12-15-2022

Building a Deep Model for Multi-class Coral Species Discrimination

Hyeong Gyu Jang
Grand Valley State University

Building a Deep Model for Multi-class Coral Species Discrimination

Hyeong Gyu Jang

A Project Submitted to

GRAND VALLEY STATE UNIVERSITY

In

Partial Fulfillment of the Requirements

For the Degree of

Master of Science in Applied Computer Science

School of Computing

December 2022

The signatures of the individuals below indicate that they have read and approved the project of Hyeong Gyu Jang in partial fulfillment of the requirements for the degree of Master of Science in Applied Computer Science.

Jonathan Leidig, Project Advisor Date

Gregory Wolffe, Project Advisor Date

D. Robert Adams, Graduate Program Director Date

Paul Leidig, Unit head Date

Abstract

The goal of this qualitative research project is to develop and optimize a multi-class discrimination model to identify different species of coral based on their digital images. Currently, there are artificial intelligence (AI) models that can distinguish between coral and other undersea objects such as sand or rocks, but to our knowledge the problem of multi-species classification has not yet been addressed. Given that coral reefs are a good indicator of overall ocean health, it is important to develop models that can classify the presence of different species in underwater images as a way to monitor the effects of climate change.

The dataset for this project consists of images of various species of coral; collected from the reef regions of offshore Florida and Bonaire, sanitized, labeled, and organized according to species. This study explores multiple options for image pre-processing, compares different model architectures, and experiments with hyperparameters such as learning rate with a goal of developing the most accurate coral species classifier. Our preliminary results: using only a portion of the complete dataset, a multi-class coral species classifier was produced that achieves 92.2% accuracy.

Introduction and Background

Within climate studies, the state of coral reef habitats serves as a good indicator of overall marine ecosystem health. To estimate the oceanic health status near coastal regions where coral reefs are formed, it is helpful to monitor the characteristics of coral habitats. To date, an automated monitoring agent has not been developed meaning that the only way to monitor coral reef health is to use human divers to manually identify species and check the state of the reef. Typical coral reefs extend many square kilometers along coastal regions. Given the size of the reefs and the limitations and time involved in using human divers, only a small portion of those reefs can be examined. Data collection becomes one bottleneck in monitoring coral reef health.

Clearly, there is a need for automated data collection and monitoring. Recent technological advancements suggest building machine-learning classifiers to fulfill this need. Researchers have built prior machine learning (ML) models—given a top-down view image of a portion of reef—that can distinguish coral from its surroundings (2). The classifier uses a Deep Learning Convolutional Neural Network (CNN) that attains ~ 97% accuracy in predicting if an object is a coral or not. The model is trained using a coral reef image dataset obtained from various reefs around the world.

Progressing one step beyond the Deep Neural Network (DNN) that distinguishes a coral object from its surroundings, this study introduces and optimizes another Deep Neural Network model, leveraging existing pre-trained image networks, that can classify specific species of coral when presented with an image. The image data used in this study is collected exclusively in the coral reefs along the coasts of Florida and Bonaire (3).

When creating an image classification model, it is common practice to leverage existing or pre-trained Deep Neural Network models. This study utilized several pre-built models during training in an effort to maximize prediction accuracy.

Among the many pre-built DNN model candidates, the VGG and ResNet network families are compared in this study. Using VGG and ResNet models is simple. The model inputs a 224 * 224 * 3 image matrix (i.e., small images with RGB color channels), and through a combination of convolution and other neural network layers, the model builds an internal representation of the image. Then, using a process called transfer learning, the last layer of the model is customized in such a way that the number of neurons matches the number of classes. In the simple case of determining if a single species is present, a binary classifier is built, in which the final layer of the model is configured with two neurons each indicating true and false. For the final layer of the multi-class classifier, eleven neurons are configured as there are a total of eleven species of coral found in the dataset.

In this study, the ML model for classifying coral species is optimized in the following areas:

1. Training data variety
2. Choice of network model
3. Learning rate selection.

Methods

2.1 Image Dataset Variation

The image data in this study is collected from coral reefs. The complete dataset contains 32 species of coral with 10,000 images of each species. Each image contains only one species, which typically fills the image. To examine the effect of background noise on prediction accuracy, the performance of the model was evaluated when trained with the dataset containing background noise (original) and without background noise (cropped).

For each dataset experiment, a binary classifier for each species is created by extending the ResNet-50 model. The training set and the test set are created by partitioning the data using a 7:3 ratio. The models are then trained over 20 epochs. After each epoch, species prediction is performed on data from the test set, and the overall prediction accuracy is calculated and collected. A set of 20-epoch training/testing experiments is repeated 10 times and averaged, so that the prediction accuracy values are plotted as the average of all 10 experiments.

2.2 Model Selection

The VGG and ResNet models all vary in terms of their depth. Within the VGG family, the VGG-11, VGG-16, and VGG-19 models are extended to build the DNNs used in this study. Similarly, the ResNet family includes ResNet-18, ResNet-32, and ResNet-50. To determine the best model for this classification problem, the prediction accuracy and runtime are compared among all models, which are trained and tested using the noise-free (i.e. cropped image) dataset. Also, to account for any possible disparity in prediction performance for different coral, a binary classification model is built for each species. As in the dataset selection experiments, a 7:3 training/test set split ratio is used. Each model is trained over 20 epochs and model accuracy is computed after each epoch of training. The 20-epoch training/testing process is repeated 10 times and averaged to produce weighted prediction accuracy values.

2.3 Learning Rate Selection

To optimize the training process for the coral species classifier, hyperparameters are also customized. Among the various network hyperparameters, learning rate is one of the most crucial factors in model training optimization. Therefore, an optimal learning rate is selected by comparing prediction accuracies. As in the previous selection experiments, the dataset used for all learning rate experiments is the cropped dataset, split into a 7:3 training/test set ratio. The network model used is ResNet-50 with varying learning rates, trained and tested over 20 epochs with the prediction accuracy value recorded after each epoch. Unlike the two previous selection processes, the prediction accuracy values are not averaged and the models created are full multi-class classifiers.

Results/Discussion

3.1 Dataset Selection

To determine if the presence of a small amount of background noise in the training dataset might guard against overfitting and hence result in an overall better classifier, the prediction accuracy values for each model built for each of the 11 species of coral is collected and plotted. See Fig. 3.1 for the binary discrimination model accuracy comparisons in a set of plots.

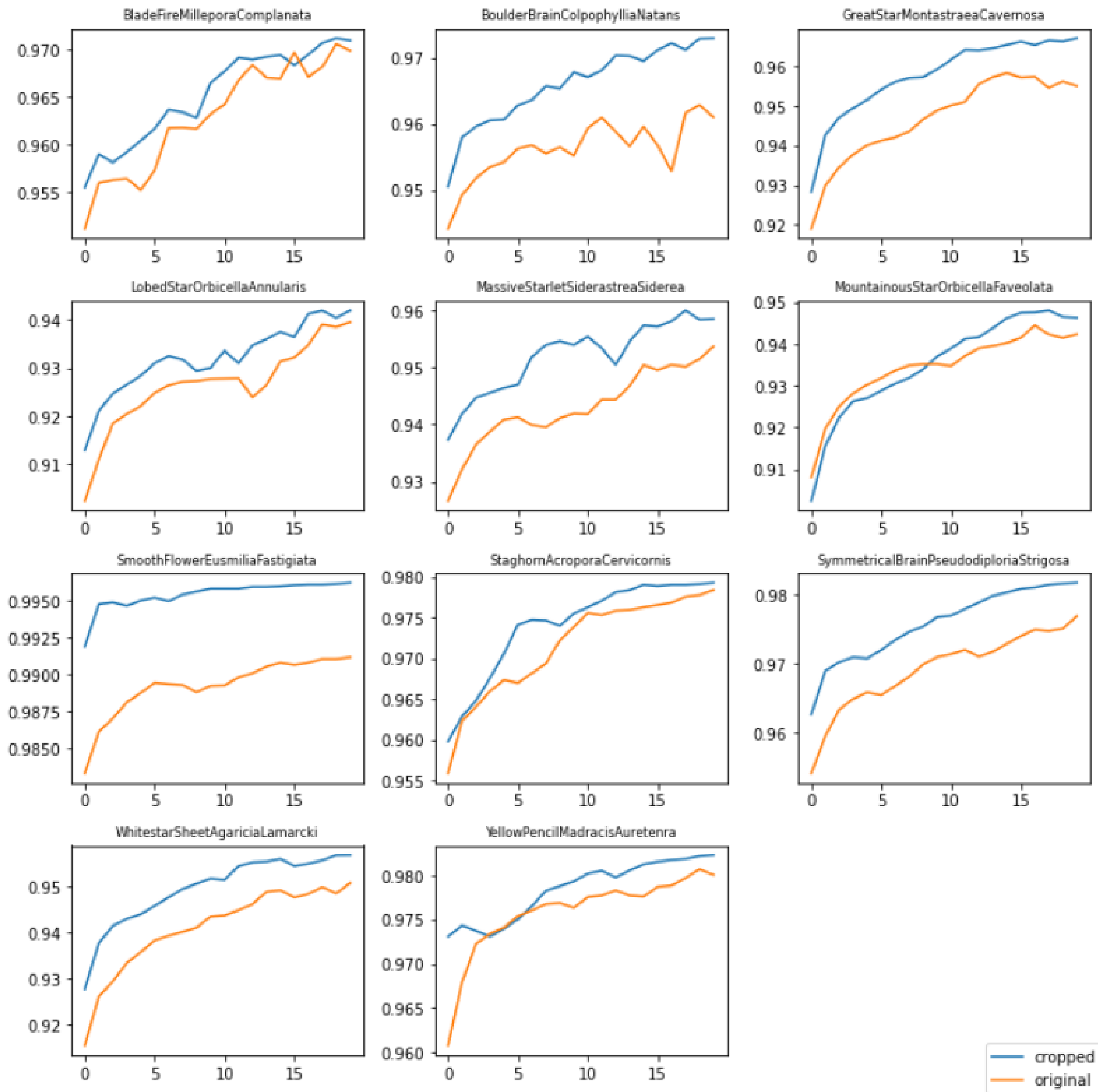


Figure 3.1 Model Prediction Accuracy Comparison for Dataset Selection

Each graph tracks the change in model prediction accuracy over the course of 20 epochs of training. Generally, the prediction accuracies acquired are excellent, where for Smooth Flower binary discrimination model, it could predict as accurately as with 99.7% accuracy. It should be noted that accuracy variation trends differ from one species to another, as the inherent shape of a coral species has an inevitable impact on model performance. A species of coral that does not entirely fill the full frame of an image might not be present in the pixels on the perimeter of the photograph. See Fig. 3.2 For a pair of images, one containing the background noise, and one without. The darker blue line in each plot gives the prediction accuracy trend for the model trained using images without background noises (cropped), whereas the lighter orange line indicates the prediction accuracy trend for the model trained with the original images including noise (original).

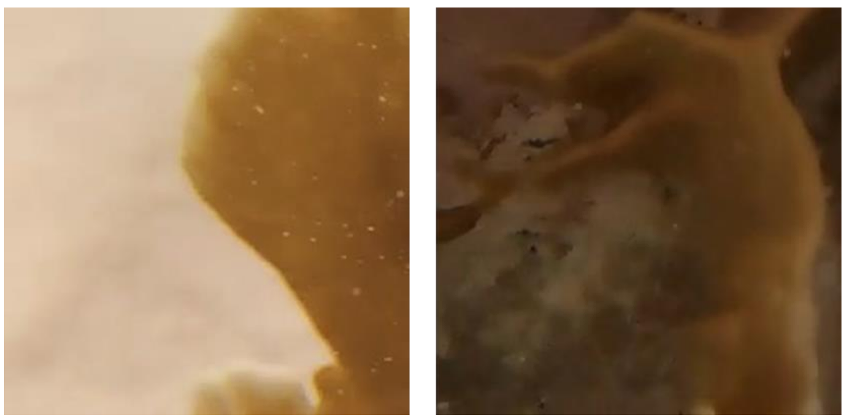


Fig 3.2 Image containing background noise (Left) and Image without background noise (Right)

It is apparent that the performance of the models trained using cropped images exceeds those of the models trained with the original images. Although in some cases (e.g. Yellow Pencil coral) the performance of the models trained with the original images displays a higher accuracy in the earlier epochs, the models trained with the cropped image set surpass the other collection in later epochs. Not all trend lines plateau, which raises the question that the models might not have fully converged, but it is still quite evident that models trained with the cropped images perform best. Therefore, for this study, it can be concluded that it is desirable to eliminate as much background noise as possible for the training set.

3.2 Model Selection

To find out which pre-built DNN model fits the data best for the problem of coral species classification, the prediction accuracy values of a binary classification model built for each species of coral are collected after each epoch of training. The below plot is generated by plotting these prediction accuracy values.

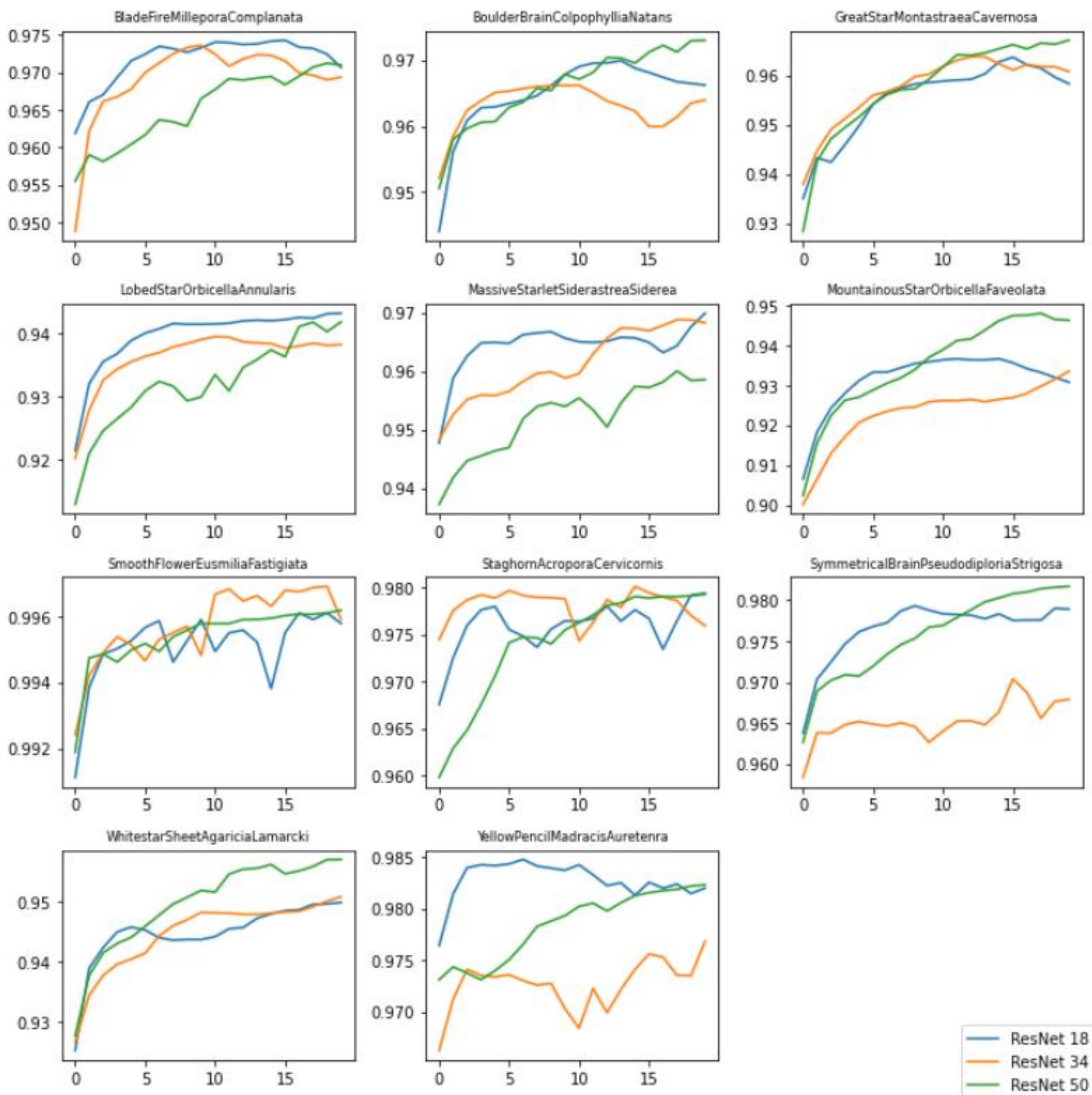


Figure 3.3 Prediction Accuracy Comparison for Model Selection

As in the dataset selection process, each subplot describes the change in prediction accuracy over 20 epochs for each coral species. The plots only display results from the ResNet DNN models, because the prediction accuracy values acquired from VGG DNN models are all significantly lower, generally falling between 80% and 90% prediction accuracies.

The conclusion here is not as evident as that from the dataset selection process. There is no single outstanding model that outperforms the other models for all species. However, among the three models in the ResNet family, ResNet-50 generally achieves a higher prediction accuracy value.

In 9 of the 11 species' binary classifier models, ResNet-50 produced the highest prediction accuracy. One of the exceptions can be seen in the Lobed Star coral species, where ResNet-18 ranks the highest in terms of accuracy value. However, the prediction accuracy of the ResNet-18 model plateaus at about the fifth epoch, whereas the prediction accuracy trend of ResNet-50 continues to increase, suggesting that it may plateau at a higher level than ResNet-18. Regardless, 95% accuracy is still sufficient performance for species specific reef monitoring purposes. The other exception is with the Massive Starlet coral. For this species' binary classification, ResNet-50 would not be desirable because it is consistently outperformed by ResNet-18. Ultimately, given the overall superior performance of ResNet-50 on the binary classification problem, it is leveraged to build the multi-class classification model. Our conclusion is that species-specific model selection is required for optimal binary classification results, with ResNet-50 serving as a baseline model for all species.

3.3 Learning Rate Selection

For this experiment a set of arbitrary learning rate values are selected and tested. The below plot shows the overall trend of prediction accuracies over 20 epochs for all models.

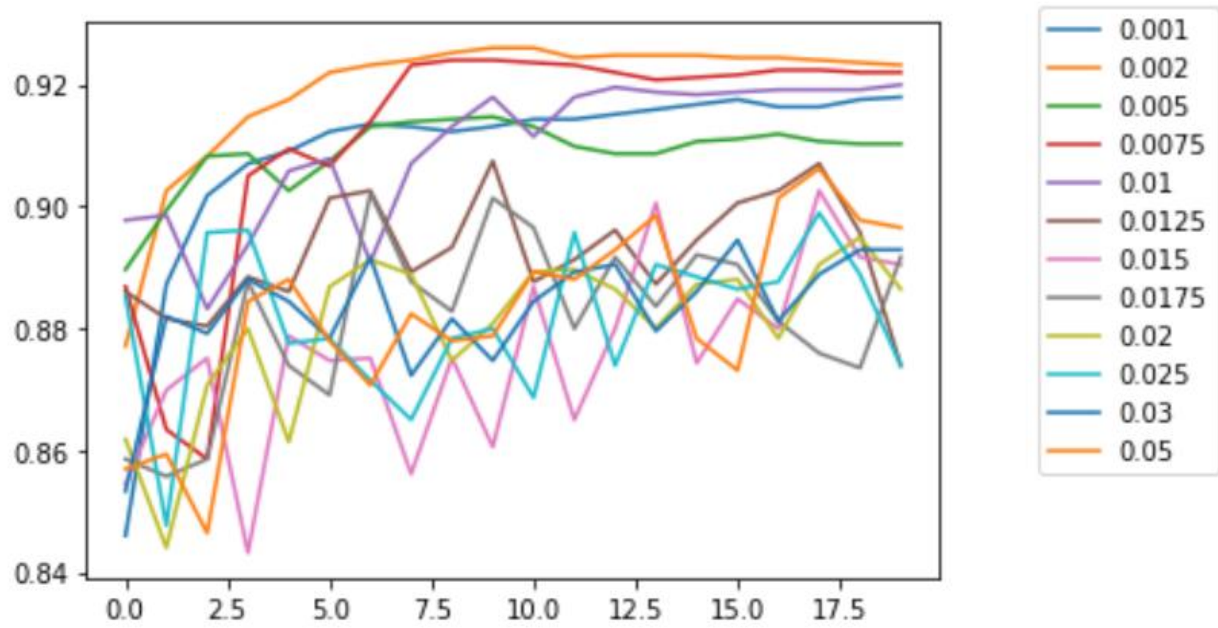


Figure 3.4 Prediction Accuracy Trend for Learning Rate Selection

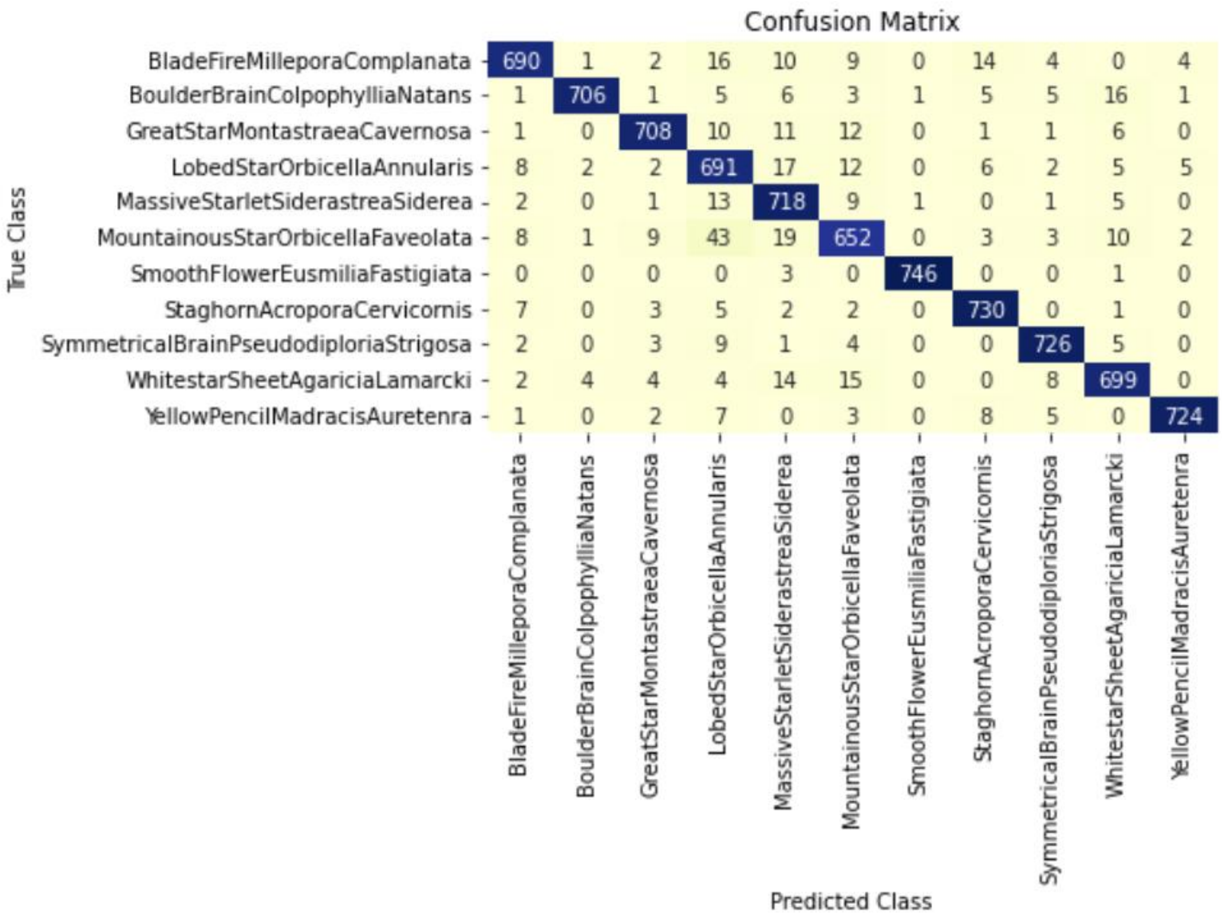
The accuracy of the learning rates ranging from 0.001 to 0.01 seem to converge within 20 epochs, whereas anything higher than 0.01 is not converged and is of lower accuracy. For the current model 0.0075, as opposed to 0.002, is selected as the learning rate used for the multi-class classification model, to maximize the model to prevent the convergence getting stuck at the local minima.

3.4 Multi-Class Classification Model

Based on the selections suggested by the experiments described above, a multi-class classification model is created. This model uses transfer learning to extend ResNet-50, is trained and tested using the cropped image dataset without background noise, and uses 0.0075 as its learning rate. The dataset is split using a 7:3 ratio for training and testing, respectively, and the model is trained over 20 epochs. The final prediction accuracy reported after the twentieth epoch is 92.2% for the multi-class identification problem.

As a further investigation, a confusion matrix is created to verify that the model is working as expected. A confusion matrix can be used to see if the classifier often misclassifies a coral species as another species that looks very similar.

Figure 3.5 Confusion Matrix for Multi-Class Classification Model



Each row represents the true species of the data, and each column represents the predicted species as generated by the classification model. Since the prediction accuracy is very high, most of the prediction instances match their true class. Among incorrectly predicted instances, the Mountainous Star coral species are often misclassified as Lobed Star. This is an expected result as these two species belong to the same genus and are quite similar in visual characteristics and growth structure. See Fig 3.6 For visual similarities between Mountainous Star and Lobed Star coral species. This provides some validation that the classifier is working as expected.

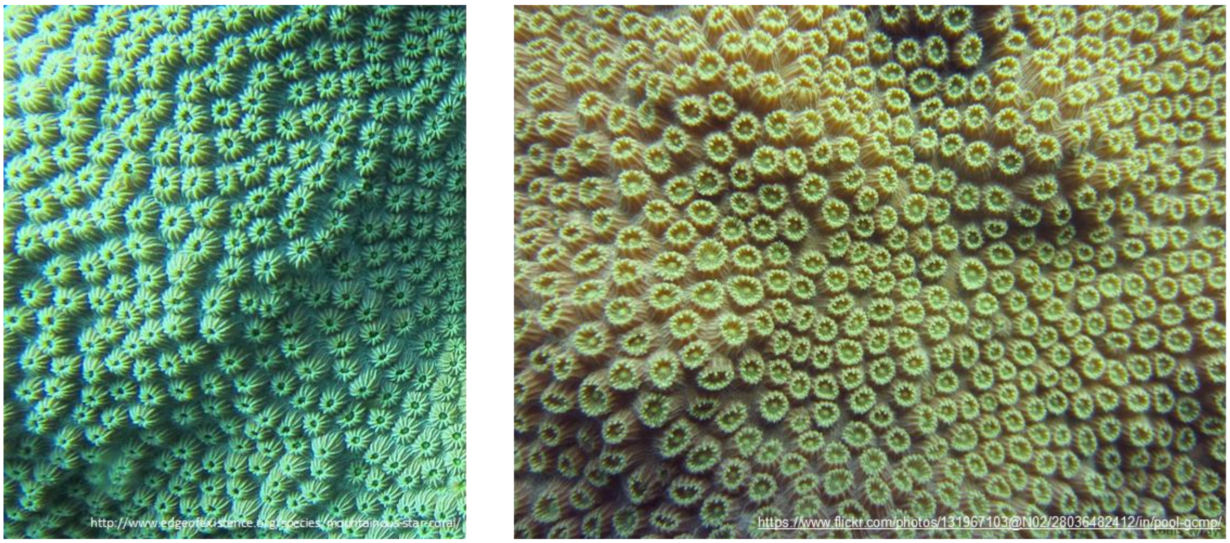


Fig. 3.6 Image of Mountainous Star and Lobed Star

Conclusion

This study describes foundational work on building a DNN coral species classifier using a newly collected image dataset. To optimize the classifier, several experiments are conducted. A dataset pre-processing experiment explores whether it is beneficial to include background noise in the model training. It is concluded that it is better to use the cropped (noise-free) dataset to train the classifier. Leveraging the availability of pre-built DNN models combined with knowledge transfer, it is concluded to extend ResNet-50 to build the coral species classifier for most species. Lastly, a range of learning rates is tested to find an optimal hyperparameter setting to be applied to the coral classifier training, and it is found that values between 0.001 and 0.01 reasonably converge within 20 epochs.

Based on these experiments and conclusions, an initial multi-class model is created—trained and tested with the cropped image dataset, extended from ResNet-50, and using a learning rate of 0.0075. This classifier is able to accurately predict the species of a coral image from the test dataset with 92.2% accuracy.

There are several possible directions for future work. Since this study represented an exploratory attempt, only 1/10 of the complete coral dataset is used to train and test the models. Provided with a more voluminous dataset, the coral classifier's performance may be improved. Furthermore, when the coral species classifier is in a more mature state, it could be combined with object detection models (2) to become more efficient in field applications covering very large areas.

Bibliography

1. Yuval, et al. Repeatable Semantic Reef-Mapping through Photogrammetry and Label-Augmentation. Remote Sens. 2021, 13, 659.
2. Leidig, J. Coral Reef Image Collections for Machine Learning, Mapping, and Monitoring. IEEE OCEANS, 2022.
3. Source: He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.