Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/21/20231135

Application and analysis of residual blocks in galaxy classification

Wenzheng Cheng

College Of Intelligence and Computing, Tianjin University, Tianjin, 300350, China

3020244227@tju.edu.cn

Abstract. In astronomy, the automated galaxy classification method based on deep learning has significantly reduced the cost of manual annotation. The degradation problem in convolutional neural networks during galaxy classification tasks limits the accuracy improvement of deep models. Therefore, to address the issue of the model being too deep, which leads to a decrease in accuracy, the paper constructs the galaxy classification model using residual block structures. Specifically, this paper uses an improved ResNet as the backbone, stacking different numbers of residual blocks to extract input features. Meanwhile, smaller and deeper fully connected layers, regularization and activation functions, and Dropout layers are used to improve the model performance. For the best-performing ResNet152 model, the paper analyzes the classification report and confusion matrix and visualizes saliency maps and GradCAM heatmaps. Finally, the experimental results show that the introduction of residual blocks can increase the accuracy of the model by over 30%, and models with more residual blocks perform better, although the influence of the number of residual blocks on accuracy improvement is small. The visualization results show that the model can accurately segment the feature focus areas and points of interest in the original image. The model also focuses more on the central points with high planetary density by stacking multi-level residual blocks.

Keywords: residual block, galaxy classification, resNet, visualization.

1. Introduction

With the development of modern astronomy and computer technology, the recognition and classification of galaxies has become one of the essential issues in astronomical research. Galaxies serve as the fundamental building blocks of celestial bodies in the universe. Studying galaxies yields insights into universe history, energy/matter distribution, and structure evolution. Manual classification of galaxies is a time-consuming and laborious task due to the wide-ranging types and intricate shapes they exhibit. Astronomers require extensive experience and professional knowledge to proficiently execute this procedure. Therefore, automated galaxy classification methods are becoming increasingly important.

Machine learning has become increasingly popular in astronomy in recent years. Specifically, it has proven to be particularly useful for tasks such as galaxy classification and identification. By utilizing multi-layer perceptrons and various machine learning algorithms, features can be extracted from data semi-supervised [1]. This approach has significantly reduced the need for manual labor and associated costs. However, previous machine learning methods have difficulty dealing with more expansive data and complex galaxy shapes. Deep learning models have been developed to address several limitations

DOI: 10.54254/2755-2721/21/20231135

of traditional machine learning approaches such as weak feature extraction ability, lack of invariance to image transformations, and high training complexity. However, most deep learning models still face the risks of gradient vanishing and explosion, overfitting, and model degradation [2, 3]. To address these issues, Kaiming He et al. proposed ResNet in 2016, a convolutional neural network structure that introduces residual blocks to extend network depth [4]. With regard to residual layer numbers, Kaiming proposed resnet50, resnet101, and even deeper resnet152, all of which achieved good classification and recognition accuracy on datasets such as CIFAR-10. There are also many ResNet-based improvement methods, such as using wider convolution kernels and depths, which can better solve overfitting problems [5]. Meanwhile, ResNet has shown substantial portability and robustness in many application areas and has also produced many application models in galaxy recognition [6]. However, the datasets and models they use are too simple and can only achieve high accuracy on specific small-scale recognition tasks [7,8]. Additionally, there has been limited horizontal comparison of ResNet models in galaxy classification datasets. Additional research is needed to compare the performance of ResNet models to other state-of-the-art (SOTA) models and assess how changing the number of residual blocks affects the outcomes [9].

The paper aims to analyze the effect of residual block structure and its superposition on galaxy identification. First, three neural networks with different numbers of residual blocks are constructed, including ResNet-18, ResNet-50, ResNet-101, and ResNest-152. Second, the classic state-of-the-art model VGG16 is implemented for comparison with ResNet. Further, the paper innovatively uses the Galaxy10 DECals dataset, a larger dataset of galaxies, for analysis. Additionally, the performance of different models can be compared and analyzed through the diverse metrics during the model training process, using the Matplotlib library for visualization. For the ResNet-101 model with the best performance on the test set, detailed analysis is conducted through heatmaps, saliency maps, and confusion matrices. According to experimental findings, the VGG model's deterioration issue is resolved when residual modules are added, considerably increasing the model's accuracy on the test set. Specifically, ResNet-18, 50, 101, and 152 have accuracy rates 31.60%, 32.22%, 34.08%, and 35.12% higher than VGG16, respectively. The number of residual modules is slightly proportional to the performance improvement. The ResNet-152 model achieves the highest performance among the tested models. Generally, the paper analyzes the effectiveness of various networks and residual blocks in galaxy recognition. The research results of the paper will help improve the accuracy and efficiency of galaxy recognition, provide more reliable data support for astronomical research, and promote the development of astronomy.

2. Methodology

2.1. Dataset description and preprocessing

The Galaxy10 DECals dataset is labeled using Galaxy Zoo, which consists of 17,736 colored images and other attributes of galaxies categorized into 10 classes [10]. The dataset is stored in the file Galaxy10_DECals.h5, which includes columns for images (image data for each sample), ans (labels for each sample), ra (right ascension values), dec (declination values), redshift (redshift values), and pxscale (the size of each pixel). In this paper, only the images and ans columns from the dataset are used. For each category, the author chooses one 256x256 RGB image at random, displays it, and offers labels, as seen in Figure 1.

DOI: 10.54254/2755-2721/21/20231135

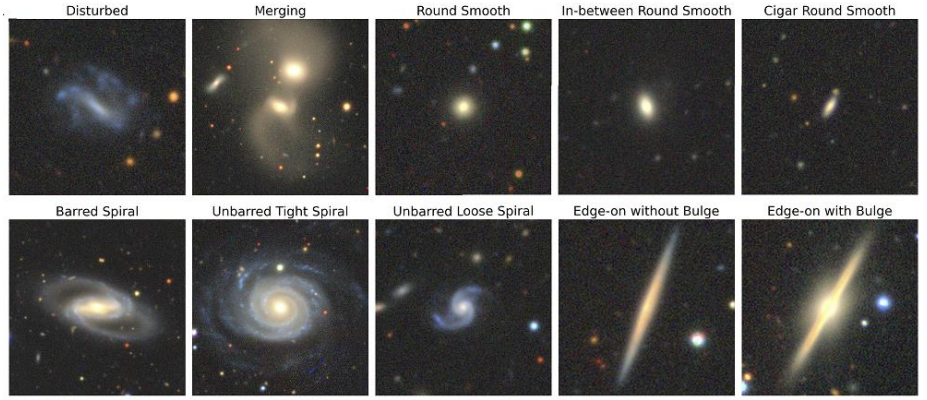

Figure 1. Example photos for each class from The Galaxy10 DECals dataset (Picture credit: Original).

The Galaxy10 DECals dataset is encoded in a 2.7GB HDF5 file, which creates a large memory burden if loaded entirely into memory as RGB value tuples, impacting training. To address this issue, the paper splits the dataset into "./galaxy10_train/" and "./galaxy10_test/" folders in a 0.8:0.2 ratio and uses image generators to load and augment data in batches, improving memory usage and model generalization. The paper employs data augmentation techniques, such as rescaling pixel values to 0-1 and applying random rotation, shearing, zooming, horizontal and vertical flipping, to expand the dataset, improve model performance, and reduce computational complexity. The paper also reserves 20% of training samples as validation data to evaluate model performance during training.

2.2. Proposed approach

To investigate the impact of residual blocks and its superposition on model performance, the paper first compares the overall performance of an improved VGG16 model with an improved series of ResNet models (ResNet-18, 50, 101, and 152). Secondly, ResNet models with different numbers of residual blocks are compared. Third, the best-performing ResNet model is further analyzed. The best-performing category is determined by analyzing the confusion matrix and classification report of the top-performing ResNet model on the test set. Then, in order to visualize the feature focus of the ResNet model with residual blocks on the original images, the saliency maps and Gradient-weighted Class Activation Mapping (Grad-CAM) heat maps of sampled images from that category are examined. To fit the dataset as well as possible and emphasize the idea of controlling variables, this paper applies the same improvement measures to all models. In order to remove the impact of non-residual block factors on the experimental outcomes, this research specifically regulates the parameter count and number of layers of the ResNet-18 and VGG16 models to a similar range. Figure 2 displays the paper's overall workflow.

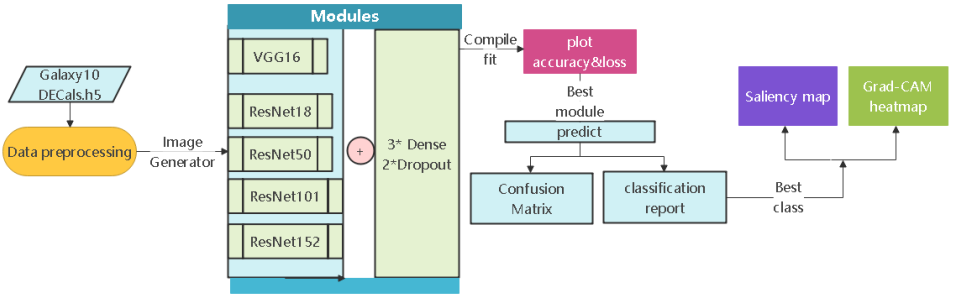

Figure 2. Overall pipeline (Picture credit: Original).

2.2.1. The improved VGG16 model. 16 weight layers comprise the original VGG16 model, including 13 convolutional layers, 3 fully connected layers, and 5 max-pooling layers. The feature extraction process is carried out by the convolutional and pooling layers, while the classification process is finished

DOI: 10.54254/2755-2721/21/20231135

by the fully connected layers. as depicted in Figure 3. The three initial fully linked layers are replaced with three new fully connected layers in the enhanced VGG16. To avoid overfitting (rate at 0.5), two dropout layers are put in between the three fully connected layers. In addition, except for the softmax activation function uses in the output layer, the other two custom layers(512 and 128 units)use the ReLU activation function and l2 regularization to prevent gradient vanishing or overfitting in the fully connected layers. The model also applies initial weights that pretrained on ImageNet for 10-class classification task to accelerate the convergence speed. The final improved VGG16 has 16 weight layers and a total parameter count of 27,627,210.

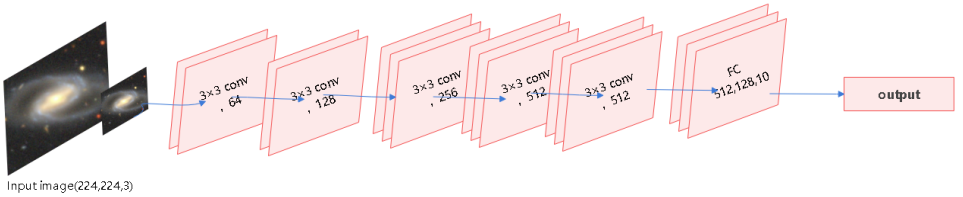

Figure 3. The improved VGG16. Conv refers to the convolution layers and the kernel size is 3×3 (Picture credit: Original).

2.2.2. The improved ResNet model. Figure 4 depicts the structure of the residual block. Identity mapping is the theory behind this. The output of the current layer is linearly added to the output of any number of preceding layers, and the result is used as the current layer's output by the activation function. Mathematically, assuming the input of the entire residual block is x, the output y is equal to:

$$y = F(x,(w_i)) + x, \tag{1}$$

where $F(x,(w_i))$ is the residual y-x between the input and output, and is also the target that the model needs to learn. A one-layer residual block specifically represents a double weight with a ReLU activation function.

$$F = w_2\delta(w_1x), \tag{2}$$

where δ refers to ReLU, and w_1 and w_2 refer to the two layers of weights in the shortcut-connection process.

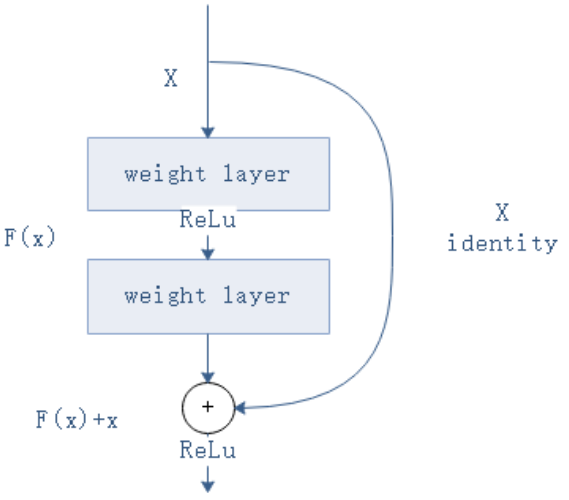

Figure 4. The structure of residual connection (Picture credit: Original).

DOI: 10.54254/2755-2721/21/20231135

In other words, as long as the residual F is fitted to be equal to 0 (which is much simpler than fitting $F(x) = \mathrm{x}$), the model is transformed into a shallow network, which to some extent ensure that the results of deep models are not worse than those of shallow models. Each residual block in the ResNet network has two or more convolutional layers, and when the output is combined with the input, a residual function is created. The ResNet network is divided into different modules based on the number of channels in the residual blocks. Each module's initial residual block downsamples the signal by halving the height and breadth and doubling the number of channels. The ResNet network structure can be changed by changing the number and width of the residual blocks to obtain different variants. For all ResNet models, the same improvement to the fully connected layers as in the previous VGG16 model are used (refer to Table 1 for the specific model structure). The final layer of the ResNet network is a global average pooling layer, which replaces the fully connected layer and reduces the number of parameters.

Table 1. The structure and parameters of all ResNet.

Layer name	Output size	18-layer	50-layer	101-layer	152-layer
Conv1	112*112	7*7, 64, stride2			
Conv2_x	56*56	3*3 max pool, stride2			
		$\begin{bmatrix}3\times3, & 64\\ 3\times3, & 64\end{bmatrix}\times2$	$\begin{bmatrix}1\times1, & 64\\ 3\times3, & 64\\ 1\times1, & 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 64\\ 3\times3, & 64\\ 1\times1, & 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 64\\ 3\times3, & 64\\ 1\times1, & 256\end{bmatrix}\times3$
Conv3_x	28*28	$\begin{bmatrix}3\times3, & 128\\ 3\times3, & 128\end{bmatrix}\times2$	$\begin{bmatrix}1\times1, & 128\\ 3\times3, & 128\\ 1\times1, & 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, & 128\\ 3\times3, & 128\\ 1\times1, & 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, & 128\\ 3\times3, & 128\\ 1\times1, & 512\end{bmatrix}\times8$
Conv4_x	14*14	$\begin{bmatrix}3\times3, & 256\\ 3\times3, & 256\end{bmatrix}\times2$	$\begin{bmatrix}1\times1, & 256\\ 3\times3, & 256\\ 1\times1, & 1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, & 256\\ 3\times3, & 256\\ 1\times1, & 1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1, & 256\\ 3\times3, & 256\\ 1\times1, & 1024\end{bmatrix}\times36$
Conv5_x	7*7	$\begin{bmatrix}3\times3, & 512\\ 3\times3, & 512\end{bmatrix}\times2$	$\begin{bmatrix}1\times1, & 512\\ 3\times3, & 512\\ 1\times1, & 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 512\\ 3\times3, & 512\\ 1\times1, & 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 512\\ 3\times3, & 512\\ 1\times1, & 2048\end{bmatrix}\times3$
	1*10	Average pool, Flatten, FC512, FC128, FC10, softmax			

In the end, the improved ResNet-18 contains 8 residual blocks(See Figure 5 for more details), ResNet-50 contains 16 residual blocks, ResNet-101 contains 33 residual blocks, and ResNet-152 contains 50 residual blocks. Notably, as depicted in Figure 5, the enhanced ResNet-18 includes a total of 11,196,042 parameters and 17 convolutional layers in addition to 3 fully connected layers. Therefore, the parameters and the number of weight layers in ResNet-18 and VGG16 are relatively close, and other factors that could affect the comparative experiments can be ignored.

DOI: 10.54254/2755-2721/21/20231135

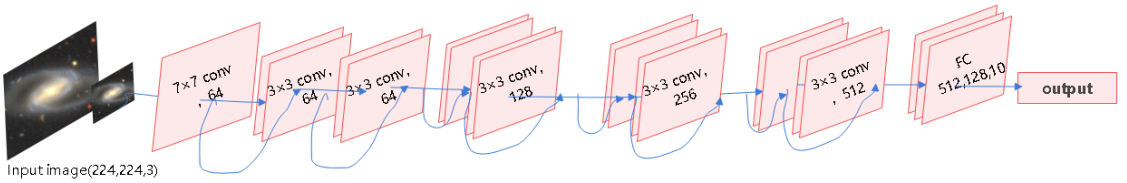

Figure 5. The improved ResNet18. The curves refer to the residual connections (Picture credit: Original).

2.3. Visualization

First, the paper uses Matplotlib functions to plot two aggregate graphs of all model fitting results. The models' performance on the validation and testing sets is depicted in the first graph as loss, and in the second graph as accuracy, both with regard to epoch. These graphs are used to analyze the performance of overall models on the galaxy classification dataset and compare the performance of ResNet models with different depths and numbers of residual blocks. Second, a thorough analysis of the best-performing ResNet-101 model's performance is provided in the study. The paper predicts the testing set data using the trained ResNet-101 model in 56 batches and draws a confusion matrix to compare the results with the true labels. The confusion matrix is then visualized as a heat map. The paper calculates the classification report based on the predicted results and selects the class with the highest f1-score for further analysis. Several images from this class are randomly selected and analyzed with saliency maps, which highlight the essential points that impact the model's output results. Finally, Grad-CAM is used to visualize the attention area of the neural network on the input image, and a heat map with annotations is generated.

2.4. Implementation details

Each model is trained for 50 epochs (except VGG, which finishes at epoch 21) with an initial epoch of 3 to accelerate convergence via an A100-40G GPU Compute Engine with a batch size of 64. Callbacks with a patience of 3 are used to stop the model training and restore the best weights to prevent overfitting. During training, the Adam optimizer is utilized with the categorical cross-entropy loss function with accuracy as the evaluation metric and the Keras learning rate decay scheduler (decay=0.00001). Early stopping and Tensorboard callbacks are also employed to monitor validation loss and record logs for visualization. The models and training logs are saved to disk using the save model and pickle dump functions, preventing data loss when the kernel is closed after 20 hours of training. The overall system RAM usage is approximately 5.8 GB, while the GPU RAM usage is about 15.6 GB. The method of storing and loading has saved over 20 GB of memory.

3. Result and discussion

Figure 6 visualizes the loss and accuracy of training and validation data on the Galaxy10 DECals dataset. ResNet models have higher loss but also higher accuracy than VGG16 models, indicating that residual blocks significantly improve performance. The ResNet performance increases with more residual blocks and layers, but at the cost of additional parameters and training time. This trade-off is due to the residual connections allowing easier training of deeper networks but causing increased loss due to gradient flow. ResNet models with more residual blocks are effective for Galaxy10 DECals classification, but the number of residual blocks and layers should be considered to achieve the best balance between accuracy and loss.

DOI: 10.54254/2755-2721/21/20231135

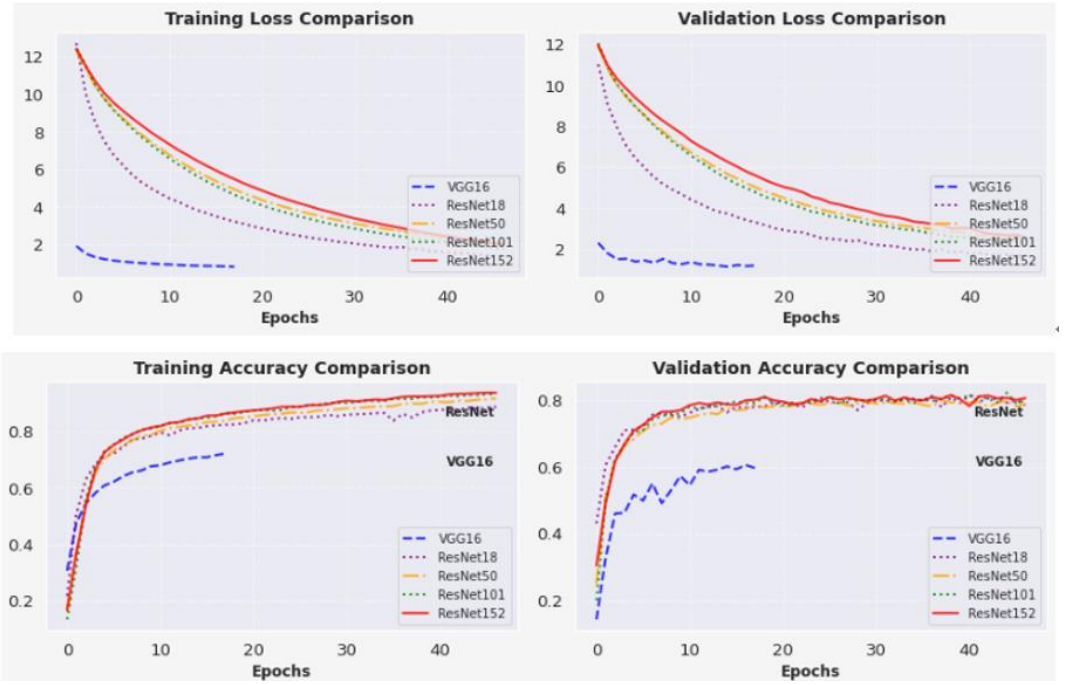

Figure 6. The performance curves of the overall models. As the epoch increases, the curve shows the changes in accuracy and loss of the five neural networks (Picture credit: Original).

Overall, the training accuracy of all models increases smoothly and reaches a relatively high level (ResNet models approximate 90%), with fluctuations at higher values indicating model convergence. The ResNet-18, 50, 101, and 152 models with residual blocks have accuracies of 78.51%, 78.88%, 79.99%, and 80.61%, respectively, on the testing set, while VGG only has an accuracy of 59.66%. Table 2 lists the number of residual blocks, training set accuracy, test set accuracy, and the percentage improvement compared to VGG for each model. This significant improvement confirms the assumption.

Table 2. Model performance comparison. The Res-block refers to the number of residual blocks, acc refers to the accuracy.

Model	Res-block	Train-acc (%)	Test-acc (%)	Improvement (%)
VGG16	0	72.21	59.66	0
ResNet18	8	88.73	78.51	31.60
ResNet50	16	91.82	78.88	32.22
ResNet101	33	93.27	79.99	34.08
ResNet152	50	93.74	80.61	35.12

ResNet152 has the highest accuracy in both the training and test sets, as shown in Table 2. This paper then aims to investigate the reason why ResNet152 extracts original image features better. Methods such as analyzing confusion matrices and classification reports, visualizing saliency maps, and Grad-CAM maps are used. Figure 7 shows the confusion matrix of ResNet-152 when predicting 56 batches of data. The diagonal elements of the matrix list the number of correct galaxy predictions. Generally, the recognition performance of the model is considerably accurate. However, for some specific types of galaxies, ResNet-152 makes more errors than others. For example, the Round Smooth Galaxy is more difficult to be identified correctly and may require more targeted training.

DOI: 10.54254/2755-2721/21/20231135

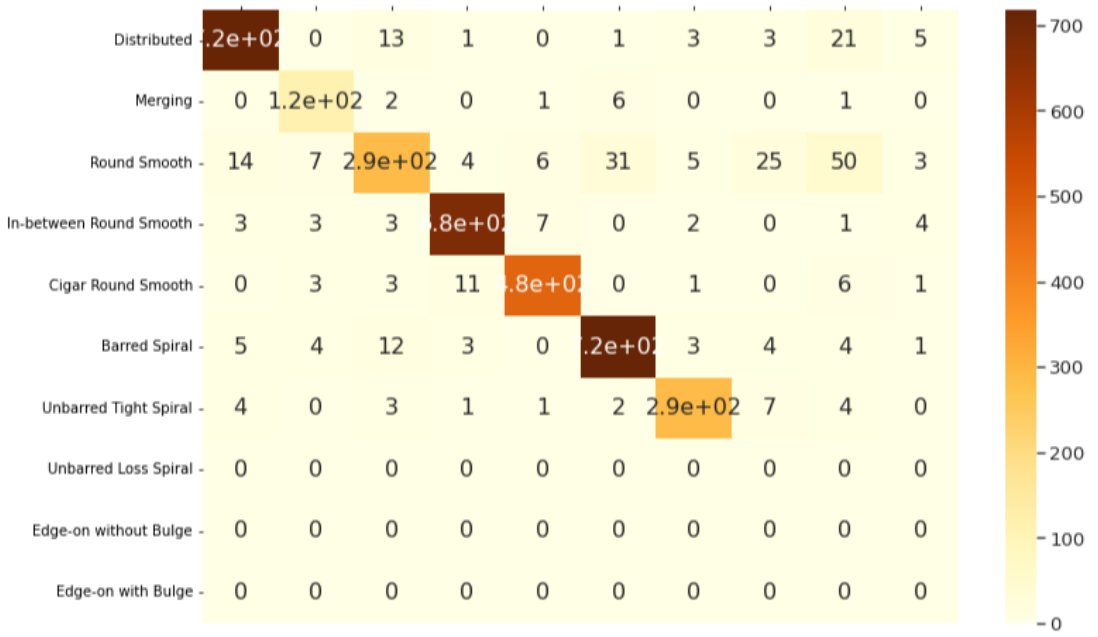

Figure 7. Confusion Matrix about classification of ResNet152. The elements on the diagonal represent the number of labels that are correctly classified (Picture credit: Original).

The predictive performance of the ResNet-152 model is further evaluated, and Table 3 shows the f1-score of the prediction. Class 7 had the highest f1-score, and the average weighted accuracy exceeded 80%, with most classes having reasonably high scores. These results demonstrate that ResNet-152 has higher efficiency and accuracy.

Table 3. Classification report with precision, recall & f1-score.

Class	Precision	Recall	F1-score	Support
0	0.80	0.81	0.80	405
1	0.80	0.85	0.82	66
2	0.49	0.61	0.54	213
3	0.87	0.91	0.89	371
4	0.92	0.86	0.89	282
5	0.94	0.83	0.88	401
6	0.93	0.88	0.91	367
7	0.88	0.93	0.91	524
8	0.67	0.63	0.65	521
9	0.71	0.72	0.71	363

To display the feature maps created by ResNet-152 on class 7 with the highest f1-score, Figure 8 shows the saliency maps superimposed over the original images of 9 example galaxies. The feature pixels can be clearly identified over the central area of each galaxy, indicating that these purple-marked pixels played an essential role in classification.

DOI: 10.54254/2755-2721/21/20231135

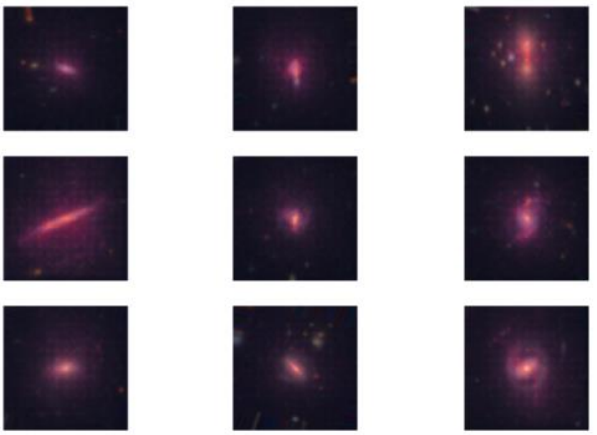

Figure 8. Saliency maps of class 7 (Picture credit: Original).

The Grad-CAM generated a class-differentiated heat map shown in Figure 9. It reveals that ResNet-152 focuses more on the information surrounding the center of the original image during feature extraction, which is more intuitive than saliency maps. Both Saliency Map and Heatmap reflect ResNet152's different attention to central and edge features of the original image. The central points of the highest planetary density in the galaxy draw more attention in Saliency Map, while the Heatmap divides the regions into patches, indicating the last layer's focus on the surrounding area of the galaxy. ResNet-152 achieves a high accuracy of 93.75% on the training data, with other ResNet models achieving over 90%, an acceptable outcome from a relatively small model with around 27 million trainable parameters. The experimental results demonstrate that residual blocks effectively address the problem of model degradation, and ResNet consistently outperforms VGG in terms of accuracy. In summary, the more residual blocks, the higher the accuracy.

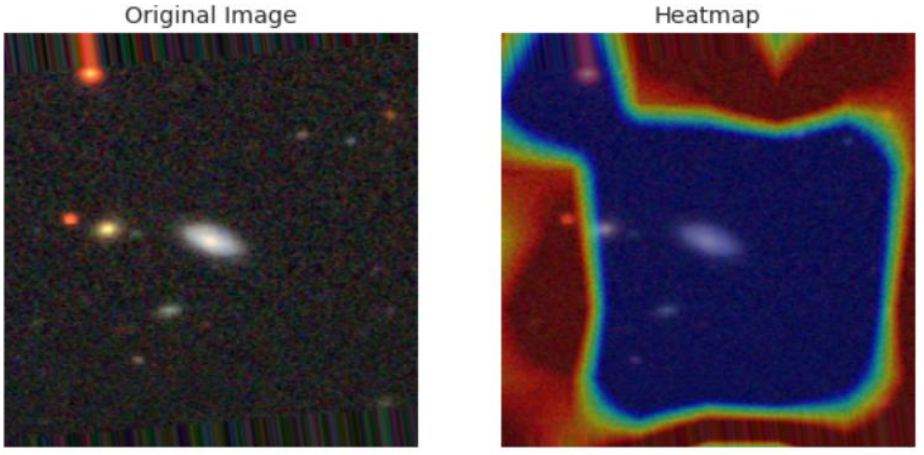

Figure 9. Heat map of class7 based on Grad-CAM (Picture credit: Original).

4. Conclusion

This study intends to examine the impact of residual block structure stacking on galaxy classification tasks in convolutional neural networks. Improved VGG16 and ResNet with different residual blocks are compared to explore the impact of residual blocks on accuracy. For the best-performing model, saliency map and GradCAM heatmap are used to analyze the model's focus on original image features and understand the essential decision-making basis in the feature extraction process. Results show that ResNet with residual blocks have a significant test accuracy improvement, meanwhile, the more residual

DOI: 10.54254/2755-2721/21/20231135

blocks stacking, the stronger the model's recognition ability. The saliency map and GradCAM heatmap of ResNet152 also show that, the model can accurately segment the feature focus points in the original image, focusing more attention on the central points with high planetary density. The study also finds that the increase in accuracy due to the number of residual blocks is small, and it is not significant enough to justify the computational and time costs. In the future, research will focus on evaluating the number of remaining blocks for optimal performance and on model lightweighting.

References

[1] Abd E et al 2014 Galaxies image classification using empirical mode decomposition and machine learning techniques International Conference on Engineering and Technology (ICET) IEEE 2014: pp 1-5
[2] Kim Edward J Robert J 2016 Star-galaxy classification using deep convolutional neural networks Monthly Notices of the Royal Astronomical Society p 2672
[3] Alhassan W Taylor A Vaccari M 2018 The FIRST Classifier compact and extended radio galaxy classification using deep Convolutional Neural Networks Monthly Notices of the Royal Astronomical Society 480(2): pp 2085-2093
[4] He K et al 2016 Deep residual learning for image recognition Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) pp 770-778
[5] Zagoruyko S Komodakis N 2016 Wide residual networks arXiv preprint arXiv 1605.07146
[6] Gupta R Srijith P K Desai S 2022 Galaxy morphology classification using neural ordinary differential equations Astronomy and Computing 38: p 100543
[7] Alawi A E B Al-Roainy A A Deep residual networks model for star-galaxy classification 2021 International Congress of Advanced Technology and Engineering (ICOTEN) IEEE pp 1-4
[8] Banerjee S Ghosh B R Gangapadhyay A et al 2021 Galaxy morphological image classification using resnet Iraqi Journal of Science pp 3690-3696
[9] Reza M 2021 Galaxy morphology classification using automated machine learning Astronomy and Computing p 100492
[10] Henry L Jo B 2022 https://github.com/henrysky/Galaxy10