Emotion Detection through Image Classification

Leigh Warner
Department of Computing, TU Dublin, Tallaght, Ireland
X00140292@myTUDublin.ie

Introduction

The project that I carried out in my final year at Technology University Dublin aimed at predicting the seven basic human emotions through the development and fine tuning of several Convolutional Neural Networks along with predicting such emotions on pre-trained Convolutional Neural Networks that already exist such like VGG-16, ResNet-50, and Inception-V3. The seven basic human emotions consist of Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprised. This project would see the prediction of such emotions through the classification of images that I have collected and augmented to further widen my data.

Data Selection

The reason why I chose this data for the training and validation set for my models is that it was the only one I could avail of. I discussed with my project supervisor that I would manually collect such data through friends and family but that came to a halt promptly due to the fact of the global pandemic. I also tried to submit for a data set to be released on the grounds that I could use it for my Data Analytic's Project but that fell through and I was left with a data set that consisted of 72 base images.

Data Preparation

To prepare the data for my project I had to expand my data set from 72 images to a sufficient amount. I availed of python's libraries such like Image Data Generator, which allowed me to augment my images in real-time. This library allowed me to change some aspects of these images and apply augmentation such like flipping, shearing, and zooming. I also applied rotation to the images, starting at 15 degrees and rotating all the way through 360 degrees which allowed me to grow my data set substantially overtime. I also made use of Python's Pillow library. This library allowed me to open, manipulate and save images when working with them throughout the data preparation stage of the project. Before implementing the data, normalization was applied to the data in that I converted the data to an numpy array and reshaped the data to fit the model input shape and then it was normalized by dividing the numpy array by 255.

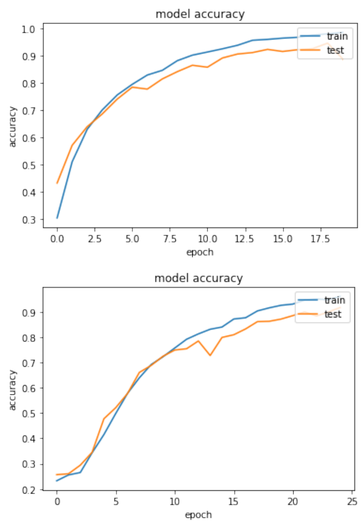
Model Development

1. Infrastructure:

For the development of my project, I availed of Google's Colaboratory. Since my own machine that I was working on from home was limited, I found myself in a sticky spot in terms of machine power for the project. Google offer a free Colab that allows anyone to write and execute python code through a web browser and avail of their free GPU service. This was ideal for me as my machine did not have sufficient power and would overheat during the development stage of the Convolutional Neural Network stage.

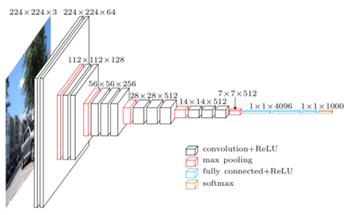
2. Development of Convolutional Neural Networks:

During the development of the Convolutional Neural Network stage, I developed and fine tuned four different models that consisted of different model architecture and different hyper-parameters. I used a validation split of 33 percent for the validation set and training used the remaining 66 percent of the data for the training of the models. All models were trained and tuned on different hyper-parameters. I developed the models based on investigations through capacity and depth. When I satisfied with the model capacity and depth I then moved onto batch size with suitable training exposure based on the amount of epochs I tested.

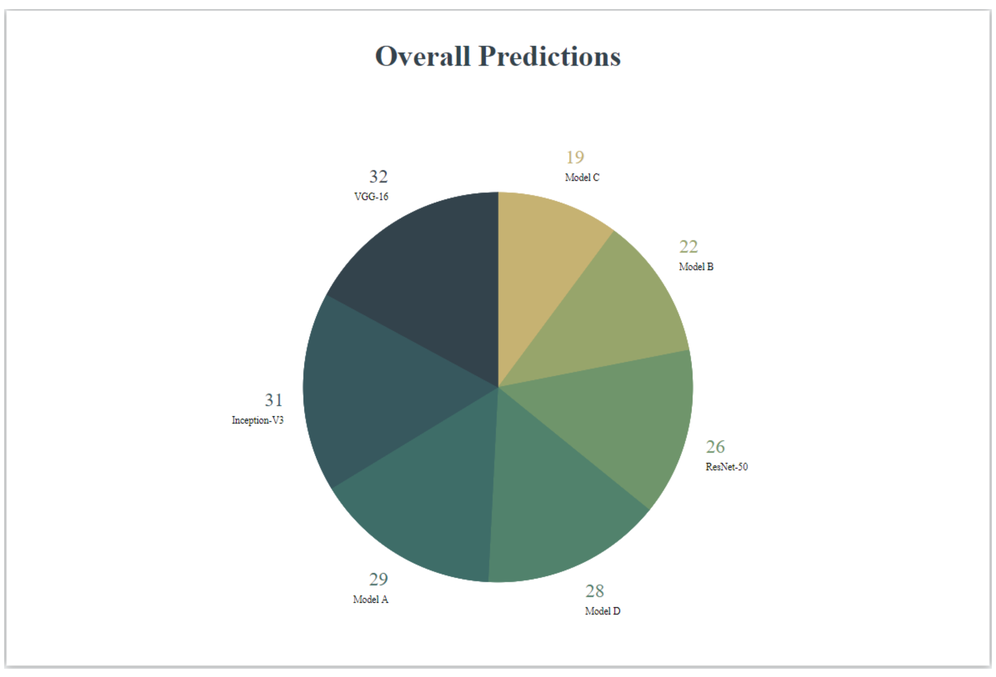


3. Existing Models:

The models that I used in the project were VGG16, ResNet-50, and Inception-V3. The reason why I used pre-trained models in my project is that I want to see how my own models would compete against already made models. These models were trained on ImageNet weights and consisted of different types of architecture. I developed and fine tuned these models until I was satisfied with their accuracy.



Results



Conclusions Work

The whole concept behind this project was to develop Convolutional Neural Networks and perform emotion detect through image classification. I would like to conclude the fact that right across the board I've seen all the models not preform well when predicting the unclassified images. We have seen the models predict well for the happy and surprised emotions as I believe that the data was sufficient for them emotions, whereas the data for the rest of the emotions was not sufficient enough for the models to learn and make good predictions on. I do believe with a more respectable data-set that all these models would perform much better.