

# 2DPCANet: Dayside Aurora Classification Based on Deep Learning

Zhonghua Jia<sup>1,2</sup>, Bing Han<sup>1,2(✉)</sup>, and Xinbo Gao<sup>1</sup>

<sup>1</sup> School of Electronic Engineering, Xidian University, Xi'an 710071, China  
bhan@xidian.edu.cn, xbgao@mail.xidian.edu.cn

<sup>2</sup> State Key Laboratory of Remote Sensing Science, Beijing 100101, China

**Abstract.** The mysterious and beautiful aurora represents various physical meaning, thus the classification of aurora images have significant scientific value to human beings. Principal component analysis network (PCANet) has achieved good results in classification. But when using PCANet to extract the image features, it transform original image into a vector, so that the structure information of the image are missing. Compared with PCA, 2DPCA is based on 2D image matrices rather than 1D vectors so that 2DPCA can use the structure information of original image more efficiently and reduce the computational complexity. Based on PCANet, we develop a classification method of aurora images, 2-dimension PCANet (2DPCANet). To evaluate 2DPCANet performance, a series of experiments were performed on two different aurora databases. The classification rate across all experiments was higher using 2DPCANet than PCANet. The experiment results also indicated that the classification time is shorter using 2DPCANet than PCANet.

**Keywords:** Aurora image · Deep learning · Principle component analysis · PCANet · 2DPCANet

## 1 Introduction

Aurora is the magnificent light when the solar wind travels the magnetosphere of high altitude areas near the north and south poles of the earth. The aurora is not only related to the earth's atmosphere and geomagnetic field, but also related to the solar eruption of high-speed charged particles. When charged particles are emitted by the sun toward the earth into the scope of the earth magnetic field, they travel along the earth's magnetic field lines into the upper atmosphere near the north and south poles under the influence of the magnetic field, and then inspire visible light after proton collisions, and finally become a high-profile, we call it aurora.

Aurora phenomenon is not only a simple optical phenomenon, but also an important way for understanding the atmospheric physics. Different forms of the aurora imply different physical meanings. Therefore, the highly efficient classification of aurora images has very important value in scientific research.

From 1964 till now, Akasofu [1], Hongqiao Hu [2], and the Chinese polar research center [3] have divided the aurora into different types. For a long time, the aurora was divided into arc aurora and corona aurora. The corona aurora was further divided

into drapery corona, radial corona and hot-spot corona. In 2015, on the basis of arc aurora and corona aurora, the Chinese polar research center considers that corona aurora contains two types aurora: drapery corona and radial corona. Hot-spot corona aurora is regarded as another aurora. So now the most significance aurora types are arc aurora, drapery corona aurora, and radial corona aurora. According to this, we classify these three types aurora effectively in this paper to discover the mechanism from them and provide an effective analysis for aurora physics research.

In 2004, Syrjasuo et al. [4] have firstly introduced the image processing and machine visual technology into the classification of the aurora image, and employ the texture feature of the aurora image to classify its shape as arc, block, omega and south-north shape. Since the texture structure of the aurora image in the complex background is not clear, the classification accuracy of the aurora image is not high. In 2007, Qian Wang et al. [5] have extracted the gray scale features of the aurora image by the principal component analysis (PCA) to divide it into the arc type, the corona type and the hybrid types, to achieve better classification efficiency. In 2008, Lingjun Gao et al. [6] have proposed the classification method of the aurora image based on Gabor transform to reduce the characteristic redundancy and to ensure the characteristic effectiveness and the classification. In 2009, Rong Fu et al. [7] has combined the analysis of the aurora image with the morphology to greatly improve the classification accuracy between the arc-like and corona aurora images. In 2010, Yuru Wang et al. [8, 9] have obtained the classification algorithm of the aurora image based on X-GLAM characteristics by modifying the GLAM field, in order to improve the classification precision, but its computational complexity is high. In 2013, Bing Han, et al. [10] employed the Salient Coding method to classify the features of aurora images, and Bing Han, et al. [11] proposed an aurora image classification method based on latent dirichlet allocation with saliency information, which improves the classification accuracy of the arc aurora.

According to the analysis of the existing algorithms, there are several critical problems on the highly effective classification of the aurora image, how to effectively extract the aurora feature, how to reduce the algorithm complexity and to improve the algorithm efficiency simultaneously.

With the development of the science and technology, convolutional deep neural network architecture [12], [13] consists of multiple trainable stages stacked on top of each other, followed by a supervised classifier. PCANet cascading two-layer PCA in [14] has been referred as a simple deep learning network to obtain the well efficiency in classification owing to that PCA in [15] has been referred as the data processing method with the advantages of revealing the essence of things and simplifying complex problems. However, in the PCA-based classification technique, the 2-dimension image matrices must be previously transformed into a vector, which ignores its structural information and leads to a high-dimensional image vector space. It is difficult to evaluate the covariance matrix accurately due to its large size and the relatively small number of training samples. While, 2-Dimension PCA (2DPCA) in [16] is a straightforward image projection technique for image feature extraction. Compared with PCA, 2DPCA is based on 2D image matrices rather than 1D vectors so that

2DPCA can use the structure information of original image more efficiently and reduce the computational complexity.

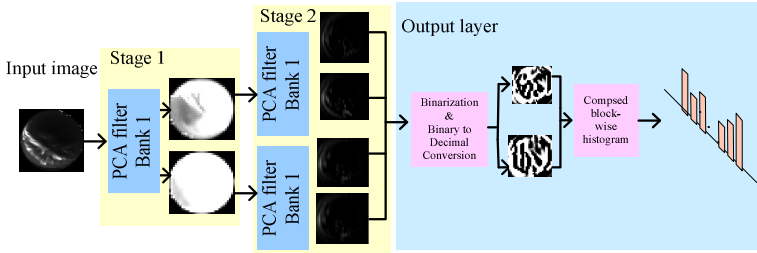
In this paper, we propose deep-learning classifications of the aurora image, two-staged 2DPCANet by full employing its structure information in order to increase the classification accuracy and to reduce the elapsed time.

The remainder of this paper is organized as follows: In section 2, the background of PCANet is reviewed. The ideal of the proposed two-staged 2DPCANet are described in section 3. In section 4, experimental results and analysis are presented. The final one is conclusion.

## 2 The PCA Network

PCANet is a deep learning network for image classification in [14]. It comprises components: cascaded PCA, binary hashing, and block-wise histograms and employed to learn multistage filter banks. This architecture can be designed and learned extremely easily and efficiently.

Assume that  $N$  input training images  $\{\mathcal{I}_i\}_{i=1}^N$  of size  $m \times n$  are given and trained in the PCANet system. The patch size is  $k_1 \times k_2$  at all stages and suppose that the number of filters in layer  $i$  is  $L_i$ .



**Fig. 1.** Illustration of how the proposed PCANet extracts features from an image through three simplest processing components

### 2.1 The First Stage: PCA

1. Take a  $k_1 \times k_2$  patch around each pixel and collect all patches of the  $i$ th input image:  $x_{i,1}, x_{i,2}, \dots, x_{i,mn} \in R^{k_1 \times k_2}$ , where  $x_{i,j}$  denotes the  $j$ th patch in the  $i$ th input image;
2. Calculate patches mean and compute mean-removed patch from each patch to get  $\bar{X}_i = [\bar{x}_{i,1}, \bar{x}_{i,2}, \dots, \bar{x}_{i,mn}]$ , where  $\bar{x}_{i,j}$  is a mean-removed patch.
3. By dealing all input image with the same steps and putting them together, we obtain  $X = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in R^{(k_1 \times k_2) \times N \times mn}$ ;
4. PCA minimizes the reconstruction error within a family of orthonormal filters, i.e.,  $\min_{V \in R^{k_1 \times k_2 \times L_1}} \|X - VV^T X\|_F^2, s.t. V^T V = I_{L_1}$ , where  $I_{L_1}$  is identity matrix of size  $L_1 \times L_1$ ;

5. The PCA filters can be expressed as  $W_l^1 = \text{mat}_{k_1, k_2}(q_l(XX^T)) \in R^{k_1 \times k_2}$ ,  $l=1, 2, \dots, L_1$ , where  $\text{mat}_{k_1, k_2}(v)$  is a function that maps  $v$  to a matrix  $W$ .

## 2.2 The Second Stage: PCA

Just like repeating the process as the first stage, the  $l$ th filter output of the first stage can be  $\mathcal{I}_i^l = \mathcal{I}_i * W_l^1, i=1, 2, \dots, N$ . Just as the first stage, collect all the patches of  $\mathcal{I}_i^l$ , compute mean-removed patch from each patch and form  $\bar{Y}_i^l = [\bar{y}_{i,l,1}, \bar{y}_{i,l,2}, \dots, \bar{y}_{i,l,mn}]$ . The matrix collecting all mean-removal patch of the  $l$ th filter output can be defined as  $Y^l = [\bar{Y}_1^l, \bar{Y}_2^l, \dots, \bar{Y}_N^l] \in R^{(k_1 \times k_2) \times Nmn}$ , and concatenate  $Y^l$  for all the filter outputs as  $Y = [Y^1, Y^2, \dots, Y^{L_1}] \in R^{(k_1 \times k_2) \times L_1 Nmn}$ . Then, obtain the PCA filters of the second stage as  $W_l^2 = \text{mat}_{k_1, k_2}(q_l(Y Y^T)) \in R^{k_1 \times k_2}$ ,  $l=1, 2, \dots, L_2$ . We can achieve  $L_2$  outputs for each input  $\mathcal{I}_i^l$  of the second stage:  $O_i^l = \{\mathcal{I}_i^l * W_l^2\}_{l=1}^{L_2}$ .

## 2.3 Output Layer

For the outputs from the second stage, we binarize these outputs and get  $\{H(\mathcal{I}_i^l * W_l^2)\}_{l=1}^{L_2}$ . Convert binary bits of each pixel from each outputs back into a decimal number. Then, we can get single integer-valued “images”:  $\mathcal{I}_i^l = \sum_{l=1}^{L_2} 2^{l-1} H(\mathcal{I}_i^l * W_l^2)$ . Partition each of the images into  $B$  blocks. Then, concatenate all the  $B$  histograms into one vector after compute the histogram of the decimal values in each block and denote as  $Bhist(\mathcal{I}_i^l)$ . After encoding process, the feature of the  $i$ th input image is defined as:  $f_i = [Bhist(\mathcal{I}_i^1), \dots, Bhist(\mathcal{I}_i^{L_1})]^T \in R^{(2^{L_2})^{L_1 B}}$ .

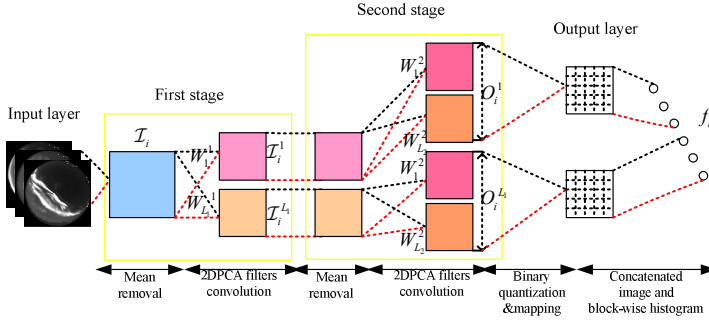
## 3 The 2DPCA Network

2DPCA is based on 2-dimension image matrices rather than 1-dimension vectors so that the structure information can be fully considered and the dimension can be reduced. In this section, we adopted the cascaded 2DPCA, the binary hashing, and block histograms to classify the aurora images. The proposed 2DPCANet model is illustrated in Fig. 2.

Suppose that  $N$  input training images  $\{\mathcal{I}_i\}_{i=1}^N$  of size  $m \times n$  are given and the number of filters in layer  $i$  is  $L_i$ .

### 3.1 The First Stage: 2DPCA

1. For each image, subtract image mean from each image to obtain  $\bar{X}_i$  and putting them together to get  $X = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in R^{(m \times n) \times N}$ ;



**Fig. 2.** The detailed block diagram of the proposed (two-stage) 2DPCANet.

2. 2DPCA minimizes the reconstruction error within a family of orthonormal filters, i.e.,  $\min_{V \in \mathbb{R}^{(m \times n) \times L_1}} \|X - VV^T X\|_F^2, s.t. V^T V = I_{L_1}$ , where  $I_{L_1}$  is identity matrix of size  $L_1 \times L_1$ ;
3. The 2DPCA filters can be expressed as  $W_l^1 = \text{mat}_{m,n}(q_l(XX^T)) \in \mathbb{R}^{m \times n}$ ,  $l=1,2,\dots,L_1$ , where  $\text{mat}_{m,n}(v)$  is a function that maps  $v$  to a matrix  $W$ , and  $q_l(XX^T)$  is the  $l$ th primal eigenvectors of matrix  $XX^T$ .

### 3.2 The Second Stage: 2DPCA

Almost repeating the same process as the first stage, set the  $l$ th filter output of the first stage is  $\mathcal{I}_i^l = \mathcal{I}_i * W_l^1, i=1,2,\dots,N$ . Before convolving  $\mathcal{I}_i$  with  $W_l^1$ , the boundary of  $\mathcal{I}_i$  is zero-padded. Just as the first stage, we define  $Y^l = [\bar{Y}_1^l, \bar{Y}_2^l, \dots, \bar{Y}_N^l] \in \mathbb{R}^{(m \times n) \times L_1 N}$  for the matrix collecting all mean-removed image of  $l$ th filter output, and concatenate  $Y^l$  for all the filter outputs as  $Y = [Y^1, Y^2, \dots, Y^{L_1}] \in \mathbb{R}^{(m \times n) \times L_1 N}$ . Then, obtain the 2DPCA filters of the second stage as  $W_l^2 = \text{mat}_{m,n}(q_l(Y Y^T)) \in \mathbb{R}^{m \times n}$ ,  $l=1,2,\dots,L_2$ . We will have  $L_2$  outputs for each input of the second stage:  $O_i^l = \{\mathcal{I}_i^l * W_l^2\}_{l=1}^{L_2}$ .

One can simply repeat the above process to build more (2DPCA) stages.

### 3.3 Output Layer

Binaries the outputs from the second stage and get  $\{H(\mathcal{I}_i^l * W_l^2)\}_{l=1}^{L_2}$ . Convert binary bits of each pixel from each outputs bake into a decimal number to get  $\mathcal{T}_i^l = \sum_{t=1}^{L_2} 2^{l-1} H(\mathcal{I}_i^l * W_l^2)$ , whose every pixel is an integer in the range  $[0, 2^{L_2-1}]$ . Partition each of the images into  $B$  blocks. Then, concatenate all the  $B$  histograms into one vector after compute the histogram of the decimal values in each block and denote as

$Bhist(\mathcal{T}_i^l)$ . After encoding process, the feature of the  $i$ th input image is defined as:

$$f_i = [Bhist(\mathcal{T}_i^1), \dots, Bhist(\mathcal{T}_i^{L_i})]^T \in R^{(2^{L_2})_{L_i} B}.$$

## 4 Experiment Results and Analysis

In this section, the proposed 2DPCANet classification scheme is evaluated by conducting several classification and comparison experiments, and the experiment results are exhibited and analyzed then.

### 4.1 Dataset

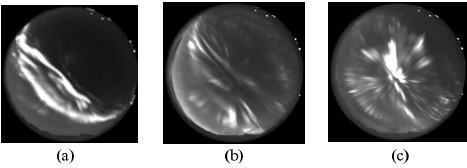
The aurora data were obtained from the all-sky imagers at Yellow River Station (YRS) in Ny-lesund, Svalbard. The optical instruments at YRS capture photoemissions at 427.8, 557.7, and 630.0nm, and the time interval between every two images is 10 second. Dayside aurora images are divided into two categories: arc aurora, and corona aurora. The corona aurora can be further divided into drapery corona and radial corona. Sample images from the three aurora shape categories are shown in Fig. 3.

Experiments in this article, we adopt two kinds of aurora database. They were concentrated on the dayside aurora and selected from December 2003 to January 2004 to constitute the two different databases. Only auroras at 557.7 nm were adopted in consideration of the image characteristics.











The database I where images are unrelated in time domain contains 2400 aurora images which have 800 arc aurora images, 1600 corona aurora images (800 drapery corona and 800 radial corona). The aurora images of this database are all similar with standard aurora in morphology. The database II where images are related in time domain is larger than the database I and contains 11133 images. In database II, images are selected due to not only their morphology similarity, but also their physical development of an aurora event. That is to say, the time interval of images with same class is so small in the database II.

Compare two kinds of database on time and show in Fig. 4. We select the images form an aurora event. And different images with different time from this aurora event belong to the database I and the database II. The numbers over each aurora image represent the time when the image was captured. When the blank space in the row of the database I are filled with a tick, it illustrates that the aurora image over the blank space are belonged to the database I. And the same to the row of the database II.

On the basis of two different aurora databases, we randomly select different numbers of aurora images used for training and testing. In addition, the numbers of aurora images for training are three times than the numbers of aurora images for testing. The labels of arc, drapery corona and radial corona are 1, 2 and 3, respectively.



**Fig. 3.** Typical categories of aurora. Columns from left to right are: (a) arc, (b) drapery corona, and (c) radial corona.

	1401	1411	1421	1431	1441	1451	1501	1511	1521	1531
An aurora event										
Database I	✓									✓
Database II	✓	✓					✓	✓		✓

**Fig. 4.** An aurora event (arc aurora).

4.2 Classification Experiment for Parameters Setting

In the experiments, we evaluate the performance of 2DPCANet-3, 2DPCANet-2, 2DPCANet-1, PCANet-2, and PCANet-1. We select 600 aurora images used in the experiments. The numbers of arc, drapery corona, and radial corona are 200, respectively. We select randomly 150 aurora images form each category used for training and the rest used for testing. In this part, the experiments are handled used this data set.

We deal with the image classification used SVM classifier. Hence, the parameters in the SVM classifier should first be selected before the classification experiments are carried out. We conduct ten times cross validation. The selection results are shown in Fig.5. According to the performance of ten times cross validation, training and testing sets with optimal SVC parameters will be employed for constructing ROC curves.

Then, we should found the optimal number of filters in the different stages in different layers. We vary the number of filters in the first stage  $L_1$  from 2 to 12 for one-staged networks. When considering two-staged networks, we set  $L_2=8$  and vary  $L_1$  from 4 to 24. At last, we set  $L_2=14, L_3=8$  and vary  $L_1$  from 4 to 20 to adjust the number of filters at the first stage in three-staged networks. The results of the number of filters are shown in Fig. 6. One can see that 2DPCANet achieves better results than PCANet. Moreover, the accuracy of 2DPCANet and PCANet (for all staged networks) increases for larger  $L_1$ . However, three-staged 2DPCANet achieves so lower classification accuracy rat.

The optimal number of filters and the optimal parameters in SVM will improve the performance of feature analysis methods.

For verifying the performance of the proposed method and the other models, we take a group of comparison experiments on the data set according to the optimal parameters that we select. Due to the performance of 2DPCANet-3, we only compare 2DPCANet-2, 2DPCANet-1, PCANet-2, and PCANet-1. The testing results of ROC curves are shown in Fig.7.

It can be seen from Fig. 7 that 2DPCANet-2 possesses the biggest area under curve which shows the best performance in classification. For getting more statistical and persuasive results, Table 1 shows the average classification accuracy rate of different methods. More intuitive results can be found in Table 1. Each method is conducted 200 times and the accuracy is the mean results of the 200 times classification procedures. We also find that 2DPCANet-2 performed better than others.

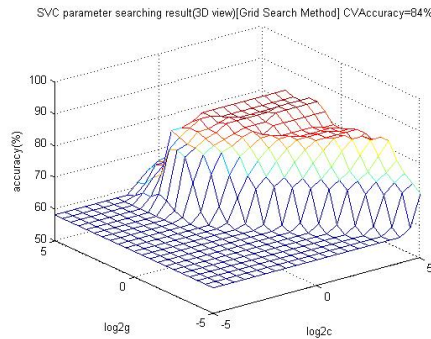


Fig. 5. SVM parameter selection with three dimensional view of (two-stage) 2DPCANet

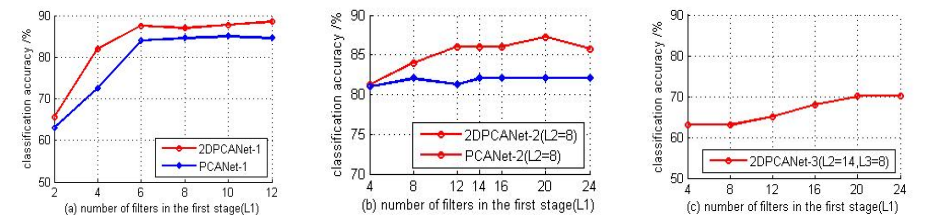


Fig. 6. Classification accuracy of 2DPCANet and PCANet for varying number of filters in the first stage

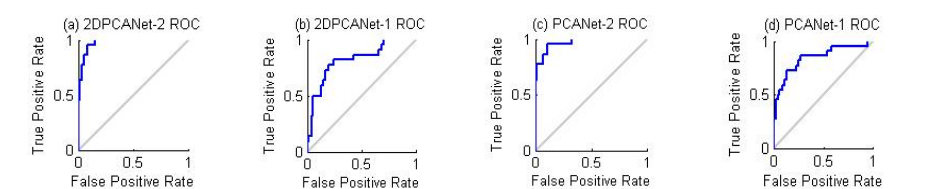


Fig. 7. The ROC curves of different classification methods



**Table 1.** The accuracy and generalized running time of different methods

Classification method	Feature dimensions	Accuracy (%)	Generalized running time(s)
2DPCANet-2	75264	<b>83</b>	<b>1.00</b>
2DPCANet-1	5376	69	<b>0.52</b>
PCANet-2	75264	80	1.10
PCANet-1	5376	64	0.61

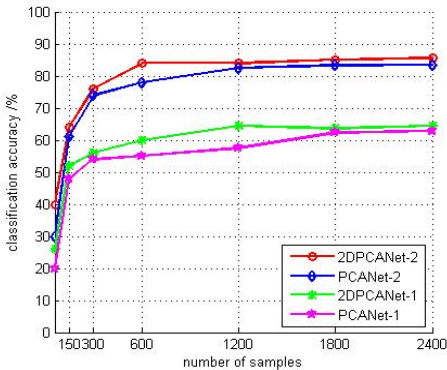
4.3 Classification Experiment on the Database I

In order to evaluate the validity of the proposed method on the database I, experiments are designed and conducted. First, we select randomly different numbers of aurora images for experiments. The image datasets are shown in Table 2. 2DPCANet-2 is compared with 2DPCANet-1, PCANet-2, and PCANet-1.

We set the optimal parameters of all method and SVM classifier. Then, the experiments are conducted 200 times, and the final results are the average of them. Fig. 8 shows the performance of our method and the other methods with different numbers dataset. We observe that, the classification accuracy of our proposed method is higher than the other methods. And 2DPCANet-2 acquired smooth faster than others. In addition, we compared our method with PCANet-2 in the generalized running time, as shown in Table 3. It illustrates that our method spend less time than others and the advantage of our method are obviously with the increasing of the numbers of dataset.

**Table 2.** Dataset of aurora images

Num of data-set	2400	1800	1200	600	300	150	60
Arc	800	600	400	200	100	50	20
Hot-spot	800	600	400	200	100	50	20
Corona	800	600	400	200	100	50	20



**Fig. 8.** Average classification accuracy of different representations on the database I

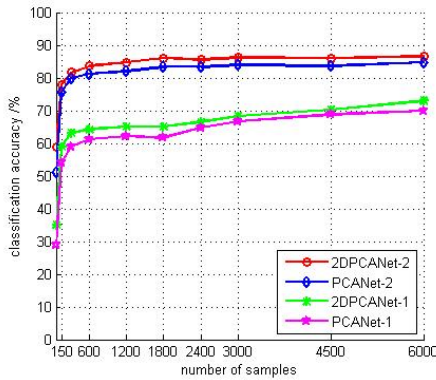
**Table 3.** The generalized running time between 2DPCANet and PCANet

Num of dataset	2400	1800	1200	600	300	150	60
2DPCANet-2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
PCANet-2	1.20	1.19	1.17	1.03	1.06	1.07	1.02

**4.4 Classification Experiment Based on the Database II**

Imitate the experiments on the database I. We also select randomly different numbers of aurora images consisting datasets. As well as selecting the 8 sets from the database I, we select three more datasets from the database II to experiment. Compared 2DPCANet-2 with 2DPCANet-1, PCANet-2, and PCANet-1, we obtain the classification accuracy shown in Fig. 9 and generalized running time shown in Table 3. Obviously, the classification accuracy of 2DPCANet-2 is higher and the running time is smaller than other methods.

In contrast to PCANet, 2DPCANet has two important advantages. First, it reduces the data dimension of the aurora image by fully considering its structure information. Second, it required less time to determine the corresponding eigenvectors.



**Fig. 9.** Average classification accuracy of different representations on the database II

**Table 4.** The generalized running time between 2DPCANet and PCANet

Num of dataset	6000	4500	3000	2400	1800	1200	600	300	150	60
2DPCANet-2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
PCANet-2	1.30	1.27	1.12	1.09	1.08	1.13	1.14	1.09	1.06	1.03

## 5 Conclusions

In this paper, the structure of the cascaded 2DPCA, the binary hashing, and block histograms has been employed to express the features of the aurora images, and the 3-type classification operation of the aurora image has been executed in terms of the features inputted into the SVM classifier. From the experimental results, the proposed deep learning model, 2DPCANet, has increased the classification accuracy of the aurora image and reduced the running time. With the increasing number of aurora image, 2DPCANet can contribute to the research on aurora.

Although the 2DPCANet processing of the aurora image increases the classification accuracy, the 2DPCA just employs part structure information of the aurora image, which must be solved in the future works. In addition, our future work also includes applying our method to other more databases.

## References

1. Akasofu, S.I.: The development of the auroral substorm. *Planetary and Space Science* **12**(4), 273–282 (1964). doi:10.1016/0032-0633(64)90151-5
2. Hu, H.Q., Liu, R.Y., Wang, J.F., Yang, H.G.: Statistic characteristics of the aurora observed at Zhongshan Station. *Antarctica. Chinese Journal of Polar Research* **11**(1), 8–18 (1999). (in Chinese with English abstract)
3. Hu, Z.J., Yang, H., Huang, D., Araki, T., Sato, N., Taguchi, M., Seran, E., Hu, H., Liu, R., Zhang, B., Han, D., Chen, Z., Zhang, Q., Liang, J., Liu, S.: Synoptic distribution of day-side aurora: Multiple-Wavelength all-sky observation at Yellow River Station in Ny-Ålesund, Svalbard. *Journal of Atmospheric and Solar-Terrestrial Physics* **71**(8–9), 794–804 (2009). doi:10.1016/j.jastp.2009.02.010
4. Syrjäso, M.T., Donovan, E.F.: Diurnal auroral occurrence statistics obtained via machine vision. *Annales Geophysicae* **22**(4), 1103–1113 (2004). doi:10.5194/angeo-22-1103-2004
5. Wang, Q., Liang, J.M., Gao, X.B., Yang, H.G., Hu, H.Q., Hu, Z.J.: Representation feature based aurora image classification method research. In: *Proc. of the 12th National Solar-Terrestrial Space Physics Academic Conf.*, vol. 71 (2007)
6. Gao, L.J.: Dayside aurora classification based on gabor wavelet transformation. MS. Thesis. Xi'an: Xidian University (2009)
7. Fu, R., Li, J., Gao, X.B., Jian, Y.J.: Automatic aurora images classification algorithm based on separated texture. In: *Proc. of the 2009 IEEE Int'l Conf. on Robotics and Biomimetics (ROBIO)*, pp. 1331–1335. IEEE (2009). doi:10.1109/ROBIO.2009.5420722
8. Wang, Y.R., Gao, X.B., Fu, R., Jian, Y.J.: Dayside corona aurora classification based on X-gray level aura matrices. In: *Proc. of the ACM Int'l Conf. on Image and Video Retrieval*, pp. 282–287. ACM (2010). doi:10.1145/1816041.1816083
9. Wang, Y.R.: Dayside aurora classification based on X-gray level aura matrices. MS. Thesis. Xi'an: Xidian University (2011) (in Chinese with English abstract)
10. Han, B., Qiu, W.L.: Aurora images classification via features salient coding. *Journal of Xidian University* **40**(6), 1001–2400 (2013). doi:10.3969/j.issn.1001-2400.2013.06.030
11. Han, B., Yang, C., Gao, X.B.: Aurora image classification based on LDA combining with saliency information. *Ruan Jian Xue Bao Journal of Software* **24**(11), 2758–2766 (2013). (in chinese) <http://www.jos.org.cn/1000-9825/4481.html>

12. Goodfellow, I.J., Warde-Farley, D., Mirza, M., Courville, A., Bengio, Y.: Maxout networks. In: ICML (2013)
13. Kavukcuoglu, K., Sermanet, P., Boureau, Y., Gregor, K., Mathieu, M., LeCun, Y.: Learning convolutional feature hierarchies for visual recognition. In: NIPS (2010)
14. Jolliffe, I.T.: Principal component analysis, 2nd edn. Springer, New York (2002)
15. Chan, T.-H., Jia, K., Gao, S., Lu, J., Zeng, Z., Ma, Y.: PCANet: A Simple Deep Learning Baseline for Image Classification? Posted on arXiv.org and submitted to TIP, August 2014
16. Yang, J., Zhang, D., Frangi, A.F., Yang, J.: Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Trans. Patt. Anal. Mach. Intell.* **26**(1), 131–137 (2004)