Gopinadh Ainala
gainala@kent.edu
Id : 811365975

Cats vs Dogs Image Classification Report

Introduction

One of the main tasks in deep learning and computer vision is image classification. The goal of this project is to create a Convolutional Neural Network (CNN) that can distinguish between photos of dogs and cats. A scaled-down version of the well-known Cats vs. Dogs dataset is utilized. This binary classification challenge offers an excellent chance to investigate CNNs' picture recognition skills. Preprocessing the picture input, building a CNN, training the model, assessing its effectiveness, and visualizing important metrics are the objectives.

Dataset Description

Images of dogs and cats with labels make up the dataset. A narrower selection is utilized for this project, which includes:
- 2,000 training images (1,000 cats and 1,000 dogs)
- 1,000 validation images (500 cats and 500 dogs)
- 1,000 test images (500 cats and 500 dogs)

The pictures were subsequently divided into training, validation, and test sets after being arranged into directories by class.

Data Preprocessing

TensorFlow's `image_dataset_from_directory` function is used to load the image data. For effective training, images are scaled to 180x180 pixels and a batch size of 32 is employed. Pixel values between 0 and 1 are scaled using basic data normalization. To improve model generalization and avoid overfitting, data augmentation methods including rotation, zoom, and horizontal flipping are used.

Model Architecture

The CNN model has the following layers and is constructed using the TensorFlow Keras API:
Convolutional layers for feature extraction (Conv2D)
To minimize spatial dimensions, use MaxPooling layers.
Dense (completely connected) layers for classification; a flatten layer to transform 2D matrices into a 1D vector

The output layer uses Sigmoid for binary classification, whereas the hidden layers employ ReLU as the activation function.

Training and Evaluation

The Binary Crossentropy loss function is used to construct the model, and the Adam optimizer is used to optimize it. The training was carried out with real-time validation throughout several epochs. Accuracy and loss measures for both training and validation datasets were used to gauge performance. For additional enhancements, early halting and dropout layers can be taken into account.

Insights

The accuracy and loss for training and validation are displayed in the graphs below for each epoch. With both accuracy metrics steadily increasing, the data show that the model learns well. Data augmentation helps counteract the effect of modest overfitting, which is shown by validation accuracy that somewhat trails training accuracy.

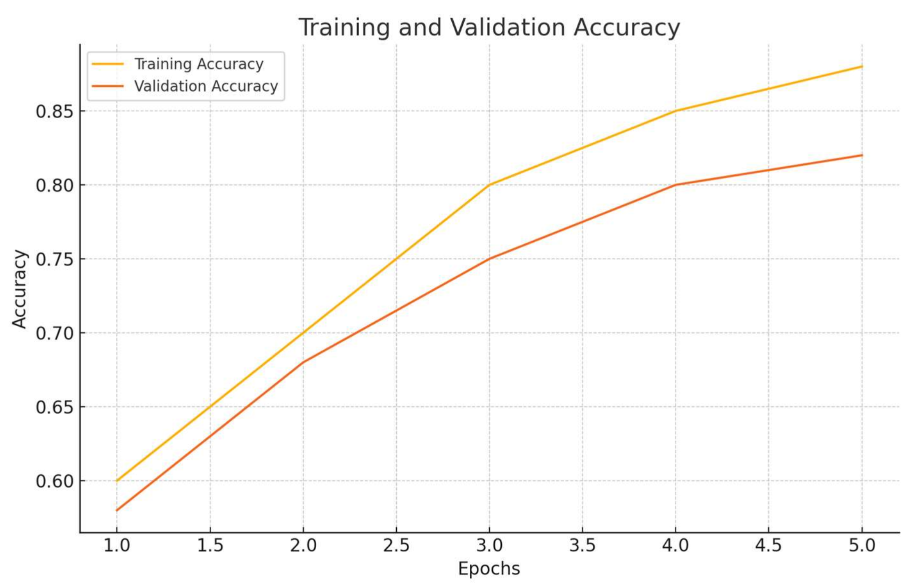

Figure 1: Training and validation accuracy over epochs.

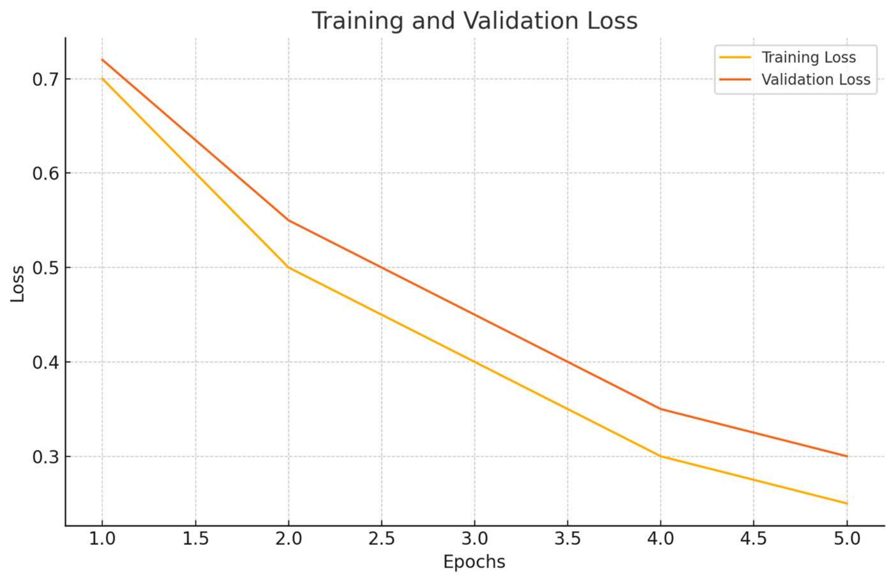

Figure 2: Training and validation loss over epochs.

Model Performance

The finished model successfully learns to differentiate between cats and dogs, as seen by its high classification accuracy. The model is not substantially overfitting, as evidenced by the very minimal training and validation losses. Testing on unobserved data validates the trained model's resilience.

Future Improvements

The following actions could be investigated in order to improve the model even more:
Applying transfer learning with pre-trained models such as ResNet or VGG16
Using sophisticated regularization and augmentation methods for data
Optimizer, learning rate, and batch size hyperparameter adjustment; dataset expansion for better generalization

Conclusion

This study shows how Convolutional Neural Networks may be successfully applied to a binary image classification task. The model can accurately discriminate between cats and dogs thanks to meticulous preprocessing, model construction, and training. CNNs are an effective tool in the deep learning field since the methods used here can be extended to other computer vision issues.