RESULTS and PLOTS

Performance metrics and heat map for confusion matrix are as follows for each model used.

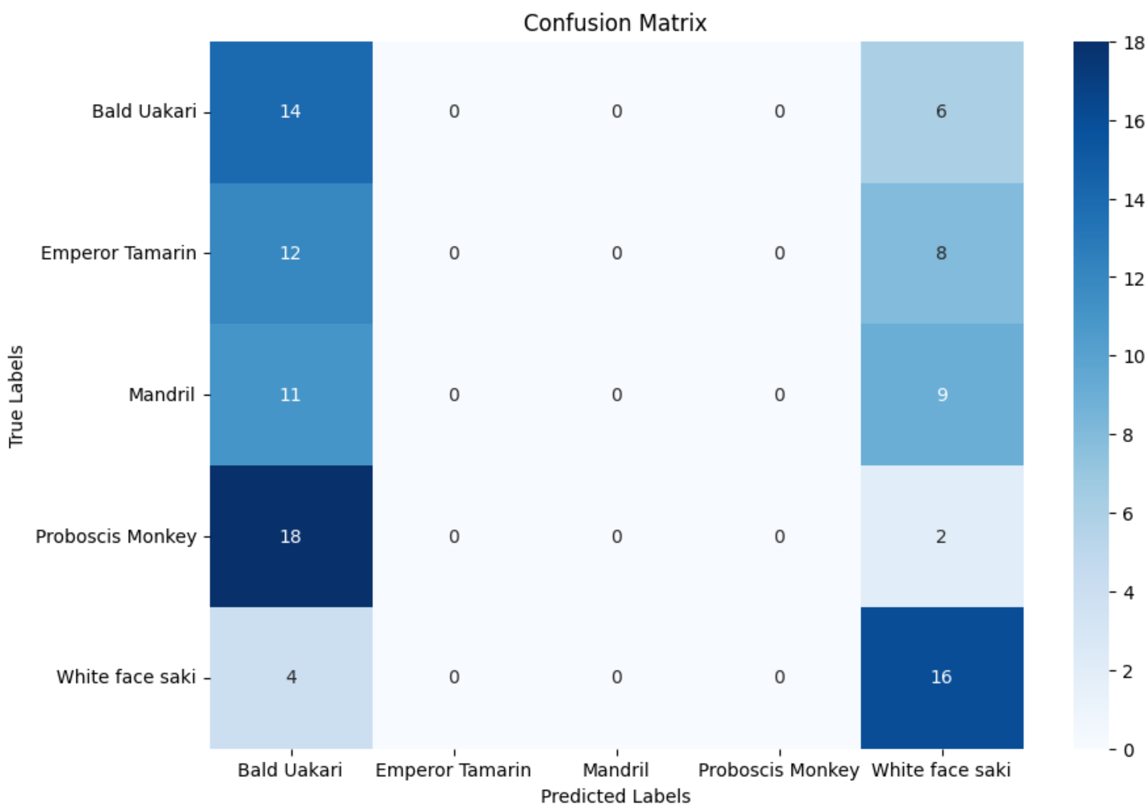
Custom CNN Results

Accuracy: 0.3

Macro Precision: 0.12550640760644896

Macro Recall: 0.3

F1 Score: 0.17580410873625235



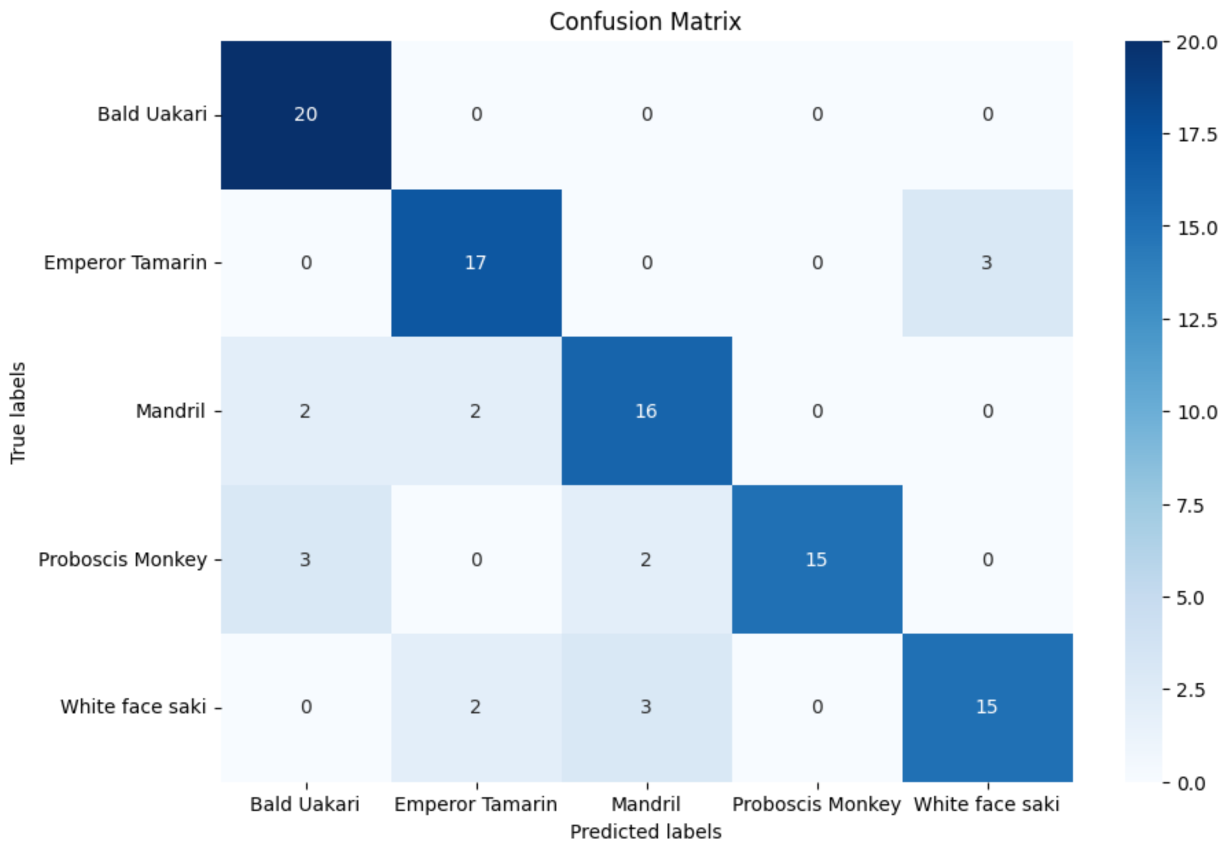
Adam Optimizer Results

Accuracy: 0.83

Macro Precision: 0.840952380952381

Macro Recall: 0.8300000000000001

F1 Score: 0.8290523055606496



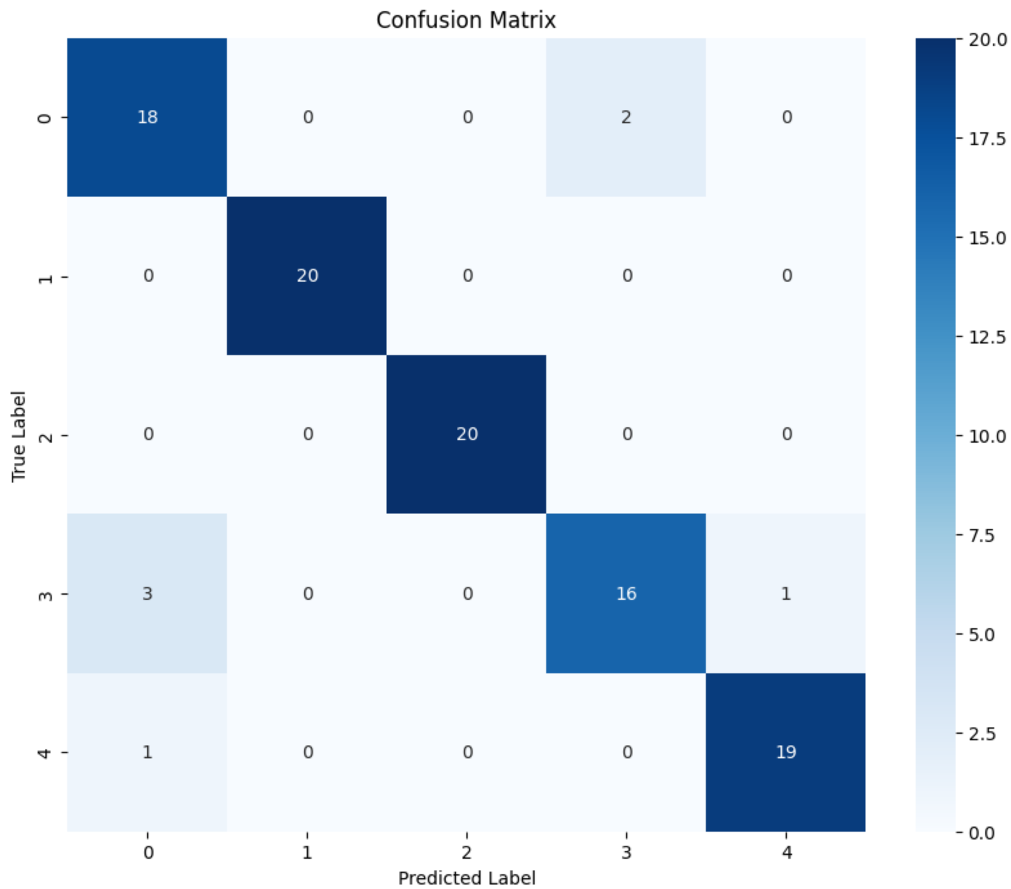
ResNet-18 Results

Accuracy: 0.9300

Macro Precision: 0.9314

Macro Recall: 0.9300

F1 Score: 0.9298



Conclusion

The other two optimization algorithms, AdamW and SGD, do much worse in comparison to the pretrained ResNet model. The ResNet model gives excellent scores in terms of accuracy, precision, recall, and F1 score: 93%, 93.14% precision, 93% recall, and 92.98% F1 score. The results from AdamW were very good: an accuracy of 83% and an F1 score of 82.91%; thus, the ResNet improved the results by about 10% in both cases. The performance metrics for SGD indicate that they are very low: 30% in accuracy and 17.58% in F1 score.

The outperformance of ResNet could be attributed to its deep architecture, which captures complex features and patterns in data very well, with a gain from transfer learning. It has a big number of layers, with pre-trained parameters on a large, diverse dataset, making it capable for fine-tuning and really beneficial for tasks other than simple classification.

Some of the benefits of using ResNet are faster convergence and learned weights that require fewer modifications to fit new tasks, hence better accuracy and generalization. Hence, ResNet becomes very useful in complex datasets where performance is driven by capturing intricate patterns with high accuracy and generalization. Using such a sophisticated model as ResNet really multiplies the benefits given to achieving better model predictions and, therefore, improving the reliability in practical use.