Faculty of Computer and Information Sciences
Computer Science Dept.
Ain Sahms University

Shehab Ahmed Bassiouni
Omar Essam Abdelhadi
Mohammed Mas'ad Saad
Mohammed Adel Abdullhamid
Ibrahim Abdelghani Mansour

Human Face Detector and Analyzer

Dr. Nermine Naguib
Dr. Assma Bahi

Abstract.
This Project aims to Develop an API application capable of detecting human face and predict Age, Gender, Ethnicity, and Emotions from it by utilizing the power of Convolutional Neural Networks.

Introduction

In recent years, the need for systems capable of detecting and analyzing the human face is increasing rapidly with the wide possible application of it in our life. Those systems can be used in many areas like marketing by analyzing customer emotions from facial image to measure the degree of satisfaction or analyzing the ages of most-visiting customers etc. it can also be integrated in surveillance systems to detect specific features. With the rabid development of machines and big datasets, applying deep learning techniques such as neural networks became possible. Starting with AlexNet, the CNN networks started to be deeper and have more parameters, with the right optimization techniques the performance of those networks can be vastly increased. In this project we built four CNN models and train and optimize them to develop models that achieve satisfactory results.

Datasets:

- Age Model: Facial-Age/UTK-FACE
- Gender Model: UTK-FACE/B3FD
- Emotions Model: FER-2013
- Ethnicity Model: Ethnicity-Aware/Arab-Celeb-Faces

Preprocessing:

- Datasets Cleaned using Haar Cascade Algorithm to remove images with no human face in it
- Data Augmentation
- Image Normalization and Resizing
- Mean Centering

Age and Gender Models CNN Archeticures:

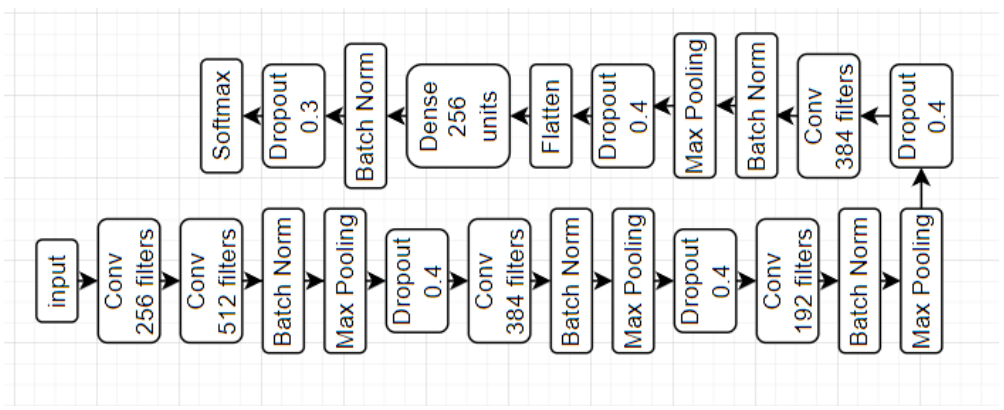
The CNN architecture proposed by Juel in [1] was used in age and gender models. In the Age model, the last layer had softmax activation with 7 units. In the Gender model, the activation changed to sigmoid activation in the last layer.

Ethnicity Model CNN Archeticures:

In ethnicity model a pre trained VGG-16 architecture with VGG-Face weights was used by fine tuning it. The architecture was fine-tuned by adding Global Average Pooling layer instead of the flatten layer, a dense layer with 1024 units and RelU activation, a dropout layer with probability 0.5 and a dense layer with 5 units and SoftMax activation.

Emotions Model CNN Archeticures:

The following architecture in figure 1 was used for emotions model. Input dimension is 48x48x1, all convolution layers have kernel size of 3 and relu activation with "same" padding, Max pooling layers have pool size of 2, Dense layer have 256 units with relu activation and softmax layer have 7 units.



Emotions Model CNN Archeticures:

The following architecture in figure 1 was used for emotions model. Input dimension is 48x48x1, all convolution layers have kernel size of 3 and relu activation with "same" padding, Max pooling layers have pool size of 2, Dense layer have 256 units with relu activation and softmax layer have 7 units.

Training:

- TensorFlow was Used
- Cross Entropy Loss
- Accuracy as Metric
- Cross Validation
- Adam Optimizer
- 0.001 Learning Rate
- Early Stopping Callback with 5 patience and 1 Verbose
- Checkpoints Callback with 1 Verbose
- Tensorboard Callback with 1 Verbose
- 32 Batch Size for Age and Gender Models
- 128 Batch Size for Ethnicity Model
- All Models Fitted for 100 epochs
- Class Weight Balancing

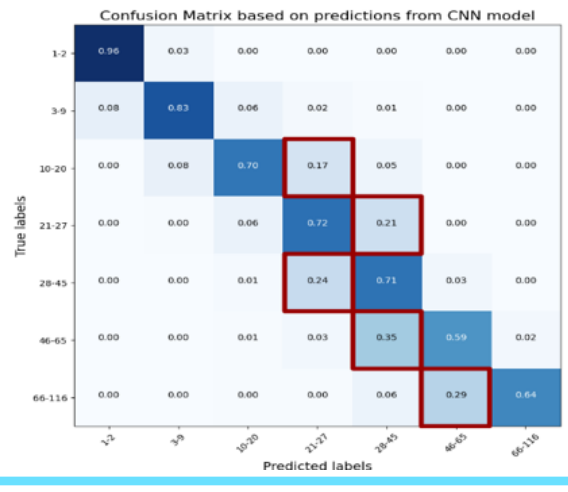
Face Detection:

The Haar Cascade algorithm was Used for face Detection.

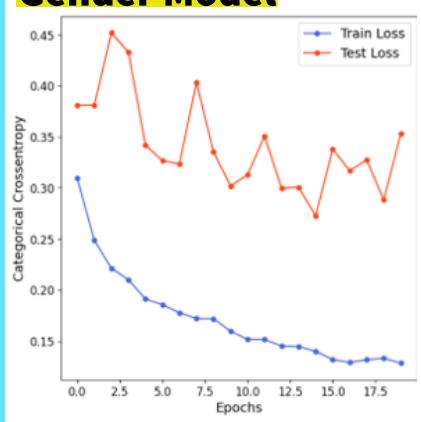
Results

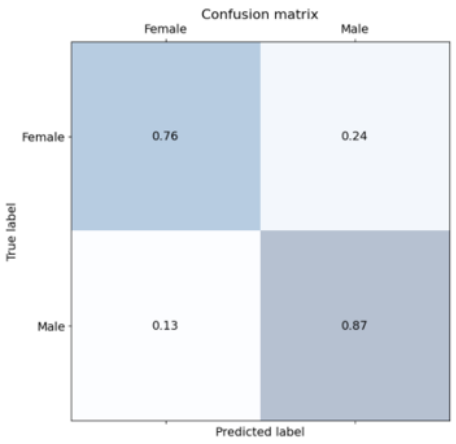
Model	*Loss*	*Accuracy*
Age Model	0.75	71%
Gender Model	0.8463	81%
Ethnicity Model	0.3744	86%
Emotions Model	0.9411	67%

Age Model

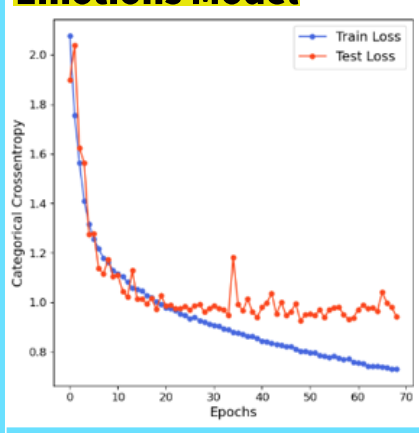


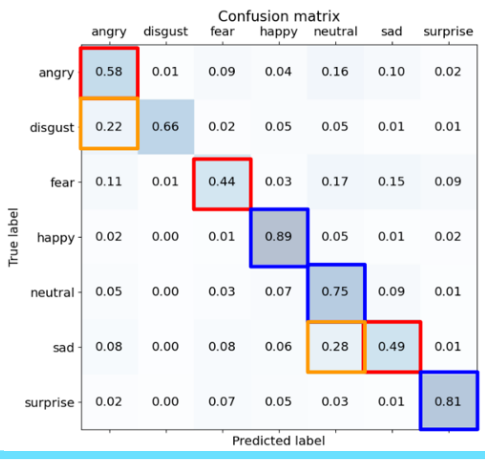
Gender Model



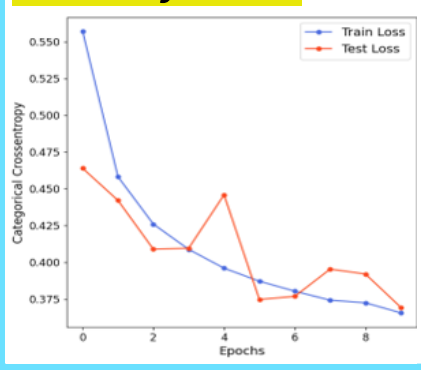


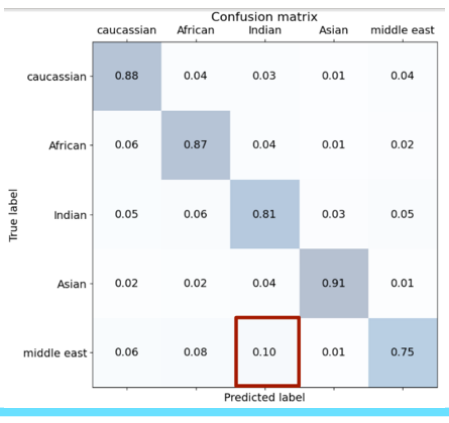
Emotions Model





Ethnicity Model





References:

Nippon Datta Nippon , Juel Sikder Juel , "An Approach Based on Deep Learning for Recognizing Emotion, Gender and Age", Rangamati Science and Technology University, 2022