Keras Assignment Report

Aaron Modiyil Joseph
22018497

University of Hertfordshire
17 January 2024

Introduction

This report analyses and critically evaluates a convolutional neural network (CNN) model for classifying animal images into three categories: cat, dog, or wild. We discuss the model's strengths and weaknesses using sufficient plots and images.

Dataset Overview

The dataset has JPG images of animals: cat, dog, or wild. It has two subsets: train and validation. The train subset has 14631 images (5154 cat, 4739 dog, 4738 wild). The validation subset has 1500 images (500 for each class). The images are in folders with numbers for class: 0 for cat, 1 for dog, and 2 for wild.

Data Processing

Data is loaded from Google Drive with tf.data.Dataset.list_files. Two functions are defined to read and process images and labels, convert to tf.float32, and resize to (64,64). Two datasets, train_ds and valid_ds, are created and batched, cached, and pre-fetched for efficient data loading and model training. The order follows TensorFlow's best practices for performance optimization.

Model Definition

Using Keras functional API, a CNN model for image classification is defined. It has an input layer, six convolutional layers, two max-pooling layers, a global average pooling layer, two dense layers, and an output layer with softmax activation function. The model is compiled with Adam optimizer (learning rate of $5e^{-5}$), sparse categorical cross-entropy loss function, and accuracy metric.

Model Training

The model is trained for 100 epochs on the processed datasets and uses callbacks to save the best weights and change the learning rate. The loss and accuracy and the learning rate are recorded for each epoch. The ModelCheckpoint callback saves the weights with the highest validation accuracy. The LearningRateScheduler callback reduces the learning rate exponentially by 0.01 after 10 epochs.

Model Evaluation

The model demonstrates exceptional performance, as depicted in Figure 1 with both training and validation losses registering very low values. Furthermore, the training and validation accuracy are observed to be proximate to 1 or 100%, indicating a high degree of precision in the model's predictions. However, a potential issue of overfitting is discernible. This is substantiated by the oscillations in validation loss and accuracy observed after approximately 80 epochs.

In the analysis of the confusion matrix heat map depicted in Figure 2, it is again evident that the model exhibits a commendable performance. The model demonstrates the highest accuracy in predicting cats, with the least number of misclassification as well. This suggests that the model has effectively learned the features of cats and seldom fails to identify these features in an image of a cat. This could be due to the larger number of cat images present in the training set. Despite having lower diagonal values for dogs and wild animals than cats, the model accurately predicts these categories. However, it is worth noting that the bottom row has the most misclassifications indicating that the model has the most difficulty predicting wild animals, frequently mistaking them for dogs. Surprisingly, the model confuses images of dogs as cats or wild animals nearly equally. To improve performance, class distinction requires further research and model refinement.

Upon close inspection of Figure 3, which presents the first 24 misclassified images from the validation dataset, several observations can be made. Wild animals such as wolves, which exhibit facial features similar to those of dogs, are misclassified as dogs. Similarly, certain breeds of dogs that bear a strong resemblance to wild animals, particularly wolves, are misclassified as wild.

In addition, images of wild animals like lions and foxes, which possess facial features similar to cats, are misclassified as cats. Moreover, dogs that are particularly furry and exhibit facial features that are similar to cats are misclassified as cats. Furthermore, cats that are brown or of a similar hue, which could potentially be confused with lions or foxes, are misclassified as wild.

In light of these observations, it becomes apparent that the model faces challenges in correctly classifying certain categories of images. The model

Figure 1: Loss and accuracy plot of the model

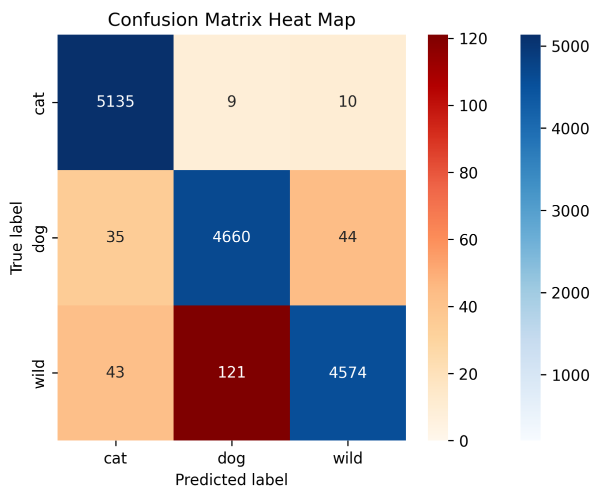

Figure 2: Confusion matrix heat map

has difficulty distinguishing between nuanced differences in facial features across different species and may be over-relying on certain features, such as fur texture, color and facial structure, in its classification process.

Conclusion

Overall, the model performs exceptionally well, but there are signs of over-fitting and it struggles to classify certain images. It excels at predicting cats but frequently misidentifies wild animals as dogs, and vice versa. Misclassifications indicate an overreliance on characteristics such as fur texture, color, and facial structure. To improve, the model should be refined to better distinguish between classes.

Figure 3: Some misclassified images