2020 29th Wireless and Optical Communications Conference (WOCC)

5G Signal Identification Using Deep Learning

Mohsen H. Alhazmi, Mofadal Alymani, Hatim Alhazmi,
Alhussain Almarhabi, Abdullah Samarkandi, and Yu-Dong Yao
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ 07030, USA
Email: malhazmi@stevens.edu, malymani@stevens.edu, halhazm1@stevens.edu,
aalmarha@stevens.edu, asamarka@stevens.edu, yyao@stevens.edu

***Abstract*—Spectrum awareness, including identifying different types of signals, is very important in a cellular system environment. In this paper, a neural network is utilized to identify 5G signals among different cellular communications signals, including Long-Term Evolution (LTE) and Universal Mobile Telecommunication Service (UMTS). We explore the use of deep learning in wireless communications systems. We consider the effects of training dataset size, features extracted, and channel fading in our study. Experiment results demonstrate the effectiveness of deep learning neural networks in identifying cellular system signals, including UMTS, LTE, and 5G.**

***Index Terms*—Deep Learning (DL), Classification, Convolutional Neural Network (CNN), Machine learning (ML), Rayleigh Fading, Fifth Generation New Radio (5G), Long-Term Evolution (LTE), Universal Mobile Telecommunication Service (UMTS).**

I. INTRODUCTION

5G has the potential to enable new applications to gain higher quality services around the world. There are several applications that use 5G such as mobile telecommunication, eHealth, autonomous vehicles, smart cities, and the IoT. The 5G frequency band is very complex, there are three 5G bands. For under 1 GHz, it is the Low-Band. For Mid-Band (sub-6 GHz), it is from 3.6 GHz to 6 GHz. For High-band (millimeter-wave), it is from 24 GHz to 40 GHz. Meanwhile, there are existing cellular system technologies using similar frequency band (5G Low-Band) such as UMTS (3G) and LTE (4G).

In the transition period from 3G, 4G, to 5G, different cellular systems may use different cellular standards (3G, 4G, or 5G). In order to achieve successful cellular operations and for the radio resource management, we may need to identify various cellular signals (3G, 4G, and 5G). Recent advancement of deep learning provides tools for identifying various types of signals. For example, Automatic Signal Identification (ASI) is used for several applications in military and commercial communications, such as spectrum surveillance, software-defined, and cognitive radios.

Deep learning has been used for several signal identifications and it has an inherited automatic feature extraction mechanism. Deep learning can easily perform various techniques in order to classify data because of the multiple stacked layers of neural networks. Moreover, deep learning has been introduced to be a useful tool in various tasks such as image classification, automatic speech recognition, and machine translation.

Moreover, for classifying signals, we can define them in two main methods. First, by Likelihood Based identification [1], it can provide us the maximum average probability of a correct classifying class, which used in [2], [3]. Second, Feature-Based technique algorithms are very useful when it comes to standardized signals. Also, there are several algorithms for Feature-Based approach. Identifying LTE signals using the Pilot Induced Cyclostationary (PIC) was used in [4] and the Gaussian Maximum Likelihood (GML) was used in [5], [6].

The convolutional neural network (CNN) model learns different matched filters for several signal-to-noise ratio (SNR) and runs on the time domain IQ data. However, the CNN model may not be systematic with unknown sampling rate data. O'Shea and West [7] have extended the investigation on the effect of CNN layer sizes and depths on classification accuracy. Furthermore, they proposed complex inception modules combining CNN and Long Short Term Memory (LSTM) modules for improving the classification accuracy. However, many of the previous studies have some limitations. For example, [8] has classifiers where each can capture different characteristics in signals and then combine those multiple binary classifiers. In most recent papers [9], [10], the authors have investigated deep learning for the identification of GSM, UMTS, and LTE signals by considering spectrum awareness and utilizing deep learning algorithms for the three different cellular system signals. This paper focuses on 5G signals and utilizes the advantage of CNN in automatic feature extraction for 5G signal identifications.

The rest of this paper is organized as follows. Section II describes the signal models for the 3G, LTE, and 5G signals. Sections III present the convolutional neural network configuration utilized for the signal identification task. The results are shown in Section IV. Finally, the paper is concluded in Section V.

II. 3G, LTE, AND 5G SIGNAL MODELS

In this section, we will introduce the three different types of signals, 3G, LTE, and 5G, as shown in Fig. 1. Our research was built in two steps, generating signals in multiple scenarios (signal types, signal-to-noise ratio levels, fading, and dataset size) and utilizing deep learning algorithms for signal type identification.

978-1-7281-6124-2/20/$31.00 ©2020 IEEE

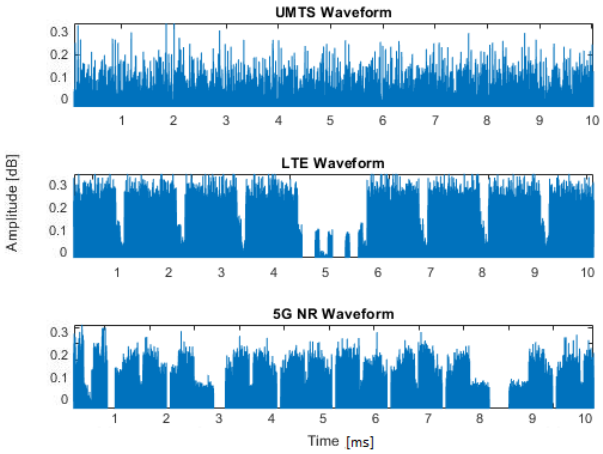

Fig. 1. 3G, LTE, and 5G signal waveforms.

A. 3G

UMTS frequency bands are radio frequencies used by third-generation (3G) wireless Universal Mobile Telecommunications System networks. 3G uses UMTS frequency band between 850 MHz and 1900 MHz [11]. Frame structure is shown in Fig. 2. The superframe consists of 72 frames, which is divided into 15 slots, and each slot has 2560 chips. The 3G frame duration is of 10 ms, and each slot is of 0.667 ms.

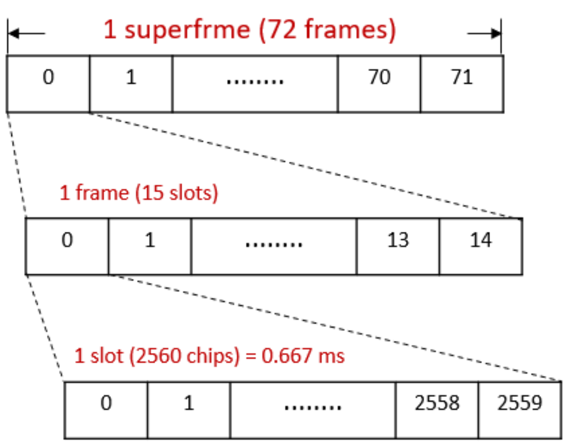

Fig. 2. 3G frame structure.

B. LTE

Long Term Evolution (LTE) telecommunications networks use frequency 700 MHz or 2600 MHz. LTE [12] frame structure has a 10 ms length frames, which is divided into two subframes. LTE frame structure is shown in Fig. 3. LTE frame duration is of 10 ms.

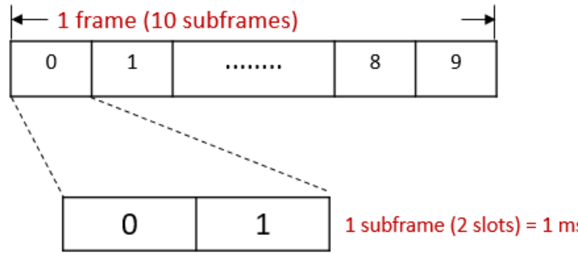

Fig. 3. LTE frame structure.

C. 5G

The frame structure of a 5G signal [13] is shown in Fig. 4. A frame has a duration of 10 ms, which consists of 10 subframes, each having 1 ms duration like LTE. Each subframe has two slots and each slot consists of 14 OFDM symbols. When we compare it to LTE numerology (subcarrier spacing and symbol length), the most outstanding difference is that 5G NR supports multiple different types of sub-carrier spacing, unlike LTE with only one type of subcarrier 15 kHz. 5G will use spectrum in the existing LTE frequency range (600 MHz to 6 GHz and in millimeter wave bands (24–86 GHz). Depending on the ranges, the maximum bandwidth and subcarrier spacing varies. In sub 6 GHz, the maximum bandwidth is 100 MHz and in millimeter wave range the maximum bandwidth is 400 MHz.

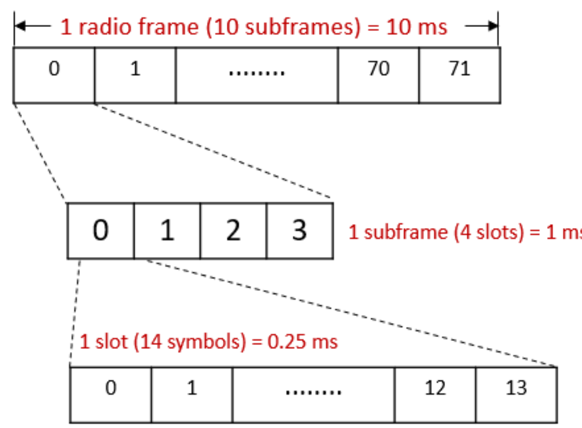

Fig. 4. 5G frame structure.

III. Signal Identification Using Deep Learning

Deep learning is a powerful tool for object classification, which gained great popularity in several application fields, especially in computer vision and image recognition. One possible approach of utilizing deep learning for modulation classification is to convert the received signal into an image form and utilize CNNs to identify the image. In this section, we will introduce how we collect different signals with different SNRs for each class. Also, we will present the CNN model and how the data was implemented to identify signals transmitted in 3G, LTE, and 5G under various deep learning parameters settings.

A. Generating Signals

We generate 1000 random signals in the form of a picture (2-D), it will give us effective method to extract the needed features, with different signal-to-noise ratio levels for each class using MATLAB UMTS function for 3G. Also, we used the "lteRMCDL" function in Matlab for 4G signals. Also, we used a 5G toolbox to generate 5G signals. We repeated this step with different SNRs 20, 15, 10, 5, 3, 0, -3, -5, -10, -15, and -20 dB. Furthermore, we add Rayleigh fading effects to the signals and repeat the same training and testing process. Also, it is so pivotal making sure that the model is set up in an unbiased manner, thus we managed generating the signals with 15 kHz for subcarrier spacing for all the three different types of signals, and OFDM modulation with one subframe. Also, we used one frame size 10 ms for the different types of signal 3G, 4G, and 5G. The signals dataset is splatted into three sets. 60% for the training set and 20% for the testing and finally 20% for the validation sets.

B. Data Processing

After we simulate AWGN channels at different noise levels starting from -20 dB to 20 dB. We collect the noisy version of signals to train and test the CNN model. Also, we further add Rayleigh fading effects to the signals and repeat the same training and testing process. We examine the identification performance using image representations. First, each time-domain signal generated is saved as a JPG image, which contains 128 x 128 pixels. Second, each generated signal is sampled into a two-dimensional vector with a length of 307200 and stored in a CSV file. Second, In order to process image representations, the convolutional layers in the model are set to be a two-dimensional convolutional layer. Also, the real part of the time domain signal is considered in order to improve the efficiency of the identification model. The signal features are observed by re-sampling the received signals in different distributed points, e.g., 1000, 5000, and 10000. Fig. 5 illustrates how the features are fed to the network.

C. CNN Model

We built our CNN model based on the LeNet-5 architecture, which consists of two sets of convolutional and average pooling layers. Then, followed by a flattening convolutional layer, and two fully-connected layers and finally a softmax classifier. Then, we used the convolutional layers with a Maxpooling layer for all of the layers separately in order to get the features out of them and lower the dimension pictures. Next, after we got 4-dim tensors from the third Max-pooling layer and we demolished the 4-dimension tensors to 2-dimension tensors. For the first two fully connected layers, we have 128 neurons for each of them and only the first layer uses Rectified Linear Units (ReLU). In the end, Softmax function was used to identify each class, 3G, LTE, and 5G,

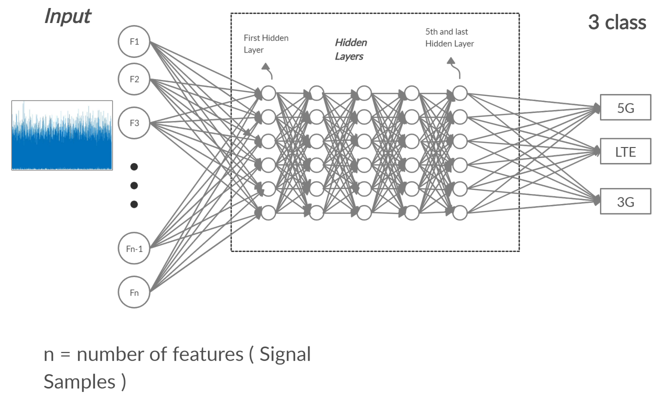

Fig. 5. Convolutional neural network.

which will be explained with more details in the following section. We choses Convolutional Neural Networks (CNN), because it requires minimal processing compared to the other deep learning methods. In addition, CNN is a feed-forwars artificial neural networks. Furthermore we used Python 3.7 version on Jupyter Notebook and Google TensorFlow library for the implementations. The architecture of our model shown in Fig. 5.

IV. Results

In this section, we present the experiment results using the CNN model in identifying signals. We add various levels of white noise to the signals in order to test the performance of the model under different signal-to-noise ratio levels. Also, we test the model using different dataset sizes and different number of features. In addition, the impact of channel fading is examined.

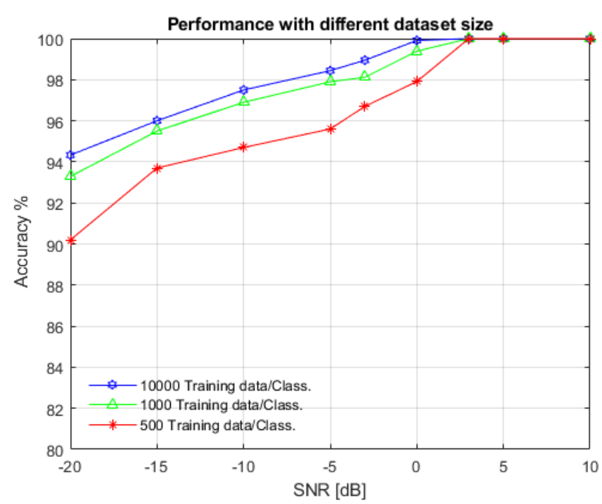

Fig. 6. Identification accuracy with different training set sizes (1000 features, AWGN channel).

A. The Impact of Dataset Size

In order to measure the performance, we consider the size of the training dataset. When we increase the number of signal training data for each class, the performance improves. As shown in Fig. 6, the increasing dataset size has noticeable impact on identification performance.

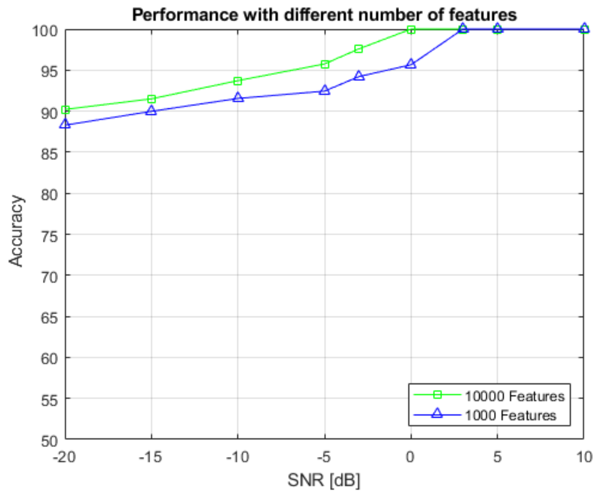

Fig. 7. The impact of the number of features (AWGN channel, 1000 training dataset).

B. The Impact of the Number of Features

Table I and Table II show the impact of the number of features on the identification performance. The accuracy of the model improves when the number of features increase. Compared with Table II (1000 features), in Table I (10000 features), the confusion matrix shows that we achieve better performance when we increase the number of features. We have improved accuracy performance 98.5% with 10,000 features extracted. In Table II, we have 1000 signals of different types such as 3G, 4G and 5G. We calculated the accuracy for 10000 features extracted at -10 dB in AWGN channel by how many signals are being captured in the right class.

TABLE I
THE CONFUSION MATRIX OF SNR = -10 dB FOR 10,000 FEATURES IN AWGN CHANNEL

Classes	**3G**	**LTE**	**5G**	**Accuracy**	**Overall Accuracy**
3G	985	11	4	98.50%	98.50%
LTE	7	981	12	98.1%	
5G	3	6	991	99.1%	

C. The Impact of Fading

Rayleigh and Rician fading channels are useful models of real-world phenomena in wireless communications. Thus, We simulate fading by using Communications Toolbox to implement fading channels. In Table III shows the confusion matrix when SNR = -10 dB for 1000 features extracted in a

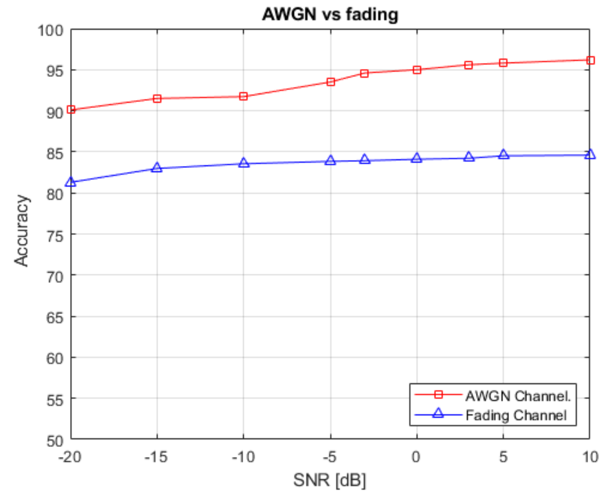

Fig. 8. Performance comparison of fading and AWGN channels (1000 features, 1000 training dataset).

TABLE II
THE CONFUSION MATRIX OF SNR = -10 dB FOR 1000 FEATURES IN AWGN CHANNEL

Classes	**3G**	**LTE**	**5G**	**Accuracy**	**Overall Accuracy**
3G	958	33	9	95.8%	97.40%
LTE	9	961	30	96.1%	
5G	8	15	977	97.7%	

Rayleigh fading channel. With 87.1%, 5G has reached better identification accuracy while 3G identification has dropped to 78.4%. In comparison, in an AWGN channel with the same number of features, 3G has reached 95.8% identification accuracy. Fig. 8 compares the identification accuracy under AWGN verses fading channel for various SNR levels.

TABLE III
THE CONFUSION MATRIX OF SNR = -10 dB FOR 1000 FEATURES IN RAYLEIGH FADING CHANNEL

Classes	**3G**	**LTE**	**5G**	**Accuracy**	**Overall Accuracy**
3G	748	147	69	78.4%	82.50%
LTE	69	821	110	82.1%	
5G	40	89	871	87.1%	

V. CONCLUSIONS

This paper investigates the identification of 5G signals by using deep learning. We examine cellular system environments with possible systems or signals such as UMTS (3G), LTE (4G), and 5G. A convolutional neural network (CNN) based on the LeNet-5 architecture is utilized. Various dataset sizes, the number of features extracted, and various SNR levels are considered in our investigation. Our experiments show that, by increasing the number of features, the 5G identification performance will be improved. Also, we examined the impact of the Rayleigh fading channel verses the AWGN channel.

REFERENCES

[1] J. Sills, “Maximum-likelihood Modulation Classification for PSK/QAM,” in *MILCOM 1999. IEEE Military Communications. Conference Proceedings (Cat. No. 99CH36341)*, vol. 1. IEEE, 1999, pp. 217–220.

[2] T. J. O’Shea, J. Corgan, and T. C. Clancy, “Convolutional Radio Modulation Recognition Networks,” in *International conference on engineering applications of neural networks*. Springer, 2016, pp. 213–226.

[3] W. Wei and J. M. Mendel, “Maximum-likelihood Classification for Digital Amplitude-Phase Modulations,” *IEEE transactions on Communications*, vol. 48, no. 2, pp. 189–193, 2000.

[4] N. Alyaoui, A. Kachouri, and M. Samet, “The Fourth Generation 3GPP LTE Identification for Cognitive Radio,” in *ICM 2011 Proceeding*. IEEE, 2011, pp. 1–5.

[5] B. Liu and K. Ho, “Identification of CDMA Signal and GSM Signal Using the Wavelet Transform,” in *42nd Midwest Symposium on Circuits and Systems (Cat. No. 99CH36356)*, vol. 2. IEEE, 1999, pp. 678–681.

[6] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. Springer Science & Business Media, 2007.

[7] N. E. West and T. O’Shea, “Deep Architectures for Modulation Recognition,” in *2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2017, pp. 1–6.

[8] Y. A. Eldemerdash, O. A. Dobre, O. Üreten, and T. Yensen, “Fast and Robust Identification of GSM and LTE Signals,” in *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2017, pp. 1–6.

[9] K. Alshathri, H. Xia, V. Lawrence, and Y.-D. Yao, “Cellular System Identification Using Deep Learning: GSM, UMTS and LTE,” in *2019 28th Wireless and Optical Communications Conference (WOCC)*. IEEE, 2019, pp. 1–4.

[10] H. Xia, K. Alshathri, V. B. Lawrence, Y.-D. Yao, A. Montalvo, M. Rauchwerk, and R. Cupo, “Cellular Signal Identification Using Convolutional Neural Networks: AWGN and Rayleigh Fading Channels,” in *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2019, pp. 1–5.

[11] K. Richardson, “UMTS Overview,” *Electronics & Communication Engineering Journal*, vol. 12, no. 3, pp. 93–100, 2000.

[12] S. Oulaourf, A. Haidine, A. Aqqal, and H. Ouahmane, “Review on Radio Resource Allocation Optimization in LTE/LTE-advanced Using Game Theory,” *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 9, no. 1, pp. 117–153, 2017.

[13] M. H. Dahri, M. H. Jamaluddin, M. I. Abbasi, and M. R. Kamarudin, “A Review of Wideband Reflectarray Antennas for 5G Communication Systems,” *IEEE Access*, vol. 5, pp. 17 803–17 815, 2017.