Deep Learning Based Sub-Nyquist Modulation Recognition

Shuangning Li
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
lishuangning@bupt.edu.cn

Shaoqing Hu, Rajagopal Nilavalan
Department of Electronic and Electrical Engineering
Brunel University London
London, UK
shaoqing.hu@brunel.ac.uk,
Nila.Nilavalan@brunel.ac.uk

***Abstract*—In this paper, we designed a Convolutional Neural Network (CNN) for Sub-Nyquist modulation recognition and compare the performance Long Short-Term Memory (LSTM) network and Convolutional Long Short-term Deep Neural Network (CLDNN) respectively. Unlike conventional modulation recognition task that operates with Nyquist sampled rate, the network architectures for Sub-Nyquist modulation recognition were specifically designed with a certain number of neurons, layers, and other hyperparameters to effectively extract key features from Sub-Nyquist sampled signals and process larger volumes of data. The simulation results demonstrate that the CNN network has the best recognition accuracy of 98.01% on the GBsense dataset, followed by the CLDNN of 96.81% and LSTM of 87.51% respectively.**

Keywords*- Sub-Nyquist *Modulation Recognition, Convolutional Neural Network, Deep learning

I. Introduction

As 6G wireless communication technology is further upgraded and expanded, higher data rates and spectrum efficiency with lower latency will be required. In order to improve spectrum efficiency, dynamic intelligent spectrum sharing techniques are used to enable users to make full use of the spectrum resources. Modulation recognition plays an essential role in spectrum resource analysis and management, primarily implemented at the receiver to classification the modulation of signals emitted from multiple sources. As a result, the transmitter needs to select adaptive coding and modulation scheme based on channel state. With the development of artificial intelligence (AI) techniques, deep learning (DL) method is gradually being applied to the task of modulation recognition, known as Automatic Modulation Recognition (AMR), which is considered to be a potentially intelligent spectrum management technique [1], [2].

Traditional methods of modulation recognition mainly include statistical pattern recognition based on feature extraction and maximum likelihood hypothesis testing based on decision theory, which require a large amount of a priori knowledge of channel parameters and have poor applicability. Recently, many studies have focused on how to use deep learning methods for the task of classifying modulation signals [3]. In [4], convolutional neural network is studied to extract features of radio signals for modulation classification, with significant performance improvements against feature-based methods. In [5], a new structure is proposed to combine the LSTM module and CNN module and takes into account the temporal characteristics of the signal The results show an effective improvement in recognition accuracy. The results shows that the CNN module can reduce the size of high-dimensional complex signal data into low-dimensional feature vectors to remove redundant information, and the LSTM module can be used to learn the temporal characteristics in low-dimensional data. The authors in [6] compare the performance of a variety of popular deep learning networks, including CNN, Residual Network (ResNet[7]), Densely Connected Network(DenseNet[8]), and Convolutional Long Short-term Deep Neural Network (CLDNN) in terms of signal recognition accuracy, the CLDNN achieving the best performance. The results of these studies show that there are great advantages in using networks applied in the image field and in speech recognition to identify the modulation of wireless signals.

However, the modulation signal data studied previous were obtained based on Nyquist sampling rule. Nowadays, broadband wireless systems need to transmit wideband signals to increase data rates, which means that high sampling rate Analog-to-Digital Converter (ADC) is required, leading a significant challenge for power consumption, and computation overhead [9]. In order to relieve the sampling burden of hardware, Compressive Sensing (CS), a new sampling theory that can only recover signals from samples with a sub-Nyquist rate, was recently introduced [10]-[12]. With the natural sparsity of the broadband signal spectrum, the sub-Nyquist sampling rate can be lower than the Nyquist sampling rate and the original redundant information in the signal can be filtered out [13]. To enhance broadband spectrum sensing performance, deep learning has attracted much attention. Motivated by this, in this paper, we present end to end DL-based frameworks for Sub-Nyquist modulation recognition. Because of high order modulation type and multiple sampling channel, it is a challenge for designing and training neural network.

II. Methodology

A. Fundamental of CNN and LSTM

A convolutional neural network mainly comprises of an input layer, convolutional layers, activation layers, and fully connected layers. The input layer is used to obtain the input information, including original data and data pre-processed by other algorithms into the convolutional neural network. For modulation signals, each frame of the sampling signal is treated as a greyscale image with a channel of 1. The number of sample points and sampling channels are treated as the length and width. The convolutional layer consists of several convolutional kernels with different size, and the convolutional operation is mainly designed to extract high-dimensional features of the data. For any input data x_i (or outputs of previous layers), the output of convolutional layer is the sum of dot product as follows:

$$Net_{out} = \sum w_i x_i \quad (1)$$

Where w_i represents weights parameters of convolutional kernel. During the training of the network each convolutional kernel learns different parameters through the back-propagation algorithm [14] so that different detailed features of the data can be extracted. By adding a scalar bias b, the output of a convolutional layer is:

$$h_{bias} = Net_{out} + b \quad (2)$$

The results of the calculations for each layer of the network are then used as input to the activation function f as follows:

$$g_{acti} = f(h_{bias}) \quad (3)$$

The activation function introduces a non-linear element to the network, thus enhancing the expressive performance of the neural network. Commonly used activation functions include sigmoid, tanh and ReLU. Since ReLU function can perform gradient descent and back propagation more efficiently, avoiding the gradient explosion and gradient disappearance problems, and simplifying the computation process making it the most popular activation function in practice.

The fully-connected layer acts as a classifier for the convolutional neural network, mapping the geographically distributed features of the data learned in the convolutional layer to each class of modulation signal.

The LSTM units are components of recurrent neural networks (RNNs) that are commonly referred to as LSTM networks. Each LSTM unit is composed of an input gate, an output gate, and a forget gate, which regulate the flow of information in and out of the cell that stores values over time intervals. LSTM networks are particularly effective in processing, classifying, and predicting time series data, as such data may contain significant temporal dependencies and patterns.

In this study, the CNN network is utilized to extract spatial features from various modulation signals, while the LSTM network is employed to examine the hidden features among temporal signals. By combining the strengths of both networks, more information can be obtained and the classification accuracy can be enhanced.

B. Designing Network Architecture

For CNN, the hyper parameters of dropout rate, number of kernels per layer and the network depth is optimized to get the best accuracy result. The CNN architecture comprises five convolutional layers for feature extraction and three fully connected layers for signal classification. The convolutional kernel size gradually decreases from large to small, allowing for a larger receptive field to capture more comprehensive feature information from the data as illustrated in TABLE I. This approach yields better results as more information is obtained from the data.

TABLE I

CONFIGURATION OF NETWORK ARCHITECTURE

Network	**Layer**	**Description and Configuration**	
		Kernel Size	***Number of Units***
CNN	Conv1	11×11	324
	Conv2	5×5	256
	Conv3	3×3	256
	Conv4	3×3	128
	Conv5	3×3	96
		LSTM Units	
LSTM	LSTM1	128	
	LSTM2	64	
	LSTM3	32	
		Fully Connected Units	
Fully Connected	Dense1	256	
	Dense2	128	
	Dense3	13	

The convolutional layer in this paper includes a zero-padding operation, which preserves the boundary information of the input signal data. Without padding, the convolution kernel would only manipulate the edge information of the input data once, while the intermediate sequence would be scanned multiple times, resulting in loss of information about the boundary features of the signal.

Furthermore, to enhance the stability and performance of the convolutional neural network, a batch normalization layer is inserted after the convolutional layer and prior to the activation function. This layer normalizes the output of the convolutional layer by utilizing the mean and standard deviation of small batches of data. As a result, the intermediate outputs of the neural network are consistently adjusted, which increases the overall stability of the network at each layer.

In the case of LSTM, the modulation signal data can be regarded as a one-dimensional sequence, which is fed as input to the network. The optimal number of hidden neurons is determined for each layer in the network. To achieve the best classification performance, three LSTM layers are designed and the fully connected layer parameters are set to the same values as those used in the CNN model. This allows for a comprehensive optimization of the network's architecture and ultimately leads to improved performance in classification tasks.

Finally, for CLDNN, five convolutional layers followed by one LSTM layer with 80 computing units and three fully connected are chosen. Different number of memory cells in the LSTM layer are tested and this configuration can give out the best performance compares to other layer settings.

III. Simulation Results And Analysis

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. Introduction to the Dataset

We use the GBsense datasets are composed of sub-Nyquist-rate time-domain samples on GHz bandwidth baseband signals as the input data of our research [15], [16]. Data can be accessed in [17]. This dataset contains 13 types of modulations given an assumption of ideal noise-free conditions. The dataset has 156,000 1024-length frames of modulation signal, where 80% for training and the remainder for testing. For the testing set, the categorization accuracy is presented. Each signal frame in the dataset is organized as the two-dimensional array with the size of 16×1024, where the eight In-phase and Quadrature (I/Q) channels modulation signal are sampled with 1024 elements.

B. Simulation results and Analysis

As network training options, parameters are randomly initialized 100 epochs using the Adam optimizer. As signal recognition can be seen as a multi-classification problem, a cross-entropy function was chosen for the loss function. The performance is measured on a system equipped by AMD Ryzen 9 5900HX CPU, 32 GB RAM, and NVIDIA GeForce RTX 3080 GPU.

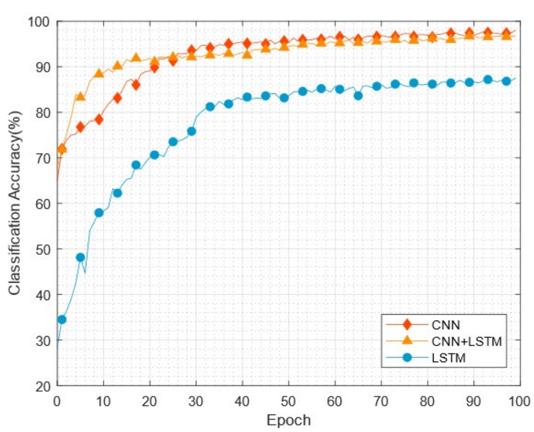


Figure 1. Accuracy of classification with different network

In the simulation experiments, the CNN network has the highest recognition accuracy of 98.01%, followed by the CLDNN 96.81% and the LSTM 87.51% respectively as shown in Figure 1. It is worth noting that the recognition accuracy of the CLDNN network is higher than that of the CNN until 20th epoch, then the accuracy is consistently lower than that of the CNN by around 1%. Although the CLDNN network can capture both spatial and temporal features of the signal, its performance does not exceed that of the CNN. The LSTM network performs the lowest accuracy, indicating that the spatial features of the signal are more important and focusing on the temporal features alone does not yield high recognition rates.

In addition, we investigate the effect of batch size on model performance. The best recognition accuracy can be achieved with a batch size of 32 for both CNN and CLDNN networks, and 64 for LSTM networks. This is due to the fact that too large or too small batch size can cause the network to easily converge to some bad local optimal points.

Figure 2 plots the confusion matrix to demonstrate the accuracy of different classes of modulation signal, where the horizontal coordinate is the type of modulation signal predicted by the trained network and the vertical coordinate is the actual class of the signal. As shown in Figure 2(a) and 2(b), the confusion matrices for CNN and CLDNN networks exhibit a clear diagonal line, indicating a high level of recognition accuracy.

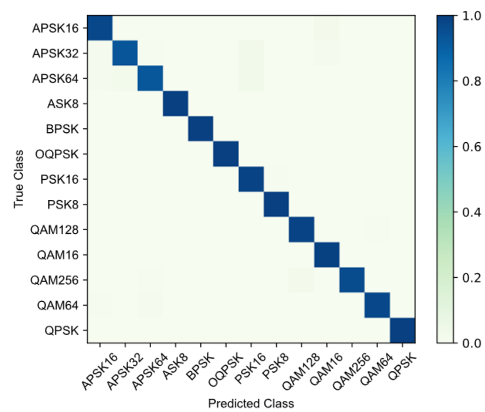


Figure 2(a). Confusion matrix of CN

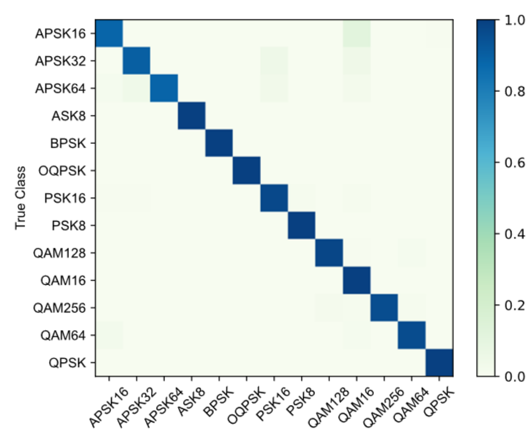


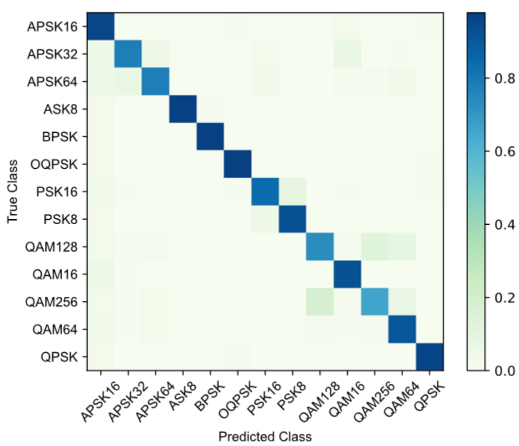
Figure 2(b). Confusion matrix of CLDNN

Figure 2(c). Confusion matrix of LSTM

Figure 2(c) shows that the LSTM network exhibits significant recognition errors for QAM256 and QAM128 modulation schemes. These errors are primarily due to the relatively small differences in amplitude and phase between symbols modulated using higher-order QAM. In scenarios with high levels of interference or noise, the recognition process becomes more challenging.

IV. Conclusions

This paper investigates the effectiveness of three different neural network models CNN, LSTM, and CLDNN in accurately identifying sub-Nyquist modulation signals. Our study focuses on identifying optimal training parameters that prevent overfitting, resulting in improved network performance. As for future work, introducing noise into the dataset will enable us to simulate a more realistic propagation environment. This, in turn, will allow for the development of more robust neural network architectures that can better handle signal interference.



References

[1] B. Jdid, K. Hassan, I. Dayoub, W. H. Lim and M. Mokayef, "Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey," in IEEE Access, vol. 9, pp. 57851-57873, 2021, doi: 10.1109/ACCESS.2021.3071801.

[2] F. Liu, Z. Zhang and R. Zhou, "Automatic modulation recognition based on CNN and GRU," in Tsinghua Science and Technology, vol. 27, no. 2, pp. 422-431, April 2022, doi: 10.26599/TST.2020.9010057.

[3] S. Peng, S. Sun and Y. -D. Yao, "A Survey of Modulation Classification Using Deep Learning: Signal Representation and Data Preprocessing," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 7020-7038, Dec. 2022, doi: 10.1109/TNNLS.2021.3085433.

[4] O'Shea, T. J., Corgan, J., & Clancy, T. C, " Convolutional radio modulation recognition networks, " in Engineering Applications of Neural Networks: 17th International Conference, EANN 2016, Aberdeen, UK, September 2-5, 2016, Proceedings 17 (pp. 213-226). Springer International Publishing.

[5] Q. Zhou, X. Jing, Y. He, Y. Cui, M. Kadoch and M. Cheriet, "LSTM-based Automatic Modulation Classification," 2020 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Paris, France, 2020, pp. 1-4, doi: 10.1109/BMSB49480.2020.9379677.

[6] 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, USA, 2017, pp. 915-919, doi: 10.1109/ACSSC.2017.8335483.

[7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[8] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[9] J. Yang, Z. Song, Y. Gao, X. Gu and Z. Feng, "Adaptive Compressed Spectrum Sensing for Multiband Signals," in IEEE Transactions on Wireless Communications, vol. 20, no. 11, pp. 7642-7654, Nov. 2021, doi: 10.1109/TWC.2021.3086952.

[10] D. L. Donoho, "Compressed sensing," in IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1289-1306, April 2006, doi: 10.1109/TIT.2006.871582.

[11] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," in IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 21-30, March 2008, doi: 10.1109/MSP.2007.914731.

[12] C. W. Lim and M. B. Wakin, "Automatic modulation recognition for spectrum sensing using nonuniform compressive samples," 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 2012, pp. 3505-3510, doi: 10.1109/ICC.2012.6364346.

[13] Song, Z., Gao, Y., & Tafazolli, R. (2021). A survey on spectrum sensing and learning technologies for 6G. IEICE Transactions on Communications, 104(10), 1207-1216.

[14] Cilimkovic, M. (2015). Neural networks and back propagation algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North Dublin, 15(1).

[15] H. Zhang, J. Yang and Y. Gao, "Machine Learning Empowered Spectrum Sensing Under a Sub-Sampling Framework," in IEEE Transactions on Wireless Communications, vol. 21, no. 10, pp. 8205-8215, Oct. 2022, doi: 10.1109/TWC.2022.3164800.

[16] Song, Z., Qi, H., & Gao, Y. (2019, October). Real-time multi-gigahertz sub-Nyquist spectrum sensing system for mmwave. In Proceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing Systems (pp. 33-38).

[17] University of Surrey, (2022, June). "Deep Sub-Nyquist Modulation Recognition Challenge," distributed by GBSense .