This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Kong, Gyuyeol; Jung, Minchae; Koivunen, Visa

Waveform Recognition in Multipath Fading Using Autoencoder and CNN with Fourier Synchrosqueezing Transform

Published in:
2020 IEEE International Radar Conference, RADAR 2020

DOI:
[10.1109/RADAR42522.2020.9114783](https://doi.org/10.1109/RADAR42522.2020.9114783)

Published: 01/04/2020

Document Version
Peer reviewed version

Please cite the original version:
Kong, G., Jung, M., & Koivunen, V. (2020). Waveform Recognition in Multipath Fading Using Autoencoder and CNN with Fourier Synchrosqueezing Transform. In *2020 IEEE International Radar Conference, RADAR 2020* (pp. 612-617). [09114783] IEEE. <https://doi.org/10.1109/RADAR42522.2020.9114783>

Waveform Recognition in Multipath Fading using Autoencoder and CNN with Fourier Synchrosqueezing Transform

Gyuyeol Kong*, Minchae Jung[†], and Visa Koivunen*

*Department of Signal Processing and Acoustics, Aalto University, Finland

Email: {gyuyeol.kong and visa.koivunen}@aalto.fi

[†]Wireless@VT, Department of Electrical and Computer Engineering, Virginia Tech, VA, USA

Email: hosaly@vt.edu

Abstract—In this paper the problem of recognizing radar waveforms is addressed for multipath fading channels. Waveform classification is needed in spectrum sharing, radar-communications coexistence, cognitive radars, spectrum monitoring and signal intelligence. Different radar waveforms exhibit different properties in time-frequency domain. We propose a deep learning method for waveform classification. The received signal is first equalized to mitigate the effect of multipath fading channels by using a denoising auto-encoder (DAE). Then, the equalized signal is processed with Fourier synchrosqueezing transform that has excellent properties in revealing time-varying behavior, rate of, strength and number of oscillatory components in signals. The resulting time-frequency description is represented as a bivariate image that is fed into a convolutional neural network. The proposed method has superior performance over the widely used the Choi-Williams distribution (CWD) method in distinguishing among different radar waveforms even at low signal-to-noise ratio regime.

I. INTRODUCTION

We propose a machine learning approach for recognizing radar waveforms based on their time-frequency properties in multipath fading channels. Recognition of the employed waveforms is very important task in various spectrum management, surveillance and cognitive radar applications [1], [2]. Some machine learning techniques for the radar waveform recognition [1], [2], [3], [4], [5], [6] based on the Choi-Williams distribution (CWD) have been introduced. However, these methods consider waveform recognition only in additive white Gaussian noise (AWGN) channel. The actual wireless channels are subject to variety of multipath and fading effects that make the classification problem very challenging. In addition, these methods have the problem of the low discrimination capability between certain waveforms, for example in the case of P1 and P4 polyphase codes used in radars. Furthermore, the CWD neglects the important phase information present in the signals since the CWD is based on real-valued second order statistic. The waveform recognition method proposed in this paper mitigates the multipath effects by using a denoising auto-encoder (DAE) as a blind equalizer, since the quality of received

signal is seriously degraded by multipath fading channels. The equalized signal is the processed using complex-valued time-frequency Fourier Synchrosqueezing Transform (FSST). The FSST has many desirable properties including rate of change, number of signal components and their strengths, describing time-varying behavior and the causality property. The FSST based recognition may achieve a better discrimination of polyphase codes since takes phase information into account unlike the CWD which is phase blind. The complex-valued output of FSST is turned into a bivariate image and reshaped to provide a sufficiently detailed time-frequency description for classification purposes. The image is used as an input to a deep convolutional neural network (CNN). The network is trained such that it is exposed to a wide variety of representative signal conditions experienced in a waveform recognition task. In the following, we describe the input signal model, FSST technique and its properties, the DAE structures for the blind equalization, representing the transform output as an image and finally the CNN configuration used for classification in more detail.

II. SYSTEM AND WAVEFORMS

In this section, we present the signal model and the radar waveforms considered in this paper. Assuming that the radar signal is transmitted over the AWGN channel. Then, the discrete-time complex-valued intercepted signal $y[k]$ can be expressed as

$$y[k] = s[k] + n[k] = ae^{j\theta[k]} + n[k], \quad (1)$$

where k is the sample index increasing every T_s for a sampling frequency f_s , $s[k]$ represents the complex radar signal, $n[k]$ is the complex AWGN with the variance σ_n^2 , a is an invariant amplitude, and $\theta[k]$ is the instantaneous phase of the radar waveforms. The instantaneous phase $\theta[k]$ can be defined as

$$\theta[k] = 2\pi f[k](kT_s) + \phi[k], \quad (2)$$

where $f[k]$ is the instantaneous frequency and $\phi[k]$ is the phase offset.

We consider classifying eight radar waveforms in this paper. There are the linear frequency modulation (LFM), Costas,

This work was partially supported by a scholarship from Maanpuolustuksen kannatusäätiö.

TABLE I
EIGHT RADAR WAVEFORMS CONSIDERED IN RECOGNITION.

Modulation	$f[k](Hz)$	$\phi_{i,j}[k](rad)$
LFM	$f_0 + k\alpha$	constant
Costas	f_j	constant
BPSK	constant	0 or π
Frank	constant	$\frac{2\pi}{M}(i-1)(j-1)$
P1	constant	$-\frac{\pi}{M}[(M-(2j-1))][(j-1)M+(i-1)]$
P2	constant	$-\frac{\pi}{2M}[2i-1-M][2j-1-M]$
P3	constant	$\frac{\pi}{\rho}(i-1)^2$
P4	constant	$\frac{\pi}{\rho}(i-1)^2 - \pi(i-1)$

binary phase shift keying (BPSK), and five polyphase codes such as Frank, P1, P2, P3, and P4 codes as shown in Table I, where α is the chirp rate, f_j is the frequency corresponding to the j th element of the Costas sequence, M is the number of frequency steps, and ρ is the pulse compression ratio. The eight radar waveforms can be divided into two pulse compression methods, FM (Frequency modulation) and PM (phase modulation). In the PM, such as the LFM and Costas codes, $f[k]$ varies while $\phi[k]$ is kept constant, and in the PM, such as the BPSK and five polyphase codes, $\phi[k]$ varies while $f[k]$ is kept constant as defined in Table I.

III. FOURIER-BASED SYNCHROSQUEEZING TRANSFORM

The intercepted time-domain radar signal is first turned into a time-frequency domain signal using the FSST. The FSST is a kind of time-frequency transform which is an effective approach to provide a concentrated representation of multicomponent signals in time-frequency plane [7], [8]. This transform belongs to the family of time-frequency reassignment methods, and operates on the time-frequency domain of the short-time Fourier transform (STFT) by reassigning the coefficients in scale or frequency. This relocation process gives sparse and sharper time-frequency representations and makes it easier to distinguish among different received waveforms. The reassignment can be used for other time-frequency representation as well, where the coefficients are reassigned in both the time and frequency components [9]. In the FSST, the coefficients are reassigned only in the frequency component in order to preserve causality.

The FSST $T_f(\omega, t)$ is based on the modified coefficients of STFT $V_f(\eta, t)$ from (η, t) to $(\hat{\omega}_f(\eta, t), t)$ described by

$$T_f(\omega, t) = \frac{1}{g(0)} \int V_f(\eta, t) \delta(\omega - \hat{\omega}_f(\eta, t)) d\eta, \quad (3)$$

where $g(0)$ is the value of a sliding window $g(t)$ at time 0, δ is the Dirac delta function, and $\hat{\omega}_f(\eta, t)$ is the instantaneous frequency defined by

$$\hat{\omega}_f(\eta, t) = \text{Re} \left(\frac{1}{j2\pi} \frac{\partial_t V_f(\eta, t)}{V_f(\eta, t)} \right). \quad (4)$$

The instantaneous frequency can be approximated by $\hat{\omega}_f(\eta, t)$, when $V_f(\eta, t) > 0$. From the FSST, we can obtain the

complex-valued bivariate image $T_f(\omega, t)$ with the size of $L \times L$. Considering the classification performance and the computational complexity, we set $L = 1024$ in this paper.

IV. PROPOSED RADAR WAVEFORM RECOGNITION SYSTEM IN MULTIPATH FADING CHANNELS

In this section, we present the structure of proposed radar waveform recognition approach for multipath fading channels. The proposed hierarchical machine learning system consists of two stages as shown in Fig. 1(a). The first stage classifies the eight different radar waveforms, and the second stage additionally distinguishes between the P1 and P4 waveforms if classifier output in the first stage is the P1 or P4 waveform. These two waveforms are very similar and most classifiers find it difficult to distinguish them. Similar hierarchical structure of classifiers is obviously applicable to any subset of classes that are hard to distinguish. In each stage, three steps are processed, the blind equalization based on the DAE, the conversion of the time-frequency representation into a bivariate image and, and the classification using the CNN. The intercepted signal from multipath fading channels is equalized using the DAE. We consider two DAE structures for the first and second stages as shown in Fig. 1(b) and 1(c). In the image preprocessing stage, the FSST transformed signal is converted to a bivariate image corresponding to the real and imaginary components of FSST. To reduce the load on CNN while maintaining a sufficient level of detail for classification, image clipping and resizing operations are performed. In the classification stage, a deep CNN structure with three convolutional layers and two fully-connected (FC) layers is constructed to perform the classification. By employing the DAEs as the blind equalizers, prominent features of the waveforms are better revealed and the constructed CNN can perform the recognition task reliably even in the face of fading.

A. Blind Equalization Based on DAE

This section describes the two proposed DAE structures. These DAEs are used to mitigate the effect of multipath fading channels and attenuate noise without the knowledge of channel state information or known training waveform. Using a representation learning through the DAE [10], [11], the blind equalization is effectively conducted by performing the dimensionality reduction and reconstruction with the encoder and decoder. As a consequence, the DAE produces a representation of data that is immune to the effect of fading channel and has reduced noise level. [10], [11]. For the hierarchical classifier, there are two types of DAEs for the eight waveforms and the P1 and P4 waveforms in the first and second stages as shown in Fig. 1(a), respectively.

The DAE for the eight waveforms equalizes and denoises the intercepted signal $\mathbf{y}$ to the estimated radar signal $\hat{\mathbf{s}}$ as shown in Fig. 1(b). It is trained by using the eight radar waveforms in Table I as a target signal and has an input layer, two FC hidden layers, and an output layer. In the input layer, the real-valued signal $\mathbf{y}$ with the size 2048×1 is used as an input of the DAE by stacking the real components with

Fig. 1. Block diagrams of the proposed waveform recognition system. (a) Two-stage radar waveform recognition, (b) DAE for eight waveforms and (c) DAE for P1 and P4 waveforms.

the size 1024×1 first followed by imaginary components with the size 1024×1 of the complex input signal with the size 1024×1 into the real-valued signal that has twice the length of the complex signal. For the complex input signal of the DAE, we first collect complex intercepted signal samples of $N \leq 1024$ observations. Then, the complex input signal with the size 1024×1 of the DAE is made by repeatedly attaching the collected signal so that it is larger than a size of 1024 and cutting the attached signal to a size of 1024. For the complex input signal, it is more effective to insert the collected signal itself repeatedly than zero in the unobserved part to maintain the performance of the DAE for $N \leq 1024$ observations. In the FC layers, there are two FC hidden layers with an exponential linear unit (ELU) [12] activation function for both the encoder and decoder. The numbers of neurons are 1024 and 512 for the first and second FC hidden layers, respectively. In the output

layer, the hyperbolic tangent activation function is employed to estimate the radar waveform $\hat{s}$ with the size 2048×1 for a multi-dimensional regression problem, since we consider the constant modulus waveforms.

The DAE for the P1 and P4 waveforms equalizes and denoises the intercepted signal y to the estimated radar signals $\hat{s}_{P1}$ and $\hat{s}_{P4}$ targeted to distinguishing between P1 and P4 waveforms as shown in Fig. 1(c), respectively. In the DAE for the P1 and P4 waveforms, a switch is added and a different type of output layer is used compared to the DAE for the eight waveforms. The switch determines whether to turn on the DAE or not based on a flag indicating whether the P1 and P4 waveforms are detected in the first stage. The output layer with the hyperbolic tangent activation function outputs two equalized signals $\hat{s}_{P1}$ and $\hat{s}_{P4}$ of the size 2048×1 targeted to the P1 and P4 waveforms, respectively. In the training procedure, the DAE for the P1 and P4 waveforms is learned by setting the target signals for $\hat{s}_{P1}$ and $\hat{s}_{P4}$ as s_{P1} and $-s_{P4}$ for the P1 waveform and $-s_{P1}$ and s_{P4} for the P4 waveform, respectively. s_{P1} and s_{P4} are the P1 and P4 waveforms with the same size, and a negative gain means the phase shift of 180 degrees. By learning the phase shift of 180 degrees of the P4 or P1 waveform together when the P1 or P4 waveform is equalized, the P1 or P4 waveform is equalized to be well separated from the P4 or P1 waveform, respectively. It makes the average of two CNN outputs in the second stage can be the highest value for the P1 or P4 waveform and the lowest value for P4 or P1 waveform, respectively.

B. Image Preprocessing

In the image preprocessing step, there are four phases to reshape the FSST output to an image, namely image clipping and resizing. The image clipping and resizing convert the complex-valued image from the FSST with the size of $L \times L$ into the real-valued image with the size of $I \times I \times 2$ for $I < L$ to make an input signal in the real-valued CNN with the reduced computational load. First, the real-valued output of the DAE is rewritten as a complex signal of half the dimension and used as an input to the FSST. We convert the real-valued output of the DAE with the size 2048×1 to the complex input signal of the FSST with the size 1024×1 . Second, we create a bivariate image of the size of 1024×1024 from the FSST based on the complex-valued equalized signal with the size 1024×1 . Third, the bivariate image is clipped to eliminate the pixels on the FSST generated by the repeatedly attached signal in the complex input signal of the DAE for improving the accuracy of the CNN. In the image clipping, the image with the size of 1024×1024 is clipped by $1024 \times N$, where N is the length of original waveform samples. Finally, we resize the clipped image with original observation vector size of $1024 \times N$ into 128×128 pixels for the input of CNN. The size of 128×128 is sufficient to capture the salient characteristics of various waveforms [4]. For the image resizing, the nearest neighbor interpolation (NNI) technique [13] is used to prevent pixel degradation due to the multiple noisy pixels. As the result, the size of image for the input of the CNN is $128 \times 128 \times 2$,

TABLE II
STRUCTURE OF THE EMPLOYED CNN WITH THE FSST INPUT IN
TERMS OF THE LAYERS AND THEIR PARAMETERS.

Layer	Dimension	Parameters	Activation
Input	(128,128,2)	-	-
Convolution 2D	(128,128,8)	152	-
Batch Normalization	(128,128,8)	16	ELU
Max Pooling 2D	(64,64,8)	-	-
Convolution 2D	(64,64,16)	1,168	-
Batch Normalization	(64,64,16)	32	ELU
Max Pooling 2D	(32,32,16)	-	-
Convolution 2D	(32,32,32)	4,640	-
Batch Normalization	(32,32,32)	64	ELU
Average Pooling 2D	(16,16,32)	-	-
Flatten	-	-	-
FC1 (Dense)	(8,192)	1,048,704	ELU
Dropout (0.5)	-	-	-
FC2 (Dense)	(128)	1,032	ELU
Dropout (0.5)	-	-	-
Dense	(8)	-	Softmax

because the NNI process is separately operated for the real and imaginary parts of image.

C. Classification Based on CNN

This section describes the proposed CNN structure used in the first and second stages. A deep CNN is used for the radar waveform recognition problem, with the network structures given in Table II. Unlike the DAEs, the first and second stages use the same CNN structure trained by using the eight radar waveforms in Table I as target signals. The CNN consists of an input layer, three convolutional layers, three batch normalization layers, three pooling layers, two FC layers, and an output layer. In the input layer, the bivariate image of the size 128×128 created in the image preprocessing stage is used as an input to the CNN. For the convolutional layer, the numbers of convolution filters for the first, second, and third layers are set to 8, 16, and 32 with the filter size 3×3 , respectively. We use the batch normalization layers between the convolutional layer and activation function [14]. For the pooling layer, we use the max and average pooling with 2×2 filter size, stride size 2, and no zero-padding. The ELU is used as the activation function layers except the output layer where Softmax was employed [12]. In the FC layers, the number of neurons for the second FC layer is 128, and random Dropout with rate 0.5 are employed between the two FC layers to avoid an overfitting problem since a large number of neurons in the FC layer. In the output layer, Softmax consisting of the same number of nodes as the categories computes probabilities to each class for a multi-class problem. Based on the probabilities calculated in Softmax, the CNN classifies it under categories including of six radar waveforms in Table I except the P1 and P4 waveforms in the first stage. In the second stage, based on the average of probabilities from two Softmax outputs, two CNNs classify it under categories including of the remaining P1 and P4 waveforms.

TABLE III
WAVEFORM AND SIMULATION PARAMETERS.

Radar Waveforms	Parameters	Value of Range
LFM	f_c	$U(f_s/6, f_s/4)$
	B	$U(f_s/20, f_s/16)$
	N	$U[512, 1024]$
Costas	FH sequence	$\{3, 4, 5, 6\}$
	f_{min}	$U(f_s/24, f_s/20)$
	N	$U[512, 1024]$
BPSK	L_c	$\{7, 11, 13\}$
	f_c	$U(f_s/6, f_s/4)$
	N_{cc}	$U[10, 13]$
Frank,P1	f_c	$U(f_s/4, f_s/3)$
	N_{cc}	$\{3, 4, 5\}$
	M	$\{6, 7, 8\}$
P2	f_c	$U(f_s/4, f_s/3)$
	N_{cc}	$\{3, 4, 5\}$
	M	$\{6, 8\}$
P3, P4	f_c	$U(f_s/4, f_s/3)$
	N_{cc}	$\{3, 4, 5\}$
	ρ	$\{36, 49, 64\}$

TABLE IV
PARAMETERS OF MULTIPATH RICIAN FADING CHANNELS.

Channel Parameters	Value of Range
Number of path delay, N	$U[1, 3]$
Step size of path delays, D	$U[200, 500] (ns)$
Step size of path gains, G	$U[-8, -4] (dB)$
K-factor, K	$U[5, 15]$
Maximum Doppler shift, S	$U[1000, 3000] (Hz)$

V. SIMULATION RESULTS

In this section, we evaluated the classification performance for the radar waveform recognition based on the CNN using the FSST in multipath Rician fading channels. In the FSST, a Kaiser window of length 256 and shape parameter $\beta = 10$ is used. For comparison purposes, we experimented with the CNN using the CWD represented as an image as an input to the CNN as in [3], [4]. The CNN classifies the collected radar signal samples with $L = 1024$ into eight different radar waveforms: the LFM, Costas, BPSK, Frank, and P1-P4 codes.

Table III describes the waveform and simulation parameters for testing the classification performance as similar to [4], where f_c , B , f_{min} , L_c , and N_{cc} is the center frequency, the bandwidth of the intercepted signal, the fundamental frequency of the Costas code, the code length, and the cycles per phase code, respectively. For the sampling frequency, we used $f_s = 3.84 MHz$. $U(\cdot)$ denotes a uniform distribution.

Table IV describes the parameters of multipath Rician fading channels. The realization of Rician fading channel is based on the path delays, path gains, Rician K-factor K , and maximum Doppler shift S [15]. The path delays and path gains with the length of N specifying the channel's average delay profile are created by increasing in step size of D (ns) and G (dB) from 0, respectively, where N is the number of path delay, D is the step size of path delays, and G is the step size of path gains. For example, when $N = 3$, $D = 300$, and $G = -5$, the path delays are $[0, 300, 600] (ns)$, and the path

Fig. 2. Classification performances as a function of SNR. (a) Overall, (b) P1 and (c) P4.

gains are $[0, -5, -10]$ (dB). These variables N , D , G , K and S are generated by uniform distribution as shown in Table IV.

In the training and validation processes of the DAE and CNN, we generated 150,000 and 15,000 signals in the DAE and CNN for each of the eight radar waveforms with signal-

to-noise ratio (SNR) ranging from -12 dB to -2 dB at a step size of 2 dB, respectively. In each signal, 80% and 20% of the signals were used as the training and validation sets, respectively. In the test process of the CNN, we generated 10,000 signals for each of the eight radar waveforms. The SNR is defined as $\text{SNR (dB)} = 10\log_{10}(\sigma_s^2 / \sigma_n^2)$, where σ_s^2 is the variance of signal, and we assume $a = 1$. The DAE and CNN were trained over 20 epochs using Adam optimizer [16]. The learning rate is updated every five epochs by multiplying with 10^{-1} from an initial learning rate 10^{-3} . In the training process, we used the mini-batch gradient descent algorithm [17] with the batch size of 64 that splits the training dataset into small batches to estimate the error gradient before the model weights are updated, and each batch was randomly shuffled at each epoch.

For the CNN using the CWD, we use the similar CNN structure described in Table II by modifying the size of input layer. In the CNN with the CWD, the size of input layer is $128 \times 128 \times 1$ since the CWD is defined as real-valued and non-negative across time and frequency. In the CWD, the scaling factor $\sigma = 1$ is used.

Fig. 2 shows the classification performances as a function of SNR for the overall, P1, and P4 cases. We evaluated the classification probabilities of the classification without the DAE and the proposed schemes, one-stage and two-stage classifications, considering the CWD and FSST, respectively. The overall classification probabilities of the proposed methods are significantly improved compared to the classification without the DAE as shown in Fig. 2(a). Even at SNR of -12 dB, the overall probabilities of the proposed methods are still more than 90%. In the proposed methods, the two-stage classification has the better overall classification probability than the one-stage classification by improving the discrimination capability between the P1 and P4 waveforms. The P1 and P4 waveforms that are usually difficult to distinguish have considerably worse classification probabilities compared to the overall classification probability as shown in Fig. 2(b) and (c). Therefore, the discrimination between these two waveforms allows for significantly improving the overall classification probability by using the two-stage classification.

Considering the CWD and FSST in the proposed methods, the FSST shows a better overall classification probability than the CWD in the two-stage classification, while the CWD and FSST have the similar classification probabilities in the one-stage classification. In [18], the FSST shows better overall classification probability than the CWD in AWGN channels. The FSST has the better classification probability of the P4 waveform than the CWD in the two-stage classification, while the CWD shows the somewhat better classification probability of the P1 waveform than the FSST. At the SNR below -10 dB, the two-stage classification has lower classification probability than the one-stage classification when the CWD is used. From the average of P1 and P4 waveforms, the FSST shows the about 1.3% higher (83.7% vs 82.4%) classification probability than the CWD in the two-stage classification at SNR of -12 dB. Also, in the P4 waveform which has the

TABLE V
CONFUSION MATRIX FOR THE TWO-STAGE CLASSIFICATION WITH
CWD AT SNR OF -10 dB. THE OVERALL CLASSIFICATION RATE
IS 95.3%.

	LFM	Costas	BPSK	Frank	P1	P2	P3	P4
LFM	97.7	1.3	0.5	0.1	0.1	0	0.2	0.1
Costas	2.1	97.3	0.5	0.1	0	0	0	0
BPSK	0.9	0.4	98.5	0	0	0	0.1	0.1
Frank	0.3	0.1	0.3	97.9	0.4	0.4	0.5	0.1
P1	0.2	0.1	0.1	0.3	91.9	0.3	0.3	6.8
P2	0.3	0.1	0.1	0.4	0.4	98.2	0.4	0.1
P3	0.4	0.1	0.2	0.4	0.4	0.4	98.0	0.1
P4	0.3	0.1	0.1	0.4	15.7	0.2	0.3	82.7

TABLE VI
CONFUSION MATRIX FOR THE TWO-STAGE CLASSIFICATION WITH
FSST AT SNR OF -10 dB. THE OVERALL CLASSIFICATION RATE
IS 95.7%.

	LFM	Costas	BPSK	Frank	P1	P2	P3	P4
LFM	97.5	1.8	0.5	0.1	0	0	0.1	0
Costas	2.1	97.6	0.2	0	0	0	0.1	0
BPSK	1.0	0.3	98.7	0	0	0	0	0
Frank	0.1	0.1	0.1	98.2	0.3	0.5	0.5	0.2
P1	0.1	0.1	0	0.4	89.0	0.3	0.3	9.8
P2	0.1	0	0	0.4	0.3	98.6	0.4	0.2
P3	0.2	0.1	0.1	0.4	0.3	0.4	98.3	0.2
P4	0.2	0.1	0	0.4	11.2	0.3	0.4	87.4

worst classification probability, the FSST has about 5.4% higher (83.6% vs 78.2%) classification probability than the CWD in the two-stage classification at SNR of -12 dB. As a result, the two-stage classification has about 1% higher overall classification probability than the one-stage classification at SNR of -12 dB when the FSST is used.

Tables. V and VI show the confusion matrices of the two-stage classification for the CWD and FSST at SNR of -10 dB, respectively. The confusion matrices indicate that the discrimination between the P1 and P4 waveforms is the most difficult. Considering the discrimination between the P1 and P4 waveforms, the FSST has the significantly improved discrimination capability of the P4 waveform than the CWD, but the FSST shows somewhat lower discrimination capability of the P1 waveform than the CWD. The overall classification probability is about 0.3% higher when using the FSST instead of the CWD.

VI. CONCLUSION

In this paper, we proposed a hierarchical deep learning method for the radar waveform recognition in multipath fading channels. The received signal is first equalized to mitigate the effect of multipath fading by using the DAE. Then, the equalized signal is converted into complex-valued FSST which provides a sharp, causal and concentrated time-frequency representation. The resulting bivariate image is used as an input to the CNN after clipping and resizing. By using the DAE as a blind equalizer, the overall classification probability of the one-stage classification is considerably improved. Also,

the two-stage classification further enhances the classification probabilities of the P1 and P4 waveforms using the additional DAE. The one-stage classification has more than 5% higher overall classification probability than the classification without the DAE when SNR is less than -8 dB. By applying the hierarchical structure, the two-stage classification shows more than 0.5% overall classification probability than the one-stage classification when SNR is less than -8 dB. Especially, the FSST has better discrimination capability between the P1 and P4 waveforms than the CWD at low SNR regime. In the proposed two-stage classification, the FSST shows more than 0.3% higher overall classification probability than the CWD when SNR is less than -8 dB.

REFERENCES

- [1] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*. Norwood, MA, USA: Artech House, 2009.
- [2] E. R. Zilberman and P. E. Pace, "Autonomous time-frequency morphological feature extraction algorithm for LPI radar modulation classification," in *Proc. IEEE Int. Conf. Image Process.* Oct. 2006, pp. 2321–2324.
- [3] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074–11082, Jun. 2017.
- [4] S.-H. Kong, M. Kim, L. M. Hoang, and E. Kim, "Automatic LPI radar waveform recognition using CNN," *IEEE Access*, vol. 6, pp. 4207–4219, Jan. 2018.
- [5] J. Lunden and V. Koivunen, "Automatic radar waveform recognition," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no.1, pp. 124–136, Jun. 2007.
- [6] M. Zhang, L. Liu, and M. Diao, "LPI radar waveform recognition based on time-frequency distribution," *Sensors*, vol. 16, no.10, pp. 1682–1702, Oct. 2016.
- [7] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based synchrosqueezing transform," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May. 2014, pp. 315–319.
- [8] R. Behera, S. Meignen, and T. Oberlin, "Theoretical analysis of the second-order synchrosqueezing transform," *Appl. Comput. Harmon. Anal.*, vol. 45, pp. 379–404, Sep. 2018.
- [9] F. Auger, P. Flandrin, Y. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Oct. 2013.
- [10] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," *ICML*, Helsinki, Finland, Jul. 2008.
- [11] S. Li, J. Kawale, and Y. Fu, "Deep collaborative filtering via marginalized denoising auto-encoder," *CIKM*, Melbourne, Australia, Oct. 2015.
- [12] D. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)," *ICLR*, San Juan, Puerto Rico, May 2016.
- [13] J. A. Parker, R. V. Kenyon, and D. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imag.*, vol. MI-2, no. 1, pp. 31–39, Mar. 1983.
- [14] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.
- [15] M. Jeruchim, P. Balaban, and K. Shanmugan, *Simulation of Communication Systems*. New York, Kluwer Academic/Plenum, 2000.
- [16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *ICLR*, San Diego, CA, May 2015.
- [17] A. Cotter, O. Shamir, N. Srebro, and K. Sridharan, "Better mini-batch algorithms via accelerated gradient methods," *NIPS*, 2011.
- [18] G. Kong and V. Koivunen, "Radar waveform recognition using Fourier-based synchrosqueezing transform and CNN," *Proc. IEEE Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process. (CAMSAP)*, Dec. 2019.