A Novel Method for Galaxy Morphological Classification - GMC Postprint

Authors: Wang Linqian, Qiu Bo, Luo Ali, Kong Xiao, Lu Yakun, Guo Xiaoyu

Date: 2021-08-20T00:00:00+00:00

Abstract

In the field of astronomical research, galaxy classification has long been a topic of great interest and challenge. In recent years, scholars have applied machine learning to simple galaxy morphological classification tasks, but a series of difficulties have emerged during the classification process, including feature selection challenges, feature omission, and classifier selection difficulties. Galaxies can be visually classified into elliptical galaxies, spiral galaxies, lenticular galaxies, and irregular galaxies. This paper proposes a galaxy morphological classification method with higher classification accuracy, GMC (Galaxy Morphological Classification), targeting photometric images of galaxies from the SDSS DR16, Galaxy Zoo2, and EFIGI catalogs. We first performed image cropping and denoising, then applied data augmentation methods including rotation, translation, and scaling, and finally constructed a galaxy morphological classification network GMC-net to classify the images. Experimental classification results show that the classification precision for spiral galaxies, elliptical galaxies, lenticular galaxies, and irregular galaxies is 98.29%, 98.49%, 99.18%, and 99.91%, respectively, with recall rates of 98.44%, 99.03%, 98.89%, and 99.34%; the classification accuracy for the four morphological types of galaxies from the EFIGI catalog alone also reached 99.34%. Experimental results demonstrate that GMC performs better compared to other classification methods and can be more effectively utilized for galaxy morphological classification.

Full Text

A New Method for Galaxy Morphological Classification: GMC

Wang Linqian[1], Qiubo[1+], Luo Ali[2], Kong Xiao[2], Lu Yakun[1], Guo Xiaoyu[1]
[1]Hebei University of Technology, Tianjin 300401, China

[2]National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

Abstract: Galaxy classification has long been a challenging and active research area in astronomy. While machine learning has been applied to galaxy morphological classification in recent years, these efforts face difficulties in feature selection, feature omission, and classifier design. Galaxies can be visually categorized into elliptical, spiral, lenticular, and irregular types. This paper proposes GMC (Galaxy Morphological Classification), a novel method achieving higher classification accuracy for photometric images from SDSS DR16, Galaxy Zoo2, and the EFIGI catalog. Our approach involves image cropping and denoising, followed by data augmentation through rotation, translation, and scaling. We then introduce GMC-net, a dedicated convolutional neural network for galaxy morphological classification. Experimental results demonstrate classification precision of 98.29%, 98.49%, 99.18%, and 99.91% for spiral, elliptical, lenticular, and irregular galaxies, respectively, with corresponding recall rates of 98.44%, 99.03%, 98.89%, and 99.34%. Classification accuracy for galaxies from the EFIGI catalog alone reached 99.34%. These results indicate that GMC outperforms alternative methods and can be effectively applied to galaxy morphological classification.

Keywords: Galaxy morphological classification; Data augmentation; Convolutional neural network

1 Introduction

Advances in observational technology and astronomical instrumentation have led to the implementation of large-scale digital sky surveys such as the Sloan Digital Sky Survey (SDSS) [?], the Cosmic Evolution Survey (COSMOS) [?], and the Large Synoptic Survey Telescope (LSST) [?]. These projects have produced an explosive growth in galaxy observational data. Galaxies are gravitationally bound systems composed primarily of stars, stellar remnants, interstellar gas, dust, and dark matter. Galaxy morphology is intimately connected to galaxy formation and evolution, serving as a crucial parameter for understanding galaxy physics.

With machine learning and deep learning demonstrating remarkable success across various domains, automated galaxy morphological classification methods have developed rapidly. Freed and Lee [?] employed multiple Support Vector Machines (SVMs) to perform three-class classification of spiral, elliptical, and irregular galaxies, achieving a maximum accuracy of 96.8%. Dieleman et al. [?] pioneered the use of convolutional neural networks for this task, training on over 50,000 galaxy images through more than 100 iterations and winning the "Galaxy Zoo Challenge" with an RMS value of 0.07492. Kim and Brunner [?] utilized 17,344 stellar images and 47,656 galaxy images from SDSS DR12, proposing an 11-layer deep convolutional neural network similar to VGG that achieved 99.52% and 99.48% accuracy for star-galaxy classification. Selim and Mohamed [?] performed four-class classification of spiral, elliptical, lenticular, and irregular galaxies from the EFIGI catalog, extracting color, texture, and shape features, selecting the most relevant features using a binary sine-cosine algorithm, and classifying with KNN to obtain accuracies of 97.43%, 100%, 79.48%, and 100%, respectively, with an average accuracy of 94.2%.

Mittal et al. [?] proposed the daMCOGCNN network for galaxy morphological classification, employing data augmentation for irregular galaxies and constructing CNNs with different activation functions to achieve 97% accuracy for elliptical, spiral, and irregular galaxies. In subsequent work, Mittal [?] combined data augmentation with deep learning to classify lenticular, elliptical, and spiral galaxies, achieving 90.2% classification accuracy and 88.3% validation accuracy. Hosny et al. [?] extracted non-redundant color features from galaxy images and proposed a method for finding optimal feature subsets, finally employing Extreme Machine Learning (EML) to classify elliptical, spiral, lenticular, and irregular galaxies with 98% accuracy.

However, current research on galaxy morphological classification still suffers from limited classification categories and severe inter-class sample imbalance, with most previous studies focusing on binary or ternary classification of elliptical, spiral, and lenticular galaxies. Existing methods demonstrate relatively low accuracy when faced with data containing more galaxy types, necessitating an approach that can accurately distinguish among more morphological classes. Our objective is to develop a method capable of automatically classifying spiral, elliptical, lenticular, and irregular galaxies, and even performing cross-database classification of these four types across different catalogs.

As shown in Figure 1 [Figure 1: see original paper], we crop and downsample galaxy images from different databases to filter out low-quality data, while applying denoising and data augmentation to mitigate the effects of image noise and inter-class sample imbalance on the classification model. We then propose GMC-net, a more efficient automated galaxy morphological classification network that circumvents the challenges of image feature extraction, selection, and classifier design, thereby achieving effective classification of the four different morphological types.

2 Data

Figure 1 presents the overall flowchart of our GMC framework. This study primarily utilizes data from SDSS DR16, Galaxy Zoo2, and the EFIGI catalog, all of which ultimately derive from the SDSS digital sky survey. While SDSS obtains raw data in five bands (u, g, r, i, z), the u and z bands contain mostly near-ultraviolet and near-infrared information with very limited useful content. The g, r, and i bands alone are sufficient to reconstruct relatively realistic galaxy images, which is why current research generally employs images synthesized from these three bands [?]-[?], [?][?].

2.1 Data Acquisition

The photometric and spectroscopic data in the EFIGI catalog [?] were obtained from SDSS DR5. The catalog classifies galaxies morphologically into elliptical, lenticular, spiral, irregular, and dwarf types, with each category further divided into subtypes. The morphological parameter T ($T \in [-6, 11]$, where T is an integer representing different galaxy types) can be used to filter galaxies of different morphologies. Table 1 shows the selection criteria for each galaxy type, yielding 920 spiral galaxies, 289 elliptical galaxies, 531 lenticular galaxies, and 248 irregular galaxies from EFIGI.

Galaxy Zoo2 [?] includes 11 tasks and 37 responses, with samples classified only when over twenty participants agreed on the classification. The dataset provides clean sample threshold ranges for each classification task and details the 11 specific tasks. To ensure higher accuracy in sample selection, we set thresholds higher than the recommended values. The detailed parameter threshold settings are explained in the notes to Table 1. Ultimately, we obtained 3,095 spiral galaxies, 4,208 elliptical galaxies, 1,805 lenticular galaxies, and 235 irregular galaxies from Galaxy Zoo2.

This study employs the latest SDSS DR16 [?] photometric data, whose catalog can be cross-matched with the Galaxy catalog in CasJobs [?] using galaxy specObjID to obtain corresponding right ascension and declination. In addition to the main query criteria described in Table 1, we applied the following settings: redshift lower limit of 0.001, redshift upper limit of 0.025, flux lower limit of 50, flux upper limit of 500, image scaling factor of 0.01, and extraction of the top 2000 data points. As the physical constraints for irregular galaxies remain unknown, we did not obtain irregular galaxies from DR16. The distribution of galaxy types in DR16 is also uneven, so we manually filtered out double, merging, and images containing many unknown objects, finally obtaining 913 spiral galaxies, 1,956 elliptical galaxies, and 805 lenticular galaxies.

Table 1: Galaxy Data Selection Criteria

Class	EFIGI Sample Selection	Galaxy Zoo2 Tasks & Threshold Setting	SDSS DR16 Main Query Criteria
Spiral	Sb($T = 3$), Scd($T = 6$)	$f_{\text{features/disk}} > 0.430$, $f_{\text{edge-on,no}} > 0.750$, $f_{\text{spiral,yes}} > 0.719$	$g.\text{lnLDeV}_g < -2000.0$, $g.\text{lnLDeV}_g + 0.1 < g.\text{lnLExp}_g$, $g.\text{lnLDeV}_r > g.\text{lnLExp}_r + 0.1$
Elliptical	cE($T = -6$), E($T = -5$), cD($T = -4$)	$f_{\text{smooth}} > 0.469$, $f_{\text{in\_between}} > 0.70$	$g.\text{lnLExp}_r > -999.0$, $g.\text{lnLDeV}_g > -999.0$

Class	EFIGI Sample Selection	Galaxy Zoo2 Tasks & Threshold Setting	SDSS DR16 Main Query Criteria
Lenticular	S0−($T=-3$), S0^0($T=-2$), S0+($T=-1$)	$f_{\text{features/disk}} > 0.630$, $f_{\text{edge-on,yes}} > 0.785$	g.lnLDeV$_r$ < g.lnLExp$_r + 0.1$, g.lnLDeV$_g + 0.1$ > g.lnLExp$_g$, -1200.0 < g.lnLDeV$_g$ < -1500.0
Irregular	Im($T=10$)	$f_{\text{features/disk}} > 0.430$, $f_{\text{edge-on,no}} > 0.715$, $f_{\text{no\_bar}} > 0.715$, $f_{\text{spiral,no}} > 0.715$, $f_{\text{No\_bulge}} > 0.750$, $f_{\text{odd,yes}} > 0.650$, $f_{\text{irregular}} > 0.715$	N/A

Note: In EFIGI sample selection, the letters (e.g., S0^0) represent galaxy morphological types, with T as the morphological parameter in parentheses. In Galaxy Zoo2, tasks T01–T11 represent 11 classification tasks; $f_{features/disk}$ denotes the frequency of a smooth, disk-like structure, $f_{edge\text{-}on,no}$ the frequency of an image without an edge-on view, $f_{spiral,yes}$ the frequency of being a spiral galaxy, etc. In SDSS DR16 physical constraints, g. is a prefix for the Galaxy library, lnLDeV$_g$ represents the likelihood of a de Vaucouleurs profile fit in the g-band, and lnLExp$_r$ represents the likelihood of an exponential profile fit in the r-band. N_{sample} is the total number of samples.

2.2 Galaxy Image Preprocessing

Convolutional neural networks exhibit stronger learning capabilities and faster training speeds on smaller-sized data [?][?]. To reduce the impact of unnecessary adjacent information in images on experimental results, we first cropped the galaxy data and then performed downsampling. Using lenticular galaxies as an example, Figure 2 [Figure 2: see original paper] shows how a 424×424 pixel image is cropped to 164×164 pixels and then downsampled to 80×80 pixels.

Noise interference occurs during camera capture, image transmission, and digital conversion, and the accumulation of noise severely degrades image quality, altering essential image features. Preserving galaxy contour and texture information is crucial for morphological classification, so we employ an edge-oriented non-local means denoising method [?]. First, we extract edges using a second-order differential Sobel operator. Next, we construct a non-local collaborative filtering framework using both edge information and the original noisy image. Finally, the edge information participates in the restoration of the noisy image. The denoising effect is shown in Figure 3 [Figure 3: see original paper], where noise points around the galaxy are removed while more pronounced edge textures become visible.

The dataset contains relatively few irregular and lenticular galaxies, and inter-class imbalance affects model reliability. We therefore apply data augmentation to increase the number of irregular and lenticular galaxies. The augmentation effects are shown in Figure 4 [Figure 4: see original paper], with the following methods employed [**?**]:

- **Rotation**: Galaxy images exhibit rotational invariance; we exploit this property by randomly rotating images within a range of 30°.
- **Scaling**: Scaling range of 0.7–1.3×.
- **Flipping**: Random flipping along vertical and horizontal axes.
- **Translation**: Objects may not be centered and can be offset in different directions; we apply random horizontal and vertical translation of 0–10 pixels per image.

3.1 GMC-net Network Architecture

As shown in Figure 5 [Figure 5: see original paper], a typical ConvNet [**?**] consists of an input layer, convolutional layers, pooling layers, fully connected layers, and a final output layer. The input layer preprocesses initialized data; convolutional layers extract features; pooling layers compress features and reduce overfitting; and fully connected layers serve as the "classifier."

Inspired by LeNet5' s advantages of having few parameters and being easy to train, and combining characteristics of different activation functions and batch normalization layers, we constructed the GMC-net network. This network not only has fewer trainable parameters but also achieves faster convergence and excellent classification accuracy due to the inclusion of BN layers.

Figure 6 [Figure 6: see original paper] illustrates the overall GMC-net framework, which comprises one input layer, five convolutional layers, one fully connected layer, and one output layer. Table 2 summarizes the parameter settings for each layer. Every convolutional layer in GMC-net is followed by a BN layer and a max pooling layer. BN layers accelerate convergence and training speed, while pooling layers compress features extracted by convolution to reduce overfitting. Additionally, GMC-net employs different activation functions in coordination: the first two layers use the hyperbolic tangent activation function (Tanh) [**?**] for better input to subsequent layers; the middle third and fourth convolutional layers use the Rectified Linear Unit (ReLU) [**?**] for stable convergence and faster computation; the fifth convolutional layer uses Leaky ReLU to suppress neuronal death. Features after the fifth convolutional layer are flattened into a one-dimensional array and fed into the first fully connected layer, which uses ReLU activation and outputs 1600 units. The output layer is configured for 4-way classification using softmax activation.

Table 2: Overview of GMC-net Architecture

Layer	Filters	Filter Size	Padding	Activation Function
Conv_1	32	3×3	Same	Tanh
Pooling_1	-	2×2	-	Max pooling
Conv_2	64	3×3	Same	Tanh
Pooling_2	-	2×2	-	Max pooling
Conv_3	128	3×3	Same	ReLU
Pooling_3	-	2×2	-	Max pooling
Conv_4	256	3×3	Same	ReLU
Pooling_4	-	2×2	-	Max pooling
Conv_5	512	3×3	Same	Leaky ReLU (alpha=0.01)
Pooling_5	-	2×2	-	Max pooling
Fully_1	-	-	-	ReLU
Output	4	-	-	Softmax

3.2 Other Classification Networks

This study also employs AlexNet [?], the convolutional neural network proposed by Dieleman et al. [?], ResNet-26 proposed by Dai et al. [?], and the C2 network specifically designed for galaxy morphological classification by Cavanagh et al. [?]. Table 3 provides a brief introduction to these networks, clearly showing the overall number of layers, layer positions, filter quantities and sizes, pooling methods, and dropout rates in their architecture diagrams. All four comparison networks use ReLU activation functions in their convolutional layers.

Table 3: Introduction to Other Classification Networks

Network	Main Structure	Overall Network Architecture
AlexNet [?]	5 Conv layers, 3 Maxpooling layers, 3 FC layers	Standard AlexNet architecture
Dieleman [?]	5 Conv layers, 3 Maxpooling layers, 3 FC layers	Custom CNN for galaxy classification
ResNet-26 [?]	26 Conv layers, 1 Maxpooling layer, 1 Averagepooling layer	Residual network with 26 layers
C2 [?]	5 Conv layers, 3 Maxpooling layers, 3 FC layers	Specialized for galaxy morphology

4 Experimental Results Analysis and Discussion

This section introduces the performance evaluation metrics, presents classification results using different networks, and compares our approach with related studies.

4.1 Evaluation Metrics

Based on the confusion matrix (Table 4), we calculate four key performance metrics: accuracy, precision, recall, and F1-score.

Table 4: Confusion Matrix

	Predicted False	Predicted True
Actual False	TN (True Negative)	FP (False Positive)
Actual True	FN (False Negative)	TP (True Positive)

- **TP**: Positive samples correctly predicted as positive
- **TN**: Negative samples correctly predicted as negative
- **FP**: Negative samples incorrectly predicted as positive
- **FN**: Positive samples incorrectly predicted as negative

Accuracy reflects the proportion of all correct predictions among total observations. **Precision** measures the proportion of correct positive predictions among all positive predictions. **Recall** measures the proportion of correct positive predictions among all actual positive samples. **F1-score** is the harmonic mean of precision and recall. The formulas are:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

4.2 Training and Validation

All programs were implemented in Python and executed on a desktop with a 2.80 GHz Intel Core i9-10900F CPU, 16 GB RAM, 64-bit Windows system, and accelerated by an RTX 2070 Super GPU. During model training, batch size depends on dataset size and GPU capability; we set it to 64 after comprehensive consideration.

We first tested classification on a combined dataset (Galaxy Zoo2, SDSS DR16, and EFIGI catalog) containing four galaxy types. Before training, we split the dataset into training and validation sets at a 7.5:2.5 ratio, applying data augmentation to both. The final dataset composition is shown in Table 5 .

Table 5: Dataset Information

Dataset	Spiral	Elliptical	Lenticular	Irregular	Total
Data set 1 (Training)	6,872	7,832	2,568	300	17,572
Data set 1 (Test)	2,291	2,611	856	100	5,858
Data set 2 (Training)	1,380	1,156	1,062	439	4,037
Data set 2 (Test)	460	385	354	147	1,346

Note: Data set 1 is the combined dataset from all three sources. Due to the relatively small number of lenticular and irregular galaxies (as shown in Table 1), we applied data augmentation to these classes to reduce inter-class imbalance. Data set 2 consists solely of EFIGI catalog galaxies, originally containing 920 spiral, 289 elliptical, 531 lenticular, and 248 irregular galaxies. To maintain balanced class ratios, we applied varying degrees of augmentation to each class.

During training and validation, we 统计了 the trainable parameters of GMC_{net}, C2, AlexNet, Dieleman' s network, and ResNet-26, as shown in Figure 7 [Figure 7: see original paper]. The number of trainable parameters reflects computational complexity and is a key factor determining training speed. Larger parameter counts indicate more complex networks, which consume more training time on the same hardware and demand higher computational performance. AlexNet and ResNet-26 have significantly more parameters than the other three networks. Dieleman' s network has approximately 3.62 million parameters, C2 has about 3.57 million, and GMC_{net} has only 2.93 million —offering a substantial advantage in training speed.

Figure 8 [Figure 8: see original paper] shows the relationship between training/validation accuracy and epochs for the five CNN architectures (all weights and biases were randomly initialized at the start of training). We present results for 20 epochs. All five networks exhibit rapid accuracy improvement followed by stabilization. AlexNet converges after approximately 10 epochs, achieving maximum training accuracy of 92.3% and validation accuracy of 90.0%. Dieleman's network stabilizes after about 7 epochs, with maximum training accuracy of 96.3% and validation accuracy of 95.2%. ResNet-26 converges more slowly, stabilizing after approximately 16 epochs with maximum training accuracy of 98.2% and validation accuracy of 97.8%. C2 stabilizes after about 6 epochs, achieving 98.5% training accuracy and 97.9% validation accuracy. GMC_{net} converges fastest, stabilizing after only 4 epochs with optimal training accuracy of 99.53% and validation accuracy of 99.18%. GMC_{net} demonstrates the highest accuracy throughout training. In terms of training time, AlexNet and ResNet-26 are the most time-consuming, while GMC_{net} is the fastest.

In summary, GMC_{net} has the fewest trainable parameters while maintaining stable and superior training and validation accuracy compared to the other four networks, with faster convergence—making it the best overall performer.

4.3 Comparison of Classification Results

Table 6 presents the confusion matrix from GMC_{net} validation on Data set 1, from which we calculate accuracy, precision, recall, and F1-score.

Table 6: Confusion Matrix for Data Set 1 Validation Set

Actual Predicted	Spiral	Elliptical	Lenticular	Irregular	Recall	F1-score	Accuracy
Spiral	2,255	18	18	0	98.44%	98.36%	
Elliptical	13	2,585	13	0	99.03%	98.75%	
Lenticular	7	2	846	1	98.89%	99.03%	
Irregular	0	0	1	99	99.34%	99.62%	
Precision	98.29%	98.49%	99.18%	99.91%			**98.93%**

The results show spiral galaxy classification precision of 98.29% and recall of 98.44% (F1-score: 98.36%); elliptical galaxy precision of 98.49% and recall of 99.03% (F1-score: 98.75%); lenticular galaxy precision of 99.18% and recall of 98.89% (F1-score: 99.03%); and irregular galaxy precision of 99.91% and recall of 99.34% (F1-score: 98.36%). The overall classification accuracy is 98.93%.

Table 7 compares the final classification results of the five networks on the 5,858 validation images from Data set 1, showing the best results from multiple repeated trials.

Table 7: Comparison of Verification Results Across Networks

Network	Accuracy	Precision	Recall	F1-score
AlexNet	91.23%	90.15%	92.34%	91.23%
Dieleman	94.92%	95.32%	93.47%	94.38%
ResNet-26	97.82%	98.36%	97.54%	97.94%
C2	98.04%	98.27%	97.96%	98.11%
GMC_{net}	**98.93%**	**98.96%**	**98.90%**	**98.94%**

AlexNet and Dieleman exhibit lower accuracy, precision, recall, and F1-score than the other networks. While ResNet-26 has higher precision than C2, it is slightly lower in accuracy, recall, and F1-score. GMC_{net} achieves the highest accuracy among all five networks, with the highest precision, recall, and F1-score values. GMC_{net}' s classification performance is superior to all comparison networks.

To further demonstrate our method' s feasibility, we separately trained and classified galaxies from the EFIGI catalog alone using GMC_{net} and compared results with other studies. To maintain balanced class ratios, we expanded the EFIGI galaxies into Data set 2 from Table 5. Based on the data descriptions in

[?] and [?], our dataset includes all samples used in those studies (covering all subclasses of the involved types).

Method [?] extracted three feature types: color features (first three order moments), texture features (gray-level co-occurrence matrix including entropy, contrast, correlation, energy), and shape features (contour moments), selected the most relevant features using a binary sine-cosine algorithm, and classified with KNN. Method [?] extracted color features from galaxy color images using quaternion polar complex exponential transform moments (QPET), performed feature selection, and classified using Extreme Machine Learning (EML).

Table 8: Comparison with Other Studies

Method	Accuracy	Precision	Recall	F1-score
[?]	91.9%	92.7%	88.68%	88.68%
[?]	-	-	98.78%	98.74%
GMC_{no}	99.04%	98.71%	98.72%	98.76%
GMC	**99.34%**	**98.88%**	**99.12%**	**98.98%**

Note: GMC_{no} differs from GMC only in that it excludes the denoising step.

Using the EFIGI catalog, method [?] achieved a best classification precision of 92.7% with F1-score of 88.68%. Method [?] achieved a best overall recall of 98.78% with F1-score of 98.74%. Without denoising, GMC_{no}’ s recall was lower than method [?]. After denoising, GMC’ s overall accuracy, precision, recall, and F1-score all improved, surpassing both methods [?] and [?].

Furthermore, methods [?] and [?] face significant challenges in feature and classifier selection, with complex processing and computation. They also suffer from severe inter-class sample imbalance, which can lead to results that do not reflect true distributions and may be highly misleading. Our approach preprocesses images with non-local means denoising to reduce noise effects and applies data augmentation to different galaxy types to mitigate impacts from small sample sizes and uneven distributions. Finally, GMC-net avoids the difficulties of feature extraction, selection, and classifier choice entirely. Overall, our classification method is highly feasible.

4.4 GMC_{net} Convolutional Feature Visualization

We employed Grad-CAM [?] to visualize and interpret GMC_{net}’ s convolutional features. Grad-CAM combines heatmaps with original images to display features of different galaxy types after convolution, reflecting the contribution distribution to network predictions. Higher scores indicate regions in the original image that elicit stronger network responses and greater contributions.

Different convolutional layers in GMC_{net} extract different features: early layers detect edges and corners, intermediate layers extract simple shapes through edge detection, and deep layers identify abstract patterns using combinations of high-level features. Using spiral galaxies as an example, Figure 10 [Figure 10: see original paper] shows that in the fourth convolutional layer, each feature map in the merged map becomes more distinguishable—exactly what a classification model requires. Grad-CAM visualization of features after four convolutional layers clearly reveals the central bulge and spiral arm structures, with feature contribution decreasing spirally from the center outward. This demonstrates GMC_{net}' s high performance in extracting and processing galaxy contour and texture features.

5 Summary and Outlook

Galaxy morphology is closely linked to galaxy formation and evolution and serves as a vital parameter for exploring galaxy physics. Current research on galaxy morphological classification still faces challenges including limited classification categories, difficulty in image feature selection, uneven sample distributions across morphological types, and relatively low classification accuracy. To address these issues, this paper proposes GMC, a convolutional neural network-based galaxy morphological classification method that achieves efficient classification of four types: spiral, elliptical, lenticular, and irregular galaxies.

Our study first preprocesses galaxy images through cropping, downsampling, denoising, and data augmentation to ensure sample diversity and balance while reducing the impact of image noise and inter-class imbalance. We then construct GMC-net, a specialized CNN for galaxy morphological classification that automatically extracts features and performs classification, avoiding the challenges of manual feature extraction, selection, and classifier design. Applying GMC to the combined dataset (SDSS DR16, Galaxy Zoo2, and EFIGI catalog) yielded classification precision of 98.29%, 98.49%, 99.18%, and 99.91% for spiral, elliptical, lenticular, and irregular galaxies, respectively, with recall rates of 98.44%, 99.03%, 98.89%, and 99.34%. Classification accuracy for the four galaxy types from the EFIGI catalog alone reached 99.34%. These results demonstrate that GMC outperforms alternative methods and can be effectively applied to galaxy morphological classification.

While this work advances galaxy morphological classification, several limitations remain for future exploration:

First, to ensure sample accuracy, we used relatively high thresholds in Galaxy Zoo2, potentially underutilizing the dataset. Additionally, since physical parameters for irregular galaxies in SDSS DR16 have not been systematically studied, we could not directly obtain irregular galaxies from DR16. Galaxy morphological classification undoubtedly requires large sample sizes, and numerous data acquisition methods exist. Future work should investigate improved database utilization and application of five-band photometric data.

Second, while GMC_{net} automatically extracts galaxy morphological features and achieves excellent classification accuracy, misclassifications among lenticular, elliptical, and spiral galaxies remain relatively common, and misclassified samples are difficult to distinguish. Future classification systems could explore hybrid models combining expert systems with neural networks—i.e., neural network expert systems—to enhance classification performance.

Acknowledgments

This work was supported by the Joint Research Fund of the National Natural Science Foundation of China and the Chinese Academy of Sciences (U1931134), the Natural Science Foundation of Hebei Province (A2020202001), and the LAMOST Major Achievement Cultivation Project of the Astronomical Big Science Center of the Chinese Academy of Sciences.

References

[1] Lupton R H, Yasuda N. SDSS Imaging Pipelines[J]. 2002.
[2] Scoville N, Aussel H, Brusa M, et al. The Cosmic Evolution Survey (COSMOS) -Overview[J]. The Astrophysical Journal Supplement Series, 2008, 172(1):1.
[3] Ivezic Z, Tyson J A, Acosta E, et al. LSST: from Science Drivers to Reference Design and Anticipated Data Products[J]. American Astronomical Society, 2008.
[4] Freed M, Lee J. Application of Support Vector Machines to the Classification of Galaxy Morphologies[C]// Fifth International Conference on Computational & Information Sciences. IEEE, 2013.
[5] Sander D, Willett K W, Joni D. Rotation-invariant convolutional neural networks for galaxy morphology prediction[J]. Monthly Notices of the Royal Astronomical Society, 2015(2):2.
[6] Kim E J, Brunner R J. Star-galaxy Classification Using Deep Convolutional Neural Networks[J]. Monthly Notices of the Royal Astronomical Society, 2016.
[7] Selim I M, Mohamed A. Automated morphological classification of galaxies based on projection gradient nonnegative matrix factorization algorithm[J]. Experimental Astronomy, 2017, 43(2):131-144.
[8] Mittal A, Soorya A, Nagrath P, et al. Data augmentation based morphological classification of galaxies using deep convolutional neural network[J]. Earth Science Informatics, 2019, 13(1).
[9] Mittal M. Morphological classification of galaxies using Conv-nets[J]. Earth Science Informatics, 2020(1):1-10.
[10] Hosny K M, Aziz M, Selim I M, et al. Classification of galaxy color images using quaternion polar complex exponential transform and binary Stochastic Fractal Search[J]. Astronomy and Computing, 2020, 31.
[11] The Sloan Digital Sky Survey: Technical Summary[J]. Astronomical Journal, 2000, 120(3):1579.
[12] Anderson S F, Arns J A, Aubourg E, et al. SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way Galaxy, and Extra-Solar Planetary Systems[J]. Astronomical Journal, 2011, 142(3):725-735.

[13] Baillard A, Bertin E, Lapparent V D, et al. The EFIGI catalogue of 4458 nearby galaxies with detailed morphology[J]. Astronomy and Astrophysics, 2011, 532.
[14] Bamford S P, Masters, et al. Galaxy zoo 2: Detailed morphological classifications for 304,122 galaxies from the sloan digital sky survey.
[15] Li N, Szalay A. CASJobs: A workflow environment designed for large scientific catalogs. IEEE, 2008.
[16] Flesch E W. Identification confusion and blending concealment in the SDSS-DR16 Quasar catalogues -40 new quasars and 82 false quasars identified[J]. 2020.
[17] Yong X, Zhong J. Down-Sampling Face Images and Low-Resolution Face Recognition[C]// International Conference on Innovative Computing Information & Control. IEEE, 2008.
[18] Zhang Y, Zhao D, Zhang J, et al. Interpolation-dependent image downsampling[J]. IEEE Transactions on Image Processing, 2011, 20(11):3291-3296.
[19] Sander D, Willett K W, Joni D. Rotation-invariant convolutional neural networks for galaxy morphology prediction[J]. Monthly Notices of the Royal Astronomical Society, 2015(2):2.
[20] Fu Bo, Wu Yuechu, Wang Liyan, Wang Ruizi. Edge Map Oriented Non-Local Means Filtering Algorithm[J]. Journal of Jilin University (Information Science Edition), 2020, 38(06):687-693.
[21] Fukushima K, Miyake S, Ito T. Neocognitron: A neural network model for a mechanism of visual pattern recognition[J]. Systems Man & Cybernetics IEEE Transactions on, 1982, SMC-13(5):826-834.
[22] Mathias A C, Rech P C. Hopfield neural network: the hyperbolic tangent and the piecewise-linear activation functions[J]. Neural Networks, 2012, 34(10):42-45.
[23] Pedamonti D. Comparison of non-linear activation functions for deep neural networks on MNIST classification task[J]. 2018.
[24] Technicolor T, Related S, Technicolor T, et al. ImageNet Classification with Deep Convolutional Neural Networks[50].
[25] Dai J M, Tong J. Galaxy Morphology Classification with Deep Convolutional Neural Networks[J]. 2018.
[26] Cavanagh M K, Bekki K, Groves B A. Morphological classification of galaxies with deep learning: comparing 3-way and 4-way CNNs[J]. Monthly Notices of the Royal Astronomical Society, 2021(1):1.
[27] Clevert, Djork-Arné, Unterthiner T, Hochreiter S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)[J]. Computer Science, 2015.
[28] Simon M, Rodner E, Denzler J. ImageNet pre-trained models with batch normalization[J]. 2016.
[29] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization[J]. International Journal of Computer Vision, 2020, 128(2):336-359.

Note: Figure translations are in progress. See original paper for figures.