Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp259-264)

MLPs for Detecting Radar Targets in Gaussian Clutter

D. de la MATA-MOYA, P. JARABO-AMORES, R. VICEN-BUENO, L. CUADRA-RODRÍGUEZ, F. LÓPEZ-FERRERAS
Departamento de Teoría de la Señal y Comunicaciones, Escuela Politécnica Superior
Universidad de Alcalá
Ctra. Madrid-Barcelona, Km. 33.600, 28805, Alcalá de Henares – Madrid
SPAIN
http://www2.uah.es/teose/

Abstract: - A neural network based coherent detector is proposed for detecting gaussian targets in gaussian clutter. Target and clutter ACF are supposed gaussian with different powers and one lag correlation coefficients. While clutter mean Doppler frequency is set to 1, the influence of target mean Doppler frequency is considered. The neural detector performance is compared to the Neyman-Pearson one. For evaluating the neural detector performance, Montecarlo Simulation and Importance Sampling Techniques are used in order to assure a low relative error with a suitable computational charge. Results show that a low complexity neural network can implement very good approximations of the Neyamn-Pearson detector for the case of study. In the presented cases, the MLP performance tends to decrease when the TSIR (Training Signal to Interference Ratio) decreases to very low values, but it is more robust when the correlation characteristics of target and clutter are varied.

Key-Words: - Neyman-Pearson detector, gaussian interference, radar detection, neural network, Importance Sampling, Levenberg-Marquardt

1 Introduction

This paper deals with the study of the Neyman-Pearson (NP) detector for radar target detection in clutter, and the possibility of using neural networks (NNs) for implementing robust approximations to this detector. The NP 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 [1]. The problem of detecting radar echoes in additive, white, Gaussian noise has been studied widely, but less attention has been paid to the problem of detecting radar targets in clutter.

In relation with the application of NNs to radar detection problems, Ruck et al. [2] and Wan [3] demonstrated that a NN can be used to approximate the optimum bayessian classifier, when trained using the least mean squared-error (LMSE) criterion. NNs have also been applied to approximate the NP detector [4,5]. In these works, multi-layer perceptrons (MLPs) with a hidden layer and one output, trained using the standard back-propagation (BP) algorithm were used to detect deterministic signals and non-fluctuating targets with zero doppler frequency considering different white interference models.

The detection of fluctuating tatgets using NNs has been studied in [6,7], assuming Swerling I and II fluctuation models [8], (target amplitude Rayleigh distributed) .

In this paper, the detection of Rayleigh targets with a one lag correlation coefficient in the range (0,1) (1 corresponds to Swerling I targets and 0 corresponds to Swerling II ones) and gaussian autocorrelation function (ACF) is studied. The clutter is modeled also as a Gaussian process with gaussian ACF. The gaussian probability density function (pdf) can be used for modeling atmospheric clutter, but the pdf of land and sea clutter only can be modeled as gaussian when the radar resolution cell, or area illuminated by the radar, is relatively large. Traditionally, clutter power spectral density (PSD) has been modeled as gaussian, but it has been demonstrated that land and rain clutter are better modeled as [8]:

$$S(f) = \frac{S_0}{1+\left(\frac{f}{f_c}\right)^n} \quad (1)$$

So, although the clutter model that assumes gausian pdf and ACF is not the better choice for many practical situations, the fact that the Neyman-Pearson detector is easy to obtain and analyze, makes it a powerful tool to prove the ability of NNs to approximate the Neyman-Pearson detector and to analyze the robustness improvement that can be obtained using NN based detectors instead the optimum one. Note that no assumption is made about the target and interference models during the training process, so the results obtained with the considered models can be generalized to prove the possibility of using NNs in practical situations where target and interference statistics are unknown and difficult to estimate.

Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp259-264)

As a previous step, the NP detector is calculated. For training the MLPs, the Levenberg-Marquardt (LM) [9] algorithm is used. To evaluate the performance of the trained MLP, its Receiver Operating Characteristic (ROC) curve is estimated and compared to the Neyman-Pearson detector one. The P_D is estimated by conventional Montecarlo simulation, but for estimating the P_{FA} Importance Sampling techniques are used in order to obtain a low relative error with a reasonably computational charge [10,11]. Finally, after a careful study of the results, the more relevant conclusions are extracted.

2 Problem Formulation

The target echo is modeled as a gaussian complex vector of dimension n, and gaussian ACF, with a covariance matrix $\mathbf{M_s}$.

$$\left(\mathbf{M}_s\right)_{h,k} = p_s \cdot \rho_s^{|h-k|^2} \cdot \exp\left[j2\pi \left(h-k\right) f_s \right] \qquad (2)$$

Where:

- p_s is the target power.
- f_s is the target mean Doppler frequency normalized to the radar pulse repetition frequency (PRF). Note that the PRF is the sampling frequency of the system-
- ρ_s is the target one lag correlation coefficient.

The interference is modeled as white, gaussian noise (thermal noise) plus gaussian clutter with gaussian ACF, so the associated covariance matrix is given by:

$$\left(\mathbf{M}_n\right)_{h,k} = p_w \delta_{hk} + p_c \cdot \rho_c^{|h-k|^2} \qquad (3)$$

Where:

- p_w and p_c are the white noise and the clutter powers, respectively.
- δ_{hk} is the Kronecker delta
- ρ_c is the clutter one lag correlation coefficient.

Without loss of generality, as detection performance is a function of the difference between the target Doppler frequency and the clutter one, it has been assumed a clutter Doppler frequency equal to zero.

As a normalization criterion, p_w is assumed equal to 2 (in-phase and in-quadrature components of the white noise interference components are of unity variance), and taking this into consideration the clutter to noise ratio (CNR) and the signal to interference (SIR) ratios can be expressed as:

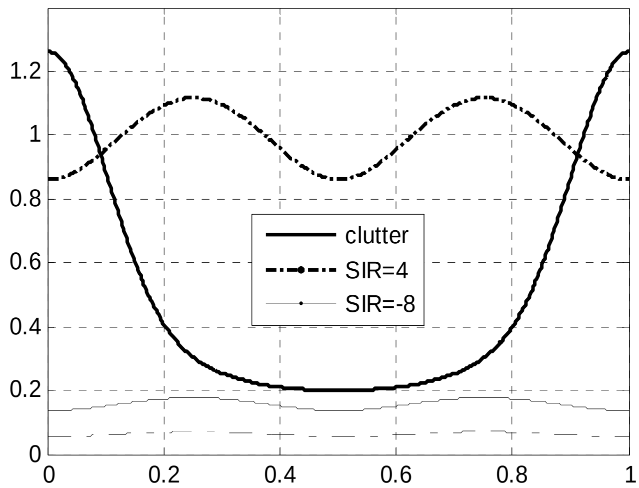

Fig. 1. PDS of target and clutter (solid thick line) for CNR=20 dB, SIR= 4, -4 and -8 dB and f_s=0,25

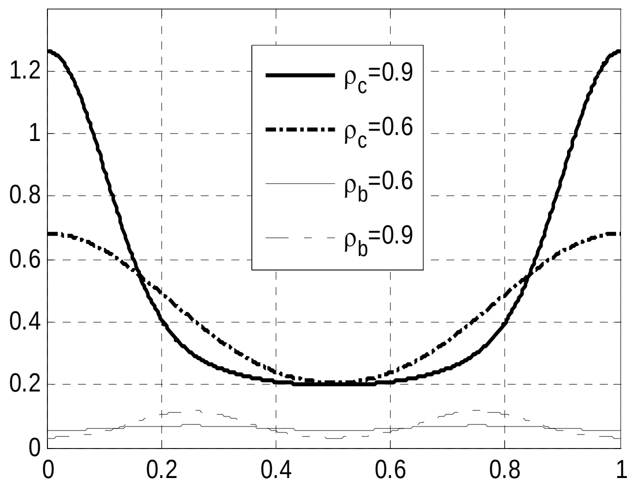

Fig. 2. PDS of target (thin lines) and clutter (thick line) for CNR=20 dB, SIR= -8 dB, f_s=0.25 and different one lag correlation coefficients.

$$\begin{aligned} CNR &= 10\log_{10}\left(cnr\right) \\ &= 10\log_{10}\left(\frac{p_c}{p_w}\right) = 10\log_{10}\left(\frac{p_c}{2}\right) \end{aligned} \qquad (7)$$

$$\begin{aligned} SIR &= 10\log_{10}\left(sir\right) \\ &= 10\log_{10}\left(\frac{p_s}{p_w + p_c}\right) = 10\log_{10}\left(\frac{p_s}{2 + p_c}\right) \end{aligned} \qquad (8)$$

In figure 1 the PSD of the clutter and the target are represented for CNR=20dB and SIR=4 and -8 dB. A CNR equal to 20 dB represents a case where clutter is not dominant. The environment dominated by clutter can be studied without considering thermal noise. The two values of SIR are the extreme ones that have been selected for presenting the results. On the other hand, the target mean Doppler frequency normalized to the radar PRF is equal to 0,25, an intermediate value between 0 (target and clutter PSDs are superimposed) and 0.5 (the higher difference between target and clutter mean Doppler frequencies). The one lag correlation coefficients of the target and the clutter are equal to 0.9.

In figure 2, clutter and target PSDs for CNR=20dB, SIR=-8 dB, f_s = 0,25 and different one lag correlation coefficients are represented.

Figures 1 and 2 show how the power, the spectral width and the mean Doppler frequency of the clutter and the target are critical parameters that determine the detection capabilities of any detection scheme, including the Neyman-Pearson detector.

In the next section, the Neyman-Pearson detector is obtained and studied as a previous step for designing the NN based detector.

2.1 Neyman-Pearson decision rule

The pdf of a *n*-dimension complex gaussian random vector, **z**, with zero mean and covariance matrix **C** is expressed as:

$$f\left(\mathbf{z}\right)=\frac{1}{\pi^{n}\det\left(\mathbf{C}\right)}\exp\left(-\mathbf{z}^{T}\mathbf{C}^{-1}\mathbf{z}^{*}\right) \qquad (4)$$

The Neyman-Pearson detection rule is the result of comparing the likelihood ratio, or any monotonic function of it, to a detection threshold fixed attending to P_{FA} requirements (5).

$$\Lambda\left(\mathbf{z}\right)=\frac{f\left(\mathbf{z}\,|\,H_1\right)}{f\left(\mathbf{z}\,|\,H_0\right)} \qquad (5)$$

The detection problem likelihood functions can be obtained substituting the covariance matrixes (2) and (3) in (4) for the alternative (H_1) and null (H_0) hypothesis, respectively. The decision rule based on the likelihood ratio is presented (6)

$$\frac{\dfrac{1}{\pi^{n}\det\left(\mathbf{M}_s+\mathbf{M}_n\right)}\exp\left(-\mathbf{z}^{T}\left(\mathbf{M}_s+\mathbf{M}_n\right)^{-1}\mathbf{z}^{*}\right)}{\dfrac{1}{\pi^{n}\det\left(\mathbf{M}_n\right)}\exp\left(-\mathbf{z}^{T}\mathbf{M}_n^{-1}\mathbf{z}^{*}\right)}\ \underset{H_o}{\overset{H_1}{\gtrless}}\ \eta_0\left(P_{FA}\right) \qquad (6)$$

Decision rule (6) can be simplified as:

$$\mathbf{z}^{T}\left[M_n^{-1}-\left(M_n+M_s\right)^{-1}\right]\mathbf{z}^{*}=\mathbf{z}^{T}\mathbf{Q}\mathbf{z}^{*}\ \underset{H_o}{\overset{H_1}{\gtrless}}\ \eta'\left(P_{FA}\right) \qquad (7)$$

2.2 Neyman-Pearson detector performance

The false alarm probability (P_{FA}) and the detection porbability (P_D) of the decision rule (7) have been calculated in [12]. They can be obtained from the pdf of the test variable q = $\mathbf{z}^T\mathbf{Q}\mathbf{z}^*$ in the hypotesis H_0 and H_1 (7):

$$\begin{aligned}\int_T^{\infty} f_q\left(q\,|H_0\right)dq &= P_{FA}\\ \int_T^{\infty} f_q\left(q\,|H_1\right)dq &= P_D\end{aligned} \qquad (7)$$

Using the inverse transform of the Fourier domain characteristic of $f_q(q)$, the probability of exceeding a threshold T can be expressed as:

$$\int_T^{\infty} f_q(q)dq=\sum_{i=1}^{N}\exp\left[-\frac{T}{\lambda(i)}\right]\times\prod_{\substack{k=1\\k\neq i}}^{N}\frac{\lambda(i)}{\lambda(i)-\lambda(k)}U\left[\lambda(i)\right] \quad (8)$$

P_{FA} can be calculated from, (8) if λ(i) are the eigenvalues of $\mathbf{M_n}^*\mathbf{Q}^*$**,** while for obtaining P_D, λ(i) must be the eigenvalues of $\mathbf{(M_n+M_s)}^*\mathbf{Q}^*$. As U(·) is the unit step, the sum is extended to the terms with positive eigenvalues.

In figures 3 and 4 the ROC curves of the Neyman-Pearson detector for the cases considered in figures 1 and 2, are presented. Note that the detection loss that is observed when the clutter spectrum is wider, is higher than that observed when the target spectrum is wider.

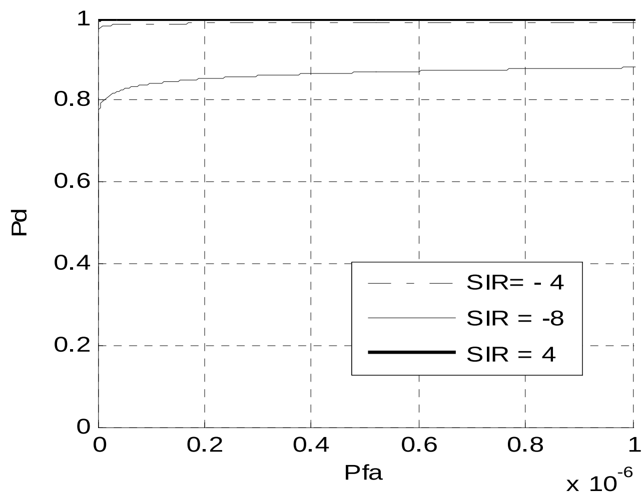

Fig. 3. Neyman-Pearson detector ROC curves for the cases considered in figure 1 (CNR=20 dB, SIR= 4 and -8 dB, ρ_c=ρ_s=0.9 and f_s=0.25)

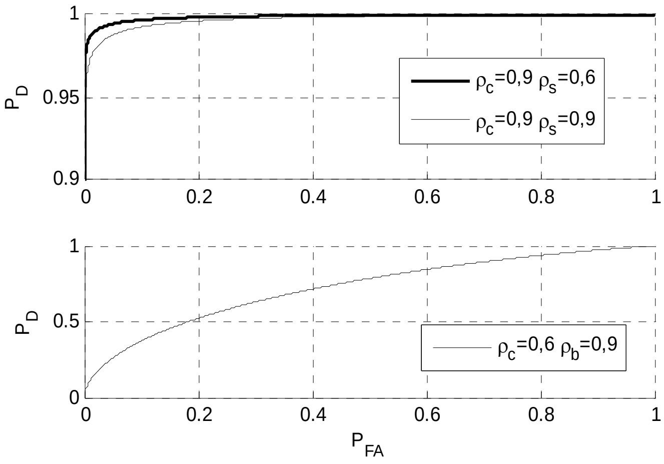

Fig. 4. Neyman- Pearson detector ROC curves for the cases considered in figure 2 (CNR=20 dB, SIR= -8 dB, fs=0.25 and different one lag correlation coefficients)

3 Design of experiments and results

In Air Traffic Control radar, the usual number of collected pulses in a scan is *n*=8. So this will be the assumed value in all the experiments.

As we have selected MLPs that woks with real inputs and real arithmetic, the complex vectors must be transformed in 2*n*-dimension real ones. It is assumed that each real input vector is composed of the real parts (the first *n* samples) and the imaginary parts (the remaining *n* samples) of the samples of the complex patterns. So MLPs with 2n=16 inputs have been trained.

MLPs are followed by a hard threshold detector whose threshold is fixed after training attending to P_{FA} requirements. If the NN output is greater than the threshold, H_1 is accepted, in other case, H_0 is accepted.

For designing the Neyman-Pearson detector, a set of parameters have to be known: CNR, SIR, f_s ρ_s and ρ_c.

The detector that implements rule (6), or any other equivalent to it, will be optimum if the parameters that characterize both hypothesis are those assumed during the design. If this is not the case, the designed detector will not be optimum and a loss in detection capability will be observed.

When training the NN, these design parameters are called training parameters. The NN approximation error is expected to be a function of the selected training parameters, and in order to evaluate this dependence different sets of training parameters have been selected. The Training Signal to Noise ratio has been varied from -8dB to 4dB, while one lag correlation coefficients ranging from 0.6 to 0.9 have been considered. Finally, only results for f_s equal to 0.25 and 0.5 are presented.

For each set of training parameters, separated training and validation sets composed of 50,000 randomly distributed patterns from H_0 and H_1 have been generated.

In all cases, results are presented for MLPs with one hidden layer of four neurons. This number has been determined after a exhaustive trial and error process as a compromise between complexity and performance improvement. In order to speed up the training process of the MLPs, the hyperbolic tangent sigmoid transfer function has been used

NNs have been trained for minimizing the LMSE, using Levenberg-Madquardt algorithm (LM) (with adaptive parameter) [9]. The LM is based on the Newton method. The weights adaptation rule of the Newton method is expressed in (9).

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta \cdot \mathbf{H}_k^{-1} \cdot \mathbf{g}_k \qquad (9)$$

where **H** is the Hessian matrix (second derivatives) of the performance index at the current values of the weights. Unfortunately, it is complex and expensive to compute the Hessian matrix for MLPs. Actually, the LM is based on the Gauss-Newton method, which has been designed specifically for minimizing the LMSE. In this case, as the error function has the form of a sum of squares, the Hessian matrix can be approximated as $\mathbf{H}=\mathbf{J}^{-T}\bullet\mathbf{J}$, where **J** is the Jacobian matrix that contains the first derivatives of the network errors with respect to the weights. The LM algorithm actualization rule is expressed in (10).

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \left(\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}\right)^{-1} \cdot \mathbf{J}^T\mathbf{e} \qquad (11)$$

When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function will always be reduced at each iteration of the algorithm.

A cross-validation technique has been used to avoid over-fitting and all NNs have been initialized using the Nguyen-Widrow method [13]. For each case, the training process has been repeated ten times. Only the cases where the performances of the ten trained networks were similar in average, have been considered to extract conclusions.

Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp259-264)

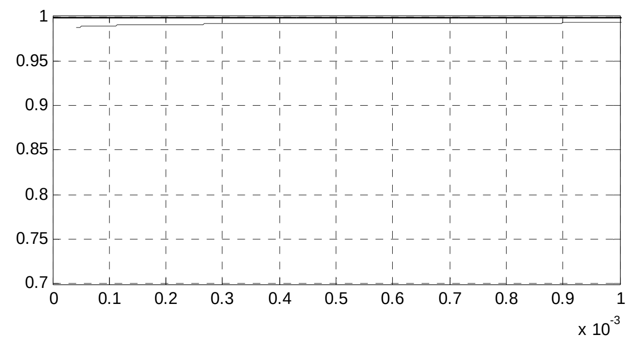

Fig. 5. ROC curves for the Neyman-Pearson detector (wider line) and the MLP one (thinner line) for CNR=20 dB, TSIR=4 dB, ρ_s=0.6, ρ_c=0.9 and f_s=0.25.

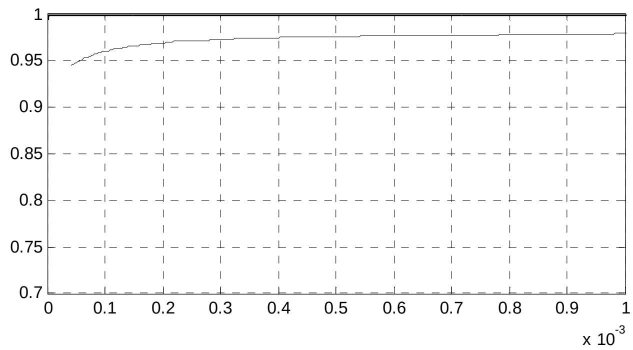

Fig. 6. ROC curves for the Neyman-Pearson detector (wider line) and the MLP one (thinner line) for CNR=20dB, TSIR=0 dB, ρ_s=0.6, ρ_c =0.9 and f_s=0.25.

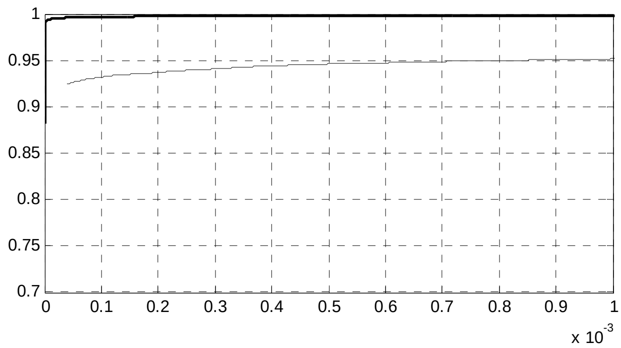

Fig. 7. ROC curves for the Neyman-Pearson detector (wider line) and the MLP one (thinner line) for CNR=20dB, SIR=-4 dB, ρ_s =0.6, ρ_c =0.9 and f_s=0.25.

Since for radar applications only low P_{FA} values are of interest, results are presented for P_{FA} lower than 10^{-3}. These values have been estimated using Importance Sampling Techniques (relative error lower than 10% in the presented results) [10,11]. The selected strategy involves modifying the covariance matrix under hypothesis Ho, multiplying the original one by a small factor K=1.1:

$$\mathbf{M}_n^{'} = K\mathbf{M}_n \tag{10}$$

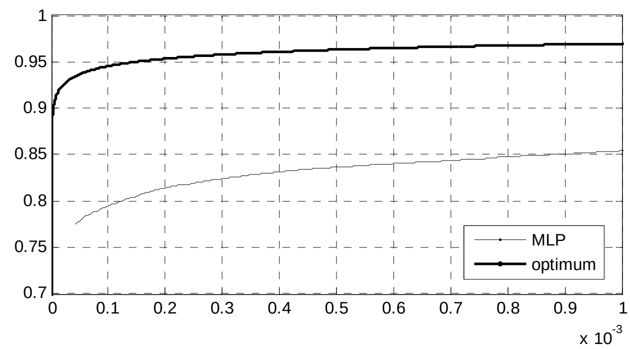

Fig. 8. ROC curves for the Neyman-Pearson detector (wider line) and the MLP one (thinner line) for CNR=20dB, TSIR =-8dB, ρ_s=0.6, ρ_c=0.9 and f_s=0.25.

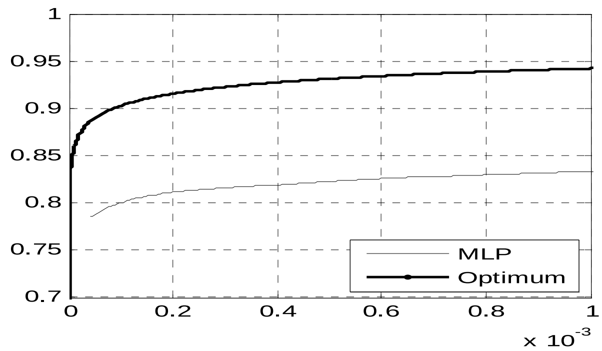

Fig. 9. ROC curves for the Neyman-Pearson detector and the MLP one for CNR=20dB, TSIR=-8 dB, ρ_s=0.9, ρ_c=0.9 and f_s=0.25.

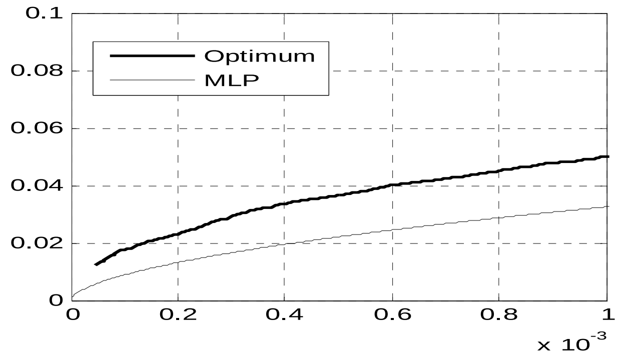

Fig. 10. ROC curves for the Neyman-Pearson detector and the MLP one for CNR=20 dB, TSIR =-8 dB, ρ_s =0.9, ρ_c =0.6 and f_s=0.25.

P_D values have been estimated using conventional Montecarlo simulation. Using sets of 80,000 patterns generated under hypothesis H_1, an estimation error lower than 1% is guaranteed in all the results.

In figures 5-8, the dependence of MLPs performance on TSIR is studied. Results show that the MLP approximation error increases when the TSIR decreases. This behaviour can be explained taking into consideration that weak targets are more difficult to characterize. The worst behavior is observed for the lowest TSIR, -8 dB: the P_D is lower for the same $P_{FA.}$

In figures 9 and 10, the influence of target and clutter one lag correlation coefficients is studied. The fixed clutter and target characteristics are: CNR=20 dB, TSIR =-8 dB and f_s=0.25. Due to the great detection performance decrease, in figure 10 the scales have had to be changed. The clutter power is much higher than the target one, so the detector performance is more sensible to variations in clutter correlation characteristics.

4 Conclusion

The application of neural networks for detecting Rayleigh radar targets with gaussin ACF in gaussian clutter with gaussian ACF plus white noise has been considered. The Neyman-Pearson detector and its ROC curves have been calculated. The influence of design parameters such as the TSIR and one-lag correlation coefficients of target and clutter has been studied.

MLPs with one hidden layer of four neurons have been trained using the LM algorithm. Results show that an MLP can approximate the Neyman-Pearson detector for very low P_{FA} values and a wide range of TSIR values, but the approximation error increases when the TSIR decreases, because weaker targets are more difficult to characterize.

When ρ_s or ρ_c are varied, results show that the sensibility of MLP performance is lower that that observed when varying the TSIR.

The obtained results prove the possibility of using NN to approximate the optimum detector without knowledge of the statistical properties of the detection problem, making use of sets of pre-classified patterns.

References:

[1] Van Trees, H.L, *Detection, estimation, and modulation theory*. Vol. 1. Willey. (1968)

[2] Ruck, D.W. et al, The multilayer perceptron as an aproximation to a Bayes optimal discriminant function, *IEEE Trans. on Neural Networks*. Vol. 1, No. 1, 1990, pp. 296-298.

[3]Wan, E.A, Neural network classification: A Bayesian interpretation. *IEEE Trans. on Neural Networks*. Vol. 1, No. 1, 1990, pp. 303-305.

[4] Gandhi, P. P., Ramamurti, V, Neural Networks for Signal Detection in Non-Gaussian Noise. *IEEE Trans. on Signal Proc.* Vol. 45, No. 11, 1997, pp. 2846-2851.

[5] Andina, D., Sanz-Gonzalez, J.L: Comparison of a neural network detector vs Ney-man-Pearson optimal detector. *Proc. of ICASSP-96*, USA, 1996, pp. 3573-3576.

[6] D. de la Mata-Moya, P. Jarabo-Amores, M. Rosa-Zurera, F.López-Ferreras, R. Vicen-Bueno. Approximating the Neyman-Pearson detector for Swerling I targets with low complexity neural networks. *Lecture Notes in Computer Science. Artificial Neural Networks: Formal Models and Their Applications*. No. 3697, 2005, pp. 917-922

[7] P. Jarabo-Amores, R. Gil-Pita, M.Rosa Zurera, F. López Ferreras, MLP and RBFN for detecting white gaussian signals in white gaussian interference, *Lecture Notes in Computer Science*. No. 2687, 2003, pp. 790-797.

[8] Skolnik, M, *Radar Handbook, Second Edition*. McGraw-Hill, Inc., 1990.

[9] Hagan, M.T., Menhaj, M.B, Training feedforward networks with the Marquardt algorithm. *IEEE Trans. on Neural Networks*, Vol. 5, No. 6, 1994, pp. 989-993.

[10] Grajal, J., Asensio, A: *Multiparametric importance sampling for simulation of radar systems*. IEEE Trans. on Aerospace and Electronic Systems, Vol. 35, No. 1,1999, pp. 123-137.

[11] Sanz-Gonzalez, J. L, Andina, D: *Performance analysis of neural network detectors by importance sampling techniques*. Neural Proc. Letters, No. 9, (1999) 257-269. S. Haykin. *Neural Networks a comprehensive foundation*. 1999.- Prentice-Hall, Inc.

[12] V. Aloisio, A. di Vito, G. Galati: *Optimum detection of moderately fluctuating radar targets*. IEEE Proc Radar, Sonar and Navigation, Vol. 141, No. 3, 1994, pp. 164-170.

[13] Nguyen, D., Widrow, B: *Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights*. Proc. of the Int. Joint Conf. on Neural Networks, Vol 3, 1990, pp. 21-26.