Modulation Recognition and classification using DBRNN and CNN modules

G. INDUMATHI, LOGESHWARAN J, SATHYA NARAYANAN G
Department of Electronics and Communication Engineering,
Mecpo Schlenk Engineering College, Sivakasi,
INDIA

Abstract: - Modulation recognition represents one of the key fields in the wireless communications field, as it is widely used in both criminal and peaceful communication forms, following the application area. Nonetheless, things get hard when there is a drop in the signal quality, especially when we deal with the low SNR (across -10dB), providing us with such poor results. This is why we offer an enlistment of the Modulation Recognition System using a Denoising Bidirectional Recurrent Neural Network (DBRNN). The initial layer examines signal restoration, which is the role of the DBRNN (Denoising Bidirectional RNN), and the goal is to reduce the effect of noise. Next, an autoencoder is used as the second phase of the system, where the signal is decomposed into the information without noise. The procedure in question involves transmission and then finally, a noiseless deletion. Finally, we use three stages of the Convolutional Neural Network (CNN) modules for real classification of the noise-reduced output of the autoencoder unit. By means of the application of our method, we can look for the modulations, spectral features or speech patterns within the noisy audio signals, which are very useful for cognitive radio systems and many other applications. The detection algorithm simulations that we conducted evidenced that our solution was showing a growing tendency to increase the efficiency of the recognition accuracy. Subsequent inspection of the diverse design with various structural groups and frameworks of the parameters helps in achieving an accuracy of 85% for classification, even for -20 dB signals , thus proving its sturdiness and efficiency.

Key-Words: - autoencoder, CNN, Denoising Bidirectional Recurrent Neural Network, low SNR, modulation recognition, noisy audio.

Received: June 21, 2025. Revised: July 23, 2025. Accepted: August 24, 2025. Published: December 31, 2025.

1 Introduction

Modulation recognition is one of the most widely used topics for research and development in wireless communications. This method is used for both military purposes like signal confirmation and spectral detection, whereas the same can be used for civilian purposes such as cognitive radio and signal verification. There are models used to do recognition for high-quality signals. However, when the quality of the signal is very low in terms of signal-to-noise ratio, that is less than -10dB, the classification process becomes difficult and results in very low accuracy. Hence, in order to compensate for that, a model is built using deep learning and the noise in the signal is reduced and suppressed by means of a model named Denoising Bidirectional Recurrent Neural Network (DBRNN) as proposed in [1]. In [1], a three-layered model is proposed to detect 11 modulations like ASK, PSK, FSK, AM, SSB, DSB,VSB, FM, n-QAM. The classification is done by means of the instantaneous features of the signal like instantaneous amplitude, phase and frequency, as described in [2].

In [3], the authors emphasize the vital significance of high-quality datasets for model development, testing, and evaluation and uses deep learning to modulation recognition for the first time. A deep neural network-based modulation recognition technique is shown in [4]. In order to maximize the number of neural network nodes and adaptively choose the ideal number of hidden nodes for the network, the program employs the particle swarm optimization technique. This increases the modulation recognition accuracy in environments with multipath interference. In [5], [6], a convolutional neural network-based modulation recognition technique is shown. To improve recognition performance, the algorithm makes use of the convolutional neural network and modulation signal constellation diagram.

From [7], [8], [9], [10], [11], [12], it is evident that the approach has very little computing cost, however noise significantly affects the recognition rate. The low SNR environment will see a significant decline in performance as a result. Therefore, a blind filtering multitude-based technique is adopted to mitigate the noise. For eleven different types of digital modulation signals, a classifier based on a feature selection technique tree-shaped smooth support vector machine is created. According to the simulation results, for the signals with SNR greater than -1dB, the classifier's identification rate surpasses 97%. suggests a novel approach to modulation classification that enhances feature extraction from the recognition modulation by utilizing a decision tree classifier to identify the appealing relationship between high-order cumulants has been demonstrated that these methods have a low computing complexity and good anti-noise performance. Nevertheless, the performance remains inadequate in low SNR environments and all approaches necessitate manually derived features. Deep learning is a useful method for automatically extracting a variety of complex features from the original data. Its outstanding capacity for self-learning and nonlinear mapping has led to its widespread application in modulation recognition. It can be used in conjunction with basic features to automatically identify more intricate features through a series of nonlinear transformations. Neural networks are capable of solving complex classification problems, such modulation recognition in low signal-to-noise ratio environments, because of their nonlinear mapping capability.

2 Data Model and Description

2.1. About Dataset

For the proposed work, “DeepSig Dataset: RadioML 2018.01A” from RF Datasets for Machine Learning by Deepsig Inc., from the Kaggle website is choosen. The dataset is of Hierarchical Data Format (HDF5) which is to write a huge quantity of data in such manner that they don’t overwrite themselves due to space concerns. The dataset has the following characteristics:

1) There are totally 24 modulations present in the dataset which are variants of shift keying methods, m-ary QAM, sideband suppressed single and double sidebands of AM, variants of Binary Phase and Amplitude Shift Keying and Frequency Modulation.
2) The dataset consists of 2,555,904 frames. Each frame has three objects, namely X, Y and Z component.
3) The X component is the actual signal itself as pairs of reference (I) and 90 degree shifted reference samples (Q). The Y component contains the modulation applied to the signal in one-hot encoded form. The index of the '1' is mapped to the modulation string from the array above. The Z component is the SNR value of the signal.
4) There exists 26 SNRs per modulation type: from -20dB to 030dB in steps of 2dB.
5) The dataset contains 4096 frames per modulation-SNR combination. For instance, there are 4096 frames in which FM modulation is applied at -10dB.
6) In each frame, there are 1024 pairs of I and Q samples, thus each signal is of (1024,2) shape.

For ease of understanding, the images of each object {‘X’,’Y’,’Z’} are attached under the section Results for further understanding.

Since the dataset is of type *hdf5* file, the tool ‘*HDFView*’ is used to visualize the dataset by viewing the objects as a whole Excel sheet-like window.

3 Network Formulation

3.1 Data Preprocessing

Since our dataset is abundant with signals, the dataset has to be split separately for the training, validating and testing process. The ratio of train ,validation and test is 80:10:10. The splitting has to be done for each SNR level so that each level signal is tested and the corresponding accuracy for each SNR level is noted as such.

For doing recognition of the signal based on modulation, the signal has to be denoised in order to effectively do the classification process, which will be discussed in a later section.

3.2 Network Structure

As described in [1], a three-layered model is proposed with some modifications applied which will be discussed later. These three layers are as follows:

- The first layer is the signal reconstruction layer which takes the given input signal and suppresses the noise present in it with the help of a bidirectional RNN.
- The second layer is the feature reconstruction layer which performs the noise suppression again in order to extract the features from the signal for classification purposes.
- The third layer performs the classification of the signals based on the modulation with the help of the extracted features from the second layer.

3.3 Signal Reconstruction Layer

In this layer, denoising is performed with the help of a Bidirectional RNN. This BRNN consists of neurons that are bidirectional by means of a hidden layer in between the input layer and the output layer, which is shown in Figure. 1. It shows the relationship between the input neuron x_i to the output neuron y_i through the hidden states *A* and *A'*, which can be mathematically expressed as:

Ai = f(Ux+WAi-1) (1)

in which, U represents the weight matrix from the input layer to hidden layer, x is the input signal vector, W stands for the weight matrix between the previous state Ai-1 and the current state Ai , which is a weight matrix of recursion and f(.) is the sigmoid activation function.

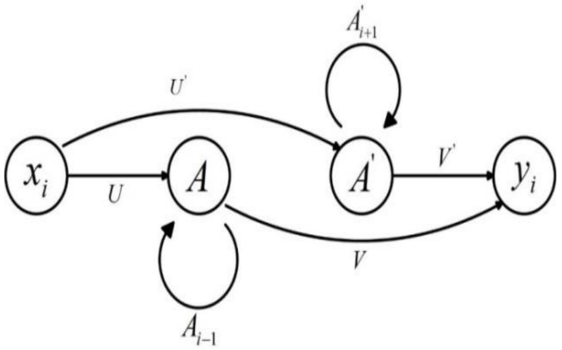


Fig 1: Structure of a single neuron in DBRNN
"Source:created by the authors"

Equation (1) tells about the forward processing of the input to the hidden state from the input layer. However, (1) alone is insufficient to explain the bidirectional nature of the structure due to the temporal correlation of the signal. Due to this, the signal propagates from the hidden layer to the input layer, which is explained as:

Ai' = f(Ux'+W'Ai+1) (2)

where, U' is the weight matrix from the hidden layer to the input layer, W' stands for the weight matrix between the next state Ai+1 and the current state Ai'. The deployed DBRNN structure is shown in Figure. 2.

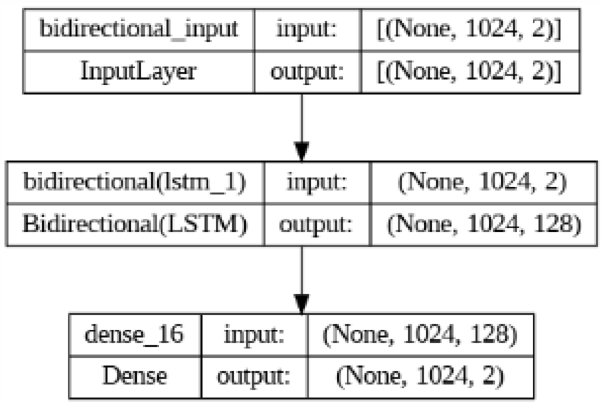


Fig. 2: Network Structure of proposed DBRNN
"Source:created by the authors"

3.4 Feature Reconstruction Layer

The signal, reconstructed after the first layer, exposes key information like amplitude, frequency, and phase. However, residual noise persists, affecting internal feature representation. This diminishes recognition accuracy and compromises the signal's representativeness. Thus, the residual noise has to be removed for classification. The input to this layer is the output of the signal reconstruction layer. The working process is described in Figure. 3.

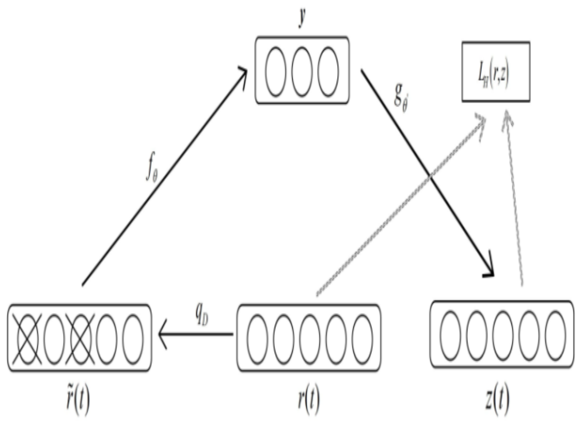


Fig 3: Structure of the feature reconstruction layer
"Source:created by the authors"

Fig. 3. is depicted as follows: the input r(t) is damaged by the process qD(t), which randomly sets some points in the signal to either 0 or 1, resulting in a damaged signal ŕ(t). This can be achieved by quantization, which is shown as:

ŕ(t) = qD(t) (3)

After this, the damaged ŕ(t) is compressed to a small size data using the compression process f0, resulting in an encoded vector y. The signal is again reconstructed by the expansion process g0 and z(t) is the obtained output. This series of operations is mathematically expressed as,

y(t) = f0(ŕ(t)) (4)

z(t) = g0(y(t)) (5)

The pair r(t) and z(t) are used as a learning pair LH(r,z) to train this layer. By choosing the desired number of bits to be quantized and size reduction, the residual noise present in the signal will be suppressed so that the resulting signal is used for classification. The hidden layers and their size are tabulated in Table 1.

Table 1: Feature Reconstruction Layer Details

Layers	Input size	Output size
Input Layer	(1024,2)	(1024,2)
Quantization Layer	(1024,2)	(1024,2)
Conv1D	(1024,2)	(1024,64)
MaxPooling1D	(1024,64)	(512,64)
Conv1D	(512,64)	(512,32)
MaxPooling1D	(512,32)	(256,32)
Conv1D	(256,32)	(256,32)
UpSampling1D	(256,32)	(512,32)
Conv1D	(512,32)	(512,64)
UpSampling1D	(512,64)	(1024,64)
Conv1D	(1024,64)	(1024,2)

“Source:created by the authors”

3.5 CNN Layer

After removing the noise from the signals, the signal is fed directly to the convolutional neural network in order to classify the signals based on modulation. The role supposed to be filled by CNN has to deal with the modalities that are already in the layers that are above it in the context of the basic classification of modulation type in the context of the action, that is, of modulation recognition. One advantage of the later is the definitive choice in selection of the modulation types through the process of scanning the input in succession with a series of convolution filters which are responsible for the separation of spatial and temporal patterns. . These filters, referred to as kernels, perform the mapping of the feature maps from the bottom to the altitude of the neural network, where the higher-dimensional information is retrieved through the summation at the element and multiplication in each layer. The information about this CNN classification layer is shown in Table 2.

Table 2 - CNN Layer Details

Layers	Input size	Output size
Input layer	(1024,2)	(1024,2)
Conv1D	(1024,2)	(1022,32)
Dropout	(1022,32)	(1022,32)
MaxPooling1D	(1022,32)	(511,32)
Conv1D	(511,32)	(509,32)
Dropout	(509,16)	(509,16)
MaxPooling1D	(509,16)	(254,16)
Flatten	(254,16)	4064
Dense	4064	32
Dense	32	24

“Source:created by the authors”

The Dropout layer used in neural networks serves as a measure to prevent overfitting by randomly demolishing certain neurons during the training process finally leading to powerful feature learning. It adds diversity of interpretation and resistance to noise using which the process of modulation recognition can be made reliable. The loss function used here is the “categorical cross entropy” function, which calculates the loss due to the misclassified data, thus indicating the degree of learning. This loss can be used to train the model up to the optimal value. The formula used to calculate the categorical cross-entropy loss is represented as:

$$\text{Loss} = -(m-1) \sum \sum y_{ij} \log(n_{ij})$$

where, m denotes the number of samples (number of signals to be classified in our case), C denotes the number of classes and the variable y_{ij} equals 1if the signal belongs to the class j or 0 if the sample does not belong to the class. The outer summation is for i up to m and the inner summation goes through C.

4 Results and Discussion

This section provides the performance of the individual layers as well as the overall structure and compares the results of various modules. The dataset is a “hierarchical data format” (.hdf5) in which each frame consists of In-phase and Quadrature-phase samples. These samples are viewed with HDFViewer tool.

Moving to the performances of the structure, the three layers are tested with varying parameters and the corresponding accuracy or loss are obtained, which are discussed in this section. The relation between epoch (iteration times) and test accuracy in various SNR situations are observed. From the graphs shown under Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8, it is understood that when the number of iterations grows, the test accuracy of the algorithm steadily increases, and when the number of iterations exceeds a particular threshold, the method's accuracy varies and becomes saturated within a short range. When the epoch (iteration) is the same, the test accuracy improves as the SNR increases. For SNR of -10 dB, the proposed model's test accuracy approaches 70%. When the SNR is 0 dB, the suggested algorithm has a test accuracy of more than 80%. This demonstrates that, when compared to previous methods, the proposed technique achieves higher recognition accuracy in low SNR environments.

Signal Reconstruction layer deploys Long Short-Term Module (LSTM), which is an advanced module for RNN. The output dimensionality is varied for different values and the corresponding accuracy is noted. These values are tabulated in Table 3.

Table 3 Output Dimensionality vs Memory Space and Accuracy

Output Dimensionality	Memory Space (KB)	Accuracy after 200 epochs
16 Units	4.88	57%
32 Units	17.76	65%
64 Units	67.51	93.57%
128 Units	263.01	94%
256 Units	1010	94.04%

“Source:created by the authors”

From Table 3, it is evident that the accuracy of the layer does not exceed 70% if the output dimensionality is less than 32 units. Once the size equals or greater than 64 units, greater than 93% accuracy is obtained. For choosing the appropriate unit size, the occupied memory space is taken into account. Occupied memory space is the space taken by the trainable parameters, such as the weight matrix, bias value and so on. After 64 units, the accuracy does not improve drastically. Thus, the bidirectional RNN with LSTM of 64 units is chosen as the optimal one.

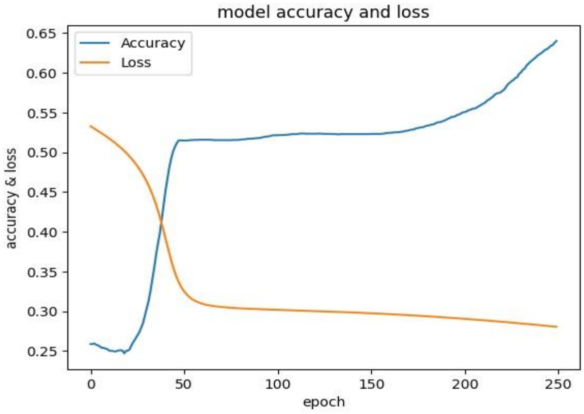


Fig. 4: Training Accuracy and Loss for 16 units
“Source:created by the authors”

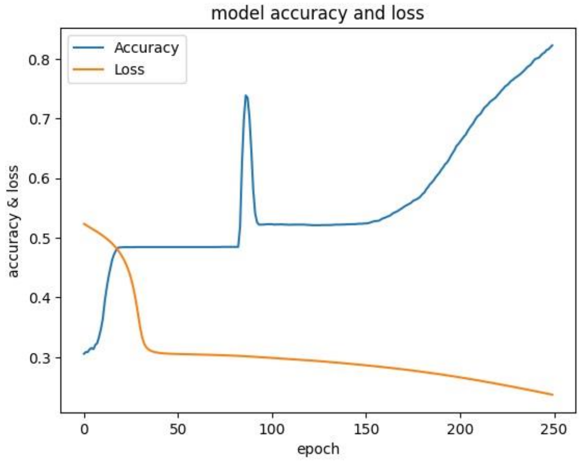


Fig. 5: Training Accuracy and Loss for 32 units
“Source:created by the authors”

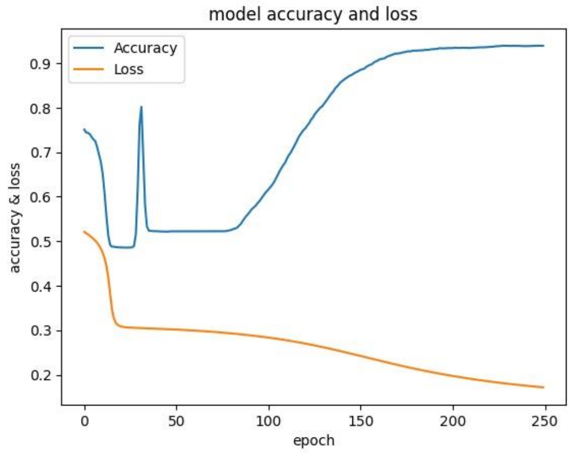


Fig. 6: Training Accuracy and Loss for 64 units
“Source:created by the authors”

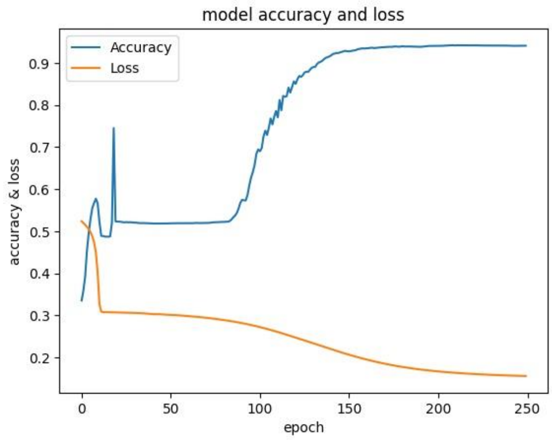


Fig. 7: Training Accuracy and Loss for 128 units
“Source:created by the authors”

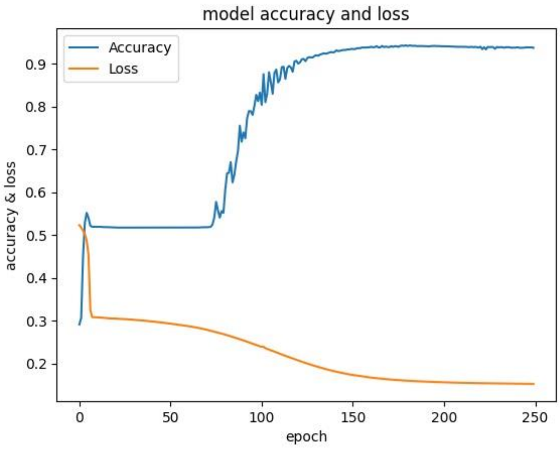


Fig. 8: Training Accuracy and Loss for 256 units
"Source:created by the authors"

In the feature reconstruction layer, the quantization process and the compression process are performed in the signal data. For quantization, number of bits retained after quantization is varied in order to effectively reduce the loss of this layer, which is shown in Table 4.

Table 4 Number of Retained Bits vs Loss

Bits to be retained after Quantization	Loss
24 bits	27.86%
20 bits	27.83%
16 bits	38.60%

"Source:created by the authors"

Each sample is a 32-bit floating-point number. When the number of quantized bits is less than 8, there is no such significant change and the value remains almost the same. When 8 bits are quantized, the value is changed in a notable way. If half bits are quantized, i.e., 16 nits are quantized, the loss obtained is more than 38%, which is not desired.

Thus, 20 bits are chosen to be retained after quantization. The size reduction is done in various patterns and the corresponding loss is noted in Table 5. By seeing Table 5, the optimal one, i.e., the first sequence is chosen as it exhibits minimal loss. For the classifying CNN layer, various models are trained and tested, and the corresponding accuracy results for three different cases (20 dB, 0 dB and -20 dB) are shown in Table 6.

Table 5 Size Comparison vs Observed Loss

Size comparison	Loss observed
1024 > 64 > 32	27%
1024 > 128 > 64 > 32	≈ 50%
1024 > 64 > 32 > 16	≈ 50%

"Source:created by the authors"

Table 6 Accuracy for Different Levels vs Models

Model	Accuracy		
	SNR=0 dB	SNR= - 20 dB	SNR=20 dB
CNN(Proposed)	82%	80%	91%
FC	45%	22%	65%
KNN	21.6%	11%	41.6%
SVM	24.3%	21.2%	33%

"Source:created by the authors"

The recognition is done in the last layer in which 24 modulations are classified. The confusion matrix for overall modulations is shown in Figure 9.

This work has considered three different scenarios (0 dB, - 20 dB and 20 dB) and their corresponding confusion matrices are shown in Figure 10, Figure 11 and Figure 12, respectively.

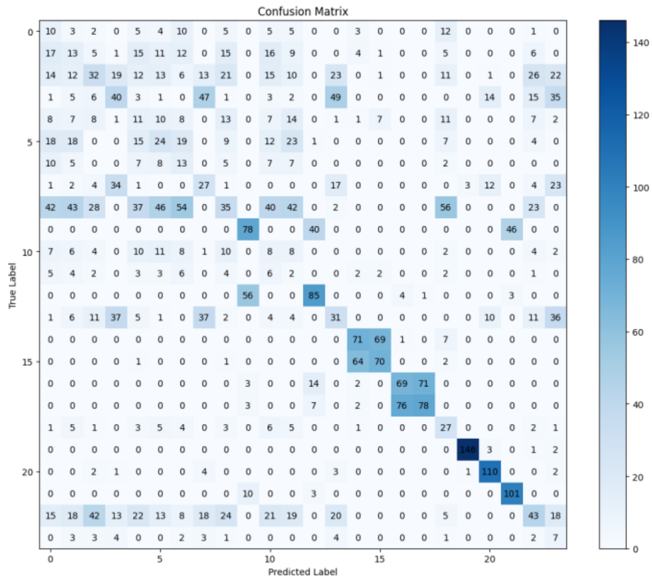
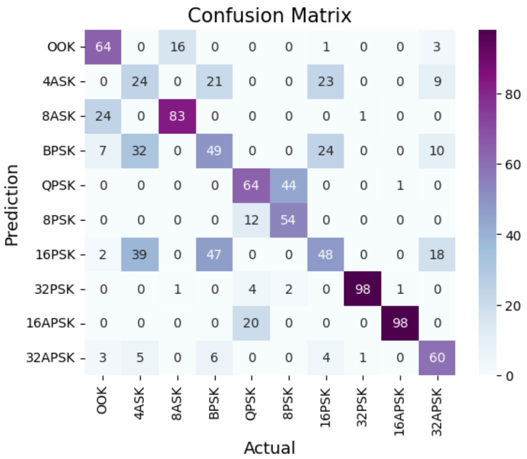


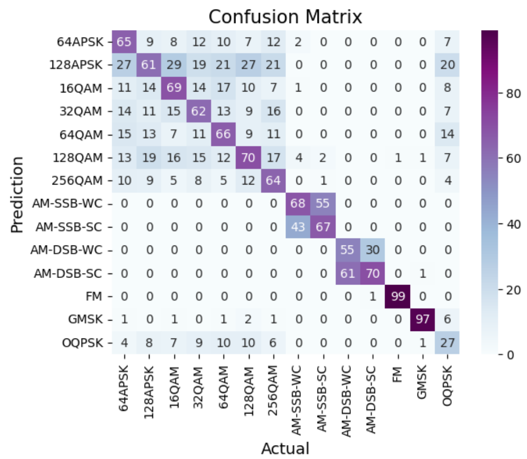
Fig. 9: Confusion Matrix of denoised signals under 24 modulations
"Source:created by the authors"

As observed from the above confusion matrices, the signals are classified under AM-SSB-SC, AM-SSB-WC, AM-DSB-WC, AM-DSB-SC, FM, GSK and OQPSK in group 2. This may be due to their closely related properties with the number of bits being used as the only difference among them. Apart from that, the diagonal elements in each of the classes are far more than the other elements in a given row of a class. Thus, we can say that the designed model works in a better manner.

From noticing the confusion matrices, it is understood that the APSK signals and the QAM signals tend to misclassify due to their similar properties. Also it is inferred, the QAM and APSK tend to misclassify due to their closely associated properties. By training the model for a desired number of iterations, the classification can be done at a relatively better rate. Thus, the proposed model works as expected by classifying the signals based on modulation.

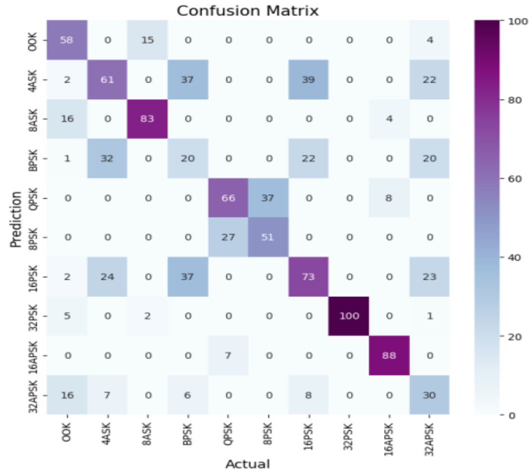


(a)

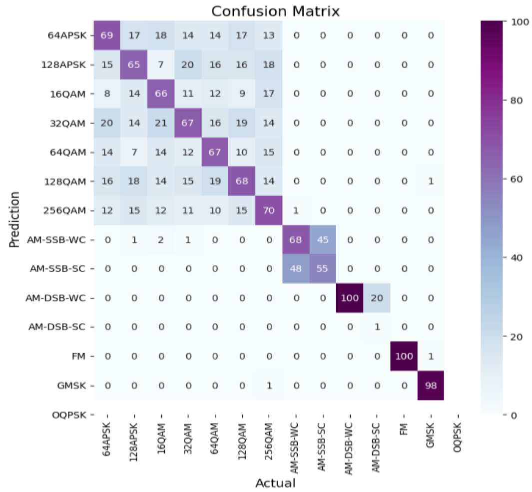


(b)

Fig 10: Confusion matrix with SNR = 0 dB
"Source:created by the authors"

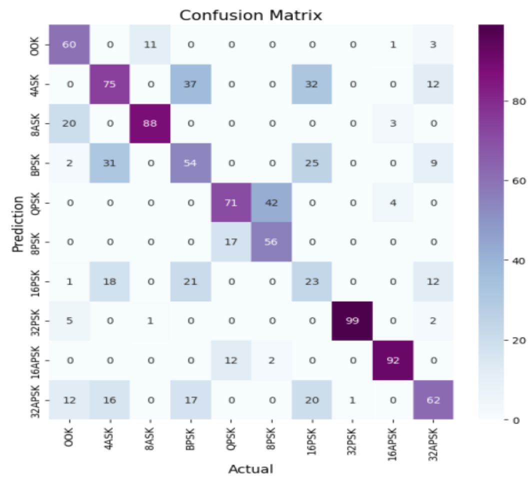


(a)

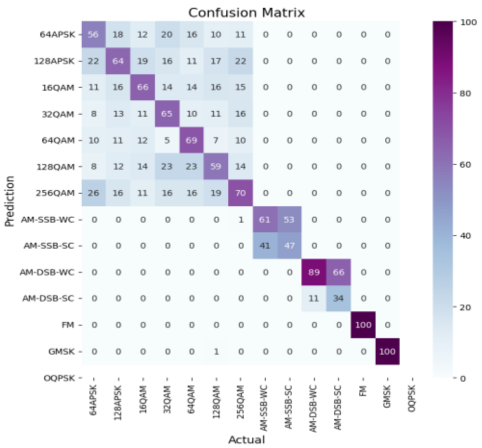


(b)

Fig. 11: Confusion matrix with SNR = -20 dB
"Source:created by the authors"



(a)



(b)

Fig. 12: Confusion matrix with SNR = 20 dB
"Source:created by the authors"

5 Conclusions

The DBRNN-based modulation recognition approach is put forward in the proposed work, which is known for having lower recognition rates in low signal-to-noise ratio (SNR) conditions. It reconstructs the signals, eliminates noise, and carries out feature optimization to enable effective classification. It achieves superior performance compared to known algorithms in noise-limited scenarios, an achievement illustrated in simulations. The method relies on DBRNN's memory capacity, which lets it identify different correlations in signals. These, in turn, improve its accuracy. As a result, it gives rise to a feasible approach of noise handling and classification improvement with effectiveness under extreme circumstances. The proposed work can be extended to cryptographic algorithms and hashing in order to do the process with additional security measures, so that concerns like eavesdropping and masquerading problems can be addressed with improvised solutions.

References

[1] Du, R., Liu, F., Zhang, L, Ji Y., Xu J., Gao F., Modulation Recognition Based on Denoising Bidirectional Recurrent Neural Network. *Wireless Pers Commun* **132**, 2437–2455 (2023). https://doi.org/10.1007/s11277-023-10725-5

[2] E. Moser, M. K. Moran, E. Hillen, D. Li and Z. Wu, "Automatic modulation classification via instantaneous features," *2015 National Aerospace and Electronics Conference (NAECON)*, Dayton, OH, USA, 2015, pp. 218-223, doi: 10.1109/NAECON.2015.7443070.

[3] S. Aer, M. Diao, X. Zhang and X. Zhang, "Automatic Modulation Classification Strategy Based on Novel Feature Processing," *2022 7th International Conference on Signal and Image Processing (ICSIP)*, Suzhou, China, 2022, pp. 74-80, doi: 10.1109/ICSIP55141.2022.9886359

[4] Y. Wang, M. Liu, J. Yang and G. Gui, "Data-Driven Deep Learning for Automatic Modulation Recognition in Cognitive Radios," in *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4074-4077, April 2019, doi: 10.1109/TVT.2019.2900460.

[5] OSHEA T J, WEST N, "Radio machine learning dataset generation with GNU radio", Vol.1, No.1, Proceedings of the GNU Radio Conference, USA (2016).

[6] Z. Qu, C. Hou, C. Hou, and W. Wang, "Radar signal intra pulse modulation recognition based on convolutional neural network and deep Q-Learning network", IEEE Access, Vol.8, 49125–49136(2020). DOI: 10.1109/ACCESS.2020.2980363.

[7] Yu W, Miao L, "Data driven deep learning for automatic modulation recognition in cognitive radios", IEEE Transactions on Vehicular Technology. 68(4), 4074-4077(2019). doi: 10.1109/TVT.2019.2900460.

[8] J Shi, S Hong, C Cai, Y Wang, H Huang, G Gui, " Deep Learning-Based Automatic Modulation Recognition Method in the Presence of Phase Offset", IEEE Access, 8, 42841 – 42847(2020). doi: 10.1109/ACCESS.2020.2978094

[9] W Shi, D Liu, X Cheng, Y Li, Y Zhao, "Particle swarm optimization-based deep neural network for digital modulation recognition", IEEE Access. 7, 104591-104600(2019) doi: 10.1109/ACCESS.2019.2932266.

[10] Ivan A, Jafar W, Abdul Sadah, Thamir R. "Recognition of QAM signals with low SNR using a combined threshold algorithm", IETE Journal of Research. 61(1), (2015).

[11] Swami A, Sadler B M, "Hierarchical digital modulation classification using cumulants", IEEE Transactions on Communications. 48(3), 416-429(2000) doi: 10.1109/26.837045.

[12] Ali AK, Erçelebi E. Algorithm for automatic recognition of PSK and QAM with unique classifier based on features and threshold levels. ISA Trans. 2020 Jul;102:173-192. doi: 10.1016/j.isatra.2020.03.002. Epub 2020 Mar 5. PMID: 32169291

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article