Deep Learning for Radar Signal Detection in Electronic Warfare Systems

Mustafa Atahan Nuhoglu[1], Yasar Kemal Alp[2]
Radar, Electronic Warfare and Intelligence Systems Division
ASELSAN A.S.
Ankara, Turkey
Email: {[1]manuhoglu, [2]ykalp}@aselsan.com.tr

Fatih Cagatay Akyon
Electrical and Electronics Engineering
Ihsan Dogramaci Bilkent University
Ankara, Turkey
Email: akyon@ee.bilkent.edu.tr

Abstract—Detection of radar signals is the initial step for passive systems. Since these systems do not have prior information about received signal, application of matched filter and general likelihood ratio tests are infeasible. In this paper, we propose a new method for detecting received pulses automatically with no restriction of having intentional modulation or pulse on pulse situation. Our method utilizes a cognitive detector incorporating bidirectional long-short term memory based deep denoising autoencoders. Moreover, a novel loss function for detection is developed. Performance of the proposed method is compared to two well known detectors, namely: energy detector and time-frequency domain detector. Qualitative experiments show that the proposed method is able to detect presence of a signal with low probability of false alarm and it outperforms the other methods in all signal-to-noise ratio cases.

Index Terms—passive systems, detection, deep learning, long-short term memory, autoencoder

I. INTRODUCTION

In an electronic warfare (EW) environment, passive systems, such as electronic intelligence (ELINT), electronic support measures (ESM), aim to extract useful information from received signals. The utilization process in these systems involves detection of signals first and then analysis of the detected signals [1]. Since neither the received signal nor its type is known before interception, application of matched filter is infeasible. In addition, general likelihood ratio tests are not also applicable since they require prior knowledge about the modulation type of the received signal [2].

In contrast to aforementioned methods, energy detector is practicable in these systems as it does not have any assumptions about modulation type, signal shape or arrival time to make a detection. In [3], energy detector is used to detect spread-spectrum signals. It is utilized to detect optimal frequency band sensing time for cognitive radios in [4]. As a drawback, performance of this method gets degraded in low signal-to-noise ratio (SNR) scenarios. Another detector type that does not have such assumptions is based on time-frequency domain analysis. In [5], a detector based on Wigner-Ville distribution is proposed to detect transient regions of signals by weighting the test statistics. Compared to the energy detector, this method suffers from high complexity.

Passive detection of radar signals is a time-series problem. Special neural network architectures called recurrent neural networks (RNN), are designed to solve time-series problem by learning in a recursive manner. They have been successfully applied to several signal processing problems of areas including computer vision, text analysis, acoustic and radar [6]–[9]. Furthermore, to solve learning of long time series, [10] proposed a special architecture for RNN, called long-short term memory (LSTM). Its bidirectional version (BLSTM) has both forward and backward pass, which leads to deduction of non-causal information as well [11]. In [12], LSTM networks are used to detect real-life voice activity. Due to generalization ability of neural networks, this approach performs well even in settings where high noise background exists. [13] utilized stacked LSTM networks to make anomaly/fault detection in different time series problems including ECG, space shuttle, power demand and multi-sensor engines. In [14], polyphonic sound events are detected in real life recording with the usage of BLSTM and [15] proposed sliding BLSTM to make detections without prior knowledge of channel state information in communication systems.

To the best of our knowledge, there exists no work regarding radar signal detection in passive systems by using a denoising neural network architecture. To fill the gap in the literature and to overcome the drawbacks mentioned in the existing state-of-the-art detectors, an RNN based denoising autoencoder is proposed in this work. The proposed method uses the last layer output of BLSTM based deep denoising autoencoder as a test statistic. To train the network in the sense of detection, a novel loss function that is based on binary cross entropy is developed. Conducted experiments show that the proposed method is capable of detecting radar signal presence automatically with low probability of false alarm even in low SNR scenarios, while outperforming the other detectors in all SNR levels.

The paper is organized as follows: Section II describes signal model, Section III introduces existing detectors, Section IV explains the proposed method. Section V demonstrates the simulation results and the last section is spared for conclusions.

II. SIGNAL MODEL

Complex IQ (Inphase / Quadrature) baseband samples of a radar pulse is defined as:

$$x(t_n) = a(t_n)e^{j\phi(t_n)} + u(t_n) \quad n = 1, 2, ..., N . \quad (1)$$

Here $\phi(t_n)$ is the instantaneous phase, $u(t_n)$ is the circularly symmetric complex Gaussian (CSCG) noise with zero mean and σ_u^2 variance, t_n denotes the sampling instants and N is the total number of samples. $a(t_n)$ is the instantaneous magnitude of the pulse and it is defined as:

$$a(t_n) = \begin{cases} 1, & T_0 \le t_n < T_0 + T_g \\ e^{-(t_n-T_0)^2/\sigma_1^2}, & t_n < T_0 \\ e^{-(t_n-T_0-T_g)^2/\sigma_2^2}, & T_0 + T_g \le t_n \end{cases} \quad (2)$$

where T_0 is the time of arrival of the pulse, T_g is the pulse width. Note that, the rising and falling regions of $a(t_n)$ are modeled as a Gaussian functions with σ_1^2 and σ_2^2 variances, respectively [16].

If the pulse has phase hoppings, which is called phase modulation on pulse (PMOP), instantaneous phase can be expressed as:

$$\phi(t_n) = \begin{cases} 2\pi f_c t_n + \phi_0, & t_n < t_{\phi_1} \\ 2\pi f_c t_n + \phi_0 + \phi_{\Delta_1}, & t_{\phi_1} < t_n < t_0 + t_{\phi_2} \\ \vdots \\ 2\pi f_c t_n + \phi_0 + \sum_{k=1}^{K-1} \phi_{\Delta_k}, & t_{\phi_{K-1}} < t_n < t_{\phi_K} \\ 2\pi f_c t_n + \phi_0 + \sum_{k=1}^{K} \phi_{\Delta_k}, & t_n > t_{\phi_K} . \end{cases} \quad (3)$$

Here, f_c is the initial center frequency, t_{ϕ_k} is the time where k-th phase hopping occurs, ϕ_0 is the initial phase in radian, ϕ_{Δ_k} is the amount of phase hopping. Similarly, intentional frequency changes in a pulse, which is called frequency modulation on pulse (FMOP), can be defined as:

$$\phi(t_n) = \begin{cases} 2\pi f_c t_n + \phi_0, & t_n < t_{f_1} \\ 2\pi(f_c + f_{\Delta_1})t_n + \phi_0, & t_{f_1} < t_n < t_0 + t_{f_2} \\ \vdots \\ 2\pi(f_c + \sum_{l=1}^{L-1} f_{\Delta_k})t_n + \phi_0, & t_{f_{L-1}} < t_n < t_{f_L} \\ 2\pi(f_c + \sum_{l=1}^{L} f_{\Delta_k})t_n + \phi_0, & t_n > t_{f_L} \end{cases} \quad (4)$$

where t_{f_k} is the time where k-th frequency hopping occurs and f_{Δ_k} is the amount of frequency hopping.

Two probabilities are of interest for detecting a radar pulse. Probability of detection, which is the probability of making a detection when the signal is present ($\mathcal{H}_1$ hypothesis), and probability of false alarm, which is making a detection in the absence of the signal ($\mathcal{H}_0$ hypothesis). In the next section, a brief review of the state-of-the-art detectors will be given.

III. EXISTING METHODS

In this paper, energy detector and time-frequency domain detector are considered for comparison.

A. Energy Detector (ED)

The test statistic for energy detector is:

$$T(x) = \frac{1}{N} \sum_{n=1}^{N} |x(t_n)|^2 , \quad (5)$$

where N is the window length. Under hypothesis $\mathcal{H}_0$, magnitude square of the received signal: $s_n = |x(t_n)|^2 = x^I(t_n)^2 + x^Q(t_n)^2$ is chi-square distributed, and its probability density function is calculated as [17]:

$$p(s|\mathcal{H}_0) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} s^{\frac{\nu}{2}-1} e^{-\frac{s}{2}}, & s \ge 0 \\ 0, & s < 0 \end{cases} \quad (6)$$

where $\nu = 2$ indicates the degree of freedom. Probability of false alarm is then formulized as [18]:

$$P_{FA} = \int_\gamma^\infty p(s|\mathcal{H}_0)ds$$
$$= e^{-\frac{\gamma}{\sigma^2}} \sum_{k=0}^{N-1} \frac{(\frac{\gamma}{\sigma^2})^k}{k!} \quad (7)$$

where γ is the detection threshold. Here, N must be carefully chosen in order not to miss signals with minimum pulse width that is required to be detected. Assuming that minimum pulse width to be detected is PW_{min}, then N should be selected as the maximum integer not greater than $PW_{min}f_s$ where f_s is the sampling frequency.

B. Time-Frequency Domain Detector (TFDD)

TFDD transforms received signal to complex valued time-frequency images and make detections over these 2D images. There are several time-frequency distribution functions, such as wavelet, Wigner-Ville and short-time Fourier transform (STFT). Among various time-frequency distributions, STFT has no cross terms and lowest complexity which is highly demanded in real hardware applications, for instance ELINT receivers [19]. Therefore, for performance comparisons in the extended simulations we conducted, the detection rule proposed in [5] is adapted to the detector scheme using STFT, which is abbreviated as STFTD. For a window with N samples, the test statistic is described as:

$$T = \sum_{n=1}^{N} \sum_{k=1}^{K} e_k(n)p_k(n) \quad (8)$$

where $p_k(n)$ is the energy of n-th transformed signal sample at k-th frequency subband of total K subbands. The corresponding weight $e_k(n)$ is defined as:

$$e_k(n) = \frac{\sum_n p_k(n)}{\sum_k \sum_n p_k(n)} . \quad (9)$$

This choice of weight function puts emphasis on subbands with higher energy content. Therefore, (8) behaves as a frequency domain matched filter.

Finally, both of the detectors make detections by comparing the test statistics to the threshold γ:

$$\begin{cases} \mathcal{H}_0, & T < \gamma \\ \mathcal{H}_1, & T \geq \gamma \ . \end{cases}$$

IV. PROPOSED METHOD (BLSTMD)

The proposed method uses denoising autoencoder output as a test statistic, which is then compared to a predefined threshold for detection. Denoising autoencoders are used for fitting a noisy version of data by learning a nonlinear function [20]. They consist of two parts: encoder and decoder. Assume that $f(.)$ is the encoding function of noisy input data $\tilde{x}$, and h is the encoded data, i.e. $h = f(\tilde{x})$, then the decoder output can be described as $y = g(f(\tilde{x}))$ where $g(.)$ is the decoding function. They minimize $\mathcal{L}(x, y)$, where x is the noiseless input data and $\mathcal{L}$ is a loss function such as the ℓ_2 norm of their difference. In the proposed method, BLSTM is utilized as both the encoding and decoding mechanisms for the autoencoder. BLSTM is acyclic, meaning that forward and backward layers are independent. The input sequence is given to the forward layer and its reversed version is given to the backward layer. Resulting outputs are concatenated and given to the next layer as input. An LSTM unit includes four gates having different purposes and operates by making the following computations:

$$g_i^t = \sigma(b_i^g + \sum_j W_{i,j}^g s_j^t + \sum_j U_{i,j}^g a_j^{t-1}) \tag{10}$$

$$f_i^t = \sigma(b_i^f + \sum_j W_{i,j}^f s_j^t + \sum_j U_{i,j}^f a_j^{t-1}) \tag{11}$$

$$\tilde{c}_i^t = tanh(b_i^c + \sum_j W_{i,j}^c s_j^t + \sum_j U_{i,j}^c a_j^{t-1}) \tag{12}$$

$$c_i^t = f_i^t \cdot c_i^{t-1} + g_i^t \cdot tanh(\tilde{c}_i^t) \tag{13}$$

$$q_i^t = \sigma(b_i^q + \sum_j W_{i,j}^q s_j^t + \sum_j U_{i,j}^q a_j^{t-1}) \tag{14}$$

$$a_i^t = q_i^t \cdot tanh(c_i^t) \ . \tag{15}$$

Here g, f, c, and q denote the results of update, forget, cell and output gates, respectively. $\tilde{c}$ is an intermediate step variable called candidate cell and a is the activation value of the unit. $W_{i,j}$ denotes the LSTM kernel parameter for i-th unit j-th feature, which is multiplied with j-th feature of input data: s_j^t. In our case, the input data s_j^t for the first layer corresponds to $|x(t_n)|$ in (5), which is the magnitude of pulse amplitude. $U_{i,j}$ is the recurrent kernel which is multiplied with the previous activation value a_j^{t-1}.

Fig. 1 shows the proposed model for denoising and detection of the received signals. The encoder part learns time series features, while the decoder part deduces information from the encoded signal so that the noisy signal is mapped to the desired signal that enables cognitive detection. The last layer, LSTM, is used to reduce the output feature number to one so that fitting is possible to the label, which is a binary indicator for signal presence. Therefore, different from the regular denoising autoencoders, the proposed network
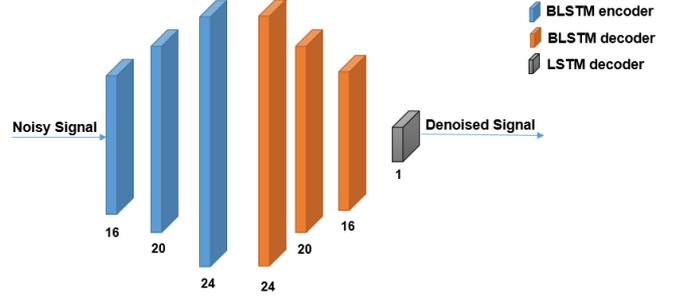

Fig. 1. Proposed denoising network scheme is shown. The numbers below the layers indicate unit sizes.

minimizes $\mathcal{L}(x^o, y)$, where x^o denotes the binary label. Note that the network can input dynamic size data, which means that it does not include a weight matrix depending on input size. Hence, using recurrent layers makes the network to be more generalizing and to be easily implementable in a real hardware.

V. RESULTS AND ANALYSIS

BLSTMD, ED and STFTD performances are evaluated on synthetic radar signals by receiver-operating characteristic (ROC) analysis for varying SNR values. Modulation types used in [21] are considered. There are 23 different modulation types in our dataset, which are shown in Table I.

TABLE I
MODULATION TYPES USED IN SIMULATION RESULT SETS

23 class set	
SCM	8-PSK
+Ramp FM	16-PSK
-Ramp FM	Frank Code
Sinusoidal FM	P1 Code
Triangular FM	P2 Code
Costas-5 FM	P3 Code
Costas-7 FM	P4 Code
Costas-10 FM	T1 Code
Barker-3 PM	T2 Code
Barker-7 PM	T3 Code
Barker-13 PM	T4 Code
QPSK	

In order to train the autoencoder network, 100000 synthetic multi pulses sampled at 100 MHz with different pulse widths from the interval $[0.04, 1000]$ μs are created and concatenated. This causes window length of the ED to obey $N \leq 4$. The best option $N = 4$ case is chosen for comparison. Rising and falling region durations of the pulses are randomly selected from the interval $[0, 0.2]$ μs. The proposed model is created and trained in TensorFlow/Python. Labels given to the autoencoder depend on whether the sample amplitude is larger than -3 dB or not. If the sample amplitude is above -3 dB, it is labeled as 1; it is labeled as 0 otherwise. ADAM optimizer,
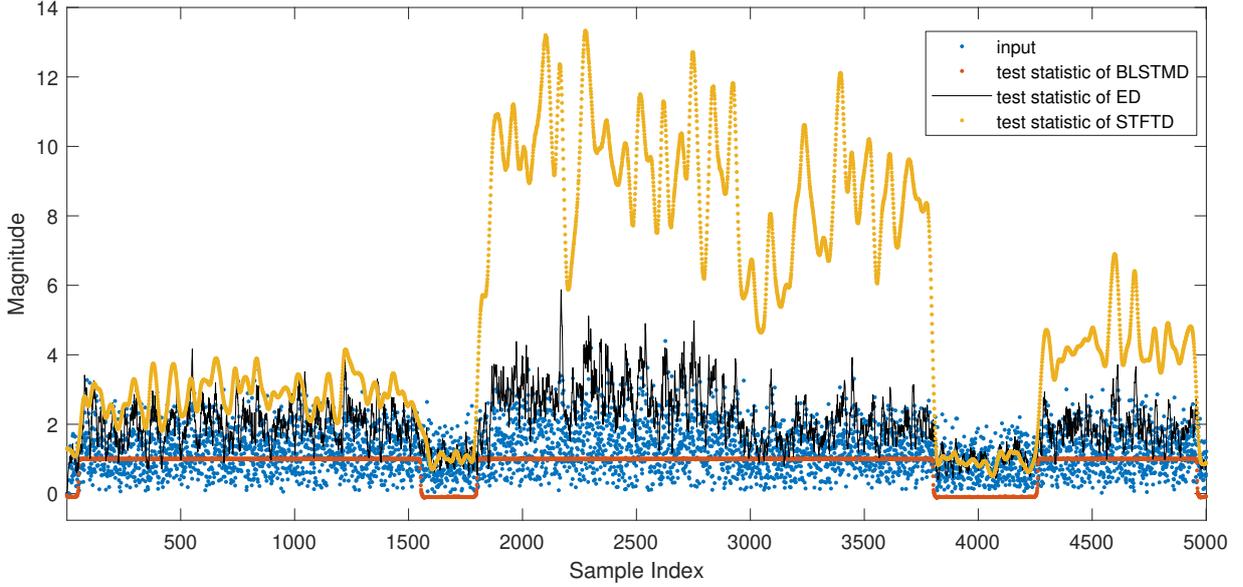
Fig. 2. Test statistics of the detectors at 5 dB SNR. Blue: received noisy signal (input), red: test statistic of BLSTMD, black: test statistic of ED, yellow: test statistic of STFTD.

which is highly used in optimization of deep neural networks [22], is preferred for training of the proposed model. Since sharp notches and peaks in the test statistic result in estimating multi-pulses out of a single pulse (in other words, cracks), we propose a new loss function that is specially designed for detection purposes. The loss function is an updated version of binary cross entropy, which is defined as:

$$\mathcal{L} = \sum_{n=0}^{N-1} \big(- x_n^o \log(y_n) - (1 - x_n^o) \log(1 - y_n)$$
$$+ \max(0, |2y_n - y_{n-1} - y_{n+1}| - 1))$$
(16)

where y_n is the output of the network (test statistic) for n-th sample of the received signal, and x_n^o is the corresponding label. The custom term in the proposed loss function includes $\max(.)$ operation, which penalizes certain combinations of the three sequential outputs of the network. Table II demonstrates all of the combinations for the three samples and the corresponding values of the custom term in $\mathcal{L}$, which is called $\mathcal{L}_{custom}$. As seen, only two combinations $(0, 1, 0)$ and $(1, 0, 1)$, which are related to undesired sharp peaks and notches in the test statistic, exhibit additional cost. Hence, the proposed loss function targets to create a Markov model-like transition matrix behaviour in a way that certain transitions are penalized. Note that, the test statistic values y_n are not actually binary due to the non-linear layers of the network. However, it is aimed to keep them close to 0 and 1 by training with binary labels. The proposed loss function penalizes sharp changes in the test statistic and makes thresholding more accurate. Lastly, batch size is chosen as 32. 80, 10 and 10 percent of 100000 data is used for training, validation and test, respectively.

TABLE II
CUSTOM LOSS VALUES FOR DIFFERENT COMBINATIONS OF SAMPLES

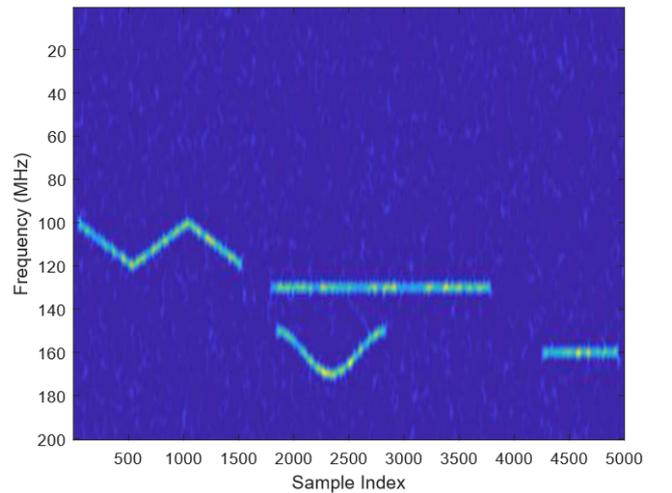| y_{n-1}	y_n	y_{n+1}	$\mathcal{L}_{custom}$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	0

Fig. 3. STFT of the input data at 5 dB SNR.

Fig. 4. ROC analysis of the detection methods at -5 dB SNR.

Fig. 5. ROC analysis of the detection methods at 0 dB SNR.

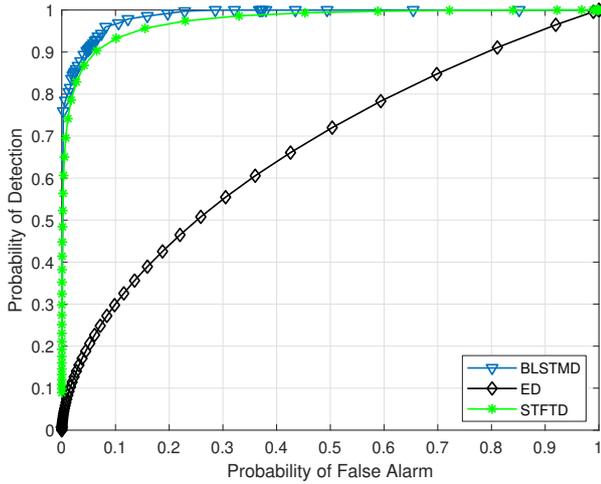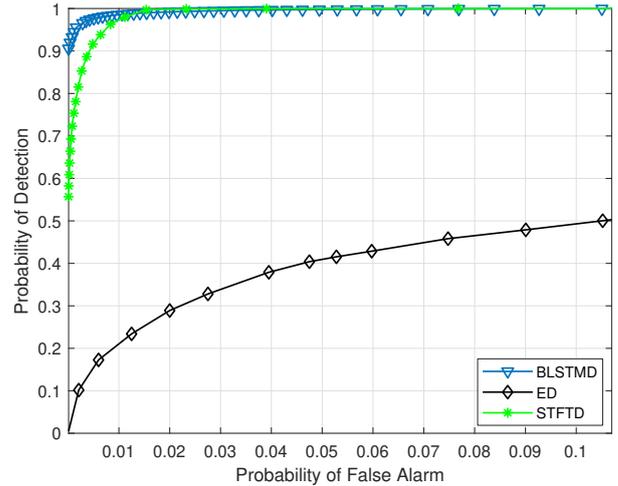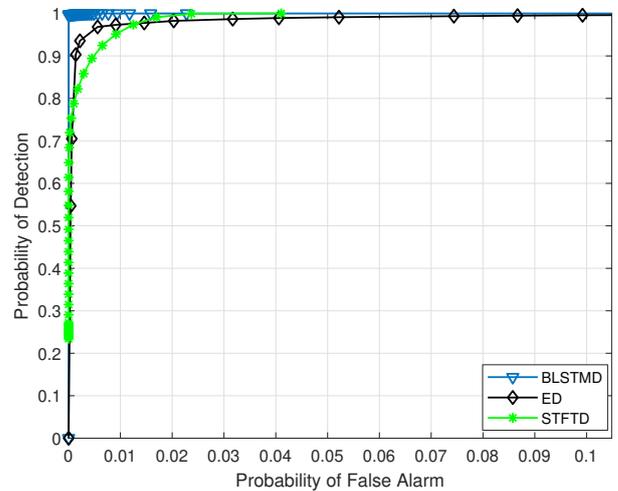Fig. 6. ROC analysis of the detection methods at 5 dB SNR.

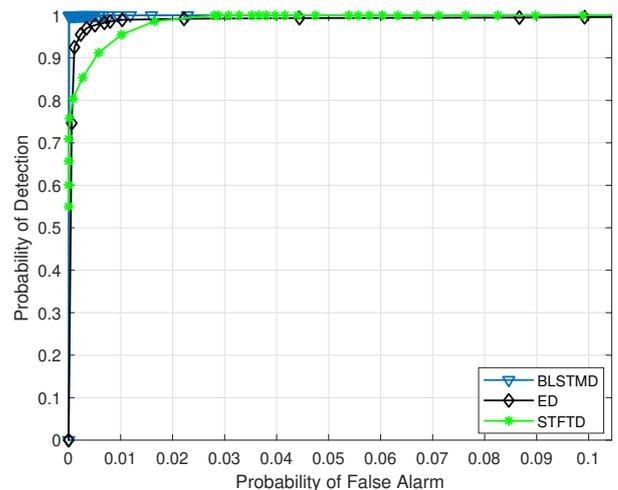Fig. 7. ROC analysis of the detection methods at 10 dB SNR.

In Fig. 2, we examine amplitude of the received noisy synthetic input data, whose STFT is demonstrated in Fig. 3, and test statistics of three detectors at 5 dB SNR. The best performing STFT window length and intersection ratios are selected for STFTD to have a fair comparison. The data includes different types of modulations. In Fig. 3, triangular and half sinusoidal shapes indicate pulses with FMOP, straight lines indicate pulses with constant frequencies. Also, pulse on pulse (POP) situation can be seen between 2000-th and 2800-th samples. As observed, frequency bands with higher energy contents results in higher test statistic for STFTD, which makes detection easier. This can be seen from Fig. 2. Notice that STFTD test statistic does not drop much just after the instant where POP situation happens. This is because pulses with constant frequencies have larger weighted energies as explained in (9). ED has a test statistic that is highly variational due to small window length. Lastly, the proposed BLSTMD correctly fits signal presence and absence to 1 and 0, respectively. Therefore, it can be seen that the proposed detector is able to denoise data without depending on frequency domain analysis or POP situation.

Fig. 4 shows ROC analysis of the detectors at -5 dB SNR. This is a very low SNR level since ELINT/ESM systems generally work around 10 dB SNR. As observed, BLSTMD clearly outperforms ED and STFTD by having the highest probability of detection at all probability of false alarm levels. The denoising feature of autoencoders provides the network a desired test statistic, resulting in easier seperation of signal presence and absence regions while test statistic of ED causes incorrect thresholding. Similarly, STFTD performs also well because it includes Fourier transform operation, which increases SNR as in the case of matched filter.

Fig. 5, 6 and 7 show ROC performance of the detectors at 0, 5 and 10 dB SNR. To focus on the methods' low probability of false alarm performances, the figures are zoomed in x-axis where probability of false alarm is below 0.1. It is observed that the performance difference between the detectors

reduces as SNR increases. STFTD performs similar to the proposed method in all scenarios, while ED performance gets dramatically degraded as SNR decreases. Overall, the proposed method shows the best performance in all SNR levels.

VI. CONCLUSION

This work focuses on the detection of radar signals in electronic warfare systems. A novel detection method utilizing bidirectional long-short term memory based deep denoising autoencoder is proposed. Moreover, to train the network, a loss function that is specially designed for detection purposes is proposed. Qualitative experiments show that the proposed method outperforms energy detector and short-time Fourier transform detector according to receiver-operating characteristic analysis in different signal-to-noise ratio (SNR) level scenarios. It is also observed that denoising neural networks are highly capable of behaving like matched filters in terms of SNR maximization.

REFERENCES

[1] R. G. Wiley, *ELINT: The interception and analysis of radar signals*. Artech House, 2006.

[2] P. Q. Ly, S. Sirianunpiboon, and S. D. Elton, "Passive detection of bpsk radar signals with unknown parameters using multi-sensor arrays," in *Signal Processing and Communication Systems (ICSPCS), 2017 11th International Conference on*. IEEE, 2017, pp. 1–5.

[3] R. A. Dillard and G. M. Dillard, "Detectability of spread-spectrum signals," *Norwood, MA, Artech House, 1989, 157 p.*, 1989.

[4] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[5] N. K. Mishra and G. Anand, "Transient signal detection using wigner-ville distribution and wavelet denoising," in *TENCON 2006. 2006 IEEE Region 10 Conference*. IEEE, 2006, pp. 1–4.

[6] M. Liang and X. Hu, "Recurrent convolutional neural network for object recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3367–3375.

[7] I. Sutskever, J. Martens, and G. E. Hinton, "Generating text with recurrent neural networks," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 1017–1024.

[8] H. Sak, A. Senior, K. Rao, and F. Beaufays, "Fast and accurate recurrent neural network acoustic models for speech recognition," *arXiv preprint arXiv:1507.06947*, 2015.

[9] V. Jithesh, M. J. Sagayaraj, and K. Srinivasa, "Lstm recurrent neural networks for high resolution range profile based radar target classification," in *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT)*. IEEE, 2017, pp. 1–6.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.

[12] F. Eyben, F. Weninger, S. Squartini, and B. Schuller, "Real-life voice activity detection with lstm recurrent neural networks and an application to hollywood movies," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 483–487.

[13] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "Long short term memory networks for anomaly detection in time series," in *Proceedings*. Presses universitaires de Louvain, 2015, p. 89.

[14] G. Parascandolo, H. Huttunen, and T. Virtanen, "Recurrent neural networks for polyphonic sound event detection in real life recordings," *arXiv preprint arXiv:1604.00861*, 2016.

[15] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *arXiv preprint arXiv:1802.02046*, 2018.

[16] Y. K. Alp and G. Gök, "Estimation of inter-pulse phase/frequency stepping parameters in electronic intelligence systems," in *2018 26th Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2018, pp. 1–4.

[17] S. M. Kay, *Fundementals of statistical signal processing*. Prentice Hall PTR, 1993.

[18] M. Abramowitz and I. A. Stegun, "Handbook of mathematical functions with formulas, graphs, and mathematical table," in *US Department of Commerce*. National Bureau of Standards Applied Mathematics series 55, 1965.

[19] V. Choqueuse, M. Benbouzid, J.-F. Charpentier, G. Barakat *et al.*, "A comparative study of time-frequency representations for fault detection in wind turbine," in *IECON 2011*, 2011, pp. 3584–3589.

[20] H. Bourlard and Y. Kamp, "Auto-association by multilayer perceptrons and singular value decomposition," *Biological cybernetics*, vol. 59, no. 4-5, pp. 291–294, 1988.

[21] F. C. Akyon, Y. K. Alp, G. Gok, and O. Arikan, "Classification of intra-pulse modulation of radar signals by feature fusion based convolutional neural networks," in *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE, 2018, pp. 2290–2294.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.