

# Photometric redshift estimation of galaxies with Convolutional Neural Network

Yong-Huan Mu<sup>1</sup>, Bo Qiu<sup>1</sup>, Jian-Nan Zhang<sup>2</sup>, Jun-Cheng Ma<sup>1</sup> and Xiao-Dong Fan<sup>1</sup>

<sup>1</sup> School of Electronics and Information Engineering, Hebei University of Technology, Tianjin 300401, China; [qiubo@hebut.edu.cn](mailto:qiubo@hebut.edu.cn)

<sup>2</sup> CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; [jnzhang@bao.ac.cn](mailto:jnzhang@bao.ac.cn)

Received 2019 April 24; accepted 2019 December 11

**Abstract** The abundant photometric data collected from multiple large-scale sky surveys give important opportunities for photometric redshift estimation. However, low accuracy is still a serious issue in the current photometric redshift estimation methods. In this paper, we propose a novel two-stage approach by integration of Self Organizing Map (SOM) and Convolutional Neural Network (CNN) methods together. The SOM-CNN method is tested on the dataset of 150 000 galaxies from Sloan Digital Sky Survey Data Release 13 (SDSS-DR13). In the first stage, we apply the SOM algorithm to photometric data clustering and divide the samples into early-type and late-type. In the second stage, the SOM-CNN model is established to estimate the photometric redshifts of galaxies. Next, the precision rate and recall rate curves (PRC) are given to evaluate the models of SOM-CNN and Back Propagation (BP). It can be seen from the PRC that the SOM-CNN model is better than BP, and the area of SOM-CNN is 0.94, while the BP is 0.91. Finally, we provide two key error indicators: mean square error (MSE) and Outliers. Our results show that the MSE of early-type is 0.0014 while late-type is 0.0019, which are better than the BP algorithm 22.2% and 26%, respectively. When compared with Outliers, our result is optimally 1.32%, while the K-nearest neighbor (KNN) algorithm has 3.93%. In addition, we also provide the error visualization figures about  $\Delta Z$  and  $\delta$ . According to the statistical calculations, the early-type with an error of less than 0.1 accounts for 98.86%, while the late-type is 99.03%. This result is better than those reported in the literature.

**Key words:** galaxies: distances and redshifts — techniques: photometric — method: data analysis

## 1 INTRODUCTION

A galaxy can be seen as a celestial system containing a variety of complex materials and galaxies are the most important class of cosmic celestial (Zhang et al. 2010). At present, humans use advanced telescopes (Wu et al. 2012), such as Anglo-Australian Telescope in the Two Degrees Field (2DF) Galaxy Redshift Survey and a 2.5 m telescope in the Sloan Digital Sky Survey (SDSS). The total number of galaxies has exceeded 100 billion.

In the visible band, the wavelength of spectrum will increase, which looks like moving toward the red side of the band because the star is flying away. This phenomenon is called “redshift.” The redshift is a very important parameter of celestial bodies, while the spectral redshift and

photometric redshift are actually calculated from two different methods based on different datasets. The spectral redshift is derived from spectra that are obtained by spectroscopic observations, while the photometric redshift is derived from photometry data of galaxies or quasars in multiple bands by a large CCD camera. Photometric redshift estimation has become an important technique in astronomy (Scranton et al. 2005; Myers et al. 2006; Hennawi et al. 2006; Giannantonio et al. 2008) because it enables us to measure the distance of far objects through a spectroscopic survey. Furthermore, the photometric redshift has a deeper limit magnitude, which also includes the complete SED (Spectral Energy Distribution) of observed objects and allows us to identify the features of certain redshift. Therefore, it is very important to study the photo-

metric redshift targeting galaxy evolution studies and the statistics of various attributes of galaxies.

There are two approaches to estimate photometric redshift: template fitting (e.g., Bruzual & Charlot 1993) and empirical training set (e.g., Coleman et al. 1980). The SED method is the most representative approach of template-fitting (Bolzonella et al. 2000). The other method is the  $x^2$  minimization technique (Wu et al. 2004). Although it is easy to estimate photometric redshift, the accuracy of this approach strongly depends on the templates. Meanwhile, the training set approach is data-driven rather than template-driven. The most common algorithms are as follows. For example, nearest neighbor algorithm (Ball et al. 2008), neural network method (Abdalla et al. 2008), linear spectral connectivity analysis (Freeman et al. 2009), ArborZ boosted decision tree method (Gerdes et al. 2010), random forest (Almosallam et al. 2016), and the support vector machine method (Cavuoti et al. 2017).

Convolutional Neural Network (CNN, Collister & Lahav 2004) has been used in astronomy widely, such as spectral classification and galaxies classification. However, this method has not been applied to the prediction of photometric redshift. A self organizing map (SOM) is able to learn independently and automatically adjust network parameters and structures according to sample characteristics, which can improve the efficiency of CNN. Therefore, in this work, SOM-CNN, which is a ‘mixed’ technique is used to estimate photometric redshift based on the SDSS dataset (Han et al. 2016).

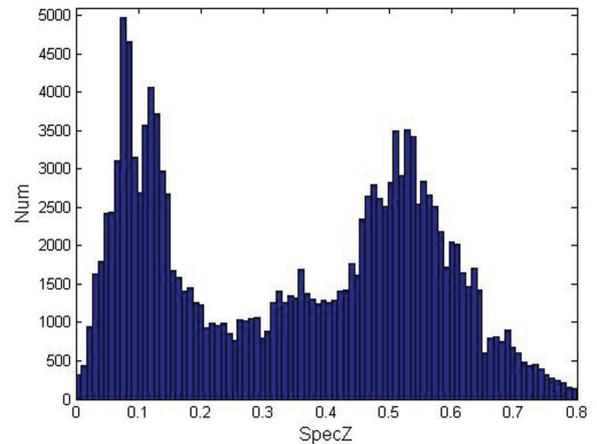
This paper provides a novel two-stage photometric redshift estimation approach (i.e., the integration of CNN and SOM), to improve the accuracy of estimation. The structure of the rest of this paper is as follows. Section 2 describes the original dataset after preprocessing. Section 3 presents a brief overview of CNN, SOM and our estimation model of CNN-SOM. The experiments and results are provided and discussed in Section 4. The conclusions are given in Section 5.

## 2 DATA

### 2.1 SDSS-DR13 Photometric Data

The data that we have used in this paper are from SDSS (York et al. 2000), which covers more than a quarter of the sky. In SDSS-DR13, the spectral data has reached more than 4 million, which include the spectral data of more than 2.4 million galaxies (Gao et al. 2017).

In this paper, 150 000 galaxies are selected from the SDSS-DR13 as the initial samples, which include the



**Fig. 1** Redshift distribution histogram of original samples of galaxies. The redshift distribution for 150 000 galaxies of the whole sample is derived from the SDSS-DR13, and the spectral redshift (Spec Z) is in the range of [0, 0.8].

spectral redshift of the galaxies and the photometric values of five bands:  $\text{petroMag}_u$ ,  $\text{petroMag}_g$ ,  $\text{petroMag}_r$ ,  $\text{petroMag}_i$ ,  $\text{petroMag}_z$ . The redshift distribution for the whole sample is shown in Figure 1.

### 2.2 Data Preprocessing

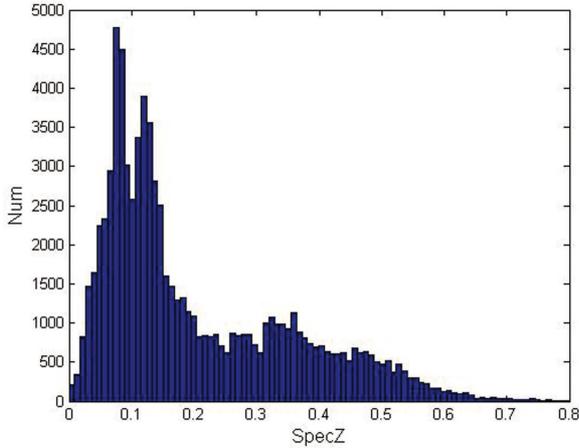
Due to the complexity of the data and the existence of various cosmic noises and errors of the observation system, data preprocessing is a very important job in the photometric redshift estimation.

For the experiment, the sample is processed for extinction at first. The extinction values at each band are  $\text{extinction}_u$ ,  $\text{extinction}_g$ ,  $\text{extinction}_r$ ,  $\text{extinction}_i$  and  $\text{extinction}_z$ . Equation (1) can be used to calculate the photometric redshift after extinction, where  $\text{petroMag}_x$  is the original photometric redshift for each band, while the  $\text{petroMag}_{x'}$  denotes photometric redshift after extinction removed

$$\text{petroMag}_{x'} = \text{petroMag}_x - \text{extinction}_x. \quad (1)$$

After extinction, the sample data are filtered with a standard that the photometric errors of each band are greater than zero, and the upper limit is set at 0.25. Preprocessing also helps to remove the bad samples of  $-9999$ , and also ensures the distribution of sample in each redshift segment. The redshift distribution histogram of the galaxies after data preprocessing is shown in Figure 2.

There is a considerable difference in Figure 1 and Figure 2, the reasons for the large difference of redshift distribution before and after cutting bad data are as follows: first, because electromagnetic waves from distant galax-



**Fig. 2** Redshift distribution histogram of the galaxies after data preprocessing. After two-step data preprocessing, the number of suitable galaxies is 81 142. There is a considerable difference of redshift distribution before and after cutting bad data.

ies are partly absorbed and scattered by interstellar gas or dust, the photometry are weakened after extinction; second, there are bad samples of  $-9999$  in the original SDSS-DR13 database. Consequently, the data is reduced after filtering the bad samples. In addition, because the samples of high-redshift galaxies are in deep air and have more complex properties, the spectral data that can be collected in the measurement process are relatively rare. Therefore, the redshift is significantly reduced after the data processing.

### 3 METHODOLOGY

To reduce the complexity of the data and improve the precision of the regression model, a new methodology is proposed according to the characteristics of galaxies. The SOM method is used at first to classify the galaxies into early-type galaxies and late-type galaxies. Then, the photometric redshift of galaxies is estimated based on the SOM-CNN algorithm. To clearly describe the methods used in our work, a flow chart is given in Figure 3.

#### 3.1 Classification with SOM

Clustering analysis is an important method in data mining (Ai et al. 2017). The clustering algorithm can effectively reduce the complexity of data especially for the case of huge data volume, large feature parameters and unknown sample categories. Clustering analysis of the sample is a key step in photometric redshift prediction, because the data of the galaxies have high dimensional and nonlinear characteristics.

There are many standards for galaxy classification, including traditional morphological classification, spectral

classification, and color classification (Shaun Cole et al. 1998). Because the color of galaxies is closely related to the spectral energy distribution, we can classify our work according to the color features.

SOM (Cho et al. 2015; Zhu 2014) is an efficient dimensionality reduction model. Furthermore, the SOM is able to learn independently and automatically adjust network parameters and structures according to sample characteristics. In the case of high-dimensional input, the data can be converted to low-dimensional for clustering, which can improve the efficiency of the algorithm. Therefore, the SOM method is adopted to classify, and the galaxies can be divided into early-type galaxies and late-type galaxies by SOM. A scatter diagram based on the color clustering in the  $u-r$  and  $r-g$  color spaces is shown in Figure 4(a). To better visualize the galaxies' classification results, the number of early and late galaxies after clustering is drawn in the  $u-r$  color space. The distribution histogram is shown in Figure 4(b).

It can be seen from Figure 4(a) that the SOM algorithm clearly divides the samples in the color space, and the boundaries between the classes are obviously approaching a linear function. In addition, Figure 4(b) gives a distribution histogram in  $u-r$  color space. Among them, the number of early-type galaxies is 63 487, while the late-type galaxies is 17 655. The clustering results are consistent with the theories that there are more early-type galaxies than the late-type galaxies.

#### 3.2 Estimation with CNN

In recent years, deep learning has been widely used. Convolutional Neural Network (CNN), which was originally conceived as a model of the brain, knowledge is obtained through the CNN learning process (Haykin 1994). CNN is an important improved multi-layer feed-forward neural network (Babu et al. 2016; Li et al. 2017; Moon et al. 2016). In addition, the CNN has sparse connections, which can simplify network parameters compared with traditional neural networks. The basic structure of CNN is shown in Figure 5.

In addition to the features of multi-feature extraction, CNN has other advantages as follows: (1) the weight sharing feature reduces the parameters of CNN learning, and so shortens the training time of the model, which helps to establish a deep neural network model; (2) the original information can be directly input, and then the CNN can automatically extract features from the training data.

CNN has been widely used in image processing, and the effect is ideal. However, by reading the relevant ar-

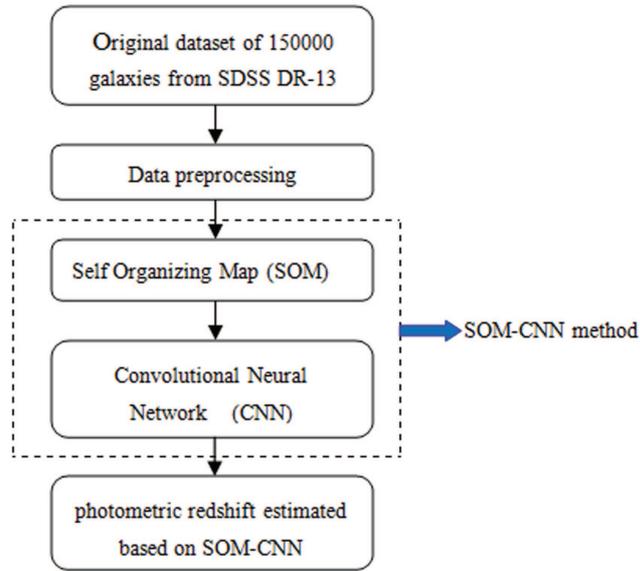


Fig. 3 A flow chart of the SOM-CNN methodology in photometric redshift estimation.

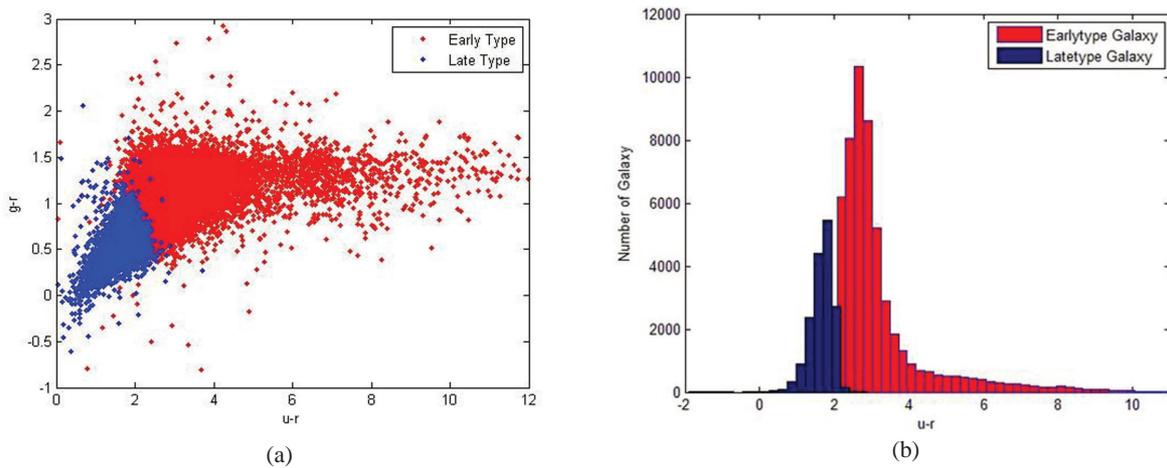


Fig. 4 Clustering results based on the SOM algorithm. The blue part indicates late-type galaxies, while the red part indicates early-type galaxies. Panel (a) indicates the scatter diagram of galaxies and panel (b) shows distribution histogram in  $u - r$  color space, which clearly pointed out that the number of early-type galaxies is more than late-type galaxies.

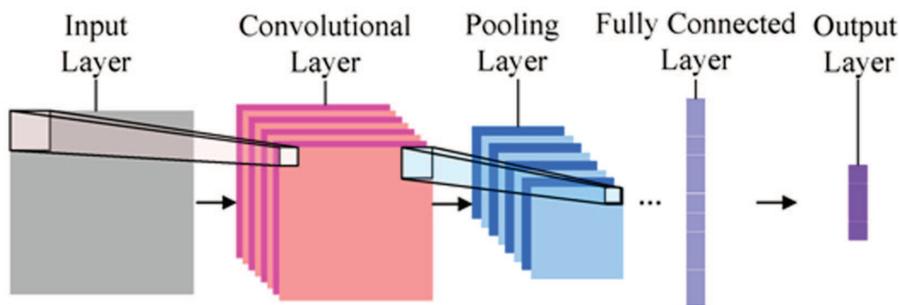


Fig. 5 Basic structure diagram of convolutional neural network. It consists of five basic layers, such as input layer, convolutional layer, pooling layer, fully connected layer and output layer. The specific representation of each layer is shown in Fig. 5.

ticles, it is known that CNN has not been used for redshift estimation. Considering the advantages of CNN, we attempt to adopt the CNN algorithm to estimate the photometric redshift of galaxies. CNN usually adopts supervised learning (Li et al. 2007). So it needs to include the desired result in training data. Here, each input data corresponds to one galaxy and the output data is the photometric redshift. When the network is trained successfully, we can give a new input into the CNN model, and then the approximate result can be obtained of the real redshift.

### 3.3 SOM-CNN Photometric Redshift Estimation

We provide a photometric redshift prediction model of CNN and SOM, which integrates algorithm of CNN and SOM to estimate photometric redshift, as follows. A seven-layer CNN will be built based on the Tensorflow framework to predict the photometric redshift of galaxies (Zhang et al. 2017). The photometric redshift prediction model based on SOM-CNN includes one input layer, two convolution layers, two pooling layers, a fully connected layer and an output layer. The prediction model structure is shown in Figure 6.

**Input layer:** A matrix consisting of the number of samples. Each sample contains 20 inputs, such as photometric data, photometric errors of five bands and 10 color features.

**Convolution layer:** The convolution kernel with a size of  $1 \times 2$  is selected to perform convolution operations. Generally, the convolution layer is composed of many feature faces, and the neurons on each feature surface are connected to the upper layer by the convolution kernel, and then the convolution operation is performed.

**Pooling layer:** The maximum pooling method is adopted and the ReLU function is selected as the activation function. Compared with other activation functions, the ReLU function is simple in calculation and does not need to normalize the input. It is the most widely used in activation functions.

**Fully connected layer:** The number of fully connected layers can be one layer or multiple layers, and each layer uses an activation function to integrate the results of feature extraction. The prediction model contains one layer in our experiment. After two sets of convolution–pooling operations, the extracted features are input to the fully connected layer, and then the predicted values are fitted based on the ReLU function.

**Output layer:** The mean square error is defined as the loss function. Training the network model by minimizing the loss function, the photometric redshift prediction results are obtained through the output layer.

## 4 EXPERIMENT AND RESULTS

In the experiment, the sample of each galaxy is divided into training sets and test sets, with 80% and 20% respectively. In addition, the 20% of the training set is used as a verification set. In every galaxy, the five bands of  $u, g, r, i, z$  are provided. So we can get 10 color features, such as  $u - g, u - r, u - i, u - z, g - r, g - i, g - z, r - i, r - z, i - z$ . In addition, we use five bands values and 10 color features as the input in our experiment, then the corresponding redshift as an estimation output. Next the experiments and results of BP and SOM-CNN are presented, then the analysis are provided in the following subsections.

### 4.1 Photometric Redshift Estimation Experiment by BP

Three layers of BP neural network structure is adopted in the experiment. First of all, the number of hidden layer neurons in early-type galaxies is 19, while the late-type galaxies is 28. Then the tansig function is used to be transfer function, and the trainlm is chosen as the training function to estimate photometric redshift. The experiment results based on BP algorithm are shown in Figure 7.

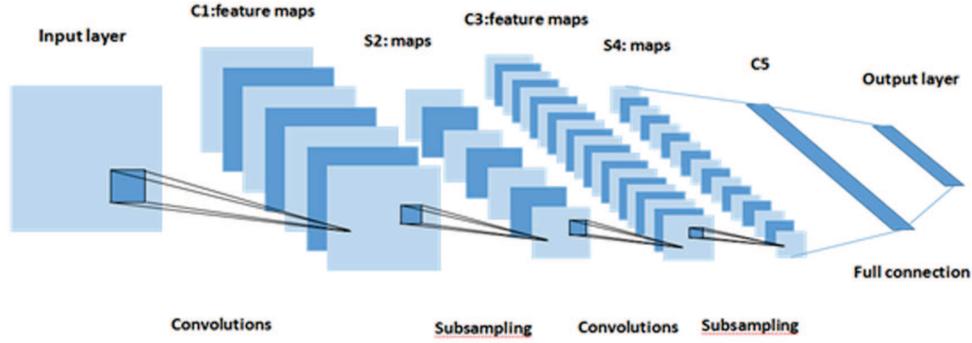
### 4.2 Photometric Redshift Estimation Experiment by SOM-CNN

Based on the SOM-CNN model described in Section 3.3, the experiment of photometric redshift estimation was carried out. First, the learning rates in early-type and late-type galaxies are set to 0.2 and 0.4, respectively. Then the number of trainings is set to 200. Finally the predicted results and expected values are plotted as a scatter diagram, in which the proportional function is used as the fitting center line, the spectral redshift is taken as the abscissa, and the photometric redshift is taken as the ordinate. Thus the optimal training results are obtained as shown in Figure 8.

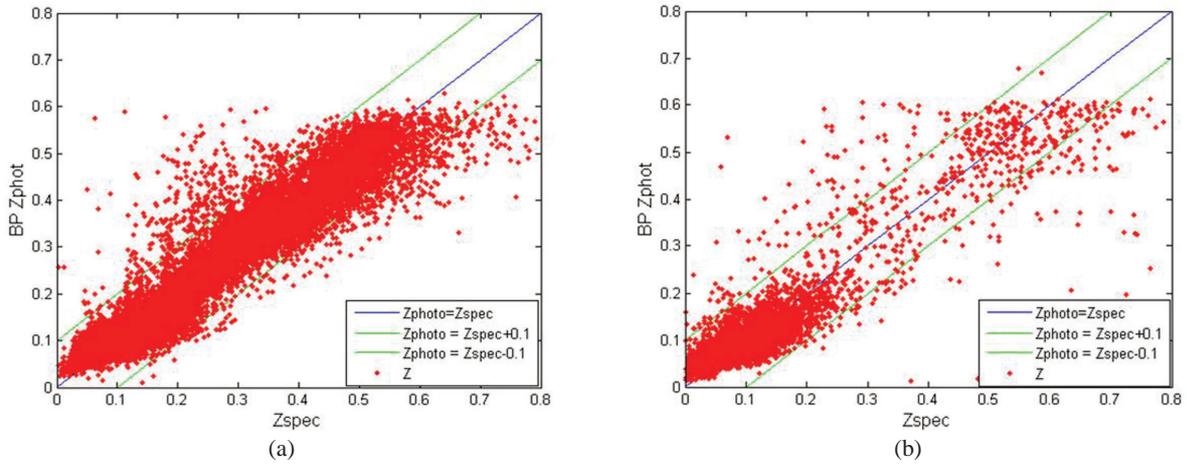
It can be seen from Figure 8 that the photometric redshift estimation model based on SOM-CNN has achieved good prediction for both early-type and late-type galaxies. Most of the dispersion is controlled within the range of 0.1. Furthermore, the two types of galaxies have better convergence and prediction in the part of  $z < 0.8$ .

### 4.3 Comparison and Analysis

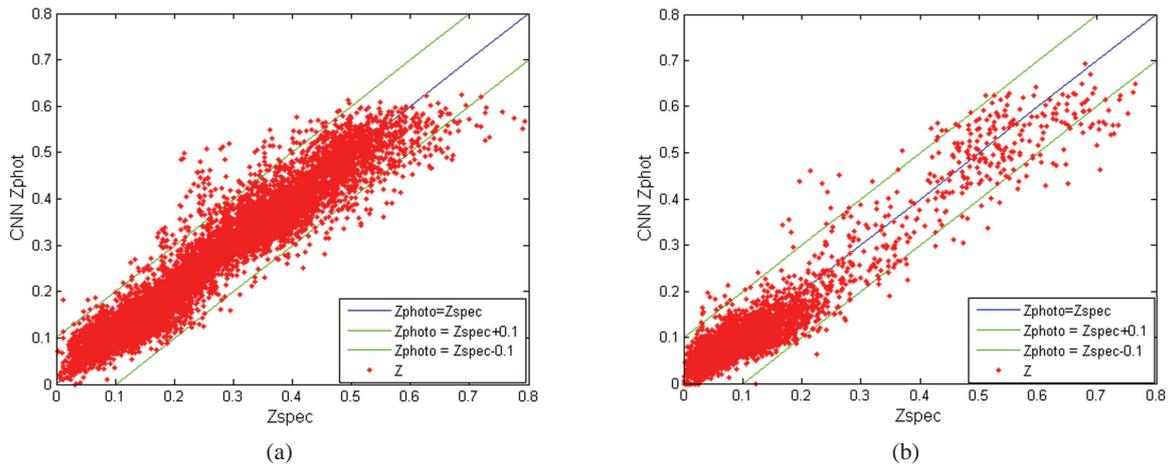
In the following part, the BP and SOM-CNN models are evaluated first. To analyze the experimental results more accurately, the error comparisons and visualizations analysis are provided of BP and SOM-CNN algorithm.



**Fig. 6** Photometric redshift prediction model. C1 and C3 are convolutional layers, S2 and S4 are pooling layers. Between C1 and S2 is a subsampling, which is the same as between C3 and S4. C5 is a full connection layer.



**Fig. 7** Photometric redshift estimation by the BP algorithm. Panel (a) represents early-type galaxies, while panel (b) shows late-type galaxies. The blue line indicates that  $Z_{\text{photo}}$  is equal to  $Z_{\text{spec}}$  and the green lines represent the error tolerance interval.

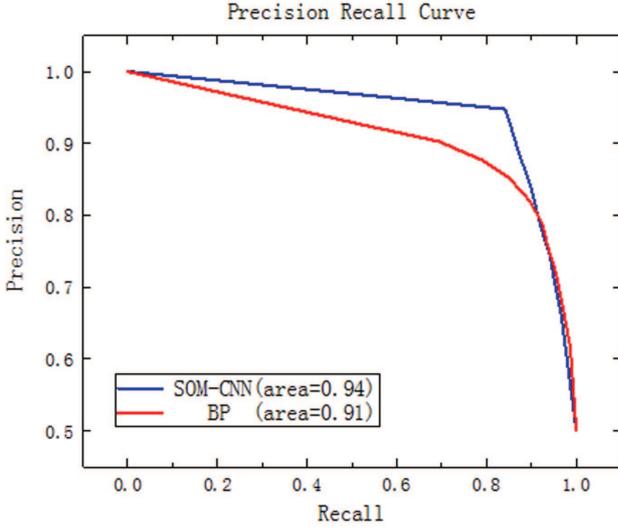


**Fig. 8** Photometric redshift estimation by the SOM-CNN algorithm. Panel (a) demonstrates the estimation results of early-type galaxies, while panel (b) represents late-type galaxies.

#### 4.3.1 Model evaluation

Usually we can evaluate the generalization error through experiments and then make a choice. This requires not only

effective and feasible estimation methods but also performance measurements. Based on BP and SOM-CNN prediction algorithms, four scenarios will be generated in the prediction process: true positive (TP), false positive (FP),



**Fig. 9** The precision rate and recall rate curves of BP and SOM-CNN models. The *blue line* shows SOM-CNN model and the *red line* represents the BP model. The PRC area of the two models are 0.94 and 0.91, respectively.

true negative (TN) and false negative (FN). Equation (2) is used to calculate the precision, while Equation (3) can be used to calculate the recall:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (3)$$

$$\text{area} = \frac{1}{2} \sum_{i=1}^{m-1} (x_{i+1} - x_i)(y_i + y_{i+1}). \quad (4)$$

Precision rate and recall rate are a pair of contradictory measurements. We have provided the precision rate and recall rate curves (PRC), which can be used to evaluate the two prediction models. The PRC of BP and SOM-CNN models is shown in Figure 9. When there is a crossover of PRC, the optimal model can be determined by comparing the area under PRC (the larger area, the better prediction result). The calculation formula of area is shown in Equation (4), where  $(x, y)$  represents the coordinates of points on the curve, and  $m$  represents the number of points. The PRC area of BP model is 0.91, while the SOM-CNN model is 0.94 by calculating. Therefore, the SOM-CNN model is better than BP model in photometric redshift estimation.

#### 4.3.2 Results analysis

To more accurately analyze the estimation results based on SOM-CNN, we introduced five indicators as follows.

(1) Mean square error (MSE). It can be used to calculate the average error, as well as the degree of change in the data.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^n (Z_{\text{phot}}(i) - Z_{\text{spec}}(i))^2. \quad (5)$$

(2) Root mean square error (RMSE), which can reflect the degree of dispersion between data.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^n (Z_{\text{phot}}(i) - Z_{\text{spec}}(i))^2}. \quad (6)$$

(3) Bias can reflect the central trend of  $\Delta Z$ .

$$\text{Bias} = \frac{1}{N} (Z_{\text{phot}}(i) - Z_{\text{spec}}(i)). \quad (7)$$

(4)  $\delta$  represents the standard deviation of  $\Delta Z$ , which can describe the distance from the average, as well as reflect the degree of dispersion of the data.

$$\delta = \frac{Z_{\text{phot}}(i) - Z_{\text{spec}}(i)}{1 + Z_{\text{spec}}(i)}. \quad (8)$$

(5) Outliers:

$$\text{Outliers} = \Delta Z > 3\delta. \quad (9)$$

According to these five error indicators, the experimental results based on SOM-CNN and BP are calculated. The results are shown in Table 1, where ETG represents Early-type Galaxy, while LTG denotes Late-type Galaxy.

To describe the experimental results more intuitively, the error visualization figures about  $\Delta Z$  ( $\Delta Z = Z_{\text{phot}}(i) - Z_{\text{spec}}(i)$ ) and  $\delta$  have provided. The  $\Delta Z$  error visualization is shown in Figure 10. It can be seen from the distribution histogram of  $\Delta Z$  that the estimation errors of SOM-CNN model for both early-type galaxies and late-type galaxies are less than 0.1. Furthermore, the error distribution is similar to the Gaussian distribution, which highlights that our method gives a fair assessment of the estimation accuracy. According to statistical calculations, the early-type galaxies with an error of less than 0.1 accounted for 98.86%, while the late-type galaxies accounted for 99.03%. It is obvious that the SOM-CNN algorithm has achieved a good estimation.

Additionally, by introducing the  $\delta$  visualization figure to better evaluate the accuracy of SOM-CNN prediction model. Spectral redshift ( $Z_{\text{spec}}$ ) is used as the horizontal axis, and the  $\delta$  is adopted as the vertical axes to plot the error maps (Freeman et al. 2009). Therefore, Figure 11 shows the  $\delta$  scatter distribution of the two galaxies. In theory, the smaller error of prediction results, the closer is to

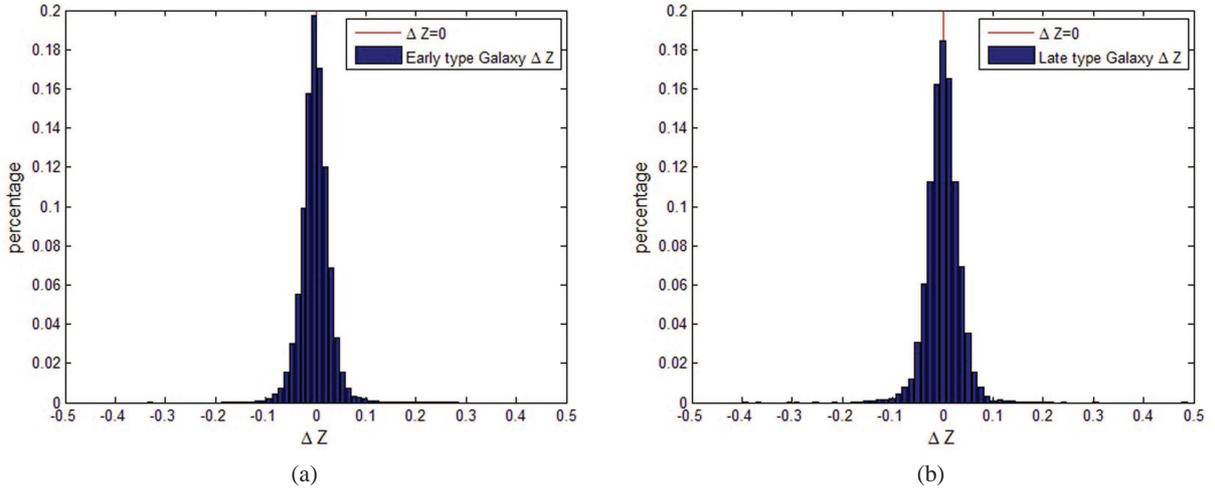
**Table 1** Comparison of Prediction Results based on SOM-CNN and BP Algorithms

Error parameter	SOM-CNN ETG	BP ETG	SOM-CNN LTG	BP LTG
MSE	0.0014	0.0018	0.0019	0.0026
RMSE	0.0343	0.0433	0.0438	0.0509
Bias	0.0057	-0.0082	-0.0011	-0.0034
$\delta$	0.0295	0.0341	0.0355	0.0396
Outliers	0.0140	0.0129	0.0132	0.0161

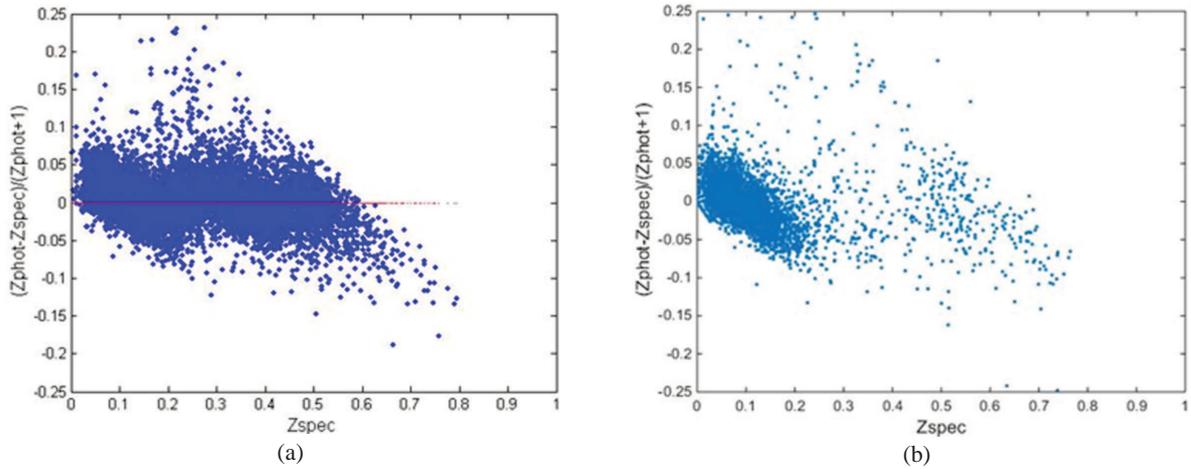
  

Statistical parameter	ETG increase percentage	LTG increase percentage
MSE	22.2%	26%
RMSE	20.7%	13.4%
$\delta$	13.4%	10.4%
Outliers	-8.5%	18%

Table 1 is divided into two parts. In the upper part, we give the standard statistical indicators (see the explanation in the text) which used to evaluate the prediction results of SOM-CNN. It is improved in each error parameter compared with the BP algorithm. To more intuitively reflect the accuracy improvement of the SOM-CNN model, we report the fraction (expressed as percentages) in the lower part of Table 1.



**Fig. 10**  $\Delta Z$  histogram distribution based on the SOM-CNN algorithm. The *red line* represents  $\Delta Z = 0$  and the  $\Delta Z$  is plotted in *black*, where panel (a) demonstrates early-type galaxies, while panel (b) indicates late-type galaxies.



**Fig. 11**  $\delta$  scatter diagrams based on SOM-CNN algorithm. The *blue dots* represent the  $\delta$  error distribution in the photometric redshift experiments, in which the early-type galaxies are shown in panel (a) and the late-type galaxies are shown in panel (b).

**Table 2** Comparison of Prediction Results based on SOM-CNN and KNN Algorithms

Error parameter	SOM-CNN ETG	KNN ETG	SOM-CNN LTG	KNN LTG
Bias	0.0057	-0.0033	-0.0011	-0.0051
$\delta$	0.0295	0.0331	0.0355	0.0549
Outliers	0.0140	0.0393	0.0132	0.0393

zero line. In the range of error less than 0.05, which indicates that the prediction results of the early-type galaxies account for 91.8% while the late-type galaxies is 93.3% according to the statistics.

To further verify the prediction effect of SOM-CNN, the SOM-CNN prediction results are compared with the KNN algorithm-based redshift prediction results used by Robert Beck (Beck et al. 2016). The comparison results are shown in Table 2.

It can be seen from Table 2 that under the condition that the prediction accuracy is equivalent, the Outliers of the SOM-CNN algorithm are optimally 1.32%, which is better than the optimal Outliers of 3.93% by KNN. In addition, using KNN algorithm for photometric redshift estimation requires millions of training samples and needs to be globally traversed to achieve high accuracy, so the algorithm is inefficient. But, the photometric redshift prediction based on SOM-CNN algorithm not only achieves good results but also has high efficiency. Therefore, the photometric redshift estimation based on SOM-CNN is successful.

## 5 CONCLUSIONS AND PROSPECTS

The accuracy of photometric redshift estimation is a problem with a long history and it exists in most photometric redshift estimation approaches. In this paper, we analyze the data before and after preprocessing from SDSS-DR13 at first, and then we provide a new estimation methodology by integration of SOM and CNN methods together. In this way, the efficiency of the algorithm has been improved due to many features can be learned independently and automatically.

The experimental results based on SOM-CNN show that the integrated approach can improve the accuracy of photometric redshift estimation (e.g., the MSE errors are 0.0014 and 0.0019, which are better than the BP algorithm 22.2% and 26%, respectively). When compared with the Outliers, our result is optimally 1.32%, while the KNN algorithm is 3.93% used by Robert Beck (Beck et al. 2016). Therefore our work successfully implements the application of SOM-CNN, and the validity of SOM-CNN is verified in redshift estimation.

In addition, the redshift extends up to about  $z = 0.8$ , with a useful coverage of  $z = 0.7$  in our work. But the error in  $z > 0.7$  is large, thus further improvement is needed. With the development of astronomical observation, there are increasing numbers of parameters. However, inappropriate parameters have no obvious redshift correlation, this maybe lead to large errors. Thus, it has become necessary to select the appropriate parameters. To improve the accuracy and the range of photometric redshift, we should consider the intersection of different astronomical observations (Wang et al. 2008, 2009), such as SDSS and TWOMASS in our following work.

**Acknowledgements** This work is supported by the Joint Research Fund in Astronomy (U1531242) under cooperative agreement between the National Natural Science Foundation of China (NSFC) and Chinese Academy of Sciences (CAS).

## References

- Abdalla, F. B., Amara, A., Capak, P., et al. 2008, MNRAS, 387, 969
- Ai, M., Zhu, M., & Fu, J. 2017, RAA (Research in Astronomy and Astrophysics), 17, 101
- Almosallam, I. A., Jarvis, M. J., & Roberts, S. J. 2016, MNRAS, 462, 726
- Babu, G. S., Zhao, P., & Li, X. L. 2016, in International Conference on Database Systems for Advanced Applications, Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life, 214
- Ball, N. M., Brunner, R. J., Myers, A. D., et al. 2008, ApJ, 683, 12
- Beck, R., Dobos, L., Budavári, T., Szalay, A. S., & Csabai, I. 2016, MNRAS, 460, 1371
- Bolzonella, M., Miralles, J. M., & Pelló, R. 2000, A&A, 363, 476
- Bruzual, A. G., & Charlot, S. 1993, ApJ, 405, 538
- Cavuoti, S., Amaro, V., Brescia, M., et al. 2017, MNRAS, 465, 1959
- Cho, Y. S., Moon, S. C., & Ryu, K. H. 2015, in Future Information Technology - II, eds. J. J. J. H. Park, Y. Pan, C. Kim, & Y. Yang (Dordrecht: Springer Netherlands), 273
- Cole, S., Hatton, S., Weinberg, D. H., & Frenk, C. S. 1998, MNRAS, 300, 945
- Coleman, G. D., Wu, C. C., & Weedman, D. W. 1980, ApJS, 43, 393
- Collister, A. A., & Lahav, O. 2004, PASP, 116, 345
- Freeman, P. E., Newman, J. A., Lee, A. B., Richards, J. W., & Schafer, C. M. 2009, MNRAS, 398, 2012
- Gao, Y.-L., Lian, J.-H., Kong, X., et al. 2017, RAA (Research in Astronomy and Astrophysics), 17, 041

- Gerdes, D. W., Sypniewski, A. J., McKay, T. A., et al. 2010, *ApJ*, 715, 823
- Giannantonio, T., Scranton, R., Crittenden, R. G., et al. 2008, *Phys. Rev. D*, 77, 123520
- Han, B., Ding, H.-P., Zhang, Y.-X., & Zhao, Y.-H. 2016, *RAA (Research in Astronomy and Astrophysics)*, 16, 74
- Haykin, S. 1994, *Neural Networks: A Comprehensive Foundation (3rd Edition)* (Macmillan)
- Hennawi, J. F., Strauss, M. A., Oguri, M., et al. 2006, *AJ*, 131, 1
- Li, L.-L., Zhang, Y.-X., Zhao, Y.-H., & Yang, D.-W. 2007, *ChJAA (Chin. J. Astron. Astrophys.)*, 7, 448
- Li, J., Liang, X., Shen, S. M., et al. 2017, *IEEE transactions on Multimedia*, 20, 985
- Moon, S., Park, Y., & Suh, I. H. 2016, in *International Conference on Neural Information Processing, Predicting Multiple Pregrasping Poses by Combining Deep Convolutional Neural Networks with Mixture Density Networks*
- Myers, A. D., Brunner, R. J., Richards, G. T., et al. 2006, *ApJ*, 638, 622
- Scranton, R., Ménard, B., Richards, G. T., et al. 2005, *ApJ*, 633, 589
- Wang, D., Zhang, Y.-X., Liu, C., & Zhao, Y.-H. 2008, *ChJAA (Chin. J. Astron. Astrophys.)*, 8, 119
- Wang, T., Huang, J.-S., & Gu, Q.-S. 2009, *RAA (Research in Astronomy and Astrophysics)*, 9, 390
- Wu, X.-B., Zhang, W., & Zhou, X. 2004, *ChJAA (Chin. J. Astron. Astrophys.)*, 4, 17
- Wu, X.-B., Zuo, W.-W., Yang, Q., et al. 2012, *RAA (Research in Astronomy and Astrophysics)*, 12, 1185
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, *AJ*, 120, 1579
- Zhang, L., Yuan, Q.-R., Zhou, X., et al. 2010, *RAA (Research in Astronomy and Astrophysics)*, 10, 1
- Zhang, K., Zuo, W., Chen, Y., et al. 2017, *IEEE Transactions on Image Processing*, 26, 3142
- Zhu, J. 2014, *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets* (Springer)