IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 21, 2020, accepted January 13, 2021, date of publication January 21, 2021, date of current version January 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3053491

Identification of OFDM-Based Radios Under Rayleigh Fading Using RF-DNA and Deep Learning

MOHAMED K. M. FADUL[1], DONALD R. REISING[1], (Senior Member, IEEE), AND MINA SARTIPI[2], (Senior Member, IEEE)

[1]Department of Electrical Engineering, The University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA
[2]Department of Computer Science and Engineering, The University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA

Corresponding author: Donald R. Reising (donald-reising@utc.edu)

ABSTRACT The Internet of Things (IoT) is here and has permeated every aspect of our lives. A disturbing fact is that the majority of all IoT devices employ weak or no encryption at all. This coupled with recent advances within the areas of computational power and deep learning has increased interest in Specific Emitter Identification (SEI) as an effective means of IoT security. Deep learning is capable of in-situ extraction of discriminating features, making it well suited to discrimination of wireless transmitters without the need for feature engineering. However, the accuracy of the deep learning model is adversely affected by time-varying channel conditions. The time-varying nature is attributed to the mobility of the transmitter, receiver, objects within the operations environment, or combinations thereof. This can result in the channel conditions changing faster than the deep learning algorithm is capable of handling. This paper assesses deep learning–based SEI using waveforms that undergo Rayleigh fading, as well as channel estimation and equalization, prior to being input into a deep learning algorithm.

INDEX TERMS AutoEncoder, convolutional neural network (CNN), Deep learning, feature engineering, radio frequency (RF) fingerprinting, specific emitter identification (SEI).

I. INTRODUCTION

The Internet of Things (IoT) is here and has permeated every aspect of our lives both personal and professional. An estimated 26.66 billion IoT devices are currently deployed, and that number is expected to reach 75.4 billion by 2025 due to the roughly 127 IoT devices being connected to the Internet every second [1]–[4]. A disturbing fact of this rapid growth is that the majority, roughly 70%, of all IoT devices fail to use encryption due to (i) on-board computation restrictions, (ii) the manufacturer's cost of implementation being too high, and (iii) implementation and management challenges that are exacerbated at scale [5]–[7]. This lack of security makes IoT devices and the corresponding infrastructure open to attacks by devices that are incorrectly authenticated—often due to their use of compromised digital credentials that have been transmitted in the clear. Thus, there is a critical need for the development and integration of more advanced security techniques as the proliferation of IoT devices continues [4]. The criticality of this need is intensified as the lack of encryption or related weaknesses of IoT devices and/or infrastructures are attacked [8]–[14].

Specific Emitter Identification (SEI) is one technique capable of filling this gap in security [15], [16]. SEI is a physical layer approach that has demonstrated success in identifying wireless transmitters by exploiting the unintentional coloration that is imparted upon every waveform during its formation and transmission [17]–[35]. This coloration is attributed to the distinct characteristics and interactions of the radio's Radio Frequency (RF) front-end components. It differs from other security approaches (e.g., encryption, passwords) in that it is inherent to every transmitter and is very difficult to mimic, which in turn makes it very challenging to circumvent.

One SEI implementation is known as RF-Distinct Native Attributes (RF-DNA) fingerprinting. RF-DNA fingerprinting extracts SEI features from the portion or portions of the waveform associated with a fixed, known sequence of symbols such as those of the IEEE 802.11a Wi-Fi preamble [36]. RF-DNA fingerprinting has demonstrated success in

The associate editor coordinating the review of this manuscript and approving it for publication was Adnan Akhunzada.

multiple research efforts within the SEI and network security communities. Most of these RF-DNA fingerprinting efforts have achieved intra-model (a.k.a., serial number) discrimination [20], [22]–[26], [34], [35], [37]–[40]. Prior RF-DNA fingerprinting has performed radio discrimination using traditional supervised or unsupervised classification algorithms. We argue that these implementations overlook recent successes within the area of deep learning, which stands to improve the accuracy and efficiency of fingerprinting-based security. These improvements are of even greater importance when SEI is conducted using waveforms that transverse a multipath environment, because the resulting interference and phase shifting makes it more difficult for the classification algorithm to learn the inherent features.

Our paper contributes the first case in which RF-DNA fingerprint–based SEI is performed using deep learning. Specifically, we perform deep learning–based SEI using:

1) Waveforms that undergo Rayleigh fading (i.e., the line-of-sight waveform is not present in the received signal) and equalization prior to RF-DNA fingerprint learning.
2) RF-DNA fingerprints in which the features are learned from the raw In-Phase & Quadrature (IQ) samples, complex Gabor coefficients, and representations thereof.
3) An autoencoder-initialized Convolutional Neural Network (CNN) to aid in convergence and improve classification accuracy.
4) An assessment under four multipath channels: noise only, three, five, and seven reflectors/paths
5) An assessment of deep learning's ability to learn SEI features that are invariant to noise or noise and multipath.

This work differs from previous RF-DNA fingerprinting works in that handcrafted features (a.k.a., feature engineering) are *not* used. In feature-engineered RF-DNA fingerprinting, the segmentation of the waveform representation (e.g., power spectral density, instantaneous phase) and calculated features (e.g., skewness, kurtosis) are done prior to and without knowledge of how they aid or hinder the discrimination of one radio from another and vice versa. Thus, handcrafted RF-DNA fingerprints rely upon an individual's knowledge and expertise, which can negatively impact SEI performance. Deep learning–based RF-DNA fingerprinting eliminates the need for (i) empirical partitioning of the waveform or its transform, (ii) calculation of features (e.g., variance, skewness, kurtosis, entropy) over each partition, and (iii) feature selection (e.g., principal component analysis) prior to SEI. Our deep learning–based RF-DNA fingerprinting approach is compared with the handcrafted RF-DNA fingerprinting approach in [40]; this prior approach is described in Sect. III-F and corresponding results shown in Sect. V.

The remaining sections of this paper are as follows: Sect. II presents a review of related works; Sect. III provides a brief overview of signal collection, detection, post-processing, and multipath generation and equalization; Sect. IV presents the CNN-based SEI approaches; Sect. V presents CNN-based SEI results; and the conclusion is in Sect. VI.

II. RELATED WORK

RF-DNA fingerprinting and similar SEI works augment wireless network security through the use of radio signal classification. Recently, multiple SEI investigations have conducted radio signal classification using deep learning [35], [41]–[54]. The application of deep learning within the SEI domain is motivated by the removal of feature engineering while achieving superior radio signal classification results. The remainder of this section presents a brief summary of related deep learning–based SEI works.

These prior deep learning–based SEI works either (i) perform analysis without any channel impairments (e.g., interference, noise, multipath) [44]–[46], [49], [50], (ii) use a noise only channel [41], [42], [47], [51], [52], (iii) use a multipath channel with unspecified or unknown characteristics [43], [48], [53], [54], or (iv) use a static multipath channel (i.e., the same multipath channel coefficients are used for every transmitted waveform) [35]. These works assume that the chosen deep learning approach will sufficiently learn the channel to mitigate its impact on the classification decisions. This assumption is convenient in that it removes the need for channel estimation and correction steps within the SEI process, but it overlooks the time-varying nature of most multipath fading conditions. In our work, CNN-based SEI is assessed using RF-DNA features learned from waveforms that have undergone channel estimation and equalization after simulated transmission through a dynamic Rayleigh fading channel under degrading SNR conditions.

III. BACKGROUND

A. SIGNAL COLLECTION, DETECTION, AND POST-PROCESSING

RF-DNA fingerprint–based SEI uses a portion of the transmitted waveform that corresponds with a fixed, known sequence of symbols used by the receiver to perform synchronization and channel equalization to facilitate demodulation. This work uses the fixed, known, Orthogonal Frequency Division Multiplexed (OFDM) symbol sequence of the IEEE 802.11a Wi-Fi waveform's preamble. Our rationale for using the 802.11a preamble is three-fold: (i) OFDM is used in 802.11ac, 802.11ad, 802.11ax, and Long Term Evolution (LTE) [55]; (ii) prior SEI research has demonstrated success with its use [18], [20], [23], [27], [33], [34], [37], [38], [40], [46], [48], [56]; and (iii) the on-hand set of preambles was used in [37], [38], [40], which facilitates comparative analysis.

All signals are collected from $N_D = 4$ Cisco AIR-CB21G-A-K9 Wi-Fi cards/devices using an Agilent E3238S-based spectrum analyzer within an office environment. This spectrum analyzer has a 36 MHz wide RF bandwidth, a frequency range of 20 MHz to 6 GHz, a sampling rate of up to 95 mega-samples/s, and a 12-bit analog-to-digital converter [57]. For each radio, a total of $N_B = 2,000$ transmissions/bursts

TABLE 1. The delays, τ_k, and variances, σ_k^2, used to generate the Rayleigh fading channel models comprised of L paths.

	Path Delays (τ_k)						
L	50 ns	100 ns	150 ns	200 ns	250 ns	300 ns	350 ns
3	✓		✓		✓		
5	✓	✓	✓	✓	✓		
7	✓	✓	✓	✓	✓	✓	✓

	Path Variances (σ_k^2)						
L	σ_1^2	σ_2^2	σ_3^2	σ_4^2	σ_5^2	σ_6^2	σ_7^2
3	0.8		0.13		0.07		
5	0.865	0.117	0.016	0.002	0.0003		
7	0.8	0.117	0.065	0.0157	0.0018	0.0003	0.0002

are selected using amplitude-based variance trajectory detection [22]. Following detection, each waveform is filtered using a fourth order low-pass Butterworth filter with a cutoff frequency of 7.7 MHz. Once filtered the preamble's IQ samples are stored for post-processing. Post-processing consists of carrier frequency offset correction and re-sampling to a sampling rate of 20 MHz.

B. MULTIPATH CHANNEL MODELING

Multipath fading is a common occurrence within wireless communications environments and leads to distortion of the received waveform. If not compensated for, this distortion adversely affects the demodulation process. In addition to interfering with receiver function, multipath distortion alters the SEI-exploited distinct and native waveform characteristics that facilitate transmitter discrimination.

For the results presented in Sect. V, the multipath channel is modeled using Rayleigh fading, which is the default, indoor model for IEEE 802.11a Wi-Fi operating environments [58]. The Rayleigh fading model is implemented through the use of a tap delay line given by

$$h(t,\tau) = \sum_{k=1}^{L} \alpha_k \delta(t - \tau_k T_s),\ \alpha_k = a_k + jb_k, \qquad (1)$$

where T_s is the sampling period; L is the total number of paths; τ_k is the time delay for the $k = 1, \ldots, L$ path; and a_k, b_k are zero mean, independent and identically distributed (iid) Gaussian random variables of variance σ_k^2. The specific values for τ_k and σ_k^2 used to construct Rayleigh fading channels of $L = [3, 5, 7]$ paths are presented in Table 1.

C. MULTIPATH ESTIMATION AND CORRECTION

RF-DNA fingerprinting is a waveform-based SEI approach; thus, estimation and correction of the fading channel delays and coefficients is performed using the waveform-based approach initially presented in [38], which uses the Nelder-Mead (N-M) simplex algorithm [59].

The first step in estimating the channel's impulse response is time synchronization. The start of the received waveform is obtained using the Least Square (LS) estimator in [60]. The resulting estimate coincides with the delay of the fading channel's $k = 1$ path. Any remaining delays are estimated in relation to the first as described in [38].

Following time synchronization, estimation of the impulse response's coefficients is performed using the N-M channel estimator developed in [38]. The N-M estimator is an iterative, direct search approach that minimizes the function given in (2):

$$f(h) = \sum_{k \in m} \left| r(m) - \sum_{k=0}^{L-1} x(m - \tau_k) h_k \right|^2, \qquad (2)$$

where $r(m)$ is the preamble of the received 802.11a waveform, $x(m)$ is the transmitted preamble, and h_k is the k^{th} coefficient to be estimated. As in [37], [38], [40], $x(m)$ is one of $N_p = 20$ "candidate" preambles that represent each of the $N_D = 4$ Wi-Fi radios. A total of five preambles are randomly selected from each Wi-Fi radio's set of $N_B = 2,000$ waveforms to serve as that radio's candidates. The use of $N_p = 20$ candidate preambles results in twenty estimated values of h_k. The best estimate h_k^b results in the smallest residual error between the received and selected candidate preamble.

Following estimation of the impulse response's coefficients, channel equalization is performed using the Minimum Mean Square Error (MMSE) equalizer in [61]. MMSE equalization accounts for the channel statistics, which makes it well suited to the low SNR conditions encountered in SEI applications.

D. CONVOLUTIONAL NEURAL NETWORKS

CNN networks are feed-forward networks that use back-propagation to minimize a loss function. In CNNs, the feed-forward multi-layer neural network is prepended with one or more convolutional and pooling layers that enable it

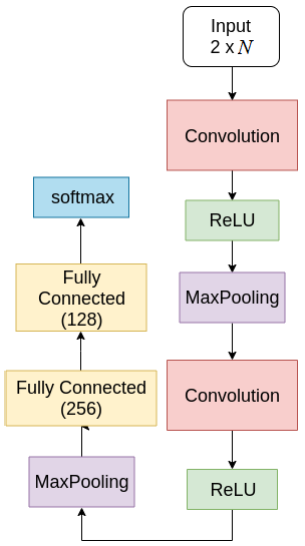


FIGURE 1. CNN architecture used for *one-dimensional* RF-DNA fingerprint representations.

to take two-dimensional data (e.g., images) as input (a.k.a., tensor) [46], [62]. CNN uses the convolutional layer(s) to extract and learn features from the input data and generate feature maps. After each convolution layer, an activation function performs a nonlinear transformation for each node in the feature map. In this work, the activation function is the Rectified Linear Unit (ReLU) function [46]. The activated feature maps are then passed to the pooling layer for dimensionality reduction. After one or more convolutional, activation, and pooling layers, dense layers extract higher-level features from the feature maps [48]. Finally, the output layer assigns the learned RF-DNA features to one of the classes that represent each of the $N_D = 4$ Wi-Fi radios.

In this work, over-fitting is minimized by using l_2 regularization to control the CNN network parameters size. Categorical cross-entropy is used for the loss function, which represents the error between the model prediction and the ground truth. ADAM optimization, with a learning rate $l = 0.0001$, is used to adjust the CNN network weights to minimize the loss function during the training process. The CNN networks are implemented using Keras with a TensorFlow backend running on NVIDIA Tesla K40m GPUs.

1) ONE-DIMENSIONAL (1D) CNN

The CNN, shown in Fig. 1, is used for 1D SEI. The input to this CNN is a $2 \times N$ real-valued tensor where N is the length of the slice/partition, as described in Sect. IV. 1D SEI is conducted using a CNN comprised of two convolutional stages. The first and second convolutional layers use 50 filters of size 1×7 and 2×7, respectively. Down sampling is implemented using 2×2 max pooling. ReLU activation is used for all layers except the output layer. The first and second fully connected layers consist of 256 and 128 neurons, respectively. A 50% dropout rate is used in the 256 layer to minimize overfitting. The output layer uses a softmax decision.

2) TWO-DIMENSIONAL (2D) CNN

All 2D SEI results are generated using the CNN network shown in Fig. 2. The input image is passed to the first convolution stage, which consists of a convolution layer constructed of 64 3×3 filters, and a 2×2 max pooling layer. The output of the first convolution stage passes through three additional convolution stages and one convolution layer prior to the flattening layer. The flattening layer transforms the passed feature maps into a 1D vector that is fed into the fully connected layers. As with the 1D CNN, a dropout rate of 50% is used to minimize over-fitting. The output of the second fully connected layer is fed to the softmax output layer where the classification decision is made. The activation function of all layers except the output layer is ReLU.

E. AUTOENCODER

An Autoencoder (AE) is a feed-forward neural network that is used to learn an efficient representation of the input data by regenerating the input at the output layer [63], [64]. The AE architecture is similar to a Multi-Layer Perceptron (MLP)

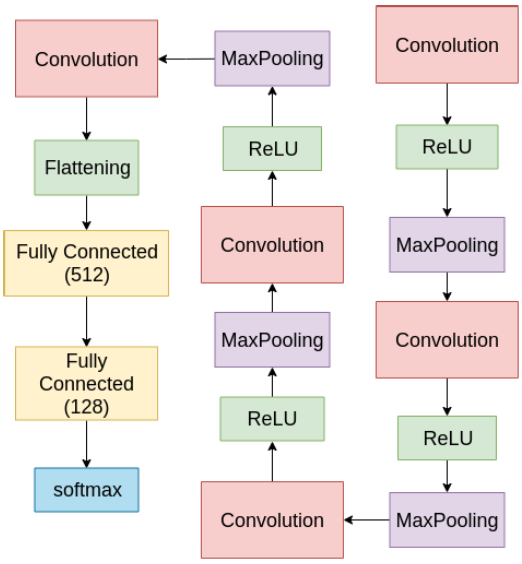


FIGURE 2. CNN architecture used for *two-dimensional* RF-DNA fingerprint representations.

network in that it has an input layer, a variable number of hidden layers, and an output layer. An AE differs from an MLP in that (i) the input and output layers have the same size (i.e., number of units), (ii) it is trained using unsupervised learning via unlabeled data, and (iii) it learns a compressed representation of the input to reproduce it at the output [62]–[64]. An AE can be used for (i) dimensionality reduction by finding an efficient, compressed representation of the input; (ii) denoising by forcing the AE to learn features from corrupted input data [52]; and (iii) initializing neural networks, especially when the number of labeled training samples is small [62]–[64]. In this paper, unsupervised initialization of a CNN is achieved by using the weights and biases of a trained Convolutional AE (CAE).

1) CONVOLUTIONAL AUTOENCODER

In CAE, the encoder is constructed using convolutional and pooling layers to extract the features while reducing the dimensionality of the input. The CAE decoder is comprised of

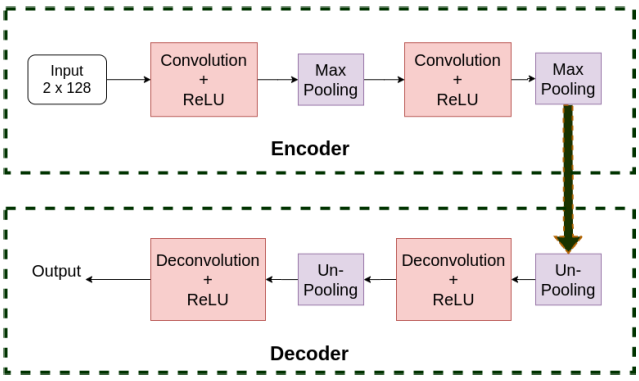


FIGURE 3. The CAE architecture used to initialize the convolutional layers' weights and biases of a *one-dimensional* CNN.

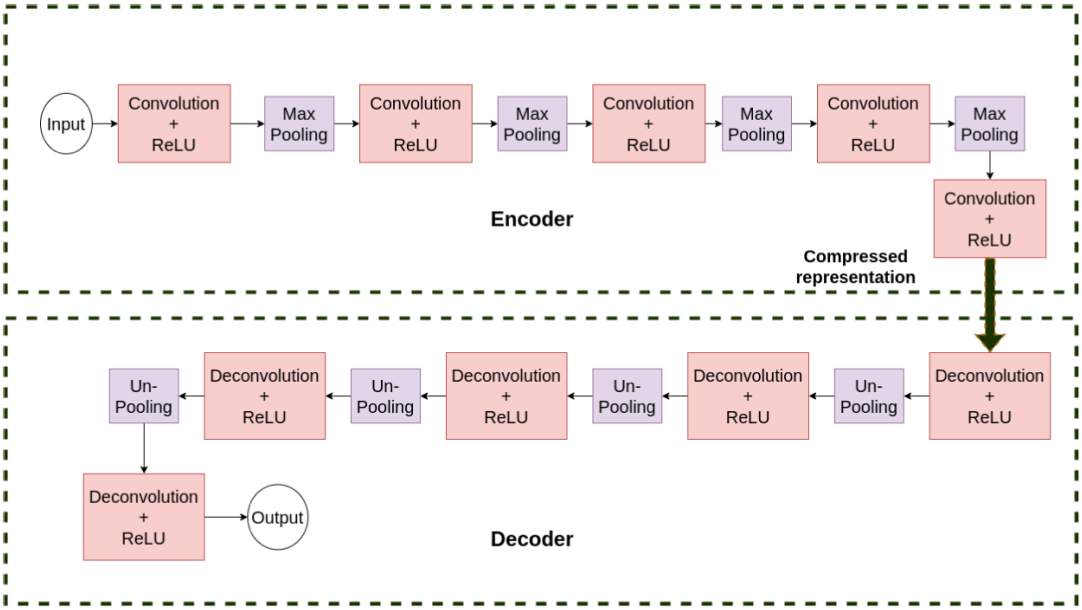


FIGURE 4. The CAE architecture used to initialize the convolutional layers' weights and biases of a *two-dimensional* CNN.

deconvolutional and unpooling layers to expand the encoder output to the same size as that of the input [63], [64]. Unpooling zeroes out all locations except those corresponding to the maximum values preserved by the pooling layers [63]. Spatial locality is preserved by using shared convolutional filter weights across all input locations [64]. If the input tensor to the CAE is x, then the resulting code after the encoder is given by

$$e_i = \sigma(x_i * W + b), \tag{3}$$

where $*$ denotes the 1D or 2D convolution depending on the input selected, W are the convolutional filter weights, b is the bias, and σ is the activation function [63]. The decoder reconstructs the output of the encoder using

$$z_i = \sigma(e_i * \tilde{W} + \tilde{b}), \tag{4}$$

where z_i, is the reconstruction of the i^{th} input, $\tilde{W}$ are the deconvolutional filter weights, $\tilde{b}$ is the bias at the decoder [63]. The CAE parameters to include W, $\tilde{W}$, b, and $\tilde{b}$ are adjusted through the use of backpropagation. Unsupervised pre-training of the CAE aims to minimize the loss function given by

$$E(\theta) = \sum_{i=1}^{m} (x_i - z_i)^2, \tag{5}$$

which measures the reconstruction error between the input x_i and the output z_i. The use of 1D and 2D RF-DNA fingerprints requires the use of 1D and 2D CAE architectures, which are shown in Fig 3 and Fig. 4, respectively.

F. FEATURE-ENGINEERED RF-DNA FINGERPRINTING

To compare the presented deep learning approach with prior feature-engineered RF-DNA fingerprinting approaches, our assessment includes results generated using the approach in [40]. The work in [40] classifies the same Cisco Wi-Fi cards mentioned in Sect. III-A and whose waveforms undergo the channel equalization process described in Sect. III-C. Following equalization, the Time-Frequency (TF) response of each waveform is calculated using the discrete Gabor transform (DGT) [65]. The DGT is computed using a Gaussian window of width $w_G = 0.015$, $M = 186$ total Gaussian window shifts, $K = 186$ total frequency values, and the window is advanced by $N_\Delta = 1$ samples between calculations, which equates to over sampling of the DGT.

The normalized magnitude-squared GT surface is calculated and subdivided into non-overlapping sub-regions that are $N_T = 12$ by $N_F = 10$ in dimension in accordance with [25], where N_T and N_F are the length of the sub-region along the time and frequency dimension, respectively. Each sub-region is reshaped into a $1 \times (N_T N_F) = 120$ length vector and statistical values of variance, skewness, and kurtosis calculated. Each sub-region's statistical features are sequentially concatenated to those of the prior sub-region. This process is repeated for the entire TF surface. The final three statistical features, composing an RF-DNA fingerprint, are calculated over the entire TF surface itself. The result is a $N_f = 363$ length feature-engineered RF-DNA fingerprint.

Classification of the feature-engineered RF-DNA fingerprints is conducted using the Multiple Discriminant Analysis/Maximum Likelihood (MDA/ML) classifier. MDA performs feature selection by linearly projecting the N_f-dimensional RF-DNA fingerprints into a $N_D - 1$-dimensional subspace that maximizes inter-class separability while simultaneously minimizing intra-class spread [66]. Following the projection, ML classification is facilitated by fitting a multivariate Gaussian distribution to

each of the classes' projected RF-DNA fingerprints. Finally, an unknown radio's projected RF-DNA fingerprints are assigned to the class whose associated likelihood function returns the largest value.

IV. METHODOLOGY

In this work, RF-DNA fingerprint–based SEI is performed using the CNN-learned features drawn from 1D and 2D representations of the 802.11a Wi-Fi preambles.

A. CNN USING ONE-DIMENSIONAL RF-DNA FEATURES

For 1D RF-DNA fingerprints, the time domain IQ samples of the 802.11a preamble are used. The CNN architecture used for 1D RF-DNA fingerprinting approaches is shown in Fig. 1.

The time-domain, channel-equalized preambles consist of $N_s = 320$ complex IQ samples due to the 16 μs duration preamble being sampled at a rate of 20 MHz. The training of deep learning networks requires large data sets, which is typically not the case when working with collected RF signal data sets [63]. Motivated by the results presented in [46], [47], this work adopts a signal partitioning scheme to provide data augmentation. This partitioning scheme has three advantages: (i) the size of the network is reduced by at least reducing the size of the input layer, (ii) the shift-invariant nature of the network is improved by training the CNN using shorter length sequences [46], and (iii) the partitioning of long sequences into multiple shorter sequences results in a larger data set for training the CNN.

Given a discrete-time RF signal comprised of N_s total samples, signal partitioning is implemented by sliding a N_b length window along the entire duration of the N_s long signal. The sliding window is advanced by one sample between consecutive windows; thus, two consecutive sub-sequences differ by one sample. Fig. 5 provides a representative illustration of the signal partitioning process applied to all 1D signals as well as their representations. The length of the sliding window is set to $N_b = 128$, which results in the $N_s = 320$ length sequence being partitioned into 193 complex-valued sub-sequences each of length $N_b = 128$. The partitioning results in shorter signals, but the number of data set entries increases to $193(2,000) = 386,000$ per Wi-Fi radio. A 2×128 real-valued tensor is input to the CNN shown in Fig. 1 Each tensor's first and second rows contain the I and Q sample values, respectively.

B. CNN USING TWO-DIMENSIONAL RF-DNA FEATURES

For 2D RF-DNA fingerprinting, the 802.11a preamble's TF representation is generated using the DGT, which is briefly explained in Sect. III-F. Selection of the DGT is due to (i) computational complexity being proportional to the sampling rate, (ii) being well-suited to degrading SNR when *over sampled*, and (iii) superior performance in prior RF-DNA fingerprinting work [25], [33], [37], [40]. For all TF-based results, the DGT is computed using the same variable settings as stated in Sect. III-F. Greater detail on the DGT, including

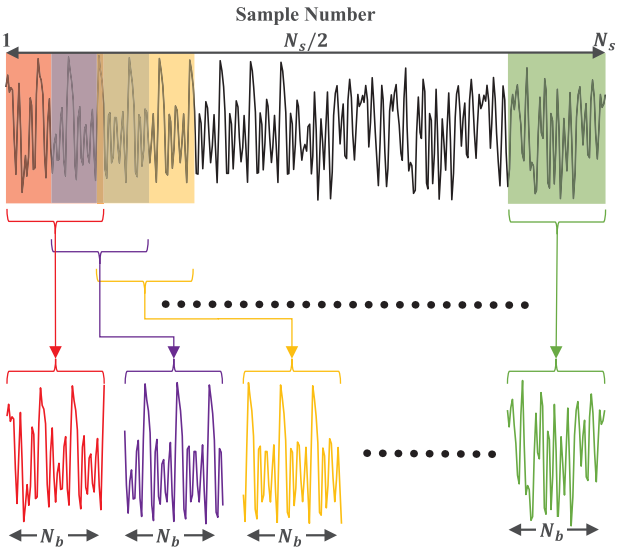


FIGURE 5. Representative illustration of the signal partitioning scheme applied to 1D sequences. In this case, the partitioning is applied to the in-phase samples of an 802.11a Wi-Fi preamble. The length of the sliding window, N_b, as well as the number of samples advanced between windows differs from that used to generate the results in Section V in an effort to improve visual clarity.

mathematical expression and use within the RF-DNA fingerprint generation process, can be found in [25].

1) IMAGE-BASED

For image-based RF-DNA fingerprinting using CNN, a 2D intensity representation is generated from the phase angle of the complex Gabor coefficients. This intensity image is 300×300 in size and is input to the CNN architecture described in Sect. III-D2 and illustrated in Fig, 2.

2) PARTITIONED TIME-FREQUENCY

For partitioned TF-based RF-DNA fingerprinting, each 300×300 image is partitioned along the time dimension. A 300×100 sliding window is applied to the original image and advanced by one sample/pixel to generate each sub-image. This process generated 201 sub-images per image, resulting in a data set of $201(2,000) = 402,000$ sub-images per Wi-Fi radio.

C. CNN INITIALIZATION USING CAE

A CNN's convolutional layers are often initialized using randomly generated weights and biases; however, such an approach can result in poor convergence and classification performance of the final model. This issue becomes an even greater detriment when using a training set comprised of a limited number of RF-DNA fingerprints/samples [63], [64]. We mitigate this issue herein by initializing the 1D and 2D CNNs using the corresponding CAE as described in Sect. III-E. The use of CAE initialization results in a two-stage training process: (i) unlabeled RF-DNA fingerprints are used to train the CAE via unsupervised learning, and (ii) the CNN is trained using labeled RF-DNA fingerprints and convolutional

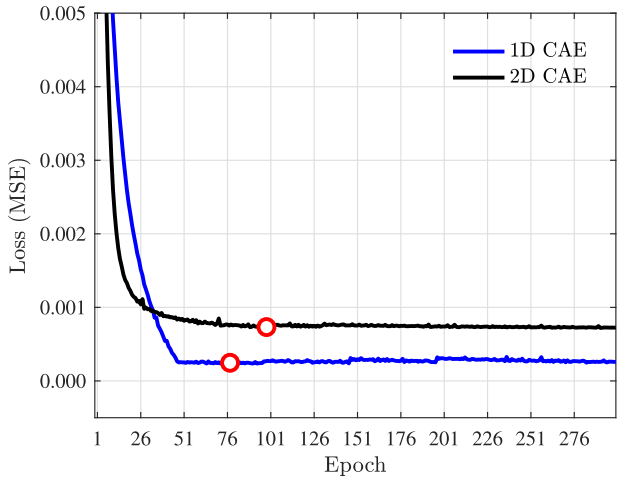


FIGURE 6. The loss function output after each training epoch of a 1D and 2D CAE at SNR = 9 dB.

layers whose weights and biases are initialized using the corresponding values of the trained CAE's encoder.

In general, the CAE training data is sampled into minibatches, and backpropagation is used to compute the gradient of the loss function over each minibatch [67]. During CAE training, the difference between the CAE's input, x_i, and the reconstructed sample, z_i, is calculated using a Minimum Mean Square Error (MMSE) loss function. In this work, minibatch sizes of 16 and 32 are used for the 1D and 2D representations of the CAE training data, respectively. These values were determined based on a grid search that is performed for three hyperparameters: (i) minibatch size, (ii) layer size, and (iii) hidden layers activation. The number of epochs, used to train a given CAE, is determined experimentally by: (i) setting the number of epochs to 300 and (ii) recording the loss function value calculated at the end of each epoch. The epoch corresponding to the smallest loss function value is selected as the total number of CAE training epochs. That process was conducted for all SNRs in the range of 9 dB to 30 dB. It was observed that the number of epochs needed to train the CAE increased as the SNR decreased. Fig. 6 shows that for an SNR of 9 dB, the minimum number of training epochs needed to achieve the desired loss value was 78 and 99 for the 1D and 2D CAE, respectively. Since an SNR = 9 dB represents the lowest SNR used in this work, all 1D and 2D CAEs are trained using a total of 78 and 99 training epochs, respectively (i.e., SNR $\geq$9 dB).

Following CAE training, the encoder's convolutional layers weights and biases are stored for CNN initialization. It is important to note that the selected (i.e., 1D or 2D) CNN's fully connected layers are always initialized randomly.

V. RESULTS

All results presented in this section are generated using 802.11a Wi-Fi preambles that have undergone Rayleigh fading as described in Sect. III-B. Each of the $N_B = 2,000$ preambles is convolved with a unique Rayleigh channel and like-filtered, scaled Additive White Gaussian Noise (AWGN) is used to achieve SNR $\in$ [9], [30] dB at 3 dB steps. Monte Carlo analysis is facilitated through the use of $N_z = 10$ AWGN realizations per SNR. Following convolution with the Rayleigh fading channel and SNR scaling, each preamble undergoes channel estimation and equalization in accordance with Sect. III-C. The resulting data set is normalized using the minimum and maximum values and then divided into two subsets: one for training of the CAE and CNN, and the other for use as a "blind" test set. A total of 200 preambles from each noise realization and radio (i.e., $200 \times 10 = 2,000$ per radio) are randomly selected to form the blind test set. The remaining preambles are used to train the 1D and 2D CAEs and CNNs.

Both of the CAEs are trained using the entire training set for each of the Wi-Fi radios without class/radio labels (i.e., unsupervised). For CNN model development (i.e., the second stage in the training process) the training set is further subdivided to facilitate five-fold cross-validation. A total of 1,800 preambles from each noise realization and radio (i.e., $1,800 \times 10 = 18,000$ per radio) is divided into five equally sized subsets. For each fold, four of the training subsets are used for model development, while the remaining subset is held out for model validation. This process is repeated five times such that each subset serves once in validating the developed CNN model. The CNN model that results in the highest classification accuracy, across all five folds and noise realizations, is selected as the "best" model. All CNN-based classification results are generated using the best CNN model and the blind test set. Classification performance is assessed using average percent correct classification, which is the average of at least 2,000 total decisions per radio at a given SNR.

A. RESULTS: CNN WITHOUT CAE INITIALIZATION

Initial assessment of CNN-based RF-DNA fingerprinting is conducted using 1D and 2D CNNs that use randomly initialized weights and biases. This approach serves two purposes:

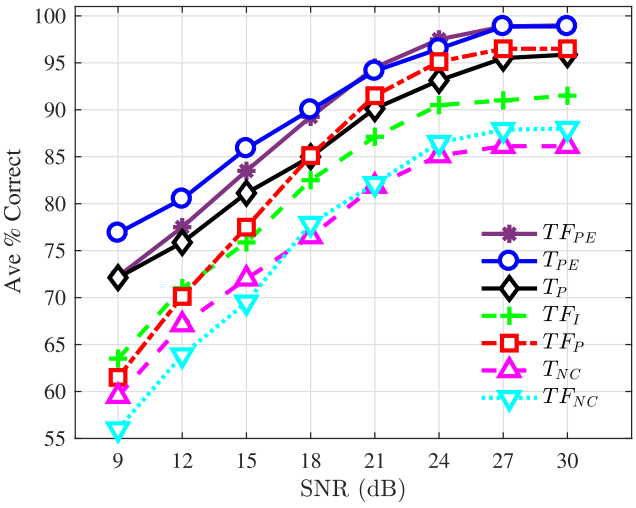
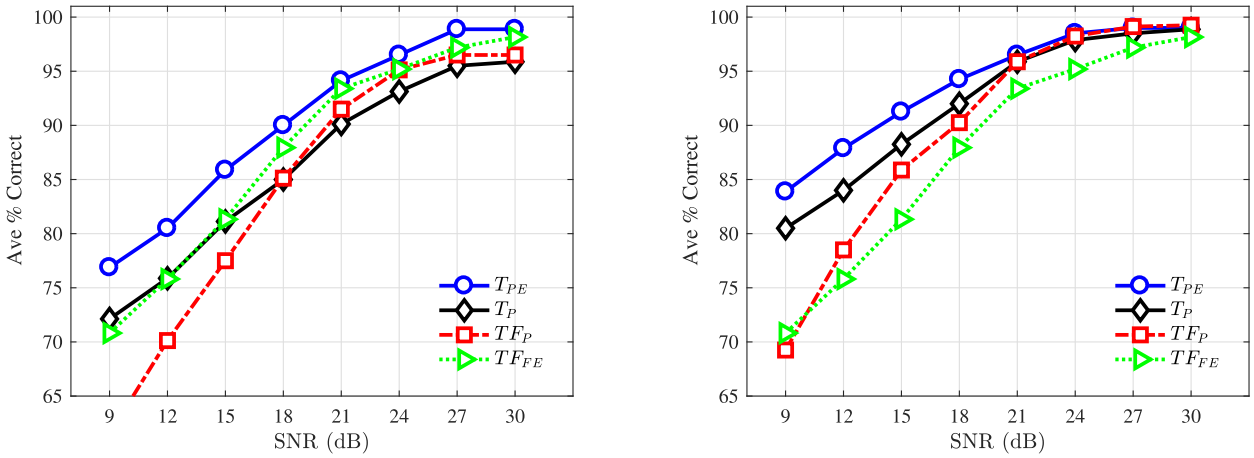


FIGURE 7. Average percent correct classification performance of 802.11a Wi-Fi preambles using seven different preamble representations and pre-processing scenarios for $L = 5$ Rayleigh fading channel conditions.



(a) *Randomly* initialized CNN results.

(b) *CAE* initialized CNN results.

FIGURE 8. Average percent correct classification performance of the T_P and TF_P results corresponding to *randomly* and *CAE*-initialized CNNs. These results are compared with the T_{PE} scenario and handcrafted RF-DNA fingerprinting, TF_{FE}, results generated in accordance with Sect. III-F and consistent with the work in [40].

(i) selection of the best RF-DNA fingerprint representation based on classification performance and (ii) facilitating comparison with prior CNN-based SEI publications.

The results in this section are generated using a Rayleigh fading channel that consists of $L = 5$ total paths. Based upon the 1D and 2D RF-DNA fingerprint representations presented in Sect. IV, three RF-DNA fingerprinting scenarios result: (i) partitioned time (T_P), (ii) TF phase images (TF_I), and (iii) partitioned TF images (TF_P). Four additional scenarios are also presented to aid in assessing the SEI performance of the developed CNN-based RF-DNA fingerprinting approaches. The first two of these additional scenarios corresponds to perfect estimation (PE) of the Rayleigh fading channel coefficients, which represents the case in which the exact value of α_k is known or achieved. Thus, PE eliminates the impact of channel estimation error from the classification results. Results for this scenario are generated using the partitioned time and time-frequency approaches and are designated as T_{PE} and TF_{PE}, respectively. It is important to note that PE is an *ideal* case that is impractical within operational systems. However, PE is included as a baseline against which the T_P, TF_I, and TF_P results can be compared. The remaining two scenarios use the 802.11a Wi-Fi preambles directly from the Rayleigh fading channel (i.e., no channel estimation nor correction (NC) is performed). These scenarios assess the CNN's ability to extract discriminating RF-DNA fingerprint features directly from signals that undergo Rayleigh fading, but not channel estimation and correction. The NC scenarios use RF-DNA fingerprints learned from the partitioned time signals T_{NC} and partitioned TF images TF_{NC}.

For the sake of brevity, clarity, ease of comparative analysis, and to facilitate selection of the best RF-DNA fingerprint approach, average percent correct classification performance is presented for each of the seven RF-DNA fingerprinting scenarios, Fig. 7. It is important to note that the results are generated using CNNs that were *not* initialized using the CAEs' encoder weights and biases. The use of CAE initialization may change the performance of an individual scenario, but the relation between scenarios is unchanged. The perfect estimation scenario T_{PE} resulted in the best average percent correct classification performance for SNR $\in$ [9], [30] dB. The partitioned TF images, TF_P, achieved average percent correct classification performance of 90% or greater for SNR $\geq$21 dB, which proved superior to the other four scenarios: T_P, TF_I, T_{NC}, and TF_{NC} at these SNRs. For SNR $\leq$18 dB, the partitioned time (T_P) scenario achieved superior average percent correct classification performance versus that of the TF_I, TF_P, T_{NC}, and TF_{NC} scenarios. The poorest average percent correct classification performance is associated with the T_{NC} and TF_{NC} scenarios across all SNRs. The poor performance is attributed to the Rayleigh fading effects that are unique for every preamble; thus, making it difficult for the CNNs to extract RF-DNA fingerprint features that are sufficiently unique and distinct to facilitate serial number discrimination. This suggests that channel correction is necessary unless channel agnostic features can be learned by the deep learning approach. Learning of channel agnostic features may be facilitated through the use of a larger training data set, alternate signal representation, or both. Based upon the results shown in Fig. 7, all subsequent results are generated using the partitioned time and TF scenarios only. Lastly, the remainder of this section is focused on presenting the results and analysis associated with modifying the presented approach for the purpose of improving classification performance.

B. RESULTS: CNN WITH CAE INITIALIZATION

Based upon the results presented in Fig. 7 and the discussion in Sect. V-A, 1D and 2D CAE-initialized CNN (CAE-CNN) RF-DNA fingerprinting assessments are conducted using the partitioned time (T_P), partitioned TF (TF_P), and perfect

estimation (T_{PE}) data sets. Fig. 8 presents average percent correct classification performance for randomly (Fig. 8(a)) and CAE (Fig. 8(b)) initialized CNNs directly overlaid with T_{PE} and feature-engineered, TF_{FE}, RF-DNA fingerprinting performance. When the CNN weights and biases are randomly initialized, TF_{FE} results achieve superior average percent correct classification performance for SNR $\geq$18 dB, Fig. 8(a). However, for SNR $\leq$15 dB, the T_P RF-DNA fingerprinting performance is as good as or marginally worse than that of the TF_{FE} results. For SNR $\geq$21 dB, TF_P-based RF-DNA fingerprinting results prove superior in average percent correct classification performance versus the T_P-based RF-DNA fingerprinting approach, Fig. 8(a).

Average percent correct classification results for CAE initialized 1D and 2D CNNs are compared with T_{PE} and TF_{FE} results in Fig. 8(b). When compared to the results in Fig. 8(a), CAE initialization of the CNNs improves the average percent correct performance of all three DL-based RF-DNA fingerprinting scenarios at each of the eight SNRs. The largest improvement occurs at SNR $=$ 9 dB and represents an approximate classification increase of 9% for the DL-based scenarios of T_P, TF_P, and T_{PE}. For SNR $\geq$21 dB, the T_P and TF_P average percent correct classification performance matches that of the perfect estimation scenario, T_{PE}. In contrast to the results in Fig. 8(a), T_P and TF_P average percent correct performance is superior to TF_{FE} for SNR $\geq$9 dB and SNR $\geq$12 dB, respectively. When excluding T_{PE}, T_P-based RF-DNA fingerprinting achieves superior average percent correct classification performance for SNR $\leq$18 dB. These results suggest that noise is more detrimental to the image-based, TF_P, RF-DNA fingerprinting approach than that of the time partitioned cases T_P and T_{PE}. This is because all three DL-based scenarios are conducted using the same set of signals (i.e., the same Rayleigh fading, estimation, and correction effects are present). Thus, when considering the architecture complexity and number of parameters associated with 2D CAE and CNN, as well as the results presented thus far, the partitioned time signal representation, T_P, is selected as the superior CNN-based RF-DNA fingerprinting approach under Rayleigh fading and noisy channel conditions. All subsequent results and investigations are performed using the T_P scenario and a CAE-CNN.

1) MODEL COMPLEXITY

Table 2 presents the number of hyperparameters and training times associated with the 1D and 2D CAE-CNNs used for the T_P, with $N_b = 64$, and TF_P RF-DNA fingerprinting scenarios, respectively. All training times are determined by running the selected architecture and scenario on a Dell i7 computer with a GTX GP1060 Graphics Processing Unit (GPU). The training time for the feature engineering, TF_{FE}, is included to complete its comparison with the DL approaches (Fig. 8(b)). The TF_{FE} training time encompasses the time needed to perform signal transformation, feature generation, and MDA/ML development using five-fold cross validation and a total of ten Monte Carlo trials at each SNR value. Based on Table 2, the TF_P scenario's 2D CAE-CNN requires the longest training time (33 hours) and most hyperparameters (roughly 5.4×10^6), representing the highest complexity of the three scenarios. Although the TF_{FE} scenario requires the least amount of training time, its performance is inferior to both DL approaches. The TF_P and TF_{FE} scenarios require the calculation of the GT coefficients. For a sampling frequency of 20 MHz, the GT requires a sequential calculation time of $\mathcal{O}\left(N_B^3\right)$ for a data set comprised of N_B preambles. When considering the training times and hyperparameters, as well as the results in Fig. 8(b), the 1D CAE-CNN architecture associated with the T_P scenario results in superior performance.

TABLE 2. Number of parameters and training times for the different architectures used in this paper.

Architecture	Scenario	Parameters	Training Time
1D CNN	T_P	81,918	11 hours
1D CAE	T_P	53,701	6 hours
2D CNN	TF_P	4,757,636	23 hours
2D CAE	TF_P	629,249	10 hours
MDA/ML	TF_{FE}	–	12 hours

2) SLIDING WINDOW LENGTH ANALYSIS

In an effort to improve percent correct classification performance, the optimal sliding window size, N_b, is determined using empirical assessment. Selection of the initial sliding window length of $N_b = 128$ was inspired by the work presented in [46]. However, the work in [46] performed SEI using IQ signals that lacked channel impairments such as noise and multipath; thus, a sliding window of length $N_b = 128$ may not facilitate maximum SEI performance when these impairments are present within the collected IQ signals/preambles. The sliding window assessment is conducted using (i) a Rayleigh fading channel consisting of $L = 5$ reflectors; (ii) an SNR $=$ 9 dB; (iii) a 1D

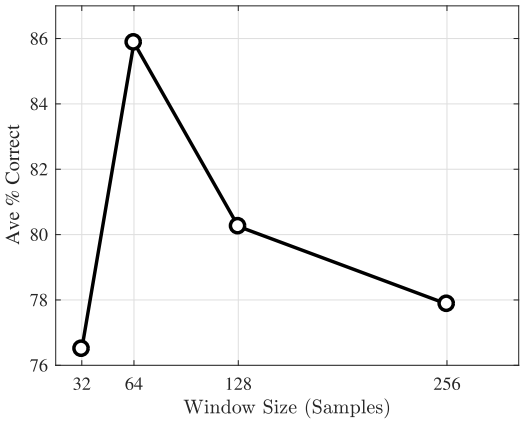


FIGURE 9. Average percent correct classification performance using time partitioned, T_P, 802.11a preambles; a Rayleigh fading channel comprised of $L = 5$ reflectors; sliding window lengths of N_b = [32], [64], [128], [256]; and 1D CAE-initiated CNN at an SNR of 9 dB.

CAE-CNN; and (iv) the 802.11a preambles' time partitioned, T_P, IQ samples that are partitioned using sliding windows of length $N_b = [32], [64], [128], [256]$. Average percent correct classification performance for each sliding window length is presented in Fig. 9. Based on these results, a $N_b = 64$ length sliding window results in superior classification performance, which results in a 6% classification performance increase versus that of the $N_b = 128$ length window.

3) ADDITIONAL FADING CHANNELS

All results presented up to this point used a Rayleigh fading channel comprised of $L = 5$ reflectors. This section presents results for three additional channel conditions of noise only and Rayleigh fading channels consisting of $L = 3$ and $L = 7$ reflectors. Fig. 10 shows average percent correct classification performance generated using the $N_D = 4$ 802.11a Wi-Fi radios' T_P preambles, partitioned using a sliding window of length $N_b = 64$, for each of the four channel conditions at SNR $\in$ [9], [30] dB in 3 dB steps. A 1D CAE-CNN is trained at each SNR of a given channel condition. Noise only classification performance is at 95% or higher for all investigated SNRs. The noise only results prove superior to those associated with the Rayleigh fading channels. This is attributed to the fact that the noise only preambles do not undergo fading nor subsequent channel estimation and correction, which can alter the SEI exploited features. For the $L = 3$ channel conditions, average percent correct classification performance of 91% or higher is achieved for all SNRs. For the $L = 5$ Rayleigh fading channel, average percent correct classification of 93% or higher is achieved for SNR $\geq$15 dB. Average percent correct classification performance is greater than 80% for the $L = 7$ Rayleigh fading channel at SNR $\geq$21 dB but fails to exceed 87% at any SNR. The $L = 7$ case represents the most challenging case in terms of classification performance. The decrease in classification performance is attributed to error that occurs within the N-M channel estimation process. The work in [40] shows that as the number of reflectors, L, increases so does the error associated with estimating the channel coefficient(s) h_k. This error is exacerbated as the SNR degrades. The inability to accurately estimate the channel coefficients carries forward into the correction stage and results in the SEI features being corrupted and/or lost, which inhibits the CNN's ability to sufficiently learn them as they change across preambles due to the time-varying nature of the channel. The estimation error can be reduced or eliminated by (i) improving the N-M channel estimator's accuracy as L increases and SNR decreases, (ii) developing or selecting a more accurate channel estimation approach, and/or (iii) developing an SEI approach capable of learning signal features that are channel invariant. The next section provides a brief investigation into channel-immutable signal features using CAE-CNNs. The first two error-reducing suggestions are left to future research.

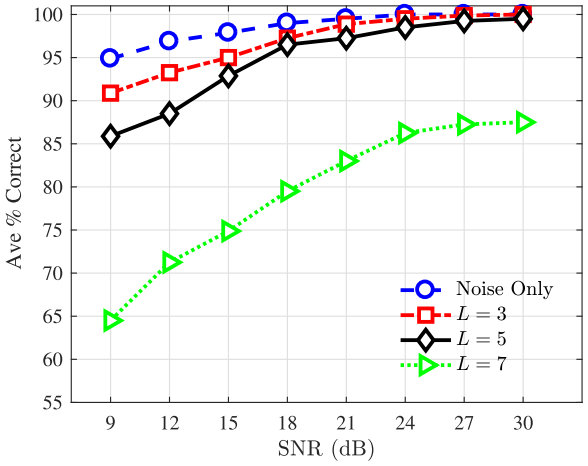
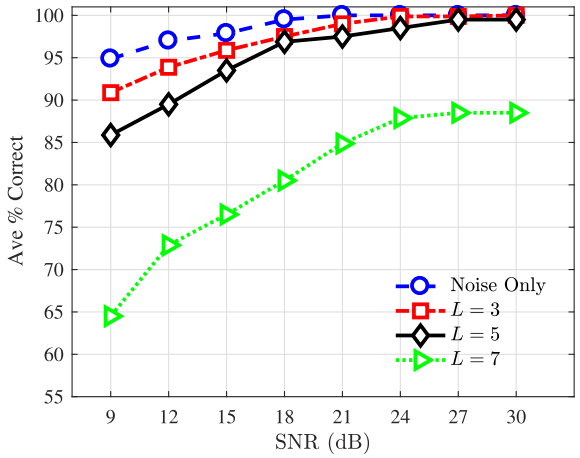
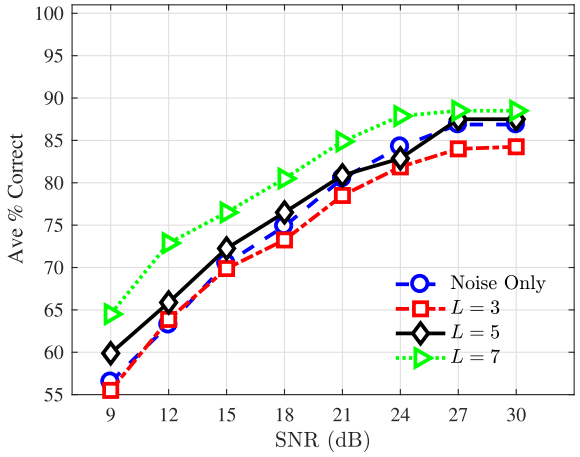


FIGURE 10. Average percent correct classification performance across the $N_D = 4$ 802.11a Wi-Fi radios using time partitioned, T_P, preambles; a sliding window of length $N_b = 64$; and CAE-CNNs for SNR $\in$[9], [30] dB in 3 dB steps.

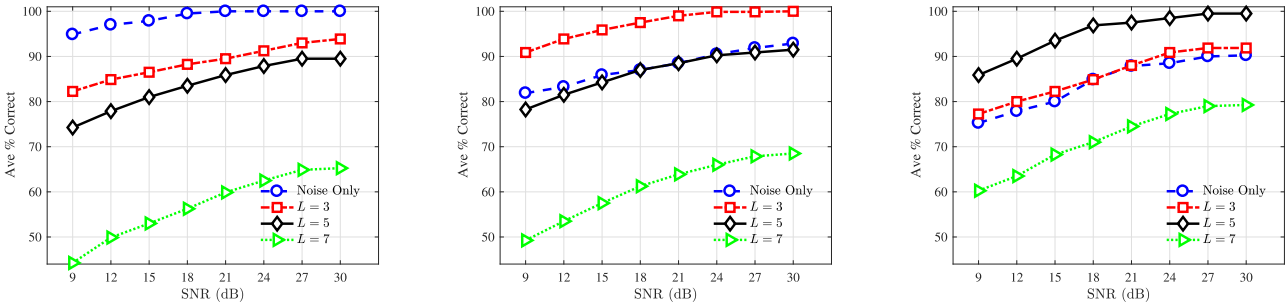


(a) Each 1D CAE-CNN trained at SNR=9 dB classifies only the T_P preambles corresponding with its same channel condition.



(b) The 1D CAE-CNN trained at SNR=9 dB and L=7 classifies the T_P preambles corresponding with every channel condition and SNR.

FIGURE 11. Average percent correct classification performance across the $N_D = 4$ 802.11a Wi-Fi radios using time partitioned, T_P, preambles, a sliding window of length $N_b = 64$, and 1D CAE-CNN.



(a) 1D CAE-CNN trained for the noise only channel at SNR=9 dB.

(b) 1D CAE-CNN trained for the L=3 channel at SNR=9 dB.

(c) 1D CAE-CNN trained for the L=5 channel at SNR=9 dB.

FIGURE 12. Average percent correct classification performance across the $N_D = 4$ 802.11a Wi-Fi radios using time partitioned, T_P, preambles, a sliding window of length $N_b = 64$, and a 1D CAE-CNN that is trained for a specific channel condition (e.g., $L = 3$) and SNR = 9 dB, but is used to classify all T_P preambles across the four channels and SNR $\in$ [9], [30] dB.

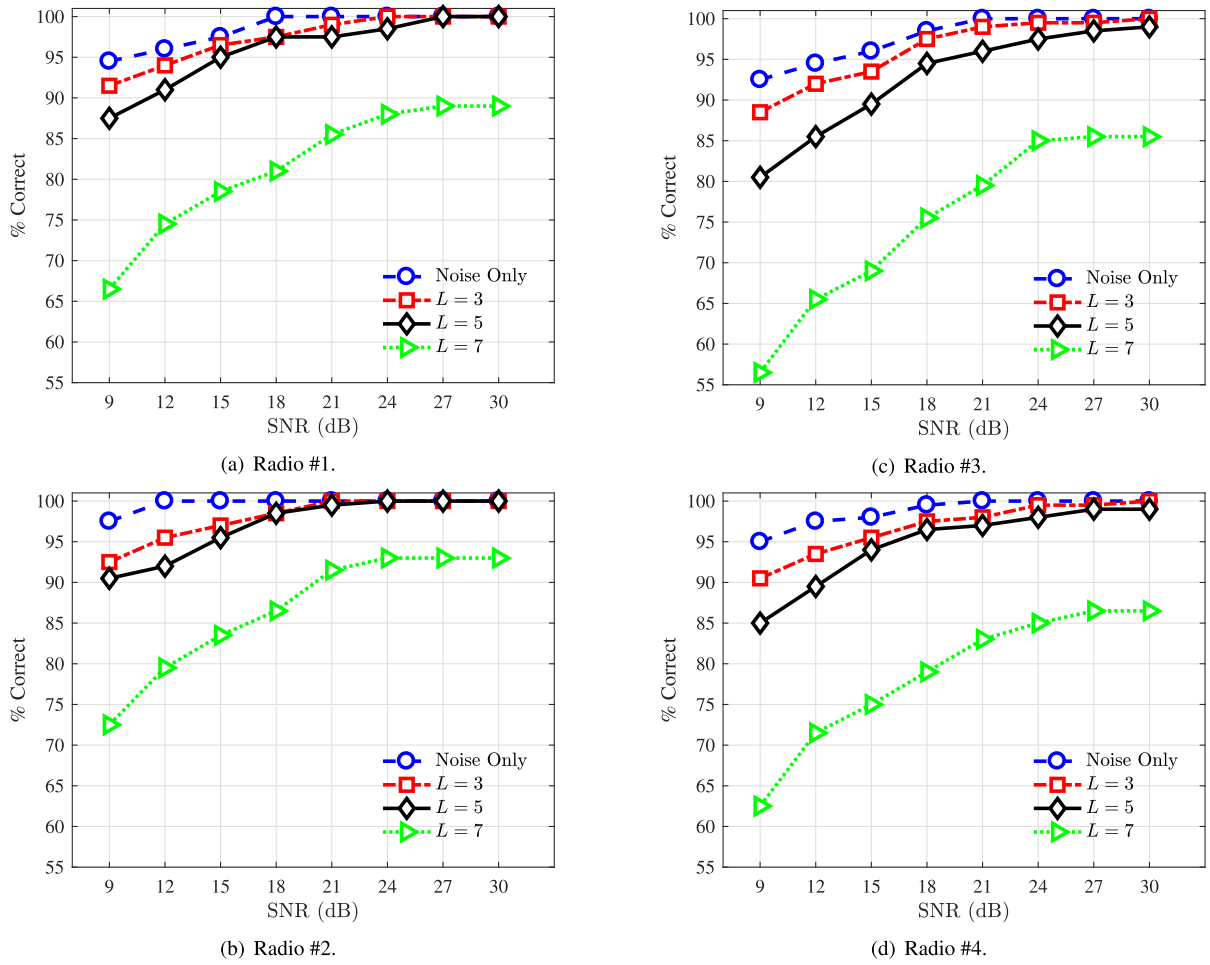
4) CHANNEL INVARIANT SEI

The results presented in Fig. 10 require training and storage of individual CAE and CNN models at each SNR for a selected channel condition. This results in a total of 16 models—eight CAE and eight CNN—for a given channel condition (e.g., $L = 3$) and 64 overall when considering all four investigated channel conditions. Thus, every CAE and CNN model is developed for a specific SNR and channel. This specificity can be problematic in two ways: (i) individual deep learning models may be hindered or prevented from learning SEI features that remain consistent as the SNR and/or channel changes, and (ii) storage of the developed deep learning models may become problematic as the number of radios and/or possible channel conditions (e.g., more multipath channel cases such as $L = [2], [9]$) increases. These issues are important to consider because SNR and multipath channel conditions are continuous, not discrete as investigated here. It would prove ideal if a single CAE-CNN pair could learn a set of SEI features that is invariant to the channel conditions (i.e., SNR and multipath).

Motivated by the results in Fig. 10 and the findings in [53], [54], the approach presented here investigates learned SEI features that are invariant to (i) noise or (ii) noise and Rayleigh fading. Fig. 11(a) presents average percent correct classification performance in which an individual 1D CAE-CNN pair is trained for each channel condition at SNR = 9 dB using the corresponding T_P preambles with $N_b = 64$, resulting in four 1D CAE-CNN trained pairs (i.e., one for each channel condition). When comparing the results in Fig. 11(a) to those in Fig. 10, average percent correct classification performance actually improves by 1% to 2% for all four channel conditions at SNR $\geq$12 dB. These results demonstrate that the SEI features learned at SNR = 9 dB remain consistent as the SNR improves. These results suggest that using noisy data to train the 1D CAE-CNN results in a more robust set of learned SEI features.

Investigation into noise- and Rayleigh fading–invariant SEI features is conducted by training a single 1D CAE-CNN using T_P preambles, with $N_b = 64$, for the $L = 7$ Raleigh fading channel at SNR = 9 dB. This approach is selected because it represents the most difficult case in terms of the channel conditions under which the CNN must learn discriminating SEI features. The resulting 1D CAE-CNN is used to classify the T_P preambles for every channel condition, noise only and $L = [3], [5], [7]$, at SNR $\in$ [9], [30] dB in 3 dB steps. Fig. 11(b) shows the average percent correct classification performance for this investigation. The results show that the SEI features learned by the 1D CAE-CNN, at $L = 7$ and SNR = 9 dB, are invariant to changes in SNR but not Rayleigh fading conditions. The poor classification performance for the $L \neq 7$ channel cases is attributed to poor channel estimation performance as discussed previously in Sect. V-B3. Since the 1D CAE-CNN never exceeds an average percent correct classification performance of 90% or higher when classifying $L = 7$ T_P preambles at any SNR, it is not surprising that it performs poorly when classifying the T_P preambles corresponding with the other three channel conditions.

Based on these observations, three additional investigations are conducted in which a 1D CAE-CNN model is developed for each of the remaining Rayleigh fading channels: (i) noise only, (ii) $L = 3$, and (iii) $L = 5$ at an SNR = 9 dB. Each of the trained 1D CAE-CNNs classifies the T_P preambles for all of the channel conditions and SNRs. The associated average percent correct classification performance associated with the (i) noise only, (ii) $L = 3$, and (iii) $L = 5$ cases are shown in Fig. 12(a), Fig. 12(b), and Fig. 12(c), respectively. The results in Fig. 12, along with those in Fig. 11(b), indicate that the RF-DNA fingerprint features learned for a particular multipath channel condition (e.g., $L = 5$) are insufficient to achieve the same level of discrimination when classifying preambles received under a different channel conditions (e.g., $L = 3$ or noisy only). Although the same level of discrimination is not achieved, the results suggest that some RF-DNA fingerprint feature commonality does exist between the preambles of two differing Rayleigh fading channels (e.g., train for $L = 5$ and classify $L = 3$). If RF-DNA feature commonality did not exist, then one would expect an average percent correct classification performance of 25% (i.e., a guess) or lower. The



(a) Radio #1. (c) Radio #3.

(b) Radio #2. (d) Radio #4.

FIGURE 13. Percent correct classification performance for each of the $N_D = 4$ 802.11a Wi-Fi radios using T_P preambles, a sliding window of length $N_b = 64$, and a 1D CAE-CNN that is trained at SNR = 9 dB for each channel condition: noise only and $L = [3], [5], [7]$. The trained 1D CAE-CNNs classify their corresponding multipath channel's T_P preambles for SNR $\in$ [9], [30] dB.

degradation in classification performance is attributed to the number of paths, L, that compose two different Rayleigh fading channels. For example, a Rayleigh fading channel with $L = 5$ paths will impact a preamble's inherent SEI features differently than a channel consisting of $L = 3$ paths. The CNN's RF-DNA feature learning is influenced by channel estimation inaccuracies as well as residual channel effects that remain after correction and differ across channel conditions. This issue persists whenever the 1D CAE-CNN is trained using T_P preambles for one channel condition (e.g., noise only) but used to classify those associated with a different channel (e.g., $L = 3$). Since the results shown in Fig. 11(b) and Fig. 12 fail to demonstrate multipath channel-invariant RF-DNA fingerprint feature learning, percent correct classification performance for each of the $N_D = 4$ 802.11a Wi-Fi radios is presented in Fig. 13 for the noise-invariant SEI feature learning case whose average percent correct classification performance is shown in Fig. 11(a).

VI. CONCLUSION

This paper presents RF-DNA fingerprint–based SEI using a CAE-initialized CNN (CAE-CNN) under Rayleigh fading and degrading SNR conditions. A total of seven RF-DNA fingerprinting scenarios were investigated to determine the approach best suited to maximizing serial number discrimination performance of four 802.11a Wi-Fi radios under noise only and three different Rayleigh fading channel conditions. These seven scenarios span both time (1D) and joint time-frequency (2D) representations of the 802.11a preambles. For the seven scenarios, SEI features learned from the preambles' partitioned, time IQ samples resulted in superior average percent correct classification performance across all four channel conditions and eight SNR values.

In an effort to maximize classification performance, the length of the sliding window, used to partition the preambles' IQ samples, was analyzed. This analysis revealed that average percent correct classification performance was optimal when a sliding window 64 samples in length is used to partition the raw IQ samples of the 802.11a preambles. The use of a 64-sample window improved classification performance by as much as 5% when compared to one of length 128.

In addition to analyzing the length of the sliding window, this work assessed the 1D CAE-CNN's ability to learn SEI

features that are invariant to noise or noise and Rayleigh fading effects. Although able to learn SEI features that are invariant to noise, the 1D CAE-CNN was not able to determine a set of features that remained invariant to changing multipath channel conditions. The inability to learn multipath-invariant SEI features is attributed to the channel estimation and correction processes used herein. One possible solution to this issue would be to train the 1D CAE-CNN using a data set comprised of signals representing each of the possible multipath channel conditions. One challenge associated with this solution would be in determining the number of signals needed to represent each of the possible channel conditions. This may not be a tenable approach when considering the continuum of possible channel conditions (e.g., number of paths, non-Rayleigh fading, etc.). Another alternative would be to replace the CNN with a Long Short-Term Memory (LSTM) architecture, because it is designed to handle sequences of data such as waveforms.

REFERENCES

[1] *Gartner Says 6.4 Billion Connected Things Will Be in Use in 2016, Up 30 Percent From 2015*, Gartner Res., Nov. 2015. [Online]. Available: https://www.gartner.com/en/newsroom/press-releases/2015-11-10-gartner-says-6-billion-connected-things-will-be-in-use-in-2016-up-30-percent-from-2015#:~:text=Gartner%2C%20Inc.,will%20get%20connected%20every%20day

[2] *Internet of Things Connected Devices to Triple by 2021, Reaching Over 46 Billion Units*, Juniper Res., Dec. 2016. [Online]. Available: https://www.juniperresearch.com/press/press-releases/%E2%80%98internet-of-things%E2%80%99-connected-devices-triple-2021

[3] Statista. (2019). *Internet of Things (IoT) Connected Devices Installed Base Worldwide From 2015 to 2025 (in billions)*. [Online]. Available: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/

[4] G. Maayan. (Jan. 2020). *The IoT Rundown For 2020: Stats, Risks, and Solutions*. [Online]. Available: https://securitytoday.com/articles/2020/01/13/the-iot-rundown-for-2020.aspx

[5] K. Rawlinson. (Jul. 2014). *HP Study Reveals 70 Percent of Internet of Things Devices Vulnerable to Attack*. [Online]. Available: https://www8.hp.com/us/en/hp-news/press-release.html?id=1744676

[6] I. Ray, D. M. Kar, J. Peterson, and S. Goeringer, "Device identity and trust in IoT-sphere forsaking cryptography," in *Proc. IEEE 5th Int. Conf. Collaboration Internet Comput. (CIC)*, Dec. 2019, pp. 204–213.

[7] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on Internet-scale IoT exploitations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2702–2733, 3rd Quart., 2019.

[8] S. Larsen. *A Smart Fish Tank Left a Casino Vulnerable to Hackers*. (Jul. 2017). [Online]. Available: https://money.cnn.com/2017/07/19/technology/fish-tank-hack-darktrace/index.html

[9] J. Wright and J. Cache, *Hacking Wireless Exposed: Wireless Security Secrets and Solutions*, 3rd ed. New York, NY, USA: McGraw-Hill, 2015.

[10] M. Stanislav and T. Beardsley, "Hacking IoT: A case study on baby monitoring exposures and vulnerabilities," Rapid7, Boston, MA, USA, Tech. Rep., pp. 1–17, 2015. [Online]. Available: https://information.rapid7.com/iot-baby-monitor-research.html

[11] J. Wright, "KillerBee: Practical ZigBee exploitation framework or 'wireless hacking and the kinetic world,'" in *Proc. 11th ToorCon Conf.*, San Diego, CA, USA, vol. 67, 2009, pp. 1–39. [Online]. Available: https://www.inguardians.com/works/

[12] S. Simon, "Internet of Things' hacking attack led to widespread outage of popular Websites," NPR, Washington, DC, USA, Tech. Rep., Oct. 2016. [Online]. Available: https://www.wbur.org/npr/498954197/internet-outage-update-internet-of-things-hacking-attack-led-to-outage-of-popular

[13] P. Shipley. (2014). *Insteon: False Security and Deceptive Documentation*. DEFCON. [Online]. Available: https://www.youtube.com/watch?v=dy1LTQLmPtM

[14] P. Shipley. (2015). *Tools for Insteon RF*. [Online]. Available: https://github.com/evilpete/insteonrf

[15] C. M. Talbot, M. A. Temple, T. J. Carbino, and J. A. Betances, "Detecting rogue attacks on commercial wireless insteon home automation systems," *Comput. Secur.*, vol. 74, pp. 296–307, May 2018.

[16] K. Sa, D. Lang, C. Wang, and Y. Bai, "Specific emitter identification techniques for the Internet of Things," *IEEE Access*, vol. 8, pp. 1644–1652, 2020.

[17] J. Dudczyk, J. Matuszewski, and M. Wnuk, "Applying the radiated emission to the specific emitter identification," in *Proc. 15th Int. Conf. Microw., Radar Wireless Commun.*, vol. 2, Jul. 2004, pp. 431–434.

[18] J. Pangm, B. Greenstein, R. Gummadi, S. Seshan, and D. Wetherall, "802.11 user fingerprinting," in *Proc. 13th Annu. ACM Int. Conf. Mobile Comput. Netw.* New York, NY, USA: Association Computing Machinery, 2007, pp. 99–110.

[19] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures," in *Proc. 14th ACM Int. Conf. Mobile Comput. Netw. (MobiCom)*. New York, NY, USA: Association Computing Machinery, 2008, pp. 116–127.

[20] W. Suski, M. Temple, M. Mendenhall, and R. Mills, "RF fingerprinting commercial communication devices to enhance electronic security," *Int. J. Electron. Secur. Digit. Forensics*, vol. 1, no. 3, pp. 301–322, Oct. 2008.

[21] B. Danev and S. Capkun, "Transient-based identification of wireless sensor nodes," in *Proc. IEEE Int. Conf. Inf. Process. Sensor Netw.*, Apr. 2009, pp. 25–36.

[22] R. Klein, M. A. Temple, M. J. Mendenhall, and D. R. Reising, "Sensitivity analysis of burst detection and RF fingerprinting classification performance," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2009, pp. 1–5.

[23] M.-W. Liu and J. F. Doherty, "Nonlinearity estimation for specific emitter identification in multipath channels," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1076–1085, Sep. 2011.

[24] I. O. Kennedy and A. M. Kuzminskiy, "RF fingerprint detection in a wireless multipath channel," in *Proc. 7th Int. Symp. Wireless Commun. Syst.*, Sep. 2010, pp. 820–823.

[25] D. Reising, "Exploitation of RF-DNA for device classification and verification using GRLVQI processing," Ph.D. dissertation, Air Force Inst. Technol., Dec. 2012. [Online]. Available: https://apps.dtic.mil/dtic/tr/fulltext/u2/a572506.pdf

[26] M. D. Williams, S. A. Munns, M. A. Temple, and M. J. Mendenhall, "RF-DNA fingerprinting for airport WiMax communications security," in *Proc. 4th Int. Conf. Netw. Syst. Secur.*, Sep. 2010, pp. 32–39.

[27] D. Takahashi, Y. Xiao, Y. Zhang, P. Chatzimisios, and H.-H. Chen, "IEEE 802.11 user fingerprinting and its applications for intrusion detection," *Comput. Math. Appl.*, vol. 60, no. 2, pp. 307–318, Jul. 2010.

[28] O. H. Tekbas, O. Ureten, and N. Serinken, "Improvement of transmitter identification system for low SNR transients," *Electron. Lett.*, vol. 40, no. 3, pp. 182–183, Feb. 2004.

[29] K. J. Ellis and N. Serinken, "Characteristics of radio transmitter fingerprints," *Radio Sci.*, vol. 36, no. 4, pp. 585–597, Jul. 2001.

[30] S. S. Soliman and S.-Z. Hsue, "Signal classification using statistical moments," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 908–916, May 1992.

[31] "Interferometric intrapulse radar receiver for specific emitter identification and direction-finding," Defence R D Canada, Ottawa, ON, Canada, Tech. Rep. REW 224, Jun. 2007.

[32] E. Azzouz and A. Nandi, *Automatic Modulation Recognition of Communication Signals*. New York, NY, USA: Springer, 2013.

[33] D. R. Reising, M. A. Temple, and J. A. Jackson, "Authorized and rogue device discrimination using dimensionally reduced RF-DNA fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 6, pp. 1180–1192, Jun. 2015.

[34] C. G. Wheeler and D. R. Reising, "Assessment of the impact of CFO on RF-DNA fingerprint classification performance," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2017, pp. 110–114.

[35] Y. Pan, S. Yang, H. Peng, T. Li, and W. Wang, "Specific emitter identification based on deep residual networks," *IEEE Access*, vol. 7, pp. 54425–54434, 2019.

[36] *Local and Metropolitan Area Networks, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std 802.11-2007, Jun. 2007.

[37] M. K. M. Fadul, D. R. Reising, T. D. Loveless, and A. R. Ofoli, "RF-DNA fingerprint classification of OFDM signals using a Rayleigh fading channel model," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–7.

[38] M. Fadul, "The Impact of Rayleigh fading channel effects on the RF-DNA fingerprinting process," M.S. thesis, Dept. Elect. Eng., Univ. Tennessee Chattanooga, Chattanooga, TN, USA, Aug. 2018.
[39] F. Kandah, J. Cancelleri, D. Reising, A. Altarawneh, and A. Skjellum, "A hardware-software codesign approach to identity, trust, and resilience for IoT/CPS at scale," in *Proc. Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData)*, Jul. 2019, pp. 1125–1134.
[40] M. Fadul, D. Reising, T. D. Loveless, and A. Ofoli, "Using RF-DNA fingerprints to classify OFDM transmitters under Rayleigh fading conditions," *IEEE Trans. Inf. Forensics Security*, Jan. 2020. [Online]. Available: https://arxiv.org/abs/2005.04184
[41] G. Baldini and R. Giuliani, "An assessment of the impact of wireless interferences on IoT emitter identification using time frequency representations and CNN," in *Proc. Global IoT Summit (GIoTS)*, Jun. 2019, pp. 1–6.
[42] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks," *Electron. Lett.*, vol. 55, no. 2, pp. 90–92, Jan. 2019.
[43] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
[44] L. J. Wong, W. C. Headley, S. Andrews, R. M. Gerdes, and A. J. Michaels, "Clustering learned CNN features from raw I/Q data for emitter identification," in *Proc. MILCOM-IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2018, pp. 26–33.
[45] G. J. Mendis, J. Wei-Kocsis, and A. Madanayake, "Deep learning based radio-signal identification with hardware design," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2516–2531, Oct. 2019.
[46] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, "Deep learning convolutional neural networks for radio identification," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 146–152, Sep. 2018.
[47] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 160–167, Feb. 2018.
[48] F. Restuccia, S. D'Oro, A. Al-Shawabka, M. Belgiovine, L. Angioloni, S. Ioannidis, K. Chowdhury, and T. Melodia, "DeepRadioID: Real-time channel-resilient optimization of deep learning-based radio fingerprinting algorithms," in *Proc. 20th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jul. 2019, pp. 51–60.
[49] H. Jafari, O. Omotere, D. Adesina, H.-H. Wu, and L. Qian, "IoT devices fingerprinting using deep learning," in *Proc. MILCOM-IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2018, pp. 1–9.
[50] L. Guyue, Y. Jiabao, Y. Xing, and A. Hu, "Location-invariant physical layer identification approach for WiFi devices," *IEEE Access*, vol. 7, pp. 106974–106986, Aug. 2019.
[51] K. Youssef, L. Bouchard, K. Haigh, J. Silovsky, B. Thapa, and C. V. Valk, "Machine learning approach to RF transmitter identification," *IEEE J. Radio Freq. Identificat.*, vol. 2, no. 4, pp. 197–205, Dec. 2018.
[52] J. Yu, A. Hu, F. Zhou, Y. Xing, Y. Yu, G. Li, and L. Peng, "Radio frequency fingerprint identification based on denoising autoencoders," in *Proc. Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2019, pp. 1–6.
[53] T. Jian, B. C. Rendon, E. Ojuba, N. Soltani, Z. Wang, K. Sankhe, A. Gritsenko, J. Dy, K. Chowdhury, and S. Ioannidis, "Deep learning for RF fingerprinting: A massive experimental study," *IEEE Internet Things Mag.*, vol. 3, no. 1, pp. 50–57, Mar. 2020.
[54] J. Robinson, S. Kuzdeba, J. Stankowicz, and J. M. Carmack, "Dilated causal convolutional model for RF fingerprinting," in *Proc. 10th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2020, pp. 0157–0162.
[55] H. Lajos, A. Yosef, W. Li, and J. Ming, *MIMO-OFDM for LTE, Wi-Fi, and WiMAX*. Hoboken, NJ, USA: Wiley, 2011.
[56] M.-W. Liu and J. F. Doherty, "Specific emitter identification using nonlinear device estimation," in *Proc. IEEE Sarnoff Symp.*, Apr. 2008, pp. 1–5.
[57] *Agilent E3238 Signal Intercept Collection Solutions: Family Overview*, Agilent Technol., Santa Clara, CA, USA, Jul. 2004.
[58] B. O'Hara and A. Petrick, *IEEE 802.11 Handbook: A Designer's Companion* (IEEE Standards Wireless Networks). Piscataway, NJ, USA: IEEE Press, 2005.
[59] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, Jan. 1965.
[60] K. Wang, M. Faulkner, J. Singh, and I. Tolochko, "Timing synchronization for 802.11a WLANs under multipath channels," in *Proc. Australian Telecommun., Netw. Appl. Conf.*, Jul. 2003, pp. 1–5.
[61] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Commun. Lett.*, vol. 9, no. 7, pp. 619–621, Jul. 2005.
[62] J. Patterson and A. Gibson, *Deep Learning A Practitioners Approach*. Newton, MA, USA: O'Reilly Media, 2017.
[63] M. S. Seyfioglu, A. M. Ozbayoglu, and S. Z. Gurbuz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1709–1723, Aug. 2018.
[64] J. Masci, U. Meier, D. Ciresan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Artificial Neural Networks and Machine Learning—ICANN 2011*, T. Honkela, W. Duch, M. Girolami, and S. Kaski, Eds. Berlin, Germany: Springer, 2011, pp. 52–59.
[65] M. J. Bastiaans and M. C. W. Geilen, "On the discrete Gabor transform and the discrete Zak transform," *Signal Process.*, vol. 49, no. 3, pp. 151–166, Mar. 1996.
[66] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. Hoboken, NJ, USA: Wiley, 2001.
[67] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

MOHAMED K. M. FADUL received the B.S. degree in electrical and electronics engineering from the University of Khartoum, Khartoum, Sudan, in 2012, and the M.S. degree in electrical engineering from The University of Tennessee at Chattanooga, Chattanooga, TN, USA, in 2018, where he is currently pursuing the Ph.D. degree in computational engineering. He is also working as a Research Assistant with The University of Tennessee at Chattanooga. His research interests include software-defined radios, wireless device discrimination using RF distinct native attribute fingerprints, and deep learning.

DONALD R. REISING (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of Cincinnati, Cincinnati, OH, USA, in 2006, and the M.S. and Ph.D. degrees in electrical engineering from the Air Force Institute of Technology, Dayton, OH, USA, in 2009 and 2012, respectively.

From 2008 to 2014, he was an Electronics Engineer with the U.S. Air Force Research Laboratory, Wright-Patterson Air Force Base, Dayton. He is currently an Associate Professor of electrical engineering with The University of Tennessee at Chattanooga. His research interests include wireless device discrimination using RF-distinct native attribute fingerprints, compressive sensing, cognitive radio, and deep learning. He is a member of Eta Kappa Nu and Tau Beta Pi.

MINA SARTIPI (Senior Member, IEEE) received the B.S. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2003 and 2006, respectively.

She is currently the Guerry Professor of computer science and engineering and the Founding Director of the Center for Urban Informatics and Progress (CUIP), The University of Tennessee at Chattanooga. Her work aims to coordinate cross-disciplinary research and strategic visions for urbanism and smart cities advancement with a focus on people and quality of life. She has conducted research on wireless communications, intelligent mobility, data acquisition, data transmission, and data analysis for more than 15 years. She has expertise in spectrum, dynamic spectrum access, and user allocation.