Chicken Disease Image Classification

Coursework for C7082 - Techniques in Machine Learning and AI

Introduction

A repo containing my coursework for C7082 module of the MSc in Data Science for global agriculture, food and the environment at Harper Adams University.

The github can be found at https://github.com/joebmitchell/C7082 and the final model can be found at https://www.dropbox.com/s/t49w8rr9jm5a36d/Models.zip?dl=0.

Background

Poultry production in the UK provides a large proportion of the UK animal protein, but commercial production is based on very fine margins and so inefficiencies must be mimisimised in order to maximise profitability. Three diseases that can cause changes in the faeces of domestic chickens are Newcastle disease, Salmonella and Coccidiosis. The aim of this project is to create a CNN that can accurately predict based on a photo of chicken faeces whether they are infected with one of these three diseases or healthy.

Data

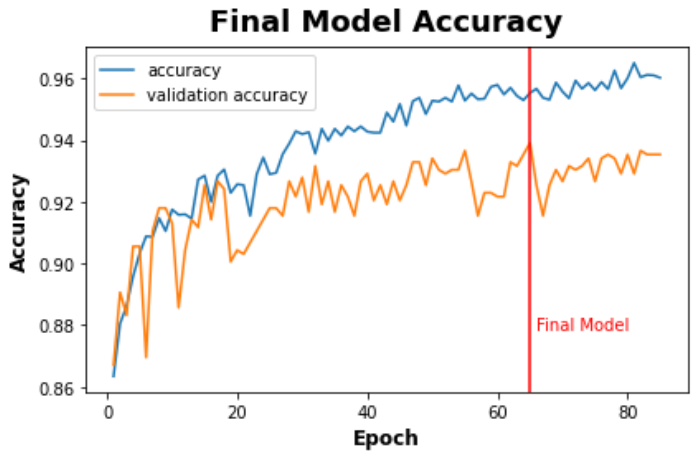
The data for this project was sourced from Kaggle and the images can be found at https://www.kaggle.com/datasets/allandclive/chicken-disease-1. It consists of 8067 jpg files that were taken in Tanzania. The data was split in an 80:10:10 ratio into train, test and validate datasets. Examples of the images used for training and their classification are included below.

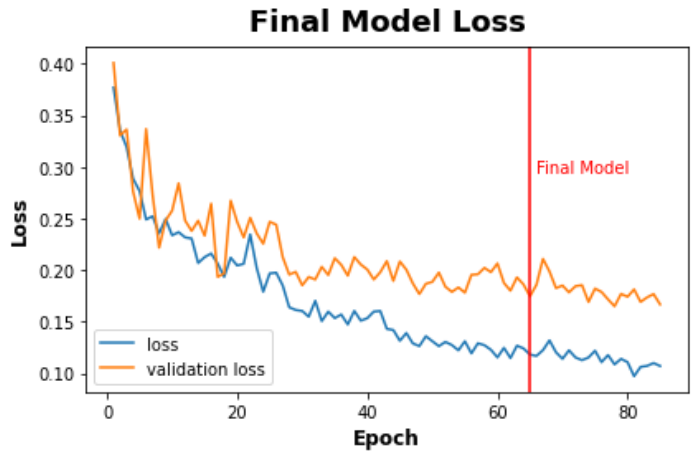
Model

Various model architectures were trialed with augmentation, callbacks, different bases for transfer learning, different heads and batch size varied in a stepwise manor in order to find the model with the best validation accuracy.

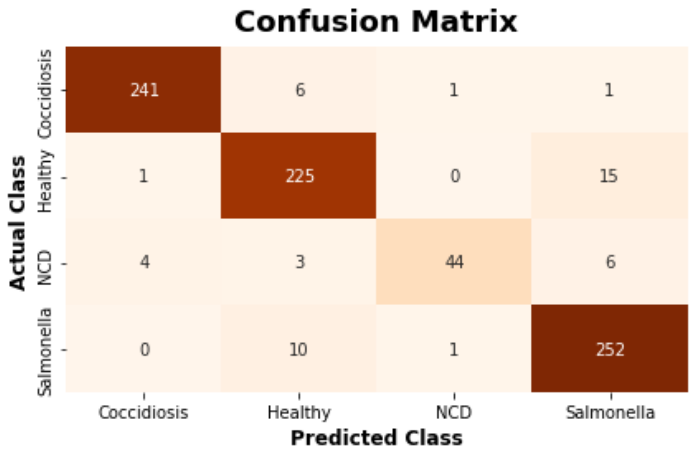
Results

The final model selected has a validation accuracy of 0.939 and a validation loss of 0.165. The training data for this model is shown below highlights the epoch at which the final model was selected





In order to assess its ability to predict the class on unseen data the model was evaluated using the model.evaluate function on the test dataset which showed an accuracy of 0.9407 and a loss of 0.1898. A confusion matrix was created using the models predictions of the test dataset in order to visualise differences between categories in prediction accuracy



Discussion

The final model has reasonable accuracy on the images in this dataset and could be used as a screening test for some clinical diseases of chickens although further real world validation would be required before this could be relied on.