AMC-IoT: Automatic Modulation Classification Using Efficient Convolutional Neural Networks for Low Powered IoT Devices

Muhammad Usman and Jeong-A Lee
Department of Computer Engineering, Chosun University, Republic of Korea
Email: usman@chousn.kr, jalee@chosun.ac.kr

***Abstract*—Automatic modulation classification (AMC) is used to identify the modulation for the received signal. IoT devices use modern communication methods which are based on multiple input multiple output (MIMO) in which the signals are received from various sources. The identification of modulation is vital. Feature based AMC methods combined with deep learning techniques has the potential to meet the latency requirement in the IoT applications. An efficient convolutional neural network based on depthwise separable convolution has been proposed to classify the modulation of the received signals. The proposed architecture has 58% less parameters than the conventional convolutional architecture and the performance is comparable.**

***Index Terms*—automatic modulation classification, convolutional neural networks, depthwise separable, IoT**

I. Introduction

THE advent of internet of things (IoT) and evolution in the mobile networks is driven by the need to satisfy the consumer's demand of enhanced performance, high speeds, seamless links elasticity, and portability in the telecommunication network. The 5th Generation technology abbreviated as 5G is the next generation telecommunication standard that assures to meet the ever growing needs of the communication applications.

IoT is a potential candidate to utilize the communication network resources and it is estimated that more than 50 billion devices are now connected to the internet [1]. Considering the current 5G communication, the receiver is expected to receive the signals from multiple directions. Therefore, the well-known problem of multi-path fading is inevitable, and would lead to the difficulty in identification of signals. Multiple-input multiple output (MIMO) shall be actively utilized in the 5G communication which means the signals would be received from several sources. Furthermore, the software defined radio (SDR) and the cognitive radio (CR) based technologies are getting immense popularity in which the devices are expected to adopt the neighboring conditions and adjust their transmission parameters, modulation schemes etc. It is therefore of utmost importance to be able to classify the modulation type at the receiver's end without needing the prior knowledge about the transmitter's parameters. Several researchers have proposed the method of automatically classifying the modulation schemes which is termed as automatic modulation classification (AMC) [2]–[4], and presented that modulation recognition can extract the digital baseband information based on very low or no prior information about the device type and transmission schemes. AMC is regarded as an intermediate step between the signal detection and demodulation at the receiving end. The AMC problem can be divided into to categories: liklihood based (LB) and feature based (FB). In the LB methods the ratio of the liklihood of the signals are measured and the decision is made on the basis of the comparison of ratios to a threshold. Particularly, various liklihood solutions with unknown modulation parameters on the basis of power spectral density (PSD) of the received signal are tested. It has been observed that the LB methods yields good results but at the cost of computational complexity. Most of the studies have focused on the accurate prediction of the modulation schemes using LB method did not consider the computation complexity [5], [6]. The FB methods, developed as a suboptimal classifier for the practical implementation, relies on the extraction of useful features from the received signal and then classify them using a classifier. The features can be of instantaneous time, transform based, statistical or based on constellation shape. Extensive research have been carried out in both of these methods and it is found that the LB method provides an optimal solution with good accuracy, but that demands high computation and prior knowledge about the signal. On the contrary, the FB method presents a sub-optimal solution but has relatively low latency since it does not require any prior knowledge. Conventional FB methods rely on the expert's knowledge and may perform better in certain situations if not all. Furthermore, to meet the latency standards of 5G as defined in [7], artificial intelligence, specially deep learning (DL) has been effectively utilized for the purpose of modulation classification [8], [9] . A trivial issue in the utilization of ML methods is the feature engineering, for which expert's experience is required. To obviate feature engineering in the FB methods, several machine learning (ML) based classifiers, particularly deep learning methods have been thoroughly studied as they present the tendency to learn complex functions with more significance compared to their shallow counterparts. With the rapid advancement in the DL technologies, many methods have been proposed that can self-learn the features, furthermore, DL is favoured because it requires large datasets which can be easily obtained in the communication systems. One of the major concern using DL could be its complexity, as DL involves the phases of training and testing of the model. However, an efficient convolutional network architecture which performs depthwise convolution has been utilized in many applications. Due the depthwise convolutions, the model size is significantly reduced without affecting the accuracy. The model has less number of parameters compared to the conventional convolution, thus is suitable to match the design requirements of the small size devices in the IoT environment. In this study, the efficacy of the depthwise convolution method is compared with the conventional convolution neural network.

The rest of the paper is organized as follows: Section II presents the review of automatic modulation classification based on deep learning based methods, the details of the proposed method is discussed in Section III. Results are presented in Section IV and the paper is concluded in Section V.

II. Background

The deep learning is the subset of artificial neural networks that tries to develop the relationship between the input and the target by mimicking the human brain's mechanism. It is a layered architecture consisting of an input layer, a hidden layer/layers in the middle and an output layer. The number of hidden layers governs the depth of the architecture and theoretically any architecture with more than one hidden layer is termed as a deep network. Deep learning algorithms have been used in variety of classification problems ranging from computer vision [10], bio-informatics [11]–[13], natural language processing [14], speech recognition [15], and signal processing [16], [17] etc. It is a branch of machine learning that deals with large data and treats it as an input which is carried into multilayered architecture containing hidden nodes. The final layer of the architecture is called the output layer which provides the desired target classes. By virtue of these hidden layered architecture, the DL-based methods outperform classical machine learning techniques. Typically the deep neural network learn the features from the data autonomously, therefore manual feature representation to the machine is not required. This makes it superior to the conventional ML approaches that require detection and extraction of effective features which is an arduous and time consuming task. Deep neural networks assemble the simple features to develop a complex feature by themselves through multiple non-linear transformations [18]. Several research has been conducted to address the modulation classification problem which has been handled through conventional signal processing in the past [19], [20].

With the success of ML and DL many researchers have opted these methods propose automatic modulation classification methods. In [21], an artificial neural network (ANN) based method coupled with the genetic algorithm (GA) has been proposed, and utilized resilient backpropagation (RPROP) for the first time to recognize digital modulation. GA was employed to extract the six discriminating features and the proposed algorithm achieved 99% recognition performance. The DL algorithm have been utilized by [22], [23] and [24], in the complex-valued temporal signal domain. Deep autoencoders (DAE) that are formed by two symmetrical deep belief networks are the unsupervised learning algorithm that generates an output similar to its input but reconstructs it using fewer dimensions. The dimension reduction is performed in the latent space which is composed of few neurons. This latent space is regularized to be sparse, which helps it to learn the most important features to reproduce the input. The discriminative features learnt by the latent space can be utilized to perform the classification task. The autoencoder is composed of three sections: encoder, decoder and latent space. Several configurations of autoencoder can be found in literature that have been tuned to perform the required tasks. Most common of them are the regularized autoencoders in which the encoding function is not linear and allows the AE to show sparsity [25], [26]. One of the most successful DL architectures, convolutional neural network (CNN), has been widely adopted to address the modulation recognition problem. For instance, in [27], CNN has been trained to classify the modulation by converting the modulated signals in to images. Images including the constellation diagrams have been used to train a famous CNN architecture named as AlexNet [28]. Some advancements towards the representation of modulated data into gridlike topology for training of CNN have been presented in [29]. In particular, methods to convert complex signals to images have been presented along with analysis of impacts of selection of conventional parameters, and finally a CNN model is trained to show the effects of the proposed study. A robust method for the modulation classification is proposed in which 15 modulation types are used to train a CNN without needing to extract the features from the signals [30]. The use of the robust features provided good accuracy results. In [31], three architectures including Convolutional Long Short-term Deep Neural Network (CLDNN), a Long Short-Term Memory neural network (LSTM), and a deep Residual Network (ResNet) have been studied with the aim of maintaining a good performance in while adopting strategies to reduce the amount of time required in training the networks. The ability to classify the modulation in short amount of time using a comparatively less complex architecture is of great interest, since the devices in the IoT, unmanned air vehicles (UAVs) and other technologies in the 5G have limited computation and power resources. A significant amount of research has been carried out by Shea et.al who have presented CNN based solutions to classify the modulation and presented open access dataset for training of the neural networks which have been adopted in several of the above mentioned literature. We also utilize the datasets from [22], and train a convolutional neural network that performs depthwise convolution and compare its results with that of a conventional neural network architecture. Hence we explore a less complex automatic modulation classification method to assist the low powered devices employed in the wireless communication. This development would be beneficial for the deployment of deep learning based solution in the wireless communication.

A. Convolutional Neural Network

CNNs are regularized form of multilayer perceptrons, inspired by the biological process of connection of neurons. They are effectively utilized in various classification tasks because require little preprocessing in comparison with other classification algorithms. A convolutional neural network condenses the input to the important features that helps in distinguishing the input. Both 2D and 3D convolution architectures are commonly employed according to the requirement of the user. It is essentially composed of four layers.

- **Convolution Layer**: The convolution layer is responsible to perform the convolution operation between the filter and the input map. After the filter is passed over the underlying input is changed in a way that certain features of the input are emphasized.
- **Pooling Layer**: Some applications require compression of the output of the convolution layer. Pooling is a way of summarizing the features by downsampling the feature map. This adds robustness for the features to the position changes. Most commonly used pooling methods are average pooling and the max pooling.
- **Fully connected layer**: The classification task is performed using the dense layers in which each neuron is usually connected to each neuron in the succeeding layer with some weights and activations.
- **Output Layer**: The output layer is the last layer which has a certain activation function to calculate the probability response.

A typical CNN architecture is shown in Fig. 1.

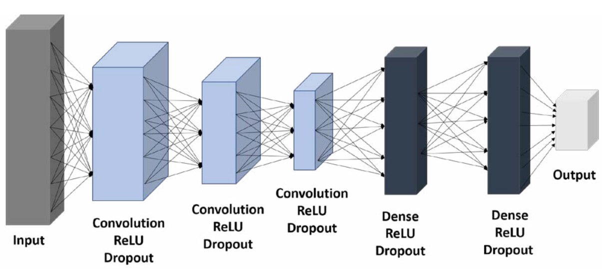

Fig. 1. Convolutional Neural Network

B. Depthwise Separable Convolution

The concept of depthwise separable convolution was first proposed in [32] and was used by the Inception models for complexity reduction in the first few layers [33]. Other methods including shrinking of the pre-trained network, compression on the basis of product quantization [34], hashing and pruning have been adopted to obtain a small network.

1) Architecture: Depthwise separable convolution is a depthwise convolution followed by a pointwise convolution as shown in Fig. 2.

- **Depthwise Convolution** uses single filter for each input channel i.e, if there are 3 input channels, there will be 3 DkxDk spatial convolutions.
- **Pointwise Convolution** is used to change the dimension of the output of the depthwise convolution and makes is similar to the output of the convention convolutional neural network output layer. It does so by applying a 1x1 convolution having a depth equal to the number of channels of the depthwise convolution layer.

In standard convolution the filtration and combination is performed in a single step where as in depthwise separable convolutions, the filtration and combination task are separated in two layers as shown in Fig. 3.

The general form of 2D convolution is shown in (1)

$$\begin{aligned} z[m,n] &= x[m,n] * k[m,n] \\ &= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} x[i,j] \cdot k[m-i, n-j] \end{aligned} \tag{1}$$

which is equivalent to (2), since the convolution operation is commutative

$$\begin{aligned} z[m,n] &= x[m,n] * k[m,n] \\ &= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} k[i,j] \cdot x[m-i, n-j] \end{aligned} \tag{2}$$

For separable convolution, the kernel is separated into [Mx1] and [1xN] as shown in (3)

$$k[m,n] = k_1[m] \cdot k_2[n] \tag{3}$$

Substituting $k[m,n]$ in (2)

$$\begin{aligned} z[m,n] &= x[m,n] * k[m,n] \\ &= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} k[i,j] \cdot x[m-i, n-j] \\ &= \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} k_1[i] \cdot k_2[j] \cdot x[m-i, n-j] \\ &= \sum_{-\infty}^{\infty} k_2[j] \left(\sum_{-\infty}^{\infty} k_1[i] \cdot x[m-i, n-j] \right) \end{aligned} \tag{4}$$

The 1D convulotion is given as (5)

$$\begin{aligned} z[n] &= x[n] * k[n] \\ &= \sum_{p=-\infty}^{\infty} x[p] \cdot h[n-p] \end{aligned} \tag{5}$$

In separable convolution, the input is convolved twice, once in the horizontal direction and once in the vertical direction. Due to associative property of convolution, the order does not matter, therefore, any of the following convolutions shown in (6) can be performed first followed by the other.

$$\begin{aligned} z[m,n] &= (h_1[m] \cdot h_2[n] * x[m,n]) \\ &= h_2[n] * (h_1[m] * x[m,n]) \\ &= h_1[m] * (h_2[n] * x[m,n]) \end{aligned} \tag{6}$$

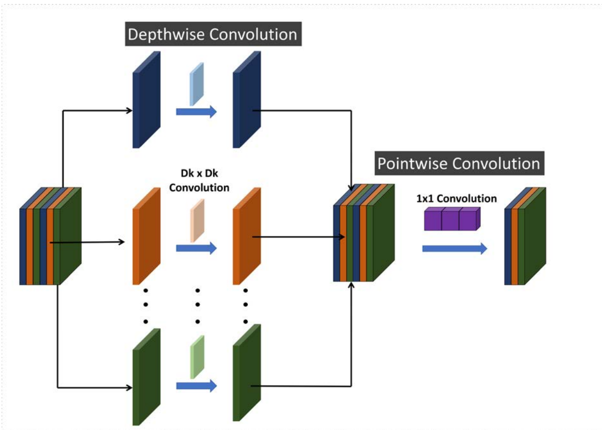

Fig. 2. Depthwise Separable Convolution

A difference between the standard and the depthwise separable convolution is depicted in Fig. 3.

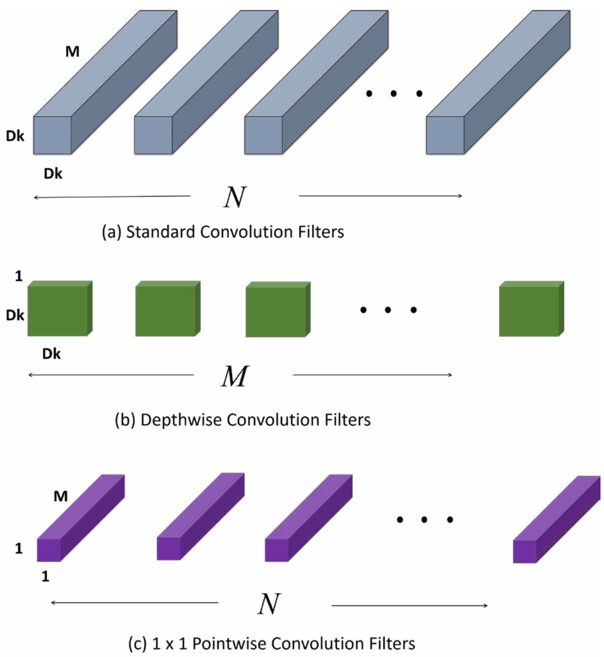

Fig. 3. Standard convolution in (a) replaced by two layers of depthwise convolution in (b) and pointwise convolution in (c)

III. Proposed Deep Modulation Classifier

In this study we use depthwise separable convolution neural network to minimize the number of parameters and consequently the number of multiplication operations during the convolution process.

- A lightweight modulation classification algorithm based on convolutional neural network has been proposed for two different datasets.
- Three models have been evaluated:
  – A conventional convolutional neural network using the Keras *Conv2D* API is trained on each dataset and performance is measured.
  – Depthwise spatial convolution using the Keras *DepthwiseConv2D*, which is the first step of the depthwise separable convolution has been utilized for the classification of each dataset.
  – Depthwise separable convolution that performs depthwise spatial convolution followed by the pointwise convolution has been used using Keras *SeparableConv2D* API to classify the two dataset.
  – The performance of the models and the corresponding number of parameters have been presented.

The implementation is done on Python using Theano/Keras with GPU RTX2070, 2560 CUDA cores, and 8GB GDDR6 VRAM on top of Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz with 64 GB DDR3 RAM.

A. Dataset

The dataset used in this work is obtained from GNU Radio [35]. The datasets are named as RML2016a and RML2016c respectively. The data is deterministically generated to match the real system's data, furthermore, the parameters of pulse shaping, modulation and carried data are made identical. The dataset named RML2016a is composed of two data sources. Continuous signals including voice and the publicly available Serial Episode #1 are used as an input for the analog modulations. For digital communication entire Gutenberg works of Shakespeare in ASCII is used. To make the digital modulated data equiprobable, block normalizer has been utilized for data whitening. The synthetic signals are then passed through several channels to get the effect of unknown scaling and translation. The GNU radio channel model block generates the final data set, of which the time series signals are sliced into the training and test dataset by 128 samples rectangular window. The dataset RML2016c has been constructed by collecting the widely used modulations. Both discrete and continuous alphabets for digital and analog modulation data is modulated over modem and is exposed to channel effects using GNU Radio. The sampling is performed and 128 samples with a shift of 64 samples are extracted. The dataset is finally labeled into 11 classes corresponding to the modulation scheme.

IV. Results and Evaluation

We utilize the two datasets mentioned in section III-A to train the three neural networks and compare their performances. The weights of the convolutional layers in all the networks are initialized using ”Glorot” initializer, and the weights in the fully connected layers of the networks are initialized using ”He” initializer. The weights are learned through Adam optimizer by keeping a learning rate of 0.01 and categorical cross-entropy loss function is minimized. The networks are trained for 100 epochs with a batch size of 1024 samples. The total number of parameters were 921611, 596491 and 385307 for Conv2D, Depthwise and Separable convolutional

Architecture	Parameters	Reduction	Accuracy	
			RML2016a	RML2016c
Conv2D	921611	–	71.30%	83.40%
Depthwise	596491	35.2%	69.90%	92.10%
Separable	385307	58.2%	71.25%	83.03%

TABLE I

COMPARISON OF ACCURACY ACHIEVED ON 18DB SNR OF BOTH DATESETS ON DIFFERECT ARCHITECTURE CONFIGURATIONS

neural networks respectively. The accuracy achieved by these architectures is shown in Table I and graphically depicted in Fig. 4 and Fig. 5 for dataset RML2016a and RML2016c respectively.

The accuracy curves for different SNR values have been plotted for both datasets trained on the convolutional neural network architectures. For RML2016a dataset, the performances of the separable convolutional network is comparable to the conventional neural network architecture as shown in Fig. 4. For RML2016c dataset, the depthwise and separable CNN configurations outperforms the conventional CNN approach as depicted in Fig. 5, suggesting efficient performance of the network with reduced parameters.

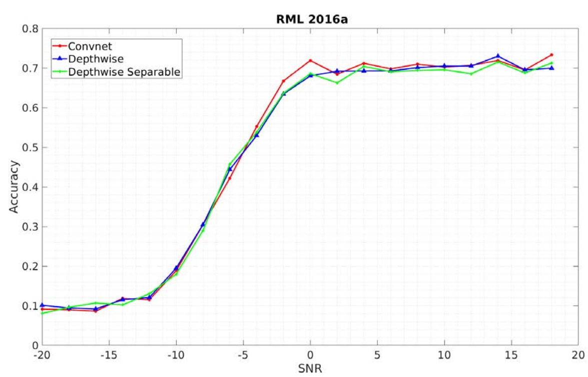

Fig. 4. Classification Accuracy of the convolutional neural network architectures on RML2016a dataset

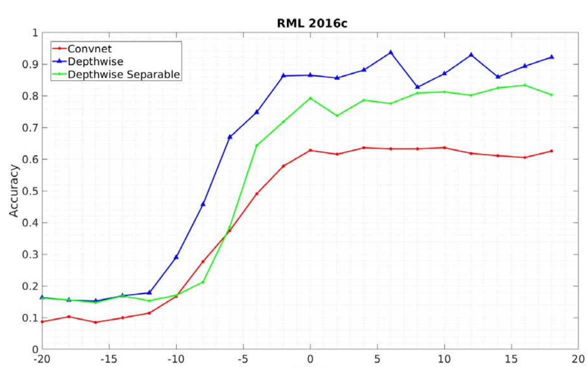

Fig. 5. Classification Accuracy of the convolutional neural network architectures on RML2016c dataset

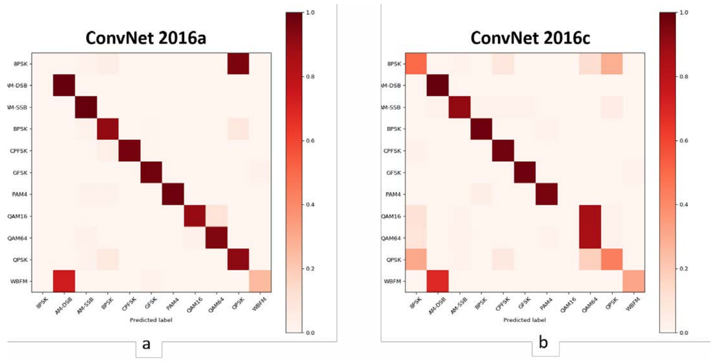

Fig. 6. Confusion matrix depicting the ability of the conventional CNN to classify several modulation schemes at SNR=18 on RML2016a dataset (a) and RML2016c dataset (b)

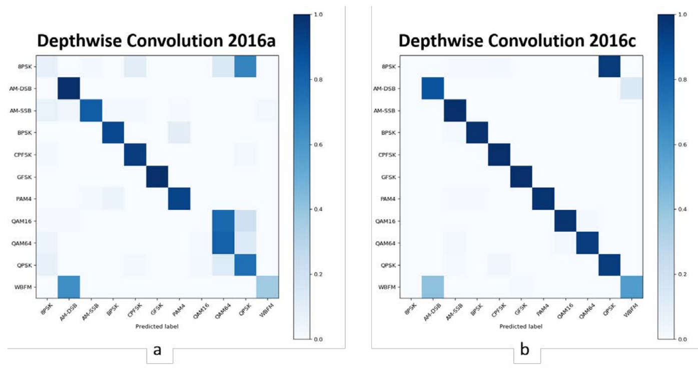

Fig. 7. Confusion matrix depicting the ability of the deptwise CNN to classify several modulation schemes at SNR=18 on RML2016a dataset (a) and RML2016c dataset (b)

Fig. 6, 7 and 8 depicts the confusion matrices of the CNN, depthwise CNN and separable CNN respectively on samples from both datasets with SNR $18dB$.

V. CONCLUSION

Depthwise and separable convolutional neural network architectures have been presented to classify the modulation schemes. The low powered IoT devices does not have the capability to process large number of parameters as in the conventional neural networks. The proposed methods achieves similar and in some cases better performance compared to the conventional convolution approaches with significantly less number of parameters. Such architectures are highly favorable for implementation in the devices that are constrained by power and area.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1I1A3063857) and in part by Korea Institute of Energy Technology Evaluation and Planning, and the Ministry of Trade, Industry and Energy of the Republic of Korea (No. 20184010201650).

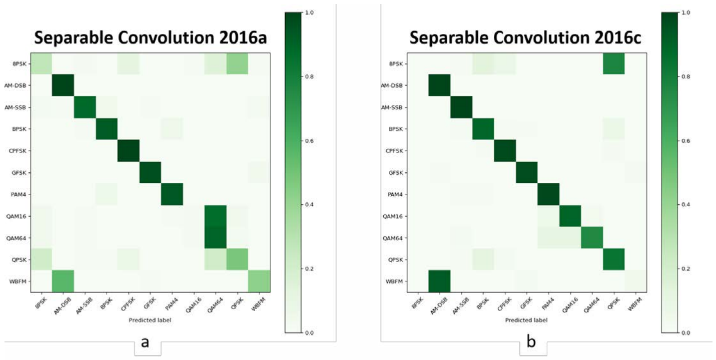

Fig. 8. Confusion matrix depicting the ability of the separable CNN to classify several modulation schemes at SNR=18 on RML2016a dataset (a) and RML2016c dataset (b)

References

[1] M. Usman, I. Ahmed, M. I. Aslam, S. Khan, and U. A. Shah, "Sit: A lightweight encryption algorithm for secure internet of things," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 1, 2017. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2017.080151

[2] W. Chen, Z. Xie, L. Ma, J. Liu, and X. Liang, "A faster maximum-likelihood modulation classification in flat fading non-gaussian channels," *IEEE Communications Letters*, vol. 23, no. 3, pp. 454–457, 2019.

[3] J. Nie, Y. Zhang, Z. He, S. Chen, S. Gong, and W. Zhang, "Deep hierarchical network for automatic modulation classification," *IEEE Access*, vol. 7, pp. 94 604–94 613, 2019.

[4] J. Jagannath, N. Polosky, D. O'Connor, L. N. Theagarajan, B. Sheaffer, S. Foulke, and P. K. Varshney, "Artificial neural network based automatic modulation classification over a software defined radio testbed," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.

[5] T. Yucek and H. Arslan, "A novel sub-optimum maximum-likelihood modulation classification algorithm for adaptive ofdm systems," in *2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No. 04TH8733)*, vol. 2. IEEE, 2004, pp. 739–744.

[6] W. Wei and J. M. Mendel, "A new maximum-likelihood method for modulation classification," in *Conference Record of the Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*, vol. 2. IEEE, 1995, pp. 1132–1136.

[7] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5g: Ran, core network and caching solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3098–3130, 2018.

[8] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4074–4077, 2019.

[9] Y. Wang, J. Yang, M. Liu, and G. Gui, "Lightamc: Lightweight automatic modulation classification via deep learning and compressive sensing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3491–3495, 2020.

[10] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: A survey," *IEEE Access*, vol. 6, pp. 14 410–14 430, 2018.

[11] M. Usman, S. Khan, and J.-A. Lee, "Afp-lse: Antifreeze proteins prediction using latent space encoding of composition of k-spaced amino acid pairs," *Scientific Reports*, vol. 10, no. 1, pp. 1–13, 2020.

[12] M. Usman and J. A. Lee, "Afp-cksaap: Prediction of antifreeze proteins using composition of k-spaced amino acid pairs with deep neural network," in *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 2019, pp. 38–43.

[13] S. M. Ibrahim, M. S. Ibrahim, M. Usman, I. Naseem, and M. Moinuddin, "A study on heart segmentation using deep learning algorithm for mri scans," in *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*. IEEE, 2019, pp. 1–5.

[14] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *ieee Computational intelligenCe magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[15] D. Amodei, S. Ananthanarayanan, R. Anubhai, J. Bai, E. Battenberg, C. Case, J. Casper, B. Catanzaro, Q. Cheng, G. Chen *et al.*, "Deep speech 2: End-to-end speech recognition in english and mandarin," in *International conference on machine learning*, 2016, pp. 173–182.

[16] M. Usman, M. S. Ibrahim, J. Ahmed, S. S. Hussain, and M. Moinuddin, "Quantum calculus-based volterra lms for nonlinear channel estimation," in *2019 Second International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT)*, 2019, pp. 1–4.

[17] S. Khan, M. Usman, I. Naseem, R. Togneri, and M. Bennamoun, "Vp-flms: a novel variable power fractional lms algorithm," in *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2017, pp. 290–295.

[18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[19] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Transactions on communications*, vol. 48, no. 3, pp. 416–429, 2000.

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

[21] M. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm," *Signal Processing*, vol. 84, no. 2, pp. 351–365, 2004.

[22] 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.

[23] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[24] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.

[25] Q. Ji, Y. Sun, J. Gao, Y. Hu, and B. Yin, "Nonlinear subspace clustering via adaptive graph regularized autoencoder," *IEEE Access*, vol. 7, pp. 74 122–74 133, 2019.

[26] W. Xu, S. Keshmiri, and G. Wang, "Adversarially approximated autoencoder for image generation and manipulation," *IEEE Transactions on Multimedia*, vol. 21, no. 9, pp. 2387–2396, 2019.

[27] S. Peng, H. Jiang, H. Wang, H. Alwageed, and Y.-D. Yao, "Modulation classification using convolutional neural network based deep learning model," in *2017 26th Wireless and Optical Communication Conference (WOCC)*. IEEE, 2017, pp. 1–5.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[29] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y.-D. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 3, pp. 718–727, 2018.

[30] S. Zhou, Z. Yin, Z. Wu, Y. Chen, N. Zhao, and Z. Yang, "A robust modulation classification method using convolutional neural networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2019, no. 1, p. 21, 2019.

[31] S. Ramjee, S. Ju, D. Yang, X. Liu, A. E. Gamal, and Y. C. Eldar, "Fast deep learning for automatic modulation classification," *arXiv preprint arXiv:1901.05850*, 2019.

[32] L. Sifre and S. Mallat, "Rigid-motion scattering for image classification," *Ph. D. thesis*, 2014.

[33] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[34] J. Wu, C. Leng, Y. Wang, Q. Hu, and J. Cheng, "Quantized convolutional neural networks for mobile devices," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4820–4828.

[35] T. J. O'shea and N. West, "Radio machine learning dataset generation with gnu radio," in *Proceedings of the GNU Radio Conference*, vol. 1, no. 1, 2016.