Galaxy Morphology Classification with CNN

Date

Dec 14, 2024

Submitted by

Mohamed Hassan 202201257

Yusuf Tamer 202201929

Course

DSAI 308 Deep Learning

Table of Contents

1. Methodology

1.1 Datasets

Galaxy DECaLS: https://astronn.readthedocs.io/en/latest/galaxy10.html

Galaxy SDSS: https://astronn.readthedocs.io/en/latest/galaxy10sdss.html

We decided to choose multiple derivatives from the Galaxy Zoo dataset, which seems to be the main dataset used for galaxy morphology.

The classes for the Galaxy DECaLS dataset are labelled as follows:

```
Galaxy10 dataset (17736 images)
├── Class 0 (1081 images): Disturbed Galaxies
├── Class 1 (1853 images): Merging Galaxies
├── Class 2 (2645 images): Round Smooth Galaxies
├── Class 3 (2027 images): In-between Round Smooth Galaxies
├── Class 4 ( 334 images): Cigar Shaped Smooth Galaxies
├── Class 5 (2043 images): Barred Spiral Galaxies
├── Class 6 (1829 images): Unbarred Tight Spiral Galaxies
├── Class 7 (2628 images): Unbarred Loose Spiral Galaxies
├── Class 8 (1423 images): Edge-on Galaxies without Bulge
└── Class 9 (1873 images): Edge-on Galaxies with Bulge
```

The classes for the Galaxy SDSS dataset are labelled as follows:

```
Galaxy10 dataset (21785 images)
├── Class 0 (3461 images): Disk, Face-on, No Spiral
├── Class 1 (6997 images): Smooth, Completely round
├── Class 2 (6292 images): Smooth, in-between round
├── Class 3 (394 images): Smooth, Cigar shaped
├── Class 4 (1534 images): Disk, Edge-on, Rounded Bulge
├── Class 5 (17 images): Disk, Edge-on, Boxy Bulge
├── Class 6 (589 images): Disk, Edge-on, No Bulge
├── Class 7 (1121 images): Disk, Face-on, Tight Spiral
├── Class 8 (906 images): Disk, Face-on, Medium Spiral
└── Class 9 (519 images): Disk, Face-on, Loose Spiral
```

1.2 Pre-Processing

The main Pre-Processing steps we did were the usual image resizing instead of (256, 256, 3) we made it (128, 128, 3) as to improve the speed of the model and not crash the session. Then we normalize the images and set the labels to_categorical. However, the data was not balanced a distribution that favoured some classes over others as can be seen in section 2.2. We added weights to each class during the training process to mitigate the problem.

1.3 Training, Models, and Architecture

Given the scope of this project as an 'Image Classification' project, we opted for using CNN models. Since the features we are trying to extract are very high ordered such as the spatial features of the galaxies and their longevity and width etc… that are usually extracted with the human's eyes to be able to classify the galaxy's morphology, we viewed the option of using deeper pre-trained models as more effective.

Six legacy pre-trained models were initially used and trained on the dataset in this project to test the difference between each of them and the highest accuracy each of them can reach on the astroNN datasets. These models are, EfficientNetB0, EfficientNetB3, EfficientNetB4, ResNet50, InceptionV3, and MobileNetV2.

These models were first trained and evaluated using a validation set and were trained multiple times using different methods such as changing the patience of the callback, using randomly initialized weights and using imagenet weights, unfreezing all layers, unfreezing half of the layers, and only training the output layer. After this training, we only decided to move forward with 4 of these models due to their higher accuracy range: EfficientNetB0, EfficientNetB3, and EfficientNetB4.

1.3.1 Architectural Modification

After the initial training and evaluation of the models, the accuracy of each one alone was still not satisfying this led us to apply an ensemble-like technique where we have a voting classifier using the four mentioned models. This leads the accuracy to increase from individual accuracies ranging from 75 to 79 into an ensemble accuracy reaching 83% on the test data.

1.4 Metrics

For the metrics, we used Accuracy and a confusion matrix to measure the precision, recall and f1_score of each label which will be shown in detail in the conclusion section.

1.5 Model Generalization

In this part, we discuss how we generalized the data to be able to test both models. We first focused on resizing the images to be compatible with the model's input layer. Then, realized that the data classes were different from each other so we mapped the similar classes between them to be able to classify the images for the Galaxy SDSS dataset using the model trained on the Galaxy DECaLS dataset.

2. Challenges

2.1 Environment

Google Colab and Kaggle had limited time usage and GPU which were an annoyance while working on the project. Google Colab also had limited RAM usage. This lack of resources led to multiple issues in the training process which we mitigated by resizing the image and using a small batch size during the training process.

2.2 Data Issues

The classes in the dataset utilized were not balanced as seen in Figure 1, samples from some classes such as class 4 were very low in count leading to a biased model if trained on these

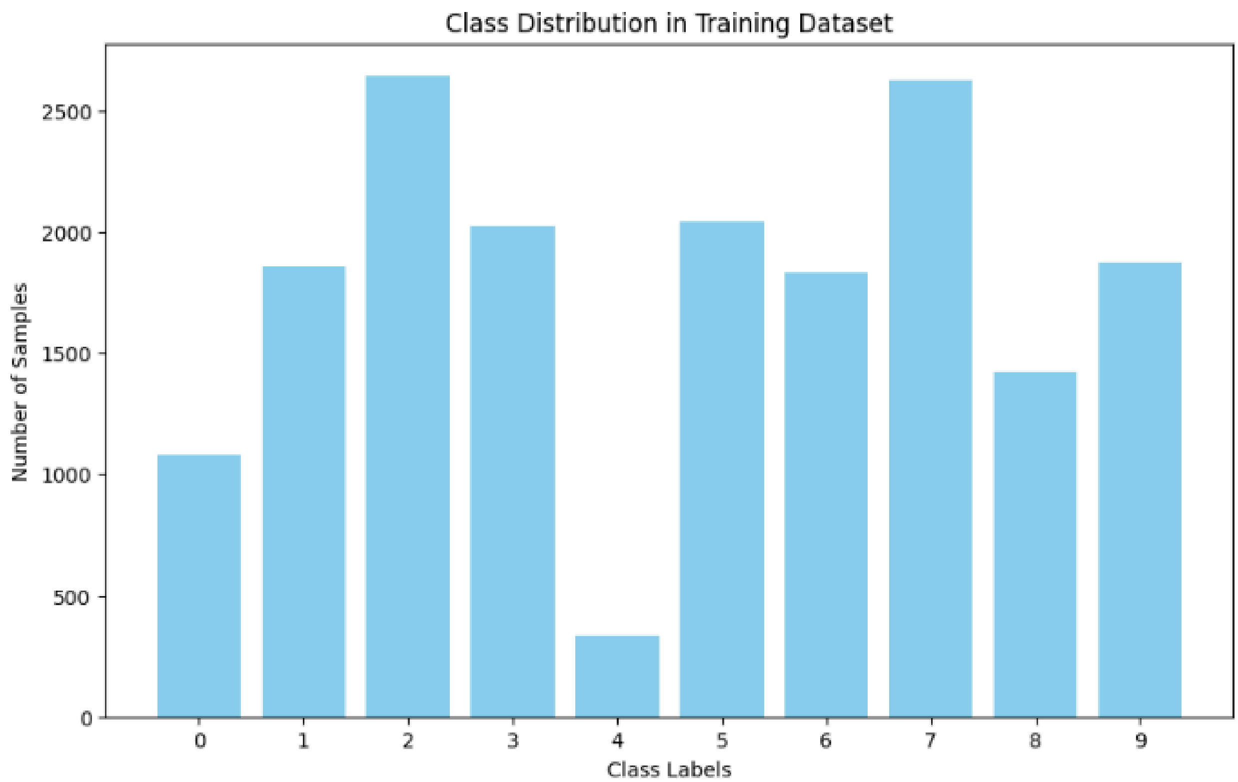


Figure 1

3. Conclusion

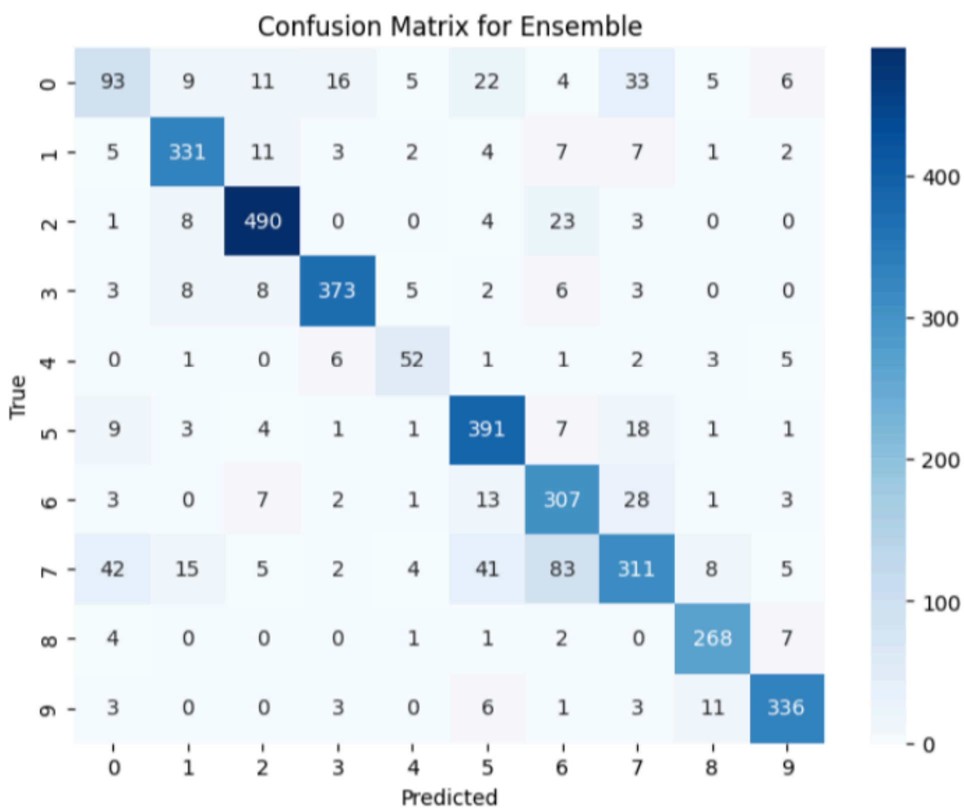
3.1 Individual Model Results

Models DECaLS	Accuracy	Loss
EfficientNetB0	77.28%	0.7621
EfficientNetB3	78.41%	0.7728
EfficientNetB4	81.22%	0.8122
ResNet50	73.80%	0.7380

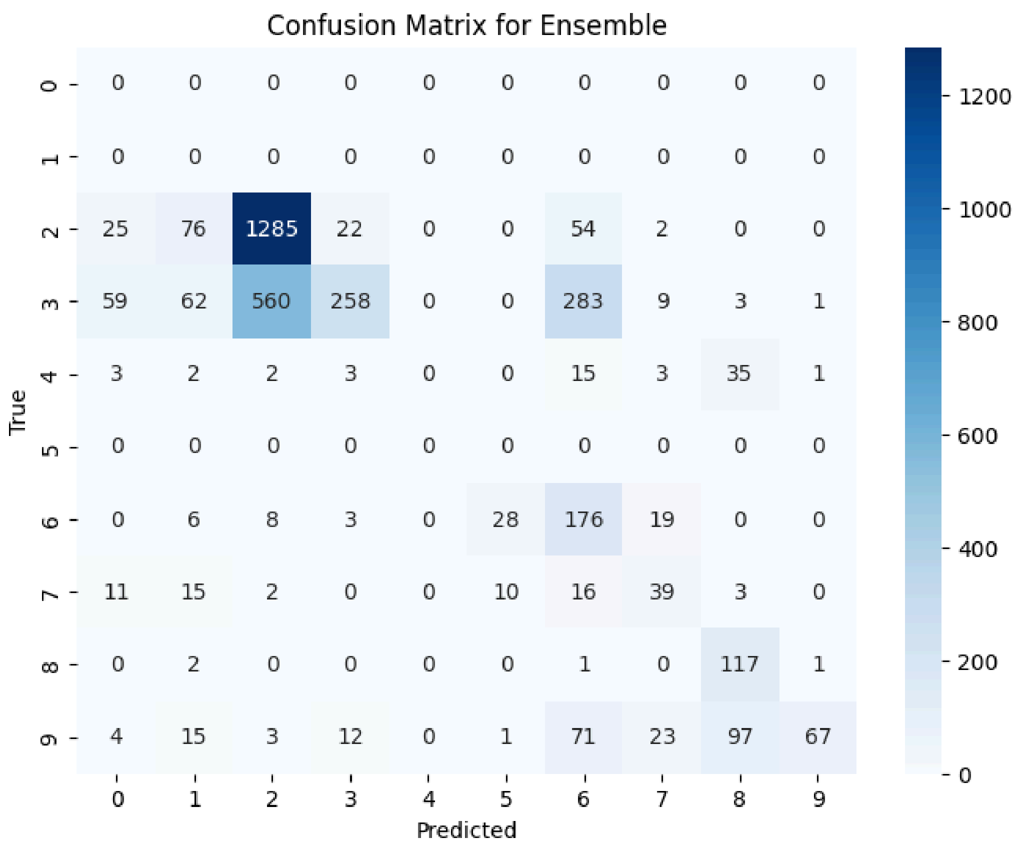
Models SDSS	Accuracy	Loss
EfficientNetB0	49.81%	1.7286
EfficientNetB3	44.37%	2.1693
EfficientNetB4	52.70%	2.0441
ResNet50	44.67%	2.4416

3.2 Ensemble Model Results

Metric	Value
Accuracy	83.2%
Precision	82.97%
Recall	83.2%
F1-Score	82.81%



Metric	Value
Accuracy	55.28%
Precision	71.86%
Recall	55.28%
F1-Score	53.20%



3.3 Predictions Example

3.3.1 Predictions on Decals Dataset

3.3.2 Predictions on SDSS Dataset

3.4 Final Conclusion

The project demonstrated an effective approach to classifying galaxy images using deep learning models with transfer learning techniques. EfficientNet architectures performed the best compared to the other models, with the ensemble model showing improved accuracy and balanced precision, recall, and F1-scores. Class weighting mitigated class imbalance, enhancing model generalization. Evaluation on the SDSS dataset indicated the models' adaptability to new data, though slight performance drops were observed. This highlights the need for further fine-tuning to achieve higher generalization. The results affirm the potential of ensemble methods for galaxy image classification. Our implementation performed better, achieving an accuracy of 83% on the test set, whereas the model provided as a baseline by the library in the GitHub repository achieved only 41% validation accuracy.

4. References

[1] H. Sky, "Galaxy10 Tutorial", *astroNN*, [Online]. Available:
https://github.com/henrysky/astroNN/blob/master/demo_tutorial/galaxy10/Galaxy10_Tutorial.ipynb. [Accessed: 14-Dec-2024].

[2] "Galaxy10", *astroNN*, v1.0.0, [Online]. Available:
https://astronn.readthedocs.io/en/v1.0.0/galaxy10.html. [Accessed: 14-Dec-2024].

[3] "Galaxy10 SDSS", *astroNN*, [Online]. Available:
https://astronn.readthedocs.io/en/latest/galaxy10sdss.html. [Accessed: 14-Dec-2024].