter gaussiano. En la Tabla 7.14 se resumen los esquemas de detección que finalmente se propusieron en el apartado 7.6 para los distintos tipos de clutter gaussiano considerados. En esta Tabla se han incluido los valores de SIR necesarios para $P_D > 80\%$ con $P_{FA} = 10^{-4}$, según los resultados mostrados en la Figura 5.14, donde se presentan las curvas ROC estimadas para los CGLRs considerados en cada caso.

Según el programa POFACETS, para un barco de unas características similares al yate Azimut 53 (Figura 7.40(b)) visto con un ángulo de elevación de 77° (correspondientes a un ángulo de incidencia de 13°) se estima una RCS media de $-2,305$ dBsm. Con esta RCS y los parámetros de clutter considerados, se han calculado las coberturas esperadas para los valores de SIR d ela

	Distancia de cobertura
$CNR = 20$ dB, $\rho_c = 0,7$ y $SIR = 13$ dB	13,02 km
$CNR = 20$ dB, $\rho_c = 0,995$ y $SIR = 1$ dB	25,98 km
$CNR = 40$ dB, $\rho_c = 0,995$ y $SIR = -14$ dB	19,53 km

Tabla 7.15: Distancias de cobertura para un blanco con $\sigma_{RCS} = -2,305$ dBsm.

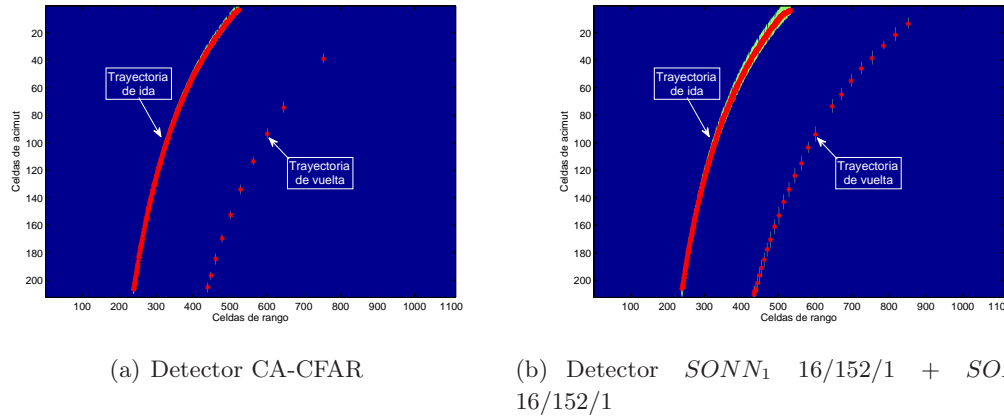


Figura 7.51: Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 2 con $CNR = 20$ y $\rho_c = 0,7$.

Tabla 7.14. Los resultados obtenidos se muestran en la Tabla 7.15. Dado que en ningún caso se cumple el requisito de cobertura igual a 18 MN (33,34 km), se ha modificado la RCS del blanco para garantizar en todos los casos los valores de SIR de la Tabla 7.14 en los puntos más alejados de la trayectoria.

Los resultados de aplicar el detector CA-CFAR incoherente sin integración de pulsos y los detectores neuronales asociados a cada tipo de clutter se presentan en las Figuras 7.51, 7.52 y 7.53. Los resultados obtenidos con detectores neuronales son mejores que los correspondientes a detectores CA-CFAR, independientemente del tipo de clutter gaussiano considerado. En las Figuras 7.51(b), 7.52(b) y 7.53(b), se demuestra que los detectores neuronales no presentan casi pérdidas de detección ni falsas alarmas. Además, según el clutter se hace más correlado y más dominante, las soluciones basadas en redes neuronales siguen manteniendo su capacidad de detección a pesar de que los correspondientes valores de SIR sean más pequeños. En las Figuras 7.52(a) y 7.53(a), se observa cómo el detector CA-CFAR no es capaz de detectar ninguna de las dos trayectorias.

7.8.5. Resultados de detección en el escenario simulado para el caso de estudio 3

Según el ángulo de incidencia se hace más pequeño y el estado del mar es más alterado, hay que asumir un modelo de clutter más impulsivo. En el apartado 7.7, se propone el diseño de soluciones basadas en inteligencia artificial para detectar blancos Swerling V dentro de un margen de pulsaciones Doppler restringido $[1, 2566; 1, 8850]$ radianes en clutter K-distribuido. El

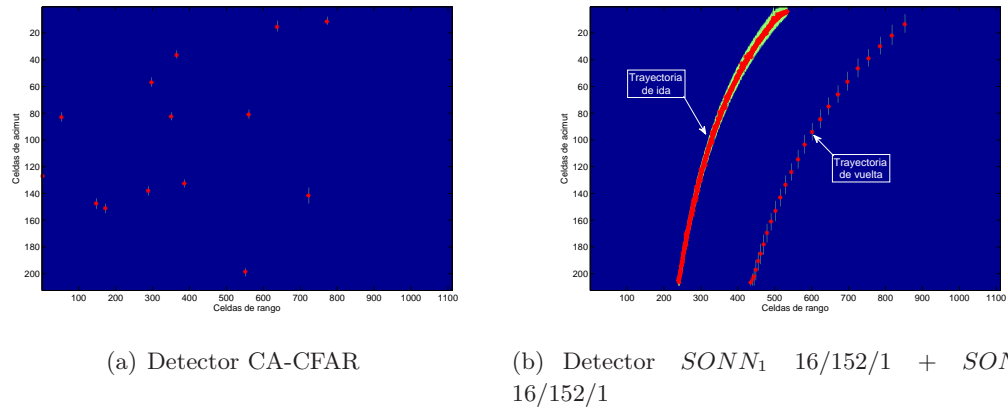


Figura 7.52: Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 2 con $CNR = 20$ y $\rho_c = 0,995$.

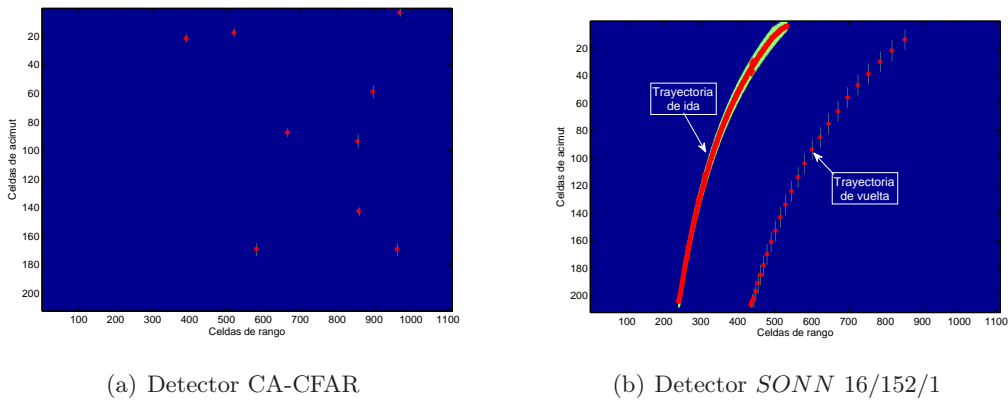


Figura 7.53: Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 2 con $CNR = 40$ y $\rho_c = 0,995$.

	Detector neuronal
$\rho_c = 0$ y $SCR = 9$ dB	<i>RBFNN</i> 16/7/1
$\rho_c = 0,9$ y $SCR = -3$ dB	<i>MLP</i> 16/17/1

Tabla 7.16: Detectores neuronales propuestos para el caso de estudio 3 en función de ρ_c .

	Distancia de cobertura
$\rho_c = 0$ y $SCR = 9$ dB	17,75 km
$\rho_c = 0,9$ y $SIR = -3$ dB	281,44 km

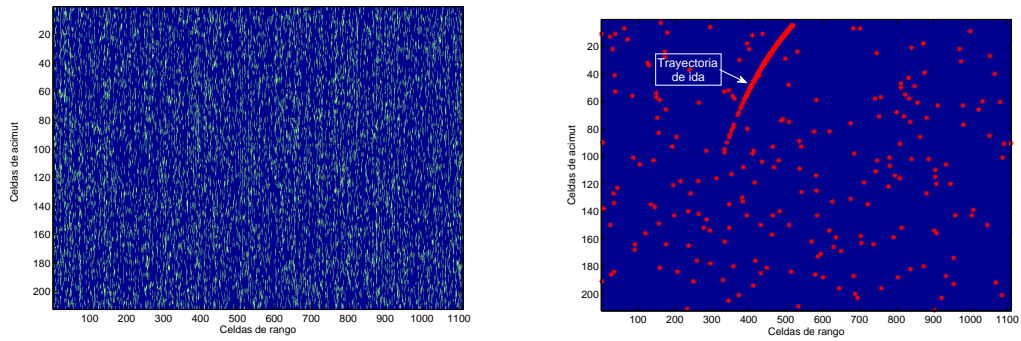
Tabla 7.17: Distancias de cobertura para un blanco con $\sigma_{RCS} = 0,226$ dBsm.

intervalo de Ω de entrenamiento coincide con parte de las velocidades radiales de la trayectoria de ida ($\Omega \in [1,0178, 1,3176]$), mientras que las velocidades de la trayectoria de vuelta son mayores ($\Omega \in [2,3593; 3,54309]$), por lo que los ecos recibidos del yate Azimut 53 no serán detectados por los detectores neuronales. En la Tabla 7.16 se presentan los detectores neuronales propuestos en el apartado 7.7 para los dos valores de coeficiente de correlación del clutter. En esta Tabla también se han incluido los valores de SCR que, según la Figura 5.17, requieren los correspondientes CGLRs para proporcionar $P_D > 80\%$ con $P_{FA} = 10^{-4}$.

Utilizando el programa POFACETS para un ángulo de incidencia de $0,1^\circ$, se ha estimado una RCS media de un barco de unas características similares al yate Azimut 53 (Figura 7.40(b)) de $0,226$ dBsm. Las coberturas estimadas para esta RCS se indican en la Tabla 7.17. Se observa que para $\rho_c = 0$ no se cumple el requisito de cobertura 18 MN (33,34 km), pero para $\rho_c = 0,9$ se supera por un amplio margen, por lo que es de esperar que pueda aumentarse el margen de variación de Ω a expensas de una reducción en la cobertura. Debido a esto, se han calculado las RCSs de los blancos necesarias para que la SCR a 18 MN sea la indicada en la Tabla 7.16.

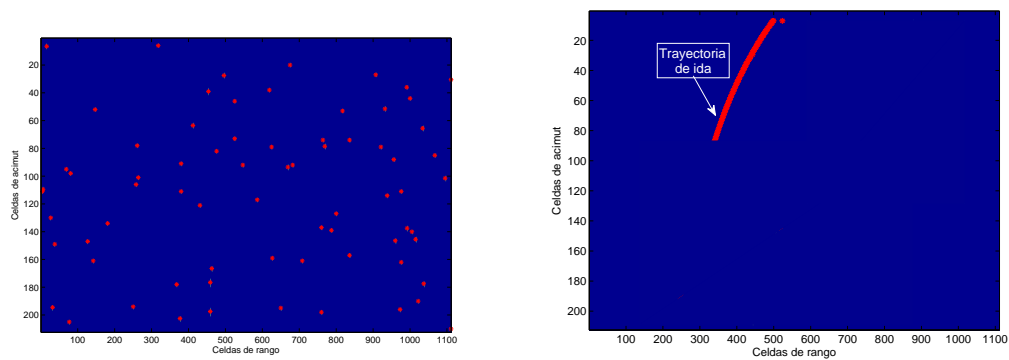
Los resultados obtenidos con los detectores neuronales son claramente mejores que los correspondientes al detector CA-CFAR que asume variables aleatorias independientes K-distribuidas, Figuras 7.54 y 7.55. En las Figuras 7.54(a) y 7.55(a) el número de falsas alarmas resultado de aplicar el detector CA-CFAR es muy elevado y no se puede identificar ninguna de las trayectorias de los barcos. Para los detectores RBFNN 16/7/1 (Figura 7.54(b)) y MLP 16/17/1 (Figura 7.55(b)), aunque también aparecen falsas alarmas, se distingue claramente el trozo de la trayectoria de ida con velocidades radiales dentro del margen de Ω asumido en el entrenamiento. En la Figura 7.54(b), no se ha aplicado la etapa de reducción de falsas alarmas. El detector RBFNN 16/7/1 sólo detecta un pulso de los 8 pertenecientes al blanco, por lo que si se aplicase la etapa de reducción de falsas alarmas, que elimina aquellos clusters que no tengan al menos $P/2$ celdas en acimut, nos se detectaría ningún cluster.

Una de las razones en las que se basa la mejora en detección de los detectores neuronales, es la especialización de éstos en una margen restringido de frecuencias Doppler. En el apartado 5.3.3, donde se estudiaban los detectores CGLR para este caso de estudio, ya se manifestaba que el diseño de un detector para blancos con $\Omega \in [0; 2\pi)$, especialmente en presencia de un clutter correlado, conllevaba una SCR muy elevada para obtener P_D de interés en aplicaciones radar, es decir, radios de cobertura muy pequeños.



(a) Detector CA-CFAR (sin centroides estimados) (b) Detector *RBFNN* 16/7/1 (con centroides estimados)

Figura 7.54: Superposición de pulsos detectados de las 905 exploraciones sin etapa de reducción de falsas alarmas para el caso de estudio 3 con $\rho_c = 0$.



(a) Detector CA-CFAR

(b) Detector *MLP* 16/17/1

Figura 7.55: Superposición de pulsos detectados y centroides estimados de las 905 exploraciones con etapa de reducción de falsas alarmas para el caso de estudio 3 con $\rho_c = 0,9$.

Caso de estudio		Detector neuronal	Nº operaciones
1 (AWGN)		RBFNN 16/5/1	256
2 (Clutter+AWGN)	$CNR = 20$ dB y $\rho_c = 0,7$	SONN 16/152/1	441
	$CNR = 20$ dB y $\rho_c = 0,995$	SONN 16/152/1	441
	$CNR = 40$ dB y $\rho_c = 0,995$	SONN 16/152/1	441

Tabla 7.18: Detectores neuronales propuestos para la detección de blanco gaussianos con $\rho_s \in [0; 1]$ en interferencia gaussiana.

7.9. Conclusiones

En este capítulo se han diseñado diversas soluciones basadas en técnicas de inteligencia artificial para cada caso de estudio planteado. Se ha realizado un análisis de sus prestaciones mediante la estimación de sus curvas ROC, tomando como referencia las soluciones basadas en el CGLR que se propusieron en la sección 5.3. Para ello se han generado bases de datos sintéticos en los que cada vector de entrada pertenece a una de las hipótesis, H_1 o H_0 , en función de la salida deseada asociada, 1 o 0, respectivamente. Finalmente, se han estudiado las capacidades de detección de los detectores neuronales propuestos en un escenario radar simulado en el que se modelan fenómenos como la pérdida de propagación, la variación del tamaño del área del clutter iluminada con la distancia y las variaciones Doppler asociadas a trayectorias de blancos reales.

A continuación se detallan las conclusiones y aportaciones más relevantes relacionadas con el diseño de detectores basados en NNs:

- **Detección de blancos gaussianos con coeficiente de correlación desconocido** ($\rho_s \in [0; 1]$) **en presencia de interferencia gaussiana.** En la Tabla 7.18 se resumen los detectores neuronales que mejor compromiso presentan entre capacidades detectoras y coste computacional asociado. Estos detectores han sido entrenados con patrones bajo H_1 con una ρ_s variable uniformemente en $[0; 1]$.

En general, las soluciones basadas en una única SONN presentan muy buen rendimiento independientemente de tipo de interferencia gaussiana. Sólo en el caso de estudio 1, detección de blancos en AWGN, un detector RBFNN 16/5/1 es capaz de obtener una buena aproximación con un coste computacional más reducido.

- **Detección de blancos gaussianos con frecuencia Doppler desconocida** ($\Omega \in [0; 2\pi)$) **en presencia de interferencia gaussiana.** En la Tabla 7.19 se resumen los detectores neuronales que mejor compromiso presentan entre capacidades detectoras y coste computacional asociado.

Las soluciones propuestas se basan en mezclas de dos SONNs especializadas en intervalos de variación de Ω restringidos ($SONN_1$ especializada en $\Omega^t \in [0; \pi/4) \cup [3\pi/4; 2\pi)$ y $SONN_2$ en $\Omega^t \in [\pi/4; 3\pi/4)$), excepto para el caso de interferencia gaussiana compuesta por AWGN más clutter gaussiano con $CNR = 40$ dB y $\rho_c = 0,995$, donde la mejora introducida por el comité de SONNs no es comparable con el aumento de complejidad computacional asociado, y se ha optado por una solución basada por una única SONN entrenada con patrones bajo H_1 con una Ω variable uniformemente en $[0; 2\pi)$.

Caso de estudio		Detector neuronal	Nº operaciones
1 (AWGN)		$SONN_1$ 16/5/1 + $SONN_2$ 16/5/1	882
2 (Clutter + AWGN)	$CNR = 20$ dB $\rho_c = 0,7$	$SONN_1$ 16/5/1 + $SONN_2$ 16/5/1	882
	$CNR = 20$ dB $\rho_c = 0,995$	$SONN_1$ 16/5/1 + $SONN_2$ 16/5/1	882
	$CNR = 40$ dB $\rho_c = 0,995$	$SONN$ 16/5/1	441

Tabla 7.19: Detectores neuronales propuestos para la detección de blanco gaussianos con $\Omega \in [0; 2\pi)$ en interferencia gaussiana.

■ **Detección de blancos Swerling V con frecuencia Doppler desconocida ($\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10)$) en clutter K-distribuido con $\nu = 0,5$.**

- Después de estudiar el detector CGLR para este caso de estudio, se ha optado por restringir el intervalo de variación de Ω a $[\pi/2 - \pi/10; \pi/2 + \pi/10]$ para obtener P_D de interés para $P_{FA} \leq 10^{-4}$ sin necesitar valores de SCR muy elevados.
- Se ha realizado un estudio de las fronteras de decisión formadas por los detectores CGLR para vectores de entrada compuestos por las componentes en fase y cuadratura de un pulso, obteniéndose arcos de una corona circular cuyo radio depende de la potencia del blanco. Las fases son el resultado de sumar la fase del blanco asumida en el diseño y la frecuencia Doppler uniformemente distribuida en $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$. Cuando el vector de entrada está formado por $P = 8$, se ha comprobado que un corte bidimensional de las hipersuperficies de separación también corresponde con un arco de corona, por lo que se podría deducir, que las superficies de decisión que se formarán en un espacio de $2P$ dimensiones posiblemente se definan con sectores de dos hiperesferas concéntricas.
- En la Tabla 7.20 se presentan los detectores neuronales con mejor compromiso entre capacidades detectoras y coste computacional para la detección de blancos Swerling V en clutter K distribuido. Estos detectores han sido entrenados con patrones bajo H_1 con una Ω variable uniformemente en $[\pi/2 - \pi/10; \pi/2 + \pi/10]$.

Para este caso de estudio, no se han obtenido buenos resultados con soluciones basadas SONNs con una única QNU. No tiene sentido aumentar el número de QNUs, porque aumentaría considerablemente la complejidad computacional y los detectores RBFNN 16/7/1, para $\rho_c = 0$, y MLP 16/17/1, para $\rho_c = 0,9$, presentan un buen rendimiento controlando el número de operaciones requerido.

Debido a las conclusiones extraídas en la sección 6.3, en relación a las capacidades de soluciones basadas en C-SVMs y 2C-SVMs, finalmente se han diseñado detectores basados en 2C-SVMs, realizando búsquedas en rejilla que permitan maximizar la P_D para una $P_{FA} \leq 10^{-4}$. Las principales conclusiones y aportaciones relacionadas con este tipo de soluciones son las que se resumen a continuación:

Caso de estudio		Detector neuronal	Nº operaciones
3 (Clutter)	$\rho_c = 0$	RBFNN 16/7/1	358
	$\rho_c = 0,9$	MLP 16/17/1	597

Tabla 7.20: Detectores neuronales propuestos para la detección de blanco Swerling V con $\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$ en clutter de distribución K.

Caso de estudio	Parámetro desconocido	C	γ	Nº operaciones
1 (AWGN)	$\rho_s \in [0; 1]$	10^3	0,7	21.172
	$\Omega \in [0; 2\pi/2)$	10^3	0,7	25.996
2 (AWGN+Clutter) $CNR = 20$ dB; $\rho_c = 0,7$	$\rho_s \in [0; 1]$	10^5	0,95	27.671
	$\Omega \in [0; 2\pi/2)$ (Swerling I)	10^3	0,99	8.911
2 (AWGN+Clutter) $CNR = 20$ dB; $\rho_c = 0,995$	$\rho_s \in [0; 1]$	10^3	0,7	13.601
	$\Omega \in [0; 2\pi/2)$ (Swerling I)	10^3	0,9	13.735
2 (AWGN+Clutter) $CNR = 40$ dB; $\rho_c = 0,995$	$\rho_s \in [0; 1]$	10^3	0,7	9.648
	$\Omega \in [0; 2\pi/2)$ (Swerling I)	10^3	0,9	18.358
3 ($\rho_c = 0$) Clutter K-dist + Swerling V	$\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$	10^5	0,999	66.665
3 ($\rho_c = 0,9$) Clutter K-dist + Swerling V	$\Omega \in [\pi/2 - \pi/10; \pi/2 + \pi/10]$	10	0,999	37.386

Tabla 7.21: Resumen de las soluciones basadas en 2C-SVMs para todos los casos de estudio.

- En la Tabla 7.21 se resumen las soluciones basadas en 2C-SVMs para todos los casos de estudio. En esta Tabla se reflejan la pareja de parámetros C y γ que permiten obtener una $P_{FA} \leq 10^{-4}$ y proporcionan la mayor P_D después de haber realizado una búsqueda en rejilla. También se presenta el coste computacional asociado a cada solución.
- La aproximación al punto de la curva ROC del detector óptimo asociado a $P_{FA} = 10^{-4}$ depende del caso de estudio y del tamaño del conjunto de entrenamiento (se ha utilizado un conjunto de entrenamiento compuesto por 500 patrones pertenecientes a las hipótesis H_1 y H_0 con la misma probabilidad). Hay ciertos valores de C para los que se han obtenido $P_{FA} = P_D = 0$, es decir, las 2C-SVMs no han convergido a una solución, debido a que el número de patrones requería de más generalización proporcionada por valores de C más pequeños, asociados a parámetros de regularización mayores.
- Cuanto mayor es el valor C , mejor es la aproximación y mayor es el coste computacional asociado. Los detectores 2C-SVMs con capaces de aproximar el punto deseado de la curva ROC con un error asociado aceptable, excepto para la detección de blancos gaussianos con Ω desconocida en interferencia gaussiana. En este caso, el error de aproximación puede deberse a que las funciones de Green de tipo cuadrático no son adecuadas.
- La principal desventaja que ofrecen los detectores basados en SVMs es que necesitan muchas más operaciones que las soluciones basadas en NNs.

En relación a la evaluación de las soluciones propuestas en un escenario radar simulado:

- Se ha simulado un escenario radar con dos trayectorias reales extraídas de [Localizado], correspondientes a dos barcos con una velocidad radial variable debido a su movimiento y posición relativos a un hipotético radar situado en el Estrecho de Gibraltar. Se han considerado los parámetros del radar IPIX [Drosopoulos1994].

En realidad, asumiendo un mismo radar, un mismo emplazamiento y las dos mismas trayectorias de los barcos, se han simulado dos escenarios radar:

- Clutter gaussiano correspondiente, en sistemas de alta resolución, a un ángulo de incidencia mayor de 10° .
 - Clutter K-distribuido con $\nu = 0,5$ correspondiente, en sistemas de alta resolución, a un ángulo de incidencia menores de 10° y estados del mar alteraros mayores de 1.
- Se han aplicado detectores CA-CFAR con el objetivo de realizar estudios comparativos, los cuales han permitido concluir que los resultados obtenidos al aplicar los detectores neuronales son mejores que los correspondientes a detectores CA-CFAR, independientemente del caso de estudio considerado.
 - En el caso de la detección de blancos en presencia de clutter Gaussiano, los detectores neuronales ofrecen grandes capacidades de detección para valores de SIR cada vez más pequeños según el clutter es más correlado y más dominante. Sin embargo, los detectores CA-CFAR necesitan más SIR para ofrecer prestaciones similares.
 - En presencia de clutter de distribución K con $\nu = 0,5$, el número de falsas alarmas asociado a los detectores CA-CFAR es tan elevado que no se pueden detectar las trayectorias de los blancos. Al reducir el margen de variación de Ω para entrenar los detectores neuronales, se consiguen mejores resultados, identificando la parte de la trayectoria de ida cuyas velocidades radiales coinciden con el intervalo de variación de Ω asumido en el entrenamiento.
 - En la Tabla 7.22 se presentan los radios de cobertura estimadas para mantener $P_D > 80\%$ para $P_{FA} = 10^{-4}$ para cada tipo de clutter, junto con la RCS del blanco estimada con el programa POFACETS para cada ángulo de incidencia. En general, las coberturas aumentan en presencia de un clutter correlado.

Por último, hay que destacar la robustez de la estrategia de entrenamiento donde todos los patrones bajo H_0 son de clutter y/o ruido y bajo H_1 son blanco más clutter y/o ruido. A la hora de aplicar las soluciones basadas en redes neuronales, se han procesado las distintas columnas de la matriz asociada a cada exploración de forma independiente, desplazando una ventana de 8 celdas de acimut consecutivas a lo largo de la dimensión de acimut, por lo que los vectores de entrada pueden contener muestras generadas bajo H_0 o H_1 en distintas proporciones. Aún así, los buenos resultados de detección validan la estrategia seguida para el diseño de conjuntos de entrenamiento, validación y test.

Ángulo	RCS	Caso	Relación	Cobertura
13° Clutter Gaussiano	-2,305 dBsm	$CNR = 20$ dB $\rho_c = 0,7$	$SIR = 13$ dB	13,02 km
		$CNR = 20$ dB $\rho_c = 0,995$	$SIR = 1$ dB	25,98 km
		$CNR = 40$ dB $\rho_c = 0,995$	$SIR = -14$ dB	19,53 km
0,1° Clutter K-distribuido	0,226 dBsm	$\rho_c = 0$	$SCR = 9$ dB	17,75 km
		$\rho_c = 0,9$	$SCR = -3$ dB	281,44 km

Tabla 7.22: Distancias de cobertura para las RCS medias obtenidas por la herramienta POFA-CETS para distintos ángulos de incidencia.

Parte III

Conclusions and Future Research Lines

Capítulo 8

Conclusions and Future Research Lines

8.1. Conclusions

This PhD. Thesis tackles the problem of the study of the suitability of Artificial Intelligence (AI) based radar detectors as solutions for the problem of detecting targets with unknown parameters in different radar clutter environments. Learning machines trained in a supervised manner using a suitable error function, were proved to be able to approximate the optimum Neyman-Pearson detector in simple hypothesis tests, where the likelihood functions of the two possible hypothesis (H_0 , target absent, and H_1 , target present) are completely known [Jarabo2009]. In practical situations, the statistical properties of the clutter can be estimated and tracked in some degree, but the parameters of the targets are very difficult to estimate. The sensitivity of the Likelihood Ratio Test (LRT) detectors to target parameters such as the correlation coefficient and the Doppler shift can be very high, giving rise to high detection losses when the values assumed in the LRT design differ from the real target ones. If the unknown parameters are random variables of known probability density function, the decision rule based on the Average Likelihood Ratio (ALR) test is an implementation of the Neyman-Pearson detector. In many situations, this solution leads to intractable integrals that can be approximated using complex mathematical approximations or alternative solutions based on the Generalized Likelihood Ratio (GLR) or constrained versions (CGLR). In this PhD. Thesis, AI based detectors are proposed for approximating the ALR detector controlling the computational cost. The work carried has been structured in the following way:

- First of all, the theoretical study presented in [Jarabo2009] has been extended to the composite hypothesis test problem (detecting targets of unknown parameters in presence of radar clutter), confirming that the sufficient condition proposed in [Jarabo2009] can be applied for testing if an error function is suitable for training learning machines in a supervised manner in order to approximate the ALR based detector for any pair of likelihood functions.
- For the analysis of the suitability of an error function for training a learning machine to approximate the Neyman-Pearson detector, the calculus of the function approximated by the learning machine after training is a key element. In [Jarabo2005a, Jarabo2009] the function approximated by a learning system such as a MultiLayer Perceptron (MLP),

a Radial Basis Function Neural Network (RBFNN) or a Second Order Neural Network (SONN) when trained using the mean-squared error and the cross-entropy, were calculated. An important contribution of this PhD. Thesis, is the theoretical study of the function approximated by a Support Vector Machine (SVM), when the error used for training is the classification error proposed by [Shawe2004]. This analysis has provided important keys to explain, from theoretical foundations, the observed performances and limitations of C-SVMs and 2C-SVMs in different detection problems presented in the literature.

- As this PhD. Thesis has been carried out in the scope of research projects funded by the Spanish Ministry, the *Comunidad de Madrid*, the University of Alcalá, and the company AMPER SISTEMAS, S.A., which focused on marine radar applications, radar clutter models proposed for sea clutter have been studied and used for generating synthetic data sets for training, validating and testing the proposed AI solutions, and for simulating a radar scenario. Three cases of study have been considered taking into account different clutter models and target unknown parameters:
 - **Case of study 1:** Detection of Gaussian fluctuating targets in presence of Additive White Gaussian Noise (AWGN). This is considered as a reference case, corresponding to target detection in clear air conditions. This case of study has been subdivided into two:
 - Detection of Gaussian targets with unknown correlation coefficient.
 - Detection of Swerling I (SWI) targets with unknown Doppler shift.
 - **Case of study 2:** Detection of Gaussian fluctuating targets in presence of correlated Gaussian clutter and AWGN. This model can be used for target detection in AWGN and sea clutter with low resolution radar systems or high resolution radar systems with incidence angle higher than 10° . Again, this case of study has been subdivided into two:
 - Detection of Gaussian targets with unknown correlation coefficient.
 - Detection of SWI targets with unknown Doppler shift.
 - **Case of study 3:** Detection of non fluctuating target in presence of spiky K-distributed clutter. This model is suitable for target detection in sea clutter with high resolution radar systems and low grazing angles in sea states higher than 2. In this case, the problem of detecting Swerling V (SWV) targets with unknown Doppler shift is considered.
- For all the aforementioned cases of study, an analysis of the sensitivity of the LRT detectors for targets of known parameters was carried out in a first step, in order to study the detection losses associated to a mismatch between the assumed design target parameters and the actual ones. After that, the ALR detector was formulated and sub-optimum approaches based on the CGLR were designed as reference solutions for the evaluation of the proposed AI based ones.
- For each case of study, detectors based on MLPs, RBFNNs, SONNs and SVMs have been designed and tested. Another important contribution of the PhD Thesis is the use of

SONNs with only one quadratic neural unit, which present a high robustness against target correlation coefficient and Doppler shift, in Gaussian interference. Mixture of experts are also designed for improving the detection capabilities and/or reducing the computational cost. Different combination techniques have been considered. Some of them, as far as the author knows, are novel contributions. A final AI based solution has been proposed as a compromise between detection performance and computational cost. These detectors have been finally evaluated in a simulated radar scenario, comparing their performances with those provided with CA-CFARs.

Following the previous structure, the main conclusions are summarized, emphasizing the more relevant contributions.

8.1.1. Regarding the Capability of Intelligent Agents for Approximating the Neyman-Pearson Detector in Composite Hypothesis Test

The sufficient condition proposed in [Jarabo2009] has been extended to the analysis of the capabilities of learning systems trained in a supervised manner to approximate the ALR in composite hypothesis test. MLPs, RBFNNs and SONNs are usually trained using error functions such as the mean-squared error or the cross-entropy, for which the function approximated after training has been calculated. MLPs have been trained using the cross-entropy error function and SONNs and RBFNNs with the mean-squared error. Although in all cases the sufficient condition is fulfilled, the final approximation error depends on the way the neural network builds the decision boundaries, the number of degrees of freedom, the training sets and the learning algorithm. Because of that, different architectures are considered attending to the problem that must be solved.

The SVMs are trained using the classification or the ϵ -insensitive cost functions [Shawe2004, Vapnik1995], for which the function approximated by the SVM has not been calculated, so far. In this PhD Thesis, the function approximated by a C-SVM trained using an infinite training set ($N \rightarrow \infty$) and allowing the regularization parameter, C , tend to ∞ , has been calculated:

- C-SVM are trained in a supervised manner to minimize the classification error [Shawe2004]:

$$E_m = \int_{\mathbf{z}} (P(H_0)u(1 + F(\mathbf{z}))f(\mathbf{z}|H_0) + P(H_1)u(1 - F(\mathbf{z}))f(\mathbf{z}|H_1)) d\mathbf{z}$$

In order to apply calculus of variations to determine the function $F(\mathbf{z})$ that minimizes E_m , the Heaviside step function has been approximated by a logistic function for a fixed value of the parameter k , that controls the sigmoid function slope in the lineal zone. The function approximated by the SVM has been calculated as a function of k :

$$F_0(\mathbf{z}) = \frac{1}{2k} \ln \left(\frac{\sqrt{P(H_0)f(\mathbf{z}|H_0)} - \exp(-2k)\sqrt{P(H_1)f(\mathbf{z}|H_1)}}{\sqrt{P(H_1)f(\mathbf{z}|H_1)} - \exp(-2k)\sqrt{P(H_0)f(\mathbf{z}|H_0)}} \right)$$

- For building the decision rule, $F_0(\mathbf{z})$ is compared to a threshold fixed attending to P_{FA} requirements $\left(F_0(\mathbf{z}) \underset{H_1}{\overset{H_0}{\gtrless}} \eta_0(P_{FA}) \right)$, and the relation between η_0 and η_{lr} (the threshold

required by the LRT for the same P_{FA}) is studied. As the final relation between both thresholds does not depend on the input vector, the proposed error function fulfills the sufficient condition for any finite k value.

- Applying the limit when $k \rightarrow \infty$, the C-SVM is able to approximate only three points of the ROC curve, corresponding to $P_{FA} = 0$, $P_{FA} = 1$ and the P_{FA} associated to the minimum error probability. In the last case, η_0 is always equal to zero. This result is a very important contributions of this PhD Thesis, because it allows us to conclude that a C-SVM trained with $N \rightarrow \infty$ and $k \rightarrow \infty$ can not approximate the NP detector for any value of P_{FA} .

In practical situations, N and C are finite. Experimentally, good approximations to the Neyman-Pearson detector have been obtained for detecting Swerling I and II targets in AWGN with N and C high enough, at the expense of a high computational cost. In all considered cases, the interval variation of η_0 is so small, that a high accuracy is required to represent the threshold according to P_{FA} requirements.

As an alternative, 2C-SVMs [Davenport2006] allow the control of the cost associated to the two possible errors. In this PhD. Thesis, the associated error function has been analyzed in order to reduce the approximation error to the Neyman-Pearson detector for a given P_{FA} , that can be different from that associated to the minimum error probability. Following a parallel reasoning as in the study of the C-SVM, the Heaviside step functions that appear in the error function are again approximated by logistic ones and the solution is obtained calculating the limit when $k \rightarrow \infty$:

$$F_0(\mathbf{z}) = \frac{1}{2k} \ln \left(\frac{\sqrt{C_{10}P(H_0)}f(\mathbf{z}|H_0) - \exp(-2k)\sqrt{C_{01}P(H_1)}f(\mathbf{z}|H_1)}{\sqrt{C_{01}P(H_1)}f(\mathbf{z}|H_1) - \exp(-2k)\sqrt{C_{10}P(H_0)}f(\mathbf{z}|H_0)} \right)$$

When $\eta_0 = 0$, the 2C-SVM trained with $N \rightarrow \infty$ and $C \rightarrow \infty$, is able to approximate the point of the NP detector ROC curve associated to $\eta_{lr} = \frac{C_{10}P(H_0)}{C_{01}P(H_1)} = \frac{\gamma P(H_0)}{(1-\gamma)P(H_1)}$, where γ is the parameter of the 2C-SVM that controls the asymmetric assignment of costs associated with the two types of error (detection loss and false alarm).

2C-SVMs detectors have been designed for detecting Swerling I and II targets in AWGN. In these cases, the Neyman-Pearson detector is completely characterized, and the value of γ can be obtained as a function of the η_{lr} value associated to a desired P_{FA} . The approximation error is reduced as C and N increases at the expense of a high computational cost.

In a general problem, the η_{lr} associated to a P_{FA} is not usually known a priori, so a grid search in C and γ should be carried out to approximate the desired point of the Neyman-Pearson detector ROC curve.

8.1.2. Regarding the Detection of Targets with Unknown Parameters

As first step, a study of the sensitivity of the optimum detector for targets with known parameters has been carried out, for the cases of study summarized in 8.1:

- In presence of Gaussian interference, the LRT detector is robust against the design Signal-to-Interference Ratio (SIR) but presents different degrees of sensitivity with respect to the

correlation coefficient (ρ_s) and the Doppler shift (Ω). For example, LRTs designed for ρ_s^d close to 1, are less robust with respect to the simulated ρ_s^s , than those LRTs designed for ρ_s^d close to 0. But for Swerling I targets, the LRT performance strongly depends on the design Doppler shift.

- In presence of K-distributed clutter, the LRT detector presents a high dependence on all the considered parameters: the Signal-to-clutter Ratio (SCR) the known phase of the target (θ_R) and the Doppler shift (Ω). The dependence on Ω is specially critical when the clutter is correlated.

CGLR based solutions have been proposed for the cases of study, that will be used as references in the design of the AI based detectors:

- When ρ_s varies uniformly in $[0; 1]$, a good approximation is obtained with a $CGLR_{P/2}$ detector, independent of the type of Gaussian interference. For $P = 8$, the $CGLR_4$ detector requires 2,235 operations.
- When the Ω of a Swerling I target varies uniformly in $[0; 2\pi)$, in presence of Gaussian interference, a $CGLR_{2P}$ is proposed, except for a dominant and high correlated clutter ($CNR = 40$ dB and $\rho_c = 0.995$), for which the improvement with respect to the $CGLR_P$ is insignificant. For $P = 8$, the $CGLR_{16}$ and the $CGLR_8$ require 8,943 and 4,467 operations, respectively.
- A reduction of the Ω variation interval is proposed to design detectors for Swerling V targets in K-distributed clutter, in order to obtain useful radar coverages. Assuming a Ω varying uniformly in $[\pi/2 - \pi/10; \pi/2 + \pi/10]$, a $CGLR_P$ presents a good approximation in uncorrelated clutter, while a $CGLR_{2P}$ is proposed in correlated clutter. For $P = 8$, the $CGLR_8$ and the $CGLR_{16}$ require 9,439 and 18,879 operations, respectively.

8.1.3. Regarding the Design of Robust AI Detectors

In Table 8.1, the NN based detectors proposed as compromise solutions between detection capabilities and computational cost are summarized. MLPs and RBFNNs have been proposed for approximating the Neyman-Pearson detector in simple hypothesis test, but the application of SONNs for approximating the Neyman-Pearson detector is a contribution of this PhD Thesis. For controlling the computational cost, SONNs with only one quadratic neural unit are considered.

Committees of NNs are also studied to improve the detection performance and/or to reduce the computational complexity. A mixture of experts are proposed combining different NNs specialized on a subspace of the observation space:

- If the mixture of experts is composed of NNs of the same type, the output is selected applying the maximum selection criterion (following the philosophy of the CGLR).
- If the mixture of experts is composed of NNs of different type, the outputs of each NN are compared with a threshold fixed according to P_{FA} requirements, and the binary outputs are combined using an OR-logic function.

These combination strategies are original contributions of this PhD Thesis.

The main conclusions that can be extracted from Table 8.1 are the following:

- SONNs present a great robustness in Gaussian interference scenarios regardless of the unknown target parameter.
- For unknown ρ_s in presence of Gaussian interference, the NNs are trained considering training patterns under H_1 with ρ_s varying uniformly in $[0; 1]$. The detectors based on a unique SONN present the best compromise between detection capabilities and computational cost, except in presence of AWGN (case of study 1) where the detector RBFNN 16/5/1 maintains the detection capabilities reducing the computational cost.
- For unknown Ω in presence of Gaussian interference, the mixture of SONNs presents a good compromise between detection capabilities and computational cost. $SONN_1$ has been using $\Omega \in [0; \pi/4) \cup [3\pi/4; 2\pi)$ and $SONN_2$ with $\Omega \in [\pi/4; 3\pi/4)$ trained.

For $CNR = 40$ dB and $\rho_c = 0.995$, the detector based on a unique SONN trained with $\Omega \in [0; 2\pi)$ presents similar performance reducing the computational cost.

- For detecting Swerling V targets in K-distributed clutter, the RBFNN detector presents better detection performance than the MLP one, when the K-distributed clutter is uncorrelated. However, when $\rho_c = 0.9$, the detection capabilities of the MLP detector are better, showing a complementary behavior.

To explain the previous result, the decision boundaries of the CGLR detector for Swerling V targets with unknown Ω in uncorrelated K-distributed clutter have been studied. The decision region associated to H_1 is defined by a sector of the region limited by two concentric hyper-spheres, whose radii depend on the SCR .

- In any case, the computational cost associated to the NN based detectors is really much lower than that associated to the proposed CGLRs.

In Table 8.2, the 2C-SVM based detectors proposed to approximate the NP detector for $P_{FA} = 10^{-4}$ are presented. From these results, the following conclusions can be extracted:

- For each case of study, a grid search in the C, γ space has been carried out using training sets composed of $N = 500$ patterns. This N value has been selected as a compromise solution between system performance and computational cost.
- The number of required operations grows with C . For a given training set size, N , C is as high as possible in order to minimize the empirical error guaranteeing good generalization capabilities.
- 2C-SVM based detectors present good approximations to the desired point of the NP detector ROC curve, except for detecting Gaussian targets with unknown Ω in Gaussian interference.
- The main disadvantage of 2C-SVM based detectors, that is shared with C-SVM solutions, is the huge number of required operations. The computational complexity is higher than that associated to CGLRs.

Case of Study	Unknown parameter	NNs Based Detector	Number of Operations
Case of study 1	$\rho_s \in [0; 1]$	RBFNN 16/5/1	256
	$\Omega \in [0; 2\pi)$	$SONN_1 + SONN_2$	882
Case of study 2 $CNR = 20$ dB $\rho_c = 0.7$	$\rho_s \in [0; 1]$	SONN 16/152/1	441
	$\Omega \in [0; 2\pi)$ Swerling I	$SONN_1 + SONN_2$	882
Case of study 2 $CNR = 20$ dB $\rho_c = 0.995$	$\rho_s \in [0; 1]$	SONN 16/152/1	441
	$\Omega \in [0; 2\pi)$ Swerling I	$SONN_1 + SONN_2$	882
Case of study 2 $CNR = 40$ dB $\rho_c = 0.995$	$\rho_s \in [0; 1]$	SONN 16/152/1	441
	$\Omega \in [0; 2\pi)$ Swerling I	SONN 16/152/1	441
Case of study 3 $\rho_c = 0$	$\Omega \in [1.0178; 1.3176]$ Swerling V	RBFNN 16/7/1	358
Case of study 3 $\rho_c = 0.9$	$\Omega \in [1.0178; 1.3176]$ Swerling V	MLP 16/17/1	597

Tabla 8.1: NN based detectors proposed for the cases of study.

Case of Study	Unknown parameter	C	γ	Number of Operations
Case of study 1	$\rho_s \in [0; 1]$	10^3	0.7	21,172
	$\Omega \in [0; 2\pi)$	10^3	0.7	25,996
Case of study 2 $CNR = 20$ dB; $\rho_c = 0.7$	$\rho_s \in [0; 1]$	10^5	0.95	27,671
	$\Omega \in [0; 2\pi)$; SWI	10^3	0.99	8,911
Case of study 2 $CNR = 20$ dB; $\rho_c = 0.995$	$\rho_s \in [0; 1]$	10^3	0.7	13,601
	$\Omega \in [0; 2\pi)$; SWI	10^3	0.9	13,735
Case of study 2 $CNR = 40$ dB; $\rho_c = 0.995$	$\rho_s \in [0; 1]$	10^3	0.7	9,648
	$\Omega \in [0; 2\pi)$; SWI	10^3	0.9	18,358
Case of study 3 ($\rho_c = 0$)	$\Omega \in [1.0178; 1.3176]$; SWV	10^5	0.999	66,665
Case of study 3 ($\rho_c = 0.9$)	$\Omega \in [1.0178; 1.3176]$; SWV	10	0.999	37,386

Tabla 8.2: 2C-SVM based detectors proposed to approximate the NP detector for $P_{FA} = 10^{-4}$. The cases for which the estimated P_D differs significantly from the desired value have been highlighted in red

Angle	RCS	Case	Relation	Coverage
13° Clutter Gaussiano	-2,305 dBsm	$CNR = 20$ dB $\rho_c = 0.7$	$SIR = 13$ dB	13.02 km
		$CNR = 20$ dB $\rho_c = 0.995$	$SIR = 1$ dB	25.98 km
		$CNR = 40$ dB $\rho_c = 0.995$	$SIR = -14$ dB	19.53 km
0,1° Clutter K-distribuido	0,226 dBsm	$\rho_c = 0$	$SCR = 9$ dB	17.75 km
		$\rho_c = 0.9$	$SCR = -3$ dB	281.44 km

Tabla 8.3: Simulated radar scenario estimated coverages for $P_D > 0.8$ and $P_{FA} \leq 10^{-4}$

8.1.4. Regarding the Application of the Proposed Robust NN Detectors to a Simulated Radar Scenario

A radar scenario has been simulated with two real trajectories from [Localizotodo], corresponding to two ships with variable radial speed due to their movement relative to a hypothetical radar located on the Strait of Gibraltar. Parameters of the IPIX radar for a range resolution of 30m have been considered [Drosopoulos1994].

For the same radar location and ship trajectories, two cases of study have been simulated:

- Gaussian clutter corresponding to incidence angles higher than 10°.
- K-distributed Clutter with $\nu = 0.5$, corresponding to low grazing angles and sea states higher than 1.

CA-CFAR detectors have been applied as reference ones. Results prove that the proposed NN based solutions present better detection performances than CA-CFAR detectors:

- In presence of Gaussian clutter, NN detectors show good detection capabilities for smaller values of SIR, as clutter is more correlated and dominant, while the CA-CFAR needs higher SIR in order to provide similar performances.
- In presence of K-distributed clutter, ships trajectories can not be detected by the CA-CFAR, due to the high number of false alarms. The proposed NN detectors, which have been trained in a restricted variation interval of Ω , are able to detect the part of the trajectories whose Ω belongs to the training variation interval.
- In Table 8.3, the coverages estimated for $P_D > 0.8$ and $P_{FA} = 10^{-4}$ are presented. In this Table, the mean RCS estimated with POFACETS tool for the considered incidence angles (13°, for Gaussian clutter, and 0.1°, for K-distributed clutter) are also indicated. In general, the coverages increase in presence of correlated clutter.

8.2. Future Research Lines

- **Design of CFAR techniques to be combined with the proposed AI detectors.** So far, each radar scan has been processed in order to compensate the propagation losses

and the variation of the illuminated area with range, but a very simple propagation model has been assumed. Therefore, in order to improve the robustness of the designed detectors, CFAR techniques can be designed.

- **Design of robust detectors against fluctuations of the statistical parameters of the assumed clutter model.** Following a methodology similar to that applied in this PhD Thesis for the design of detectors for unknown target parameters, the sensitivity with respect to clutter ones could be studied in order to apply AI based detectors based on a unique agent or mixtures of experts.
- **Design of banks of detectors controlled by clutter automatic classification techniques.** Attending to the sensitivity studies performed in the previous research line, mixtures of experts could be designed, whose outputs should be selected by an automatic classification stage that will determine the type of clutter present.
- **Sensitivity of the CGLR detectors against the variation of multiple target parameters and proposal of robust detectors.** To perform sensitivity studies with respect to the variation of multiple parameters of the target such as the correlation coefficient, the Doppler shift, or the statistical model.
- **Application to extended target models.** Extended target models could be implemented in order to evaluate the performance of the proposed solutions, and investigate new strategies for designing AI solutions capable, not only of providing good centroids for the tracking stages, but also defining suitable clusters of hints to be used for classification purposes if the radar resolution is high enough.
- **Creation of a solid basis for the development of cognitive radar systems.** All the results and conclusions extracted from this PhD Thesis and the related future research lines, constitute a solid theoretical and experimental basis for the future research line of the High Frequency Technologies Group in the design and development of cognitive radar systems.

Parte IV

Bibliografía

Bibliografía

- [Localizado] URL <http://www.localizado.com>.
- [Ahmed1994] AHMED, W., D.M. HUMMELS, y M.T. MUSAVI. «Adaptive RBF neural network in signal detection.» *Proceedings of the IEEE International Symposium on Circuits and Systems*, 6, (1994), 265–268.
- [Aloisio1994] ALOISIO, V., A. DI VITO, y G. GALATI. «Optimum detection of moderately fluctuating radar targets.» *IEE Proceedings on Radar, Sonar and Navigation*, 141, nº 3, (1994), 164–170.
- [Amoozegar1994] AMOOZEGAR, F., y M.K. SUNDARESHAN. «A robust neural network scheme for constant false alarm rate processing for target detection in clutter environment.» *Proceedings of the American Control Conference*.
- [Andina1995d] ANDINA, D. *Optimización de detectores neuronales. Aplicación a radar y sonar*. Tesis doctoral, E.T.S.I. Telecomunicación, U.P.M., Diciembre 1995. Dirigida por J.L. Sanz González.
- [Andina1996a] ANDINA, D., y J.L. SANZ-GONZÁLEZ. «Comparison of a neural network detector vs. Neyman-Pearson optimal detector.» *Proceedings of the 1996 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6, (1995), 3573–3576.
- [Andina1995c] —. «On the problem of binary detection with neural networks.» *38th Midwest Symposium on Circuits and Systems*, 1, (1995), 554–557.
- [Andina1996b] —. «Quasi-optimum detection results using a neural network.» *IEEE International Conference on Neural Networks*, 4, (1996), 1929–1932.
- [Andina1995a] ANDINA, D., J.L. SANZ-GONZÁLEZ, y J.A. JIMÉNEZ-PAJARES. «A comparison of criterion functions for a neural network applied to binary detection.» *IEEE International Conference on Neural Networks*, 1, (1995), 492–495.
- [Andina1995b] ANDINA, D., J.L. SANZ-GONZÁLEZ, y O.A. RODRÍGUEZ-MARTÍN. «Performance improvements for a neural network detector.» *IEEE International Conference on Neural Networks*, 1, (1995), 492–495.

- [Aref1994] AREF, M.R. «Likelihood-ratio detection.» *Proceedings of the 1994 IEEE International Symposium on Information Theory*, (1994), 270.
- [Artyomov2004] ARTYOMOV, E., y O. YADID-PECHT. «Modified high-order neural network for invarinat pattern recognition.» *Pattern Recognition Letters*, 26, nº 6, (2004), 843–851.
- [Awad2010] AWAD, M. «Optimization RBFNNs Parameters Using Genetic Algorithms: Applied on Fuction Approximation.» *International Journal of Computer Science and Security*, 4, nº 3, (2010), 295–307.
- [Bach2006] BACH, F., D. HECKERMAN, y E. HORVITZ. «Considering cost asymmetry in learning classifiers.» *Journal Machine Learning Research*, 7, (2006), 1713–1741.
- [Baker1991] BAKER, C.J. «K-distributed coherent sea clutter.» *IEE Proceedings F Communications, Radar and Signal Processing*, 138, nº 2, (1991), 89–92.
- [Barrett1987] BARRETT, M. «Error Probablity for Optimal and Suboptimal Quadratic Receivers in Rapid Rayleigh Fading Channels.» *IEEE Journal on Selected Areas in Communications*, 5, nº 2, (1987), 302–304.
- [Barton1985] BARTON, D.K. «Land clutter models for radar design and analysis.» *Proceedings of the IEEE*, 73, nº 2, (1985), 198–204.
- [Billingsley2002] BILLINGSLEY, J.B. *Low-Angle Radar Land Clutter*. William Andrew, 2002.
- [Billingsley1999] BILLINGSLEY, J.B., A. FARINA, F. GINI, M.V. GRECO, y L. VERRAZZANI. «Statistical analyses of measured radar ground clutter data.» *IEEE Transactions on Aerospace and Electronic Systems*, 35, nº 2, (1999), 579–593.
- [Bishop1995] BISHOP, C.M. *Neural networks for pattern recognition*. Oxford University Press, Gran Bretaña, 1995, 1 ed^{ón}.
- [Blacknell1994] BLACKNELL, D., y R.G. WHITE. «Optimum classification of non-Gaussian processes using neural networks.» *IEE Proceedings.- Vision, Image and Signal Processing*, 141, nº 1, (1994), 56–66.
- [Breiman1994] BREIMAN, Leo. «Bagging predictors.» *Inf. Téc. Technical Report 421*, Department of Statistics, University of California at Berkeley, 1994.
- [Brennan1968] BRENNAN, L.E., I.S. REED, y W. SOLLFREY. «A comparison of average-likelihood and maximum-likelihood ratio tests for detecting radar targets of unknown Doppler frequency.» *IEEE Transactions on Information Theory*, IT-14, nº 1, (1968), 104–110.
- [Bucciarelli1993a] BUCCIARELLI, T., G. FEDELE, y R. PARISI. «Neural networks based signal detection.» *NAECON 93*, (1993), 814–818.

- [Bucciarelli1993b] BUCCARELLI, T., y R. PARISI. «Radar detection by multilayer perceptrons.» *International Conference on Signal Processing*, 1, (1993), 1135–1138.
- [Burges1998] BURGES, C.J.C. «A tutorial on support vector machines for pattern recognition.» *Data Mining and Knowledge Discovery*, 2, (1998), 121–167.
- [Burian1999] BURIAN, A., P. KUOSMANEN, y J. SAARINEN. «Neural detectors with variable threshold.» *IEEE International Symposium on Circuits and Systems*, 5, (1999), 599–602.
- [Carretero2010] CARRETERO-MOYA, J., J. GISMERO-MENOYO, A. BLANCO DEL CAMPO, y A. ASENSIO-LÓPEZ. «Statistical Analysis of a high-resolution sea-clutter database.» *IEEE Transactions on Geoscience and Remote Sensing*, 48, n^o 4, (2010), 2024–2037.
- [Casasent2003a] CASASENT, D., y X.W. CHEN. «New training strategies for RBF neural networks for X-ray agricultural product inspection.» *Journal of Statistical Computation and Simulation*, 36, n^o 2, (2003), 535–547.
- [Casasent2003b] —. «Radial basis function neural networks for nonlinear Fisher discrimination and Neyman-Pearson classification.» *Neural Networks*, 16, n^o 5-6, (2003), 529–535.
- [Cheikh2004] CHEIKH, K., y S. FAOZI. «Application of neural networks to radar signal detection in K-distributed clutter.» *First International Symposium on Control, Communications and Signal Processing*, (2004), 295–298.
- [Cheikh2006] CHEIKH, K., y F. SOLTANI. «Application of neural networks to radar signal detection in K-distributed clutter.» *IEE Proceedings on Radar, Sonar and Navigation*, 153, n^o 5, (2006), 460–466.
- [Conte1997] CONTE, E., M. DI-BISCEGLIE, C. GALDI, y G. RICCI. «A procedure for measuring the coherence length of the sea texture.» *IEEE Transactions on Instrumentation and Measurement*, 46, n^o 4, (1997), 836–841.
- [Conte1987] CONTE, E., y M. LONGO. «Characterisation of radar clutter as a spherically invariant random process.» *IEE Proceedings F Communications, Radar and Signal Processing*, 134, n^o 2, (1987), 191–197.
- [Conte1991b] CONTE, E., M. LONGO, y M. LOPS. «Modelling and simulation of non-Rayleigh radar clutter.» *IEE Proceedings F Communications, Radar and Signal Processing*, 138, n^o 2, (1991), 121–129.
- [Conte1991] CONTE, E., M. LONGO, M. LOPS, y S.L. ULLO. «Radar detection of signals with unknown parameters in K-distributed clutter.» *IEE Proceedings Radar and Signal Processing*, 138, n^o 2, (1991), 131–138.

- [Conte1994] CONTE, E., M. LOPS, y G. RICCI. «Radar detection in K-distributed clutter.» *IEE Proceedings Radar and Sonar Navigation*, 141, n^o 2, (1994), 116–118.
- [Conte2004a] CONTE, E., y A. DE MAIO. «Mitigation Techniques for Non-Gaussian Sea Clutter.» *IEEE Journal of Oceanic Engineering*, 29, n^o 2, (2004), 284–302.
- [Conte2000] CONTE, E., A. DE MAIO, y C. GALDI. «Signal detection in compound-gaussian noise: Neyman-Pearson and CFAR detectors.» *IEEE Transactions on Signal Processing*, 48, n^o 2, (2000), 419–428.
- [Conte2004b] —. «Statistical analysis of real clutter at different range resolutions.» *IEEE Transactions on Aerospace and Electronic Systems Proceedings*, 40, n^o 3, (2004), 903–918.
- [Conte1998] CONTE, E., y G. RICCI. «Sensitivity study of GLRT detection in compound-gaussian clutter.» *IEEE Transactions on Aerospace and Electronic Systems*, 34, n^o 1, (1998), 308–316.
- [Cortes1995] CORTES, C., y V. VAPNIK. «Support-Vector Networks.» *Machine Learning*, 20, n^o 3, (1995), 273–297.
- [Cristianini2000] CRISTIANINI, N., y J. SHAWE-TAYLOR. *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, Cambridge, U.K., 2000.
- [Cybenko1989] CYBENKO, G. «Approximation by superpositions of a sigmoidal function.» *Mathematics of Control, Signals and Systems*, 2, (1989), 303–314.
- [Davenport2006] DAVENPORT, M.A., y R.G. BARANIUK. «Controlling false alarms with support vector machines.» En *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Toulouse, France, 2006, 589–592.
- [Davenport2010] DAVENPORT, M.A., R.G. BARANIUK, y C.D. SCOTT. «Tuning Support Vector Machines for Minimax and Neyman-Pearson Classification.» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32, n^o 10, (2010), 1888–1898.
- [Debnath1990] DEBNATH, L., y P. MIKUSINSKI. *Introduction to Hilbert Spaces with Applications*. Academic Press, 1990.
- [Drosopoulos1994] DROSOPOULOS, A. «Description of the OHGR Database.» *Inf. Téc. Technical Note 94-14*, Defense Research Establishment Ottawa, December 1994.
- [Eaves1987] EAVES, J.L., y E.K. REEDY. *Principles of modern radar*. Van Nostrand Reinhold, 1987.

- [El-Jaroudi1990] EL-JAROUDI, A., y J. MAKHOUL. «A new error criterion for posterior probability estimation with neural nets.» En *Proceedings of the International Conference on Neural Networks IJCNN*. U.S.A., 1990, tomo 1, 185–192.
- [Ertel2011] ERTEL, W. *Introduction to Artificial Intelligence*. Springer-Verlag, London, 2011.
- [Eryildirim2011] ERYILDIRIM, A., y I. ONARAN. «Pulse Doppler Radar Target Recognition using a Two-Stage SVM Procedure.» *IEEE Transactions on Aerospace and Electronic Systems*, 47, n^o 2, (2011), 1450–1457.
- [Evgeniou2000] EVGENIOU, T., M. PONTIL, y T. POGGIO. «Regularization networks and support vector machines.» *Advances in computational mathematics*, 13, (2000), 1–50.
- [Farina1997] FARINA, A., F. GINI, M.V. GRECO, y L. VERRAZZANI. «High resolution sea clutter data: statistical analysis of recorded live data.» *IEE Proceedings on Radar and Sonar Navigation*, 144, n^o 3, (1997), 121–130.
- [Findlay1968] FINDLAY, A.M. «Sea-clutter measurements by radar-return sampling.» *NRL Rept. 6661*.
- [Freund1996] FREUND, Y., y R.E. SCHAPIRE. «Experiments with a new boosting algorithm.» *Machine Learning: Proceedings of the Thirteenth International Conference on Machine Learning*.
- [Friedman1998] FRIEDMAN, J., T. HASTIE, y R. TIBSHIRANI. «Additive logistic regression: a statistical view of boosting.» *Inf. Téc. Technical Report*, 1998.
- [Fukuda2001] FUKUDA, S., y H. HIROSAWA. «Support vector machine classification of land cover: application to polarimetric SAR data.» En *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium. IGARSS '01*. Sydney, Australia, 2001, 187 – 189.
- [Gandhi1997] GANDHI, P.P., y V. RAMAMURTI. «Neural networks for signal detection in non-gaussian noise.» *IEEE Transactions on signal processing*, 45, n^o 11, (1997), 2846–2851.
- [Gelfand2000] GELFAND, I., y S. FOMIN. *Calculus of Variations*. Mineola, NY: Courier Dover, 2000.
- [Giles1987] GILES, C.L., y T. MAXWELL. «Learning, invariance, and generalization in high-order neural networks.» *Applied Optics*, 26, n^o 23, (1987), 4972–4978.
- [Gini1997] GINI, F. «Sub-optimum coherent radar detection in a mixture of K-distributed and Gaussian clutter.» *IEE Proceedings on Radar and Sonar Navigation*, 144, n^o 1, (1997), 39–48.

- [Gonzalez2003] GONZALEZ, J., I. ROJAS, J. ORTEGA, H. POMARES, F.J. FERNÁNDEZ, y A.F. DÍAZ. «Multiobjective evolutionary optimization of the size, shape, and position parameters of radial basis function networks for function approximation.» *IEEE Transactional on Neural Networks*, 14, nº 6, (2003), 1478–1495.
- [Gonzalez2002] GONZALEZ, J., I. ROJAS, H. POMARES, J. ORTEGA, y A. PRIETO. «A new clustering technique for function approximation.» *IEEE Transactional on Neural Networks*, 13, nº 1, (2002), 132 – 142.
- [Gonzalez2005] GONZÁLEZ-GARCÍA, J.E., J.L. SANZ-GONZÁLEZ, y F. ÁLVAREZ VAQUERO. «Nonparametric permutation tests versus parametric tests in radar detection under K-distributed clutter.» *IEEE International 2005 Radar Conference*, (2005), 250–255.
- [Grajal1999] GRAJAL, J., y A. ASENSIO. «Multiparametric importance sampling for simulation of radar systems.» *IEEE Transactions on Aerospace and Electronic Systems*, 35, nº 1, (1999), 123–137.
- [Greco2004] GRECO, M., F. BORDONI, y F. GINI. «X-Band Sea-Clutter Nonstationarity: Influence of Long Waves.» *IEEE journal of oceanic engineering*, 29, nº 2, (2004), 269–283.
- [Greco2000] GRECO, M., F. GINI, A. FARINA, y J.B. BILLINGSLEY. «Analysis of clutter cancellation in the presence of measured L-Band radar ground clutter data.» En *Proceedings of the IEEE 2000 International Radar Conference*. Alexandria, USA, 2000, 422 – 427.
- [Hagan1994] HAGAN, M.T., y M.B. MENHAJ. «Training feedforward networks with the Marquardt algorithm.» *IEEE Transactions on Neural Networks*, 5, nº 6, (1994), 989–993.
- [Hanson1998] HANSON, G.A., y R.A. ILLIS. «Near optimal detection of complex signals with unknown parameters.» *Conference record of the 31st Asilomar Conference on Signals, Systems and Computers*, 1, (1997), 530–534.
- [Harpham2004] HARPHAM, C., C. W. DAWSON, y M. R. BROWN. «A review of genetic algorithms applied to training radial basis function networks.» *Neural Computing and Applications*, 13, nº 3, (2004), 193–201.
- [Haykin1999] HAYKIN, S. *Neural networks. A comprehensive foundation*. Prentice Hall International, Inc., U.S.A., 1999, 2 ed^{ón}.
- [Haykin2002] HAYKIN, S., R. BAKKER, y B.W. CURRIE. «Uncovering nonlinear dynamics-the case study of sea clutter.» *Proceedings of IEEE*, 90, nº 5, (2002), 860–881.

- [Hecht-Nielsen1989] HECHT-NIELSEN1989, R. «Theory of the backpropagation neural network.» *International Joint Conference on Neural Networks*, 1, (1989), 593–605.
- [Hovanesian1984] HOVANESSIAN, S. A. *Radar system design and analysis*. Artech House, Inc., U.S.A., 1984.
- [Hummels1995] HUMMELS, D.M., W. AHMED, y M.T. MUSAVI. «Adaptive detection of small sinusoids in non-gaussian noise using a RBF neural network.» *IEEE Transaction on Neural Networks*, 6, nº 1, (1995), 214–219.
- [Hwang2006] HWANG, S.I., S. ISHII, S. SAYAMA, y M. SEKINE. «Suppression of Sea-ice Clutter Observed by a Millimeter Wave Radar.» *Int. Journal Infrared Millimeter Waves*, 28, nº 1.
- [IEEE-Std1997] SYSTEMS PANEL OF THE IEEE AES SOCIETY, Radar. «IEEE standard radar definitions.» *Inf. Téc. IEEE-Std, 686-1997*, IEEE Aerospace and Electronic Systems Society, March 1998.
- [Jakeman1976] JAKEMAN, E., y P.N. PUSEY. «A model for non-Rayleigh sea echo.» *IEEE Transactions on Antennas and Propagation*, 24.
- [Jao1984] JAO, J.K. «Amplitude distribution of composite terrain radar clutter and the K-distribution.» *IEEE Transactions on Antennas and Propagation*, AP-32.
- [Jarabo2005a] JARABO-AMORES, M.P. *Desarrollo y optimización de un detector neuronal de blancos radar en presencia de interferencia aditiva*. Tesis doctoral, Escuela Politécnica Superior, U.A.H., 2005. Dirigida por F. López-Ferreras.
- [Jarabo2005b] JARABO-AMORES, M.P., M. ROSA-ZURERA, R. GIL-PITA, y F. LÓPEZ-FERRERAS. «Sufficient condition for an adaptive system to approximate the Neyman-Pearson detector.» En *IEEE/SP 13th Workshop on Statistical Signal Processing*. Bordeaux, France, 2005, 295 – 300.
- [Jarabo2003] JARABO-AMORES, P., R. GIL-PITA, M. ROSA-ZURERA, y F. LÓPEZ-FERRERAS. «MLP and RBFN for detecting white gaussian signals in white gaussian interference.» *Lecture Notes in Computer Sciences (Springer-Verlag)*, 2687, (2003), 790–797.
- [Jarabo2006] JARABO-AMORES, P., D. MATA-MOYA, M. ROSA-ZURERA, J.C. NIETO-BORGE, y F. LÓPEZ-FERRERAS. «A neural network approach to improve radar detector robustness.» En *Proceedings of the 14th European Signal Processing Conference*. Florence, Italy, 2006.

- [Jarabo2009] JARABO-AMORES, P., M. ROSA-ZURERA, R. GIL-PITA, y F. LOPEZ-FERRERAS. «Study of Two Error Functions to Approximate the Neyman-Pearson Detector Using Supervised Learning Machines.» *IEEE Transactions on Signal Processing*, 57.
- [Kailath1966] KAILATH, T. «The complex envelope of white noise.» *IEEE Transactions on Information Theory*, IT-12, n^o 4, (1966), 397–398.
- [Karayiannis1997] KARAYIANNIS, N.B., y G.W. MI. «Growing radial basis neural networks merging supervised and unsupervised learning with network growth techniques.» *IEEE Transactions on Neural Networks*, 8, n^o 6, (1997), 1492–1506.
- [Karayiannis1995] KARAYIANNIS, N.B., y A.N. VENETSANOPOULOS. «On the training and performance of high-order neural networks.» *Mathematical Biosciences*, 129, n^o 2, (1995), 143–168.
- [Khairnar2008] KHAIRNAR, D. G., S. N. MERCHANT, y U. B. DESAI. «Radar Signal Detection In Non-Gaussian Noise Using RBF Neural Network.» *Journal of Computers*, 3, n^o 1, (2008), 32–39.
- [Lanczos1961] LANCZOS, C. *Linear differential operators*. Van Nostrand Company Inc., Princeton, New Jersey, 1961.
- [Leung2002] LEUNG, H., N. DUBASH, y N. XIE. «Detection of small objects in clutter using a GA-RBF neural network.» *IEEE Transactions on Aerospace and Electronic Systems*, 38, n^o 1, (2002), 98–118.
- [Lazaro2003] LÁZARO, M., I. SANTAMARÍA, y C. PANTALEÓN. «A new EM-based training algorithm for RBF networks.» *Neural Networks*, 16, n^o 1, (2003), 66–77.
- [Marcum1948] MARCUM, J.I. «A Statistical theory of target detection by pulsed radar.» *IRE Transactions of Information Theory*, IT-6, n^o 2, (1960), 59–267. Originalmente publicado como Research Memorandum RM-753, Rand Corp., Julio 1948.
- [Mata2011a] MATA-MOYA, D., P. JARABO-AMORES, y J. MARTIN DE NICOLÁS-PRESA. «High Order Neural Network Based Solution for Approximating the Average Likelihood Ratio.» En *Proceedings of IEEE Statistical Signal Processing Workshop (SSP)*. Niza, Francia, 2011, 657 – 660.
- [Mata2008] MATA-MOYA, D., P. JARABO-AMORES, J.C. NIETO-BORGE, y F. LÓPEZ-FERRERAS. «Robust detectors for targets of unknown correlation in clutter.» En *Proceedings of the International Conference on Radar*. Adelaide, Australia, 2008, 247 – 252.

- [Mata2009] MATA-MOYA, D., P. JARABO-AMORES, M. ROSA-ZURERA, J.C. NIETO-BORGE, y F. LÓPEZ-FERRERAS. «Combining MLPs and RBFNNs to Detect Signals With Unknown Parameters.» *IEEE Transactions on Instrumentation and Measurement*, 58, nº 9, (2009), 2989 – 2995.
- [Mata2007] MATA-MOYA, D., P. JARABO-AMORES, M. ROSA-ZURERA, R. VICEN-BUENO, y J.C. NIETO-BORGE. «Neural Network Based Approaches for Detecting Signals With Unknown Parameters.» En *IEEE International Symposium on Intelligent Signal Processing. WISP'2007*. Alcalá de Henares, Spain, 2007, 1 – 6.
- [Mata2008b] MATA-MOYA, D., P. JARABO-AMORES, R. VICEN-BUENO, J.C. NIETO-BORGE, M. ROSA-ZURERA, y F. LÓPEZ-FERRERAS. «MLP Solutions for Approximating the Average Likelihood Ratio Detector in Radar Applications.» En *Proceedings of IEEE Radar Conference. RADAR'08*. Rome, Italy, 2008, 1 – 6.
- [Mata2006] MATA-MOYA, D., P. JARABO-AMORES, R. VICEN-BUENO, M. ROSA-ZURERA, y F. LÓPEZ-FERRERAS. «Neural Network Detectors for Composite Hypothesis Test.» *Lecture Notes in Computer Science*, 4224, (2006), 298–305.
- [Mata2011b] MATA-MOYA, D., J. MARTIN DE NICOLÁS-PRESA, P. JARABO-AMORES, y R. GIL-PITA. «Statistical blind classification of terrain surfaces in SAR images.» En *Proceedings of International Conference on Space Technology (ICST)*. Atenas, Grecia, 2011.
- [Mert2011] MERT, A., H.J. DE WIND, J.E. CILLIERS, y P. HERSELMAN. «Breast Cancer Classification by Using Support Vector Machines with Reduced Dimension.» En *Proceedings of ELMAR 2011*. Zadar, Croacia, 2011, 37 – 40.
- [Michalewicz1996] MICHALEWICZ, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, 1996.
- [Miller1988] MILLER, R.J. «A coherent model of radar weather clutter.» En *Proceedings of the 1988 IEEE National Radar Conference*. Ann Arbor, USA, 1988, 246 – 248.
- [Moody1989] MOODY, J., y C.J. DARKEN. «Fast learning in networks of locally-tuned processing units.» *Neural Computation*, 1, nº 3, (1989), 284–294.
- [Musavi1992] MUSAVI, M.T., W. AHMED, K.H. CHAN, K.B. FARIS, y D.M. HUMMELS. «On the training of radial basis function classifiers.» *Neural Networks*, 5, nº 4, (1992), 595–603.
- [Nathanson1999] NATHANSON, F.E. *Radar Design Principles. Signal Processing and the Environment*. MacGraw-Hill, 1999.

- [Nayebi1996] NAYEBI, M.M., M.R. AREF, y M.H. BASTANI. «Detection of coherent radar signals with unknown Doppler shift.» *IEE Proceedings on Radar, Sonar and Navigation*, 143, n^o 2, (1996), 79–86.
- [Neyman1933] NEYMAN, y PEARSON. «On the problem of the most efficient test of statistical hypotheses.» *Phil. Trans. Roy. Soc., A* 231, n^o 9, (1933), 492–510.
- [Orsak1993] ORSAK, G.C. «A note on estimating false alarm rates via importance sampling.» *IEEE Transactions on Communications*, 41, n^o 9, (1993), 1275–1277.
- [Osuna1997] OSUNA, E., T. FREUND, y F. GIROSI. «Support vector machines: Training and Applications.» *Inf. Téc. 1602*, MTI Artificial Intelligence Laboratory, 1997.
- [Paisitkriangkrai2008] PAISITKRIANGKRAI, S., C. SHEN, y J. ZHANG. «Performance evaluation of local features in human classification and detection.» *IET Computer Vision*, 2, n^o 4, (2008), 236–246.
- [Papoulis1991] PAPOULIS, A. *Probability, random variables, and stochastic processes*. McGraw-Hill Int. Ed., Singapur, 1991, 3 ed^{ón}.
- [Park2000] PARK, S., M.J.T. SMITH, y R.M. MERSEREAU. «Target recognition based on directional filter banks and higher-order neural networks.» *Digital Signal Processing*, 10, n^o 4, (2000), 297–308.
- [Pentini1992] PENTINI, F.A., A. FARINA, y F. ZIRILLI. «Radar detection of targets located in a coherent K distributed clutter background.» *IEE Proceedings F Radar and Signal Processing*, 139, n^o 3, (1992), 239–245.
- [Rangaswamy2003] RANGASWAMY, M. «Spherically invariant random processes for modeling non-Gaussian radar clutter.» En *Proceedings of the Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*. 1993, 1106 – 1110.
- [Redlapalli2003] REDLAPALLI, S., M.M. GUPTA, y K. SONG. «Development of quadratic neural unit with applications to pattern classification.» En *Proceedings of the Fourth International Symposium on Uncertainty Modeling and Analysis (ISUMA'03)*. College Park, USA, 2003, 141 – 146.
- [Rich1983] RICH, E. *Artificial Intelligence*. McGraw-Hill, New York, 1983.
- [Rosenblatt1959] ROSENBLATT, F. «The Perceptron: A probabilistic model for information storage and organization in the brain.» *Psychological Review*, 65, (1959), 386–408.
- [Ruck1990] RUCK, D.W., S.K. ROGERS, M. KABRISKY, M.E. OXLEY, y B.W. SUTER. «The multilayer perceptron as an approximation to a Bayes optimal

- discriminant function.» *IEEE Transactions on Neural Networks*, 1, n^o 4, (1990), 296–298.
- [Rumelhart1986] RUMELHART, D.E., G.E. HILTON, y R.J. WILLIAMS. «Learning representations of back-propagation errors.» *Nature*, 323, (1986), 533–536.
- [Saerens2002] SAERENS, M., P.LATINNE, y C.DECAESTECKER. «Any reasonable cost function can be used for a posteriori probability approximation.» *IEEE Transactions on Neural Networks*, 13, (2002), 1204–1210.
- [Sanz-Gonzalez1999] SANZ-GONZALEZ, J.L., y D. ANDINA. «Performance analysis of neural network detectors by importance sampling techniques.» *Neural Processing Letters*, 9, (1999), 257–269.
- [Sayama2001] SAYAMA, S., y M. SEKINE. «Weibull, log-Weibull and K-distributed ground clutter modeling analyzed by AIC.» *IEEE Transactions on Aerospace and Electronic Systems*, 37, n^o 3.
- [Sayama2002] —. «Amplitude statistics of sea-ice clutter using a millimeter wave radar.» *Int. Journal Infrared Millimeter Waves*, 23, n^o 11.
- [Sayama2004] —. «Suppression of sea-ice clutter observed by a millimeter wave radar using a new log-weibull/CFAR system.» *Int. Journal Infrared Millimeter Waves*, 25, n^o 10.
- [Schapire1990] SCHAPIRE, R.E. «The strength of weak learnability.» *Machine Learning*, 5, n^o 2.
- [Scholtz1976] SCHOLTZ, R.A., J.J. KAPPL, y N.E. NAHI. «The Detection of Moderately-Fluctuating Rayleigh Targets.» *IEEE Transactions on Aerospace and Electronic Systems*, AES-12, n^o 2.
- [Schwartz1956] SCHWARTZ, M. «Effects of signal fluctuation on the detection of pulsed signal in noise.» *IRE Transactions on Information Theory*, 2, n^o 2.
- [Sekine1990] SEKINE, M., y Y. MAO. *Weibull radar clutter*. Peter Pregrinus Ltd., London, 1990.
- [Sekine1980] SEKINE, M., T. MUSHI, T. IRABU, E. KIUCHI, T. HAGISAWA, y Y. TOMITA. «Non-Rayleigh weather clutter.» *IEE Proceedings F Communications, Radar and Signal Processing*, 127, n^o 6.
- [Sekine1979] SEKINE, M., T. MUSHI, Y. TOMITA, T. IRABU, y E. KIUCHI. «On Weibull-Distributed weather clutter.» *IEEE Transactions on Aerospace and Electronic Systems*, AES-15, n^o 6.
- [Sekine1981] SEKINE, M., S. OHTANI, T. MUSHI, T. IRABU, E. KIUCHI, T. HAGISAWA, y Y. TOMITA. «Suppression of ground and weather clutter.» *IEE Proceedings F Communications, Radar and Signal Processing*, 128, n^o 3.

- [Sharkey1999] SHARKEY, A. *Combining Artificial Neural Nets*. Springer-Verlag, 1999.
- [Shawe2004] SHAWE-TAYLOR, J., y N. CRISTIANINI. *Kernel methods for pattern analysis*. Cambridge University Press, Cambridge, U.K., 2004.
- [Skolnik2002] SKOLNIK, M.I. *Introduction to radar systems*. McGraw-Hill Science, USA, 2002, 3 ed^{ón}.
- [Skolnik2008] —. *Radar handbook*. McGraw-Hill, Inc., USA, 2008, 3 ed^{ón}.
- [Spirkovska1993] SPIRKOVSKA, L., y M.B. REID. «Coarse-Coded Higher-Order NEural Networks for PSRI object recognition.» *IEEE Transactions on Neural Networks*, 4, n^o 2, (1993), 276–283.
- [Srinath1996] SRINATH, M. D., y P. K. RAJASEKARAM. *An introduction to statistical signal processing with applications*. Prentice-Hall Inc., USA, 1996, 2 ed^{ón}.
- [Swerling1954] SWERLING, P. «Probability of detection for fluctuating targets.» *IRE Transactions on information theory*, IT-6, n^o 2, (1960), 246–308. Originalmente publicado como Research Memorandum RM-1217, Rand Corp. Marzo de 1954.
- [Swerling1997] —. «Radar probability of detection for some additional fluctuating target cases.» *IEEE Transactions on Aerospace and Electronic Systems*, 33, n^o 2, (1997), 698–709.
- [Thomson2001] THOMSON, A.D., y E.S. RISEBOROUGH. «Evaluation of a eather clutter simulation.» *IEE Proceedings on Radar, Sonar and Navigation*, 148, n^o 3.
- [Tikhonov1977] TIKHONOV, A.N., y V.Y. ARSENIN. *Solutions of Ill-posed Problems*. W.H. Winston, Washington, U.S.A., 1977.
- [Van1968] TREES, H.L. Van. *Detection, estimation, and modulation theory*, tomo 1. Wiley, 1968.
- [Trunk1972] TRUNK, G.V. «Radar properties of non-Rayleigh sea clutter.» *IEEE Transactions on Aerospace and Electronic Systems*, AES-8, n^o 2.
- [Trunk1970] TRUNK, G.V., y S.F. GEORGE. «Detection of targets in non-Gaussian sea clutter.» *IEEE Transactions on Aerospace and Electronic Systems*, AES-6, n^o 5.
- [Turin1960] TURIN, G.L. «The characteristic function of Hermitian quadratic forms in complex normal variables.» *Biometrika*, 47, (1960), 199–201.
- [Vapnik1995] VAPNIK, V. *The nature of statiscal learning theory*. Springer-Verlag, New York, 1995.
- [Veropoulos1999] VEROPULOS, K., N. CRISTIANINI, y C. CAMPBELL. «Controlling the sensitivity of support vector machines.» En *Proceedings of the International Joint Conference on AI*. 1999, 55–60.

- [Villalobos1993] VILLALOBOS, L., y F.L. MERAT. «Learning capability assessment and feature space optimization for higher-order neural networks.» *IEEE Transactions on Neural Networks*, 6, n^o 1, (1993), 267–272.
- [diVito1999] DI VITO, A., y M. NALDI. «Robustness of the likelihood ratio detector for moderately fluctuating radar targets.» *IEE Proceedings on Radar, Sonar and Navigation*, 146, n^o 2, (1999), 107–112.
- [Wan1990] WAN, E.A. «Neural network classification: a bayesian interpretation.» *IEEE Transactions on Neural Networks*, 1, n^o 4, (1990), 303–305.
- [Wan1996] WAN, L., y L. SUN. «Automatic target recognition using higher order neural network.» En *Proceedings of the IEEE 1996 National Aerospace and Electronics Conference*. Dayton, USA, 1996, 221 – 226.
- [Wang2005] WANG, L. *Support Vector Machines: Theory and Applications*. Springer-Verlag, Heidelberg, 2005.
- [Ward1990] WARD, K.D., C.J. BAKER, y S. WATTS. «Maritime surveillance radar I. Radar scattering from the ocean surface.» *IEE Proceedings F Radar and Signal Processing*, 137, n^o 2, (1990), 51–62.
- [Ward2006] WARD, K.D., R.J.A. TOUGH, y S. WATTS. *Sea clutter: Scattering the K-Distribution and radar performance*. The Institution of Engineering and Technology, London, 2006.
- [Watts1985] WATTS, S. «Radar detection prediction in sea clutter using the compound K-distribution model.» *IEE Proceedings F Communications, Radar and Signal Processing*, 132, n^o 7, (1985), 613–620.
- [Watts1987a] —. «Radar detection prediction in K-distributed sea clutter and thermal noise.» *IEEE Transactions on Aerospace and Electronic Systems*, AES-23, n^o 1, (1987), 40–45.
- [Watts2000] —. «The performance of cell-averaging CFAR systems in sea clutter.» *The Record of the IEEE 2000 International Radar Conference*, (2000), 398 – 403.
- [Watts1990] WATTS, S., C.J. BAKER, y K.D. WARD. «Maritime surveillance radar II. Detection performance prediction in sea clutter.» *IEE Proceedings F Radar and Signal Processing*, 137, n^o 2, (1990), 63–72.
- [Watts1987b] WATTS, S., y K.D. WARD. «Spatial correlation in K-distributed sea clutter.» *IEE Proceedings F Communications, Radar and Signal Processing*, 134, n^o 6, (1987), 526–532.
- [Weinstock1952] WEINSTOCK, R. *Calculus of variations with applications to Physics and Engineering*. McGraw-Hill, New York, 1952.

- [Whitehead1996] WHITEHEAD, B.A., y T.D. CHOATE. «Cooperative-Competitive genetic evolution of radial basis function centers and widths for time series prediction.» *IEEE Transactions on Neural Networks*, 7, n^o 4, (1996), 869–880.
- [Yamaguchi2002] YAMAGUCHI, H. «Target detection in ground clutter with W-band stepped frequency polarimetric radar system.» En *Proceedings of RADAR 2002*. Edinburgh, UK, 2002, 258 – 262.
- [Yao1999] YAO, X. *Evolutionary Computation: Theory and Applications*. World Scientific, Singapore, 1999.
- [Zhang2004] ZHANG, S.J., Z.L. JING, y J.X. LI. «Fast learning high-order neural networks for pattern recognition.» *Electronics Letters*, 40, n^o 19, (2004), 1207 – 1208.

Lista de Símbolos y Acrónimos

Símbolos y Acrónimos con Caracteres Latinos

$a(t)$	Modulación de amplitud de la señal transmitida por el radar de impulsos
A_C	Área iluminada en una celda radar
A_e	Área efectiva de la antena del radar
ALR	Cociente de verosimilitud promediado
AWGN	Ruido gaussiano blanco aditivo
B_n	Ancho de banda de ruido de los circuitos de pre-detección
BW	Ancho de haz en acimut de la antena
c	Velocidad de propagación de la onda
$\tilde{\mathbf{c}}$	Vector complejo de observación de clutter
c_p	Componente en fase de la señal compleja recibida de clutter
c_q	Componente en cuadratura de la señal compleja recibida de clutter
C	Parámetro de regularización de las SVM
\mathbf{c}_i	Centros de las RBFs
$\tilde{\mathbf{c}}$	vector de clutter
C_{10}	Coste asociado a las pérdidas de detección
C_{01}	Coste asociado a las falsas alarmas
CALR	Cociente de verosimilitud restringido
CA-CFAR	CFAR basado en el promedio de las celdas de referencia
CFAR	Probabilidad de falsa alarma constante
CGLR	Cociente de verosimilitud generalizado restringido
CNR	Relación señal a ruido
D	Operador Diferencial Lineal
$e(n)$	Función de error instantáneo
E	Función de error
E_m	Valor medio del error
$E[\cdot]$	Operador esperanza matemática
E_{RC}	Energía de la señal recibida de clutter
E_{RS}	Energía de la señal recibida de blanco
E_T	Energía de la señal transmitida por el radar de impulsos
$f_p(\tilde{\mathbf{z}})$	FDP multivariable
f_0	Frecuencia nominal del sistema
$F(\mathbf{z})$	Función aproximada por un agente inteligente
$F_N(\mathbf{z})$	Función aproximada con un conjunto de entrenamiento de tamaño N

FDP	Función de densidad de probabilidad
$F_0(\mathbf{z})$	Función aproximada óptima con un entrenamiento perfecto
FDP	Función de densidad de probabilidad
$g(I(\mathbf{x}))$	Representa la función de activación de la neurona de la red
G	Ganancia de la antena del radar
$G(\mathbf{z}_i, \mathbf{z}_j)$	Función de Green
GLR	Cociente de verosimilitud generalizado
H_0	Hipótesis de blanco ausente
H_1	Hipótesis de blanco presente
HONN	Red neuronal de orden superior
k	Parámetro que controla la pendiente de una función logística
K	Número de detectores LR que componen el CGLR
$K_x(y)$	Función de Bessel modificada de segunda especie
L	Número de celdas de referencia utilizado en los detectores CA-CFAR
L_c	Longitud de coherencia
LR	Cociente de verosimilitud
M	Número de neuronas ocultas
MAP	Criterio de máxima probabilidad a posterior
MLP	Perceptron multicapa
MTI	Indicación de blancos en movimiento
N	Tamaño del conjunto de entrenamiento
$\tilde{\mathbf{n}}$	Vector complejo de observación de ruido
N_{op}	Número de operaciones
N_{sv}	Número de vectores soporte
N_0	Densidad espectral de ruido
NN	Red neuronal artificial
NP	Neyman-Pearson
P	Número de ecos de un blanco recogidos en un scan
p_c	Varianza de las componentes en fase y cuadratura de clutter
p_n	Varianza de las componentes en fase y cuadratura de ruido
p_s	Varianza de las componentes en fase y cuadratura de blanco
P_D	Probabilidad de detección
P_{FA}	Probabilidad de falsa alarma
P_t	Potencia transmitida por el radar
$P(H_i \mathbf{z})$	Probabilidad de la hipótesis H_i condicionada al vector de observación \mathbf{z}
$P(H_i)$	Probabilidad a priori de H_i
PRF	Frecuencia de repetición de pulsos
q	Forma cuadrática
QNU	Unidad neuronal cuadrática
R	Distancia
R_{est}	Riesgo estructural
R_{emp}	Riesgo empírico
RBF	Función de base radial

RBFNN	Red neuronal con funciones de base radial
ROC	Característica de operación del receptor
$s_R(t)$	Señal recibida por el sistema radar
$s_T(t)$	Señal transmitida por el sistema radar de impulsos
s	proceso aleatorio asociado a la componente de textura del clutter
\mathbf{s}_i	Vectores soporte
$\tilde{\mathbf{s}}$	Vector complejo de observación de blanco
SAR	Radar de apertura sintética
SCR	Relación señal a clutter
SIR	Relación señal a interferencia
SIRP	Proceso aleatoria invariante esféricamente
SIRV	Vector aleatorio invariante esféricamente
SNR	Relación señal a ruido
SONN	Red neuronal con funciones de segundo orden
SVM	Máquina de vectores soporte
T_0	Periodo de repetición de pulsos
T_q	Umbral adaptado según técnicas CA-CFAR
T_s	Temperatura de ruido del sistema
t_{H_i}	Salida deseada de la red neuronal cuando el vector de entrada se genera con la hipótesis H_i
$u[x]$	Función escalón
V_C	Volumen iluminado en una celda radar
v_R	Velocidad radial del blanco
\mathbf{w}	Matriz de pesos de conexión
$\tilde{\mathbf{z}}$	Vector complejo de observación de observación
\mathcal{Z}	Conjunto de patrones de entrenamiento

Símbolos y Acrónimos con Caracteres Griegos

α	Coefficientes de Lagrange
δ_{ij}	Delta de Kronecker
η_r	Umbral de detección asociado al LR
η_{ALR}	Umbral de detección asociado al ALR
η_{GLR}	Umbral de detección asociado al GLR
η_{CGLR}	Umbral de detección asociado al CGLR
η_0	Umbral de detección asociado a la red neuronal
η_s	Umbral de detección asociado a un estadístico suficiente
γ	Parámetro de coste asimétrico de las 2C-SVMs
Γ	Función Gamma
λ	Parámetro de regularización
$\Lambda(\mathbf{z})$	Cociente de verosimilitud
ν	Parámetro de forma de la FDP de distribución K
Ω	Frecuencia Doppler normalizada
Ω^d	Frecuencia Doppler normalizada asumida en el diseño
Ω^s	Frecuencia Doppler normalizada utilizada en la simulación
ω_{ij}	Pesos de conexión de la capa i con la capa j
ω_g	Velocidad angular de giro de la antena
ω_p	Pulsación de la portadora transmitida por el radar de impulsos
ϕ	ángulo de incidencia
ψ_R	Fase recibida del clutter
ρ_c	Coefficiente de correlación del clutter
ρ_s	Coefficiente de correlación del blanco
ρ^d	Coefficiente de correlación asumido en el diseño
ρ^s	Coefficiente de correlación utilizado en la simulación
σ_s	Sección radar de blanco
σ_c	Sección radar de clutter superficial
$\sigma^{(0)}$	Sección radar de clutter por unidad de superficie
Σ	Matriz de covarianza
τ	Duración del pulso transmitido por el radar de impulsos
θ_R	Fase recibida del blanco
φ	Parámetro desconocido del blanco
ϖ_D	Frecuencia Doppler
ξ	Margen de error