The Application of Deep Learning in Communication Signal Modulation Recognition

Invited Paper

Yun Lin[1], Ya Tu[1], Zheng Dou[1,*], Zhiqiang Wu[2]
1. College of Information and Communication Engineering, Harbin Engineering University, Harbin, China
2. Department of Electrical Engineering, Wright State University, Dayton, USA
douzheng@hrbeu.edu.cn

***Abstract*—Automated Modulation Classification (AMC) has been applied in various emerging areas such as cognitive radio (CR). We also notice that Deep Learning (DL) is a powerful classification tool that has gained great popularity in various field. This article focuses on DL and aims at using it to solve communications problems. We propose a new data conversion algorithm in order to gain a better classification accuracy of communication signal modulation. This paper will show that our new method will bring significant improvement in signal modulation classification accuracy. Besides, AlexNet and GoogLeNet, two well-known DL network models, ResNet and VGG, will be utilized in this task to compare with each other.**

Keywords-Deep Learning; Convolutional Networks; modulation recognition; AlexNet； GoogLeNet；VGG；ResNet

I. Introduction

With the increasing demand for wireless band-width of radio spectrum, it is imperative to have better approaches of using the radio spectrum [1]. Hence, more and more attentions has been paid to developing opportunistic spectrum access techniques, known as Cognitive Radio (CR).In Dynamic Spectrum Access (DSA), the first priority is to avoid radio interference and optimize spectrum allocation[2]. Due to the fact that modulation pattern become one of the significant sorting parameters of signal characteristics of intercepted emitters, Modulation Recognition is a way to understand what type of communications scheme and emitter is presented [2]. Hence, great attention should be attached to finding a smart method to recognize the modulation.

Deep Learning is a new field of Machine Learning research. Using a cascade of many nonlinear processing units layers, Deep Learning are capable of feature extracting and transforming. The output, comes from the previous layers, will be fed into the successive layers. What's more, unsupervised learning of features and representations of the data is fundamental to Deep Learning [3]. Derived from lower level features, higher level features represent hierarchically.

The world has witnessed a great achievement accomplished by Deep Learning. For example, when it comes to Image recognition, In October 2012, Krizhevsky and Hinton won the ImageNet competition by using a similar system [4]. In November 2012, Ciresan et al.'s system, which could analyze the large medical images of cancer detection, also won the ICPR contest [5], in 2013 and 2014, followed by a similar trend in large-scale speech recognition, the error rate on the ImageNet task was decline by using Deep Learning. These achievements were publicized by The Wolfram Image Identification project [6].

Since Deep Learning is booming everywhere, it also attracts a lot of attention in communication field, it evolves so rapidly day by day that we can't afford to ignore its potential in other complicated problems. Besides, DNN will extract the feature and use them by itself, which will greatly free our job in feature extraction and reduce the demand of our prior knowledge [7]. Due to those advantages, we will make full use of deep networks in modulation recognition.

In the next section, we will give an overview on the architectures and models of different CNN. In section III, compared to the article named Deep Learning and Its Applications in Communications Systems: Modulation Classification deep learning model deep learning model [7], we present a new idea about how to process Constellation Diagram to make our training set more discernible to get a better classification accuracy. In section IV, we apply two new model, ResNet VGG, which is come up with Microsoft and University of Oxford. Recently in this task. Lastly, we will analysis our result and try to figure out what really counts in this task in section V.

II. Deep Learning Architectures and Models

A. Deep Neural Network

Deep Neural Network (DNN) can present complex nonlinear relationships. By using the fusion of features from lower layers, the extra layers potentially use complex data with fewer units than a similarly performing shallow network [8]. Deep Neural Network has some basic network architecture, such as CNN, LSTM and so on.

B. Convolutional Network

Convolutional Network is widely known model in architecture. Three basic components are needed to define a basic Convolutional Network.

- After applying a convolution operation to the input, Convolutional layers pass the result to the next layer. This operation involves Convolutional core. This layer has the property of sparse connectivity and weight sharing

- Pooling layer are not necessary. But when the images are somehow large, we would wish to reduce the number of parameters. Then, Pooling layers is designed to periodically introduce between subsequent convolution layers. Pooling is done just for the purpose of reducing the spatial size of the image.
- Fully connected layer. The convolution and pooling layers would only be able to extract features and reduce the number of trainable parameters. However, we need fully connected layer to be a classifier.

C. AlexNet

AlexNet has made a great contribution to CNN's application in computer vision with a large and deep architecture. It consists of 5 convolutional layers and 3 fully-connected layers with a 1000-ways Softmax layer [4]. AlexNet describes two primary ways in which we could combat overfitting. The first measure is Data Augmentation, and the second is "Dropout" [4]. Another great contribution is the use of ReLU, which eliminate the problem of vanishing gradient somehow.

D. GoogLeNet

GoogLeNet, a 22-layers CNN and the winner of ILSVRC'14, is an efficient deep neural network model designed for computer vision. Its pioneer work was the development of an Inception Module that significantly r lessen the scale of parameters in the network (4M, versus AlexNet with 60M). Moreover, considering the notorious problem of vanishing/exploding gradients, additional classifiers such as Softmax to compensate the loss of gradients [9].

E. VGGNet

The runner-up in ILSVRC 2014 was the network from Karen Simonyan and Andrew Zisserman that became known as the VGGNet. Its main contribution was showing that the depth of the network is a critical component for good performance [10].

A downside of the VGGNet we should keep in mind is that it is more expensive to evaluate and uses a lot more memory and parameters (140M).

F. ResNet

Residual Network developed by Kaiming He et al. was the winner of ILSVRC 2015. Many visual recognition tasks have greatly benefited from very deep models. But Kaiming He and his team found networks have not demonstrated accuracy gains with extremely increased depth due to the degradation problem [11]. Kaiming He and his team proposed a novel reformulation named residual block may help to the problem [11]. In order to optimize the training speed, they used to tune the layer sizes carefully in order to balance the computation between the various model sub-networks from ResNet-v1 to ResNet-v4 [11].

III. Modulation Data Conversion for Deep Learning

It is well-know that Deep Learning only accept three dat a formation: image, sound and text. In this task, we need to f eed CNN with picture. So our first job is to create image dat aset. Similar to Deep Learning and Its Applications in Com munications Systems: Modulation Classification [7], we als o assume our system is operating at the same environment. We generate a random time domain sequences and then poll ute it by additive white Gaussian noise.

A. Constellation Diagram

By displaying the signal as a two-dimensional xy-plane scatter diagram in the complex plane, Constellation diagram has been widely applied to represent a modulated signal . In a more abstract sense, it represents the possible symbols that may be selected by a given modulation scheme as points in the complex plane.

The number of sample point should be elaborately planned. If number of sample point are selected too large, valuable compute resource will be wasted. On the contrary, if number of sample point are selected too small, some signal samples severely polluted by the noise may be excluded from the image and abandoned. So we select 10000 sample point per picture empirically.

B. 3-Channel Image

Nowadays, prevalent CNN, which aims to recognize objects in high-resolution image, are designed to handle with RGB image. As is shown in Deep Learning and Its Applications in Communications Systems: Modulation Classification [7], gray images is only one channel is utilized and the capacity of these models is not fully exploited [7]. So the result will be better if we use colored image. Thanks to their job, it proves CNN's learns better on colored image training set than gray image. Compared to their work, we now propose a new method to make the colored image training set more discernible.

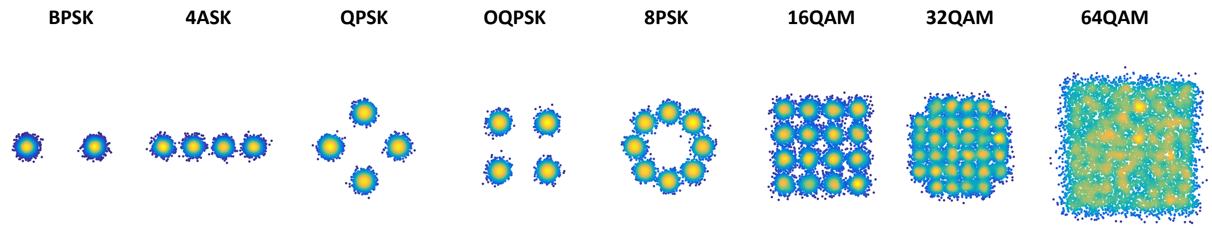
In the article Deep Learning and Its Applications in Communications Systems: Modulation Classification deep learning model deep learning model [7], they use the exponential decay model to draw Stellar Image. They treat each sample point as a star, which emits light and illuminates nearby pixels. The brightness of a sample point will be influenced by its neighbor sample point [7].

Though they have proved their method works very well, since they obtain much higher test accuracy on 3-channel Stellar Image on CNN (AlexNet and GoogLeNet) than other traditional methods, such as SVM and Cumulent. However, they still use almost the same color (mainly dominated by black and blue colors) in the image, where a good question arises: Can we use different color to increase the diversity of the image feature?

We introduce a new image called contour stellar image, which means this image will display different color somehow. In Constellation diagram, caused by Gauss white Noise, signal sample point won't get gather in one point, different area in the image has different density of sample point.

The first thing we need to do is convert complex signal into Constellation Diagram. Then a small square, we called density window, will slide on the picture. On sliding on the picture, density window will sum how many dots is in its field. Different result means different density, we will use different color to mark different density. Yellow means higher density, Green means middle density and Blue means lower density. It seems that the sample point are more likely to gather in the middle of the clusters.

In the article Deep Learning and Its Applications in Communications Systems: Modulation Classification deep learning model deep learning model [7], their dataset use 8 category signals: BPSK, 4ASK, QPSK, OQPSK, 8PSK, 16QAM, 32QAM, 64QAM, we will generate the same dataset to perform the compare test. As is shown in Fig. 1, Fig .1(a) and Fig. 1(b). Through the same process, the contour Stellar Image of at the SNR of 14dB, 4dB will be presented as Fig .1.



(a) 8 category signals at SNR = 14dB

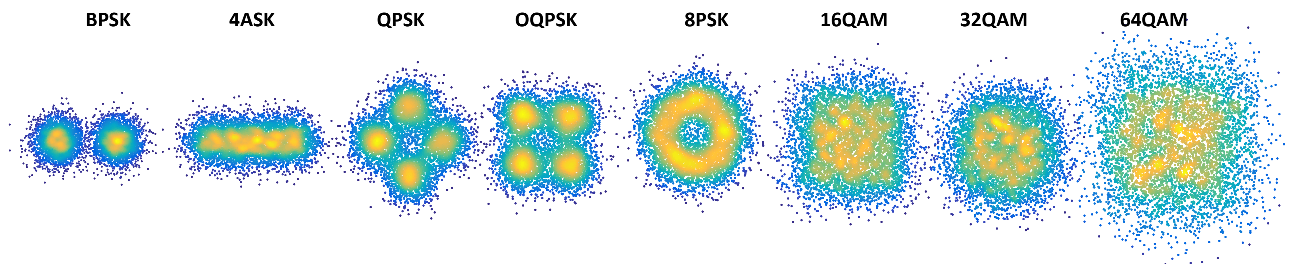
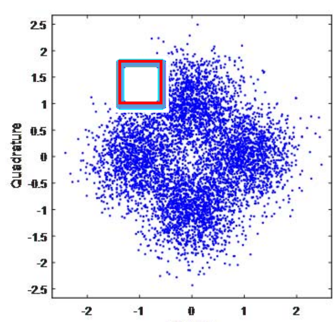


Figure 1. (b) 8 category signals at SNR = 4dB

Taking QPSK at SNR =4dB as an example, the method will be shown in Fig.2. The red square (density window) will count how many dots in the square and then convert it to contour stellar image depended on the dots density.



(a)QPSK Constellation Diagram at SNR=4dB

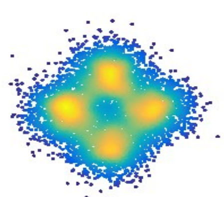


Figure 2. (b) QPSK Contour Stellar Image at SNR=4dB

As is depicted in the Fig .1, feature in contour stellar image become more diversity. While keeping the principle of dealing with training set fairly, we add more feature to it. Because we believe the separability in dataset plays an important role in the task of classification. In the following of this article, we will feed AlexNet and GoogLeNet, VGG, ResNet models with contour stellar images to implement modulation classification. Every signal will be distributed 9000 contour stellar images in training set, 1000 contour stellar images in validation set and 1000 contour stellar images for test set. Without data augmentations, the article, Deep Learning and Its Applications in Communications Systems: Modulation Classification deep learning model deep learning model, expects that larger training dataset could be helpful on eliminating overfitting and improving the predictive accuracy too [7]. Hence, they generate 100 thousand images and 1000 images per types for training and testing, respectively. Actually, we just use 1/10 scale of training set and try to figure out if diversity feature dataset could defeat large dataset.

IV. EXPERIMENTS

In last chapter, we have discussed how to convert signal sample point into Constellation diagram. Thanks to MATLAB 2017a comprehensive and plentiful toolbox in signal processing, it is good tool for us for implementing this procedure. In order to compare with the article Deep

Learning and Its Applications in Communications Systems: Modulation Classification deep learning model Deep Learning model [7] result, we generate the same type signals ,which is BPSK, 4ASK, QPSK, OQPSK, 8PSK, 16QAM, 32QAM, 64QAM. Every signal will be disturbed 9000 contour stellar images for training set, 1000 for validation set and 1000 for test set. The SNR will range from -6dB from 14dB, with the stride is 2dB.

A. ConvNets Classification

To build our ConvNets model, Tensorflow, a well-known framework, is employed to this work. Due to its strong modularity and friendly interface with Python. It greatly facilitate our work in constructing four CNN model,which are AlexNet and GoogLeNet-v2 ,VGG-16, ResNet-v4.

Refer to the idea of Transfer Learning, we slightly modify a little parameter to speed up our training while prevent information loss. Considering the trade-off between available computing resources on GPU and training efficiency, we set our batch size is 32.

B. AlexNet-based Classification

We just modify one layers from the AlexNet model come up with Alex Krizhevsky[4], The number of output in the layer #8 is reduced from the default of 1000 to 8 which is mean to match the number of modulation types in our task. Eventually, we fix our contour stellar image resolution to 227x227, which is compatible with the size of receptive field in AlexNet. The parameter in Local Response Normalization, beta is change from 0.75 to 1, learning rate is 0.001 and loss function is categorical cross-entropy, other is default. After our 20 times experiment of tuning parameter, it is found that this parameter will bring slight better performance.

C. GoogLeNet-based Classification

Based on the article [9], we construct GoogLeNet-v2. Like what we do with AlexNet, the number of output in the layer which is fully connected layer behind inception block 5b is reduced from the default of 1000 to 8 to match the number of modulation types in our task. Later we fix our contour stellar image resolution to 224x224 to satisfy the demand of input layer. The parameter learning rate is 0.001 and loss function is categorical cross-entropy, other is default. After our 20 times experiment of tuning parameter, it is found that this parameter will bring better performance.

D. VGG-based Classification

Based on the article [10], we construct VGG-16. There are several places have been changed based on default configuration of VGG-16. Firstly, just like what we have done in AlexNet and GoogLeNet. We modify the number of output in the layer from 1000 to 8 and fix image resolution to 224x224. Then we change 0.001 to 0.0001 as our learning rate and RMSprop algorithm as our optimizer.

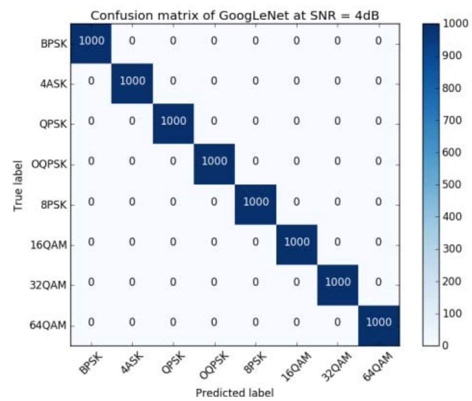
E. ResNet-based Classification

Based on the article [11], we also establish ResNet-v2, as what we have done in AlexNet, GoogLeNet and VGG-16. We change the output of the last average pooling layer from 1000 to 8 to meet the demand of our task fix image resolution to 299x299. Learning rate also be set to 0.001, and RMSprop algorithm is employed in this task. Dropout possibility has been change to 0.8.

V. RESULT

A. contrast experiment

Our paper mainly introduce a new data conversion to improve the classification accuracy. Here we give some experiment results to show the effect of data conversion. In our experiment, 1000 tests are implemented to evaluate the classification performance. Since the article [7] we want to compare with gives a result in GoogLeNet-v2 at the SNR = 4dB in Fig .3, we will also give our result about GoogLeNet-v2 in Fig .3.



(a) Contour stellar image dataset

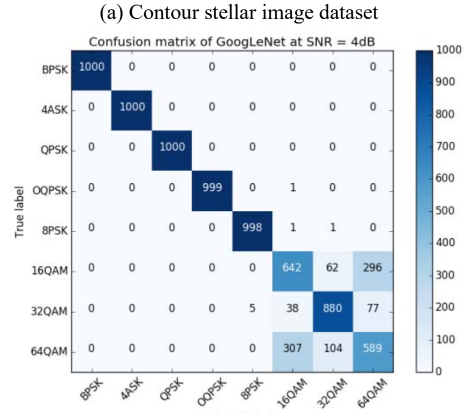


Figure 3. (b) Stellar image dataset.

We are little surprised to find that we get 100% accuracy at the SNR=4dB by using GoogLeNet. The result implies classification performance turns significantly better in our contour stellar image dataset.

B. Average accuracy comparison

For comparison, we consider four different model AlexNet, GoogLeNet, VGG-16 and ResNet, Fig. 6 depicts the average classification accuracy of five methods versus SNR. For each modulation category, 1000 tests are implemented. By averaging the performance of eight modulation categories, we will get the average accuracy.

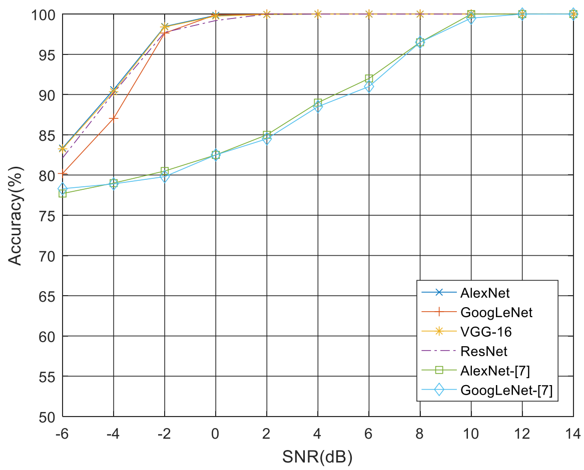


Figure 4. Average accuracy vs SNR

As shown in Fig. 4, following observations are obtained:

- With SNR increasing, classification accuracy improves.
- Contour stellar image based methods are much better than the method in the article we want to compare with.
- The accuracy of four Deep Learning based methods are nearly equivalent. It implies those four model has extracted enough features to classify reliably.

VI. Conclusion

In this paper, we briefly talk about Deep Learning, including its concepts, architectures and models. Nowadays, the basic data structure, which Deep Learning are capable of dealing with, is image, sound and text. In order to make full use of Deep Learning to complete modulation classification task, we propose a new method to create contour stellar image to convert complex signals into images. The new method actually increase the separability in dataset. Since the accuracy of four Deep Learning based methods are nearly equivalent, we can draw a conclusion that separability in dataset plays a decisive role in classification task. It is confident to say traditional machine learning will also obtain a better classification result. So a natural question arise now: What measure should be taken to add separability in dataset? Deep Learning Visualization may give us answer, because it will tell us what CNN really learns so that we could strength or denoise specific feature to increase the separability in dataset. We strongly believe it is a promising future work.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (61771154, 61301095, 61603239), the Key Development Program of Basic Research of China(JCKY2013604B001), the Nature Science Foundation of Heilongjiang Province of China (F201408) and the Fundamental Research Funds for the Central Universities (No. HEUCF100814 and HEUCF100816).
Meantime, all the authors declare that there is no conflict of interests regarding the publication of this article.

References

[1] S. Haykin, “Cognitive radio: brain-empowered wireless communications, ”IEEE Journal on Selected Areas in Communications, vol. 23, no. 2, pp. 201-220, 2005.

[2] Mendis, Gihan J. Mendis, Jin Wei, and Arjuna Madanayake. “Deep Learning-based automated modulation classification for cognitive radio,” IEEE International Conference on Communication Systems, 2016.

[3] Deng, Li, and Dong Yu. “Deep learning: methods and applications,”Foundations and Trends® in Signal Processing, vol.7, no. 3-4, pp. 197-387, 2014.

[4] A. Krizhevsky, I. Sutskever and, G. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” International Conference on Neural Information Processing Systems, pp. 1097-1105, 2012.

[5] Cires an, D. C., Giusti, A., Gambardella, L. M., and Schmidhuber, J, “Mitosis detection in breast cancer histology images with deep neural networks,” Medical Image Computing & Computer-assisted Intervention: Miccai International Conference on Medical Image Computing & Computer-assisted Intervention, pp .411-418, 2013.

[6] “The Wolfram Language Image Identification Project”, https://www.imageidentify.com/

[7] Shengliang Peng, Hanyu Jiang,Huaxia Wang,Yu-Dong Yao, “Deep Learning and Its Applications in Communications Systems: Modulation Classification”, http://personal.stevens.edu/~hwang38/DL.pdf

[8] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., and Rabinovich, A, “Going deeper with convolutions,” IEEE Conference on Computer Vision and Pattern Recognition, pp .1-9, 2015.

[9] Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A, “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning,” AAAI, pp. 4278-4284, 2017.

[10] Simonyan, Karen, and Andrew Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv: 1409.1556, 2014.

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