*International Journal of Innovative Computing, Information and Control*
Volume **19**, Number **2**, April **2023**

ICIC International ©2023 ISSN 1349-4198
pp. **579–596**

# AUTOMATIC CLASSIFICATION ALGORITHM OF ASTRONOMICAL OBJECTS BASED ON IMPROVED RESNET

Xiang Gao, Liangping Tu*, Juan Li and Xin Li

College of Science
University of Science and Technology Liaoning
No. 185, Qianshan Road, Lishan District, Anshan 114031, P. R. China
x419993789@163.com; { 1802681501; 779762386 }@qq.com
*Corresponding author: tuliangping@ustl.edu.cn

Received July 2022; revised November 2022

Abstract. *With the development of various large-scale sky survey projects at home and abroad, the amount of astronomical data has also shown a rapidly increasing trend. How to classify astronomical objects efficiently and quickly has become a difficult problem for astronomers. However, the advantages of deep learning can solve this problem. An automatic classification algorithm named ResNet50-GQS for astronomical objects with improved ResNet was proposed. We designed this algorithm based on ResNet by adding SE modules, modifying convolution kernel size, adjusting residual structure, reducing activation functions and normalizations. By comparing with some popular deep learning algorithms such as GoogLeNet, DenseNet, VIT, and Swin Transformer, we proved that the algorithm can classify galaxies, quasars and stars more accurately, and its accuracy rate reaches 92.41%, which is higher than other algorithms by 0.91% to 5.79%, indicating that it can be used in the study of astronomical objects classification in large-scale sky survey projects.*

**Keywords:** Astronomical objects, Image classification, Deep learning, ResNet

1. **Introduction.** Astronomical object is the general term of all objects in the universe. Classifying the observed astronomical objects had been a basic problem since human began observing astronomical objects. In recent years, with the advancement of science and technology, astronomical observation has developed to an unprecedented stage. Large-scale sky survey projects around the world such as Sloan Digital Sky Survey (SDSS) [1], Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) [2], Large Sky Area Multi-Object Fiber Spectroscopy Telescope (LAMOST) [3], and Large Synoptic Survey Telescope (LSST) [4] are in full swing. The development of these large-scale sky survey projects has resulted in an exponential increase in the amount of astronomical data. For such a huge amount of data, traditional manual classification methods have been unable to meet the actual needs.

As machine learning shines in various fields, many astronomical researchers have also turned their attention to machine learning, hoping to use artificial intelligence to achieve automatic classification of astronomical objects. Related researches include artificial neural network [5], support vector machine [6], decision tree [7], classifier combination strategy [8], random forest [9], stacking ensemble learning model [10], etc. Traditional machine learning is a semi-automatic learning method, because the features learned by the machine need to be determined by industry experts and then hand-coded. For example, commonly used features of astronomical objects include magnitude, size, shape, etc., and the selection of features directly affects the accuracy of astronomical object classification.

DOI: 10.24507/ijicic.19.02.579

Deep learning is a branch of machine learning, which is developed from artificial neural networks. The most significant difference between deep learning and traditional machine learning is that features are not designed manually, but automatically learned from data through neural networks. Deep learning extracts features by converting low-level features into higher-level and more abstract features to make classification or prediction easier. The specific process is as follows: When a galaxy image is input into a deep learning model, the features learned by the first layer may be the outline and orientation of the galaxy. It may be circular or elliptical. If it is an ellipse, its long axis direction may be from upper left to lower right or lower left to upper right. The features learned by the second layer may combine the outline and orientation learned by the first layer into more complex shapes. Subsequent layers continue to abstract the previously learned features to obtain classification results.

Up to now, deep learning has formed a variety of learning frameworks, such as convolutional neural networks, deep belief networks, and recursive neural networks. They have been applied in the fields of computer vision, speech recognition, natural language processing, audio recognition and bioinformatics with excellent results [11]. In recent years, convolutional neural networks have become a popular method in computer vision. Hafiyan et al. [12] combined convolutional neural network with extreme learning machine to classify DNA damage in buccal mucosa samples. Harras et al. [13] proposed a vehicle classification method using data replication and transfer learning in deep neural networks. Wang et al. [14] proposed a multi-branch adaptive squeeze and excitation residual network to solve the facial expression recognition problem. With the exponential growth of astronomical data and the rise of deep learning, traditional machine learning algorithms can no longer meet the needs of astronomical researchers. Therefore, deep learning attracts the attention of astronomers with its true sense of automatic classification. Kim and Brunner [15] used deep convolutional networks to construct a star-galaxy classification framework. They experimented on data from SDSS and CFHTLS, and demonstrated that deep convolutional neural networks can accurately classify stars and galaxies, which is more competitive than traditional machine learning methods. However, due to the insufficient number of training images with spectroscopic follow-up, the model may overfit the data. Khalifa et al. [16] proposed a deep convolutional neural network for morphological classification of galaxies. The network can classify galaxies into 3 categories according to their morphological characteristics: elliptical galaxies, spiral galaxies and irregular galaxies. After training 1356 images, the test accuracy rate reaches 97.272%. Compared with other relevant works of the same period, the proposed network structure has some advantages in the test accuracy, but it lacks the process of data processing and data augmentation, and too few training images will lead to the overfitting problem. Dai and Tong [17] proposed an improved residual network (ResNet-26) for the morphological classification of galaxies. Their dataset includes 28790 galaxy images from the Galaxy Zoo 2. They classified galaxies into 5 categories; except cigar-shaped galaxies, the other 4 categories have an accuracy rate of more than 94%. In this paper, the data are introduced and processed in detail, and the galaxies are divided in more detail. However, it is only compared with two models, which is not convincing. Xu et al. [18] completed the design and optimization of the RS-CNN model according to the characteristics of the mainstream CNN models, and used the customized RS-CNN to conduct classification experiments on spiral galaxies and stars. The inspiration of RS-CNN model comes from residual module and SE module, but the paper does not mention how to use these two modules. He et al. [19] applied a deep learning source detection network based on YOLO v4 object detection framework to detecting sources, and designed a deep learning classification network named APSCnet to conduct classification experiments on galaxies, stars and quasars. The above research has guiding significance for our work.

This paper aims to design an efficient automatic classification algorithm of astronomical objects for astronomical researchers. Firstly, we used 9 deep learning classification algorithms to classify astronomical objects. The model with the best performance was selected by comparison. We then made a series of improvements to the selected model and made comparative experiments on different versions of the improved model. Finally, an algorithm model suitable for automatic classification of astronomical objects is designed.

We studied the classification algorithm of galaxies, quasars and stars in SDSS photometric images based on convolutional neural networks. The specific workflow is as follows.

1) We first downloaded photometric images from the official website of SDSS[1], and downloaded the catalogs of galaxies, quasars and stars in the photometric image through the CasJobs server[2]. The catalogs include some information such as right ascension (ra), declination (dec), magnitude, and class.

2) Next, we synthesized the images of the i, r, and g bands into RGB images, and used Python to convert the ra and dec of galaxies, stars and quasars to pixel coordinates. Then we cropped them from the synthetic photometric images. The cropped astronomical objects were divided into training set, validation set and test set according to the ratio of $8:1:1$. Before training, we performed a series of data augmentation such as resizing, random rotation, and random flipping on the images.

3) Next, through comparative experiments, we verify that ResNet50 has the best classification performance for galaxies, quasars and stars in the existing traditional deep learning algorithms.

4) Finally, we proposed ResNet50-GQS network model which is suitable for the classifications of galaxies, quasars and stars by adding SE modules, modifying convolution kernel size, improving residual structure, reducing activation functions and normalizations in ResNet50.

The organization of this paper is as follows: Section 2 introduces the ResNet and the improved algorithm structure; Section 3 introduces data preparation, data preprocessing, data set division and data augmentation; Section 4 introduces how we selected the model with the best classification performance among the 9 deep learning algorithms, and the experimental results of the improved model; Section 5 presents the conclusion and prospect of this paper.

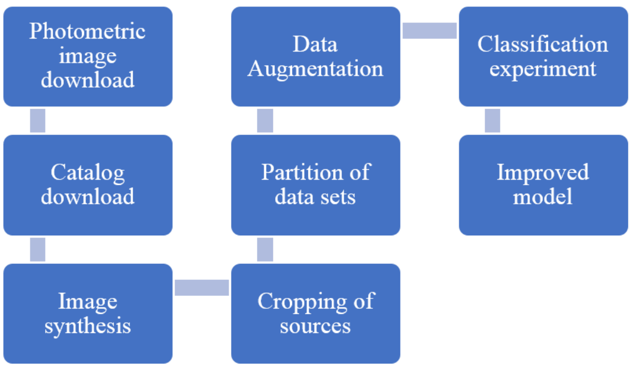



FIGURE 1. The workflow chart of this article

[1]https://dr16.sdss.org/
[2]http://skyserver.sdss.org/CasJobs/

## 2. **Algorithm.**

2.1. **Convolutional neural network.** Convolutional Neural Network (CNN) is a kind of deep learning algorithm specially designed for image classification and recognition [20]. It is a deep neural network with convolutional structure. Similar to the structure of artificial neural network, CNN has input layer, hidden layer and output layer. The difference is that CNN adds convolutional and pooling layers to the hidden layers. The function of the convolutional layer is to extract features from data, and the pooling layer is to downsample the output of the previous layer to remove redundant information, improve the robustness of feature extraction and prevent overfitting.

This paper used 9 deep learning algorithms to classify galaxies, quasars and stars, as follows: VGGNet [21], GoogLeNet [22], ResNet [23], MobileNetV2 [24], DenseNet [25], EfficientNet [26], VIT [27], Swin Transformer [28] and ConvNeXt [29].

2.2. **Deep residual network.** In order to solve the degradation problem (with the increase of network depth, the training accuracy reaches saturation and then decreases rapidly) caused by the network being too deep, He et al. [23] proposed a deep residual network. The core of ResNet is the residual structure. Figure 2 and Figure 3 are the two residual structures in ResNet. Assuming that the input of the network is $x$, after a series of layer structures the output is $H(x)$. The information learned in this process is $F(x) = H(x) - x$. In this case, the original output becomes $H(x) = F(x) + x$. As shown in Figure 2 and Figure 3, the input $x_l$ first passes through two paths, then merges, and finally outputs. One of the paths has convolutional layers, and the other path performs identity mapping, that is, the input $x_l$ is directly transmitted to the final output without any operation, and then added to the output of the other path. This structure is called shortcut connection. In ResNet, the residual structure that skips 2 convolutional layers is called Basicblock, and the residual structure that skips 3 convolutional layers is called Bottleneck. The residual structure can be expressed as

$$x_{l+1} = f\left(W_s x_l + F(x_l)\right) \tag{1}$$

where $x_l$ and $x_{l+1}$ are the input and output of the $l$th residual structure; $F$ is residual function; $W_s$ is used to match the dimensions of $x_l$ and $F(x_l)$; $f$ is activation function, which in ResNet is Linear Rectification Function (ReLU) [30]. According to the number of network layers, ResNet can be divided into ResNet18, ResNet34, ResNet50, ResNet101 and ResNet152.

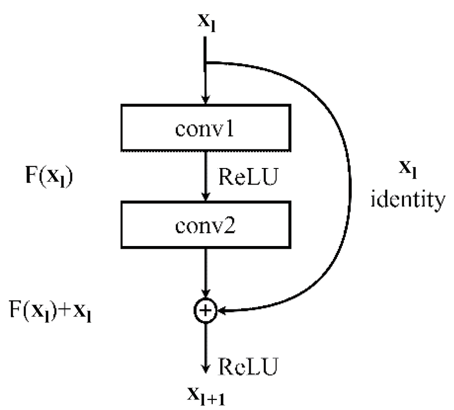



FIGURE 2. Basicblock structure

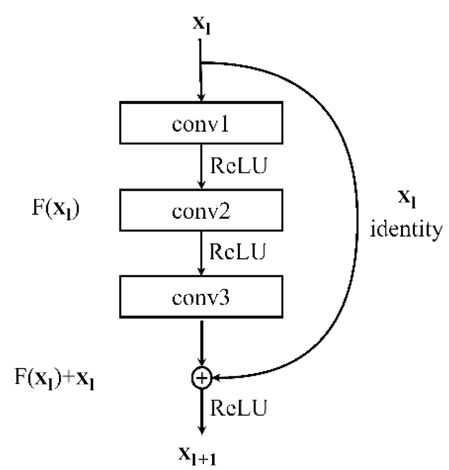



FIGURE 3. Bottleneck structure

2.3. **Introduction of the improved algorithm.** On the basis of ResNet, we propose Res-Net50-GQS network model which is suitable for classifications of galaxies, quasars and stars by adding attention mechanism, modifying the size of convolution kernel, improving residual unit, reducing activation function and normalization.

TABLE 1. Description of some notations

| Notations | Meanings |
|---|---|
| conv | Layer of convolution |
| $X$ | Input value |
| $C$ | Number of channels |
| $W$ | The width of the image |
| $H$ | The height of the image |
| $F_{tr}$ | Operation of convolution |
| $F_{sq}$ | A squeeze operation (global average pooling) |
| $F_{ex}(\cdot, W)$ | An excitation operation (learning $W$ to explicitly model channel association) |
| $F_{scale}$ | Reweighting the feature maps |
| $GAP$ | Global average pooling |
| $k$ | The size of 1D convolution |
| $s$ | Stride |
| $p$ | Padding |
| $d$ | Dimension |
| $FC$ | Full connection layer |

We first used 9 deep learning algorithms to classify galaxies, quasars and stars, and selected the model with the best classification performance – ResNet50 (see Section 4.3 for the specific experimental results). Then, to explore whether other models in the ResNet family have similar performance, we compared ResNet50 with ResNet18 and ResNet34, and found that ResNet18 performed similarly to ResNet50. On this basis, we add attention mechanism to ResNet18 and ResNet50, including SE module [31], ECA module [32] and CBAM module [33]. SE module and ECA module are the channel attention mechanism, CBAM is the attention mechanism combining channel and space. The channel attention mechanism focuses on which channel features are more meaningful, and the spatial attention mechanism focuses on which parts of the space are more meaningful. We added SE module and ECA module after the last convolution layer in the residual structure. CBAM modules were added after the initial downsampling module of the network and before the last fully connected layer of the network. The comparison results show that ResNet50-SE has the best performance. The subsequent improvements are made on this basis. The above experimental results can be found in Section 4.3.

The initial downsampling module of ResNet consists of a convolution layer with kernel size of $7 \times 7$ and a max pooling layer with kernel size of $3 \times 3$, which can downsample the

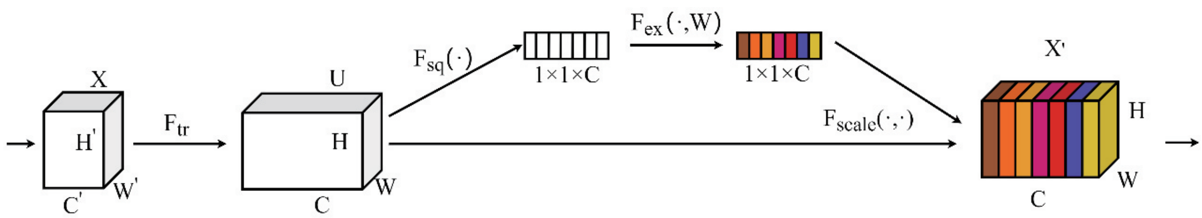



FIGURE 4. SE module

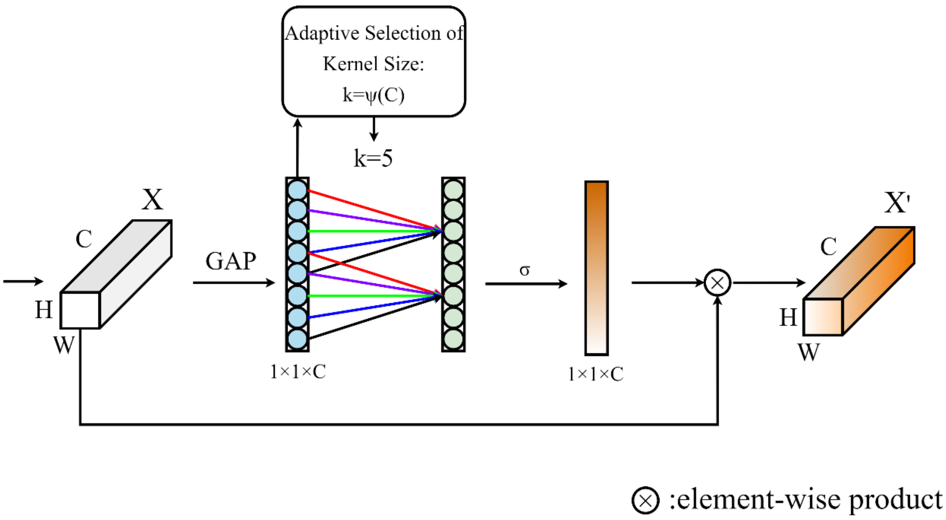



FIGURE 5. ECA module

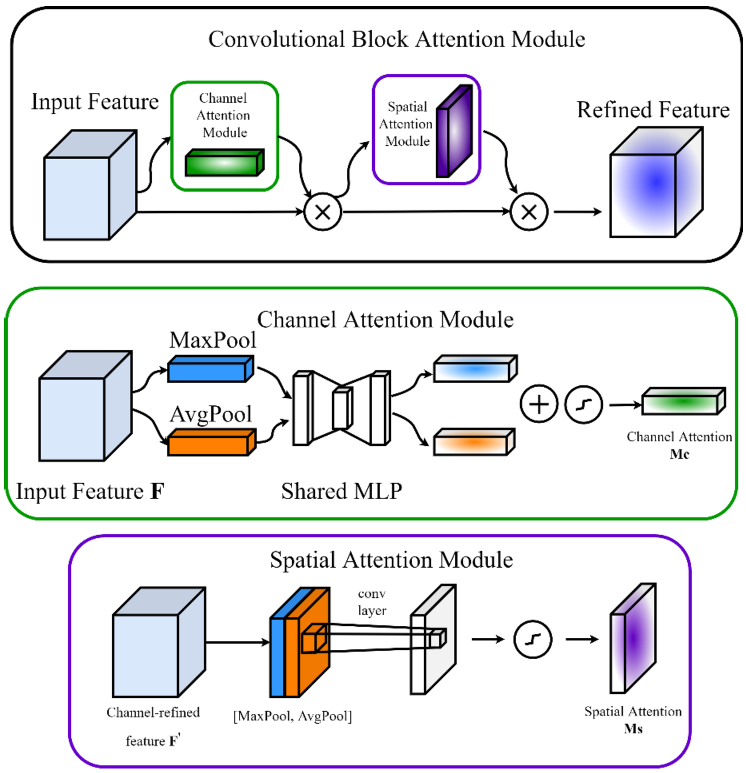



FIGURE 6. CBAM module

original image by 4 times. In Swin Transformer [28], downsampling is generally performed through a convolution layer with a large convolution kernel and no overlap between adjacent sliding windows (the size of the convolution kernel is equal to the stride). ConvNeXt [29] adopted this method with a small performance improvement. Therefore, we also applied it to our improved network.

In order to keep the downsampling multiple unchanged, we use a convolution layer with stride 4 and kernel size of $4 \times 4$ to achieve 4-fold downsampling. In ResNet, the initial numbers of channels of the residual structure from stage2 to stage5 are 64, 128, 256 and 512, respectively. The author of ConvNeXt referred to the number of channels in the

Swin Transformer and changed it to 96, 192, 384 and 768, again improving network performance. So, we also tried changing the initial numbers of channels of residual structure to 96, 192, 384 and 768. The improved ResNet50-GQS network structure is shown in Table 2 and Figure 7.

TABLE 2. Improved ResNet50-GQS network architecture

| Stage | Input size | Operator | Output size |
|---|---|---|---|
| Conv1 | $50 \times 50 \times 3$ | $4 \times 4$, 96, $s = 4$, $p = 2$ | $13 \times 13 \times 96$ |
| Conv2 | $13 \times 13 \times 96$ | $\begin{bmatrix} 3 \times 3, & 96 \\ 1 \times 1, & 384 \\ 1 \times 1, & 384 \end{bmatrix} \times 3$ | $13 \times 13 \times 96$ |
| Conv3 | $13 \times 13 \times 96$ | $\begin{bmatrix} 3 \times 3, & 192 \\ 1 \times 1, & 768 \\ 1 \times 1, & 768 \end{bmatrix} \times 4$ | $7 \times 7 \times 768$ |
| Conv4 | $7 \times 7 \times 768$ | $\begin{bmatrix} 3 \times 3, & 384 \\ 1 \times 1, & 1536 \\ 1 \times 1, & 1536 \end{bmatrix} \times 6$ | $4 \times 4 \times 1536$ |
| Conv5 | $4 \times 4 \times 1536$ | $\begin{bmatrix} 3 \times 3, & 768 \\ 1 \times 1, & 3072 \\ 1 \times 1, & 3072 \end{bmatrix} \times 3$ | $2 \times 2 \times 3072$ |
| Classifier | $2 \times 2 \times 3072$ | average pool, 3-d FC | $1 \times 1$ |

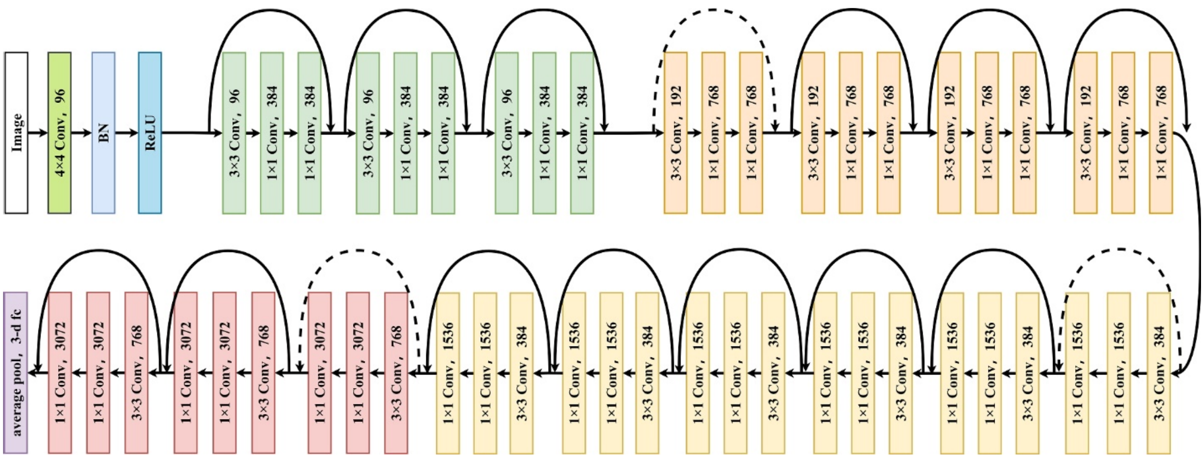



FIGURE 7. Improved ResNet50-GQS network architecture

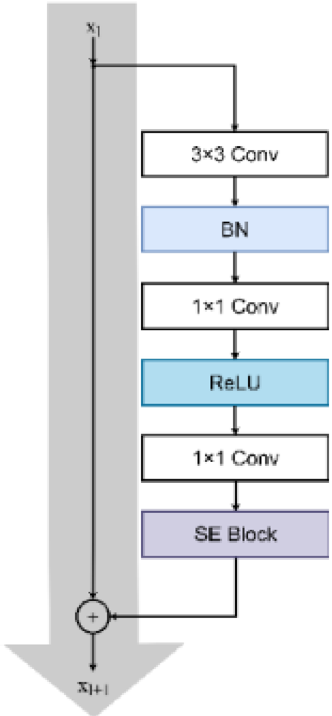



FIGURE 8. Improved residual unit

The improved residual unit is shown in Figure 8. We first moved the convolution layers with kernel size of 3×3 to the front, because a prerequisite for exploring large convolutional kernels is to move up the position of deep convolutional layers [29]. Each convolutional layer in the original ResNet is followed by a ReLU and Batch Normalization (BN) [34], while in Swin Transformer [28], both activation functions and normalization are rarely used. Therefore, we also reduce the use of ReLU and BN. We connected a BN after the first convolution layer ($3 \times 3$ conv), connected a ReLU after the second convolution layer ($1 \times 1$ conv), and added SE module after the last convolution layer ($1 \times 1$ conv). Since ReLU operation on low-dimensional information is easy to cause information loss, while information loss is less when ReLU operation is performed on high-dimensional information [24], we increased the number of channels in convolution layer connected to ReLU to be equal to the number of channels in the third convolutional layer, which is 4 times the number of the first convolution layer.

## 3. **Data.**

3.1. **Data preparation.** Sloan Digital Sky Survey (SDSS) is one of the most successful sky surveys in the history of astronomy, creating the most detailed three-dimensional map of the universe, including deep color images of one-third of the sky, and more than 3 million spectrums of astronomical objects. There are 5 stages of SDSS so far (SDSS I-V). SDSS-V began observing in October 2020, and its first data release is expected in late 2022. In this paper, we used convolutional neural network to study the classification algorithms of galaxies, stars and quasars on photometric image data obtained from SDSS Data Release 16 (DR16) [35]. SDSS DR16 is the fourth data release of the Sloan Digital Sky Survey Stage 4 (SDSS-IV). It contains SDSS data as of August 2018, covering the sky area of 14555 deg2, with photometry for five bands of u, g, r, i, and z, resulting in a catalog of more than 300 million stars and galaxies photometric data. These data can be obtained through CasJob server [36].

SDSS has a huge amount of data, including photometric images and spectral images. The photometric image catalog (PhotoObj) in SDSS DR16 classifies a subset of sources. In PhotoObj, the parameter 'type' indicates the selected source type, which classifies sources into unknown sources (type = 0), cosmic rays (type = 1), galaxies (type = 3), stars (type = 6), etc. However, quasars are not classified. We know that stars are point sources, galaxies are surface sources, and quasars are also point sources; therefore, stars and quasars cannot be accurately separated according to their morphology. The precise source classification method is presented in SDSS DR16's spectral image catalog (SpecPhoto). In SpecPhoto, the parameter 'class' classifies sources into galaxies, stars and quasars, which have been calibrated by the spectrum. Since the spectrum is obtained by photographing a single light source with a spectrometer, it contains more information about the source than the photometric image, so deep learning algorithms can be used to achieve more accurate source classification [19].

```
SELECT
    p.objID, p.run, p.camcol, p.field, p.type,
    p.ra, p.dec, s.z, s.class,
    p.u, p.g, p.r, p.i, p.z, p.psfMag_r, p.cModelMag_r
from photoObj p
JOIN SpecPhoto s ON s.ObjID = p.objID
WHERE s.class = 'GALAXY'
AND s.zWarning = 0
AND p.run = 3430
```

Figure 9. The SQL query statement that outputs the catalog of all galaxies in run = 3430

SDSS stores all photometric images in Flexible Image Transport System (FITS) format in the Science Archive Server (SAS)[3]. We downloaded FITS files from SAS's PhotoObj with parameters 'run' of 16 regions, including 3430, 3434, 3437, 3438, etc. There are a total of 21276 photometric images, and the resolution of each photometric image is $2048 \times 1489$. We wrote Structured Query Language (SQL) in the CasJobs server to cross the PhotoObj and SpecPhoto of 16 regions, and then set zWarning = 0 to exclude the source of redshift anomalies. Finally, the catalogs were output in CSV format. We obtained a total of 55486 sources, including 30934 galaxies, 10038 quasars and 14514 stars. The range of r-band magnitudes is $13.1 <$ modelMag_r $< 23.2$.

3.2. **Image synthesis.** The photometric images we downloaded from SAS were saved in FITS format files according to the 5 bands of u, g, r, i, and z. Images in each band may contain different information, so in order to use these images, it is often necessary to combine images of several of bands. If we directly synthesize images of different bands of the same photometric image, the synthesized image will deviate from the real image to a certain extent. Because CCD cameras with different filters are not exposed at the same time when shooting, there is a shooting time difference between the 5 bands of the same photometric image. For example, a source might have pixel coordinates $(256, 256)$ on an i-band image, but it might have pixel coordinates $(250, 256)$ on an r-band image. We applied the program RGB synthesis written by He et al. [19] for image synthesis. Due to the higher peaks of the response curves of the i, r, and g bands, astronomical researchers usually use the images of these bands to synthesize the corresponding RGB pseudo-color images. SDSS also uses i, r, and g bands to synthesize pseudo-color images, and confirmed

[3]https://data.sdss.org/sas/

FIGURE 10. Photometric image synthesized by i, r and g bands

that the pseudo-color images are very close to real colors. We also followed the method to synthesize the images of i, r, and g bands. The synthesized photometric image is shown in Figure 10.

3.3. **Cropping of sources.** The classification network requires the input image size to be a fixed value, but the source image size is different, so we convert all the data set images into fixed size images. We can query the information of the sources we need through the CasJobs server and output it in catalogs. The information includes ra, dec, magnitude and class of the sources. Python's astropy.wcs library has a function called all_world2pix to convert ra and dec to pixel coordinates. Therefore, we converted ra and dec of all sources in catalogs to their pixel coordinates $(x, y)$ in the photometric images through Python. In this way, we were able to find the center positions of the sources in photometric images. Then, we started cropping the sources. During the cropping process, we found that we could not determine the size of the crop because some sources may be located near the edge of the photometric images. For example, if we crop an image with a size of $30 \times 30$ (at this point, the vertical distance between the center of the source and the crop border is 15), and some sources may have a center position coordinate $x$ (or $y$) value less than 15. To avoid this situation, we introduced a parameter a0. a0 represents the vertical distance between the center position of the source and the cropped bounding box. When the value of $x$ (or $y$) is less than a0, a0 will be replaced with the value of $x$ (or $y$) smaller than a0. We need to both crop the source completely and crop the noise around the source as little as possible. After a series of cropping experiments, we determined that the cutting size of $50 \times 50$ is the most appropriate. It not only ensures that each source is cropped down completely, but also minimizes the noise cut down together. So, we initialized a0 to 25. As a result, most of our cropped images are $50 \times 50$ in size. The cropped images are shown in Figure 11.

FIGURE 11. Cropped images (galaxy, quasar, star)

3.4. **Partition of dataset.** We got a total of 55486 sources, but only 10038 quasars, far fewer than galaxies and stars. To balance each sample, we try to keep the number of galaxies and stars consistent with the number of quasars. Therefore, we only used the galaxies in the first 8 'runs', and the stars in the first 11 'runs'. The numbers of galaxies and stars are 11512 and 10206. Finally, the number of sources we got was 31756, and Figure 12 shows the histogram of the number of galaxies, quasars and stars. In order to maximize the training set, we set the ratio of training set to 80% while ensuring the same ratio of validation set and test set. Finally, we divided the dataset into training set, validation set, and test set in a ratio of 8 : 1 : 1. Figure 13 and Table 3 show the number distribution of galaxies, quasars, and stars in the training set, validation set, and test set.

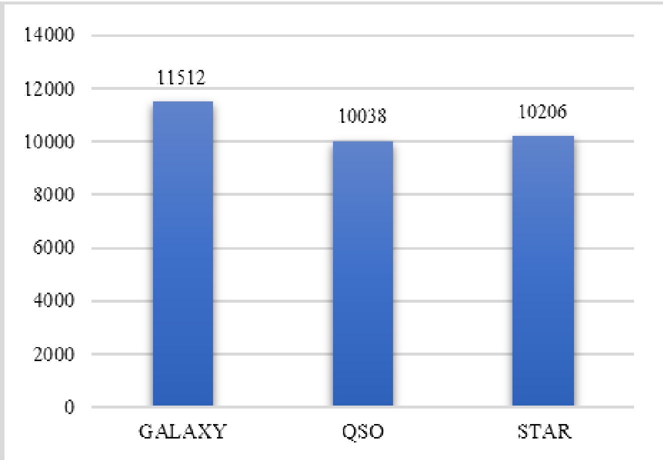



FIGURE 12. The histogram of the number of galaxies, quasars and stars

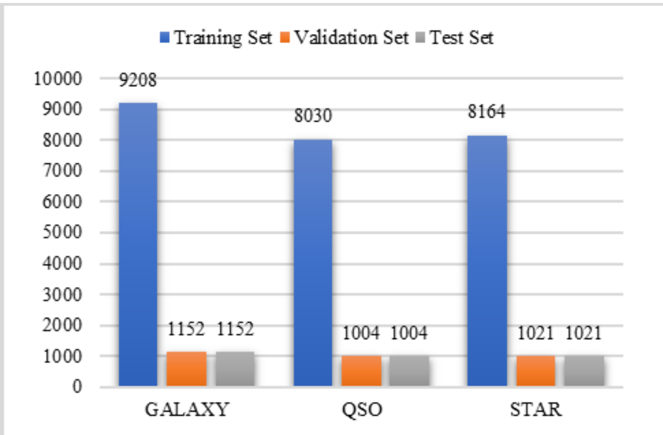



FIGURE 13. The number distribution of galaxies, quasars, and stars in the training set, validation set, and test set

TABLE 3. The number distribution of galaxies, quasars, and stars in the training set, validation set, and test set

| | Training set | Validation set | Test set | Total |
|---|---|---|---|---|
| GALAXY | 9208 | 1152 | 1152 | 11512 |
| QSO | 8030 | 1004 | 1004 | 10038 |
| STAR | 8164 | 1021 | 1021 | 10206 |

3.5. **Data augmentation.** In order to improve the generalization ability and the robustness of the model, we preprocessed the training data before inputting the training data into the model in the following ways.

1) Resize: reset the image resolution. We scaled the image to $50 \times 50$. The reason for this is that when the sources were cropped, some of the sources were cropped to a size other than $50 \times 50$, so here we scaled them all to $50 \times 50$.

2) Random rotation: rotating the image does not change the type of the sources. We took advantage of this to randomly rotate the image. The rotation angle is set to 90°, that is, randomly rotating between $(-90°, 90°)$.

3) Random flip: ① Random horizontal flip: flip left and right along the center axis of the image; ② Random vertical flip: flip the image up and down along the horizontal axis of the image. We took the default value of 0.5 for the flip probability.

For the validation set and test set, we only performed steps 1).

## 4. Analysis and Discussion of Experimental Results.

4.1. **Performance evaluation indexes for algorithms.** After training, the generalization performance of the algorithm needs to be evaluated. The following performance evaluation indexes were adopted in this paper: accuracy, precision, recall, F1-score, ROC curve and AUC, which can be obtained through confusion matrix.

TABLE 4. Confusion matrix

| Reference | Prediction | |
|---|---|---|
| | Positive | Negative |
| Positive | True Positive (TP) | False Negative (FN) |
| Negative | False Positive (FP) | True Negative (TN) |

Accuracy is the proportion of samples correctly classified by classifier to the total samples:

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{2}$$

Precision is the proportion of samples that are themselves positive among all samples classified as positive:

$$P = \frac{TP}{TP + FP} \tag{3}$$

Recall is the proportion of all positive samples that are correctly classified as positive:

$$R = \frac{TP}{TP + FN} \tag{4}$$

F1-score is the harmonic mean of precision and recall, combining the results of precision and recall:

$$F1 = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2 \times P \times R}{P + R} \tag{5}$$

ROC curve is the receiver operating characteristic curve. Its vertical axis is True Positive Rate (TPR), and the horizontal axis is False Positive Rate (FPR):

$$TPR = \frac{TP}{TP + FN} \tag{6}$$

$$FPR = \frac{FP}{TP + FP} \tag{7}$$

The closer TPR is to 1 and FPR is to 0, the better the classification effect is.

AUC is the area under the ROC curve. The larger AUC is, the better the classification effect of the classifier is.

4.2. **Experimental environment.** The experiments in this paper were all performed on a server with Intel(R) Xeon(R) E5-2640 v4 @2.40GHz CPU, NVIDIA RTX 3080Ti 12G GPU, 32GB memory and Windows 64-bit. The experimental platform was PyCharm. The models were all implemented using PyTorch 1.9 deep learning framework. In addition, astropy.wcs, OpenCV, os and Matplotlib libraries are also used in this paper.

4.3. **Training and model selection.** We first used 9 deep learning algorithms to classify galaxies, quasars, and stars, and selected the model with the highest accuracy on the test set. These 9 models are the most classic and most used models in deep learning. So our comparison is persuasive. The loss function used by all algorithms is CrossEntropyLoss, and the training epoch was set to 60. Table 5 shows other hyperparameter settings of the 9 deep learning algorithms.

Table 5. Hyperparameter settings for 9 deep learning algorithms

| Model | Optimizer | Optimizer momentum | Weight decay | Base learning rate | Learning rate schedule | Batch size |
|---|---|---|---|---|---|---|
| VGG16 [21] | Adam | None | None | 0.0001 | None | 256 |
| GoogLeNet [22] | Adam | None | None | 0.0001 | None | 256 |
| ResNet50 [23] | Adam | None | None | 0.0001 | None | 256 |
| MobileNetV2 [24] | Adam | None | None | 0.0001 | None | 256 |
| DenseNet121 [25] | SGD | 0.9 | 0.0001 | 0.0001 | cosine decay | 256 |
| EfficientNet_b0 [26] | SGD | 0.9 | 0.0001 | 0.0001 | cosine decay | 256 |
| VIT_base_patch16 [27] | SGD | 0.9 | 0.00005 | 0.001 | cosine decay | 32 |
| Swin Transformer_tiny_patch4 [28] | AdamW | None | 0.05 | 0.0001 | None | 32 |
| ConvNeXt_tiny [29] | AdamW | None | 0.0001 | 0.0001 | cosine warmup | 256 |

Table 6 shows the classification results of 9 deep learning algorithms. As can be seen from Table 6, except for the AUC value, the other 4 indexes of ResNet50 are the highest among the 9 algorithms, with accuracy of 91.50%, precision of 91.44%, recall of 91.30%, and F1-score of 0.9130. Only the AUC value is 0.9748, which is slightly lower than the highest value of 0.9797 in the table. Therefore, we will improve the ResNet50 next. The accuracy of most of the algorithms in the table is around 90%, and the lowest accuracy can even reach 86.62%, showing the good performance of these algorithms for classifying galaxies, quasars and stars.

Since ResNet50 has the highest classification accuracy for galaxies, quasars, and stars among the 9 deep learning algorithms, we continued to explore whether other models of ResNet have similar performance. Table 7 shows the comparison of the classification performance of ResNet50 with ResNet18 and ResNet34 on test set. It can be seen from the table that the classification performance of ResNet18 is similar to that of ResNet50, and their accuracy, recall and F1-score are all the same. The precision of ResNet18 is slightly higher than that of ResNet50, but the AUC value is lower than that of ResNet50. The

TABLE 6. Comparison of classification performance of 9 deep learning algorithms

| Model | Accuracy | Precision | Recall | F1-score | AUC |
|---|---|---|---|---|---|
| VGG16 [21] | 90.97% | 91.18% | 90.73% | 0.9073 | 0.9729 |
| GoogLeNet [22] | 91.25% | 91.10% | 91.01% | 0.9102 | **0.9797** |
| ResNet50 [23] | **91.50%** | **91.44%** | **91.30%** | **0.9130** | 0.9748 |
| MobileNetV2 [24] | 89.58% | 89.58% | 89.28% | 0.8929 | 0.9620 |
| DenseNet121 [25] | 88.04% | 88.05% | 87.71% | 0.8774 | 0.9565 |
| EfficientNet_b0 [26] | 88.01% | 88.22% | 87.66% | 0.8767 | 0.9565 |
| VIT_base_patch16 [27] | 90.90% | 90.71% | 90.67% | 0.9068 | 0.9706 |
| Swin Transformer_tiny_patch4 [28] | 91.06% | 91.01% | 90.81% | 0.9082 | 0.9719 |
| ConvNeXt_tiny [29] | 86.62% | 86.53% | 86.27% | 0.8624 | 0.9191 |

TABLE 7. Comparison of classification performance between ResNet50 and other models of ResNet

| Model | Accuracy | Precision | Recall | F1-score | AUC |
|---|---|---|---|---|---|
| ResNet50 [23] | **91.50%** | 91.44% | **91.30%** | **0.9130** | **0.9748** |
| ResNet18 [23] | **91.50%** | **91.51%** | **91.30%** | **0.9130** | 0.9680 |
| ResNet34 [23] | 90.87% | 90.69% | 90.52% | 0.9055 | 0.9728 |

accuracy of ResNet34 can also reach more than 90%, indicating that ResNet is suitable for classification tasks of galaxies, quasars and stars.

Through layers of screening, ResNet18 and ResNet50 were selected from a series of models, which are more suitable for galaxy, quasar and star classification tasks. Because the attention mechanism can focus on important information with a high weight, ignore irrelevant information with a low weight, and constantly adjust the weight, so that important information can be selected under different circumstances. Therefore, we tried to add attention mechanism in ResNet18 and ResNet50. Table 8 shows the classification performance comparison after adding SE module, ECA module and CBAM module into the model, respectively. Among them, the 5 indexes of ResNet50-SE are higher than other models, which increase by 0.32%, 0.27%, 0.29%, 0.0031 and 0.0014 compared with those before the SE module was added. The performance of ResNet18 also improved slightly after adding the SE module and the ECA module, but the classification accuracy decreased after adding the CBAM module. The classification accuracy of ResNet50 decreased after

TABLE 8. Comparison of classification performance between ResNet18 and ResNet50 after adding attention mechanism

| Model | Accuracy | Precision | Recall | F1-score | AUC |
|---|---|---|---|---|---|
| ResNet18 [23] | 91.50% | 91.51% | 91.30% | 0.9130 | 0.9680 |
| ResNet18-SE | 91.60% | 91.46% | 91.36% | 0.9137 | 0.9752 |
| ResNet18-ECA | 91.56% | 91.54% | 91.37% | 0.9138 | 0.9720 |
| ResNet18-CBAM | 91.38% | 91.21% | 91.15% | 0.9117 | 0.9719 |
| ResNet50 [23] | 91.50% | 91.44% | 91.30% | 0.9130 | 0.9748 |
| ResNet50-SE | **91.82%** | **91.71%** | **91.59%** | **0.9161** | **0.9762** |
| ResNet50-ECA | 91.44% | 91.37% | 91.21% | 0.9122 | 0.9719 |
| ResNet50-CBAM | 91.44% | 91.37% | 91.21% | 0.9123 | 0.9754 |

adding the ECA module and CBAM module. It shows that the addition of SE module is helpful to the performance improvement of ResNet50.

4.4. **ResNet50-GQS classification results.** After learning the Transformer model and ConvNeXt model experience, we designed the ResNet50-GQS network model suitable for galaxies, quasars and stars classification by adding SE module, modifying convolution kernel size, improving residual unit, reducing activation function and normalization, etc. The confusion matrix of ResNet50-GQS on test set is shown in Figure 14.

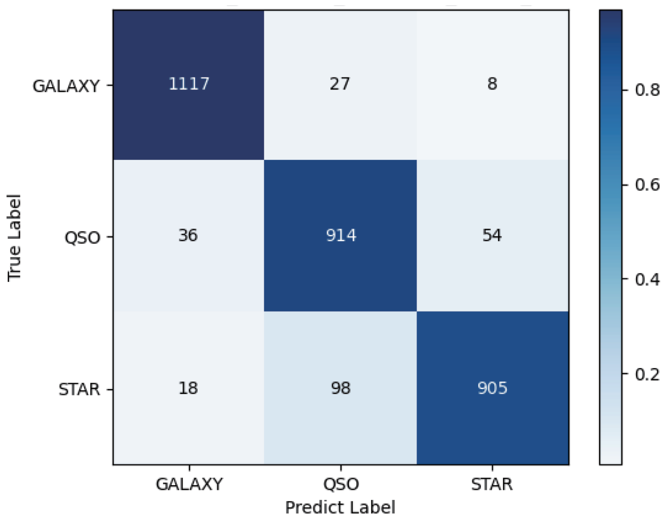



FIGURE 14. Confusion matrix of ResNet50-GQS on test set

As can be seen from the figure, 1117 galaxies, 914 quasars and 905 stars are correctly classified, from which it can be calculated that the accuracy of ResNet50-GQS on test set is 92.41%. In addition, each class has a number of misclassifications. The number of galaxies misclassified is smaller, with 54 quasars misclassified as stars and 98 stars misclassified as quasars. Because some quasars and stars are morphologically similar, they are misclassified in large numbers. The precision, recall, F1-score and their average values of galaxies, quasars and stars are shown in Table 9. Among them, the 3 indexes of galaxies are the highest values, indicating that the model has the best effect on galaxies classification.

TABLE 9. Precision, recall, and F1-score of galaxies, quasars, and stars on test sets

| Class | Precision | Recall | F1-score |
|---|---|---|---|
| GALAXY | **95.39%** | **96.96%** | **0.9617** |
| QSO | 87.97% | 91.04% | 0.8948 |
| STAR | 93.59% | 88.64% | 0.9105 |
| Average | 92.32% | 92.21% | 0.9223 |

Figure 15 shows the ROC curves and AUCs of ResNet50-GQS on test set. The curve of each color represents a class of astronomical objects. As can be seen from the figure, the best predictions are galaxies with the largest AUC value of 0.9922, followed by stars and quasars. The average AUC value of the model is 0.9704, indicating that the model has good prediction performance.

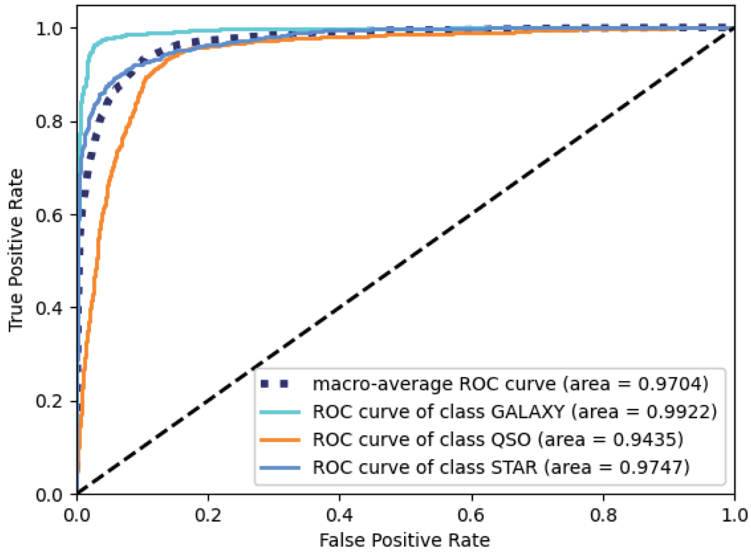



FIGURE 15. ROC curves and AUCs of galaxies, quasars and stars on test set

5. **Conclusion.** With its true automatic classification, deep learning algorithms solve a series of problems such as low classification accuracy and slow speed for astronomers. We first compared the classification performance of 9 deep learning algorithms for galaxies, quasars and stars, and selected the algorithm with the best classification performance – ResNet50. Then, through a series of operations such as adding SE module, modifying the size of the convolution kernel, improving residual unit, and reducing ReLU and BN, we designed the ResNet50-GQS network model which is more suitable for galaxies, quasars and stars classification. The performance evaluation indexes of ResNet50-GQS, such as accuracy, precision, recall and F1-score are better than all other models in this paper. However, for the classification between quasars and stars, ResNet50-GQS still has shortcomings: the number of misclassifications is large. In the future, we will try to start with the data itself, process the data in more detail, and focus on improving the classification accuracy between quasars and stars.

**Acknowledgment.** This work is supported by the Joint Fund of Astronomy of the National Natural Science Foundation of China (Grant. U1731128).

## REFERENCES

[1] R. Lupton, J. E. Gunn, Ž. Ivezić et al., The SDSS imaging pipelines, *Astronomical Data Analysis Software and Systems*, 2001.
[2] C. Heymans, L. V. Waerbeke, L. Miller et al., CFHTLenS: The Canada-France-Hawaii telescope lensing survey, *Monthly Notices of the Royal Astronomical Society*, vol.427, no.1, 2012.
[3] X. Cui, Y. Zhao, Y. Chu et al., The large sky area multi-object fiber spectroscopic telescope (LAMOST), *Research in Astronomy and Astrophysics*, vol.12, no.9, 1197, 2012.
[4] Z. Ivezic, J. A. Tyson, E. Acosta et al., LSST: From science drivers to reference design and anticipated data products, *American Astronomical Society*, 2008.
[5] S. C. Odewahn, E. B. Stockwell, R. L. Pennington, R. M. Humphreys and W. A. Zumach, Automated star/galaxy discrimination with neural networks, in *Digitised Optical Sky Surveys. Astrophysics and Space Science Library*, H. T. MacGillivray and E. B. Thomson (eds.), Dordrecht, Springer, 1992.
[6] R. Fadely, D. W. Hogg and B. Willman, Star-galaxy classification in multi-band optical imaging, *Astrophysical Journal*, vol.760, no.1, DOI: 10.1088/0004-637X/760/1/15, 2012.
[7] I. Sevilla-Noarbe and P. Etayo-Sotos, Effect of training characteristics on object classification: An application using boosted decision trees, *Astronomy & Computing*, no.11, pp.64-72, 2015.

[8] E. J. Kim, R. J. Brunner and C. K. Matias, A hybrid ensemble learning approach to star-galaxy classification, *Monthly Notices of the Royal Astronomical Society*, vol.453, no.1, DOI: 10.1093/mnras/stv1608, 2015.

[9] Y. Bai, J. F. Liu, S. Wang et al., Machine learning applied to star-galaxy-QSO classification and stellar effective temperature regression, *Astronomical Journal*, vol.157, no.1, DOI: 10.3847/1538-3881/aaf009, 2018.

[10] L. I. Chao, W. H. Zhang, L. I. Ran et al., Research on star/galaxy classification based on stacking ensemble learning, *Chinese Astronomy and Astrophysics*, vol.44, no.3, pp.345-355, 2020.

[11] Y. LeCun, Y. Bengio and G. Hinton, Deep learning, *Nature*, vol.521, pp.436-444, 2015.

[12] Y. T. Hafiyan, Afiahayati, R. D. Yanuaryska, E. Anarossi, V. M. Sutanto, J. Triyanto and Y. Sakakibara, A hybrid convolutional neural network-extreme learning machine with augmented dataset for DNA damage classification using comet assay from Buccal Mucosa sample, *International Journal of Innovative Computing, Information and Control*, vol.17, no.4, pp.1191-1201, 2021.

[13] A. Harras, A. Tsuji, S. Karungaru and K. Terada, Enhanced vehicle classification using transfer learning and a novel duplication-based data augmentation technique, *International Journal of Innovative Computing, Information and Control*, vol.17, no.6, pp.2201-2216, 2021.

[14] X. Wang, X. Hao and K. Wang, Facial expression recognition based on multi-branch adaptive squeeze and excitation residual network, *International Journal of Innovative Computing, Information and Control*, vol.17, no.3, pp.735-751, 2021.

[15] E. J. Kim and R. J. Brunner, Star-galaxy classification using deep convolutional neural networks, *Monthly Notices of the Royal Astronomical Society*, 2016.

[16] N. Khalifa, M. Taha, A. E. Hassanien et al., Deep galaxy: Classification of galaxies based on deep convolutional neural networks, *arXiv Prints*, arXiv: 1709.02245, 2017.

[17] J. Dai and J. Tong, Galaxy morphology classification using deep residual networks, *Progress in Astronomy*, vol.36, pp.384-397, 2018.

[18] Z. Xu, T. Hu and Z. Peng, Research on classification of spiral galaxies and stars based on convolutional neural network, *Journal of Physics: Conference Series*, vol.1684, no.1, DOI: 10.1088/1742-6596/1684/1/012032, 2020.

[19] Z. He, B. Qiu, A. L. Luo et al., Deep learning applications based on SDSS photometric data: Detection and classification of sources, *Monthly Notices of the Royal Astronomical Society*, 2021.

[20] J. Gu, Z. Wang, J. Kuen et al., Recent advances in convolutional neural networks, *Pattern Recognition*, vol.77, pp.354-377, 2018.

[21] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv Prints*, arXiv: 1409.1556, 2014.

[22] C. Szegedy, W. Liu, Y. Jia et al., Going deeper with convolutions, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-9, DOI: 10.1109/CVPR.2015.7298594, 2015.

[23] K. He, X. Zhang, S. Ren et al., Deep residual learning for image recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770-778, DOI: 10.1109/CVPR.2016.90, 2016.

[24] M. Sandler, A. Howard, M. Zhu et al., MobileNetV2: Inverted residuals and linear bottlenecks, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.4510-4520, DOI: 10.1109/CVPR.2018.00474, 2018.

[25] G. Huang, Z. Liu, V. Laurens et al., Densely connected convolutional networks, *IEEE Computer Society*, 2016.

[26] M. Tan and Q. V. Le, EfficientNet: Rethinking model scaling for convolutional neural networks, *arXiv Prints*, arXiv: 1905.11946, 2019.

[27] A. Dosovitskiy, L. Beyer, A. Kolesnikov et al., An image is worth 16×16 Words: Transformers for image recognition at scale, *arXiv Prints*, arXiv: 2010.11929, 2020.

[28] Z. Liu, Y. Lin, Y. Cao et al., Swin Transformer: Hierarchical vision transformer using shifted windows, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp.9992-10002, DOI: 10.1109/ICCV48922.2021.00986, 2021.

[29] Z. Liu, H. Mao, C. Y. Wu et al., A ConvNet for the 2020s, *arXiv Prints*, arXiv: 2201.03545, 2022.

[30] V. Nair and G. E. Hinton, Rectified linear units improve restricted Boltzmann machines, *Proc. of the 27th International Conference on Machine Learning (ICML-10)*, Haifa, Israel, 2010.

[31] J. Hu, L. Shen and G. Sun, Squeeze-and-excitation networks, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.7132-7141, DOI: 10.1109/CVPR.2018.00745, 2018.

[32] Q. Wang, B. Wu, P. Zhu et al., ECA-Net: Efficient channel attention for deep convolutional neural networks, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[33] S. Woo, J. Park, J. Y. Lee et al., *CBAM: Convolutional Block Attention Module*, Springer, Cham, 2018.
[34] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, *International Conference on Machine Learning (ICML)*, pp.448-456, 2015.
[35] R. Ahumada, C. A. Prieto, A. Almeida et al., The 16th data release of the sloan digital sky surveys: 1st release from the APOGEE-2 southern survey and full release of eBOSS spectra, *The Astrophysical Journal*, vol.249, DOI: 10.3847/1538-4365/ab929e, 2020.
[36] N. Li and A. R. Thakar, CasJobs and MyDB: A batch query workbench, *Computing in Science & Engineering*, vol.10, no.1, pp.18-29, 2008.

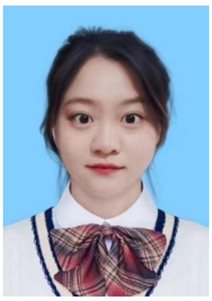
## Author Biography

**Xiang Gao** received the B.S. degree in applied statistics from Xuzhou University of Technology, China in 2020. He is a postgraduate student at University of Science and Technology Liaoning, China. His research interest is machine learning and astronomical information processing.

**Liangping Tu** obtained his Ph.D. degree in pattern recognition and intelligent systems from Institute of Automation, Chinese Academy of Sciences, Beijing, China in 2010; he worked at National Astronomical Observatory of Chinese Academy of Sciences as the postdoctoral from 2010 to 2014.

Prof. Tu is currently the dean of the College of Science, University of Science and Technology Liaoning, China. His main research interests include theory and application of pattern recognition, theory and application of machine learning and astronomical information processing. He has published over 30 papers on well-known journals, of which more than 20 were indexed by SCI. He participated in 3 national projects and a number of provincial and ministerial projects.

**Juan Li** received the B.S. degree in mathematics and applied mathematics from Shanxi Datong University, China in 2020. She is a postgraduate student at University of Science and Technology Liaoning, China. Her research interests include machine learning and theory and application of pattern recognition. She has published 1 paper on well-known journal and thepaper was indexed by SCI.



**Xin Li** received the B.S. degree in mathematics and applied mathematics from Shangluo University, China in 2019. She is a postgraduate student at University of Science and Technology Liaoning, China. Her research interest is astronomical information processing.