National University of Sciences and Technology
School of Electrical Engineering and Computer Science
Department of Computing

CS-405 Deep Learning

Fall 2023

Lab 4

Deep Learning Model for Ethnicity Classification

Date: 11th Oct 2023

Name: Alina Nasir

CMS ID: 342350

Instructor: Dr Daud Abdullah

Dataset Creation and Data Splitting:

For the dataset creation part, the irrelevant features like age and gender were removed from the dataset to make the dataset simpler to learn. Since our problem was based on predicting ethnicity only, the feature that was important for classification was the pixels. The dataset contained both the input labels that were the pixels and the target class of ethnicity, therefore we had to create a separate data frame for target class, while dropping it in the previous data frame. The pixels' array was then converted to numpy array so that it can be passed to the model. It was then normalized per channel of RGB by dividing it with 255 which is the range for RGB.

For data splitting, the data was split according to the criteria mentioned in lab that was 70% for the training dataset, 15% for validation and 15% for test dataset.
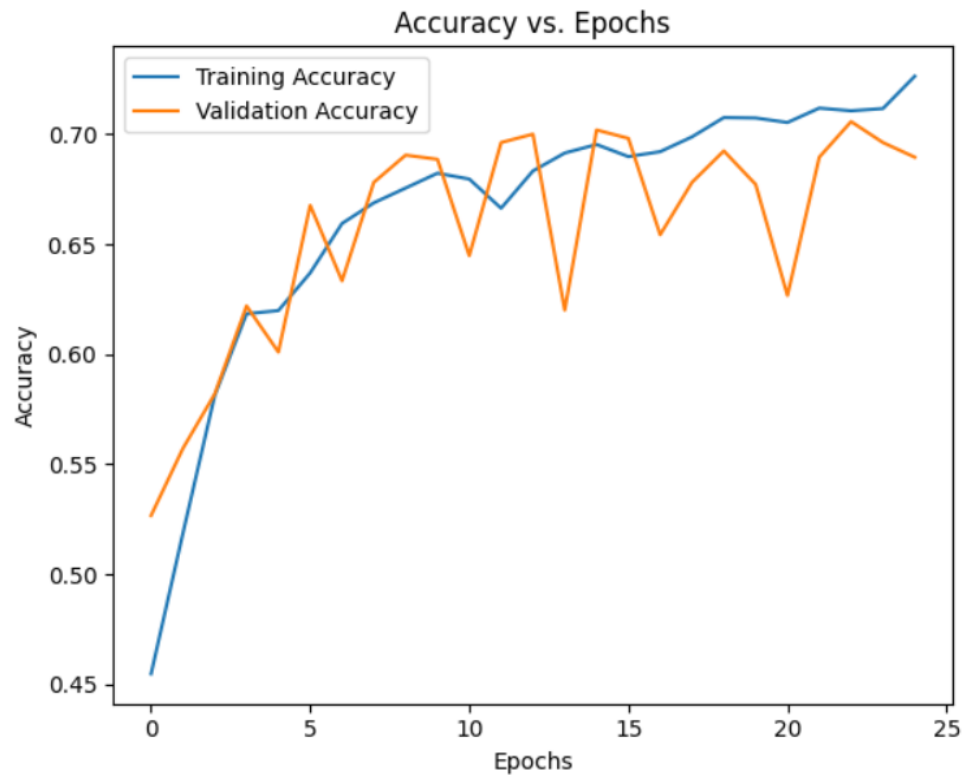
Network Architecture:

The neural network build was sequential model of Keras with input layer of size 2304, that is the length of the pixels' array. There are three hidden layers of the model with first of 256 neurons, the second of 512 neurons with dropout of 0.1 and the third with 512 neurons. The activation function at each hidden layer is reLu. The output layer is of size 5 because there are 5 output classes and activation function used was softmax. The softmax was used as an activation function because it is a multi-class classification problem and softmax is used to predict probability of each class.

Training Process:

The number of epochs for training was kept 25 with batch size of 32, the accuracy, loss and validation loss was calculated at each epoch. The results are shown below:

Accuracy vs Epochs:

Loss vs Epochs: