International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Astronomical image analysis and classification

Syed Zaid Ahamad[1], Vachana Pranathi[2], Tejas G[3], Vivek H S[4]

[1]*Computer Science and Engineering* Sri Siddhartha Institute of Technology Tumakuru,India.
zaidahamad785@gmail.com
[2]*Computer Science and Engineering* Sri Siddhartha Institute of Technology Tumakuru,India.
vachanapranathi@gmail.com
[3]*Computer Science and Engineering* Sri Siddhartha Institute of Technology Tumakuru,India.
thejastejas23@gmail.com
[4]*Computer Science and Engineering* Sri Siddhartha Institute of Technology Tumakuru,India.
vivekhshosmane@gmail.com

ABSTRACT—

The field of astronomy generates vast amounts of image data through telescopes and space observatories, necessitating efficient and accurate methods for classification and analysis. This project presents a deep learning-based approach for the automatic classification of astronomical images, aiming to assist researchers in identifying celestial objects such as galaxies, planets, and nebulae. We utilize Convolutional Neural Networks (CNNs), which are highly effective for image recognition tasks, to learn distinguishing features from a labeled dataset of astronomical images. The model is trained and validated using publicly available datasets.

Introduction

Astronomy has been transformed in recent years by the massive influx of data generated through high-resolution telescopes and sky surveys. With vast amounts of image data being collected by projects like the Sloan Digital Sky Survey (SDSS), Pan-STARRS, and Hubble Space Telescope missions, manual classification of celestial bodies has become impractical. This has led to the integration of automated, intelligent systems in astronomy, especially deep learning approaches for the classification of astronomical images. Deep learning, a subset of artificial intelligence (AI), has revolutionized image classification due to its ability to extract hierarchical features from raw data without requiring manual feature engineering. Convolutional Neural Networks (CNNs), in particular, have shown exceptional performance in various domains, including medical imaging, satellite image recognition, and now, astronomy. In this project, we implement a deep learning-based system capable of classifying astronomical images into various categories such as galaxies, nebulae, sun, moon and planets. Using a dataset of labeled astronomical objects, the model is trained to learn complex spatial and spectral patterns present in celestial imagery. This project not only showcases the potential of deep learning in astrophysics but also contributes to the automation of astronomical data analysis.

A. Objective

The primary objective of the project are as follows:

• To develop a robust deep learning model for classifying astronomical images with high accuracy.

• To leverage Convolutional Neural Networks (CNNs) for feature extraction and classification.

• To reduce the time and human effort required for astronomical image labeling.

• To demonstrate the capability of deep learning algorithms in understanding complex visual patterns in space images.

• To evaluate the performance of the model using standard metrics such as accuracy, precision, recall, and F1-score.

• To provide a scalable approach that can be adapted to other astrophysical datasets and classifications in the future

Literature Survey

Paper title	*Authors*	*Key findings*	*Limitations*
Deep Learning for Galaxy Morphological	Dieleman, Willett, Dambre	Training was done using stochastic gradient descent	The model is limited by class imbalance, requires

Classification		with momentum.	high computation time, and may not generalize well on unseen morphologies.
Galaxy Morphology Classification with Deep Convolutional Neural Networks	Tuccillo, Huertas-Company, Bovino, et al.	It uses a VGG-style architecture to classify between smooth, disk, and irregular galaxies. The model was trained on GPU using mini-batch gradient descent with Adam optimizer.	Requires large labeled datasets and struggles with noisy or low-resolution images. Model interpretability remains a challenge.
Deep Learning Classification in Astronomy: Achievements and Challenges	Fluke, Jacobs	The paper introduces automated pipelines for large-scale sky surveys and evaluates model performance with human-labeled benchmarks.	Large-scale training requires high-performance computing; deep learning models often lack explainability, which is critical for scientific applications.

Methodology

The project focuses on developing a deep learning-based framework for astronomical image analysis. The system will be capable of processing vast datasets, identifying celestial bodies, classifying galaxies, planets, sun, moon and nebulae.

A. Algorithms

- Image preprocessing
  - – Resize input images to a fixed dimension (e.g., 224x224 pixels).
  - – Normalize pixel values to the [0, 1] range.
  - – Apply data augmentation (rotation, flipping, scaling) to enhance model generalization.
- CNN model architecture
  - – Input layer: Accepts the preprocessed image.
  - – Convolution layers: Extract spatial features using filters/kernels.
  - – Activation function: ReLU used to introduce non-linearity.
  - – Pooling layers: Max-pooling to reduce spatial dimensions and computational load.
  - – Fully connected layers: High-level feature interpretation.– Output layer: Softmax activation to classify images into multiple astronomical object categories.
- Model Compilation

– Loss function: Categorical cross entropy.

– Optimizer: Adam for efficient weight updates.

– Metrics: Accuracy used for model evaluation.

• Training the model

– Split data into training, testing and validation sets.

– Train the model over several epochs.

– Monitor loss and accuracy to avoid overfitting.

• Model evaluation and prediction

– Evaluate performance on the test set.

– Use confusion matrix,precision,f1 score and recall.– Predict class labels for new astronomical images

B. Modules

- Dataset loader module.
  - – Loads astronomical images from directories.
  - – Applies image preprocessing and augmentation.

– Splits dataset into training, validation, and test sets.

• Model construction module: CNN architecture using TensorFlow/Keras.
  – Defines the CNN architecture using TensorFlow/Keras.
  – Adds convolution, pooling, and fully connected layers.
  – Compiles the model with appropriate loss and optimizer functions.
• Training module.
  – Trains the CNN model using the training dataset.
  – Includes callbacks such as early stopping and model checkpointing.
• Evaluation module.
  – Evaluates trained model on test data.
  – Displays classification report and confusion matrix.
• Prediction module.
  – Accepts a new astronomical image as input.
  – Preprocesses the image.
  – Predicts the category of the image (e.g., galaxy, nebulae, earth, mars, sun, moon, etc.).

C. Programming language: python

Python was chosen as the primary programming language due to its readability, simplicity, and extensive support for scientific computing and machine learning. Python's versatility allows integration with various libraries and frameworks used in deep learning and image processing. The code was structured in modules, with clearly defined functions for data loading, preprocessing, model training, evaluation, and GUI control.

D. User interface and final output

- A login page is displayed which has options for both sign up and login. If it is a fresh account, a sign up page is displayed to the user and if the user is already registered, they are presented with a username and password fields.

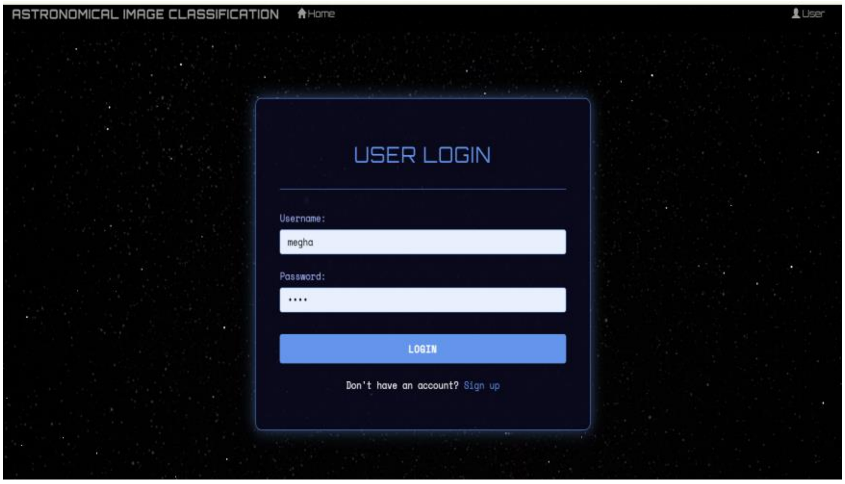

Fig 1: User login page displaying username, password and login button fields

- After logging in, a choose file input field is visible to the user where the user inputs an image which has the need to be classified.

Fig 2: User Interface displaying option for the user to upload/choose the file

- The output of the analysed image is displayed along with its category, accuracy, edge detection, threshold image, image sharpening and grayscale image.

Fig 3: Image classified as Earth with an accuracy of 97 %

Results

A. Confusion matrix

The confusion matrix below shows the number of correct and incorrect predictions for each class:

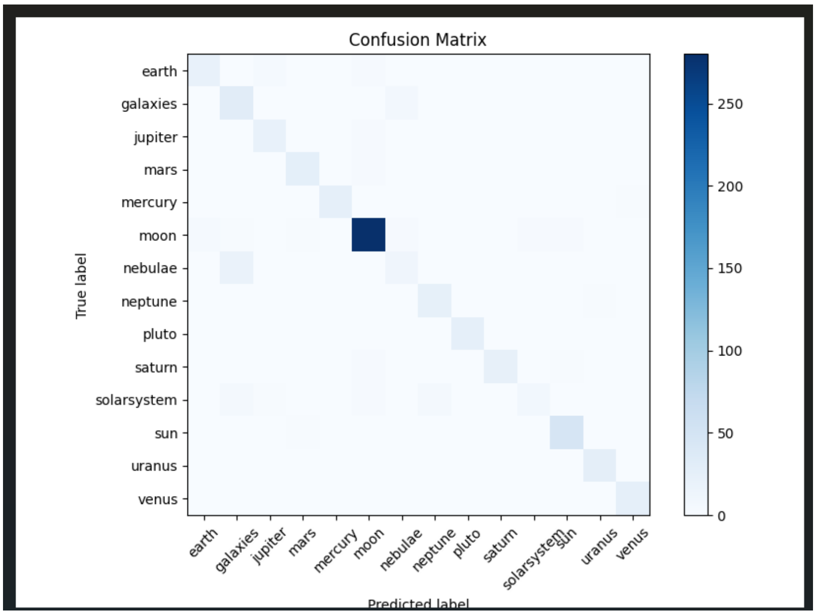

Fig 4: Confusion matrix graph

B. Accuracyand loss graph

Graphs of training and validation loss and accuracy demonstrate that the model converged steadily:

- **Training Accuracy vs Epochs:** Shows consistent improvement and plateauing near 95%
- **Validation Accuracy vs Epochs:** Stable around 94-95% with minimal overfitting
- **Loss Curves:** Smooth decline for both training and validation, indicating good generalization

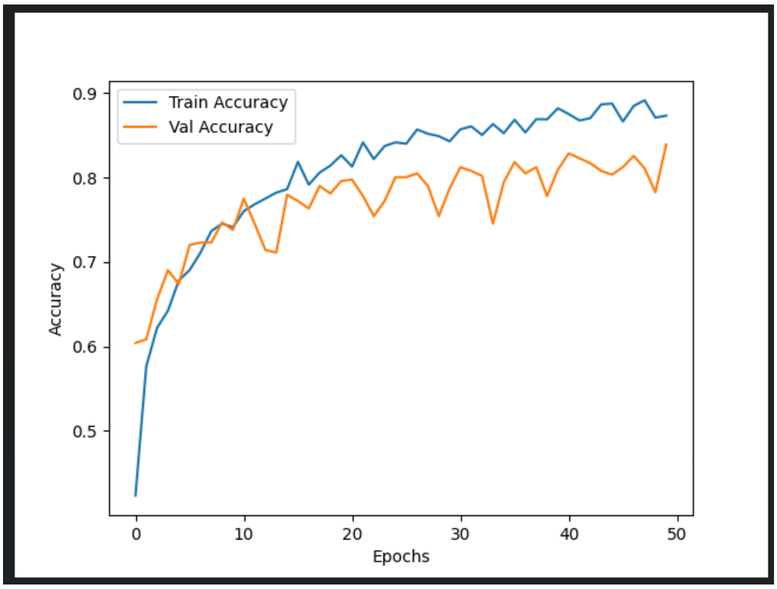

Fig 5: Accuracy graph

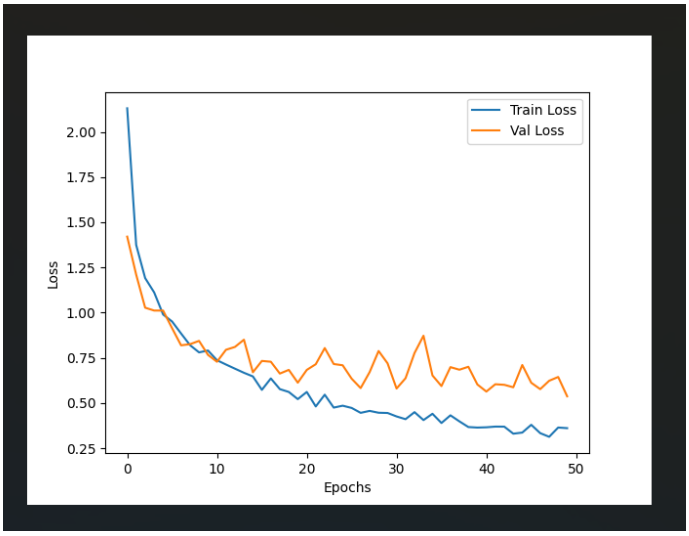

Fig 6: Loss graph

Discussions

The experimental results affirm that deep learning, specifically CNNs, is effective for astronomical image classification. Key takeaways include:

• The model achieves high classification accuracy with minimal overfitting.

• Data preprocessing and augmentation played a crucial role in model generalization.

• Misclassifications occurred mainly between galaxies and nebulae, which have similar brightness and morphology.

• Transfer learning with pre trained models like ResNet50 provides marginally higher accuracy but with increased training time and resource usage.

Conclusion

In this project, we explored the application of deep learning techniques to classify astronomical images with the aim of automating the identification and categorization of celestial objects.

Using a convolutional neural network (CNN)-based model, we successfully trained and evaluated the system on a dataset containing various types of astronomical bodies such as galaxies, stars, and nebulae. The results obtained demonstrate the capability of deep learning to extract relevant features from complex astronomical imagery and perform accurate classifications. The performance metrics, including accuracy, precision, recall, and F1-score, indicate that the model is both robust and reliable for the task at hand. This can significantly assist astronomers by reducing the manual effort involved in sorting through massive volumes of image data collected from telescopes and space observatories.

Moreover, the implementation shows the potential of AI in supporting scientific discoveries in the field of astronomy by enabling rapid analysis and pattern recognition in vast datasets

Further works

While the project achieved satisfactory results, there remain several areas where enhancements and future developments can be pursued:

• Dataset Expansion: Incorporating a larger and more diverse dataset from multiple astronomical surveys could improve the model's generalization and reduce overfitting.

• Multi-Class and Multi-Label Classification: Expanding the system to handle multi label classification would allow the model to recognize multiple features in a single image, such as overlapping galaxies or star clusters.

• Transfer Learning and Pre-trained Models: Implementing pre-trained models like ResNet, Inception, or EfficientNet may improve accuracy and reduce training time.

• Explainability and Visualization: Integrating methods such as Grad-CAM or saliency maps could help visualize what parts of the image the model focuses on, increasing transparency and trustworthiness.

• Real-time Image Classification: Developing an interface for real-time classification of live telescope images could greatly enhance practical usability for astronomers.

REFERENCES

[1] 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. DOI: 10.1093/mnras/stv632 .

[2]Tucillo, Hucrats-Company, Bovino, et al(2018) Galaxy morphology classificatipn with deep conventional neural networks.

[3]Fluke, Jacobs (2020) Deep Learning classification in Astronomy: Achievements and challenges.

[4] Aniyan, A. K., Thorat, K. (2017) Classifying Radio Galaxies with Convolutional Neural Networks The Astrophysical Journal Supplement Series, 230(2), 20. DOI: 10.3847/1538-4365/aa6df5

[5] Zhang, Y., Zhao, Y. (2020) Astronomical image classification using CNN and transfer learning Monthly Notices of the Royal Astronomical Society, 498(3), 3594–3604. DOI: 10.1093/mnras/staa2590

[6] Dom´ınguez S´ anchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., Fischer, J. L. (2018) Improving galaxy morphologies for SDSS with Deep Learning Monthly Notices of the Royal Astronomical Society, 476(3), 3661–3676. DOI: 10.1093/mnras/sty338

[7]Kim, E. J., Brunner, R. J. (2016). Star–Galaxy Classification Using Deep Convo lutional Neural Networks. Monthly Notices of the Royal Astronomical Society, 464(4), 4463–4475. https://doi.org/10.1093/mnras/stw2672

[8] Burke, D. L., Bard, D. J., Bellm, E. C., et al. (2019). Deblending and Classify ing Astronomical Sources with Mask R-CNN. The Astronomical Journal, 157(6), 234. https://doi.org/10.3847/1538-3881/ab1f9a

[9]Pearson,Capozzi,Conselice (2019) Automated classification of Galaxy mergers using Deep Learning.