J Inf Process Syst,
http://dx.doi.org/10.3745/JIPS

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Enhanced Multi-scale Coded Convolutional Neural Network

Bin Jiang*, Weiyu Li*, Haozhan Sun*, Ziliang Zhao*, Yunpeng Zhong*,
Donglai Wei*, Qingwei Li*, Jiyu Wei*, Weihong Wang*, Hao Huang*

Abstract

Spectral data classification is of great significance for astronomical data processing. The best classification network has the ability to extract the most features without noise and spectral individual features. Such a network has better generalization capabilities and can extract sufficient features for classification.

This paper tries to use a variety of classification networks of 1D and 2D to classify spectral data in two dimensions. We try to perform one-dimensionalization on the current advanced 2D classification network and by folding 1D spectral data to two dimensional. The aim is to explore the feasibility of using one-dimensionalized 2D network to classify 1D data and the feasibility of doing classification on the 2D data converted from 1D with 2D network.

First, the experimental results show that the fully connected neural network cannot extract enough features. Second, although the convolutional neural network with strong feature extraction capability (e.g. VGGNet and ResNet) can quickly achieve satisfactory results on the training set, we found that CNN is vulnerable to over-fitting, which cannot be solved during our experiments. Third, spectral data with different signal-to-noise ratios (SNR) also have an effect on the performance of classification of the same network. To solve the problems above, the Enhanced Multi-scale Coded Convolutional Neural Network (EMCCNN) is designed, which can perform denoising on spectral data, and feature extraction of denoised spectral data at different scales. The EMCCNN network also combines features in different scales to improve the performance of feature extraction, which can maximize the ability to extract common features, and minimize the noise and individual features. The thorough experimental results show that EMCCNN outperforms other well-known standard classification models such as ResNet or VGG in terms of classification performance.

Keywords

Multi-scale, Coder, Spectral Classification, Convolutional Neural Network

1. Introduction

The Sloan Digital Sky Survey (SDSS) began in 2000 and provides a wealth of valuable data for spectroscopy research. Accurate and efficient classification of acquired spectral images is an important task in astronomical data processing. The data set classified in this paper is the SDSS-DR14 data set of M dwarf.

Spectral data has different SNR, and data with different SNR often lead to different results for the same

Manuscript received Month 06, 20XX; accepted Month 06, 20XX.

Corresponding Author : Hao Huang (huanghao_sdu@mail.sdu.edu.cn)

* School of Mechanical, Electrical & Information Engineering in Shandong University, Weihai, China (jdxy@sdu.edu.cn)

method. This paper explores the influence of SNR on classification accuracy under various methods, and the influence of SNR on classification accuracy under different dimensions of the same method, and gives different classification network structures under high SNR and low SNR conditions, which greatly improves the classification accuracy. The paper also gives an in-depth discussion on the degree of feature extraction and overfitting of spectral classification, and proposes corresponding solutions.

2. The related work

In 2014, Pavel first introduced a multi-layer convolutional neural network based on the learnable kernel into the field of spectral classification. He tried to fold the one-dimensional spectral image, but did not achieve satisfactory results. Later, Wen-Yu W et al. attempted to use the one-dimensional convolution network to automatically identify the white dwarf main sequence double star [5], which proved that the one-dimensional convolution network plays a significant role in the spectral classification field.

With the continuous development of deep learning technology, Convolutional Neural Network (CNN) is giving excellent performance for feature extraction and combination. In the field of 2D image classification, there are many outstanding models. As a classic deep convolutional network, VGGNet [1] is widely used in 2D image classification tasks. Resnet [2] is the network structure proposed by Kaiming He in 2015. Its innovative use of residual learning allows deep networks to extract more effective features. DenseNet [3] first proposed that the output of each layer is the input of the next layer, allowing the network to extract features of more dimensions. For the spectral classification task, it is actually extracting different types of spectral features, so this paper experiments the availability of the above networks for spectral classification tasks in two dimensions.

3. Method

The goal of spectral classification is to maximize common features and minimize noise and spectral individual features. Therefore, the effect of spectral classification P should satisfy the following formula:

P is the final effect of spectral classification and C stands for common feature. S and N represent the individual features of the spectrum and the noise.

3.1 Convolutional Neural Network (CNN)

We use convolutional neural networks to classify M dwarfs, because convolutional neural networks perform well in 2D image classification, and it can also handle 1D classification work. The feature extraction capabilities of convolutional neural networks can help us improve the accuracy of traditional classification methods. Because M dwarfs in the SDSS dataset are stored in computers as 1D vectors, we use a 1D convolutional neural network to classify them. Considering the good performance of the convolutional neural network for 2D images, quadratic interpolation is performed on the data to facilitate the folding of the spectral data to form a 2D spectral matrix.

Using the above method above, we increase the 3522 dimensional data to 5000 dimensions, and fold it into a 50*100 spectral matrix.

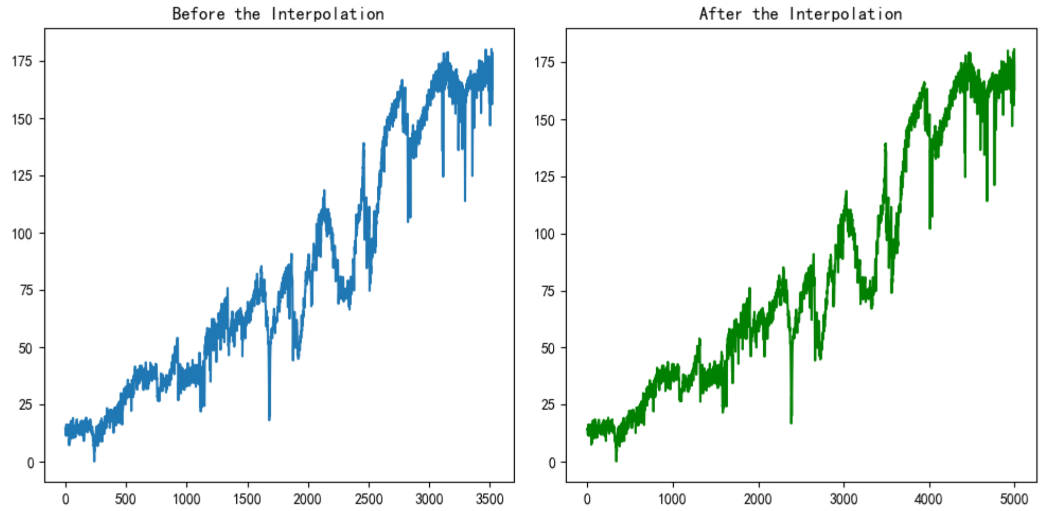

Fig.1. Comparison of spectral data before and after interpolation

As a special form of convolutional neural network, 1D convolutional neural network has certain applications in the field of signal processing. We adopted the network structure shown in the following figure.

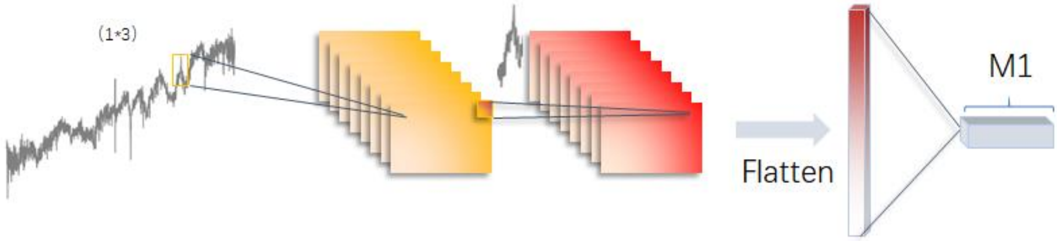

Fig.2. The network structure of convolutional neural network

In the specific network construction, the backpropagation algorithm is adopted. The 1D matrix information composed of spectral data is continuously trained to fit the parameters in each layer by gradient descent. Our loss function can be defined as the mean square error：

Using the total error to calculate the partial derivative of the parameter, the magnitude of the influence of a parameter on the overall error can be obtained, which is used to correct the parameter in the back propagation. Since the construction of the network uses a linear arrangement, the total error calculated by the final output layer is used to perform the partial derivative calculation of the parameters in all layers.

*

In a convolutional neural network, each neuron is only connected to the neurons of the previous layer, so the calculation of the local gradient requires a forward recursive calculation of the gradient of each subsequent layer of neurons. After defining the linear output of each layer and the parameters in the structure, the output of the activation function is: Out=φ(v), then the recursive formula for each gradient can be derived as:

In the successive calculation of the above formula, the gradient calculated by each parameter under the total error is used as the basis to achieve the goal of correcting the parameters in the direction of the minimum loss function.

3.2 VGGNet and ResNet

VGGNet is the most widely used in 2D data. VGGNet's generalization ability for different datasets is very prominent, so we aim to make the VGG network one-dimensional, and also try to fold the spectral data. Considering that our spectral dimensions are not particularly large, in order to prevent overfitting on the training set, the shallow VGG16 network is used in VGGNets to perform our experiments.

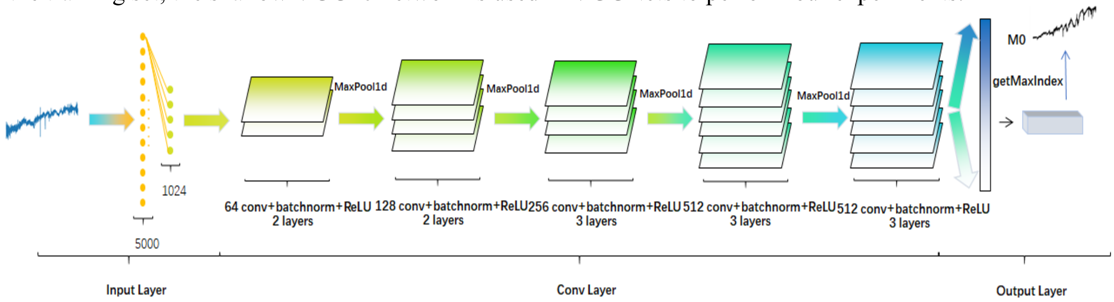

Fig.3. The network structure of VGG16

The biggest improvement of VGGNet is to convert large and short convolutional layers into small and deep neural networks by reducing the scale of convolution kernels. In the convolutional layer, the number of extracted features for each layer output can be calculated using the following formula:

With the continuous development of the classification network, the network is constantly deepening for better extraction and feature combination, but the deep network makes it very difficult to train and potential of loss of information. ResNet proposes residual learning to solve these problems, allowing information to be transmitted over the layer to preserve information integrity, while learning only the residuals of the previous network output:

This allows us to increase the number of layers in the convolutional layer, which is a good way to train a very deep network.

The biggest characteristic of the ResNet is that in addition to the result of the conventional convolution calculation in the final output, the initial input value is also added. Therefore, the result of network fitting will be the difference between the two, thereby we can obtain the calculation formula of each layer of ResNet is as follows:

Then the gradient calculation formula of the neural network is changed on the basis of conventional ones:

Compared with the traditional network, the extra value "1" makes the calculated gradient value difficult to disappear, which means the gradient calculated from the last layer can be transmitted back in the reverse direction, and the effective transmission of the gradient makes our spectral data features more efficient in the training of neural networks. However, considering that our spectral features are limited and high noise interference, the efficient feature extraction of the network may lead to significant overfitting, which makes the trained network generalization capability poor.

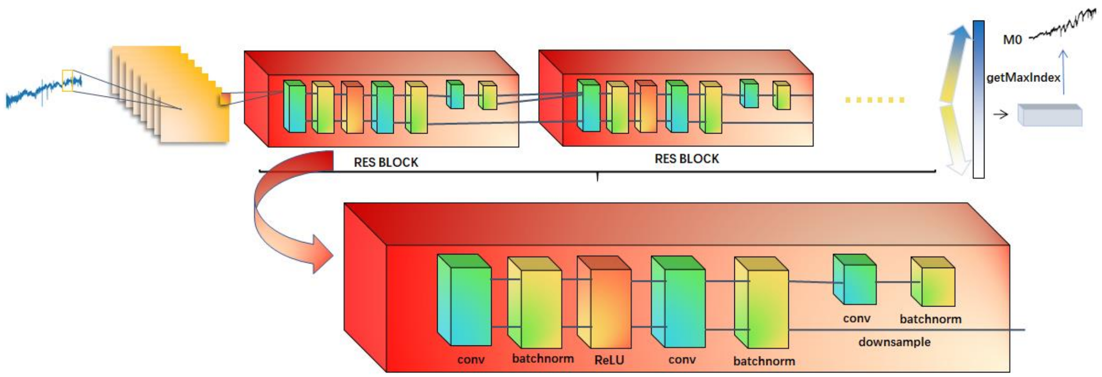

Fig.4. The network structure of ResNet

We previously suppose that the residual network is the leading classification network in the field of 2D image classification. It is worth expecting the effect of spectral classification after one-dimensionalization.

3.3 EMCCNN

The Enhanced Multi-Scale Coding Convolutional Neural Network (EMCCNN) is a network designed for the characteristics of spectral data, and has achieved good results on different SNR data. We find that for spectral data, a deep convolutional layer can lead to significant overfitting of the data and a very poor generalization capability of the model. And directly convolution of spectral data will extract a lot of noise and individual features, which is not a good way to classify spectral data. Unsupervised denoising methods tend to remove some spectral feature peaks in the data, which is also not an ideal method to maximize the common features of the spectral type.

Therefore, we try to add a supervised denoising network before the convolutional neural network, which can help the identification of features and noise. It allows the network to extract spectral features instead of noise. On the other hand, the spectral feature peak types are not consistent. For different types of feature peaks, the different convolution kernel sizes may be able to extract different quality features. The extraction of some features may be better for larger convolution kernels. Others may be more friendly to small-scale convolution kernels. We decide to let convolution kernels of different scales learn features simultaneously, then combine these features, and obtain different weights for features through the fully connected layer. A better feature extraction network EMCCNN can be obtained.

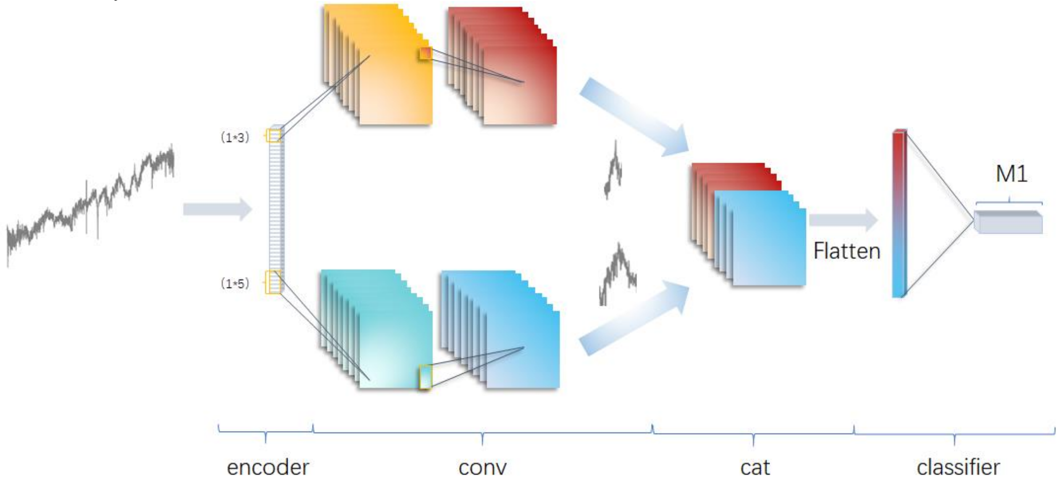

Fig.5. The structure of Enhanced Multi-scale Coded Convolutional Neural Network

4. Experiment

4.1 Environment and datasets

We use the dataset of DR14 in SDSS. The data is classified according to the SNR and the type of M dwarfs. The SNR is divided into 5-10, 10-15, and above 15. The type of M dwarfs is divided into five classes M0-M4. The specific data is shown in the figure below.

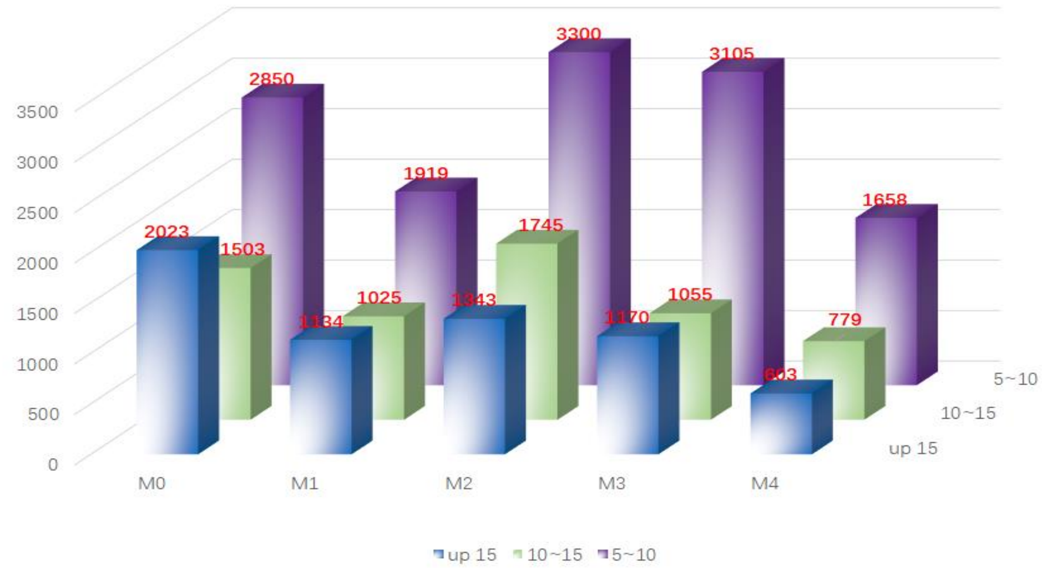

Fig.6. Distribution of the M dwarf dataset

4.2 Experimental process and results

4.2.1 1D convolution experiment and results

We first try to compare the spectral classification of the four hidden layers of DNN and one-dimensional CNN. The optimizer used was a random gradient descent, the learning rate was set to 1e-4, and the activation function chosen is ReLU. We find that the CNN network can quickly extract the corresponding features, but it will also overfit quickly, which has an accuracy of 60% in the test set, and the generalization capability is extremely poor.

CNN has strong feature extraction capabilities, but it is not a good thing for spectral data because it also extracts quite a lot of individual features and noise. And this situation gradually becomes apparent as the SNR increases.

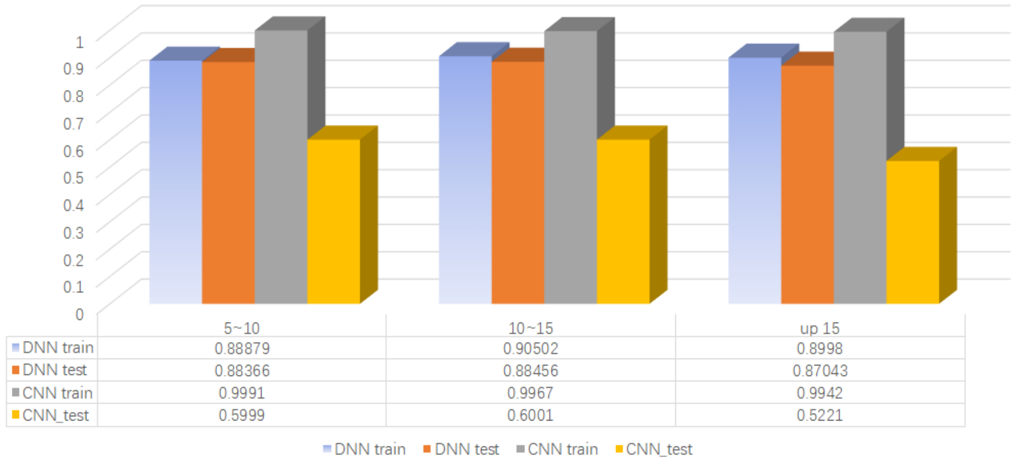

Fig.7. Comparison of classification accuracy (%) between DNN and CNN

Considering that VGG16 is composed of several convolutional layers, the direct convolution of the spectral data will be the same as the direct use of CNN, leading to the overfitting of the training data. So we try to add the encoder to the VGG16 network to denoise. We found that after denoising, VGG16 is not as bad as CNN for data on the range of 5<=sn<=15, but it is also not satisfactory for data with sn>15. It is considered that the data features of high SNR are more obvious. The spectral individual features and noise will have a strong interference effect on a very deep network. On the other hand, the denoising encoder has too long feedback process, which is not easy to identify the difference of individual features, noise and common features.

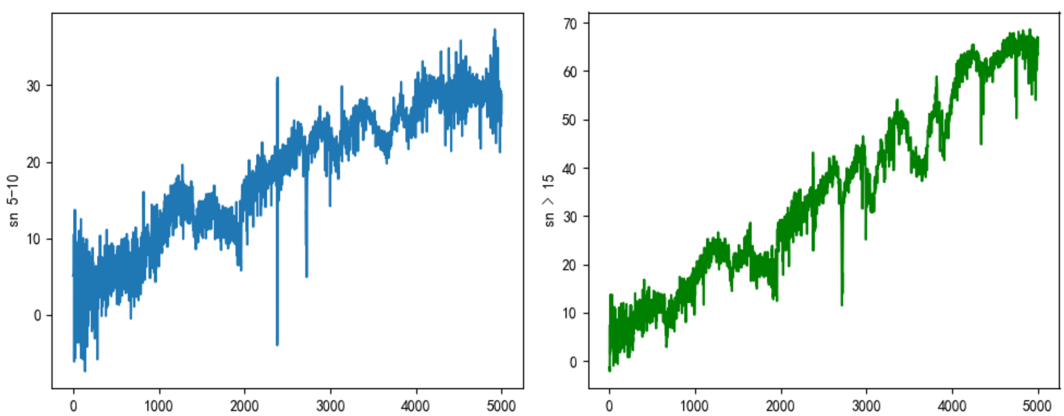

Fig.8. Spectral contrast of different signal-to-noise ratios (SNR)

ResNet's interlayer transfer mechanism might help us solve the problem of overfitting, but the effect is not very satisfactory. ResNet and CNN also perform well on the training data, and reached more than 99% accuracy on the three SNRs. It does not perform well in the test set, although not as serious as CNN. We try to add the dropout layer to help us solve the problem of overfitting, but this makes our training very slow and the final result is not satisfactory as well.

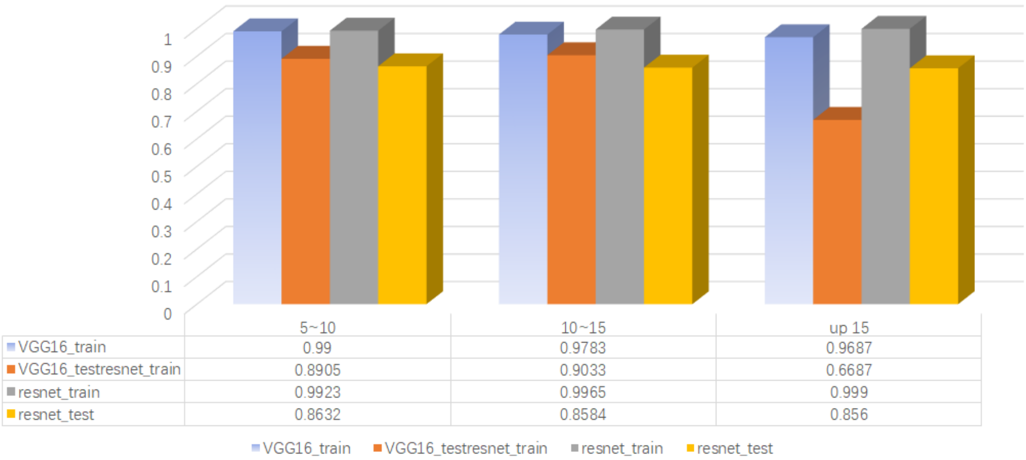

Fig.9. Comparison of classification accuracy (%) between VGG16 and ResNet

Through experiments, it is noticed that the network that is too deep is not suitable for the classification of spectral data, which will make the individualized features of noise and spectrum fully extracted, and the denoising of the encoder before these networks will make the feedback time too long, leading to a bad denoising result. Based on the above conclusions, we propose EMCCNN, which is not too deep, and is beneficial to the training of denoising capability of encoder.

We find that EMCCNN achieves the best results at all three SNRs, which proves the correctness of our ideas and is consistent with our expected experimental results.

Table.1. Classification accuracy (%) of different network models

sn / model	5~10		10~15		up15	
	train	test	train	test	train	test
DNN	0.88879	0.88366	0.90502	0.88456	0.8998	0.87043
MyCNN	0.9991	0.5999	0.9967	0.6001	0.9942	0.5221
VGG16	0.99	0.8905	0.9783	0.9033	0.9687	0.6687
ResNet	0.9923	0.8632	0.9965	0.8584	0.999	0.856
EMCCNN	0.9914	0.9217	0.9942	0.9277	0.9659	0.8916

Finally, we used SVM to directly process spectral data for classification and compared with the results of EMCCNN. Through comparison, we find that SVM can quickly fit the training set of data, but overfitting is serious.

Table.2. Classification accuracy (%) of SVM and EMCCNN

sn / model	5~10		10~15		up15	
	train	test	train	test	train	test
DNN	1	0.6802	1	0.7213	1	0.741
EMCCNN	0.9914	0.9217	0.9942	0.9277	0.9659	0.8916

4.2.2 2D convolution experiment and results

We apply the classification network that is currently applied to 2D data in the field of spectral classification to explore the classification feasibility of the two-dimensional folding of the one-dimensional data. We fold the data into a 50x100 matrix and put it into a two-dimensional classification network for experiments. The feature peaks of spectral data become inconspicuous after folding. The spectral data is only related on the same line after folding. The data correlation for each column is not obvious, but from the picture we can clearly see that each type of spectrum is different after folding, which brings the possibility of applying the 2D classification network. We aim to fully apply the classification network so that satisfactory performance can be achieved.

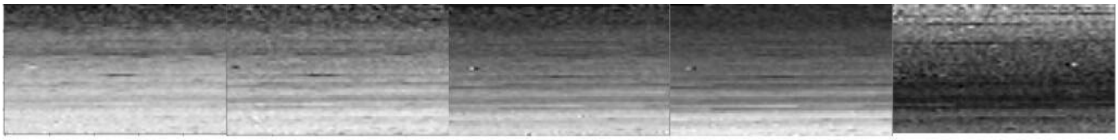

Fig.10. Spectral data after two-dimensional folding processing

We conducted three network comparison experiments (CNN, VGG16 and Res18 respectively). We find that the two-dimensional network and the one-dimensional network each has its own merits on extracting spectral features and combination features. Among them, CNN overfits heavily on 1D network. Thus in comparison, the performance on 2D data has a significant improvement over 1D data.

Among them, VGG16 performs quite well in the range of 10<sn<15 SNR, and ResNet also performs better than 1D data in the range of 5<sn<15. This discovery is inspiring, because it proves that although 1D data is not directly related to the top and bottom after folding, but the 2D classification network can still achieve satisfactory results, which proves that the 2D classification of one-dimensional data is feasible, even for some data can produce results better than one-dimensional classification.

Table.3. 2D-classification accuracy (%) of different network models

sn / model	5~10		10~15		up15	
	train	test	train	test	train	test
MyCNN_1d	0.9991	0.5999	0.9967	0.6001	0.9942	0.5221
MyCNN_2d	0.9971	0.8288	0.9977	0.8271	0.998	0.731
VGG16_1d	0.99	0.8905	0.9783	0.9033	0.9687	0.6687
VGG16_2d	0.9846	0.9103	0.9722	0.932	0.99	0.8504
ResNet_1d	0.9923	0.8632	0.9965	0.8584	0.999	0.856
ResNet_2d	0.9949	0.8738	0.9977	0.8801	0.9938	0.8132

5. Conclusions

The goal that we propose spectral classification is to maximize the extraction of common features and to minimize the extraction of noise and spectral individual features, which enhances the generalization capabilities of the network. We also find that the simple DNN network cannot extract the features of the spectrum well, but the CNN with strong feature extraction capability can lead to significant overfitting. It also comes out that the deep network does not suit for the denoising training of the encoder, because its feedback process could be very long. This makes the encoder difficult to train and achieve good denoising result.

EMCCNN is not deep, which is good for the encoder to achieve good denoising result. On the other hand, convolution kernels of different scales can extract spectral peaks of different quality, which provides sufficient options for the classifier. By weighting the quality of features, the network can select high-quality spectral type features. This design makes EMCCNN achieve the best results in the three SNRs of 1D data.

Through the classification of 1D data with 2D folding, we can find that the 1D spectral data has relevance in 1D scale. Although it is not vertically related after folding, the 2D classification network can still achieve good results. This proves that the 2D classification of 1D data is feasible, especially for the case that the 1D classification is extremely easy to overfit. In this case, folding the data into 2D can effectively eliminate the appearance of overfitting.

Acknowledgement

This paper is supported by the National Natural Science Foundation of China (11473019), Shandong Provincial Natural Science Foundation, China (ZR2017MA046, ZR2018MA032).

References

[1] Simonyan K , Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

[2] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[J]. 2015.

[3] Huang G , Liu Z , Weinberger K Q . Densely Connected Convolutional Networks[J]. 2016.

[4] Yue J , Zhao W , Mao S , et al. Spectral–spatial classification of hyperspectral images using deep convolutional neural networks[J]. Remote Sensing Letters, 2015, 6(6):468-477.

[5] Wen-Yu W , Ge-Lin G , Chun-Yu M A , et al. Extracting Convolutional Features of WDMS Spectra with Anti Bayesian Learning Paradigm[J]. Spectroscopy and Spectral Analysis, 2018.