Human Presence Detection via Deep Learning of Passive Radio Frequency Data

Jenny Liu, Asad Vakil
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI
jennyliu443@gmail.com, avakil@oakland.edu

Robert Ewing
Center for Innovative Radar Engineering
Air Force Research Laboratory
WPAFB, OH
robert.ewing.2@us.af.mil

Xiaoping Shen
Department of Mathematics
Ohio University
Athens, OH
shenx@ohio.edu

Jia Li
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI
li4@oakland.edu

***Abstract*— Human presence detection is a critical field in certain circumstances such as natural disasters and surveillance systems. This paper presents a new approach that utilizes software defined radio to passively collect radio frequency data and applying deep learning neural network to detect human presence. It provides a low cost and environment friendly solution. The long term goal of this study is to develop a deep learning based spectrum monitoring system.**

Keywords—human presence, deep learning, passive radio frequency, software defined radio

I. Introduction

For thousands of years, our ancestors have been continuously exploring and innovating human detection technology. The field of human detection has many important applications, ranging from surveillance to security and is a critical component in the fields such as disaster relief. Even in less extreme applications, such as assisted living, hospitals, or smart homes, simply detecting the presence of a person is almost always the first step to any monitoring system. Human detection technology increases the efficiency of these systems, which can be lifesaving in many situations. For any form of security application, the ability to contact police sooner is crucial for the users and for assisted living applications, providing medical treatment sooner can be lifesaving. During any kind of natural disaster, the faster survivors are located and rescued, the higher chance of survival for the rescuee. While for that particular application, there are other analog methods that can be applied to the problem, this does not mean that the ability to autonomously detect human presence is not an integral technology.

In other applications human detection is an indispensable part of most forms of security and surveillance systems. As actual sentries being used for human detection is not practical from a cost standpoint, most home security systems are traditionally reliant on autonomous systems. The system becomes vulnerable when human detection is not reliable, and such vulnerabilities are exploitable. This however becomes more difficult when the structure and preexisting detection architecture is damaged or destroyed, and normal communication methods are broken off or facing interference. Dependable human detection is a key technology to actively protect and save lives, both privately and nationally.

Many solutions have been implemented to solve the age old problem of human detection. However, existing human presence detection technologies such as camera, lidar, radar, and ultrasonic sensors, all have their individual strengths and weaknesses. Cameras, for example, can provide visualized images, which are easier to process with existing methods, but can be restricted by factors such as lighting. In addition to that, optical modalities like cameras can be considered invasive which is less desirable in private applications. Lidar and radar systems both require signal emitters, which can be expensive and can cause interference with the existing wireless system. Ultrasonic sensors and all the technologies mentioned above are constrained by factors such as installation angle and position. These modalities can also be blocked or disabled by physical interference or jamming. An environment friendly, passive, and low cost human presence detection solution is required to compensate for the deficiencies of these exiting technologies.

This paper explores feasibility of identifying the presence of one or more people inside a closed space by using passive radio frequency (RF) signals via deep learning neural network, which to the best of our knowledge, is the initial research in this aspect. The main contributions of this paper are: a new environment friendly and low cost approach to detect human presence in a closed space by collecting passive RF wireless signals from surrounding environment, description of a system built during the experiment to implement our idea, and a convolutional neural network (CNN) model to classify human presence that takes wireless RF raw data as input and produces detection results, and experimental results as an illustration of the feasibility of our proposed approach.

In the following sections, Section II introduces related researches had been done and the advantages of our proposal. Section III explains our technical approach and experimental

978-1-7281-1416-3/19/$31.00

design. Section IV demonstrates the experimental results. Conclusion and future research are discussed in Section V.

II. BACKGROUND

A. Related work

Deep learning has shown its effectiveness in many fields such as automatic speech recognition, image recognition, visual art processing, natural language processing, customer relationship management, recommendation systems, financial fraud detection, etc. Recently, some researchers have initialized the study of radio signal modulation recognition and wireless interference identification by using convolutional neural network (CNN) through the collected passive RF data. In [1], experiment was conducted to classify different modulation formats. Paper [2] presented the research work of deep learning based radio signal classification by comparing CNN and residual neural network (RNN). However the studies in [1] and [2] primarily focused on the characteristics of wireless signals themselves instead of their applications. Authors of [3] introduced an approach to detect and identify a specific radio transmitter uniquely among other similar devices by using software defined radio (SDR) and CNN. Researchers of [4] have also conducted an experiment to classify the emitter of the wireless signal. Article [5] depicted the experiments of using CNN and deep neural network (DNN) to identify rogue RF transmitters. But[3], [4] and [5] focused on the scope of the wireless system. The study conducted in [6] showed a CNN system being used to assess the quality of human actions via RF wireless signals. However, the research in [6] used an active emitter to send wireless signals rather than using passive RF. Human presence detection is addressed by research work in [7] where RFID tags were used in the experiment for human detection and behavior classification instead of passive RF. The research of [8] and [9] are focused on the analysis of human activities by using deep learning to process wireless RF signals. However, active radio signals were still used in these experiments. Passive RF data was utilized to detect objects in paper [10] but deep learning was not used in this study. By utilizing a deep learning neural network for wireless signals classification, the network can potentially achieve better performance in a complex wireless signal environment. None of the studies mentioned above and papers mentioned in [11] used wireless passive RF signals to classify the human presence inside a closed space through a deep learning neural network. Based on the existing research, the feasibility of using deep learning to analyze passive RF data to detect human presence in an area of concern, is addressed in this paper.

B. Advantages

The proposed approach has several benefits. Firstly, the usage of passive RF data shares some of the same traits with passive radar systems in which no actively transmitted signals are required, and the object is detected through third party emitters. In addition to that, both passive radar and the proposed solution have low power consumption and are difficult to detect. Both solutions can be used to find a moving target and monitor an air space when the target is not visually observable. Because the solutions do not use an active emitter and only collect passive RF signals from the surrounding environment, the solution does not introduce radio spectrum pollution into the increasingly crowded wireless space. This approach does not generate any interference with the existing wireless system due to only collecting passive RF data. A desirable trait as wireless signals transmission is restricted in certain areas.

Due to the nature of the modality, the system possesses a larger detection coverage and is not as limited by factors such as installation angle and position, unlike other methods. Because the solution is reliant on passive RF, the installation costs and complexity are greatly reduced. Ambient RF signals exist everywhere, which can be utilized for human subject detection. Therefore, this approach is not limited by location. Nor is it limited by factors such as light or weather conditions either. Further investigation of the impact of extreme weather conditions such as thunder and lightning to the system is still required. In addition, the solution also costs less without active emitter present.

III. TECHNICAL APPROACH AND EXPERIMENT DESIGN

This research is conducted under assumption that human subjects will produce signatures in the collected RF signals of the corresponding location. The presence of human subjects, the size and the speed of the subjects will alter the RF signals, and the subtle variation can be detected by the neural network. In this experiment, the presence of one or more people in a closed space such as an office room or a home study room is addressed.

Traditional signal processing algorithms are not suited for processing very complex patterns; no existing formula or algorithm has been attested to solve this problem; there is no evidence to prove this is a linear problem. Deep learning is noted for having excellent pattern recognition capabilities and excellent performance for solving nonlinear problems with unknown relationships. Motivated by recent advances and the remarkable success of CNN, this research focuses on applying convolutional neural network to solve this problem. Shared weights and biases greatly reduce the number of parameters involved in a convolutional neural network. The convolutional layer will reduce the number of parameters it needs to get the same performance as the fully-connected model. It will result in faster training for the convolutional model, and ultimately help to build deeper networks. The pooling layers simplify the information in the output from the convolutional layer. In detail, a pooling layer takes each feature map output from the convolutional layer and prepares a condensed feature map. With the computation capability of CNN, it can be trained with enormous data by consuming less time comparing to the fully connected deep neural network [12].

In order to teach CNN model to detect human presence, adequate training data needs to be collected. Software defined radio (SDR) is adopted by our research to collect passive RF signals. SDR is a radio communication system where components that have been implemented in hardware are implemented by software on a personal computer or embedded system. SDR defines a collection of hardware and software technologies where some or all the radio's operating functions are implemented through modifiable software or firmware operating on programmable processing technologies. There are several benefits of using SDR to collect the RF raw data, such as being easy to process with software programs, having a wide range of utility, and providing a cost effective means of

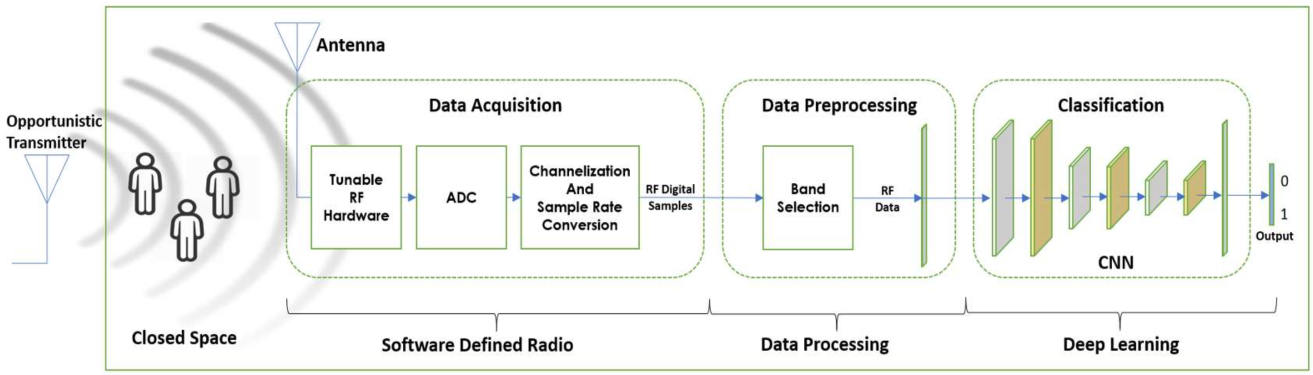

Figure 1. Passive RF signal human presence detection system

implementing software upgrades [13]. SDR has been successfully used for RF monitoring systems that use kurtosis and energy based spectrum detection [14]. In [15], an inexpensive and generic spectrum alerting system based on SDR is designed and implemented to discover rogue or unidentified RF signals.

Passive RF signal human presence detection system was developed during our experiment and is described in Figure 1. It is composed of three subsystems: data acquisition, data preprocessing, and classification. The antenna collects the passive RF signals in a closed space sent by opportunistic transmitters. These signals are in turn preprocessed by software defined radio (SDR) and then converted from analog signals to digital raw stream data. From there, the raw stream data is then preprocessed before it is fed into CNN model. Finally, the person presence probability is calculated by CNN model and the classification result is sent through the output layer. The details of the experiment are given in the following subsections, including RF signal acquisition, RF signal pre-processing, experimental scenarios design, CNN model training and human presence detection.

A. *RF signal acquisition*

Passive RF raw data collection is described in Table 1. RTL2832U is used to collect RF raw data at two separate locations, a home study room and an office, with and without human presence. Labels are assigned to RF raw data automatically during data collection. The SDR continuously scans the spectrum from the lowest frequency 24MHz to the highest frequency 1760MHz. RF raw data is collected, with and without known primary signals such as FM, TV, and cellular passive signals, at the locations of interest. Selective frequency band and full frequency band RF raw data is collected. A total of 197 selective bands are chosen by adaptive step, meaning that small scan steps are used for active bands and large scan steps are used for inactive bands. Step size is set based on FCC Table of Frequency Allocations [16], observation of frequency spectrum at collecting location through SDR and local radio station frequency list. Full band includes all frequency bands with an even step size of 1.2MHz. Both selective bands and full band is scanned with the same setting of sample rate, duration and period as listed in Table 1.

TABLE 1. PASSIVE RF DATA COLLECTION

Items	Description
Collection Device	RTL2832U
Location	Closed space: an office and a home study room
Human Presence	0: No person is in a closed space 1: One or more person are in a closed space
Data Labelling	Automatically assign scenario ID (0 or 1) and location ID to collected RF raw data
Frequency Range	From 24MHz to 1760MHz
Frequency Band Selection	Selective Band: small step for active bands, large step for inactive bands Full band: even step 1.2 MHZ
Sample Rate	2.4MHz
Period	Continually collecting for a few hours each time
Duration	2 milliseconds per frequency band

B. *RF signal pre-processing.*

The RF raw data collected at the 197 selective bands is fed to neural network directly with required format and no further frequency band data extraction is needed. Data preprocessing is then applied on full band RF raw data to extract band data of interest. These extraction bands are: active bands including and excluding cell network bands, inactive frequency bands including and excluding cell network bands, and random frequency bands. The number of each frequency band is listed in Table 2 and the extraction method is described as below.

TABLE 2. FREQUENCY BAND SELECTION

Frequency Band Group	# of Band
Selective Band	197
Active Band	76
Active Band Excluding Cell Network Band	53
Inactive Band	137
Inactive Band Excluding Cell Network Band	94
Random Band	128

In order to determine what bands are active and inactive, a continuous 48 hours full band RF raw data is collected at home study room and this data is used to calculate average power in the spectrum. Sample rate is 2.4MHz and one collection duration per frequency band is 2 milliseconds. So, the number of samples per frequency band, noted by N, is 4800. p_f is the notation of average power per frequency band per collection duration and it is calculated as below. s_i is raw data sample value received by SDR.

$$p_f = 10 * \frac{log_{10}\left(\sum_{i=1}^{N}\left(\frac{s_i}{127.5}-1\right)^2\right)}{\frac{N}{2}} \quad (1)$$

Let M be the number of collections per frequency band within these 48 hours. p_a, which is the average power value per frequency over M collection, is calculated by $p_a = (\sum_{i=1}^{M} p_{\mathrm{fi}})/M$. The average frequency band power in the spectrum ranges from 24MHz to 1760MHz, within these 48 hours as shown in Figure 2.

Frequency bands with peak average power in the spectrum are selected as active bands. Frequency bands with valley average power in the spectrum are selected as inactive bands. AMPD algorithm [17] is then used to automatically detect the peaks and valleys in the spectrum. Active and inactive bands are selected according to the detection results. Cell network bands are then excluded from the active bands and inactive bands to form active bands excluding cell network bands and inactive bands excluding cell network bands. Random bands consist of 128 randomly selected bands from full band.

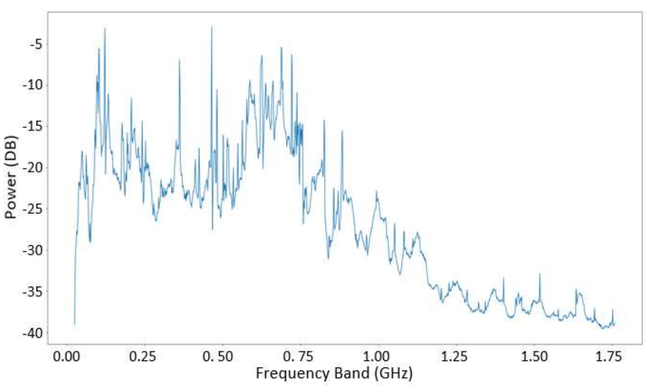

Figure 2. Average frequency band power in the spectrum

C. Experimental scenarios design

A total number of 12 experimental scenarios are designed and listed in Table 3.

TABLE 3. EXPERIMENTAL SCENARIO DESIGN

Name	Bands	Location	Time
ActH	Active Band	Home	-
ActHNCell	Active Band Excluding Cell Network Band	Home	-
InH	Inactive Band	Home	
InHNCell	Inactive Band Excluding Cell Network Band	Home	-
RndH	Random Band	Home	-
RndO	Random Band	Office	-
SelHO	Selective Band	Home & Office	-
SelH	Selective Band	Home	-
SelO	Selective Band	Office	-
ActHT1	Active Band	Home	6AM to 12PM
ActHT2	Active Band	Home	12PM to 6PM
ActHT3	Active Band	Home	6PM to 12AM

These scenarios cover human presence detection, accuracy and sensitivity tests against band selection, location diversity, and time difference. The scenarios are then categorized into 3 groups as listed in Table 4, band, location and time. These band sensitivity tests consist of 6 scenarios listed under the Band category. ActH is designed to train and test the CNN model with 76 active frequency bands RF raw data collected at home. Scenario ActHNCell is designed to train and tests the CNN model with 53 active frequency band data excluding cell network band data collected at home. Scenario InH is designed to train and test CNN model with 137 inactive frequency bands RF raw data collected at home. Scenario InHNCell is designed to train and test CNN model with 94 inactive frequency bands data excluding cell network bands data collected at home. Scenario RndH uses randomly selected 128 band RF raw data collected at home to train and test CNN model. Scenario RndO uses the same 128 frequency band to extract RF raw data collected at office. Location sensitivity test consists 3 scenarios listed under Location category. The 197 selected bands RF raw data collected at home and office are used to train and test CNN model. SelHO consists raw data of home and office, SelH only uses data of home and SelO only uses data of office. Time sensitivity test consists 3 scenarios listed under Time category. 76 active band RF raw data collected at home is used to train CNN model. ActHT1 uses RF raw collected from 6am to 12pm to test CNN model, ActHT2 uses data from 12pm to 6pm for testing and ActHT3 uses data from 6pm to 12am for testing.

TABLE 4. NUMBER OF BANDS USED IN DIFFERENT SCENARIOS

Category	Experimental Scenarios	# of Band
Band	ActH	76
	ActHNCell	53
	InH	137
	InHNCell	94
	RndH	128
	RndO	128
Location	SelHO	197
	SelH	197
	SelO	197
Time	ActHT1	76
	ActHT2	76
	ActHT3	76

D. *Training Data*

RF raw data is split into training dataset, validation dataset, and test dataset. The number of training, validation and test samples of each scenario is listed below in Table 5.

TABLE 5. DATASET

Experimental Scenarios	# of Training Samples	# of Validation Samples	# of Test Samples
ActH	2400	600	170
ActHNCell	2400	600	170
InH	2400	600	170
InHNCell	2400	600	170
RndH	2400	600	170
RndO	1200	300	92
SelHO	12480	3120	820
SelH	4560	1140	300
SelO	7920	1980	520
ActHT1	2512	327	86
ActHT2	2512	327	86
ActHT3	2512	327	86

E. *CNN Architecture and training*

The CNN consists of one 2D input layer, four 2D convolutional layers, one flatten layer, one fully connected layer and one output layer. The same CNN structure is used across all experimental scenarios except for the input layer row number. The input matrix consists K rows, which corresponds to frequency band number listed on Table 2, and 4800 columns, which is the sample number per frequency per one collection duration. The value of input matrix is RF raw data collected by SDR.

1D vector kernel is used to extract features from the frequency band raw data. The same 1D kernel shape $[1 \quad 4 \quad 8 \quad 8]$ is then used across these four convolutional layers along with the same stride step $[1 \quad 1 \quad 1 \quad 1]$. ReLU activation function $f(x) = \max(0, x)$ is used across all these four convolutional layer and fully connected layer. After the convolutional layers is the flatten layer. Connected to the flatten layer is the fully connected layer. The output layer has two perceptron which represents the human presence status. The values of the two binary numbers, indicate if human presence is detected or not. Other CNN architectures have been designed, trained and tested as well. But they did not achieve better performance than the one described above.

The CNN model is trained and evaluated for each experimental scenario listed in Table 3. The trained CNN model is used to process RF raw test data and detects the human presence in the closed space.

IV. EXPERIMENTAL RESULT

The expected overall experiment result of the initial phase is that CNN can distinguish human presence in a closed space by collected passive RF signals. In order to determine if this is the case, an F1 Score needs to be calculated in order to quantify the overall accuracy of the neural network, measuring the precision and recall of the results. The actual performance is evaluated by a confusion matrix with the equations below.

$$accuracy = \frac{TP+TN}{TP+FN+TN+FP} \tag{2}$$

$$precision = \frac{TP}{TP+FP} \tag{3}$$

$$recall = \frac{TP}{TP+FN} \tag{4}$$

$$F1 = \frac{2 \times precision \times recall}{precision+recall} \tag{5}$$

The overall experimental accuracy is shown in Figure 3. Both accuracy and F1 score from 10 experiments out of 12 is more than 90%. The accuracy and F1 score corresponding to the scenarios of ActH, ActHNCell, SelO and ActHT1 are higher than 95%.

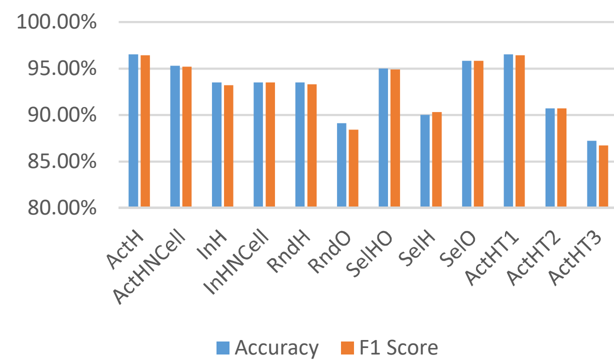

Figure 3. Overall Accuracy and F1 Score

The band sensitivity test results are shown in Figure 4. The experiments compare scenarios without cell network band data vs with cell network band data. Both scenarios achieve relatively close performances. For example, both accuracy and F1 score differences between ActH and ActHNCell is 1.2%. However further research is required to determine why the inactive band scenarios InH and InHNCell achieve similar performance as the active band scenarios ActH and ActHNCell.

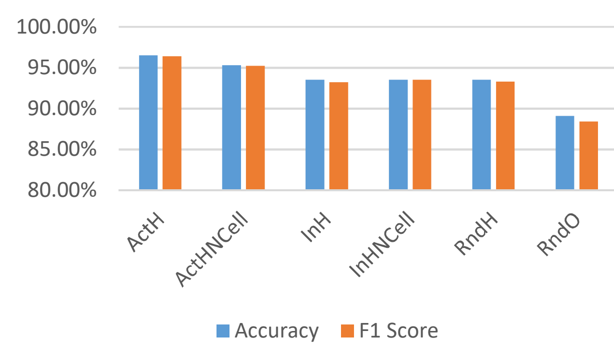

Figure 4. Band Sensitivity

The location sensitivity test result is shown in Figure 5. It can be seen the performance of SelH is slightly lower than the other two scenarios. The performance difference among locational test scenarios is less than 6%, which means the system is not very sensitive to location difference.

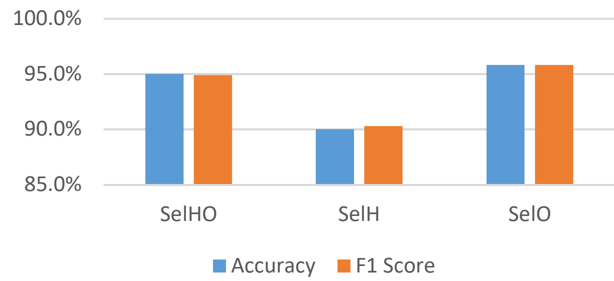

Figure 5. Location Sensitivity.

The time sensitivity test result is shown in Figure 6. The performance is the best in the 6am to 12pm time period and the worst in the 6pm to 12am time period. The cause of the difference is not clear at the moment. It might be due to the small test sample size or the variation of noise level with time. Further investigation is needed to improve the robustness over time.

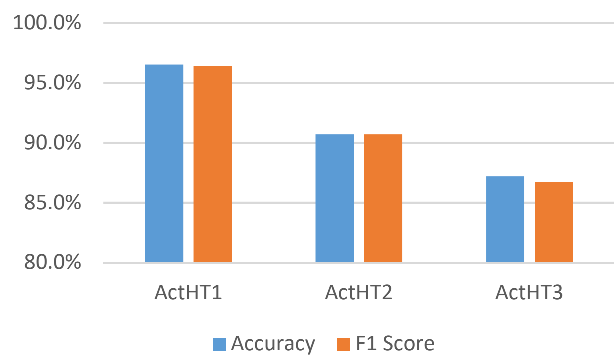

Figure 6. Time Sensitivity.

V. Conclusion

The results of this experiment indicate that human presence can be detected by passive RF wireless signals via deep learning neural network in a closed space. Robustness is verified by testing against different frequency bands, locations and time periods. Further studies such as finding out what the human signatures are in the RF data, experiment with different RF data features such as spectrum, doppler, amplitude and phase, and data collection on higher frequency bands, will be experimented with in the near future. Research into the detection of human presence in an open space, target differentiation, and speed estimation will proceed in the following phases.

Acknowledgment

This research is supported by AFOSR grant FA9550-18-1-0287.

References

[1] M. Kulin, T. Kazaz, I. Moerman, E. De Poorter, "End-to-end Learning from Spectrum Data: A Deep Learning approach for Wireless Signal Identification in Spectrum Monitoring applications," *IEEE Access*, vol. 6, pp. 18484–18501, 2018

[2] T. J. O'Shea, T. Roy, T. C. Clancy, "Over the Air Deep Learning Based Radio Signal Classification," IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 168-179, 2018

[3] S. Riyaz, K. Sankhe, S. Ioannidis, K. Chowdhury, "Deep Learning Convolutional Neural Networks for Radio Identification," IEEE Communications Magazine, vol. 56, no. 9, pp. 146–152, 2018

[4] K. Youssef, L. Bouchard, K. Haigh, J. Silovsky, B. Thapa, C. V. Valk, "Machine Learning Approach to RF Transmitter Identification", IEEE Journal of Radio Frequency Identification, vol. 2, no. 4, pp. 197–201, 2018

[5] Debashri Roy, Tathagata Mukherjee, Mainak Chatterjee, Eduardo Pasiliao, "Detection of Rogue RF Transmitters using Generative Adversarial Nets," Proc. of 2019 IEEE Wireless Communications and Networking Conference, in press, 2019

[6] S. Lv, Y. Lu, M. Dong, X. Wang, Y. Dou, W. Zhuang, "Qualitative Action Recognition by Wireless RadioSignals in Human-Machine Systems," IEEE Transactions on Human-Machine Systems , vol. 47, no. 6, pp. 789–800, 2017

[7] H. Li, C. Wan, R. C. Shah, A. P. Sample, S. N. Patel, "IDAct: Towards Unobtrusive Recognition of User Presence and Daily Activities," Proc. of 2019 IEEE International Conference on RFID, pp. 1–8, 2019

[8] M. Zhao et al., "Through-Wall Human Pose Estimation Using Radio Signals," Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7356–7365, 2018

[9] Yonglong Tian, Guang-He, Hao He, Chen-yu Hsu, Dina Katabi, "RF-Based Fall Monitoring Using Convolutional Neural Networks," Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 2, no. 3, Article 137, 2018

[10] M. Edrich, A. Schroeder, "Multiband Multistatic Passive Radar System for Airspace Surveillance: A Step towards Mature PCL Implementations", Proc. of International Conference on Radar, pp. 218–223, 2013

[11] Shuangquan Wang, Gang Zhou, "A review on radio based activity recognition," Digital Communications and Networks, vol. 1, no. 1, pp. 20-29, 2015

[12] Michael A. Nielsen, "Neural Networks and Deep Learning," Determination Press, 2015

[13] RLT-SDR.com. (2017) About RLT-SDR. [Online]. Available: http://www.rtl-sdr.com/about-rtl-sdr/

[14] Prasetiyo, R. V. W. Putra, T. Adiono, A. H. Salman, "Kurtosis and energy based spectrum detection for SDR based RF monitoring system," Proc. of International Symposium on Intelligent Signal Processing and Communication Systems, pp. 1–5, 2016

[15] D. Ball, N. Naik, P. Jenkins, "Spectrum Alerting System Based on Software Defined Radio and Raspberry Pi," Proc. of Sensor Signal Processing for Defence Conference, pp. 1–5, 2017

[16] "FCC Online Table of Frequency Allocations, " https://transition.fcc.gov/oet/spectrum/table/fcctable.pdf, Revised on May 7, 2019

[17] Felix Scholkmann, Jens Boss, Martin Wolf, "An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals," Algorithms, vol. 5, no. 4, pp. 588–603, 2012