Extraction and Denoising of Human Signature on Radio Frequency Spectrums

Shiyu Yang
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI
syyang@oakland.edu

Liangqi Yuan
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI
liangqiyuan@oakland.edu

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

***Abstract*—This paper proposes an innovative machine-learning-based method to extract compact, accurate, and adequate human radio frequency signature in residential environment. Our research created a shielded environment by using electromagnetic fields blocking material to attenuate strong signals in the background. SHapley Additive exPlanations were utilized to identify the most human impacted frequency range, which ensures the spectrums acquired later contain adequate information. In order to extract the spectrum that contains mainly human signature information, a conditional-generative-adversarial network has been trained to model the shielding effect. The proposed method can generate the spectrum containing the human signature that was originally buried in the background. In addition to simulating the denoising effect that is established by the physical shielding, the trained generator model is applied for the second time to achieve multi-stage denoising, which further improves the signal to noise ratio in the spectrum. As the result shown, the proposed model successfully generates the spectrum with root mean square distance of 0.027 when comparing with the physically shielded spectrum. Furthermore, a Support Vector Machine model is trained to evaluate the performance of the conditional generative adversarial network model. The experimental results show that the extracted human signature in the synthesized spectrum can be identified by the support vector machine classifier with 100% accuracy while the physically shielded spectrum yields 93.5% accuracy.**

Keywords—Human signature, passive radio frequency (PRF) spectrum, software defined radio (SDR), conditional generative adversarial network (CGAN), adaptive spectrum sensing

I. INTRODUCTION

Human signatures have individualized patterns and are automatic identifications by quantifying the biological characteristics exist in the human signature [1]. Pioneering studies have sensed and captured human signatures in many forms to serve the purpose of human detection and identification. In light of the rapid advances in machine learning (ML) techniques, different forms of human signature have been used in various applications, such as human detection, positioning, identification, etc. Many popular active sensing modalities for human signature collection include camera [2], LiDAR [3], radar [4][5], infrared [6], WiFi-based system with cognitive radio [7], and even device-free in WiFi sensing [8], but all with their own constraints and limitations [9]. Instead of using any of the active sensing methods aforementioned, our research focused on passive radio frequency (PRF) sensing method via two cognitive radio devices. Due to the advantages of low cost, energy saving, and harmless to human health, especially the ability of allowing customized sensing frequency, the PRF sensing method is ideal for collecting human radio frequency signatures. Our goal of this paper is not only to capture the PRF spectrum containing human signature, but also extract and denoise the human signature from the RF spectrum in order to better serve the purpose of human identification.

Since the human signature buried in the RF spectrum tends to be weak comparing to other existing RF signals within a normal environment, we created a shielded environment to block the interference from external electromagnetic fields (EMF). The RF spectrums acquired in the shielded environment are considered to be the spectrums that consist of mainly human signature, and used as the target of generating the synthetic human signature.

In this paper, we propose an innovative method of extracting and denoising human signature from the RF spectrum captured in a residential environment via software defined radio (SDR) devices. Prior to the training datasets acquisition, explainable artificial intelligence (XAI) technique is applied to identify the most human impacted RF frequency range to ensure the captured spectrums carry human signature information. To extract the human signature, a conditional generative adversarial network (CGAN) is adapted to simulate the shielding effect and then synthesize RF spectrum that contains mainly the human signature. Comparing with the target spectrums, our experimental result shows that the synthesized spectrums yield a root mean square distance of 0.027. At the end, a support vector machine (SVM) classifier is trained using the physically shielded spectrum data. The synthesized RF spectrums of five categories are sent into the trained classifier for human subject identification to evaluate the quality of the synthetic spectrum. The result shows that the synthetic spectrum obtains 1.0 as the F1-score, which indicates the generated spectrum does contain effective human signature and can be identified by the classifier with 100% accuracy.

This paper consists of five sections. In Section II, the spectrum measurement setup is explained, as well as the machine learning methods used in this paper which are the Shapley Additive Explanations (SHAP), CGAN and SVM classifier. Section III provides information regarding the hardware and experiment setup. Section IV presents the experiment results, the discussion of data frequency range selection result, CGAN training results, and SVM classification scores on the CGAN generated dataset and the physically shielded dataset. Conclusion of this research is given in Section V.

2023 IEEE International Conference on Consumer Electronics (ICCE) | 978-1-6654-9130-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCE56470.2023.10043377

978-1-6654-9130-3/23/$31.00 ©2023 IEEE

II. METHODOLOGY

A. Spectrum Measurement

In this research, the data acquisition process solely relies on software defined radio (SDR) devices for spectrum sensing. Allowing user to configure the sensing frequency range, step size, sampling rate, and the number of samples in In-phase and Quadrature (IQ) channelsare some of its major advantages, making the SDR device a convenient tool for our research. Specifically, the RTL2832U model has been selected as the sensing hardware throughout our experiment.

Human signature in RF spectrum is normally buried in the noisy background RF signals and difficult to be extracted. In our experiment, we setup a shielded environment by using EMF blocking materials to minimize the strong RF signal contaminations. The established physical shielding environment allows the SDR devices to capture human signatures in a greater signal to noise ratio (SNR) in terms of average power at each frequency that is impacted by the presence of human subjects. The spectrum recorded inside of the shield, named as shielded spectrum, contains the extracted human signature. Thus, it is used as target for training the CGAN. Our approach is to learn and recreate the physical shielding effect through a neural network so that human signatures can be extracted and their SNR boosted.

The experiment datasets acquisition process consists of three phases. Phase 1 and phase2 are constructing D_{SHAP-} and D_{SHAP-2} used in the two rounds of SHAP analysis, and phase 3 dataset is constructed for training the CGAN model and SVM classifier. The first phase dataset D_{SHAP-} is recorded at a wide frequency band of 2.4-1000MHz in order to fully cover the most human impacting frequencies. Then the SHAP result from the first phase weighed in and narrowed down the frequency band of the second phase to 200-800MHz. Based on the two phases of SHAP result, phase 3 dataset is obtained on the 500-600MHz frequency band. Two datasets were constructed during phase 3, which are the shielded spectrum dataset D_W captured in the EMF shielded area, and the without shield dataset D_{WO} captured at the same location but while the shield is removed. There are five categories of data taken for each dataset involved in this research, which include four human subjects and an unoccupied scenario. The detailed dataset structure and parameter settings are summarized in Table I.

TABLE I. DATASET STRUCTURE AND SDR PARAMETERS

	Phase 1	**Phase 2**	**Phase 3**
Frequency Range	24-1000MHz	200-800MHz	500-600MHz
Sampling Step Size	2.4MHz	1.2MHz	2.4MHz
Sampling Rate	2.4MHz	2.4MHz	2.4MHz
Background Condition	With Shield Only	With Shield Only	With & Without Shield
Total # of Spectrum Samples	320	400	2000
Subjects	Human Subject 1 - 4, & Unoccupied		
Data Prepared For	SHAP	SHAP	CGAN & SVM
Dataset Notation	D_{SHAP-1}	D_{SHAP-2}	D_w & D_{wo}

B. Data Pre-proccessing

For every RF spectrum collected in our experiment, the sample number per frequency denoted as N, is set to be 4096. Each datapoint on the spectrum is returned as a sampled complex IQ signal ($I + jQ$) represented by two voltage samples. Thus, there are $N/2$ pairs of IQ data collected at each selected center frequency. The average power P in dB at each center frequency f, which is also the amplitude of RF spectrum collected in our experiments, can be calculated as:

$$P(f) = 10\log_{10}\frac{\sum_{i=1}^{N} p_f(i)}{N/2}, \quad (1)$$

where the average spectrum power P is a function of f. p_f is the power at each center frequency f.

Fig. 1 shows one of the pre-processed spectrum samples.

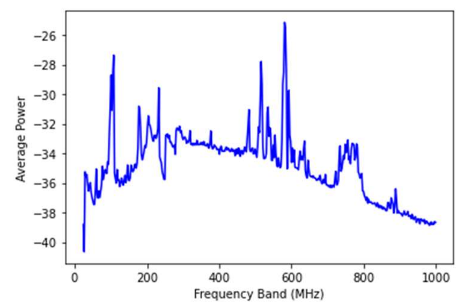

Fig. 1. Pre-processed full frequency band spectrum plot.

C. SHAP on Frequency Band Selection

Our group previous research result indicates that the RF spectrum on the full frequency range 24-1760MHz contains adequate information to achieve human occupancy detection [9]. However, due to the data recording and CGAN processing time being too long with the full range spectrum, we aim on developing a process to identify the most human impacting frequency band to improve the CGAN training efficiency.

SHAP, as a common solution of finding the most impact factors in a machine learning process, is used in this research to find how the impact of the human signature varies at each frequency, and more importantly, to identify the most human impacting frequencies. There are two phases of dataset prepared for the SHAP as aforementioned. They all contain five categories of data, including human subject 1-4 and an unoccupied category. The SHAP results mainly explain the impact factors in classifying these five categories of data.

Prior to the two major rounds of SHAP analysis, full frequency range samples were collected to identify the preliminary frequency band range, which is used to set the frequency band of D_{SHAP-} . Phase 1 SHAP analysis is then conducted with D_{SHAP-1} as input. The second round of data acquisition and SHAP is performed with narrower frequency range while using the first round SHAP result as the reference, in order to verify the result and further narrow down the frequency range of phase 3 dataset.

D. CGAN Model Structure

CGAN is the main method used in this research for creating a digital shielding effect, in order to achieve the human signature extraction and denoising.

The CGAN model structure consists of one discriminator and one generator. The goal of using the CGAN is to simulate the shielding effect and then generate the spectrum contains extracted human signature from the environment without any physical shielding setup. In our experiment, D_W is used as the target to train the discriminator, and D_{WO} is fed into the generator as the input. The condition used in training this CGAN model is one-dimensional label generated based on the experiment subject category and does not contain any feature information from the subject. Once the CGAN is fully trained, the generator model is saved and used separately to generate the spectrum contains extracted and denoised human signature with the testing data in D_{WO} as the input. The detailed structure of discriminator and generator of the proposed CGAN model are shown in Fig. 2.

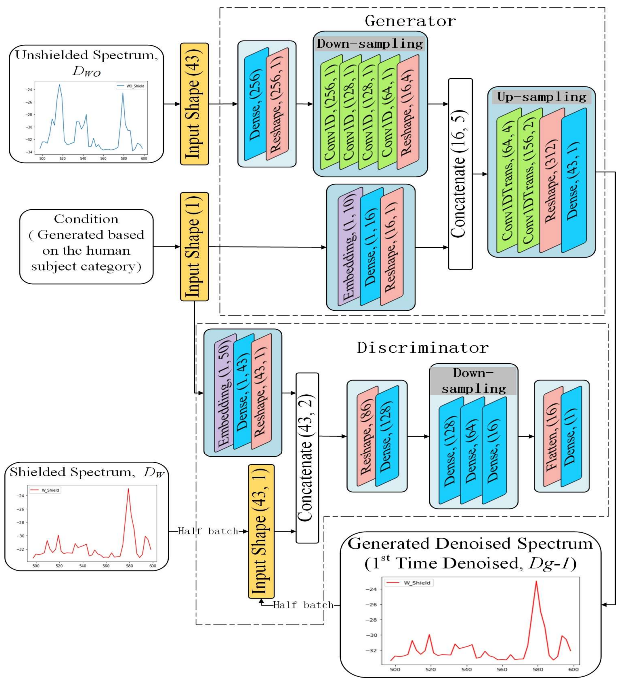

Fig. 2. CGAN model structure.

One thing worth noting is that all the Dense, Convolutional, and Convolutional Transpose layers are followed by a Leaky ReLU layer immediately. At the data processing level, the generator takes the training data in D_{WO} as input, and then send it to a down-sampling process to achieve the feature extraction. Fig. 2 shows the down-sampling consists of four Convolutional 1D layers and a Reshape layer at the end. The conditioning data is then concatenated with the output from the down-sampling and sent into an up-sampling process, which consists of two Convolutional 1D Transpose layers. A Reshape layer and one Dense layer are applied at the end to make the output in the same shape as the samples in D_W, which is (43,1). The purpose of adding this up-sampling process is to handle the data imbalance. Based on the features extracted from the down-sampling, during the up-sampling process, the neural network will alter the data points and give the useful features a more equal weight. This up-sampling process is critical in generator training since it ensures all the important features are weighed correctly, which will lead to a more accurate output with all the useful features kept.

In the discriminator, the input data has the same shape as the generator output. The condition data is embedded and then reshaped into the same shape as the input data, thus it can be concatenated with the input data. The discriminator contains down-sampling process only since it mainly serves as a binary classifier in the CGAN. After going through three Dense layers, the discriminator model is trained to achieve the feature extraction in order to classify the synthesized data and real data. One thing worth noting is that our discriminator up-sampling process uses Dense layers instead of the Convolutional layers due to our spectrum data type. As the main difference between the Convolution layer and the Dense layer being that Convolutional layer tends to force the input share the parameters when learning the relationship of the input and output, it uses fewer parameters by doing so. In contrast to the Convolutional layer, the Dense layer forms every output and corresponding input into a function since it uses linear operation. In our case, as aforementioned the input sample has only 43 data points, meaning there is limited quantity of information and parameters exist in the input data. Also, the dataset sample quantity is relatively small comparing to other CGAN applications since this experiment is not dealing with complicated data type. Considering all these factors, we decided to use Dense layers in the discriminator instead of Convolutional layers since the discriminator does not require heavy loaded learning ability as the generator requires.

E. Evaluation System of the CGAN Model

To evaluate the quality of the generated denoised spectrum, a SVM classifier is trained on the dataset D_W and then used to identify the four human subjects and unoccupied category with the generated spectrums. In Fig. 3, the details of utilizing the SVM classifier is further demonstrated.

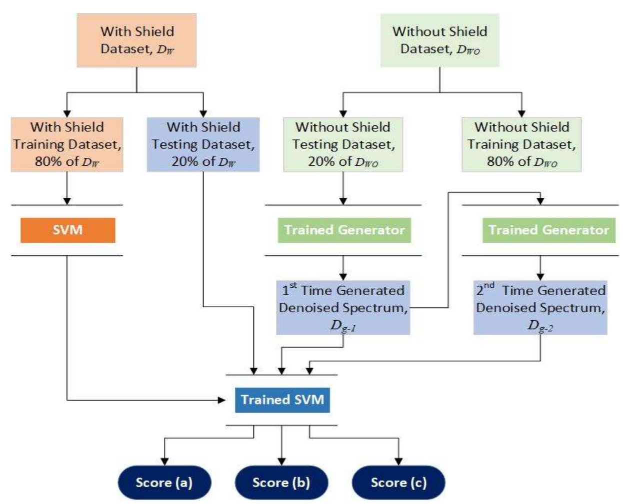

Fig. 3. Evaluation system flow chart.

As Fig. 3 shows, the SVM model is trained on D_W, and later on, the trained SVM model is used to classify three sets of data:

a) The testing samples from D_W,

b) The generated denoised spectrum D_{g-1},

c) The generated additionally denoised spectrum D_{g-2},

where D_{g-1} is the first time generated denoised spectrum with D_W being the input, and D_{g-2} is the second time generated denoised human spectrum with D_{g-1} as the input.

In order to generate the D_{g-2}, the generated D_{g-1} is fed back into the trained generator and the output is considered as second time denoised. This process simulates a better shielding effect than the physically shielded environment.

The SVM classification results are expected to have same F1-scores on all three sets of data to prove the success of CGAN training. Higher F1-score on the synthesized datasets indicates that the trained generator has the ability of extracting real human signature from the measured spectrum in a residential environment without physical shield.

III. Hardware and Experiment Setup

In the proposed research, the hardware includes two SDR devices for receiving RF data, one set of RF signal shield, and two pieces of EMF shielding fabric. The detailed product information is further explained in this section.

A. SDR Devices

Fig. 4 shows the SDR device used in the experiments. This device is named RTL-SDR as it is manufactured by Real Tech, with one of their own RTL2832U chipset to support the software defined radio feature. The receiving RF band of this device ranges from 500kHz up to 1.7GHz and could be adjusted by software.

Fig. 4. RTL-SDR device used in the experiment.

B. RF Signal Shield

Fig. 5 shows the RF signal shield used throughout our experiment. The shield is made by coated silver fiber, with a size of 120×220×220cm^3. The main purpose of using this shield is to reduce RF signal interference from the environment at a residential location, in order to create an experiment space to obtain the target data used in the training of CGAN model. The human subjects and SDR receivers were set inside of the shield through half of the data acquisition process, while the other half were taken without the shield and used as input data to the CGAN model.

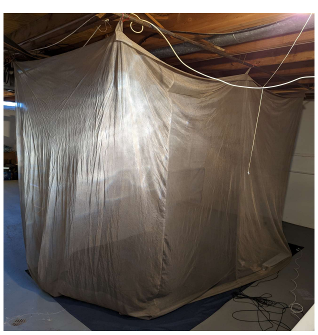

Fig. 5. RF signal shield setup.

C. Experiment Setup

The experiment was designed to be setup at a residential location to avoid unnecessary EMF interference, which commonly exist in a laboratory environment due to the usage of electronic devices. Two SDR devices are placed inside of the shield and were maintained with a fixed distance to the shielding boundary and the same pointing direction throughout the experiment as shown in Fig. 6.

Fig. 6. The setup for human subject data acquisition inside the shield. Two SDR devices were fixed on either side of the human subject and pointing to the human subject.

IV. Experiment and Results

A. Human Impacted Frequency Band

Initially, the spectrum data were taken on full frequency range supported by the SDR, which is 2.4-1760MHz, with 1.2MHz as the step size and sensing frequency set to 2.4MHz. To improve the efficiency of data acquisition, we adopted SHAP to identify the preliminary frequency range. The SHAP result of this preliminary process shows the top twenty of most human impacting frequencies fall within 100-850MHz, which lead us to set the phase 1 dataset frequency band to 24-1000MHz, just to add in some buffer.

The phase 1 SHAP result is shown in Fig. 7. The vertical axis shows the ranking of the top twelve most human impacting frequencies, while the horizontal axis indicates the Shapley value of the average power at the corresponding frequency. The color of each point represents how the impact value changes as the average power changes. Per the legend at the right side, the datapoints being marked in red meaning the impact value increases while the sensed spectrum average power increases at the indicated frequency on the left side, and vice versa. The Shapley values are minimal after twelve frequencies, therefore, only the top twelve impacting frequencies are listed and ranked. This result indicates that the most human impacting frequencies fall within the range of 513.2-753.2MHz.

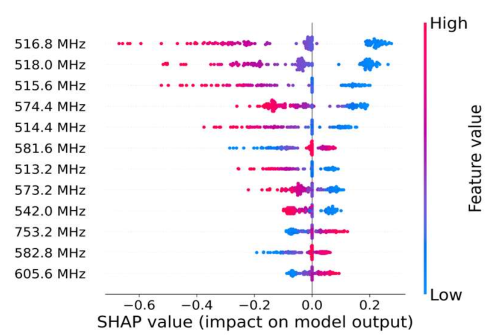

Fig. 7. Phase 1 SHAP analysis result.

SHAP analysis was performed again to verify the frequency band selection by using the phase 2 dataset which was acquired at a different time. This phase 2 dataset frequency range has been narrowed down to 200-800MHz, based on our phase 1 SHAP result. As shown in Fig. 8, all the human impacting frequencies are within 513.6-597.6MHz in the phase 2 analysis. The phase 3 dataset frequency range is then set to 500-600MHz based on the phase 2 SHAP result. The CGAN training later on proves that this frequency range does contain adequate human signature and provides better efficiency in the training process.

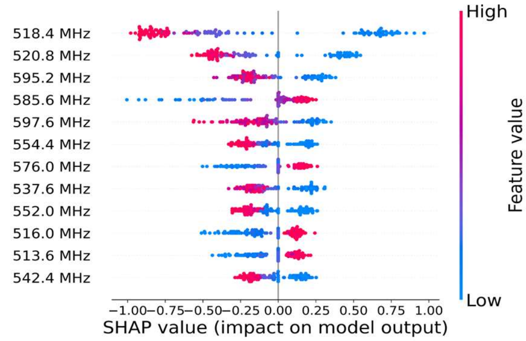

Fig. 8. Phase 2 SHAP result.

B. CGAN Training and Synthesized Spectrum

As aforementioned, the unshielded spectrums in D_{WO} contains human signature, but it is also mixed with strong background RF signals. The trained CGAN is able to simulate the shielding effect in order to generate the shielded spectrum from unshielded spectrum. This is a rough human signature extraction process.

Once the datasets are constructed based on the target human impacting spectrum frequency range, the CGAN model is trained on dataset D_W and D_{WO}. A total of 800 samples in D_W were fed into the discriminator as input and 800 samples in D_{WO} were fed into the generator. The samples in both D_W and D_{WO} datasets are labeled in five categories, which are used as condition data when training the CGAN model. The CGAN is trained on 270 epochs with a batch number of 22.

Fig. 9 shows the loss curve of the training process. The blue and orange lines represent the loss of the discriminator on dataset D_W and D_{g-1}, and the green line represents the loss of the generator, respectively. It can be seen that the discriminator and generator loss were strongly oscillating within 60 epochs of training, which indicates the CGAN model is setup correctly with the generator and discriminator fighting each other at the beginning. The model slowly converges at 60-180 epochs range.

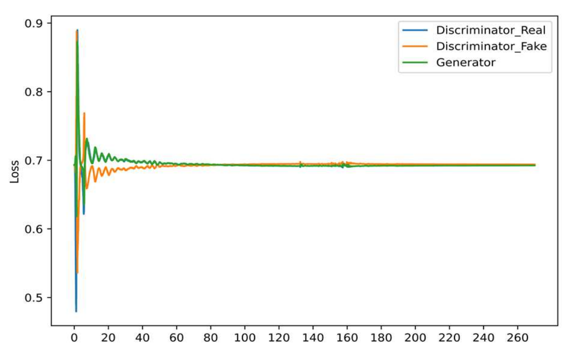

Fig. 9. The CGAN training loss curve.

Fig. 10 shows the result of the generated denoised spectrum along with the sample from D_W which represents the denoised target spectrum, and the sample from D_{WO}, which is the generator input. These generated first time denoised spectrum construct the dataset D_{g-1}. As it can be easily visualized, the generated spectrum is quite close to the target shielded spectrum in terms of shape and value, but only with small deviations at a few human impacting frequencies. As discussed later on in Section IV.C, the evaluation result implies that the deviations are improvements on the SNR of human signature as D_{g-1} yields better classification result.

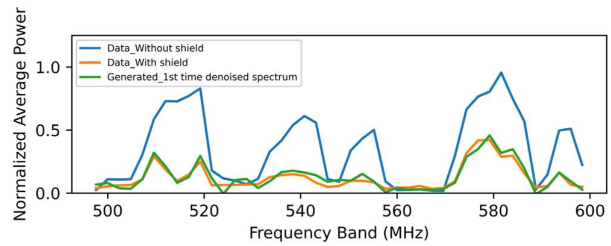

Fig. 10. Comparison of sample spectrums from D_w, D_{wo} and the synthesized when CGAN is applied the 1st time.

The root mean square distance (RMSD) is calculated to measure the difference between the synthesized 1st time denoised spectrum and the target shielded spectrum. With normalized samples in D_{g-1} and D_W being the input, the RMSD calculation result shows that the average distance between data points on the same frequency from these two datasets is 0.027.

The constructed D_{g-1} is fed back into the generator to generate the second time denoised spectrums and construct the dataset D_{g-2}, and this process simulates a better shielding effect than the physically shielded environment. The red line in Fig. 11 shows the second time denoised spectrum. Fig. 11 summarizes the CGAN generation results, including spectrum samples from D_{WO}, D_W, D_{g-1} and D_{g-2}. The purpose of computing this graph is to visualize the improvement of the first time and second time denoised spectrum performance.

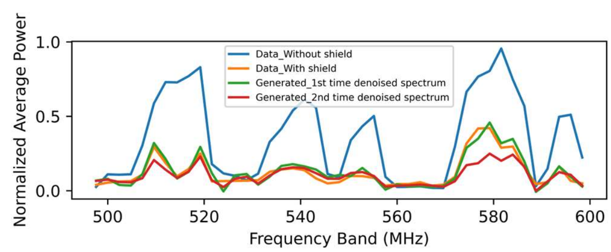

Fig. 11. Overview of the generation results. It shows the comparison of sample spectrums from D_w, D_{wo} and the synthesized when CGAN is applied the 1st time and the 2nd time.

Based on the result shown in Fig. 11, it is obvious that the generated spectrum average power has decreased when CGAN is applied the second time, and the decreased amount varies at different frequencies. Our initial goal was to train the generator to detect which portion of the spectrum contains the critical human signature and how much noise should be filtered out from the spectrum input based on the pattern that the generator has learnt from the shield. This result satisfies our initial expectation for the generator. After the CGAN is applied the second time, the SNR of the spectrum is further denoised. Multiple stages of denoising is commonly used in signal processing pipeline. To the best of our knowledge, it is the first time that CGAN has been trained and used to achieve the two times denoising function.

C. Evaluation of Synthesized Spectrum

The SVM classifier is applied to evaluate the quality of the synthesized spectrum D_{g-1} and D_{g-2}. The SVM is trained on the training dataset in D_W to classify four human subjects and unoccupied category. Fig. 12 shows the trained SVM classification result of the testing dataset in D_W, which has a F1-score of 0.935.

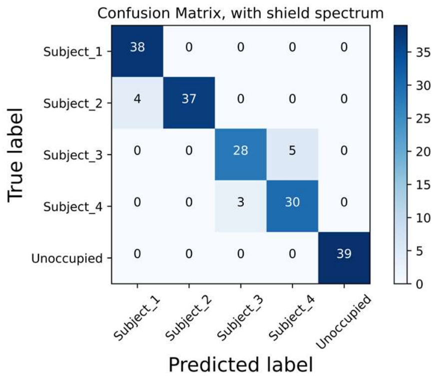

Fig. 12. SVM classifier confusion matrix on D_W testing dataset.

However, the two generated spectrum datasets D_{g-1} and D_{g-2} with the same trained SVM classifier obtained higher F1-score, 1.0 for both datasets, when classifying the five categories. This indicates that the synthesized spectrum outperformed the physical shielded spectrum in terms of subject classification accuracy.

The SVM classifier is intended and trained to evaluate the generator performance based on the human subject identification results. The initial expectation of the synthesized spectrum $D_{g-1}$and D_{g-2} is to maintain the same level of F1-score as dataset D_W. However, the result shows that the two synthesized spectrum datasets obtained higher F1-score from the classifier, which implies that the CGAN training is successful and the generator performance exceeds the expectation.

The SVM model classifies the spectrums of human subjects and unoccupied based on the pattern of their human signatures, hence the classification result is highly depending on the SNR of the human signature in the spectrum. The better quality and greater proportion of human signature exists in the spectrum would result in higher classification accuracy. In our research, the CGAN achieved the human signature extraction by simulating the physical shielding effect as expected. Furthermore, with the synthesized spectrum obtaining higher classification score, it implies that the SNR in the synthesized spectrum is higher than in the physically shielded spectrum. This may be caused by the additional denoising feature that is achieved during the generator training process due to the nature of the powerful machine learning technic and the way we setup the generator structure. In the real world, the data we captured are often heavily imbalanced, which always result in errors on classification. The up-sampling process in the generator structure has the ability of adjusting the data weights based on the features that are extracted from the down-sampling process, which could lead to the useful features weigh differently than in the target spectrum. The experiment result implies that this data points weight altering behavior has a positive impact to the spectrum generation, thus, we can come to a conclusion that our CGAN model grants the spectrum an improvement on the SNR. With the feature extraction and denoising feature, the synthesized spectrum in D_{g-1} and D_{g-2} both contain human signature with higher SNR than the shielded spectrum in D_W, which explains the higher F1-score gained on the SVM classification.

V. Conclusion

In this paper, a CGAN model is proposed to learn from physical shielding effect and create digital shielding effect to extract and denoise human signature on RF spectrum for human subject identification. In order to present the human signature in a compact and efficient way, the RF spectrum is analyzed by SHAP method to obtain the most human impacted frequency range. The CGAN can be applied multiple times to achieve different levels of denoising. The synthetic RF spectrums are evaluated by SVM classifier and proven to be accurately generated, effective to human identification, and containing adequate human signature information. This proposed method provides low energy consumption solutions to human identification related applications such as security monitoring in prison, patient and senior people caring based on human existence, smart devices control in different areas of building, climate control in factory and isolated chambers for energy consumption and comfort trade-off, etc.

Acknowledgement

This research is supported by Air Force Office of Scientific Research through Fund FA9550-21-1-0224.

References

[1] A.K. Jain, K. Nandakumar and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities. " *Pattern Recognit. Lett*, 2016, 79, 80–105.

[2] C. Zeng and H. Ma, "Human detection using multi-camera and 3D scene knowledge," *2011 18th IEEE International Conference on Image Processing*, 2011, pp. 1793-1796, doi: 10.1109/ICIP.2011.6115810.

[3] T. Taipalus and J. Ahtiainen, "Human detection and tracking with knee-high mobile 2D LIDAR," *2011 IEEE International Conference on Robotics and Biomimetics*, 2011, pp. 1672-1677, doi: 10.1109/ROBIO.2011.6181529.

[4] V. C. Chen, "Detection and analysis of human motion by radar," *2008 IEEE Radar Conference*, 2008, pp. 1-4, doi: 10.1109/RADAR.2008.4721059.

[5] S. Chang, N. Mitsumoto and J. W. Burdick, "An algorithm for UWB radar-based human detection," *2009 IEEE Radar Conference*, 2009, pp. 1-6, doi: 10.1109/RADAR.2009.4976999.

[6] K. Hashimoto, N. Yoshiike, and K. Morinaka, "Human occupancy detection method and system for implementing the same," *U.S. Patent* No. 5,703,367. 30, Dec. 1997.

[7] T. Zhang, T. Song, D. Chen, T. Zhang and J. Zhuang, "WiGrus: A Wifi-Based Gesture Recognition System Using Software-Defined Radio," in *IEEE Access*, vol. 7, pp. 131102-131113, 2019, doi: 10.1109/ACCESS.2019.2940386.

[8] R. Zhang, X. Jing, "Device-Free Human Identification Using Behavior Signatures in WiFi Sensing." *Sensors,* 21.17 (2021): 5921–. Web.

[9] J. Liu, H. Mu, A. Vakil, R. Ewing, X. Shen, E. Blasch, J. Li, "Human Occupancy Detection via Passive Cognitive Radio," *Sensors*, vol. 20, no. 15, 4248, Jul. 2020