**A Cosmic Zoo: Classification of Galaxies using Transfer Learning**

Trevor Ahlberg

**ABSTRACT**

The key problem being tackled in this paper is that of galaxy classification. Scientists are constantly looking for the most efficient way to classify the myriad of galaxies in the known universe. The proposed solution is to create a convolutional neural network in order to identify the galaxy shape from images taken by and classify it into one of ten classification categories. However, accurately classifying a large data set of galaxies requires not only a lot of computational power, but also a well-trained model to work with. Therefore, one of the main goals of this project is to examine the capabilities of various transfer learning models in the application of galaxy classification. The results propose that the use of the MobileNet transfer-learning model provides the best baseline for designing an effective galaxy classification model.

1**Introduction**

There is much we understand about the universe around us, however there is furthermore that we have yet to discover. An important aspect of understanding that which we have yet to figure out is to gain a thorough foundation of the mechanics that govern what we do understand. When it comes to the universe, a key feature is that of galaxies. Galaxies hold a large wealth of information regarding the origins of our universe, and the forces that maintain it. A clue given to us by galaxies is the formation and shape that the galaxies hold. By observing the shape of galaxies, we can create inferences on the cosmic sources that form them, as well as discern additional secrets that may be hidden by the galaxy’s shape.

The problem being addressed is that of the vast amounts of galaxies needing to be classified. While there are volunteer projects classifying them, using neural networks to automate the classification problem would make processing the vast number of galaxies far faster.

With this, however, is the need for an accurate model. In order to be able to rely on a neural network to classify such a large number of galaxies, the accuracy and precision of the model needs to be high. The apparent solution is to utilize a convolutional neural network for the classification. One such benefit of convolutional neural networks is the plethora of available transfer-learning models available for training. Therefore, a multitude of transfer models are tested against a control model in order to test the effectiveness of utilizing transfer-learning for galaxy shape classification.

The paper is organized as such. Section 2 summarizes the problem being addressed in the paper. Section 3 summarizes the system design and construction of the models. Section 4 covers the methodology of the project, as well as details the quantitative results. Section 5 compares the findings against similar works. Section 6 presents a summary of the results of the paper. Section 7 details the contributing authors of the paper. Lastly, Section 8 provides a reflection of what was learned during the project.

2**Problem Formulation**

The structure of the problem is to take a large sample size of images from the Galaxy10 DECals dataset, generated from the Galaxy Zoo Data Release 2, made available via astroNN, and construct a convolutional neural network model with high accuracy and precision that can classify the galaxy images into the appropriate classification. The galaxy images contained within the dataset were 256x256 RGB images depicting images of galaxies taken from telescope satellites, which are then downscaled to 128x128 for processing. The images are then fed into the MobileNet, ResNet50, and Inception-v3 transfer-learning models, using an independent, non-transfer convolutional neural network as a control with the goal of determining which transfer-learning model is most effective for galaxy classification.

3**System/Algorithm Design**

The goal of the system was to attempt to compare various transfer learning modules and assess their usability for the classification of galaxy structures. The various transfer models tested are MobileNet, ResNet50, and Inception-v3 models. These models are tested against an independent, non-transfer model to assess the power of the model.

The overall design constitutes preprocessing the data into a dataset containing the downsized images, which is then fed into multiple individual models, with the model then compared upon various statistics. The first model to be trained was the control convolutional neural network, before the transfer models were set up and trained on the same data.

3.1**Data Preprocessing**

The data preprocessing involves loading the dataset via astroNN, and then beginning the processing of the data. The labels are one-hot encoded for the classification output, and both the images and labels are casted as floats for easier manipulation. The images and labels are split into train and test sets, and the images are promptly downsized from 256x256 to 128x128 images for easier processing. The newly downsized images are normalized before being fed into the various training models.

3.2**Control Model**

The control model is an independent model constructed using a sequential model with various convolutional and pooling layers. The neuron counts and kernel sizes were tested, and ultimately failed to make meaningful improvement on the model. However, testing the activation function proved that the ReLU function outperformed the other functions on a consistent basis.

The model was compiled using categorical cross entropy with an Adam optimizer. SGD did not provide any improvements to the model over the Adam optimizer. Early stopping was employed during the model fit to prevent overfitting.

3.3**MobileNet Model**

The MobileNet model was utilized for transfer learning to create another sequential model. To create the model, the MobileNet model was downloaded, and the layers were transferred onto the new sequential model, before being made untrainable. The model was flattened and an additional dense layer was added to create outputs. In order to retain a comparison with the control model, the same compiler and fitting settings are utilized as the control model.

3.4**ResNet50 Model**

The next model to be trained is the ResNet50 model. The sequential model was built using the ResNet50 model being added in its entirety, as the format of the transfer model restricted building the sequential model by layer. To make this inclusion, the transfer model was loaded and then made untrainable. The model was then added to the sequential model as one unified layer before flattening the layer and creating a dense output layer. The compiler and fitting functions match the independent model for consistency.

3.5**Inception-v3 Model**

The last transfer model utilized in the project is the Inception-v3 model. The model had similar constraints to the ResNet50 model, as the model was unable to be added by layer. Therefore the sequential model was constructed by first making each layer in the model untrainable, before then adding the model to the sequential model in one layer. With the model added to the sequential model, the model is flattened and a dense layer is added. To be consistent with the previous models, the compiler and fitting functions remain identical.

4**Experimental Evaluation**

4.1**Methodology**

The data used was the Galaxy10 DECals data set, created by the Galaxy Zoo Data Release 2 containing galaxy images from the DESI Legacy Imaging Surveys (DECals). The data set contains 177736 images at 256x256 resolution that are categorically sorted into 10 labels. The given classification outputs for this data set are: disturbed galaxies, merging galaxies, round smooth galaxies, in-between round smooth galaxies, cigar shaped smooth galaxies, barred spiral galaxies, unbarred tight spiral galaxies, unbarred loose spiral galaxies, edge-on galaxies without bulge, and edge-on galaxies with bulge.

During preprocessing, the data was normalized and converted into float32 format for easy processing. The images were additionally downsized to 128x128 resolution to accommodate hardware limitations. The entire dataset was split into train and test sets, with 75% of the data being used to train, and the remaining 25% of the data used for the tests.

The key experimental aspect was the viability of the various transfer learning models. Testing the effectiveness of various transfer models allowed for identifying the most effective model for the task. The points of comparison for the various models are the accuracy, precision, and f1-scores of each model. These metrics give the best insight as to the performance of the models on the dataset, and are a good indicator as to which transfer model can perform best at the task.

To conduct the experiment, a base, independent convolutional neural network was created and trained on the dataset. This provided a control for the transfer models to compare against. The transfer models used in this experiment were the MobileNet, ResNet50, and Inception-v3 models. The models were all trained on the same data, and utilized the same compiler and fitting functions. The resulting metrics of each model were calculated for comparison of each model.

4.2**Results**

The results provided a clear comparison of the effectiveness of the various models in the context of the problem. There was a distinct difference between each model’s ability to handle the data set, and the accuracy of the resulting data.

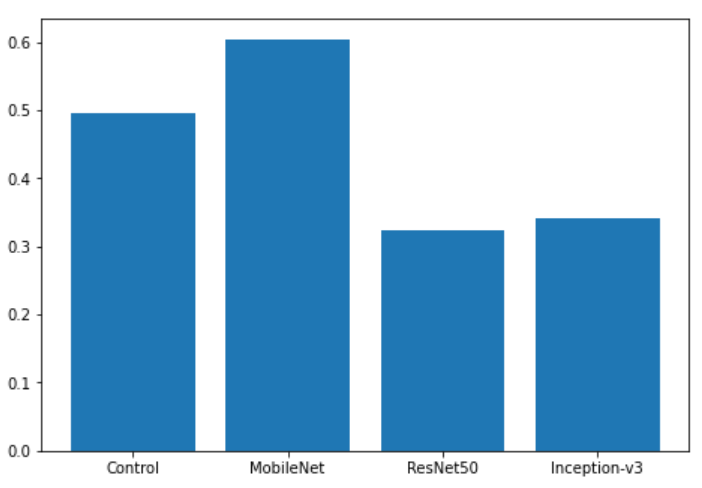


Figure 1: **F1-scores of the models tested.**

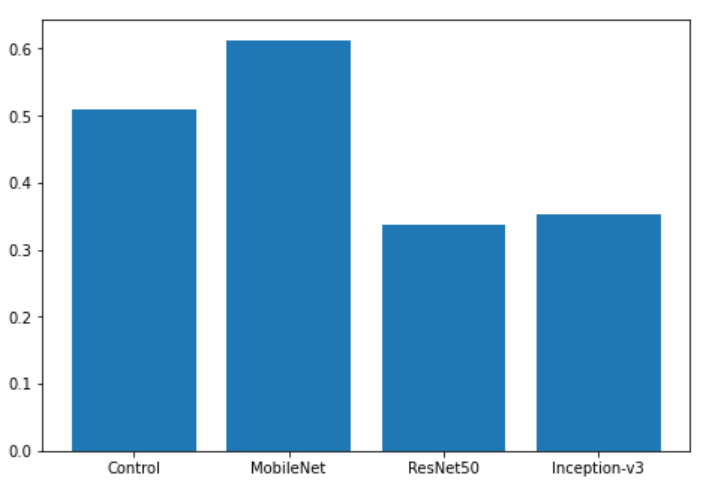


Figure 2: **Accuracy scores of the models tested.**

The demonstrated graphs (figure 1, figure 2) show that, while the unaugmented models were unable to reach adequate numbers to provide consistent galaxy labeling, there was an obvious difference between the capabilities of various models. The control model was only able to achieve an average and f1-score of approximately 0.5.

The ResNet50 and Inception-v3 models did not perform well against their competition Both models were below 0.5 in both accuracy and f1-scores, which was far below the control model’s metrics.

The MobileNet model performed admirably, being the only model to exceed an average and f1-score of 0.6, it was the best performance of all the models tested. The lightweight structure also made it ideal to work with, as it was not intensive on resources during training or testing in comparison to the other models.

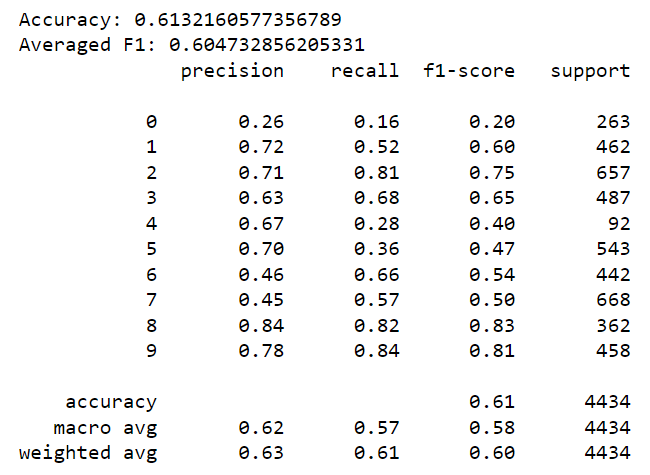


Figure 3: **Output table of metrics for the MobileNet model.**

The resulting metrics of the MobileNet model (figure 3) gives some insight as to the performance of the model. It is seen that the model excelled with some galaxy structures, but struggled with others.

5**Related Work**

The related works took similar approaches to the same problem, however different tools were utilized in each case.

In the first related work, “Morphological Classification Using Machine Learning” by S. Kasivajhula et al., the paper utilizes alternative machine learning algorithms in order to achieve the given results. The aim of the paper is similar, to compare different models for use in galaxy classification, utilizing Support Vector Machines, Random Forests, and Naïve Bayes as the comparison [1]. This differs as they utilized distinctly different machine learning algorithms to achieve their results, rather than using the same technology, but comparing various models. This paper seeks to improve upon this baseline by building upon the potent tool that is convolutional neural networks, and its ability to more easily decipher noise.

Another related work was “Deep Galaxy: Classification of Galaxies based on Deep Convolutional Neural Networks.” by N. Khalifa et al., which took a more similar approach. The paper utilizes the power of convolutional neural networks to create a highly accurate method of determining galaxy morphology. The paper gives a detailed overview of the model architecture and implementation [2]. While working on the same baseline technology, their work simply presents a model as a solution to the issue of galaxy classification, rather than comparing various potential tools. Comparing various approaches to the problem can help to further optimize existing tools, allowing for more accurate results. Additionally, transfer-learning acts as a fantastic baseline expanding a model to fit a certain dataset. By finding a transfer-model that works for the given dataset, that decreases the number of layers that need to be trained and tuned for the dataset.

6**Conclusion**

The data gave a clear indication that the MobileNet transfer-learning model provided the best performance for the classification of galaxy shapes. The model serves as a good baseline to expand upon for a perfected model.

Both the ResNet50 and Inception-v3 models performed poorly, but they did have the similar issues of not being able to be added to the model by each layer, which could potentially have impacted the model’s performance.

For future work, the overall data preprocessing system could use drastic improvement. Layers could be utilized to clean the images of extraneous noise, such as bright stars and other cosmic disruptions that the convolutional neural networks might struggle to decipher [1]. Additionally the galaxies could be reorientated to be parallel across the same axis, removing galaxy tilt as a factor [2]. Eliminating these factors can reduce confusion in identifying consistent patterns in the galaxy morphologies.

While the preprocessing had a large impact on data accuracy, the resulting data on the effectiveness of various transfer-learning holds true. Using the proposed solutions, and the basis of the MobileNet model, a highly accurate model could be developed.

7**Work Division**

The paper has one sole author and contributor, Trevor Ahlberg.

8**Learning Experience**

There was a lot to learn from this project. The limitations of hardware capability and processing time became very apparent early in the project. Some factors of the project had to be limited due to this technological constraint. The importance of proper image processing was also apparent in the effect it had on the performance of the models.

There is also the aspect of

**REFERENCES**

[1] Kasivajhula S., Raghavan N., Hemal S., “Morphological Classification Using Machine Learning.” 2007

[2] Khalifa N., Taha M., Hassanien A., Selim I., “Deep Galaxy: Classification of Galaxies based on Deep Convolutional Neural Networks.” 2017