Robust Deep Radio Frequency Spectrum Learning for Future Wireless Communications Systems

DAMILOLA ADESINA, (Graduate Student Member, IEEE),
JOSHUA BASSEY, (Student Member, IEEE), AND
LIJUN QIAN, (Senior Member, IEEE)
Center of Excellence in Research and Education for Big Military Data Intelligence (CREDIT Center), Prairie View A&M University, Texas A&M University System, Prairie View, TX 77446, USA

Corresponding author: Damilola Adesina (dadesina@student.pvamu.edu)

ABSTRACT Intelligent capabilities are of utmost importance in future wireless communication systems. For optimum resource utilization, wireless communication systems require knowledge of the prevalent situation in a frequency band through learning. To learn appropriately, it is imperative for practitioners to select the right parameters for building robust data-driven learning models as well as use the appropriate algorithms and performance evaluation methods. In this paper, we evaluate the performance of deep learning models against the performance of other machine learning methods for wireless communication systems. We explore the different wireless communication scenarios in which deep learning can be used given Radio Frequency (RF) data, and evaluate its performance in various scenarios. Furthermore, we express it as a distribution alignment problem in which deep learning models do not perform well when learning from RF data of a particular distribution and evaluating on RF data from a different distribution. We also discuss our results in the light of how signal quality affects deep learning model leveraging on the knowledge from computer vision domain. The effect of Signal-to-Noise Ratio (SNR) selection for training on the model performance as it relates to practical implementation of deep learning in communications systems is also discussed. From our analysis, we conclude that the design and use of RF spectrum learning must be tailored to each specific scenario being considered in practice.

INDEX TERMS Radio frequency learning, signal-to-noise ratio, training and testing strategy, spectrum data, convolutional neural network.

I. INTRODUCTION

There has been a lot of interest in the development of artificial intelligence for wireless communication systems using radio frequency (RF) data. This is because the ability to learn and intelligently respond to dynamic and complex operating conditions will be of utmost importance in future wireless systems. It is envisioned that the knowledge of current operating conditions and environment leveraging on wireless big data analytics [1]–[4] will allow communication systems to make the best opportunistic decisions. To achieve this, a lot of work has been done to develop efficient algorithms and methods to extract meaningful information from complex and massive datasets obtained from communication systems.

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

For example, machine learning and deep learning models are developed for:

1) *Spectrum situation awareness, spectrum sensing and spectrum occupancy prediction* [5]–[7] to identify users, monitor the spectrum as well as provide information about the communication system and radio environment. This will in no small measure be an important part of 5G systems for spectrum resource management and optimization.

2) *Device identification and Intrusion detection* [8]–[11] to identify the transmitting device in an Internet-of-Things (IoT) environment and also detect the presence of intruders in a crowded electromagnetic spectrum where secure transmission of friendly signals as well as jamming of unwanted signals are crucial.
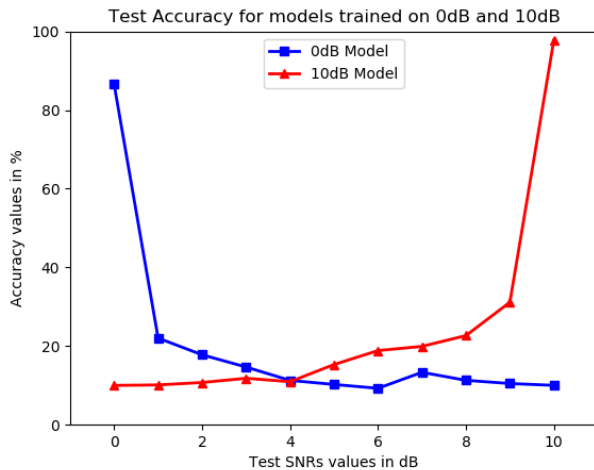
FIGURE 1. Test accuracy for models trained using RF data at 0dB and 10dB and tested using RF data at various SNR values.

3) *Modulation recognition* [12]–[14] to detect modulation scheme of detected signals before demodulation.
4) Other applications include *interference identification* [15] and *spectrum monitoring* [16].

Research in these areas has gained tremendous attention in recent time particularly the application of machine learning and deep learning to solve wireless communication problems using RF data. Most research work in this area have an assumption that the same training dataset that enables good accuracy will suffice to train robust models that will generalize on unseen data [17]. Given the stochastic nature of the wireless channel which directly affects the communication signal and the proven knowledge that perturbations can cause a decline in the ability of neural networks [18]. It becomes important to explore *how to develop robust deep learning models that generalize well on unseen data for different wireless communications scenarios in practice using RF data.*

Motivated by this, we investigate the potential of using deep learning for wireless communication systems by utilizing RF data. This approach is herein referred to as RF learning. We also look at deep learning application development for wireless communication systems taking into account the practical implementation requirements. Specifically, we evaluate the effect of signal-to-noise ratio in the development of deep learning models using radio frequency data. This practical study is motivated by the need to develop DL models that are robust and will generalize well given the stochastic nature of wireless communication systems. As a motivating example, a DL model trained on 10dB data for automatic modulation recognition may receive a signal of 8dB when deployed in a dynamic RF environment. Such models may not generalize well in this situation. Figure 1 shows the test accuracies for models developed using 0dB and 10dB dataset and tested on $0dB, 1dB, 2dB, \ldots, 10dB$. We observe high classification accuracy when train and test dataset is from same SNR level and very low accuracy otherwise. Testing the 10dB model with 10dB dataset gave 97.81% accuracy,

while testing with 8dB dataset gave 22.66% accuracy. In this work, we experimentally show the effect of train and test data selection for model robustness and generalization. We identified three scenarios where deep learning can be used for RF spectrum learning, investigated the deep learning training strategies for these scenarios and came up with performance evaluation strategies for deep learning models in practical RF learning, focusing on the effect of signal-to-noise ratio (SNR).

In our previous work [19], we formulated a 3-class classification problem for interference identification to study practical considerations for deep learning in wireless communication systems. In this work, we present a more comprehensive approach by looking at three unique scenarios for wireless communication problems namely: automatic modulation classification, type and number of users identification in a shared spectrum and spectrum monitoring. We also explain our observations and conclusions using the distribution alignment problem and observations in the more developed computer vision domain.

From our analysis, it can be deduced that to achieve robust and practicable RF learning in different scenarios, there must be unique problem formulation with special consideration for SNR selection and performance evaluation methods. Factors such as SNR step-size for training as well as testing dataset selection must be carefully studied in all cases for RF learning. Most of the work in literature focus on developing good models and identifying suitable models for unique problems. Our work differs from others because we consider what obtains in real life scenarios and the uniqueness in the data selection for development of deep learning models that are robust and generalize well even on data unseen during training.

The contributions of this work include:

1) Highlight of RF learning problem formulation for different wireless communications scenarios;
2) The use of one fixed SNR level to build a model for RF learning cannot be used in many practical scenarios where SNR values may vary. To build a practical RF learning model that captures the variations in SNRs, multiple fixed SNRs with appropriate step sizes must be used in training;
3) Evaluation and testing strategy for RF learning must be carefully designed. For instance, depending on SNR variations and need for generalizations, it is important to test on varying SNR levels.

The remainder of this paper is organized as follows. Section II provide details of the data generation process. Section III presents the background to deep RF learning detailing the concept of and the motivation behind using deep neural network on RF dataset and the choice of parameters for deep learning model development. Section IV explores the different deep learning training and testing strategies for RF learning. We discuss RF learning scenarios, the development of deep learning for three unique scenarios and

TABLE 1. Dataset parameters for each scenario.

Parameters	Scenario 1	Scenario 2	Scenario 3
Number of transmit USRP	1	2	1
Number of receive USRP	1	1	1
SNR level considered(dB)	0,1,2,3,4,5,6, ...12	0,1,2,3,4,5,6, ...12	0,0.5,1,1.5,2,2.5,3, ...12
Number of training IQ samples	80 Million	32 Million	90 Million
Train SNR (dB)	10	2, 4, 6, 8	model1: 0,0.5,1,2,2.5,3, model2: 0,1,2,4,5,6, model3:0,2,4, 8,10,12
Test SNR (dB)	10	2, 4, 6, 8	model1: 0,0.5,1,1.5,2,2.5,3 model2: 0,1,2,3,4,5,6 model3: 0,2,4,6,8,10,12
Comment	Train test on same SNR level used for training	Train test on same multiple fixed SNR level	Test singly on SNRs including SNR not seen in training

model evaluation in section V. Analysis of results, insights and observations are discussed in section VI while related research is discussed in Section VII. Section VIII concludes the paper.

II. DATA GENERATION

Over-the-air dataset generated using National Instruments Universal Software Radio Peripheral (USRP) based testbed for distributed spectrum monitoring and surveillance as well as LabVIEW is used in this work. The USRP is a tunable radio frequency transceiver and Software Defined Radio Device with high-speed analog-to-digital converter and digital-to-analog converter for streaming baseband IQ signals to a host PC over 1/10 Gigabit Ethernet saved in TDMS file format. The USRP offers frequency ranges up to 4.4 GHz with up to 20 MHz of instantaneous bandwidth. Specifically, 2.4GHz carrier frequency with 1MHz bandwidth was used in this experimental setup. The testbed consists of a USRP transmitter - receiver pair connected to local computers. At the transmitter end, the LabView software is used to generate modulated signals and sent over-the-air via the USRP. Radio Frequency (RF) IQ traces were captured by NI USRP 2932 in receiver mode and LabVIEW software. Since the data is "over-the-air", phenomena like noise and distortion are well accounted for. For the three scenarios, we collected real data transmitted over-the-air via our USRP setup. The details of the parameters of the data in each scenario are given in Table 1.

III. BACKGROUND OF DEEP RF LEARNING

Supervised learning using deep neural network for RF dataset is referred to as deep RF learning. It is expressed as a function $f()$ that models the mapping from the in-phase component (I) and the quadrature (Q) component of the RF front-end denoted as the input data X to the class label Y. The mapping, $f()$ mathematically expressed in equations 1 and 2, can adequately predict the label $\hat{y}$ for a new data sample generated from the same underlying stochastic process. $f()$ is the trained RF learning model, X is the IQ training data, ϵ is additive noise, Y is the class label and $\hat{y}$ is predicted label [6].

$$Y = f(X) + \epsilon \qquad (1)$$
$$\hat{Y} = f(X_{new}) \qquad (2)$$

To learn this input - output relationship, a weight matrix W is estimated. A loss function $l(x, y, w)$ is computed as a point-wise measure of error between the model prediction $f(X)$ and the observed ground truth Y for each value of W where x is a sample data point in RF data X and y is a label in class label Y. To estimate W for all the data points in the dataset, a cost function $J(W)$ which is the average loss over all points in X is computed using equation 3 [6], [20]:

$$J(W) \equiv J(X, W) = \frac{1}{n} \sum_{(x_i, y_i) \in X} l(x_i, y_i, w) \qquad (3)$$

where $i = 1, 2, \ldots, n$ represents the number of training data. The model is derived by minimizing the cost function $J(W)$ by:

$$\underset{w \in \mathbb{R}^n}{\arg\min} J(W) \qquad (4)$$

A. MOTIVATION FOR USING DEEP LEARNING

Although research has been done using traditional machine learning to solve problem in wireless communications in the past such as [21]–[23], the results did not come close to the guaranteed and accurate results from systems and channel models developed based on information theory and signal processing.

O'Shea et al showed that deep learning algorithms trained on RF In-phase and Quadrature (IQ) data outperformed the traditional methods used for modulation recognition based on expert features [13]. The authors compared Decision Trees, Naive Bayes, K-Nearest Neighbors, Support Vector Machines (SVM), Deep Neural Networks and Convolutional Neural Networks. The convolutional neural network has the highest accuracy of about 87.4% and significantly outperforms the accuracy of the machine learning algorithms on expert features. In addition, authors of [24] compared SVM and deep neural networks for RF transmitter identification and reported that the deep neural network outperforms the SVM.

Furthermore, machine learning models, specifically, Decision Trees, Random Forest, Adaboosted Decision Trees were developed on the datasets used for Spectrum Occupancy Prediction in [7]. The results are detailed in Figure 2. The Convolutional Neural Network shows better accuracy than the traditional machine learning algorithms.
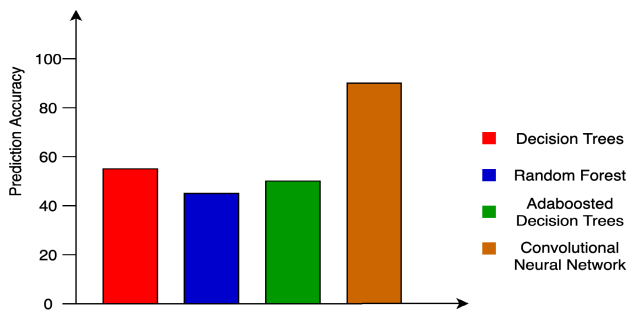
FIGURE 2. Machine learning and deep learning results.

TABLE 2. CNN and RNN(LSTM) parameters.

Parameters	CNN	RNN(LSTM)
Number of Classes	10	10
Number of Parameters	312554	102710
Training Time	219.8 secs	691.5 secs
Accuracy	97.81	98.31

Traditional machine learning defines a set of programmed features on the data and extract these features as part of the machine learning pipeline. The key differentiator of deep learning is the ability to learn the underlying features directly from the data as opposed to being hand engineered. In many practical situations, hand engineering these features can be extremely brittle. Thus, deep neural network can do better feature representation than classical machine learning models. In a very complex wireless environment, the features are complicated and the deep neural network are able to represent it better than traditional machine learning algorithms by doing automatic extraction of discriminative information from the data.

B. CNN VS RNN

To determine the DL model to be used in this work, we compared the Long Short Term Memory (LSTM) a variant of Recurrent Neural Network (RNN) known to perform well on sequence data to the Convolution Neural Network known to perform well on grid-like data for automatic modulation classification. The dataset used in Section V-A is used for this comparison. Table 2 shows the number of parameters and training time to achieve similar classification accuracy.

From the comparison in Table 2 the CNN performance in terms of accuracy is comparable to that of the LSTM given that the train and test data size, activation function, optimizer and other parameters of the network are the same. We use the CNN in this work as it shows comparable results in a shorter training time. The CNNs have shown considerable success on raw IQ traces [9], [13], [15].

C. CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks (CNNs) are a type of feed-forward neural network known to perform convolution operation. A CNN model composes of multiple processing

layers to learn different level features. Combining these hierarchy features preserves extremely discriminative and effective deep representations. Given an input $X(m, n)$ and a filter $K(i, j)$ the result of the convolution of the input and filter, $S(i, j)$ called the feature map is given by Equation 5. The convolution operation is achieved by taking the dot product of two inputs over a finite number of samples. It is an integral that expresses amount of overlap of K as it is shifted over X.

$$S(i, j) = (K * X)(i, j) = \sum_{m} \sum_{n} X(i-m, j-n)K(m, n) \quad (5)$$

The convolutional layers includes a set of neurons. Each neuron is a set of learnable weights and bias. Neurons in this layer take the local receptive fields of feature maps in the previous layers as input and identify local patterns [20]. Assuming S_j^l is the j-th feature map in l-th layer, and $S_j^{l-1}(m = 1, \ldots, M)$ are the outputs of the l-1th layer, S_j^l is calculated by:

$$S_j^l = \delta \left(\sum_{m=1}^{M} w_{jm}^l * S_m^{l-1} + b_j^l \right) \quad (6)$$

where w_{jm}^l is the weight connected to the m-th feature map in the previous layer, b_j^l is the j-th bias of the l-th layer, and $\delta()$ is the rectified linear unit [25]. In general, several pooling layers are periodically inserted in between successive convolutional layers to progressively decrease output scale of the intermediate activation maps. In the fully-connected (FC) layer, neurons have connections to all activations of the previous layer.

The CNN architecture used in this work is modeled as the network used for modulation classification in [13], [19]. It is a 4-layer CNN model consisting of 2 convolutional layers and 2 dense fully connected layers. The layers use Rectified Linear Unit (ReLU) activation function except for a Softmax activation at the output layer. The convolutional layers have 2 dimensional zero padding at its input to preserve the spatial size of the input volume so the input and output width and height are the same. *Adam optimizer* and *categorical cross entropy* loss is used in this model. *Glorot uniform initialization* was used for kernel initialization of all convolution layers and *He normal initialization* for the dense layers. The output layer is a softmax activation function for classification. The Convolutional Neural Network is implemented in Keras using TensorFlow [26].

D. SIZE OF DATA IN TRAINING

Developing DL models using adequate data that captures the relationship between the features and labels in a supervised learning context is an important consideration for DL model development. This is because insufficient training data might make a model inaccurate. Numerous research work split available data in ratio 80:20 for train and test respectively. Although this seems as the generally accepted ratio, of more importance is the size of the training data in terms of number of samples. The authors in [27] investigated the correlation between training dataset size and classification

TABLE 3. CNN architecture.

Parameters	Value
Batch Size	256
Learning Rate	1×10^{-4}
Maximum Pooling	2 x 2
Strides	(1 1)
Convolutional layer FIlter	256
Kernel Size	1 x 3
Number of Classes	10
Training & Test Sizes	80% & 20%
Number of Parameters	312554
Optimizer	Adam Optimizer
Activation Function	ReLU

TABLE 4. Classification accuracy when training and testing using RF data with the same SNR using CNN and RNN.

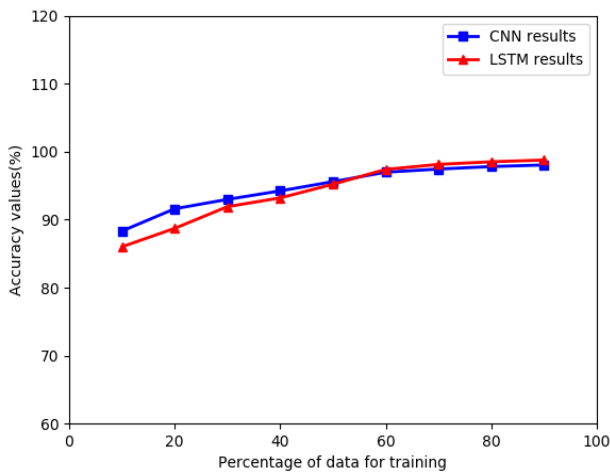| Train ratio	Test ratio	CNN Accuracy(%)	RNN Accuracy(%)
90	10	98.04	98.77
80	20	97.81	98.31
70	30	97.44	98.13
60	40	96.98	97.39
50	50	95.59	95.21
40	60	94.23	93.20
30	70	92.98	91.91
20	80	91.60	88.72
10	90	88.32	86.01

FIGURE 3. Accuracy plot for variation in size of training data.

accuracy for transmitter classification applications. This is done by investigating whether the rules-of-thumb used in neural network research applies in CNN-based transmitter task. Using 100 million IQ samples for the 10-class classification problem in section V-A, each class with 10 million samples, we train DL models using various ratio for training and testing as shown in Table 4 to establish the effect of data size in developing DL models. It is observed that the accuracy is above 95% for both CNN and RNN when 50% or more of the data was used for training. Figure 3 shows the accuracy plot for variation in size of training data.

IV. DEEP LEARNING TRAINING AND TESTING STRATEGIES

The goal here is not to design a new deep learning model but rather to explore the proper training and testing of deep learning models such as CNN for RF spectrum learning under
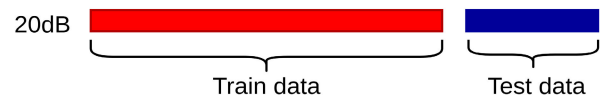

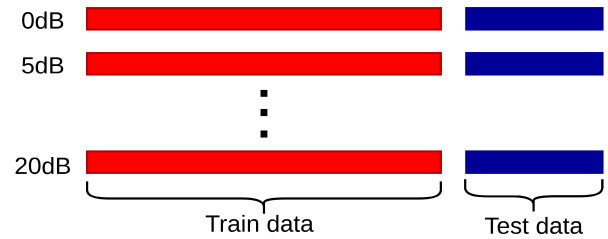**FIGURE 4.** Train and Test using RF data at a fixed SNR.

FIGURE 5. Train and Test using RF data at multiple fixed SNR.

different scenarios in wireless communications. The variation in SNR values of the training and testing datasets is a means to model the varying attributes of the communication signal due to interference, channel effects, and general degradation of the quality of the received signal. The SNR which is the measure of information content compared to noise, is a key attribute to consider during training and testing of deep learning models for communication systems. In this section, we study and analyze how SNR selection for training and testing impacts RF learning for practical systems and how it relates to the scenarios highlighted in Figure 7.

A. TRAINING AND TESTING ON ONE FIXED SNR

We consider training and testing a deep neural network on a single SNR level. This is relevant when there is an intended transmitter - receiver pair with closed loop power control or other means to maintain the SNR level as described in Figure 7A. Owing to the fact that the SNR is maintained, it is possible for the deep learning module developed on this fixed SNR level on a receiver to properly identify modulation used from the received signal in practical situations. Figure 4 shows scheme of training and testing on the same 20dB SNR data.

B. TRAINING AND TESTING ON MULTIPLE FIXED SNRs

Training on multiple fixed SNR levels and testing on the same SNR level used in training is achieved by splitting the data into training and testing sets as illustrated in Figure 5. This approach has been adopted by various researchers using different stepsizes. For instance, [13], [15], [28] used a step-size of 2dB while [7], [9] used a step-size of 5dB. As discussed in section V, this is with the notion of capturing all variations in SNR of a received signal.

C. TRAINING ON MULTIPLE FIXED SNRs AND TESTING ON SNR NOT SEEN IN TRAINING

Training is similar to the method described in section IV-B but testing on the same set of SNR may not be practical for
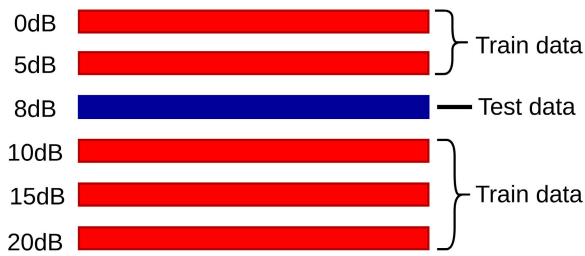
FIGURE 6. Train using RF data at multiple fixed SNRs and test using RF data at a different SNR from that used during training.

many real-life applications as the SNR level of a received signal may vary due to channel effects. Figure 6 shows an example where training is done on 0, 5, 10, 15 and 20dB and testing is done on 8dB. This is a more practical scenario for RF learning.

V. RF LEARNING IN DIFFERENT SCENARIOS

For an area of research as broad as wireless communications systems, it is important to clearly show the scenarios being considered for a particular problem. In this section, we highlight three unique scenarios for RF learning as shown in Figure 7 in which the RF learning problem must be uniquely designed [19]. This is of course not an exhaustive list but we use it to highlight various possibilities and differences in RF learning.

Figure 7A is an intended transmitter-receiver pair with closed loop power control system to keep the received signal-to-noise ratio fairly constant. Scenario 7B represent a shared spectrum where there is co-existence of multiple systems in the same unlicensed band such as LTE-U and Wifi [29], [30]. As shown in the diagram, there is interference from the LTE-U system to the Wifi receiver. Figure 7C describes a spectrum monitoring system in a network with the aim of coordinating the entire communication system for resource management and optimal performance. RF learning implementation and proper performance evaluation for these three scenarios are discussed here.

A. SCENARIO 1: INTENDED TRANSMITTER-RECEIVER PAIR

1) INTRODUCTION

This scenario is depicted in Figure 7A. For instance, the power control scheme in 4G Long Term Evolution (LTE) uplink is used to maintain a constant SNR at the receiver. Specifically, eNodeB estimates the SNR of the received signal and compares it to a target SNR value. Based on the comparison, transmit power control notifies the User Equipment (UE) to adjust the uplink transmission power to ensure a fixed SNR at eNodeB [31].

2) PROBLEM STATEMENT

Here we consider Automatic Modulation Classification (AMC) problem where the RF data for training and testing is at the same SNR level as we have in 4G LTE. The

goal of AMC is to recognize the modulation scheme of a detected signal. AMC is the process between the detection of a signal and its demodulation which is an important step towards developing an intelligent radio receiver [32].

3) DATASET DESCRIPTION

Over-the-air dataset generated using National Instruments Universal Software Radio Peripheral (NI USRP) based testbed [33] and LabVIEW is used following the setup described in Section II. Data for 10 modulation schemes, namely BPSK, QPSK, 8PSK, 16PSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM and OQPSK, were generated in LabVIEW and transmitted over-the-air via a USRP 2932 transmitter-receiver pair in the lab. In order to model a realistic channel where noise and channel impairment are unavoidable, we moved the USRP around as well as walked through the line of sight. We collected datasets of varying SNRs between 0dB and 12dB with a step-size of 0.5dB. For each modulation type, we collected a total of 10 million IQ samples.

4) DEEP LEARNING MODEL

Convolutional Neural Network model described in Section III-C is used. It is developed by training and testing on a 10dB SNR level with 80% of data used for training and 20% used for testing. The CNN model consists of 312,554 trained parameters. The algorithm for AMC is given in Algorithm 1.

Algorithm 1 CNN Model Development for AMC

Input: RF data for 10 modulation schemes at fixed SNR X and corresponding labels Y

Split dataset into **training (80%)** and **testing (20%)**

Normalize the data $z = \frac{x - min(x)}{max(x) - min(x)}$

Train the CNN model:

1: initialize $\boldsymbol{\theta}$

   Training Process

2: **for** *numberofepochs* **do**

3:   CNN forward pass $\longrightarrow Y_i'$

4:   calculate loss function $L \longrightarrow |Y - Y_i'|$

5:   perform back propagation $\longrightarrow \frac{\partial L}{\partial \theta_i}$

6:   update CNN weights, $\theta_i \longrightarrow \theta_{i+1} + \eta \frac{\partial L}{\partial \theta_i}$

7: **end for**

8: **return** $\boldsymbol{\theta}$

Test the CNN model: *Testing Process*

9: **for** *Z in ModulationType(fixedSNR)* **do**

10:   CNN forward pass $\longrightarrow Y_i'$

11:   Accuracy $\longrightarrow \frac{\sum_{i=1}^{N} Y_i == Y_i'}{\sum_{i=1}^{N} Y_i}$

12: **end for**

13: **return** *Accuracy*

---

5) RESULTS

Figure 8 shows the confusion matrix when training and testing is done on 10dB SNR with a classification accuracy of

A: Intended Receiver   **B: Co-existence of multiple systems**   **C: Spectrum monitor**
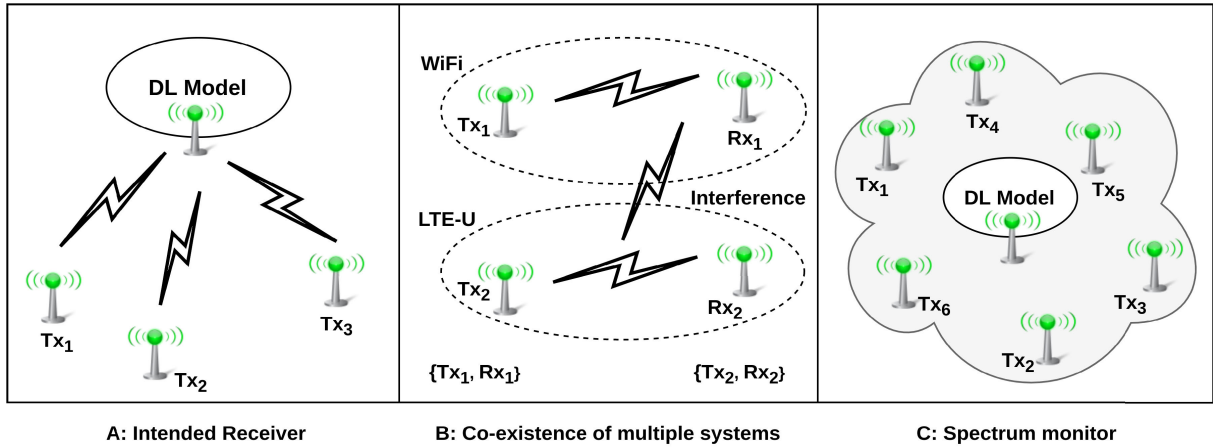
FIGURE 7. Three scenarios considered for radio frequency learning.

97.81%. Furthermore, another CNN model is trained using 0dB RF data. For both 0dB and 10dB CNN models, testing was done using SNR levels from 0dB to 10dB with 1dB increment. These include SNR levels not seen in training. Figure 1 shows the testing accuracies for 0dB and 10dB model. We observed very good testing accuracies when training and testing on the same SNR level. However, testing on another SNR level not seen in training resulted in poor accuracies. As an example, testing the 10dB model with 6dB data gave an accuracy of only 18.85% comparing to testing with 10dB data with an accuracy of 97.81%. Accuracy from 10dB model is generally higher than that of 0dB model. This is due to the better data quality when SNR is higher.

Good classification accuracy is achieved when training and testing on RF data with the same SNR but accuracy drops drastically when testing on RF data with a different SNR level. The confusion matrix in Figure 8 is diagonally dominant as expected. However, it is observed that 16QAM being misclassified as 64QAM, and 64QAM misclassified as 128QAM. This is because 16QAM could not be easily distinguished from 64QAM, as 64QAM constellation points are traversed by 16QAM points. Similar situation occurs for the 64QAM and 128QAM datasets.

B. SCENARIO 2: CO-EXISTENCE OF MULTIPLE SYSTEMS IN THE SAME UNLICENSED BAND SUCH AS LTE-U AND WiFi

1) INTRODUCTION

The need for better spectrum utilization has triggered spectrum sharing such as the coexistence of WiFi and LTE in unlicensed bands [29], [30] as described in Figure 7B. To achieve opportunistic access and ensure fair share of the spectrum without causing undesired interference, improved sensing and signal identification methods to detect and pinpoint spectrum users as well as interferer is important. Such methods and algorithms will be beneficial to spectrum sensing, dynamic spectrum access, and cognitive radio [34]. In this work, we consider LTE and WiFi coexisting in the same frequency band. Traditional power estimation and spectrum sensing can only detect whether the spectrum is occupied
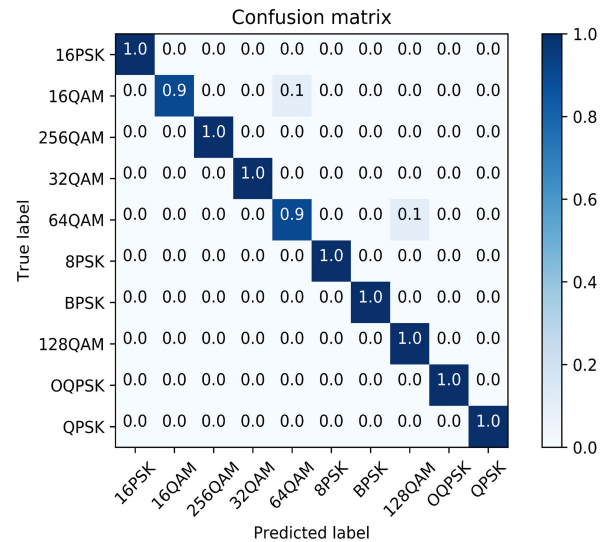

FIGURE 8. Confusion matrix for training and testing using RF data at 10dB fixed SNR.

or not as in energy detection [35]. To know the waveform occupying the spectrum, methods such as matched filter detection or cyclostationary detection methods require some prior knowledge [36]. The test statistics for these methods are generated by using model-based features such as eigenvalues from the sample covariance matrix and signal energy therefore detection ability largely depends on presumed model. Furthermore, these statistics may not adequately exploit the potential of signal sample, thus, new data-driven deep learning-based detectors with test statistics automatically generated from samples of the signals are proposed [37]. The authors in [35] gave good background and related work to spectrum sensing for cognitive radios.

2) PROBLEM STATEMENT

In this case, a 4-class classification problem is formulated to identify the type and number of users in a shared spectrum.

The four classes are:

- idle
- system1 (LTE)
- system2 (WiFi)
- system1 + system2 (LTE + WiFi)

idle implies no system is transmitting, i.e., only background noise is in the measurements. *system1* denotes that LTE is transmitting, while *system2* denotes that WiFi is transmitting. *system1 + system2* denotes both systems are transmitting simultaneously.

3) DATASET DESCRIPTION

For the four different transmission scenarios considered, RF traces were collected using NI USRP and LabVIEW software. Specifically, RF data for the four different coexisting scenarios described above were collected from the testbed for SNR values of 0dB to 12dB with increment of 1dB. We collected 10 million IQ samples for each setup considered.

4) DEEP LEARNING MODEL

Convolutional Neural Network model architecture described in Section III-C is used. It is developed by training and testing using RF data of the same set of SNR levels. For training, we used 2 million samples from each of the 4 SNR levels considered across the 4 co-existence scenarios giving a total of 32 million samples. We used the same data combination approach for testing with 400 thousand samples selected for each SNR level giving a total of 6.4 million samples. The CNN model has a total of 276,132 trained parameters and the algorithm for co-existence of multiple systems in the same unlicensed band is given in Algorithm 2.

Algorithm 2 CNN Development for Co-Existence of Multiple Systems in the Same Unlicensed Band

---

Input: RF data for 4 classes considered at multiple fixed SNRs X and corresponding class labels Y
  Data selection and representation for train and test
  Normalize the data $z = \frac{x - min(x)}{max(x) - min(x)}$
Train the CNN model:
  1: *Same Training Process as Algorithm 1*
Test the CNN model:
  2: Load saved weights θ
    Testing Process
  3: **for** $Z in Coexistence Scenario$ **do**
  4:    CNN forward pass $\longrightarrow Y_i'$
  5:    Accuracy $\longrightarrow \frac{\sum_{i=1}^{N} Y_i == Y_i'}{\sum_{i=1}^{N} Y_i}$
  6: **end for**
  7: **return** *Accuracy*

---

5) RESULTS

Results for training and testing on the same set of SNRs are detailed in Table 5. High accuracies are observed in all three cases. Specifically, an accuracy of 99.35% is obtained when

TABLE 5. Classification accuracy when training and testing using RF data with the same (multiple) fixed SNR.

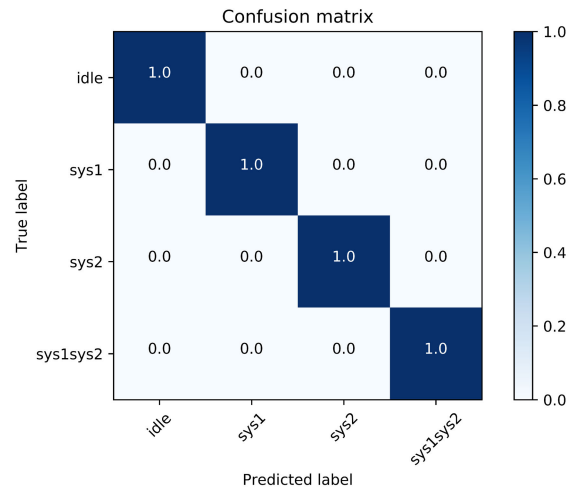| Train SNR(dB)	Test SNR(dB)	Accuracy(%)
1, 2, 3, 4	1, 2, 3, 4	98.36
2, 4, 6, 8	2, 4, 6, 8	99.35
0, 4, 8, 12	0, 4, 8, 12	98.97

FIGURE 9. Confusion matrix for training and testing on multiple fixed SNR values.

training on 2dB, 4dB, 6dB, 8dB and testing on the same set of SNRs. Figure 9 shows the confusion matrix. When training and testing using RF data with the same (multiple) fixed SNR, we obtained excellent classification accuracy. In this specific example, the classification accuracy is even higher than that in scenario 1. This may be because there is a smaller number of classes and/or there exists better distinguishable classes compared to the case in scenario 1 where there are classes that are less distinguishable, e.g., 16 QAM vs. 64 QAM.

C. SCENARIO 3: SPECTRUM MONITORING

1) INTRODUCTION

Challenges of spectrum sharing include how to enact a policy which is mutually beneficial such that all parties involved can deliver their core capabilities and also ensure that missions are protected. For instance, the use of 1697-1710MHz band is shared by National Oceanic and Atmospheric Administration (NOAA) for its weather satellite operation in the downlink while commercial wireless companies share their uplink transmission for their user equipment in this band [16]. For these sharing systems, there is a need for spectrum monitoring as discussed in [38]. It is noted that FCC did not set any limit on the technology that can be deployed within the band. New technology such as 5G systems and narrow band internet of things (NB-IoT) can be deployed which can significantly alter the propagation model used to design and develop the monitoring system [16]. Deep learning based model can be developed for spectrum monitoring to detect, identify source and classify signals in a bid to enable beneficial spectrum

TABLE 6. Prediction accuracy of CNN Model 1 (trained using RF data at SNR levels from 0dB to 3dB with step size of 0.5dB, and leave the RF data at 1.5dB out during training).

Train SNR(dB)	Test SNR(dB)	Accuracy(%)
0, 0.5, 1, 2, 2.5, 3	0	96.82
	0.5	98.48
	1	97.79
	1.5	94.47
	2	98.94
	2.5	98.73
	3	99.08

sharing and achieve optimum spectrum utilization as depicted in Figure 7C.

2) PROBLEM STATEMENT

For spectrum monitoring, we use a modulation recognition problem as an example since interest is in the identification of the modulation type of the transmitter. Since the SNR of the received signal at the spectrum monitor may vary across a wide range, we test the performance of the deep learning model using RF data at both SNR levels seen in training and SNR levels not seen in training as discussed in Section IV-C.

3) DATASET DESCRIPTION

The dataset described in Section V-A3 for automatic modulation recognition is used. It consists of 10 modulation types with varying SNR levels.

4) DEEP LEARNING MODEL

Convolutional Neural Network model described in Section III-C is used. Three CNN models (with the same model architecture) were trained using RF data comprising SNR levels of different step sizes with one SNR level left out, as listed in Tables 6, 7, and 8. Testing is done using RF data at SNR levels seen in training as well as SNR levels not seen in training. CNN Model 1 is trained using RF data at SNR levels 0dB, 0.5dB, 1dB, 2dB, 2.5dB, and 3dB, while it is tested using RF data at SNR levels from 0dB to 3dB with step size of 0.5 dB, including the RF data at 1.5dB. Similarly, CNN Model 2 is trained using RF data at SNR levels from 0dB to 6dB with step size of 1dB, and leave the RF data at 3dB out during training, while CNN Model 3 is trained using RF data at SNR levels from 0dB to 12dB with step size of 2dB, leaving out the RF data at 6dB during training. 1.5 million samples of each SNR level used in developing a model is aggregated to form a total of 9 million samples for each class and a total of 90 million samples of IQ data for training. 2 million samples of IQ data for each SNR level across 10 classes is aggregated to give 20 million samples for testing at each SNR level. Testing is done on one SNR level at a time. Each CNN model has a total of 312,554 trained parameters. The algorithm for spectrum monitoring is given in Algorithm 3.

Algorithm 3 CNN Model Development for Spectrum Monitoring

Input: RF data for 10 classes considered at varying SNRs X and corresponding class labels Y
  Data selection and representation for train and test
  Normalize the data $z = \frac{x - min(x)}{max(x) - min(x)}$
Train the CNN model:
 1: *Same Training Process as Algorithm 1*
Test the CNN model:
 2: Load saved weights θ
    Testing Process
 3: **for** $Z in ModulationType(varying SNR)$ **do**
 4:   CNN forward pass $\longrightarrow Y_i'$
 5:   Accuracy $\longrightarrow \frac{\sum_{i=1}^{N} Y_i == Y_i'}{\sum_{i=1}^{N} Y_i}$
 6: **end for**
 7: **return** *Accuracy*

TABLE 7. Prediction accuracy of CNN Model 2 (trained using RF data at SNR levels from 0dB to 6dB with step size of 1dB, and leave the RF data at 3dB out during training).

Train SNR(dB)	Test SNR(dB)	Accuracy(%)
0, 1, 2, 4, 5, 6	0	99.02
	1	98.86
	2	98.59
	3	97.55
	4	99.70
	5	99.14
	6	100

5) RESULTS

The testing results for the three CNNs are given in Tables 6, 7, and 8, respectively. It is observed that the prediction accuracy is high for all the RF data with the same SNR levels included in the training as expected. However, the prediction accuracy varies a lot for RF data with SNR not included in training. There is an accuracy of 94.47% when testing is done with 1.5dB dataset not seen in training CNN Model 1, 97.55% when testing CNN Model 2 with 3dB data and a very low 43.63% when testing CNN Model 3 with 6dB data. It seems that the step size (0.5dB, 1dB, and 2dB) affects how well the deep learning model generalizes. The results also suggest that there exist a sweet spot of step size such that

TABLE 8. Prediction accuracy of CNN Model 3 (trained using RF data at SNR levels from 0dB to 12dB with step size of 2dB, and leave the RF data at 6dB out during training).

Train SNR(dB)	Test SNR(dB)	Accuracy(%)
0, 2, 4, 8, 10, 12	0	97.37
	2	98.01
	4	98.44
	6	43.63
	8	97.84
	10	99.12
	12	99.82

TABLE 9. K - S Statistic and p-value.

		RF data for testing									
	SNR level	1 (dB)	2 (dB)	3 (dB)	4 (dB)	5 (dB)	6 (dB)	7 (dB)	8 (dB)	9 (dB)	10 (dB)
RF data for training	1dB	0.0, 1.0	0.038, 0.0	0.078, 0.0	0.111, 0.0	0.138, 0.0	0.175, 0.0	0.197, 0.0	0.230, 0.0	0.256, 0.0	0.278, 0.0
	2dB	0.038, 0.0	0.0, 1.0	0.042, 0.0	0.077, 0.0	0.105, 0.0	0.144,0.0	0.168, 0.0	0.204, 0.0	0.232, 0.0	0.255, 0.0
	3dB	0.078, 0.0	0.042,0.0	0.0, 1.0	0.035, 0.0	0.064, 0.0	0.107, 0.0	0.133, 0.0	0.173, 0.0	0.203, 0.0	0.229, 0.0
	4dB	0.111, 0.0	0.077, 0.0	0.035, 0.0	0.0, 1.0	0.030, 0.0	0.075, 0.0	0.102, 0.0	0.144, 0.0	0.177, 0.0	0.206, 0.0
	5dB	0.138, 0.0	0.105, 0.0	0.064, 0.0	0.03, 0.0	0.0, 1.0	0.046, 0.0	0.074, 0.0	0.118, 0.0	0.152, 0.0	0.182, 0.0
	6dB	0.175, 0.0	0.144, 0.0	0.107, 0.0	0.075, 0.0	0.046, 0.0	0.0, 1.0	0.029, 0.0	0.077, 0.0	0.113, 0.0	0.146, 0.0
	7dB	0.197, 0.0	0.168, 0.0	0.133, 0.0	0.102, 0.0	0.073, 0.0	0.028, 0.0	0.0,1.0	0.049, 0.0	0.086, 0.0	0.120, 0.0
	8dB	0.230, 0.0	0.204, 0.0	0.173, 0.0	0.145, 0.0	0.118, 0.0	0.077, 0.0	0.049, 0.0	0.0, 1.0	0.040, 0.0	0.075, 0.0
	9dB	0.256, 0.0	0.232, 0.0	0.203, 0.0	0.177, 0.0	0.152, 0.0	0.113, 0.0	0.086, 0.0	0.040, 0.0	0.0, 1.0	0.036, 0.0
	10dB	0.278, 0.0	0.255, 0.0	0.229, 0.0	0.206, 0.0	0.182, 0.0	0.146, 0.0	0.120, 0.0	0.075, 0.0	0.035, 0.0	0.0, 1.0

the training data may cover the underlying distribution well and allow the deep learning model to learn the distribution with high accuracy and be able to generalize well. The step size of the SNR of the dataset used in training is an important factor to consider for RF learning. Our observation remains consistent when deeper models have been used with more parameters to train on.

VI. FURTHER DISCUSSIONS

It is generally assumed that datasets with the same distribution are used in training and testing machine learning models. Empirically, the uniform convergence theory states that under this condition, the training and testing errors are close in values [39]. However, these conditions are not necessarily fulfilled in reality as training is done in a domain separate from the testing domain. Moreover, wireless signals are impaired by a number of time-varying effects such as noise, fading, and channel impairments, yet detailed information about the noise and the wireless channel may not be available in practice.

A. DISTRIBUTION ALIGNMENT

When training and testing on RF data at the same SNR level, higher classification accuracy is observed compared to when the machine learning model is tested on RF data at SNR levels not seen in training. This observation supports the general understanding that discriminative learning methods such as convolutional neural networks perform very well when training and testing data sets are drawn from the same distribution [39]. This is known as distribution alignment problem because deep learning models have an intrinsic bias to data seen in training and this does not allow the model to generalize well to unseen test data. Researchers and practitioners highlight that the increase in generalization error of supervised models is directly proportional to an increase in the variance of the training and testing distributions [40].

Kolmogorov - Simirnov (K-S) statistic and p-value [41] may be applied to examine the relationship between data distributions. K-S statistic follows the hypothesis that the distribution of two samples are the same if the K-S statistic value is small and p-value is high. Table 9 shows the (K-S statistic, p-value) pair under different training and testing data. It is observed that RF datasets at the same SNR are of the same distribution. The farther the SNR values are apart from each other, the higher the K-S statistic value, which indicates a bigger difference in distribution. Therefore, the generalization observed in CNN Model 1 and CNN Model 2 can be explained based on the K-S statistics and p-value, in other words, datasets with fine granularity (small step size) appear to be closer in distribution, thus better classification accuracy can be achieved than datasets with coarse granularity in SNR levels.

Furthermore [17] studied robust generalizations in the context of adversarial learning where small perturbations can cause state-of-the-art DL classifiers to produce incorrect prediction. This is compared to standard generalization in which there is no perturbation in the data. Our work directly relates to this in terms of SNR variations and analysis to develop robust models. Adversarial learning is unique in wireless communications systems because unlike the computer vision domain that assumes that adversarial and legitimate inputs are received 'as is'' by the classifier, wireless communication signals are subjected to perturbation due to the channel. This can cause a significant change in the distribution of the received signal. The authors studied the effect of sample complexity of standard generalization compared to that of adversarially robust generalization. The study established that even for a simple data distribution such as a mixture of two-class conditional Gaussians, the sample complexity for robust generalization is significantly larger than that of standard generalization regardless of model and learning algorithm (see [17] for mathematical formulation). The results corroborate our conclusion that SNR levels with fine granularity should be used for model development thus increasing sample complexity to adequately cover the underlying distribution. This allows the deep learning model to learn the distribution with high accuracy and generalize well.

B. COMPARISON WITH COMPUTER VISION DOMAIN

Gleaning from the computer vision domain, computer vision systems are trained and tested on images of high quality for image recognition tasks, yet quality of the input images cannot be pre-determined in practical applications. There are several recent study on characterizing the outcome of image quality on computer vision systems, for example in [42]–[44]. They studied the effect of image quality distortions on deep neural network models. Five types of quality

distortions, namely contrast, noise, blur, JPEG2000 compression and JPEG compression were considered in [42]. Specifically, in the experiment of [42], Gaussian noise was added to each color component (segment) of each pixel separately and the standard deviation of the noise was varied from 10 to 100 in steps of 10. They trained on high quality images and tested on images with varying noise level. The accuracy of the deep neural networks decreased significantly as the noise level increased. For instance, at the noise standard deviation of 90, the network performance became less than 20%. This is also observed in [43] who noted that the performance of image recognition models degrades greatly when testing on corruptions such as noise unseen in training.

The notion of adding distortion such as noise in computer vision applications can be directly related to SNR in RF learning where we compare high quality images to RF data at high SNR levels, and images with high distortion to RF data at low SNR levels. From our tests, we observe a trend similar to that of the computer vision domain when training on data from one SNR level and evaluating on another SNR level not seen in training.

C. INFERENCE IN A DYNAMIC RF ENVIRONMENT

In practical RF environments, our pre-trained model performs inference which is a feed forward computation with no iteration, thus it is very fast. Training the model on the other hand requires iterations but it is done offline. For inference, only typical edge computing devices are needed. For instance, we measured a latency of 68 μ sec in inference time on NVIDIA Tesla P100-DGX1-32GB GPU using our model with 312554 parameters. A bigger model such as *MobileNet* [45] with 4.2 million parameters has a latency of 2.4 m sec on Quad-core Cortex-A53 @ 1.5GHz + Edge TPU according to benchmark results from Coral [46].

VII. RELATED WORKS

Previous publications have shown that deep learning on RF datasets have the potential to transform the communication problems as it has done in computer vision and speech recognition. Several research groups have started exploring the capabilities of deep learning in building applications for wireless communications systems by using RF data employing state-of-the-art software and hardware tools focusing on various objectives. Among these objectives are device identification and intrusion detection [8], [9], [24], [47] where work is done to determine the transmitter and detect if there is an unwanted transmitter in the system. Similar to this is modulation recognition, identification and classification [12], [14], [48] where work is done to identify the modulation scheme in use. This may find applications in interference management and opportunistic mesh networking, thereby improving the overall radio efficiency. Spectrum sensing and adaptation is another area where work has been done [5], [7], [49]. The work identifies the combination of multiple-user coexisting in license-free frequency bands and predicts the waveform combination presented. Wireless

interference identification had been studied in [15] to determine the source of interference in a coexisting transmission scenario for coexistence management. Furthermore, signal identification was considered in [6] for spectrum monitoring. Such monitoring systems are used for coexistence management, regulatory purposes, standardization and in defense applications.

In the papers discussed above, many of them trained and tested on RF data at the same sets of fixed SNR values. The various scenarios in this paper present different considerations for the selection of SNR values of RF data used in training and testing. Training can be done on a fixed SNR if the received SNR can be controlled at a pre-fixed level. This will not generalize to cases where there is variation in the received SNR values, as described in Figure 7B and 7C where we have unintended transmitter-receiver and spectrum monitor, respectively. From our analysis in previous sections, the SNR step size selection is an important consideration for building a model that capture all SNR variations in practical applications. The data generation procedure, communication system modeling, data selection for training and testing RF learning models are all factors that significantly affect the practicality of the DL model for RF learning.

VIII. CONCLUSION

The availability of wireless big data and state-of-the-art deep learning techniques makes it possible to explore RF learning for the optimization of future communication systems. In this work, we study various use cases for RF learning in future wireless systems, examine different training and testing strategies, and under what conditions these strategies should be used. Our analysis show that to achieve practical RF learning, it is important to understand the scenario for which the model is developed, use appropriate training and performance evaluation strategies.

There are many potential applications for RF learning in wireless communication systems. The need to develop models that will generalize well on these applications has motivated us to do a detailed analysis to understand different scenarios and their requirements. Many previous studies in the literature have been done by training and testing on RF data at the same fixed SNR values. To the best of our knowledge, there are very few considerations for the SNR step sizes used in building RF learning models as well as how testing is performed in specific scenarios. Developing models that generalize well for RF learning such that data from target distributions can be correctly inferred when the training data is from another distribution is key for future intelligent communications systems.

Most of the current studies have used synthetic data generated from various devices with different setup to model the communications systems. Further study may be needed to understand how synthetic and real over-the-air data may affect the performance of RF learning.

IX. ACKNOWLEDGMENT

REFERENCES

[1] S. Zheng, S. Chen, L. Yang, J. Zhu, Z. Luo, J. Hu, and X. Yang, "Big data processing architecture for radio signals empowered by deep learning: Concept, experiment, applications and challenges," *IEEE Access*, vol. 6, pp. 55907–55922, 2018.

[2] L. Qian, J. Zhu, and S. Zhang, "Survey of wireless big data," *J. Commun. Inf. Netw.*, vol. 2, no. 1, pp. 1–18, Mar. 2017, doi: 10.1007/s41650-017-0001-2.

[3] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE Access*, vol. 6, pp. 32328–32338, 2018.

[4] M. Kulin, T. Kazaz, I. Moerman, and E. de Poorter, "A survey on machine learning-based performance improvement of wireless networks: PHY, MAC and network layer," 2020, *arXiv:2001.04561*. [Online]. Available: https://arxiv.org/abs/2001.04561

[5] S. Zheng, S. Chen, P. Qi, H. Zhou, and X. Yang, "Spectrum sensing based on deep learning classification for cognitive radios," *China Commun.*, vol. 17, no. 2, pp. 138–148, Feb. 2020.

[6] M. Kulin, T. Kazaz, I. Moerman, and 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.

[7] O. Omotere, J. Fuller, L. Qian, and Z. Han, "Spectrum occupancy prediction in coexisting wireless systems using deep learning," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–7.

[8] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification using differential constellation trace figure," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1091–1095, Jan. 2020.

[9] H. Jafari, O. Omotere, D. Adesina, H.-H. Wu, and L. Qian, "IoT devices fingerprinting using deep learning," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2018, pp. 1–9.

[10] J. Bassey, D. Adesina, X. Li, L. Qian, A. Aved, and T. Kroecker, "Intrusion detection for IoT devices based on RF fingerprinting using deep learning," in *Proc. 4th Int. Conf. Fog Mobile Edge Comput. (FMEC)*, Jun. 2019, pp. 98–104.

[11] Y. Luo, H. Hu, Y. Wen, and D. Tao, "Transforming device fingerprinting for wireless security via online multitask metric learning," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 208–219, Jan. 2020.

[12] D. Li, R. Yang, X. Li, and S. Zhu, "Radar signal modulation recognition based on deep joint learning," *IEEE Access*, vol. 8, pp. 48515–48528, 2020.

[13] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Engineering Applications of Neural Networks*, C. Jayne and L. Iliadis, Eds. Cham, Switzerland: Springer, 2016, pp. 213–226.

[14] 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.

[15] M. Schmidt, D. Block, and U. Meier, "Wireless interference identification with convolutional neural networks," 2017, *arXiv:1703.00737*. [Online]. Available: http://arxiv.org/abs/1703.00737

[16] P. Amirshahi, P. Ransom, S. Grippando, and I. Navarro, "Comparison of centralized and distributed spectrum monitoring methods for enabling spectrum sharing between weather satellites and terrestrial networks," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Nov. 2019, pp. 1–8.

[17] L. Schmidt, S. Santurkar, D. Tsipras, K. Talwar, and A. Madry, "Adversarially robust generalization requires more data," in *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2018, pp. 5014–5026. [Online]. Available: http://papers.nips.cc/paper/7749-adversarially-robust-generalization-requires-more-data.pdf

[18] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, and D. Song, "Robust physical-world attacks on deep learning models," 2017, *arXiv:1707.08945*. [Online]. Available: http://arxiv.org/abs/1707.08945

[19] D. Adesina, J. Bassey, and L. Qian, "Practical radio frequency learning for future wireless communication systems," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2019, pp. 311–317.

[20] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[21] M. Ibnkahla, "Applications of neural networks to digital communications—A survey," *Signal Process.*, vol. 80, no. 7, pp. 1185–1215, Jul. 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016516840000030X

[22] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136–1159, 3rd Quart., 2013.

[23] M. Zorzi, A. Zanella, A. Testolin, M. De Filippo De Grazia, and M. Zorzi, "Cognition-based networks: A new perspective on network optimization using learning and distributed intelligence," *IEEE Access*, vol. 3, pp. 1512–1530, 2015.

[24] 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.

[25] Z. Yu, E.-L. Tan, D. Ni, J. Qin, S. Chen, S. Li, B. Lei, and T. Wang, "A deep convolutional neural network-based framework for automatic fetal facial standard plane recognition," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 3, pp. 874–885, May 2018.

[26] F. Chollet *et al.* (2015). *Keras*. [Online]. Available: https://github.com/fchollet/keras

[27] T. Oyedare and J.-M.-J. Park, "Estimating the required training dataset size for transmitter classification using deep learning," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Nov. 2019, pp. 1–10.

[28] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.

[29] Qualcomm-Technologies. (Jun. 2014). *LTE in Unlicensed Spectrum: Harmonious Coexistence With Wi-Fi*. [Online]. Available: https://www.qualcomm.com/media/documents/files/lte-unlicensed-coexistence-whitepaper.pdf

[30] *Enabling Innovative Small Cell Use in 3.5 GHZ Band NPRM Order*, Federal Communications Commission, document FCC 12-148, Dec. 2012.

[31] B. Muhammad and A. Mohammed, "Uplink closed loop power control for LTE system," in *Proc. 6th Int. Conf. Emerg. Technol. (ICET)*, Oct. 2010, pp. 88–93.

[32] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.

[33] O. Omotere, W. Oduola, N. Zou, X. Li, L. Qian, and D. Kataria, "Distributed spectrum monitoring and surveillance using a cognitive radio based testbed," in *Proc. IEEE 37th Sarnoff Symp.*, Sep. 2016, pp. 100–105.

[34] J. Tan, L. Zhang, Y.-C. Liang, and D. Niyato, "Intelligent sharing for LTE and WiFi systems in unlicensed bands: A deep reinforcement learning approach," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2793–2808, May 2020.

[35] B. Soni, D. K. Patel, and M. Lopez-Benitez, "Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics," *IEEE Access*, vol. 8, pp. 97437–97451, 2020.

[36] A. Bagwari, G. S. Tomar, and S. Verma, "Cooperative spectrum sensing based on two-stage detectors with multiple energy detectors and adaptive double threshold in cognitive radio networks," *Can. J. Electr. Comput. Eng.*, vol. 36, no. 4, pp. 172–180, 2013.

[37] J. Xie, J. Fang, C. Liu, and L. Yang, "Unsupervised deep spectrum sensing: A variational auto-encoder based approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5307–5319, May 2020.

[38] *Manual of Regulations and Procedures for Federal Radio Frequency Management* September 2015 Revision of the May 2013 Edition, National Telecommun. Inf. Admin., Washington, DC, USA, March 2016.

[39] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," *Mach. Learn.*, vol. 79, nos. 1–2, pp. 151–175, May 2010. [Online]. Available: http://www.springerlink.com/content/q6qk230685577n52/

[40] B. Sun and K. Saenko, "Subspace distribution alignment for unsupervised domain adaptation," in *Proc. Brit. Mach. Vis. Conf.*, 2015, pp. 1–24.

[41] F. Wang, R. Xu, and Z. Zhong, "Low complexity Kolmogorov-Smirnov modulation classification," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2011, pp. 1607–1611.

[42] S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks," in *Proc. 8th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, Jun. 2016, pp. 1–6.

[43] S. Karahan, M. Kilinc Yildirum, K. Kirtac, F. S. Rende, G. Butun, and H. K. Ekenel, "How image degradations affect deep CNN-based face recognition?" in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2016, pp. 1–5.

[44] D. M. Bin Le, T. W. Rondeau, and C. W. Bostian. (2005). *Modulation Identification Using Neural Networks for Cognitive Radios*. [Online]. Available: https://www.wirelessinnovation.org/assets/Proceedings/2005/2005-sdr05-2-4-01-le.pdf

[45] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: https://arxiv.org/abs/1704.04861

[46] Coral. (2020). *Edge TPU Performance Benchmarks*. [Online]. Available: https://coral.ai/docs/edgetpu/benchmarks/

[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] E. Rusak, L. Schott, R. Zimmermann, J. Bitterwolf, O. Bringmann, M. Bethge, and W. Brendel. (2020). *Increasing the Robustness of NNS Against Image Corruptions by Playing the Game of Noise*. [Online]. Available: https://deepai.org/publication/increasing-the-robustnessof-dnns-against-image-corruptions-by-playing-the-game-of-noise

[49] O. Omotere, L. Qian, R. Jantti, M. Pan, and Z. Han, "Big RF data assisted cognitive radio network coexistence in 3.5GHz band," in *Proc. 26th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2017, pp. 1–8.

DAMILOLA ADESINA (Graduate Student Member, IEEE) received the B.Sc. degree in electronic and electrical engineering from Obafemi Awolowo University, Nigeria, and the M.Sc. degree in telecommunication engineering from the University of Vaasa, Finland. He is currently pursuing the Ph.D. degree in electrical engineering with Prairie View A&M University, Texas, USA. He is also a Research Assistant with the Center of Excellence in Research and Education for Big Military Data Intelligence (CREDIT Center), Prairie View A&M University. His research interests are big data analytics, wireless communications, mobile networks, and artificial intelligence.

JOSHUA BASSEY (Student Member, IEEE) received the B.Eng. degree from the University of East London, U.K., and the master's degrees from Cranfield University, Bedfordshire, U.K., and Prairie View A&M University. He is currently pursuing the Ph.D. degree with Prairie View A&M University. He is also a Graduate Research Assistant with the Center of Excellence in Research and Education for Big Military Data Intelligence (CREDIT Center), Prairie View A&M University. His research interests are big data analytics, artificial intelligence and deep learning, wireless communications, signal processing, and mathematical modeling.
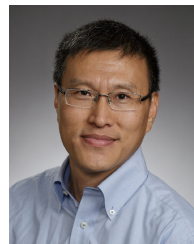
LIJUN QIAN (Senior Member, IEEE) received the B.E. degree from Tsinghua University, Beijing, China, the M.S. degree from the Technion-Israel Institute of Technology, Haifa, Israel, and the Ph.D. degree from Rutgers University, NJ, USA. He was a Technical Staff Member of Bell-Labs Research, Murray Hill, NJ, USA. He is currently a Regents Professor and holds the AT&T Endowment with the Department of Electrical and Computer Engineering, Prairie View A&M University, Prairie View, TX, USA, where he is also the Director of the Center of Excellence in Research and Education for Big Military Data Intelligence (CREDIT Center). He was a Visiting Professor with Aalto University, Espoo, Finland. His research interests are in the area of big data processing, artificial intelligence, wireless communications and mobile networks, network security and intrusion detection, and computational systems biology.

● ● ●