The effect of Human body Shadowing in ZigBee Radio Frequency Fingerprinting Identification

Raya Alhajri
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
r.alhajri@liverpool.ac.uk

Alan Marshall
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
alanm@liverpool.ac.uk

Guanxiong Shen
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
Guanxiong.Shen@liverpool.ac.uk

Miguel López-Benítez
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
M.Lopez-Benitez@liverpool.ac.uk

Valerio Selis
Department of Electrical Engineering and Electronics
University of Liverpool)
Liverpool, United Kingdom
V.Selis@liverpool.ac.uk

Junqing Zhang
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
junqing.zhang@liverpool.ac.uk

***Abstract*—Security is a major concern for the Internet of Medical Things (IoMT), however, many of these devices are limited in their capabilities. Physical layer (PHY) security measures can be used to prevent unauthorized access by exploiting intrinsic emitter characteristics such as Specific Emitter Identification (SEI), known as Radio Frequency Fingerprinting Identification (RFFI). RFFI at the IoMT is a promising approach to secure wearable Body Sensor Network (WBSN) devices. In this paper, we evaluated the effect of human body shadowing on Radio Frequency Fingerprinting Identification (RFFI) systems. Results show that shadowing has a serious impact on RFFI models. We also show that it can be mitigated by applying log-normal shadowing augmentation. Results obtained from simulations and experimental trials show that the classification accuracy increases when the multipath channel model and shadowing block size of 640 are used. A new system model for classifying devices using RFFI is then proposed. The proposed model achieved better classification accuracy when evaluated using unseen shadowed data.**

***Index Terms*—RF impairment, Deep learning, radio frequency fingerprint, emitter identification, ZigBee, device authentication.**

I. Introduction

THERE has been massive growth in Internet of things (IoT) embedded devices in various applications, including smart cities, healthcare, agriculture, industry, etc. [1]. However, this has been accompanied by substantial growth in security vulnerabilities. Cyber threats against IoMT are critical in that human life can be harmed. Therefore, calls for urgent response are required to secure these devices and save human lives. The lack of computational capacity in many IoT devices leads to challenges in the device authentication procedure. Furthermore, using passwords and Media Access Control(MAC) as standard authentication techniques are susceptible to attacks such as impersonation [2]. Hence, using a non-cryptographic technique as a device identification solution at the physical layer, such as radio frequency fingerprinting identification (RFFI), is a new option to eliminate spoofing attacks. RFFI recently considered recognizing authenticated transmitters using the unique variations imposed by hardware impairments such as I/Q imbalance, phase noise, frequency offsets, etc. [3].

Handcrafted feature extraction and deep learning-based approaches deploy RFFI as a passive physical layer authentication technique. The former approach extract features manually, which requires protocol domain knowledge, while the latter simplifies RFFI application by identifying hardware features automatically regardless of the protocols used. Higher accuracy and better feature extraction performance have been noticeably shown in deep learning compared to the handcrafted approaches [4]. Data representation [5], [6], network architectures [7]–[10], and different hardware imperfections [11]–[13] would be the most widely used techniques to extract unique features and obtain better classification results.

However, the attributes of an emitted signal change when it flows across a wireless channel, leading to degradation of the

We acknowledge the funding support received from the Department of Information Technology, University of Technology and Applied Sciences-Sur, Sultanate of Oman.

979-8-3503-9676-8/22/$31.00 ©2022 IEEE

RFFI accuracy. [14], was the first work that investigated the effects of channel distortions on analog signal fingerprints. They extracted Power Spectral Density (PSD) as a device identifier fingerprint and investigated a multipath channel fading with a constant Doppler shift using the Rician fading channel model. They concluded that the accuracy of RFFI under low multipath fading is mostly unaffected compared to medium and high multipath fading. In 2019, ORACLE was proposed to mitigate the channel effect in a dynamic environment towards adopting a channel-robust RFFI systems [13]. The researchers proposed to deploy demodulated data representation instead of raw data, in addition to the feedback-driven.

Data augmentation has recently gained high consideration in developing channel robustness RFFI models. For instance, multipath channel, including Rayleigh and path loss, was used in [15] with a constant doppler shift. Soltani et al. [16] proposed another channel robustness approach without prior knowledge of the transmitted waveform, where data augmentation was applied to the raw IQ data samples at both the transmitter and receiver sides. [17] considered doppler shift as a fixed parameter while Shen et al. [18] investigated Multipath with a variation in doppler shift. Authors in [19] evaluated the effect of feeding neural networks with datasets from different channels, indicating that their model achieves better accuracy when different channels are seen at the training stage.

However, there is no work, as far as we know, that has investigated the effect of human body shadowing on the RFFI models. Therefore, this paper investigates the effect of shadowing on the performance of RFFI models.

This paper is organized as follows. Section II presents an overview of the RFFI system, including the different steps of building the RFFI model. Section III shows the implemented experimental design. The experimental results and analysis are given in section IV, while section V concludes the paper.

II. RFFI System

RFFI is considered as an authentication technology exploiting hardware imperfections as a unique intrinsic characteristic. ZigBee emitters operating at 2.4GHz use the Orthogonal Quadrature Phase Shift Keying (OQPSK) technique for chip modulation. The modulated signal enters a Digital-to-Analog Converter (DAC) and then passes through quadrature upconversion mixers, power amplifiers (PA) and the antenna before it is transmitted into the wireless channel. The receiver captures the transmitted signal, downconverts, and demodulates it. Within this process, oscillators, amplifiers, mixers, etc., produce frequency deviations that can be used as device identifiers. The variations of crystal oscillators produce a carrier frequency offset (CFO) and phase noise where mixers produce I and Q mismatch and power amplifiers generate nonlinear distortion. As shown in Fig.1, RFFI model includes different steps after receiving emitter signals: A) Preprocessing, B) Data Augmentation, C) Training and feature extraction, and D) Classification.

A. Preprocessing

The basic steps in RFFI require preprocessing the received signal via: a) a synchronization and preamble extraction. b) The CFO is compensated and estimated to maintain system stability. c) The signal is passed through a normalization process. Then the data is ready for deep learning training and classification.

- **Synchronization and Preamble extraction:** To prepare the RFFI dataset, first the beginning of a packet should be determined to prevent performance degradation and then a region of interest (ROI) is extracted as an input to the deep neural networks. In this paper, we extracted the Synchronization Header(SHR) as the ROI.
- **CFO Compensation:** Many researchers have investigated CFO and proved that CFO compensation can help to achieve accurate classification [17], [18], [20].
- **Normalization:** Normalizing the extracted preamble helps neural networks to preserve and learn device-specific characteristics. We compute the normalization as follows:

$$y[n] = y[n]/rms(y[n]) \tag{1}$$

Where y[n] denotes the preamble part.

B. Data Augmentation

Data augmentation techniques are used to synthetically enhance the input datasets fed at training deep learning models by transforming the existing samples into new augmented samples. Data augmentation proved its efficiency in strengthening the RFFI model against variations in channel and noise during the inference phase [15]–[18].

C. Training Neural Network and feature extraction

Once the data is ready, input is fed to a neural network for learning the unique features that allow a trained network to distinguish between devices. The uniqueness of each transmitter's RF fingerprint results from manufacturing imperfections is extracted as a feature vector.

D. Classification

RFFI model classification is a process of using features to distinguish and identify the correct signal's emitter. Classification algorithms may be supervised or unsupervised such as The K Nearest Neighbor(KNN), Support Vector Machine(SVM), Neural Network(NN), k-Means clustering, and Decision Tree.

III. Experimental Design

The physical layer specification at ZigBee devices is defined by IEEE 802.15.4 [21]. Low-power, low-cost, short-range characteristics allow ZigBee devices to be utilized in a variety of IoT applications, such as medical care systems [22], home automation [23], Intelligent agriculture [24], sensor networks [25],etc. Quadrature Phase Shift Keying (QPSK) data modulation and Direct Sequence Spread Spectrum (DSSS) technologies are used in ZigBee devices with a 32-chip length when the device operates at 2.4 GHz.

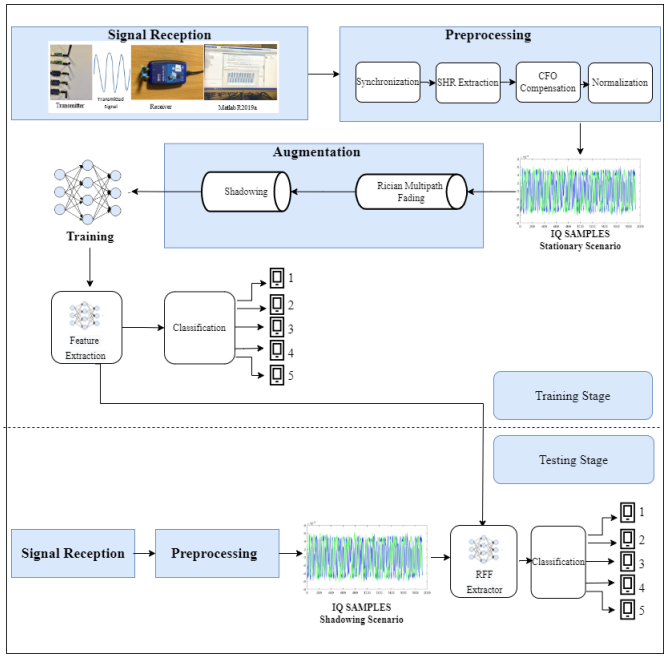

Fig. 1. RFFI System Model

A. Setup

This experiment consists of five CC2531 ZigBee emitters, which operate at 2.4 GHz, and one Adalm-Pluto SDR, As shown in Fig. 2, is used to capture the transmitted packets in Matlab employing Communications Toolbox Support Package for USRP Radio. The center frequency at both transmitter and receiver is set to 2.450 GHz (Channel 20). The Sampling Rate of RF signals is configured at 12 Msamples/s which is ten times oversampling compared to the 2 Mchip/s rates at the 2.450 GHz ZigBee band. The Synchronization header (SHR) field of the ZigBee physical format consists of two parts (preamble (4 bytes) and SFD (1 byte)). SHR field with ten symbols length of the ZigBee Physical layer waveform was identified and used to extract RF fingerprint patterns.

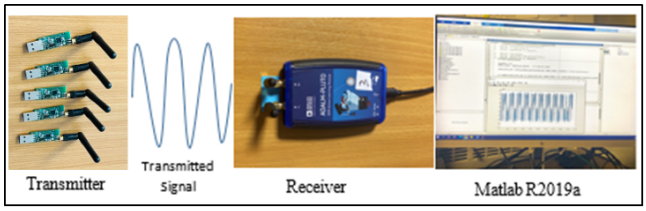

Fig. 2. RFFI experimental setup

B. Data Collection

Data were collected in three different scenarios (described below), where transmitters are located within one meter from the receiver, and CC2531 RF signal transmission is employed by Texas Instruments SmartRF Studio 7 software. The receiver captured 3000 packets with approximately 40K samples per frame. The normalized received signal is shown in Fig. 3. The datasets of each transmitter were captured on different days and saved for further processing in Matlab.

1) Stationary scenario: Transmitter and receiver placed on a table in a Line of Sight(LOS). The height from the floor was 75 cm. This data was used to train a CNN network;
2) Shadowing scenario.1: Transmitter and receiver nodes were stationary on a table, but the transmission was blocked by a human body;
3) Shadowing scenario.2: Transmitter and receiver nodes were stationary on a chair at 55 cm height, but the transmission was captured while a human body was walking back and forth between them;

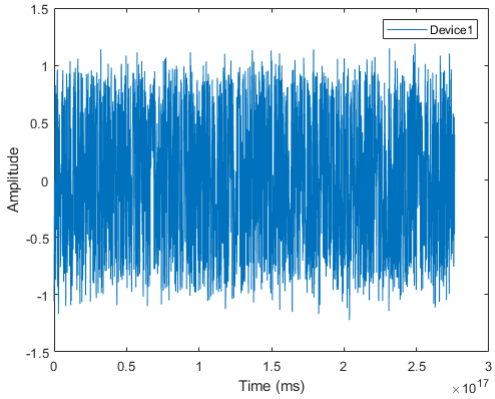

Fig. 3. Normalized Received ZigBee waveform

C. Data Augmentation

In this paper, we use a Shadowing and Rician Multipath channel model as described below:

1) Rician Multipath: This model characterizes multipath by using the exponential Power Delay Profile (PDP) as follows:

$$P(p) = \frac{1}{\tau_d} e^{-pT_s/\tau_d}, \quad p = 0, 1, \cdots, p_{max}, \tag{2}$$

Where the RMS delay spread is denoted as τ_d,

Following the multipath parameters used in [18], we used a uniformly distributed RMS delay spread in a range of [10,300]ns and the Rician K-factor in a range of [0,10]. Additive white Gaussian noise (AWGN) was included when only the multipath channel model was selected, where it ranged from zero to 60 dB. The signal will pass through the multipath model first, and then the output from this is input to the shadowing model.

2) Shadowing: Log-normal fading was used in this paper to emulate shadowing with the mean μ and standard deviation of σ. Therefore, the input signal is divided in blocks where the mean average Signal Noise Ratio (SNR) in each block will change according to the log-normal distribution. In this paper, Shadowing augmentation is carried out as follows:

- A vector generation of mean SNR values from zero to 60 dB;

Parameter / Dataset	Blocks	Stationary						Shadowed (Scenario.1)						Shadowed (Scenario.2)
Experiment No.		***Exp.1***	***Exp.2***	***Exp.3***	***Exp.4***	***Exp.5***	***Exp.6***	***Exp.7***	***Exp.8***	***Exp.9***	***Exp.10***	***Exp.11***	***Exp.12***	***Exp.13***
Multipath			Yes	Yes	Yes	**Yes**	Yes		Yes	Yes	Yes	**Yes**	Yes	**Yes**
Shadowing	**348**			Yes						Yes				
	480				Yes						Yes			
	640					**Yes**						**Yes**		**Yes**
	960						Yes						Yes	
Classification Accuracy		**99%**	95.78%	95.42%	96.06%	**97.8%**	98.54%	76%	79.4%	69.70%	74.30%	**83.3%**	75.80%	**83.4%**

TABLE I
CLASSIFICATION ACCURACY VERSUS AUGMENTATION PARAMETER

- Standard deviation (σ) distributed randomly in a range of [1,12]dB, which is configured according to the level of shadowing;
- Set the size of shadowing blocks in each signal. We investigated the use of 384, 480, 640, and 960 blocks;
- Divide the number of signal samples (1920) by the number of blocks configured to get the number of samples in each block;
- The value of SNR per block is calculated by adding one value from the mean SNR vector to the randomly selected σ value;
- The previous step is then repeated for the rest of the mean SNR values;

D. CNN Design

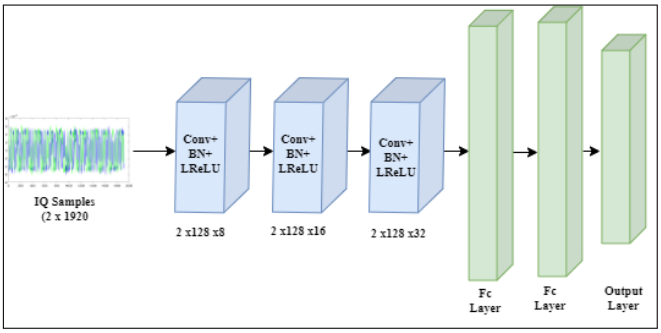

Fig. 4. CNN Architecture

Neural networks are trained to verify the genuineness of a single transmitting device amongst a group of otherwise similar devices. Due to the outstanding performance in image identification and computer vision, Convolutional Neural Network (CNN) has attracted several researchers to use this approach in RFFI systems [13], [26]–[28]. CNN is also adopted in this paper, as shown in Fig. 4, consisting of three convolutional layers and two fully connected layers. Each Convolutional (Conv) layer is followed by batch normalization (BN) and leaky ReLU activation (LReLU) layer. The filter sizes of the three Conv. layers were 2 x 128 with 8, 16, and 32 filters, respectively. The output of the third convolutional layer is input into a fully connected layer for classification, with the softmax activation function used to identify the signal with its corresponding device. In the Matlab Deep Learning Toolbox, an Adam optimizer with an initial learning rate of 0.0001 and 30 epochs is used for training the CNN.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

According to IoMT devices, including wearable devices, we demonstrated classification accuracy under static and dynamic conditions where shadowing was tested. As shown in Table I, the trained network performs better when the trained and tested data are stationary with a classification accuracy above 95%. However, it degrades sharply to 76% when shadowed data is used at the inference phase where unseen datasets are used. Therefore, to enhance the RFFI model performance, static data is augmented and trained using the Rician multipath and shadowing model. When data is augmented using a multipath channel model only, the performance increases by only 3% when tested with the unseen dataset. However, passing the augmented data into a shadowing model assists in classifying signals correctly to their corresponding emitters. We have augmented and trained a network in 348, 480, 640, and 960 shadowing blocks to identify and achieve the best classification accuracy. As shown in Figure.5, the best classification accuracy is achieved when multipath and shadowing channel models with 640 shadowing blocks are used. We have further validated the model using new datasets when shadowing performed by walking between the transmitter and receiver, achieving 83.4%.

V. CONCLUSION

In this work, we propose an RFFI system model that is able to distinguish devices in static and dynamic conditions where human body shadowing affects the transmission. We have evaluated the model in different scenarios: a) trained and tested with static dataset achieves above 95%, b) trained with static and tested with shadowed dataset achieves 83.4% when the multipath and shadowing with 640 blocks are used. Therefore, we have demonstrated that human body shadowing severely affects the RFFI systems. Further experiments are carried out to enhance the proposed model and achieve better classification accuracy in shadowing scenarios. However, future work is to look at how the model can be applied to other radio communication systems.

VI. ACKNOWLEDGMENT

This work is supported by the Ministry of Higher education, Research, and Innovation- Sultanate of Oman under the National Postgraduate Scholarship Programme.

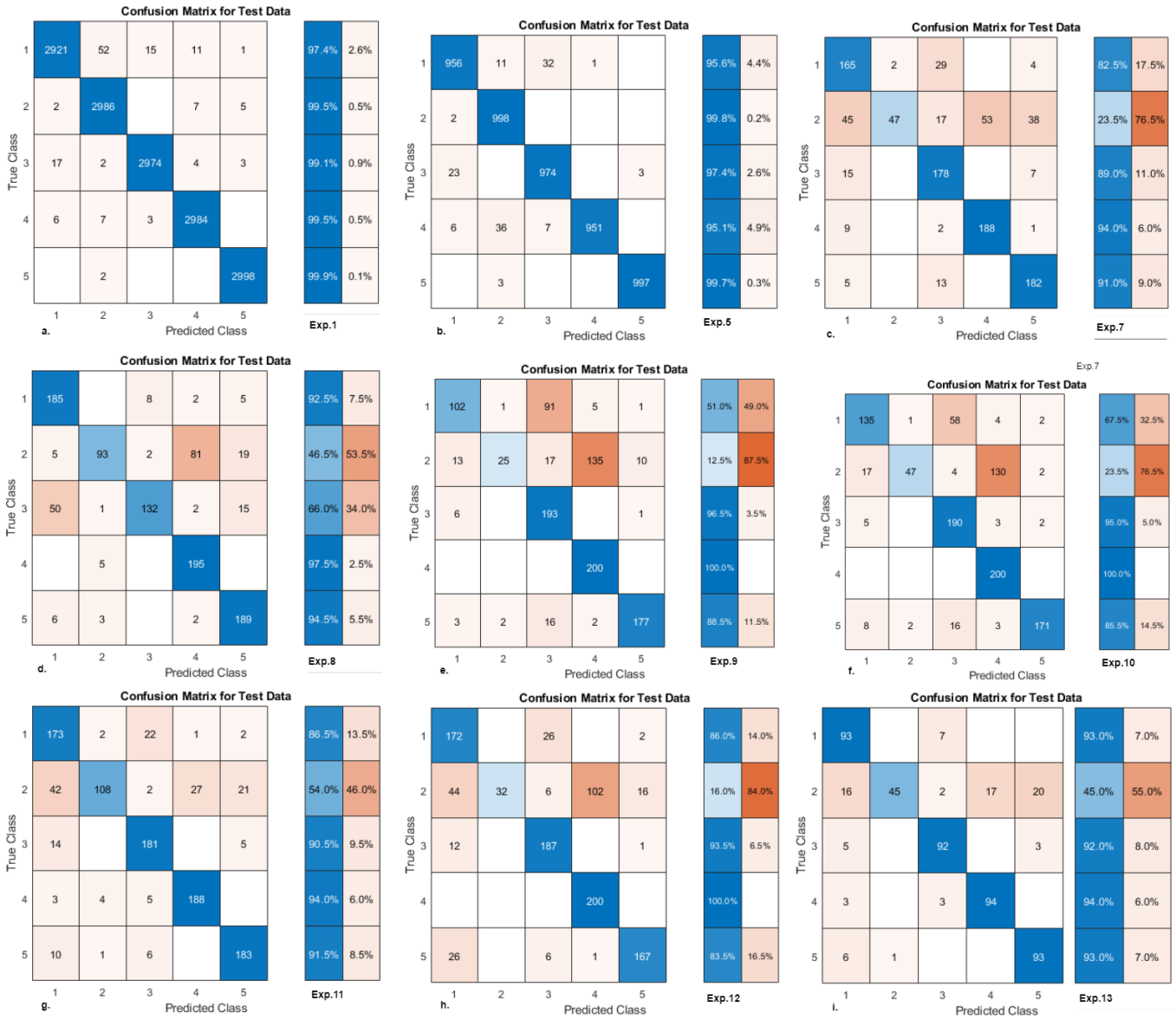

Fig. 5. CNN confusion Matrix for: a. Stationary Datasets without Multipath or shadow, b. Stationary Datasets with Multipath& 640 shadow blocks, c. Shadowing Datasets without Multipath or shadow, d. Shadowing Datasets with Multipath only, e. Shadowing Datasets with Multipath & 348 shadow blocks, f. Shadowing Datasets with Multipath & 480 shadow blocks, g. Shadowing Datasets with Multipath & 640shadow blocks, h. Shadowing Datasets with Multipath & 960 shadow blocks, i. Shadowing Datasets with Multipath & 640shadow blocks (Scenario.2)

REFERENCES

[1] A. Bahga and V. K. Madisetti, "Internet of things: A hands-on approach," 2014.

[2] J. Hall, M. Barbeau, and E. Kranakis, "Radio frequency fingerprinting for intrusion detection in wirless networks," *IEEE Transactions on Dependable and Secure Computing*, 2005.

[3] H. Yuan and A. qun Hu, "Preamble-based detection of wi-fi transmitter rf fingerprints," *Electronics Letters*, vol. 46, pp. 1165–1167, 2010.

[4] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, "Deep learning convolutional neural networks for radio identification," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 146–152, 2018.

[5] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks," *Electronics Letters*, vol. 55, no. 2, pp. 90–92, 2019.

[6] G. Baldini, R. Giuliani, and F. Dimc, "Physical layer authentication of internet of things wireless devices using convolutional neural networks and recurrence plots," *Internet Technology Letters*, vol. 2, no. 2, p. e81, 2019.

[7] I. Agadakos, N. Agadakos, J. Polakis, and M. R. Amer, "Deep complex networks for protocol-agnostic radio frequency device fingerprinting in the wild," *arXiv preprint arXiv:1909.08703*, 2019.

[8] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for rf device fingerprinting in cognitive communication networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.

[9] H. Jafari, O. Omotere, D. Adesina, H.-H. Wu, and L. Qian, "Iot devices fingerprinting using deep learning," in *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*. IEEE, 2018, pp. 1–9.

[10] C. Morin, L. S. Cardoso, J. Hoydis, J.-M. Gorce, and T. Vial, "Transmitter classification with supervised deep learning," in *International Conference on Cognitive Radio Oriented Wireless Networks*. Springer, 2019, pp. 73–86.

[11] L. J. Wong, W. C. Headley, and A. J. Michaels, "Emitter identification using cnn iq imbalance estimators," *arXiv preprint arXiv:1808.02369*, 2018.
[12] S. Hanna, S. Karunaratne, and D. Cabric, "Deep learning approaches for open set wireless transmitter authorization," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5.
[13] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "Oracle: Optimized radio classification through convolutional neural networks," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 370–378.
[14] S. U. Rehman, K. W. Sowerby, S. Alam, I. T. Ardekani, and D. Komosny, "Effect of channel impairments on radiometric fingerprinting," in *2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 2015, pp. 415–420.
[15] K. Merchant and B. Nousain, "Enhanced rf fingerprinting for iot devices with recurrent neural networks," in *MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM)*. IEEE, 2019, pp. 590–597.
[16] N. Soltani, K. Sankhe, J. Dy, S. Ioannidis, and K. Chowdhury, "More is better: Data augmentation for channel-resilient rf fingerprinting," *IEEE Communications Magazine*, vol. 58, no. 10, pp. 66–72, 2020.
[17] M. Cekic, S. Gopalakrishnan, and U. Madhow, "Robust wireless fingerprinting: Generalizing across space and time," *arXiv preprint arXiv:2002.10791*, 2020.
[18] G. Shen, J. Zhang, A. Marshall, and J. R. Cavallaro, "Towards scalable and channel-robust radio frequency fingerprint identification for lora," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 774–787, 2022.
[19] J. A. Gutierrez del Arroyo, B. J. Borghetti, and M. A. Temple, "Considerations for radio frequency fingerprinting across multiple frequency channels," *Sensors*, vol. 22, no. 6, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/6/2111
[20] S. Andrews, "Extensions to radio frequency fingerprinting," 2019.
[21] S. C. Ergen, "Zigbee/ieee 802.15. 4 summary," *UC Berkeley, September*, vol. 10, no. 17, p. 11, 2004.
[22] M.-S. Pan and Y.-C. Tseng, "Zigbee and their applications," in *Sensor Networks and Configuration*. Springer, 2007, pp. 349–368.
[23] A. Wheeler, "Commercial applications of wireless sensor networks using zigbee," *IEEE communications magazine*, vol. 45, no. 4, pp. 70–77, 2007.
[24] Z. Xiaojing and L. Yuanguai, "Zigbee implementation in intelligent agriculture based on internet of things," *EMEIT*, 2012.
[25] M. Terada, "Application of zigbee sensor network to data acquisition and monitoring," *Measurement Science Review*, vol. 9, no. 6, p. 183, 2009.
[26] R. Xie, W. Xu, Y. Chen, J. Yu, A. Hu, D. W. K. Ng, and A. L. Swindlehurst, "A generalizable model-and-data driven approach for open-set rff authentication," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4435–4450, 2021.
[27] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for lora using spectrogram and cnn," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 2021, pp. 1–10.
[28] J. Zhang, R. Woods, M. Sandell, M. Valkama, A. Marshall, and J. Cavallaro, "Radio frequency fingerprint identification for narrowband systems, modelling and classification," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3974–3987, 2021.