AI-ASSISTED CLASSIFICATION OF CELESTIAL OBJECTS USING TRANSFER LEARNING ON CONVOLUTIONAL NEURAL NETWORKS

Rohan Vishwanath Chatse
Adarsh College, Hingoli Dist. Nanded
rohancrchatse@gmail.com

Pallavi Ramdas Jadhav
Adarsh College, Hingoli Dist. Nanded
pallavijadhav@gmail.com

Abstract

Astronomical data collection has rapidly expanded with the advent of large-scale surveys and advanced telescopes. The classification of celestial objects such as stars, galaxies, and nebulae is crucial for understanding the structure and evolution of the universe. Traditional manual classification by astronomers is time-consuming and prone to human error. This study presents an artificial intelligence (AI)-based approach for the automated classification of celestial objects using transfer learning on convolutional neural networks (CNNs). The proposed method utilizes the pre- trained VGG16 model, fine-tuned on the publicly available Galaxy Zoo dataset consisting of thousands of labeled galaxy images. Experimental results demonstrate that the transfer learning approach achieves an overall accuracy of 89.7% on test data, outperforming standard CNN models trained from scratch. This work highlights the potential of AI and deep learning techniques in modern astronomy, enabling faster and more accurate identification of celestial bodies.

***Keywords**: Artificial Intelligence, Astronomy, Convolutional Neural Networks, Transfer Learning, Galaxy Classification, VGG16*

1. Introduction

Astronomy has entered a data-driven era where vast amounts of observational data are generated daily by telescopes and space missions. Projects like the Sloan Digital Sky Survey (SDSS), Hubble Space Telescope, and James Webb Space Telescope have produced millions of high-resolution celestial images. Analyzing and classifying these objects manually is an arduous process, demanding substantial time and expertise. To address this challenge, artificial intelligence (AI) and deep learning techniques have emerged as powerful tools for automating tasks in astronomy.

Machine learning (ML) models, especially convolutional neural networks (CNNs), have demonstrated exceptional performance in image recognition and pattern analysis. Their application to astronomical image classification enables researchers to automatically categorize galaxies, stars, and nebulae with high accuracy. However, training deep CNNs from scratch requires large datasets and computational resources. Transfer learning, which leverages knowledge from pre-trained models, offers an efficient alternative for astronomical image analysis with limited data.

This research aims to develop and evaluate a transfer learning-based CNN model using the VGG16 architecture for classifying celestial objects. The model's performance is analyzed using standard evaluation metrics, and the results are compared to baseline CNN approaches. The findings contribute to the growing field of AI-assisted astronomy, demonstrating the potential of deep learning in automating celestial image classification.

2. Literature Review

Several studies have applied AI techniques in the field of astronomy, focusing on automated classification and analysis of celestial images. Ball and Brunner (2010) discussed the early applications of machine learning for astronomical data mining and object recognition. Dieleman et al. (2015) introduced rotation-invariant convolutional neural networks for galaxy morphology prediction using Galaxy Zoo data, achieving significant accuracy improvements over traditional methods.

Krizhevsky et al. (2012) demonstrated the power of CNNs in visual recognition through the ImageNet classification task, paving the way for their use in astronomical image processing. Domínguez Sánchez et al. (2018) utilized deep learning models for morphological classification of galaxies using data from SDSS, achieving near-human accuracy levels.

More recent studies have explored transfer learning to enhance model performance in astronomy. Huertas-Company et al. (2019) employed pre-trained models such as VGG and ResNet to classify galaxy morphologies efficiently. Similarly, Walmsley et al. (2020) used transfer learning for classifying radio galaxies and found it highly effective for small datasets. These studies validate the growing relevance of deep learning, especially transfer learning, in astronomical data analysis.

However, there remains a gap in lightweight, reproducible, and resource-efficient AI models suitable for small research institutions. This study aims to address this gap by fine-tuning the VGG16 model for galaxy classification and demonstrating its practical use for student-level and academic research.

3. Methodology

The proposed system employs transfer learning on a convolutional neural network architecture to classify celestial images into three categories: spiral galaxies, elliptical galaxies, and irregular galaxies. The overall methodology involves dataset preparation, data preprocessing, model design using transfer learning, training, and evaluation.

3.1 Dataset

The dataset used in this research is the Galaxy Zoo dataset, publicly available on Kaggle.

It contains thousands of galaxy images categorized by volunteer classifications. The dataset includes spiral, elliptical, and irregular galaxies. The data is divided into 80% for training and 20% for testing. Images are resized to 128×128 pixels to standardize input dimensions.

3.2 Preprocessing

Data preprocessing is crucial to improve the performance of CNN models. The steps include:

Image resizing and normalization (pixel values scaled between 0 and 1)

Data augmentation (rotation, flipping, zooming) to reduce overfitting

Splitting the dataset into training and validation sets

3.3 Model Architecture

This study uses the VGG16 architecture, a pre-trained deep CNN originally trained on the ImageNet dataset. Transfer learning enables reuse of its convolutional layers for feature extraction while modifying the fully connected layers for new classification tasks.

The modified model architecture includes:

Pre-trained VGG16 base with frozen convolutional layers

Flattening layer

Fully connected Dense layer (128 neurons, ReLU activation)

Dropout (0.5) to prevent overfitting

Output layer (3 neurons, Softmax activation) for multi-class classification

The model is compiled using:

Optimizer: Adam

Loss Function: Categorical Cross-Entropy

Metrics: Accuracy

3.4 Implementation

The model is implemented using TensorFlow and Keras libraries in the Google Colab environment, utilizing GPU acceleration. The model is trained for 25 epochs with a batch size of

32. The training and validation accuracy are plotted to analyze model performance.

4. Results and Discussion

After training, the model achieved an overall accuracy of 89.7% on the test dataset. The confusion matrix indicates that the model correctly classifies most spiral and elliptical galaxies, with minor misclassifications in irregular galaxies due to their diverse structures.

Metric

Trining Accuracy -93.4%

Validation Accuracy- 88.9%

Test Accuracy- 89.7%

Precision -0.90

Recall - 0.88

F1-Score - 0.89

The results show that transfer learning using VGG16 significantly outperforms a standard CNN trained from scratch, which achieved only 82.3% accuracy under identical conditions.

The feature extraction capability of pre-trained networks enhances the learning process, enabling the model to recognize complex structures in galaxy images even with a smaller dataset. Figure 1 illustrates the training vs. validation accuracy plot, showing smooth convergence without significant overfitting.

These findings align with previous research demonstrating that transfer learning is effective for astronomical image classification. Moreover, this work demonstrates that students and small research labs with limited computational resources can achieve high performance using pre-trained CNN models.

5. Conclusion and Future Scope

This research successfully demonstrates the use of artificial intelligence and deep learning for automating celestial object classification. By applying transfer learning with the VGG16 model, the proposed approach achieved an impressive accuracy of 89.7% on the Galaxy Zoo dataset. The results confirm that transfer learning can significantly reduce computational cost and training time while maintaining high accuracy.

Future Scope:

Future research can extend this work by:

1. Implementing advanced architectures like ResNet, EfficientNet, or Vision Transformers for improved performance.
2. Incorporating additional celestial categories, such as quasars, nebulae, and exoplanets.
3. Using larger and more recent datasets from NASA and the James Webb Space Telescope.
4. Exploring Explainable AI (XAI) methods to interpret model predictions and visualize decision regions.
5. Deploying real-time classification tools for observatories and astronomy research centers.

AI-assisted astronomy continues to transform how we study the universe. With ongoing advances in deep learning, the integration of AI into astronomical image analysis promises faster discoveries and a deeper understanding of celestial phenomena.

References

1. Ball, N. M., & Brunner, R. J. (2010). Data mining and machine learning in astronomy. International Journal of Modern Physics D, 19(7), 1049–1106.
2. Dieleman, S., Willett, K. W., & Dambre, J. (2015). Rotation-invariant convolutional neural networks for galaxy morphology prediction. Monthly Notices of the Royal Astronomical Society, 450(2), 1441–1459.
3. Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer, J. L. (2018). Improving galaxy morphology classification with deep learning. Monthly Notices of the Royal Astronomical Society, 476(3), 3661–3676.
4. Huertas-Company, M., et al. (2019). A deep learning approach to galaxy morphology classification. Astronomy & Astrophysics, 625, A149.
5. Walmsley, M., et al. (2020). Radio Galaxy Zoo: Machine learning for radio source host galaxy cross-identification. Monthly Notices of the Royal Astronomical Society, 491(2), 1554–1574.
6. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097– 1105.