Journal of Advanced Engineering Trends
ISSN : 2682 - 2091

Vol. 44, No. 1, January 2025

Automatic Modulation Classification for Enhanced Cognitive Radio for IoT Systems based on Deep Learning

Engy M. Ali[1,*], Gerges M. Salama[1], Khalil A. A.[1], Mai Ezz-Eldin[2]
[1] *Electrical Engineering Dep., Faculty of Engineering, Minia University, Minia, Egypt.*
[2] *Electrical and Communication Engineering Dep., Future High Institute of Engineering, Fayoum, Egypt*
* *Corresponding author(s) E-mail:* engymohamedali1997@gmail.com

A R T I C L E I N F O

Article history:
Received:5 September 2024
Accepted: 21 October 2024
Online: 21 February 2025

Keywords:
Automatic Modulation Classification
Cognitive Radio
Deep Learning
Depthwise Separable Convolution
IoT Systems

A B S T R A C T

Automatic Modulation Classification (AMC) plays a crucial role in Cognitive Radio (CR) systems, especially within Internet of Things (IoT) devices where spectrum efficiency and flexibility are paramount. Traditional modulation classification methods often rely on feature extraction and machine learning (ML) algorithms, which need a lot of complex calculations and may struggle with complex modulation schemes and noisy channels. Deep Learning (DL), particularly Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), has a positive impact on AMC due to their ability to automatically learn and extract discriminative features from the sequences of raw I/Q received signals. So, this paper proposes DL AMC model is built and investigated using Radioml2016 data for enhancing spectrum management for CR in IoT systems. The proposed model has reduce model parameters by 35% by using depthwise separable convolutional and traditional convolution to create the model architecture while increasing the accuracy of the model to 84 % at high SNR . The reduction in the model parameters led to a reduction in the prediction time to achieve the requirement in cognitive radio systems for IoT devices.

1. Introduction

The rapid increase in IoT devices necessitates efficient spectrum utilization, driving the need for intelligent CR systems. AMC is a critical task within CR, enabling devices to identify and adapt to the modulation schemes of incoming signals dynamically. DL has demonstrated significant capabilities in handling diverse and complex data patterns, making it an attractive candidate for AMC tasks in IoT CR systems.

The CR systems enable unlicensed Secondary Users (SUs) to utilize idle channels of Primary Users (PUs) without disrupting their operations. As a result, an SU intending to access a free channel band must employ a robust Spectrum Sensing (SS) technique to accurately detect whether the PU is present or absent in the channel, as illustrated in Fig. 1 of the cognitive cycle of CR. This allows the SU to effectively manage the spectrum and make appropriate decisions [1]. Some other functions that can be performed by a SU are Adaptive Coded Modulation (ACM) and Modulation Classification (MC)[2]. A CR is expected to be able to correctly recognize or classify the type of modulation scheme of the received signal rapidly without any latency so as to know there is a PU in the channel based on the fact all primary users employ one modulation technique for the transmission over the frequency channel and also to apply the appropriate demodulation process in the receiving side.

The automatic modulation classification (AMC) methods are divided into traditional methods and advanced methods. In Traditional methods, there are two primary approaches, likelihood-based and feature-based methods, have very important roles in determining the modulation type of received signals. While advanced methods for AMC depend on deep learning (DL) [3]. Likelihood-based methods rely on statistical principles to compute the likelihood that a received signal corresponds to each modulation scheme within a predefined set [4]. These methods typically involve maximum likelihood estimation or Bayesian inference to assess the probability of observing the signal under different modulation hypotheses [5]. In contrast, feature-based methods extract discriminative features from the received signal, such as constellation points, signal-to-noise ratio estimates, or higher-order statistics, and use these features to train classifiers like neural networks or SVMs for modulation classification [6]. While likelihood-based methods excel in accuracy under ideal statistical assumptions, feature-based methods offer robustness in varying signal conditions and are computationally more efficient. The choice between these methods often hinges on the specific requirements of the application, including computational complexity, robustness to noise and channel conditions, and availability of training data.

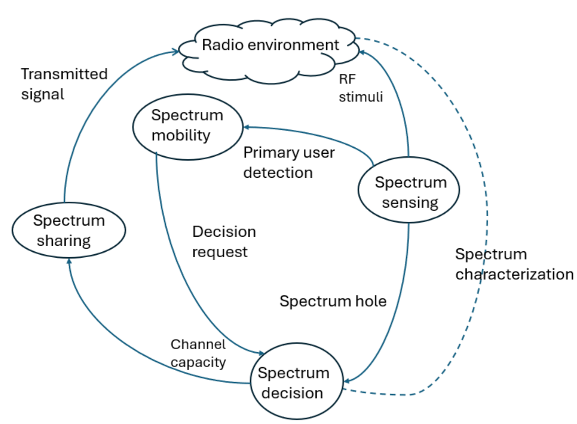


Figure 1: Cognitive cycle of CR.

Moreover, Software Defined Radio (SDR), which is core for CR systems gaining immense popularity in Which the neighbor devices are expected to adopt Conditions and adjust transmission and adjust their transmission parameters, modulation schemes, etc[7]. It is one of the most important to be able to classify the modulation type at the receiver's end without needing prior knowledge about the transmitter's system.

Recent advancements in deep learning (DL) technologies have led to the development of methods that can autonomously learn features. DL is particularly advantageous due to its ability to leverage large datasets, which are readily available in communication systems. One challenge associated with DL is its complexity, involving both training and testing phases. However, recent architectures utilizing different types of convolutions, such as separable convolutions, have shown promise. These architectures significantly reduce model size without compromising accuracy. By decreasing the number of parameters compared to traditional convolutional methods, separable convolutions are well-suited for small devices in the Internet of Things (IoT) environment. This study compares the effectiveness of separable convolutions against conventional convolutional neural networks.

This paper is organized as follows: Section1 introduction; Section 2 reviews automatic modulation classification techniques; Section 3 provides an in-depth discussion of deep learning in AMC. Section 4 is the most significant part in as it present the implementation of our model. Section 5 presents the results, while Section 6 concludes the paper and outlines directions for future work.

2. Related work

DL represents a cutting-edge research frontier within the realm of machine learning. By assimilating the underlying patterns and attributes of data samples, DL systems emulate human-like analytical and learning processes, thereby enabling advanced capabilities in prediction and classification [8]. The rapid evolution of digital computing technology has dramatically expanded both data volumes and computational power, fostering accelerated advancements in DL methodologies [9]. Contemporary deep learning algorithms find application across a diverse array of classification challenges, including but not limited to computer vision [10], bioinformatics [11], natural language processing [12], visual recognition [13], and signal processing [14]. This domain of machine learning leverages extensive datasets to train models, allowing them to extract features and effectively classify or predict novel data points [15]. The swift progress in deep learning has led to the development of sophisticated communication applications, such as modulation classification.

Numerous researchers have developed methods for the automatic classification of modulation schemes. Specifically, various likelihood-based approaches, which rely on the power spectral density (PSD) of the received signal, have been explored. It has been noted that LB methods achieve favorable results, though they come with significant computational complexity. While many studies have concentrated on the accurate prediction of modulation schemes using LB methods, they often overlook the associated computational demands [16]. FB methods, designed as suboptimal classifiers for practical use, focus on extracting valuable features from the received signal and classifying them using a classifier. These features might include instantaneous wavelet transforms, as seen in [17], or constellation diagram plots, as illustrated in [18]. Research indicates that while LB methods offer optimal solutions with high accuracy, they require considerable computation and prior knowledge about the signal. In contrast, FB methods provide a suboptimal solution with lower latency, as they do not need prior signal knowledge. Additionally, to satisfy the requirements of a cognitive radio system as outlined in [1], deep learning techniques, particularly CNN networks, have proven effective for modulation classification.

O'Shea [19] was proposed the first model using CNNs for classifying modulation types of raw signal samples generated with GNU Radio. He also introduced an open-access dataset for training neural networks, which has been widely utilized in subsequent research. In a later publication [20], a more comprehensive radio (OTA) dataset was introduced, featuring a broader array of modulation types in real-world settings. Research in [21] explored three different architectures—Convolutional Long Short-Term Memory Deep Neural Network (CLDNN), Long Short-Term Memory Network (LSTM), and Deep Residual Network (ResNet)—focusing on maintaining high performance while implementing strategies to minimize the number of parameters required for network training. Conversely, [22] developed an Automatic Modulation Classification (AMC) system based on CNN and LSTM to enhance accuracy and reduce overall training time. Additionally, [23] presented a hybrid signal recognition model combining CNN with Gated Recurrent Units (CnGr) to decrease model size, facilitating easier deployment and maintaining high recognition accuracy.

In this study, we introduce a CNN architecture designed for rapid automatic modulation classification of raw input IQ sequences data from received signals across various SNR levels, with the aim of enhancing model accuracy as possible. The CNN model was developed and analyzed to boost accuracy and minimize the total number of parameters.

3. Deep Learning Approaches for Modulation Classification

3.1. Convolutional Neural Network structure

CNNs are a specialized type of multilayer perceptron inspired by the way neurons connect in the biological brain. They are designed to handle and interpret visual data, such as images and videos, by simulating the brain's processing mechanisms [9]. Unlike conventional neural networks that often require extensive manual feature extraction, CNNs automatically learn spatial hierarchies and patterns through their convolutional layers. They are highly effective for various classification tasks because they demand less preprocessing compared to other classification techniques. A CNN consists of a series of layers, each of which transforms its input volume into a different volume using a differentiable function [8]. A typical CNN architecture, as depicted in Fig. 2, generally includes:

- **Convolutional Layers:** These layers are responsible for feature extraction from the input data. They utilize a set of learnable filters, also known as kernels, to process the input images. The filters or kernels are compact matrices, typically with dimensions like 3×3, 3×2, or 8×2. They slide over the input image and compute the dot product between the kernel weights and the corresponding patches of the image. The result of this operation is called feature maps.
- **Activation Layer**: After the convolution operation, an activation function is applied to introduce non-linearity into the network. The widely used activation function is the Rectified Linear Unit (ReLU),and SoftMax.
- **Pooling layer:** Serves to simplify and condense the data by summarizing information in localized regions of the feature map. This process aids in reducing computational demands, improving efficiency, and helping the network to focus on significant features rather than intricate details.
- **Flattening:** After the convolution and pooling layers, the feature maps are converted into a one-dimensional vector through flattening.
- **Fully Connected Layers:** These layers take the flattened input and perform the final classification or regression tasks. In these layers, each neuron is interconnected with every neuron from the previous layer. The output layer generates the ultimate results of the network.
- **Output Layer:** The ultimate layer give us the output. For classification tasks, used SoftMax layer that outputs probabilities for each class. For regression tasks, this might be a linear layer producing continuous values.
- **Normalization Layers (optional):** Batch normalization (BN) layers can be inserted to normalize the inputs of each layer to improve training speed and stability.
- **Dropout Layers (optional):** DO acts like a regularization technique, improving the network's ability to generalize by ensuring that it doesn't become overly dependent on specific neurons or patterns in the training data.

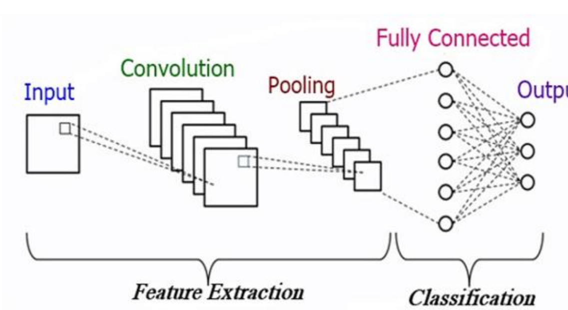


Figure 2: A typical CNN architecture.

3.2. convolution types

Convolution is a mathematical operation utilized to integrate two sets of information, producing a composite result. In the realm of CNN this operation is applied to input data to discern features and create a feature map. This operation involves a convolution kernel or filter, which is a matrix of learnable parameters, typically of dimensions. During the convolution process, the kernel traverses the input image, executing element-wise matrix multiplication. The result of this multiplication for each receptive field (the specific region of the image under the kernel) is recorded in the resulting feature map.

1.Traditional convolution

When the input feature matrix has a depth of three, four convolution kernels are used for the convolution process in traditional convolution, as illustrated in Fig. 3. Additionally, every kernel has a depth of 3, which is equal to the input feature matrix's depth. As a result, the depth of the output feature matrix will be 4 based on the quantity of convolution kernels. Considering that the convolution process has a stride of 1, the standard convolution computation cost is computed as follows in equation (1).

$$conv_t = D_k . D_k . M . N . D_f . D_f \quad (1)$$

where M: Represents the count of input channels, N: Indicates the count of output channels, D_k : Refers to the size of the convolutional kernel, and D_f : Denotes the dimensions of the feature map.

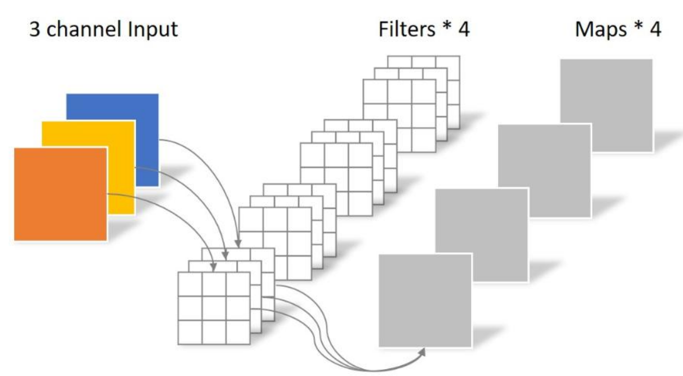


Figure 3: Traditional convolution process.

2. Depthwise Separable Convolution DSC

Depthwise separable convolution (DSC) is an optimized variant of standard convolution aimed at reducing computational complexity and model size. It decomposes the convolution process into two separate layers: depthwise convolution and pointwise convolution. The concept of depthwise separable convolution was initially introduced in [24] and later utilized by MobileNet models to decrease model complexity [25].Similar to spatially separable convolution, depthwise separable convolution involves splitting the kernel into two distinct components: depthwise convolution and pointwise convolution [26].

- Depthwise convolution applies an individual filter to each input channel independently, as demonstrated in Fig. 5. For example, with 3 input channels, there will be $3 * D_k * D_k$ spatial convolutions. Each filter interacts only with its corresponding channel,as resulting in equation (2).

$$complexity = D_k . D_k . M . D_f . D_f \quad (2)$$

where M: Represents the count of input channels, D_k: Refers to the size of the convolutional kernel, and D_f : Denotes the dimensions of the feature map.

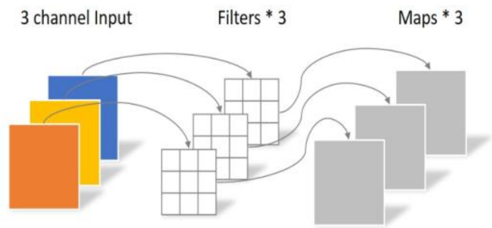


Figure 4: Depthwise convolution process.

- Pointwise convolution is employed to adjust the dimensions of the output from the depthwise convolution, aligning it with the output dimensions of a traditional convolutional neural network, as illustrated in Fig. 5. A 1×1 convolution is used to integrate the outputs from the depthwise convolution, effectively merging and combining features across various channels.

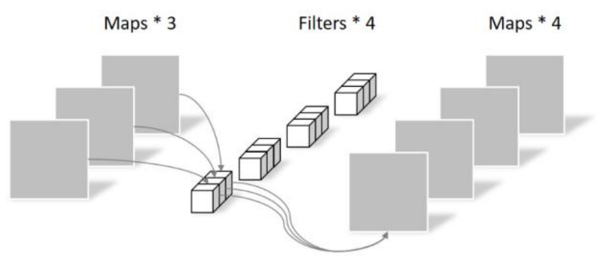


Figure 5: Pointwise convolution process.

In conventional convolutional operations, the processes of filtering and combining are executed simultaneously within a single step. In contrast, depthwise separable convolutions decompose these tasks into two distinct stages, as illustrated in Fig. 6. Specifically, DSC partition the convolutional operation into a DC layer and a PC layer. The cumulative computational complexity of depthwise separable convolution is derived in equation (3) from the sum of the complexities associated with each of these two layers [23].

$$conv_{DSC} = D_k.D_k.M.D_f.D_f + M.N.D_f.D_f \qquad (3)$$

where M: Represents the count of input channels, N: Indicates the count of output channels, D_k: is size of the convolutional kernel, and D_f: Denotes the dimensions of the feature map.

The ratio of computational cost between the traditional convolution and depthwise separable convolution is obtained by equation 4.

$$\frac{conv_{DSC}}{conv_t} = \frac{1}{N} + \frac{1}{{D_k}^2} \qquad (4)$$

When the convolution kernel size is 3, the equation simplifies to 1/N + 1/9 In theory, traditional convolution has a computational cost that is 8 to 9 times greater than that of DSC. DSC reduces the parameters involved in convolution operations. However, smaller models might yield suboptimal results if depthwise separable convolution replaces conventional 2D convolution. Nevertheless, when used appropriately, depthwise separable convolution can be highly efficient without compromising the model's performance.

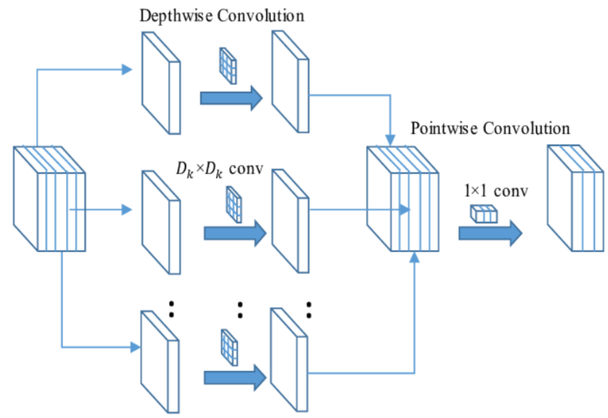


Figure 6: Depthwise separable convolution process.

4. Methodologies and Implementations

This paper seeks to identify the correct modulation scheme among 11 different modulated signals. It is assumed that the receiver is an IoT device with an antenna , receiving signals from a source within a cognitive radio system. The received signal for the CR system can be represented by the mathematical formula shown in equation (5).

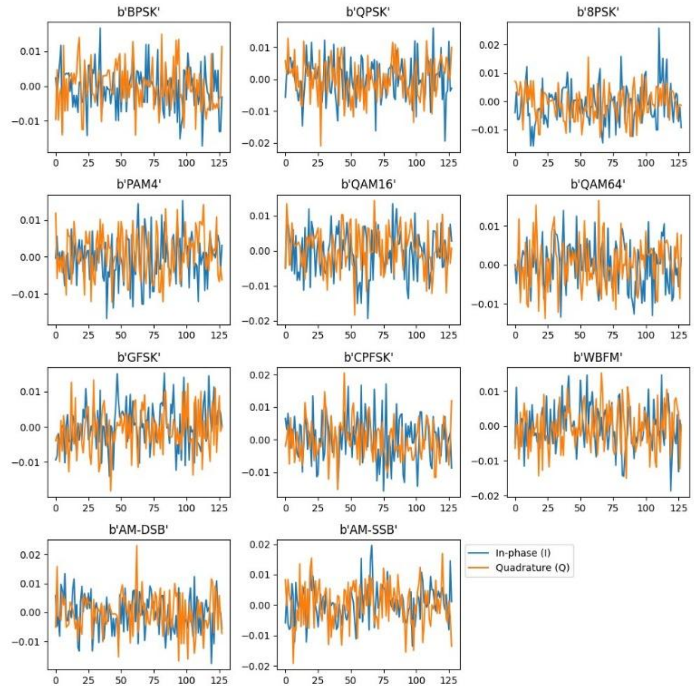
$$v(t)=s(t)*h(t)+g(t) \qquad (5)$$

where $v(t)$ denote the observed signal at the receiver, $s(t)$ signifies the transmitted modulated waveform, $h(t)$ encapsulates the channel's impulse response, representing the spectral characteristics and distortive effects imparted by the wireless medium, and $g(t)$ embodies the stochastic white Gaussian noise introduced by the channel, which follows an adaptive statistical distribution.

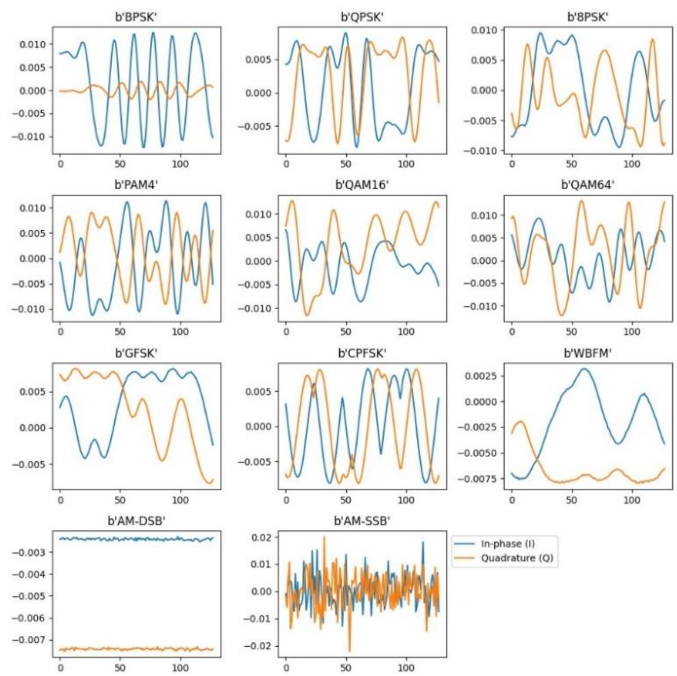
In fact, using a cognitive radio spectrum, you can transmit two signals at the same time One of the signals is known as I (in-phase), the other as Q (quadrature). The two sinusoids that have the same frequency and a relative phase shift of 90°. The I/Q data encompasses all the characteristics of the baseband signal and is represented in a complex form.

4.1. Data Preparation:

Raw IQ samples or spectrogram representations are used as inputs. In this work, we test our model over (RADIOML 2016.10A) dataset [19]. This dataset includes 11 modulation category (comprising 8 digital and 3 analog) and spans 20 different SNR values ranging from -20 dB to 18 dB. The complete set of 11 modulation classes. The data is structured in a dictionary format, where each key is a tuple representing (modulation class, SNR value), and each corresponding value contains the associated IQ samples. Initially, during data preprocessing, we visualized the data across various SNR levels. Subsequently, we restructured the dataset into a dictionary with a single key representing the entire dataset.



a. GNU Radio ML dataset at SNR= -18 dB.



b. GNU Radio ML dataset at SNR= 18 dB.

Figure 7: GNU Radio ML dataset at: (a) SNR = -18 dB (b) SNR = 18 dB.

Initially, all modulation classes were assigned numeric labels. These labels were then converted into a one-hot encoding vector, which is a 1 × N matrix used to differentiate each class within a dictionary. In this vector, only one element corresponding to the class number is set to 1, while all other elements are set to 0. This transformation is beneficial because one-hot encoding enhances the representation of categorical data [16]. Many machine learning algorithms cannot directly process categorical data, necessitating its conversion into a numerical format. Consequently, the updated records include all data from each class and SNR value, organized into a list of [one-hot encoded labels]. The modulation dataset is divided into a training set (65%), a validation set (15%), and a test set (20%). Fig. 7 illustrates the GNU Radio ML dataset for various modulation classes at two different SNR levels.

4.2 Model Architectures:

CNNs, primarily recognized for their application in image recognition, can also be adapted for wireless signal recognition by depicting the signal's time-domain information as a time-domain graph. In this study, we implement a depthwise separable convolutional neural network to optimize the number of parameters, reduce model size, and decrease overall training time. It is engineered to achieve highest possible accuracy while minimizing computational complexity. The model architecture consists of four convolutional layers: two conventional convolutional layers and two layers employing depthwise separable convolution, as illustrated in the architecture of our model in Fig. 8.

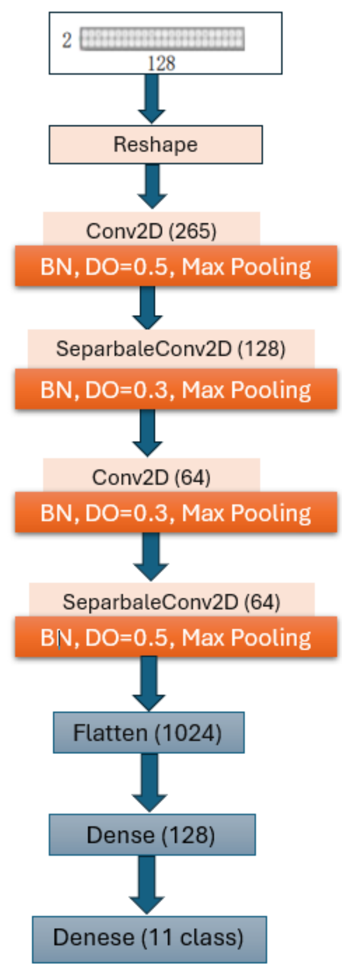


Figure 8: Proposed model architecture.

4.3 Training and testing

The models were trained and tested using the Keras library with a TensorFlow backend and executed on Google Colab. Optimal weights and coefficients were selected based on a machine learning parameter optimization scheme. The learning rate was configured to 0.001, and the Adam optimizer was utilized for the training process. A batch size of 400 was used, and the ReLU activation function was applied to all layers except the final dense layer, which utilized the softmax activation function. The convolutional layers employed a filter size of (3*3) and a stride of 1 step. The categorical cross-entropy loss function quantifies the discrepancy between the probability distributions predicted by the model and the true probability distributions, effectively measuring the error in classification performance. A callback function was set to halt training early if the accuracy of the model showed no improvement.

5. Experimental Results

In this section, we examine the impact of our proposed network architecture. To assess the performance of our model, we compared its classification accuracy against three state-of-the-art models: CNN2 [27], Depthwise Separable [28], and ResNet-LSTM [29]. For an equitable comparison, all models were trained using the same dataset (RADIOML 2016.10A) and evaluated on the identical holdout test dataset. Figure 9 provides a detailed analysis of the recognition accuracy. The results indicate that our model performs more consistently at SNRs greater than 0 dB.

When deploying the model in cognitive radio systems and IoT devices, it is essential to consider not only its accuracy but also its complexity. Key factors for evaluating model complexity include the number of training parameters and the training time, as summarized in Table 1.

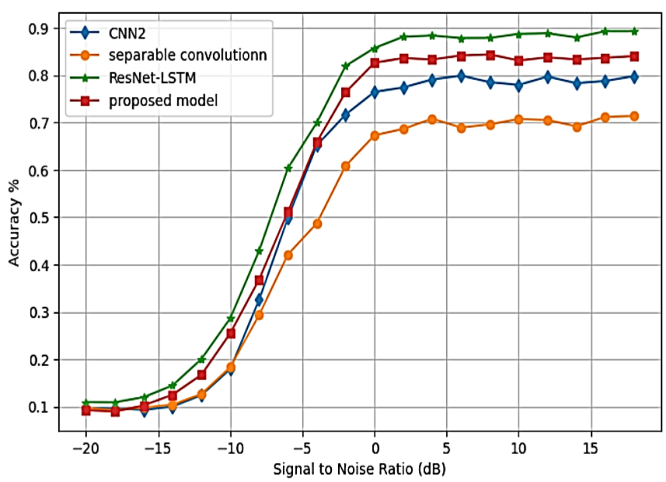


Figure 9: Accuracy overview - RadioML2016.10a.

From the analysis of the models illustrate in table 1, we see that the proposed model achieves the requirement goals to deployment in cognitive radio system for IoT devices as its reduction system parameters with accuracy 84%. On the other hand, the ResNet-LSTM achieves the highest accuracy which was 90% but increases parameters by 68 %. The reason of this difference is that the ResNet-LSTM model depends on traditional convolution in addition to using more numbers of layers that had more parameters compared by our proposed model that depends on DSC. However, as the number of parameters grows, the model may take longer to process and detect. Long Short-Term Memory (LSTM) networks are designed to manage variable-length input sequences by dynamically adjusting their internal state. This is useful when dealing with signals as in our case. So, we work on another model combining between depthwise separable convolution and LSTM to obtain the best model with low parameters and highest accuracy.

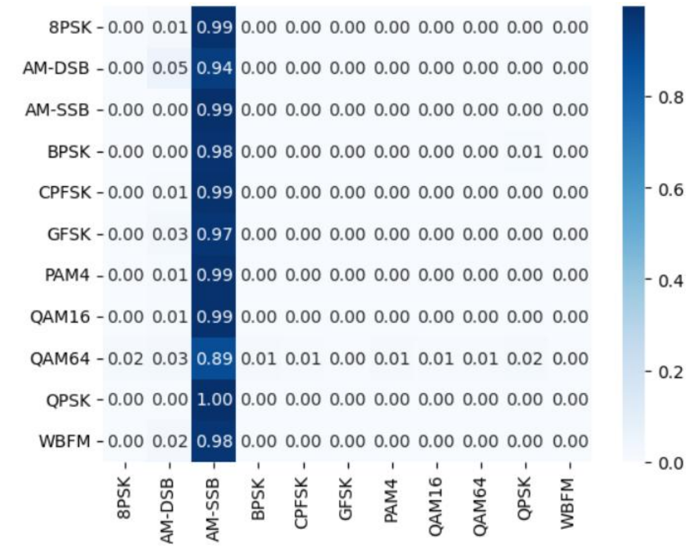
Model	Parameters Model size	Accuracy at SNR = 18
CNN2	858,123	80%
Depthwise separable	385,307	71%
ResNet-LSTM	770,378	90%
Proposed model	251,467	84%

Table 1: Comparison of accuracy achieved on (18dB) SNR and total parameters for each model.

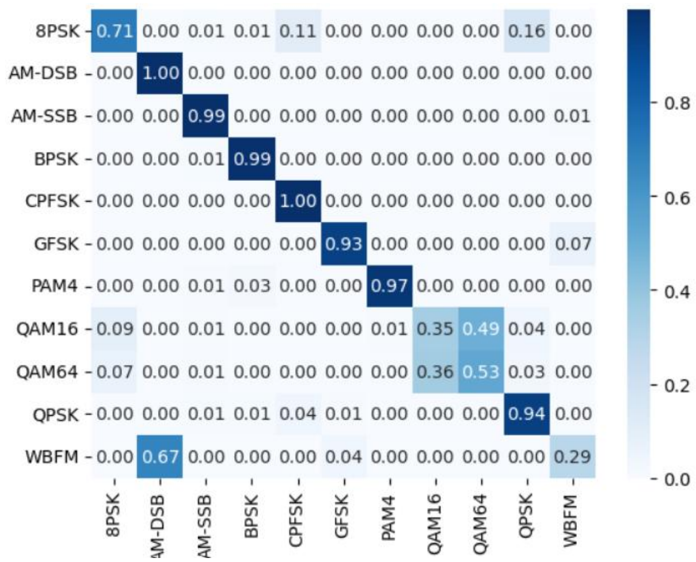
To evaluate the effectiveness of the proposed model in identifying various modulation modes, Fig. 10 displays the confusion matrices for different recognition methods across three distinct SNR levels. The horizontal axis indicates the predicted class labels assigned by the model, while the vertical axis represents the actual class labels from the dataset. The diagonal entries of the matrix indicate correctly identified classes. It was observed that at low SNR (−20 dB), the model tends to classify all modulation modes as AM-SSB. This is because noise and other interference predominantly affect the received signal, making it resemble Gaussian noise. As the SNR improves, the impact of noise diminishes, and the confusion matrix becomes more distinct, with each signal category being classified accurately. Nonetheless, even with an SNR of 18 dB, WBFM may still be incorrectly identified as AM-DSB due to the signal's quiet periods [30]. Despite some recognition errors, the proposed method achieves high accuracy while maintaining a lightweight network. There are also some misclassifications between AM-DSB and WBFM, as well as confusion between 16QAM and 64QAM, due to 16QAM being a subset of 64QAM.

6. Conclusion & future work

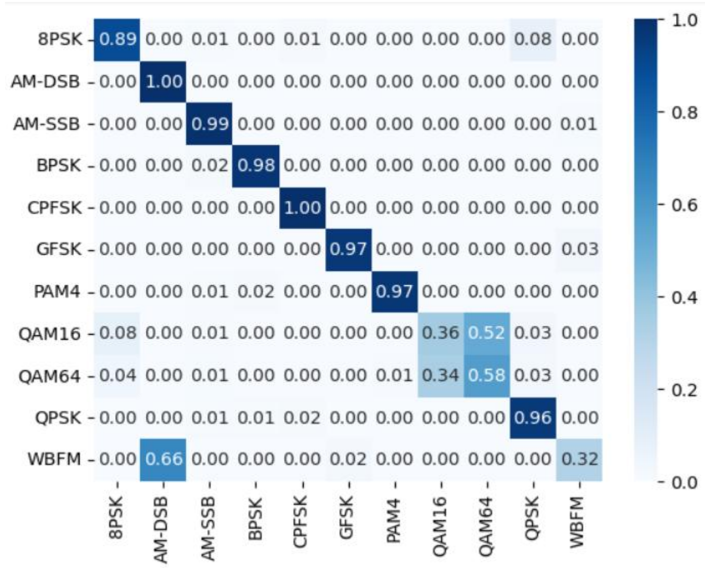
Both depthwise separable and traditional convolutional neural network architectures have been explored for classifying modulation schemes. Low-power IoT devices are often incapable of handling the large number of parameters typical in conventional neural networks. In contrast, the proposed model utilizing depthwise separable convolutions achieves a comparable accuracy to traditional convolutional methods, while significantly reducing the number of parameters, model size, and training time. This approach minimizes computational costs while maintaining high performance, making it ideal for real-time signal Classification tasks. The model's hyperparameters and architecture should be adjusted according to the specific



a. SNR = -20 dB.



b. SNR = 0 dB.



c. SNR = 18 dB.

Figure 10: The Confusion matrix for the proposed model at: (a) SNR = -20 dB (b) SNR = 0 dB (c) SNR = 18 dB.

characteristics of the dataset and the performance outcomes. These models are particularly suitable for implementation in low-power IoT devices and Software-Defined Radios (SDR) within cognitive radio systems, which face constraints related to power and time. Future research will investigate integrating additional lightweight operations and exploring recurrent neural networks, like Long Short-Term Memory (LSTM) networks, to develop even more efficient classification models with lowest possible parameters while enhancing overall accuracy. Develop our model in real life IoT devices.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

There is no fund has been received for conducting the research.

References

[1] Alias DM, Raghesh GK, "Cognitive radio networks: a survey," in: 2016 international conference on wireless communications, signal processing and networking, pp 1981–1986.

[2] Salama, G., Mohamad, Y., Saif, M. (2024). 'The Role of Sigma-Delta ADCs and Zero-Forcing Estimator in Massive MIMO Channel Estimation', Journal of Advanced Engineering Trends, 43(2), pp. 445-454. doi: 10.21608/jaet.2024.243344.1262

[3] Wang, T.; Yang, G.; Chen, P.; Xu, Z.; Jiang, M.; Ye, Q. A Survey of Applications of Deep Learning in Radio Signal Modulation Recognition. Appl. Sci. 2022, 12, 12052. [4] Smith, Anna H., and Katlyn S. Spadino. "Alzheimer's Disease: Knowledge of Early Detection, Risk Factors, and Treatment." (2016).

[4] James A. Sills. Maximum-likelihood modulation classification for psk/qam. MILCOM 1999. IEEE Military Communications. Conference Proceedings (Cat. No.99CH36341), 1:217–220 vol.1, 1999.

[5] Ghauri, S.A. KNN based classification of digital modulated signals. IIUM Eng. J. 2016, 17, 71–82.

[6] Uys, L.Y.; Gouws, M.; Strydom, J.J.; Helberg, A.S.J. 'The Performance of Feature-based Classification of Digital Modulations under Varying SNR and Fading Channel Conditions.' In Proceedings of the 2017 IEEE AFRICON, Cape Town, South Africa, 2017; pp. 198–203.

[7] Definitions of Software Defined Radio (SDR) and Cognitive Radio System (CRS), ITU-R Report SM, 2009, no. 2152.

[8] Goodfellow, Y. Bengio, and A. Courville, Deep learning book. MIT press, 2016.

[9] N. Lepri et al., "In-Memory Computing for Machine Learning and Deep Learning," IEEE Journal of the Electron Devices Society, vol. 11, pp.587–601, 2023.

[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] Yousef, Malik and ALLMER, Jens "Deep learning in bioinformatics," Turkish Journal of Biology: Vol.47: No. 6, Article 3, 2023.

[12] H. Li, “Deep learning for natural language processing: advantages and challenges,” Natl. Sci.Rev., 2017.

[13] Alaa, M., Salama, G., Galal, A., Hamed, H. (2022). 'A Robust Lane Detection Method for Urban roads', Journal of Advanced Engineering Trends, 41(1), pp. 13-26. doi: 10.21608/jaet.2020.37172.1025

[14] Ngo, Thinh, Brian T. Kelley, and Paul Rad. ”Deep Learning Based Prediction of Doppler Shift for Mobile Communications.” In 2021Telecoms Conference (ConfTELE), pp. 1-6. IEEE, 2021..

[15] Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN COMPUT. SCI. 2, 420 (2021).

[16] 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.

[17] Wu, Hong & Hua, Xiaodong. (2020). A Review of Digital Signal Modulation Methods Based on Wavelet Transform. 1123-1128. 10.1109/ICMA49215.2020.9233682.

[18] Ma, P.; Liu, Y.; Li, L.; Zhu, Z.; Li, B. A Robust Constellation Diagram Representation for Communication Signal and Automatic Modulation Classification. Electronics 2023, 12, 920.

[19] O’Shea, T.J., Corgan, J., Clancy, T.C.: Convolutional radio modulation recognition networks. In: International Conference on Engineering Applications of Neural Networks, pp. 213–226. Springer, Heidelberg (2016)

[20] O’Shea, T.J., Roy, T., Clancy, T.C.: Over-the-air deep learning based radio signal classification. IEEE J. Sel. Top. Signal Process. 12, 168–179 (2018)

[21] 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.

[22] Z. Wang, P. Wang and P. Lan, "Automatic Modulation Classification Based on CNN, LSTM and Attention Mechanism," 2022 IEEE 8th International Conference on Computer and Communications (ICCC), Chengdu, China, 2022, pp. 105-110,

[23] Wang, F.; Shang, T.; Hu, C.; Liu, Q. Automatic Modulation Classification Using Hybrid Data Augmentation and Lightweight Neural Network. Sensors 2023, 23, 4187

[24] Chollet, Francois. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. 1800-1807. 10.1109/CVPR.2017.195.

[25] 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,” Comput. Sci., 2017.

[26] Gao, Jingpeng & Ji, Xiangyu & Chen, Geng & Guo, Ruitong. (2023). Main–Sub Transformer With Spectral–Spatial Separable Convolution for Hyperspectral Image Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. PP. 1-17. 10.1109/JSTARS.2023.3342983.

[27] Tekbıyık, Kürşat & Ekti, Ali & Gorcin, Ali & Karabulut Kurt, Gunes & Kececi, Cihat. (2020). Robust and Fast Automatic Modulation Classification with CNN under Multipath Fading Channels. 1-6. 10.1109/VTC2020-Spring48590.2020.9128408.

[28] Usman, Muhammad & Lee, Jeong A. (2021). AMC-IoT: Automatic Modulation Classification Using Efficient Convolutional Neural Networks for Low Powered IoT Devices. 10.1109/ICTC49870.2021.9289261.

[29] M.M. Elsagheer, S.M. Ramzy, "A hybrid model for automatic modulation classification based on residual neural networks and long short term memory," Alex. Eng. J., vol. 67, pp. 117–128, 2023.

[30] Pijackova, Kristyna, and Tomas Gotthans. "Radio Modulation Classification Using Deep Learning Architectures." 2021 31st International Conference Radioelektronika (RADIOELEKTRONIKA). IEEE, 2021.

Abbreviation and symbols

AMC	Automatic Modulation Classification
CR	Cognitive Radio
IoT	Internet of Things
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
RNN	Recurrent Neural Networks
SDR	Software-Defined Radio
AI	Artificial Intelligence
FB	Feature-based
LB	Likelihood-based
LSTM	Long Short-Term Memory
DSC	Depthwise Separable Convolution
DC	Depthwise Convolution
PC	Pointwise Convolution
SNR	Signal to Noise Ratio
BN	Batch normalization layer
DO	Dropout Layer