Digital Modulation Classifier

by

Rohit Khandal 202011064

Alex Tamboli 202011071

Tushar Agarawal 202011074

Nishesh Jain 202011050

Vivek Borole 202011018

Department of Computer Science and Engineering

Indian Institute of Information Technology, Vadodara - International Campus Diu, INDIA

November 2022

Contents

1 Introduction

Digital communication systems continue to develop, with improvements to radio spectrum usage efficiency becoming vital. To respond to this requirement, dynamic spectrum access (DSA) is an important tool requiring spectrum sensing and signal classification. In this case, the modulation classification performs a significant role and can be widely employed in a variety of applications, such as software defined radio system and radar communication in the military. There is a high demand for radio frequency (RF) bands. In a crowded spectrum situation, the automatic modulation classification (AMC) technology can respond to the requirements optimising signal detection and subsequent demodulation when multiple complex/unknown signals are to be handled, or for cognitive radio primary-user detection.AMC can also be used by cognitive radio systems to identify the presence of primary users.

1.1 Literature Survey

This work concentrates on modulation classification provided by different methods in the deep learning area. Signals can appear like noises at very low signal to noise ratio (SNR). The purpose of this work is to develop the accuracy of classification in the low SNR area with more efficient methods. With the improvement of Artificial Intelligence (AI) technology, many areas have achieved new progress by this novel area. For the traditional statistical methods of machine learning, classification of the modulation types by the statistical features of the signals is common. Based on the previous research, the CI model can classify the modulation types in a dynamic system without phase lock and frequency lock at low SNR levels.

In previous published research, Maximum-likelihood decision theory is used as a critical method. PSK and QAM signals are distinguished with accuracy of 90% above 9 dB SNR . For the pattern recognition algorithms, the features of signals are involved in the models, where high order cumulants play a critical role in the AMC algorithm. After extracting the efficient features, a support vector machine (SVM) is applied in the recognition process. The accuracy can reach 96 % at 10 dB SNR for 200 samples . However, the probability of correct classification is between 50% and 70% at circa 0 dB SNR, suggesting a need for more improvement. Compared to the previous classifiers, the binary hierarchical polynomial classifiers are also proposed with the probability of correct classification of 56% at 0 dB SNR with 1000 symbols. Hence, there is still the challenge to optimise the AMC system, especially at low SNR. .

1.2 Objectives and Scope of work

Automatic modulation classification (AMC), which aims to blindly identify the modulation type of an incoming signal at the receiver in wireless communication systems, is a fundamental signal processing technique in the physical layer to improve the spectrum utilization efficiency. Motivated by deep learning (DL) high-impact success in many informatics domains, including radio signal processing for communications, numerous recent AMC methods exploiting deep networks have been proposed to overcome the existing drawbacks of traditional approaches. DL is capable of learning the underlying characteristics of radio signals effectively for modulation pattern recognition, which in turn improves the modulation classification performance under the presence of channel impairments. In this work, we first provide the fundamental concepts of various architectures, such as neural networks, recurrent neural networks, long short-term memory, and convolutional neural networks as the necessary background. We then convey a comprehensive study of DL for AMC in wireless communications, where technical analysis is deliberated in the perspective of state-of-the-art deep architectures. Remarkably, several sophisticated structures and advanced designs of convolutional neural networks are investigated for different data types of sequential radio signals, spectrum images, and constellation images to deal with various channel impairments. Finally, we discuss some primary research challenges and potential future directions in the area of DL for modulation classification.

2 System Model

In this section, the model of the received signal, the constellations and the statistical features are introduced. The constellations of the modulated signals are utilised for the DL model.

2.1 Signal Model

The received signal in baseband is defined by r(t) and it is given by (1):

$$(t) = s(t) + n(t) \tag{1}$$

Here, (t) is the original signal which transmits through the additive white Gaussian noise (AWGN) channel, and $n(t)$ represents the noise applied to the signals.

SNR is also an important metric to describe the noise and represent the signals in the real world, here defined by (3):

$$SNR = powerofsignal/varianceofnoise \tag{2}$$

In this work, signals are considered from -10 to 20 dB SNR. Note that the focus of previous works was on the classification above 10 dB SNR [20]. SNR values less than 0 dB are very important and should also be tested, to improve the performance in distinguishing between modulation types. We have developed our models from -10 to 20 dB SNR to provide a comprehensive dataset.

2.2 Modulation Model

The process of converting into electrical/digital signals for transferring that signal over a medium is called modulation. It increases strength for maximum reach of the signals. The process of extracting information/data from the transmitted signal is called demodulation. A Modem is a device that performs both modulation and demodulation processes. The various forms of modulation are designed to alter the characteristic of carrier waves. The most commonly altered characteristics of modulation include amplitude, frequency, and phase.

Carrier signal: The signals which contain no information but have a certain phase, frequency, and amplitude are called carrier signals.

Modulated signals: The signals which are the combination of the carrier signals and modulation signals are modulated signals. The modulated signal is obtained after the modulation of the signals.

2.2.1 Modulation Types

The two types of modulation: analog and digital modulation techniques have already been discussed. In both the techniques, the baseband information is converted to Radio Frequency signals, but in analog modulation, these RF communication signals are a continuous range of values, whereas in digital modulation these are prearranged discrete states.

Analog Modulation:

1. Amplitude modulation (AM)
2. Frequency modulation (FM)
3. Phase modulation (PM)

Digital Modulation:

1. Binary phase shift keying
2. Phase-shift keying
3. Differential quadrature phase shift keying
4. Offset quadrature phase shift keying
5. Frequency Shift Keying
6. Minimum shift keying
7. Gaussian minimum shift keying
8. Trellis coded type of modulation

2.2.2 Modulation Constellation

For the DL model, the main principle applied is that of image classification. In applying DL, the modulated signals are transferred to the CI representations. We feed the network significant datasets, which are captured by increasing number of samples in each constellation. Figure 1 shows the constellations collected from the lab as examples of the four modulation types (BPSK, QPSK, 8PSK, QAM16) at 10 dB SNR.

To detect these kinds of constellations, we divided the data into small groups which can be observed in Figure 2. To improve the accuracy of classification, we also gradually increase and plot the samples from 200 to 500 samples in each small group which can be observed in Figure 2.

2.3 Experimental Setup

2.3.1 Where to get dataset

You can get the dataset from below link;-
https://www.kaggle.com/datasets/dimitrisf/constellation-dataset
Content:-
The first line contains the modulation type, the order of the modulation, the SNR and the rotation of the installation points. The new rows contain the I and Q coordinates of the constellation images. There are around 135k constellations of 9 classes (PAM4, PAM16, QPSK, PSK8, APSK32, APSK64, QAM16, QAM32, QAM64) and some additional testing constellation samples.

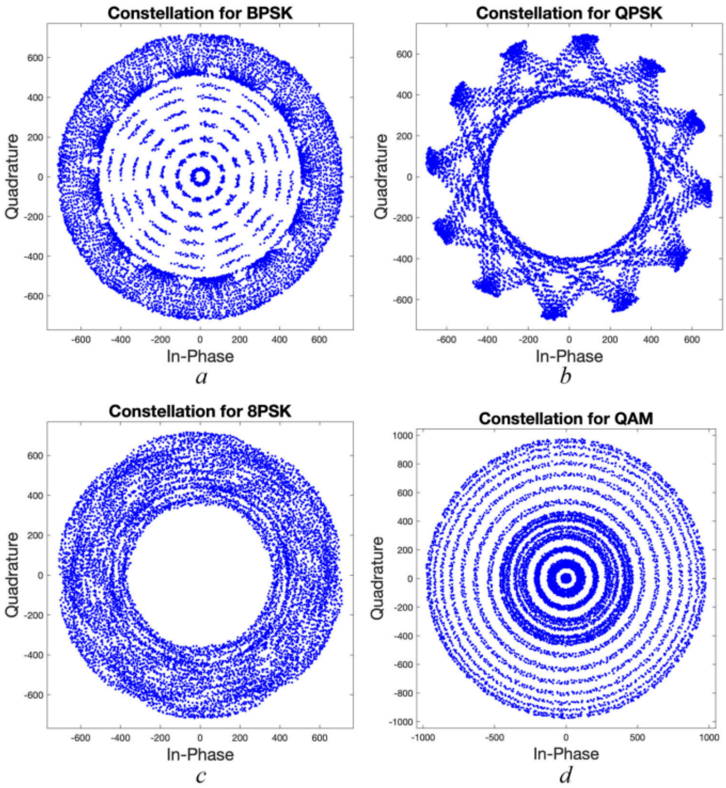

Figure 1: Lab signals collected from SDR without phase and frequency lock (a) BPSK, (b) QPSK, (c) 8PSK, (d) QAM16

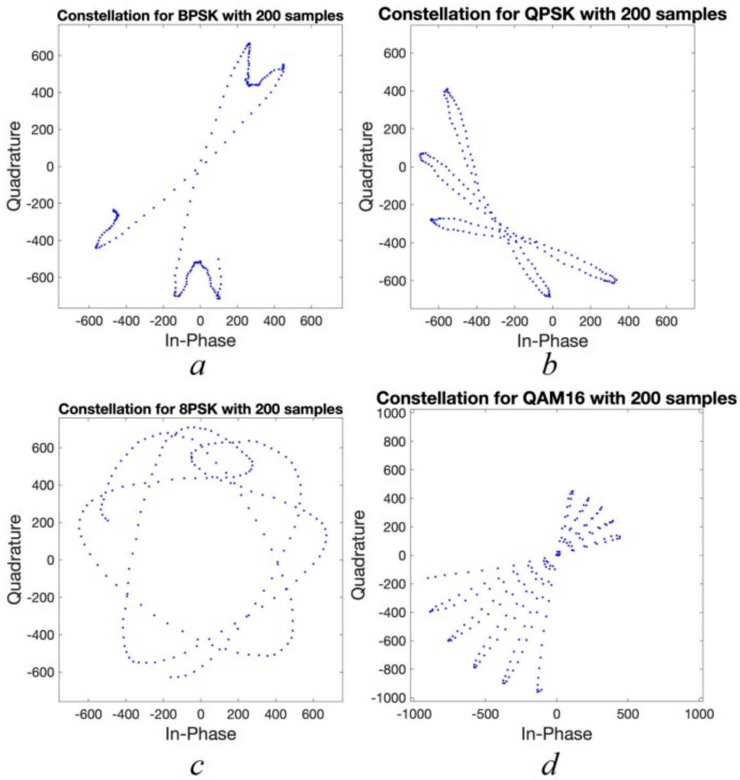

Figure 2: Constellations of signal samples in small groups (a) BPSK, (b) QPSK, (c) 8PSK, (d) QAM16

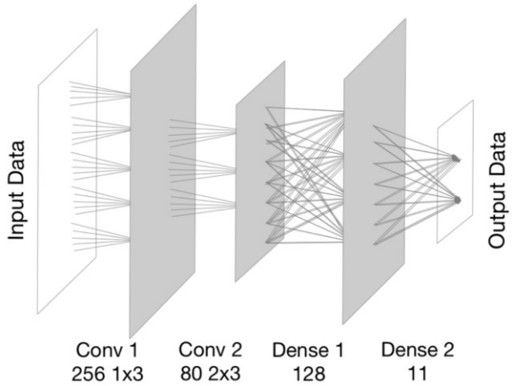

Figure 3: Two-convolutional-layer model

2.3.2 How to run the code:-

To get the code to gtihub link below:-
https://github.com/Rohit2kh/AMC–Project
then download requirement.txt and upload the dataset whereever you want to run the code and then run specific neural network technique like CNN, ResNet , CLDNN. It can take time hours to train the model ,different-different time will be taken for different-different techniques.

3 Neural Network Architecture

Various types of neural network architectures have been studied for image classification tasks, which are robust to the images rotation, occlusion, scaling and other noise conditions. Therefore, we applied several neural networks here to improve the blind modulation classification task which faces similar feature variations. We randomly choose half of the 1200000 examples for training and the other half for testing in each experiment.

3.1 CNN

3.1.1 Architecture

CNNs are feed forward neural networks that pass the convolved information from the inputs to the outputs in only one direction. Modules stack on top of each other and form a deep network. Either one or two fully connected layers follow the convolutional modules for the final outputs. Based on the framework proposed in [1], we build a CNN model with similar architecture but different hyper-parameters (Figure 3).

In this pipeline, the raw vectors are input directly into a convolutional layer consisting of 256 filters that have the size of 1×3 each. Each filter convolves with 1×3 elements in the input vector and slides one step to the next 1×3 elements. Outputs of the first convolutional layer are then fed into the second convolutional layer that utilizes 80 filters with the size of 2×3. The outputs of the convolutional module are passed to the fully connected layers with 128 neurons and 11 neurons, with respect to order.

In our model, all layers before the last one use rectified linear (ReLU) functions as the activation functions and the output layer uses the Softmax activation function to calculate the predicted label. ReLU was proposed by Nair and Hinton [56] in 2010 and popularized by

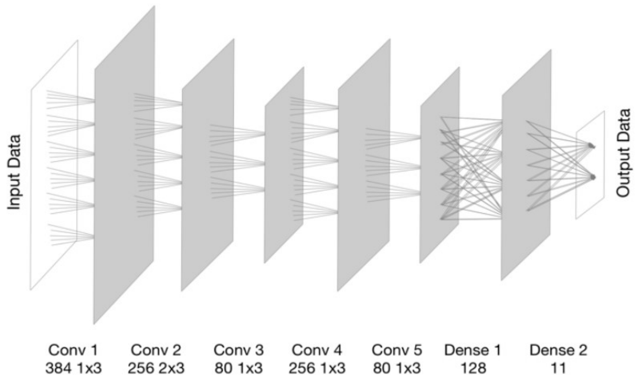

Figure 4: Five-convolutional-layer model

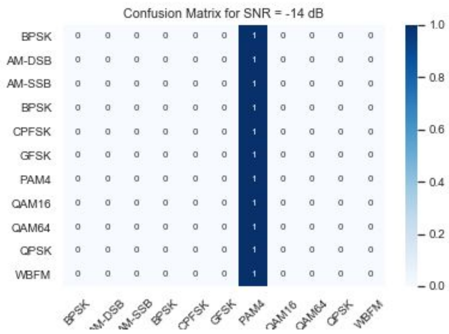

Figure 5: Confusion matrix at -18dB SNR

Krizhevsky et al. [52]. The ReLU is given by f(x) = max(x, 0), a simplified form of traditional activation functions such as sigmoid and hyperbolic tangent. The regularization technique to overcome overfitting includes normalization and dropout. Here, we set the dropout rate to 0.6 so that each hidden neuron in the network would be omitted at a rate of 0.6 during training. During the training phase, each epoch takes roughly 71s with the batch size of 1024. We do observe some overfitting as the validation loss inflects as the training loss decreases. We set the patience at 10 so that if the validation loss does not decline in 10 training epochs, the training would be regarded as converging and end. The total training time is roughly two hours for this model with Adam [57] as the optimizer. The average classification accuracy for this model is 72% when the SNR is larger than 10dB. To further explore the relationship between the neural network architecture and success rate, we adjust the first model to a new one as illustrated in Figure 4.

We build a five-layer CNN model based on the one in Figure 4, but add another convolutional layer with 256 1×3 filters in front of the convolutional module. The average accuracy at high SNR is improved by 2%. The seven-layer CNN that performs best is produced by the architecture in Figure 4. As the neural network becomes deeper, it also gets harder for the validation loss to decrease. Most eight-layer CNNs see the validation loss diverge, and the only one that converges performs worse than the seven-layer CNN. The training time rises as the model becomes more complex, from 89s per epoch to 144s per epoch.

Figure 6: Confusion matrix at 0dB SNR

3.1.2 Results

We use 600000 samples for training and 600000 samples for testing. The classification results of our first model, four-layer CNN, is shown in forms of confusion matrix. In situations that signal power is below noise power, as for the case when the SNR is -18dB (Figure 5), it is hard for all neural networks to extract the desired signal features, while when SNR grows higher to 0dB, there is a prominent diagonal in the confusion matrix, denoting that most modulations are correctly recognized. As mentioned above, the highest average classification accuracy is produced by the CNN with four convolutional layers. In its confusion matrix (Figure 6), there is a clean diagonal and several dark blocks representing the discrepancies between WBFM and AM-DSB, QAM16 and QAM64, and 8PSK and QPSK. The training and testing data sets contain samples that are evenly distributed from -20dB SNR to +18 dB SNR. So we plot the prediction accuracy as a function of SNRs for all our CNN models. When the SNR is lower than -6dB, all models perform similar and it is hard to distinguish the modulation formats, while as the SNR becomes positive, there is a significant difference between deeper models and the original ones. The deepest CNN which utilizes five convolutional layers achieves 81% at high SNRs which is slightly lower than the 83.8% produced by the four-convolutional-layer model.

3.1.3 Discussion

Blank inputs or inputs that are exactly the same but with different labels can confuse neural networks since the neural network adjusts weights to classify it into one label. The misclassification between two analogue modulations is caused by the silence in the original data source. All samples with only the carrier tone are labeled as AM-DSB during training, so silence samples in WBFM are misclassified as AMDSB when testing. In the case of digital signal discrepancies, different PSK and different QAM modulation types preserve similar constellation maps so it is difficult for CNNs to find the different features.

For neural networks deeper than eight layers, the large gradients passing through the neurons during training may lead to having the gradient irreversibly perish. The saturated and decreasing accuracy as the depth of the CNN grows is a commonly faced problem in deep neural network studies. However, there should exist a deeper model when it is constructed by copying the learned shallower model and adding identity mapping layers. So we explored a new architecture as discussed below.

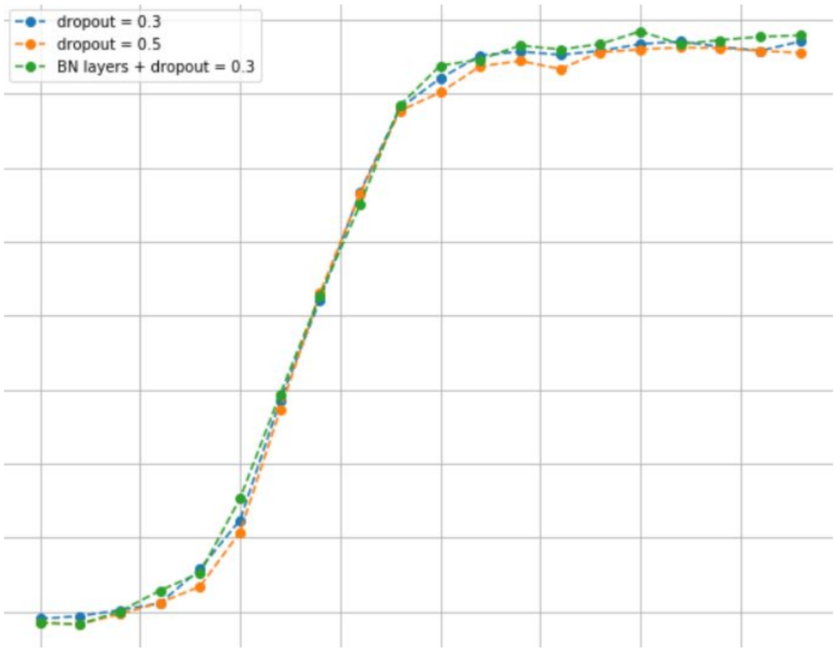

Figure 7: Classification performance vs SNR

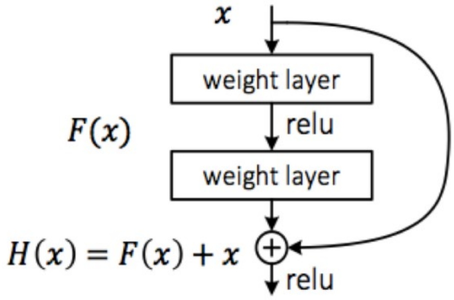

Figure 8: A building block of ResNet

3.2 ResNet

3.2.1 Architecture

Deep residual networks [59] led the first place entries in all five main tracks of the ImageNet [58] and COCO 2015 [60] competitions. As we see in the previous deep CNN training, the accuracy saturates or decreases rapidly when the depth of a CNN grows.

The residual network solves this by letting layers fit a residual mapping. A building block of a residual learning network can be expressed as the function in Figure 8, where x and $H(x)$ are the input and output of this block, respectively. Instead of finding the mapping function $H(x) = x$ which is difficult in a deep network, the ResNet adds a shortcut path so that it now learns the residual mapping function $F(x) = H(x) - x.F(x)$ is more sensitive to the input than $H(x)$ so the training of deeper networks becomes easy. The bypass connections create identity mappings so that deep networks can have the same learning ability as shallower networks do. Our neural network using the residual block is shown in Figure 9 It is built based on the six-layer CNN that performs best. Limited by the number of convolutional layers in the CNN model, we add only one path that connects the input layer and the third layer. The involved network

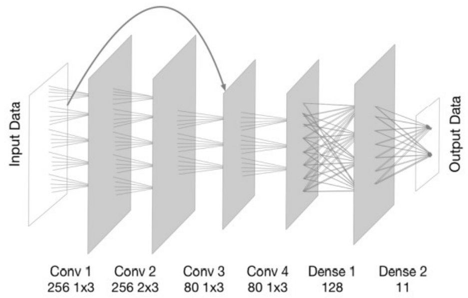

Figure 9: Architecture of six-layer ResNet

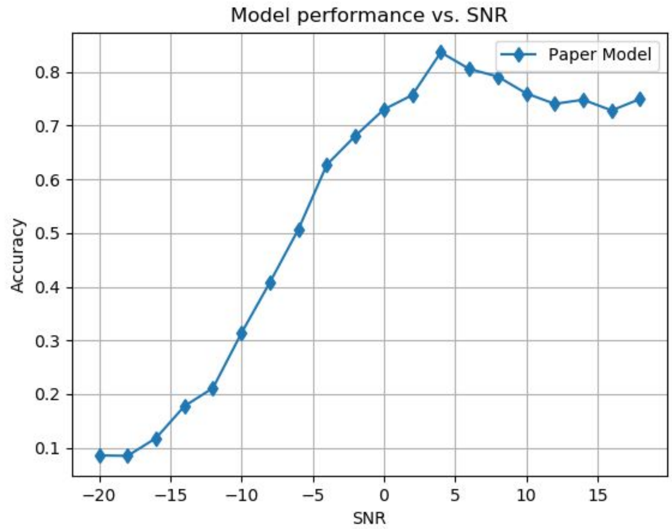

Figure 10: ResNet Model Accuracy vs SNR

parameters are increased due to the shortcut path so the training time grows to 980s per epoch.

3.2.2 Results

The classification accuracy as a function of SNR of the ResNet model displays the same trend as the CNN models. At high SNR, the best accuracy is 83.5% which is also similar to the six-layer CNN. However, when the depth of the ResNet grows to 11 layers, the validation loss does not diverge as the CNN model does, but produces a best accuracy of 81%.

3.2.3 Discussion

ResNet experiments on image recognition point out that the advantages of ResNet is prominent for very deep neural networks such as networks that are deeper than 50 layers. So it is reasonable that ResNet performs basically the same as CNNs when there are only six or seven layers. But it does solve the divergence problem in CNNs by the shortcut path. We tried another architecture that also uses bypass paths between different layers.

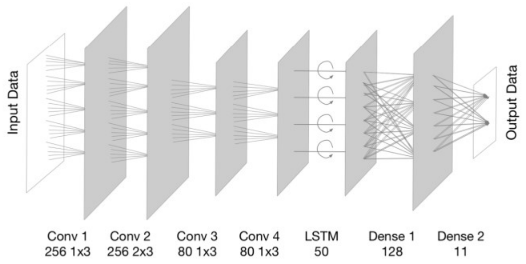

Figure 11: Architecture of the CLDNN model

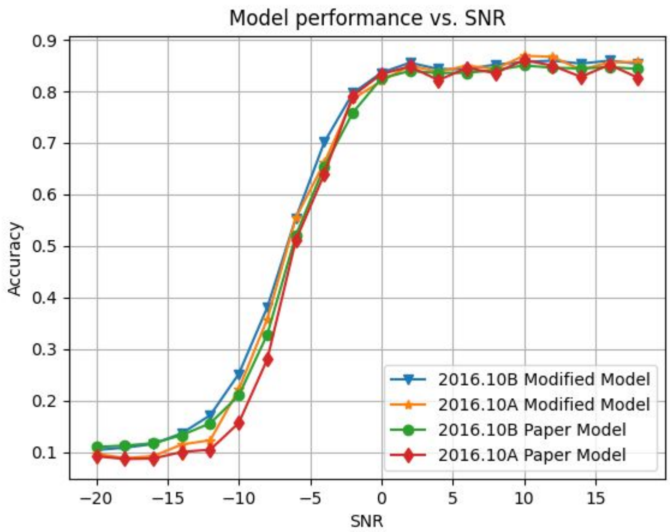

Figure 12: Architecture of the CLDNN model

3.3 CLDNN

3.3.1 Architecture

The Convolutional Long Short-Term Memory Deep Neural Network (CLDNN) was proposed by Sainath et al. [61] as an end-to-end model for acoustic learning. It is composed of sequentially connected CNN, LSTM and fully connected neural networks. The time-domain raw voice waveform is passed into a CNN, then modeled through LSTM and finally resulted in a 3% improvement in accuracy. We built a similar CLDNN model with the architecture in Figure 11, where a LSTM module with 50 neurons is added into the four-convolutional CNN. This architecture that captures both spacial and temporal features is proved to have superior performance than all previously tested architectures.

3.3.2 Results

The best average accuracy is achieved by the CLDNN model at 88.5%. In Figure 12, we can see that CLDNN outperforms others across almost all SNRs. The cyclic connections in LSTM extract features that are not obtainable in other architectures.

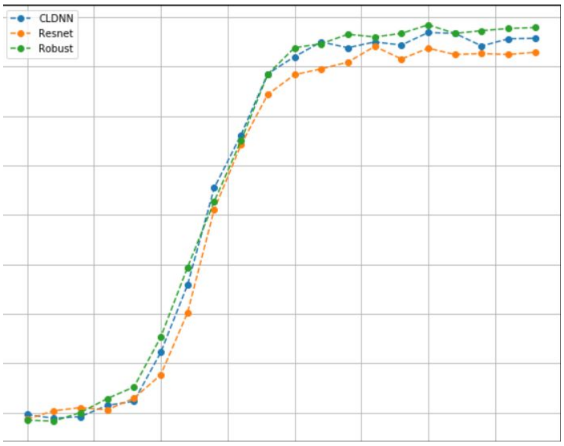

Figure 13: Architecture of the CLDNN model

3.3.3 Discussion

The CNN module in CLDNN extracts spacial features of the inputs and the LSTM module captures the temporal characters. CLDNN has been highly accepted in speech recognition tasks, as the CNN, LSTM and DNN modules being complementary in the modeling abilities. CNNs are good at extracting location information, LSTMs excels at temporal modeling and DNNs are suitable for mapping features into a separable space. The combination was first explored in [62], however the CNN, LSTM and DNN are trained separately and the three output results were combined through a combination layer. In our model, they are unified in a framework and trained jointly. The LSTM is inserted between CNN and DNN because it is discovered to perform better if provided higher quality features. The characteristic causality existing in modulated signals that is the same as the sequential relationship in natural languages contributes the major improvements of the accuracy.

Although there is a significant accuracy improvement for all modulation schemes in the confusion matrix of the CLDNN model, there are still few significant confusion blocks existing off the diagonal. The quantified measures for these discrepancies are formed . The confusion between WBFM and AM-DSB has the more prominent influence on the misclassification rate, but this is caused by the original data source and we cannot reduce it by simply adjusting neural networks. So we focus on improving the classification of QAM signals.

4 Conclusion

This thesis have implemented several deep learning neural network architectures for the automatic modulation classification task. Multiple classifiers are built and tested, which provide high probabilities of correct modulation recognition in a short observation time, particularly for the large range of the SNR from -20dB to +18dB. The trained models outperform traditional classifiers by their high success rates and low computation complexities. The CNN serves as a basic end-to-end modulation recognition model providing nonlinear mapping and automatic feature extraction. The performance of CNNs are improved from 72% [1] to 83.3% by increasing the depth of CNNs. ResNet were used to build deeper neural networks and enhance the information flow inside the networks. The average classification accuracy reaches 83.5% for ResNet . Although the best accuracies are limited by the depth of network, they suggest that the shortcut paths between non-consecutive layers produce better classification accuracies. A CLDNN model combines a CNN block, a LSTM block and a DNN block as a classifier that can automatically extract the spacial and temporal key features of signals. This model produces the highest accuracy for time domain IQ inputs and can be considered as a strong candidate for dynamic spectrum access systems which highly relies on low SNR modulation classifications. The two-layer LSTM model was proposed with different time domain format inputs. The results reach roughly 100% for all digital modulations. The experiments of time domain IQ and amplitude phase inputs also emphasize the importance of preprocessing and input representation. These models are capable to recognizing the modulation formats with various propagation characteristic, and show high real-time functionality.

5 Refrences

- [1]Xiaoyu Liu, Diyu Yang, and Aly El Gamal,
  ”Deep Neural Network Architectures for Modulation Classification”,Purdue University 2018
- [2]N. E. West and T. J. O’Shea,
  “Deep architectures for modulation recognition,” CoRR, vol. abs/1703.09197, 2017.
- [3]Sun Y., Ball E.A.: Automatic modulation classification
  using techniques from image classification. IET Commun. 16, 1303-1314 (2022).
- [4]Corinna, C.Vladimir:Support-Vector-Machine,mach.Learning 20273-(297)(1989)