Creating a CNN-Based App for Classifying Galaxies

Tucker Adeyemi, Mathematical and Computational Sciences Dept.
Advised by Prof. Sofia Visa

Project overview

Deep learning models, particularly Convolutional Neural Networks (CNNs), have shown immense success in visual tasks, but specialized astronomical applications are limited. Our research trained a CNN to identify galaxy images into five distinct classes. This evaluated its overall utility, ultimately building an app for classifying galaxies in real time.

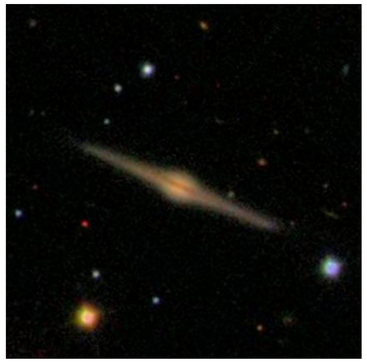

Figure 1. An example of an image used in the training of the CNN.

Dataset and contouring

The research utilized the extensive Galaxy Zoo 2 dataset, comprising over 60,000 galaxy images classified by crowdsourcing. A classification script traversed a binary decision tree to automatically sort images into five primary morphology classes.

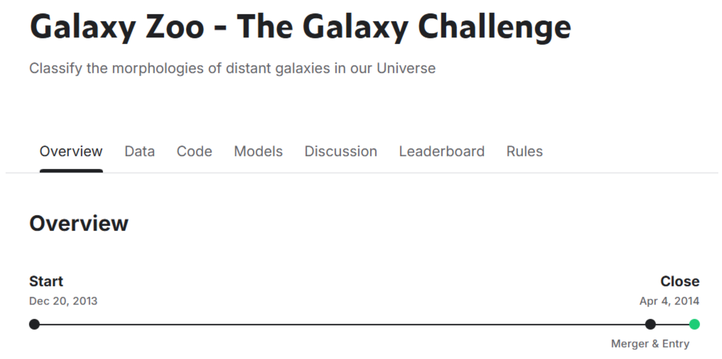

Figure 2. The front page of the Galaxy Zoo 2 competition.

Model Architecture

We developed a specialized sequential CNN using TensorFlow and Keras, following a common VGG-style structure but with optimized hyperparameters. Features include multiple convolutional layers (increasing filter depth 32->64->128), Max-pooling, and Dropout layers for regularization.

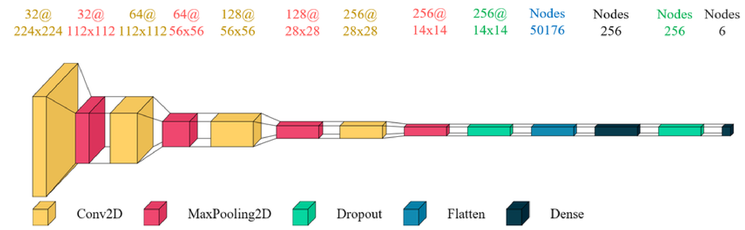

Figure 3. Example of a Sequential CNN Model Architecture

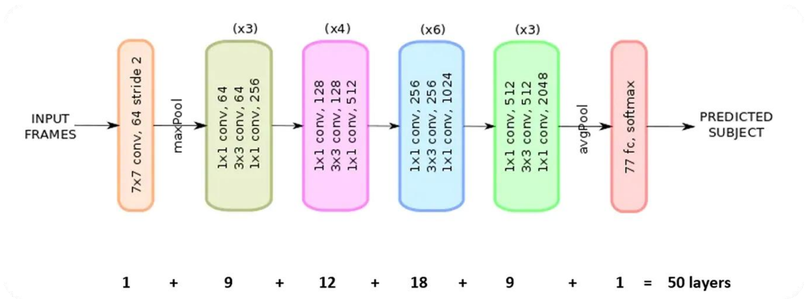

Figure 4. Example of a ResNet-50 Architecture.

Results and App Implementation

Our sequential CNN achieved validation accuracies consistently between 70% and 90% across classes. Manual testing of 50 images confirmed the high predictive general accuracy, with errors mostly on edge cases. This model was deployed in a user-facing web app.

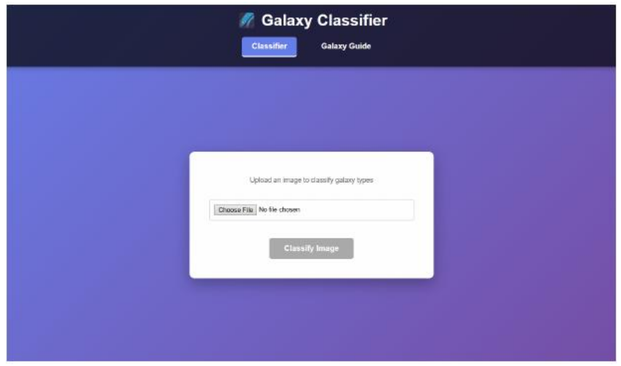

Figure 5. Figure showing the layout of the user-facing application.

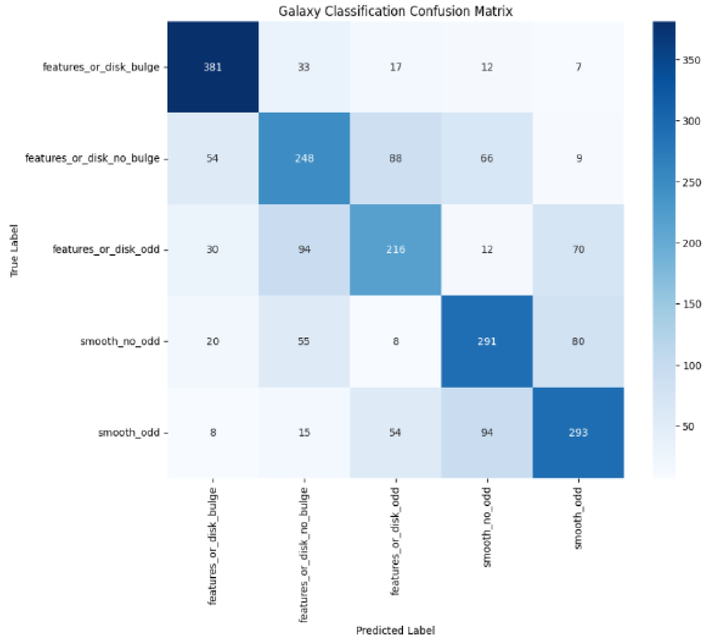

Figure 6. Confusion matrix for the final sequential CNN.

Results and future work

Future work includes expanding the training dataset, especially with more varied and irregular galaxy shapes. We aim to refine the app's performance on diverse image sources and incorporate multi-modal data.

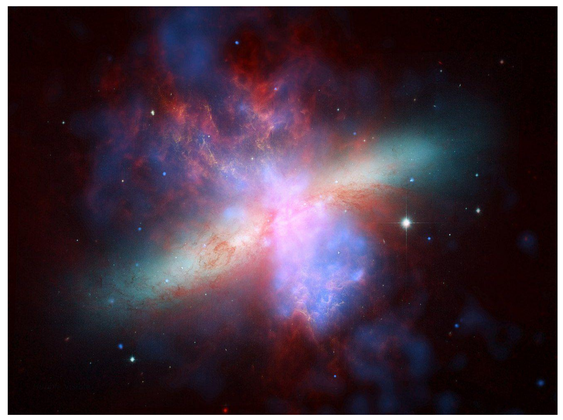

Figure 7. An example of an irregular galaxy.

Acknowledgment: Thank you to my parents and my advisor for support throughout the project.

References
[1, 2] Kaggle. *Galaxy Zoo: The Galaxy Challenge*. https://www.kaggle.com/competitions/galaxy-zoo-the-galaxy-challenge/overview
[3] Ahmed, Faisal & Islam, Md & Ahsan, Sk Md Masudul. (2021). Building Infrastructure Classification with Hybrid Convolutional Neural Network Architecture. 1-6.
[4] Minakshi Vharkate, Rajani Sajjan, and Janhvi Kadam. "Comparative Study of Pretrained Models for Remote Sensing Image Classification". In: International Research Journal on Advanced Engineering Hub (IRJAEH) 3 (2025), pp. 3270–3279.
[7] Millis, J. P. (2025, May 8). *Irregular galaxies: Oddly shaped mysteries of the universe*. ThoughtCo. https://www.thoughtco.com/irregular-galaxies-mysteries-of-the-universe-3072046