Autonomous Modulation Classification using Single Inception Module based Convolutional Neural Network

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

20-12-2020 / 12-01-2021

CITATION

Mumtaz, Zeeshan; Khurram, Muhammad; Adnan, Muhammad; Fazil, Adnan (2020): Autonomous Modulation Classification using Single Inception Module based Convolutional Neural Network. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.13455719.v3>

DOI

[10.36227/techrxiv.13455719.v3](https://doi.org/10.36227/techrxiv.13455719.v3)

Autonomous Modulation Classification using Single Inception Module based Convolutional Neural Network

Muhammad Zeeshan Mumtaz^{*} , Muhammad Khurram[†], Muhammad Adnan[‡] and Adnan Fazil[§]

Institute of Avionics and Aeronautics,
Air University, Islamabad

Email: ^{*}zeeshan2708@hotmail.com, [†]khurram.jadoon@mail.au.edu.pk, [‡]madnan@mail.au.edu.pk, [§]adnan630@hotmail.com

Abstract—Modulation recognition plays a crucial role in *non-cooperative communication* in which receivers have no prior information regarding transmitter modulation scheme. This paper presents a novel convolutional neural network with single Stem and Inception module for autonomous modulation classification from raw I/Q received channels. This addresses the computationally intensive problem of conversion of I/Q channels to constellation image processing. The proposed system is capable of classifying 11 standard modulation schemes with both 2D and 3D input array configurations for varying SNR conditions. The performance of proposed system has been evaluated for a real-time communication system simulated with Rician fading channel and AWGN noise model, providing realistic distortion effects. The proposed CNN design achieves an average accuracy of 90% at 10 dB and 99% at 20 dB SNR with reduced network learnables and lower training and testing time, which makes it computationally efficient. These attributes make Inception module based CNN a viable solution for modulation classification in practical low cost and portable yet reliable communication systems.

I. INTRODUCTION

The evolution of wireless communication is latent in the development of efficient signal modulation as it enables long-haul propagation of base-band message. Various types of modulation schemes have been developed to cope up with the contrasting real world scenarios [1]. Modulation schemes are segregated mainly into two classes, analog and digital modulation.

The knowledge of modulation scheme is crucial for flawless recovery of intended message at receiver domain. Communication systems are divided into two types based on the information of modulation scheme. *Cooperative communication* requires prior information available at receiver regarding the modulation technique, either predetermined before communication session or transmitted via separate channel. Whereas, *Non-cooperative communication* does not necessarily needs any beforehand knowledge of transmission modulation technique at receiver [2]. Generally, these systems have a modulation recognition block intermediate between signal reception portion and signal demodulation as shown in Fig. 1. Non-cooperative communication systems have found their application in cognitive radio links, radio spectrum management and signal intelligence (SIGINT).

Classification algorithms can be split into two groups, Likelihood based (LB) algorithms and Feature based (FB)

Fig. 1. Generic Non Cooperation Communication.

algorithms [3]. Likelihood based algorithms focus on hypothesis testing based on the probability density function of the received signal which provides sufficient information regarding recognition [4]. Owing to their analytical nature, LB methods provide optimal classification on the basis of signal and channel estimates. However this optimization comes at the cost of intense computational complexity, which is not desirable in rapid modern communication systems. Feature based algorithms involve the pattern recognition from instantaneous features of received signal and aim for sub-optimal classification by utilizing classifiers with low processing requirements. Traditional FB methods are based on two components, feature extraction and classifier. Feature extraction obtain patterns from the concerned signal by using time domain waveforms like cyclic spectrum, wavelet coefficients and high order cumulant [5]. Then different classifiers such as shallow neural network, K-nearest neighbourhood and decision tree for support vector machine (SVM) are used to identify modulation modes.

Deep Learning has emerged as an innovative tool for effective modulation classification. Several variants of neural networks have used to serve the purpose in communication domain such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), residual networks (ResNet) and densely connected networks (DenseNet). However, most of DL algorithms developed for modulation recognition are based on image classification capabilities of deep learning networks which converts classical communication problem to image processing dominion. Typically, constellation diagrams are produced from the raw inphase and quadrature phase (I/Q)

components of received signal and then, neural networks specialized in image feature extraction are used to classify modulation based on their standard constellation characteristics. These algorithms have good accuracy for nominal SNR values, however they consume exhaustive computational resources as image based DL methods deal with training of very large number of learnable weights and biases. In [10], AlexNet has been used to analyze 227×227 pixel constellation diagrams with 60 million parameters with over 600 million connections using 10 million image dataset for 1000 received samples. InceptionResNetv2 with transfer adaptation has been implemented in [2] for constellation image layer ($224 \times 224 \times 3$) with 55.8M network parameters. In [11], two options have been considered namely, AlexNet (60M parameters) and GoogleNet (7M parameters), and their performance is compared. All of the above discussed image based classification algorithms achieve accuracy beyond 95% but involve intense training hyperparameters which require training time in order of tens of hours. There is an additional drawback in constellation based classification that the complex plane is infinite while the image scope has to be limited in space domain. Therefore, selection of size of centric region is bound by a trade off. Too small size can lead to missing sample entries while too large size can burden network resources.

This research work is motivated by the efforts in classification based on sampled RF time-series data, the raw I/Q channels. The incentive in this approach is that time series data can be analyzed frame by frame with very low input dimensions. In [1], a sequential CNN with stacking of 5 convolution layers has been implemented for classification for 7 modulation modes. In [6], the authors have implemented CNN for 10 modulation classes with just 2 convolution layers. A sequential network based on 2 LSTM layers of 128 units each has been proposed in [7], which is based on time domain amplitude and phase information of input signal instead of I/Q channel conversion. As a comparison to LSTM performance, in [8] researchers have implemented 2 layer-gated recurrent network (GRU) which is derived from basic RNN structure, directly applied on I/Q samples. An important contribution has been made in [9] by utilizing Deep neural networks (DNNs) for time domain modulated signal recognition. Convolutional Long Short-term Deep Neural Network (CLDNN) has been developed alongwith Residual Network (ResNet) and Densely Connected Network (DenseNet), which will be compared with our proposed Inception module CNN in subsequent sections. The authors in [19] have applied time and discrete orthonormal stockwell transform (DOST) domain channeling for pre-processing of raw RF I/Q data by converting it into pseudo image transforms. Then a series of 5 convolutional layers based network is trained and tested for different modulation schemes. Their performance is discussed in detail in the results section of this document. The networks used for raw I/Q inputs have been generally sequential based CNNs which focus on deeper and deeper networks with exponentially increasing parameters. Additionally, classification procedure considers only individual inphase and quadrature phase vectors and does

not cater for possible correlation between I/Q channels.

The principal contribution of this research is introduction of parallel convolutional layer structure in form of Inception module based convolutional neural network as it can address the issue of overfitting for lower dimensional datasets. Additionally, the proposed design investigates the potential association between I/Q channels alongwith individual feature analysis at same network stage.

II. PROBLEM FORMULATION

A. Signal Model

Consider a transmitter with constant transmission power, modulating baseband message with 11 modulation schemes. The signal is subjected to Rician fading channel $\mathbb{H}$ and AWGN noise channel $\mathbb{N}$ for realistic representation of distortion effects. The signal received by modulation recognition system can be modelled by the following equation.

$$\mathbf{y}_i(\mathbf{n}) = A e^{j(2\pi f_0 n T + \theta_n)} \left(\mathbb{H} \sum_{k=-\infty}^{\infty} x(k) \right) + \mathbb{N} \quad (1)$$

A is the amplitude factor, θ_n is the phase component for I/Q representation while f_0 is the carrier center frequency, which for analog modulation schemes is selected to be 100 MHz and for digital modulation techniques to be 902 MHz. Rician fading channel $\mathbb{H}$ provides good approximation for channel effects considering line-of-sight component. However, in this research, we have applied zero-forcing equalizer to mitigate the channel effects by perfect knowledge of channel state information. The equalized signal can be approximated as

$$\tilde{\mathbf{y}}_i(\mathbf{n}) = \tilde{A} e^{j(2\pi f_0 n T + \tilde{\theta}_n)} \sum_{k=-\infty}^{\infty} x(k) + \tilde{\mathbb{N}} \quad (2)$$

Now we define the modulation scheme sample space $\mathbb{L}$ which represents the modulation identity index l_i as

$$\mathbb{L} = [l_1, l_2, \dots, l_i, \dots, l_{M-1}, l_M] \quad (3)$$

where M is total classes of modulation schemes out of which we have to identify the desired technique. For the proposed system, sample space $\mathbb{L}$ consists of 11 modulations schemes ($M = 11$) divided into following 2 classes:

1) Digital Modulation Schemes:

- Binary phase shift keying (BPSK)
- Quadrature phase shift keying (QPSK)
- 8-ary phase shift keying (8-PSK)
- 16-ary quadrature amplitude modulation (16-QAM)
- 64-ary quadrature amplitude modulation (64-QAM)
- 4-ary pulse amplitude modulation (PAM4)
- Gaussian frequency shift keying (GFSK)
- Continuous phase frequency shift keying (CPFSK)

2) Analog Modulation Schemes:

- Broadcast FM (B-FM)
- Double sideband amplitude modulation (DSB-AM)
- Single sideband amplitude modulation (SSB-AM)

B. Probabilistic Classification

Consider the transmission modulation scheme l and the classified modulation for a particular frame of size N samples by classifier at receiver is l' , then the probability of true classification is

$$P|_{true,N} = \frac{1}{M} \sum_{i=1}^M P_N(l'_i | l_i, l'_i = l_i) \quad (4)$$

This relates to the fact that true classification occurs for situations where identified classification is identical to the transmission classification. Overall true probability of classifier is dependent upon individual conditional probabilities of each modulation technique from the sample space $\mathbb{L}$. These conditional probabilities are linked to conditional probabilities for each sample in a frame to the correct classification.

$$P_N(l'_i | l_i) = \prod_{n=1}^N P(l'_{i,x_n} | l_{i,x_n}, l'_{i,x_n} = l_{i,x_n}) \quad (5)$$

For any given test signal y with probability density function $p_Y(y)$, we anticipate the expected value of probability of true classification is obtained by combination of equation (4) and equation (5).

$$P|_{true,y} = \mathbb{E}_y[P_N] \quad (6)$$

$$= \int_y P|_{true,N} p_Y(y) dy \quad (7)$$

$$= \frac{1}{M} \int_y \sum_{i=1}^M \prod_{n=1}^N P(l'_{i,x_n} | l_{i,x_n}) p_Y(y) dy \quad (8)$$

False classification leads to error in demodulation phase as the identified modulation is not identical to transmission modulation. The probability of false classification is represented by

$$P|_{false,N} = \frac{1}{M} \sum_{i=1}^M P_N(l'_i | l_i, l'_i \neq l_i) \quad (9)$$

For given test signal y , the expected value of false classification is derived from equation (9) as

$$P|_{false,y} = \mathbb{E}_y[\bar{P}|_N] \quad (10)$$

$$= \int_y P|_{false,N} p_Y(y) dy \quad (11)$$

$$= \frac{1}{M} \int_y \sum_{j \neq i} \prod_{q \neq n} P(l'_{j,x_q} | l_{i,x_n}) p_Y(y) dy \quad (12)$$

Indepth analysis of equation (8) and equation (12) relate to the fact that it is extremely complex to formulate a closed form expression for expected values of true and false classification for a test signal, which makes analytical solution for Likelihood based classification challenging. Therefore, DL becomes an intuitive choice for solving this problem due to its hidden layer architecture with no need for manual feature selection.

III. NETWORK ARCHITECTURE

Convolutional neural networks (CNNs) derive etymology from convolutional layer which accepts feature maps from preceding layer and performs convolution operation (either 2 or 3 dimensional) over input feature maps along with a set of learnable filters. The process develops a new feature map which is fed into the next layer, making the architecture modular in nature. The mathematical representation of this iterative process is

$$X_{n+1} = f(W_n * X_n) + b_n \quad (13)$$

For n^{th} layer, X_n is the input feature map, X_{n+1} is the output feature map, W_n represents convolutional filters and b_n indicate bias values. Whereas f is the activation function which decides on the basis of objective function.

Numerous neural networks have been implemented for addressing communication enigmas like CNNs, RNNs, ResNets and DenseNets. However, the inspiration for Inception module in this research owes to under-utilization of parallel functionality of convolution processes in communication realms. Inception CNN versions have been successfully used to diversify feature base in image processing. We are emulating this capability in classical communication problem of modulation recognition.

A. Proposed CNN Design

The proposed convolutional neural network has two modules, *Stem* and *Inception* modules as shown in Fig. 2. This network imitates the renowned Inceptionv1 [12] and Inception v3 networks [13] which are generally used for image classification. Traditionally, these networks have been applied to constellation images for modulation classification as discussed earlier. However, we have modified the Inception module for raw I/Q time series data as network input.

DL methods have been used in raw I/Q processing algorithms with sequential designs by stacking multiple convolutional layers in series. The novel architecture in this research is fusion of both series (*Stem*) and parallel (*Inception*) convolutional layers as shown in Fig. 3.

Fig. 2. Inception CNN concept

Before deliberating on each module, input data configuration requires some consideration. The received modulated data is divided into frames of 128 individual symbols, each of them comprising of 8 samples. Consequentially, each frame has 1024 samples ($N = 1024$) which can be stacked in two configurations for convolutional processing.

- *2-D configuration* ($2 \times 1024 \times 1$) : In this arrangement, 2 rows of I/Q channels with 1024 samples stacked in horizontal axis with inherent depth size of unity.
- *3-D configuration* ($1 \times 1024 \times 2$) : I/Q channels are arranged in depth dimension and present vector convolution option only.

Both these configurations have been implemented for proposed CNN and their performance comparison has been presented in subsequent sections.

Fig. 3. Proposed CNN Architecture

B. Stem and Inception Modules

The conceived fusion of Stem and Inception modules for modulation classification with network parameter details is shown in Fig. 3. Stem module is a simple stacked sequence of 4 blocks of convolutional layer, batch normalization, ReLU activation layer and maxpooling layer. The filter size of each of the convolutional layer increase in block succession as 16, 24, 32 and 64 filters of 1×8 filter size. These convolutional

layers perform feature mapping symbol by symbol (8 samples simultaneously) and extract patterns for modulation identification. Batch normalization standardizes the inputs to single layer over the concerned symbol. The training process is both stabilized and accelerated by normalization layer along with mitigation of network overfitting [15]. Activation functions perform mathematical operation on a node and determine the output of a CNN. We have used ReLU (rectified linear unit) function which is represented by following piecewise expression [16].

$$f(x) = \max(0, x) = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases} \quad (14)$$

Pooling layer selects single value out of a certain pool of inputs depending upon a predefined conditionality such as choosing the maximum value out of pool or averaging the values of all elements of pool. This helps in reduction of feature size to lower dimensions, hence, reduction in parameter count without sacrificing the network accuracy drastically.

Proposed inception module draws its foundation from 4 branch inception block in Inceptionv3 network [13]. The structure is similar to the referred block but the filter sizes have been modified according to concerned input data requirements. It is to be noted that all the convolutional filters in the complete network focus on the individual channel characteristic in a particular symbol except one ($2 \times 8, 24$ filters) in the second branch of Inception module which explores any possible correlation between both input inphase and quadrature channels. Moreover, average pooling in each branch of the module has been used to reduce the dimensions of feature maps before classification layers which form output classification vector ($1 \times 1 \times 11$).

C. Network Training

Owing to very low parametric volume ($\sim 0.2M$ learnables), network was trained using Matlab Deep Learning Toolbox in CPU configuration [14]. Dataset is generated by modulation of 1000 frames for each type of 11 modulation scheme and then subjected to Rician channel, AWGN noise model and zero-forcing equalization. This dataset is split into training (80%), validation (10%) and testing (10%). The network training parameters are tabulated in Table I.

For most of the SNR range, the training and validation accuracy values have conformed to each other throughout the training process, which relates to judicious curve fitting by the proposed CNN. The training process for 10 dB and 20 dB SNR with 2-D input configuration is shown in Fig. 4 and Fig. 5 with training and validation accuracy increasing monotonically while the cross-entropy loss reduces exponentially with incremental epoch index.

IV. PERFORMANCE ANALYSIS

The trained Inception module based CNN has been tested over a wide SNR range for two input configurations, 2-D and 3-D I/Q channel stacks. Individual modulation classification accuracy for both configurations are evaluated and confusion

TABLE I

Training Hyperparameters	
Total network learnables	190903 (~ 0.2 million)
Classification outputs	11
Sample length or Frame size	128 symbols
Samples per symbol (sps)	8 samples
Optimizer	Stochastic Gradient Descent with Momentum (SGDM)
Training data size	800 frames/modulation class
Validation data size	819,200 samples/modulation class
	100 frames/modulation class
	102,400 samples/modulation class
Sampling rate	200kHz
Drop out rate	0.5
Epochs	15
Batch size	256
Maximum iterations	510
Initial learning rate	0.02
Final learning rate	0.002
Training Processor	Intel i7-8550U clocked 3.00GHz (8 GB RAM)
Average training time	25 minutes

Fig. 4. Training process for 2-D input configuration at 10 dB SNR

Fig. 5. Training process for 2-D input configuration at 20 dB SNR

matrices have been formulated. Moreover, average accuracy for modulation sample space ($\mathbb{L}$) is analyzed which points out that beyond 10 dB SNR ratings, accuracy of proposed network surpasses 90% benchmark. The test parameters for proposed CNN are given in Table II.

A. Individual Modulation Classification Accuracy

For entire range of SNR values under consideration, the individual classification accuracy for modulation sample space

TABLE II

Test Parameters	
SNR range	[-25,+25]
Testing data size	100 frames/ modulation class
	102,400 samples /modulation class
Average testing time	33 seconds (for complete test)
	2.7 ms/symbol

has been investigated for both 2-D and 3-D input configurations and are presented in Fig. 6 and Fig. 7 respectively.

Fig. 6. Individual modulation classification accuracy for $2 \times 1024 \times 1$ - 2-D input configuration Fig. 7. Individual modulation classification accuracy for $1 \times 1024 \times 2$ - 3-D input configuration

Careful analysis of Fig. 6 and Fig. 7 leads to the fact that network accuracy rapidly improve for all modulation techniques as the SNR crosses the 0 dB threshold, which is intuitive as the signal strength becomes dominant with respect to noise level. For 3 modulation schemes (GFSK, CPFSK and B-FM), network performs exceedingly well as it classifies these modulation with accuracy beyond 80% for even -10 dB SNR. The reason behind this performance is that these three modulation has least pattern resemblance with other modulation schemes. Moreover, at SNR level 20dB, 9 out of 11 modulation sample space elements have 100% classification accuracy for 2-D configuration except DSB-AM and SSB-AM.

These corollaries are further validated by the confusion matrices which provide visualization of accuracy of predicted class given a certain transmission modulation. Confusion matrices for both configurations at 10 dB SNR are shown in Fig. 8 and Fig. 9.

Fig. 8. Confusion matrix for $2 \times 1024 \times 1$ - 2-D input configuration

Fig. 9. Confusion matrix for $1 \times 1024 \times 2$ - 3-D input configuration

A preview of these confusion matrices leads to the trade off between two pairs of modulation schemes, 16-QAM $\leftrightarrow$ 64-QAM and DSB-AM $\leftrightarrow$ SSB-AM. The reason behind this confusion of network for these pairs is that one of the pair elements is subset of the other.

B. Average Network Accuracy

The average network accuracy for both input configurations can be calculated by equation (4) as conditional probabilities for all the modulation schemes are available in confusion matrices over complete SNR range as shown in Fig. 10.

Fig. 10. Average test accuracy for proposed network

The average accuracy curves indicate the superior performance of 2-D input configuration compared to 3-D version for most of the SNR range because it offer more cross-section for convolutional operations (plane versus vector incidence),

hence, better feature extraction is possible in planar convolution.

C. Performance Comparison with DL Algorithms for Time-Series I/Q Inputs

In this section, we consider the performance of proposed Inception module based CNN with other DL algorithms classifying time-series raw I/Q channels data. It is unjustified to compare with constellation image based classification algorithm because they involve massive number of parameters and large training and testing time, which certainly provide better accuracy at cost of intricate computational architecture and resources. In [1], Meng et al. have simulated 5 sequentially stacked convolutional layers using complex maximum likelihood (ML) loss function for classification of 7 modulation schemes with 2M parameters with no reference to training time and achieves 95% average accuracy at 10 dB. The authors in [19] have used time and stockwell channeling for automatic DL modulation classification with 1.8M hyperparameters and claim network accuracy of 97% at just 8 dB SNR. However, the training process was heavily biased as 85% of the total dataset was used for training, 10% for validation and just 5% for testing purpose.

For performance analysis of the proposed Inception module based CNN, some comparable research works have been subjected to similar input conditions. 2-D Input configuration $2 \times 1024 \times 1$ and 11 output modulation classes are set to be benchmarks for various architectures developed by different researchers. The CNN model presented in [6] by Shea et al. which uses just 2 series convolutional layers CNN of 64 and 128 filters with readjusted network parameters 3M for $2 \times 1024 \times 1$ input configuration for 11 modulation techniques. The average accuracy for this network is limited to 83% as SNR reaches 10 dB and remains constant till 20 dB value. The proposed Inception module based CNN performs much better with average accuracy of 90% at 10 dB and 99% at 20 dB with around 40% lesser parameters by utilizing parallel convolutional layers. Sequential LSTM structure (Fig. 11) proposed by Rajendran et al. in [7] has been implemented by 2 layers of 128 LSTM cells in which first layer takes parallel feed from I/Q inputs and returns complete sequences while second layer has sequential feeding from adjacent cells and only last cell returns last state to dense layer. This networks utilizes 1.25M parameters for 2-D input samples and provides impressive network accuracy beyond 0 dB. 2-layer GRU network with 64 cells in each layer devised by Hong et al. in [8] provides comparative accuracy performance with 0.57M parameters as shown in Fig. 12. Finally, we draw comparison with three deep neural networks presented by Liu et al. in [9]. The designed Resnet and DenseNet both have similar 7-layer structures involving 2.3M parameters as shown in Fig. 13 and 14. Whereas, CLDNN model is a sequential combination of CNN network similar to ResNet and DenseNet and LSTM with 50 units with 33.2M parameters for $2 \times 1024 \times 1$ input configuration as shown in Fig 15.

Fig. 11. Rajendran et al 2-layer LSTM network for $(2 \times 1024 \times 1)$ input configuration

Fig. 12. Hong et al 2-layer GRU network for $(2 \times 1024 \times 1)$ input configuration

Fig. 13. Liu et al 7-layer ResNet model for $(2 \times 1024 \times 1)$ input configuration

The comparative results for proposed Inception CNN and the above discussed network models is based on two metrics: accuracy and hyperparameter volume. Number of parameters required to be trained are least for presented Inception CNN as shown in Table III, while test accuracy is comparable with other contemporaries networks at low SNR condition and even

Fig. 14. Liu et al 7-layer DenseNet model for $(2 \times 1024 \times 1)$ input configuration

Fig. 15. Liu et al 8-layer CLDNN model for $(2 \times 1024 \times 1)$ input configuration

surpasses at higher SNR values as shown in Fig 16. The reason for this superior performance is the parallel architecture of Inception module which allows high modulation classification accuracy with very low parameter volume (just 0.2M parameters).

TABLE III

Parametric Volume of Modulation Classification Networks (in millions)	
Inception module CNN	0.2
Single Layer LSTM	0.54
2-Layer GRU	0.57
2-Layer LSTM	1.25
7-Layer ResNet	2.3
7-Layer DenseNet	2.3
8-Layer CLDNN	33.2

V. CONCLUSION

In this research, a novel Inception module based CNN has been developed for autonomous modulation recognition for 11 different modulation schemes from raw I/Q time-series data. The proposed classification network possesses two components, *Stem* and *Inception* modules which fuse series stacking of convolutional layers with 4 branch parallel structure of convolution operation inspired by *Inceptionv3* architecture. Owing to low parametric bulk (~ 0.2 million learnables), the network requires minimal computational resources for training and inference purposes. For network training, a sample set

Fig. 16. Comparison of Inception Module with various modulation classification networks at $(2 \times 1024 \times 1)$ input configuration

of 11.3 million samples for 11 modulation classes has been generated and transmitted over Rician fading and AWGN noise channels. Signal conditioning is performed by receiver by using zero-forcing equalizer. Dataset is segregated in training (80%), validation (10%) and inference (10%) subsets. The trained Inception module based network has been analyzed with inference accuracy curves and confusion matrices over wide range of SNR values. The proposed network has presented impressive inference performance as it achieves average accuracy of 90% at 10 dB and 99% at 20 dB SNR which compete with test accuracy valuations for other high parametric DL algorithms used for modulation classification. These characteristics of the proposed CNN make it an ideal choice for low cost solutions for modulation recognition devices in time constrained communication scenarios yet not compromising reliability where expensive computational resources (such as GPUs) are deemed as exorbitant selection. Hardware implementation on SDR for RF domain of this modulation classification system can be considered as a possible future endeavour.

REFERENCES

- [1] F. Meng, P. Chen, L. Wu and X. Wang, "Automatic Modulation Classification: A Deep Learning Enabled Approach," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10760-10772, Nov. 2018.
- [2] Jiang K, Zhang J, Wu H, Wang A, Iwahori Y. A Novel Digital Modulation Recognition Algorithm Based on Deep Convolutional Neural Network. *Applied Sciences*. 2020; 10(3):1166.
- [3] X. Xie, Y. Ni, S. Peng and Y. Yao, "Deep Learning based Automatic Modulation Classification for Varying SNR Environment," *2019 28th Wireless and Optical Communications Conference (WOCC)*, Beijing, China, 2019.
- [4] J. Zheng and Y. Lv, "Likelihood-Based Automatic Modulation Classification in OFDM With Index Modulation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8192-8204, Sept. 2018.
- [5] Zhang D, Ding W, Zhang B, Xie C, Li H, Liu C, Han J. Automatic Modulation Classification Based on Deep Learning for Unmanned Aerial Vehicles. *Sensors*. 2018; 18(3):924.
- [6] T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563-575, Dec. 2017.
- [7] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders and S. Pollin, "Deep Learning Models for Wireless Signal Classification With Distributed Low-Cost Spectrum Sensors," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433-445, Sept. 2018.

- [8] D. Hong, Z. Zhang and X. Xu, "Automatic modulation classification using recurrent neural networks," *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, 2017.
- [9] X. Liu, D. Yang and A. E. Gamal, "Deep neural network architectures for modulation classification," *2017 51st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2017.
- [10] S. Peng, H. Jiang, H. Wang, H. Alwageed and Y. Yao, "Modulation classification using convolutional Neural Network based deep learning model," *2017 26th Wireless and Optical Communication Conference (WOCC)*, Newark, NJ, 2017.
- [11] S. Peng et al., "Modulation Classification Based on Signal Constellation Diagrams and Deep Learning," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 718-727, March 2019.
- [12] C. Szegedy et al., "Going deeper with convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015.
- [13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016.
- [14] The MathWorks, I., 2019. *Deep Learning Toolbox*, Natick, Massachusetts, United States.
- [15] Sergey Ioffe and Christian Szegedy. 2015. "Batch normalization: accelerating deep network training by reducing internal covariate shift." In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37 (ICML'15)*. JMLR.org, 448-456.
- [16] H. Ide and T. Kurita, "Improvement of learning for CNN with ReLU activation by sparse regularization," *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, 2017.
- [17] Murphy, K. P. *Machine Learning: A Probabilistic Perspective*. The MIT Press, Cambridge, Massachusetts, 2012.
- [18] P. Szymak, "Selection of Training Options for Deep Learning Neural Network Using Genetic Algorithm," *2019 24th International Conference on Methods and Models in Automation and Robotics (MMAR)*, Miedzydroje, Poland, 2019, pp. 24-29.
- [19] S. M. Hiremath, S. Behura, S. Kedia, S. Deshmukh and S. K. Patra, "Deep Learning-Based Modulation Classification Using Time and Stockwell Domain Channeling," *2019 National Conference on Communications (NCC)*, Bangalore, India, 2019, pp. 1-6.