Python ML Project

Created by

Shaik Abdul Khadar

Project Overview:

Your goal is to build a machine learning model that can accurately classify images as either a cat or a dog.(you can use any other dataset from internet)

Data Set:

I have used Kaggle Dogs vs Cats Dataset for this project

Dataset Link: https://www.kaggle.com/datasets/salader/dogs-vs-cats

Process:

After Downloading and importing the data set I have done data exploration where I found there were 10,000 dog images and 10,000 cat images for training set of the data, as the PC will take very long time to train the model I have taken 1000 each of dog and cat images from the data sets and stored in the new file, and displayed the random images of cats and dogs in Notebook.

After creating new folder i have resized the every image to (224,224) through loop and checked for the random image.

Then I have created labels for the classes (dogs and cats) in the dataset as

Cats > 0

Dogs >1

After labelling I have converted all the images into arrays which will be used to train the model

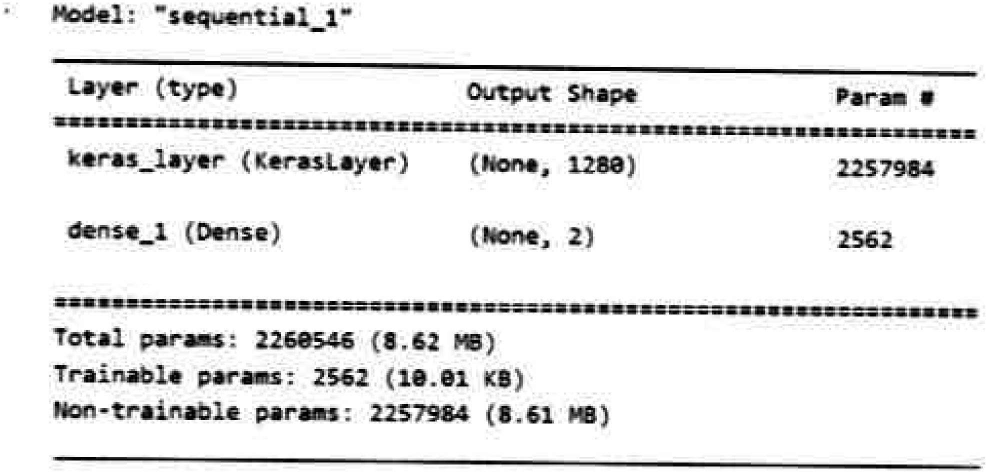
With the help of cv2 and glob I have created the images into arrays where my arrays shape was **(2000, 224, 224, 3),** where count of the dogs and cats where 2000 with the size of 224,224 and 3 refers to the RGB colors.

After the Data Exploration and Preprocessing I have defined X variable as arrays and Y variable as labels and done **Train Test Split** with the help of **sci-kit learn** where I have splitted 80% of data into training and 20% of data into testing after splitting my training data was (1600, 224, 224, 3) (400, 224, 224, 3).

Building Neural Network:

Models used:

- Tensorflow
- Keras
- Mobilenet V2
- Sequential model with Dense Layer



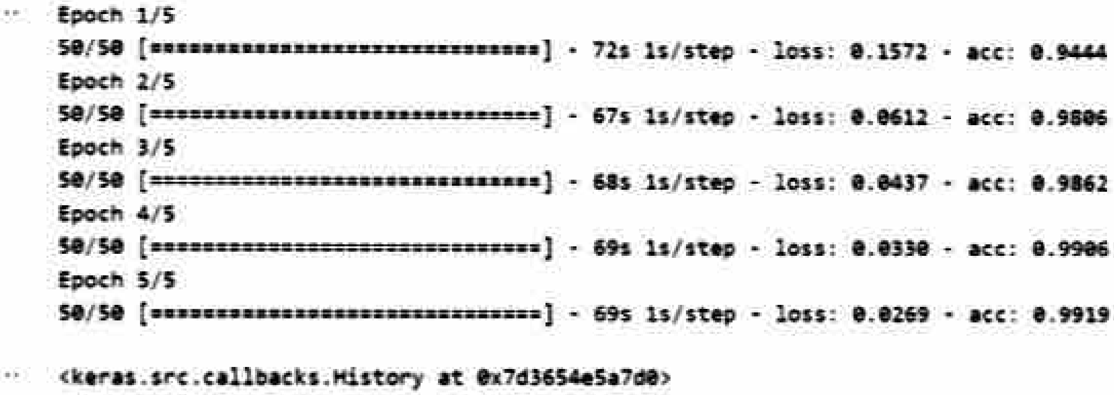
```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 keras_layer (KerasLayer)    (None, 1280)              2257984

 dense_1 (Dense)             (None, 2)                 2562

=================================================================
Total params: 2260546 (8.62 MB)
Trainable params: 2562 (10.01 KB)
Non-trainable params: 2257984 (8.61 MB)
_________________________________________________________________
```

After building the model I used **Adam Optimizer** for compiling and **SparseCategoricalCrossentropy for loss** with **Accuracy for metrics** in the model

Then fitted the model with 5 epochs and results for the training is here:



```
Epoch 1/5
50/50 [==============================] - 72s 1s/step - loss: 0.1572 - acc: 0.9444
Epoch 2/5
50/50 [==============================] - 67s 1s/step - loss: 0.0612 - acc: 0.9806
Epoch 3/5
50/50 [==============================] - 68s 1s/step - loss: 0.0437 - acc: 0.9862
Epoch 4/5
50/50 [==============================] - 69s 1s/step - loss: 0.0330 - acc: 0.9906
Epoch 5/5
50/50 [==============================] - 69s 1s/step - loss: 0.0269 - acc: 0.9919

<keras.src.callbacks.History at 0x7d3654e5a7d0>
```