PES UNIVERSITY

(Established under Karnataka Act No. 16 of 2013)
100-ft Ring Road, Bengaluru – 560 085, Karnataka, India

Report on

‘Age, Gender & Emotion Detection’

Submitted by

Rutwik Ganagi (PES120180000627)
Harshit Raj (PES1201801364)
Akshay Parashar (PES1201801846)

Aug - Dec 2021

under the guidance of

Internal Guide

Raghavendra M J
Associate Professor
Department of ECE
PES University
Bengaluru -560085

FACULTY OF ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

PROGRAM B.TECH

Table of Contents

Chapter -1

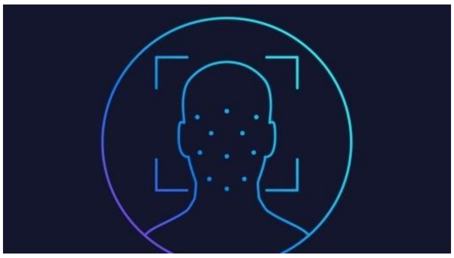
Introduction

- Automatic gender classification and age estimation from facial image has become an attractive research area in the field of machine learning.
- Real and apparent age estimation of human face has attracted increased attention due to its numerous real-world applications like medical diagnosis, etc.
- More recently, Deep learning with Convolutional Neural Networks (CNNs) methods have been extensively used for these kind of classification tasks due to their excellent performance in facial analysis.
- In our project we propose a simple convolutional net architecture that can be used even when the amount of learning data is limited.

Chapter -2

Problem Statement

To design and implement a deep learning model, capable of detecting Age, Gender and facial Emotion of a person in real-time.



Chapter -3

Literature survey

3.1 A Gender Recognition System from Facial Image

Tawhid, Md. Nurul Ahad & Dey, Emon. (2018). International Journal of Computer Applications.

- **Problem Statement:** Approach for classifying gender from facial images using the feature extraction method named as tCENTRIST.
- **Methodology:** Proposed method for the gender classification task are as follows:
  - ✓ Image Enhancement
    - Bilateral Histogram Equalization with Pre-processing (BHEP)
  - ✓ Feature Extraction and dimension reduction
    - tCENTRIST for feature extraction
    - PCA for dimensionality reduction
  - ✓ Classifier Training
    - Support Vector Machine(SVM**)**
- **Results/Conclusion:** Accuracy(%) on different dataset.

Feature extraction using tCENTRIST	Accuracy (%)		
	LFW	Adience	Color FERET
Proposed method without BHEP	93.75	82.79	97.24
Proposed method with BHEP	94.29	84.99	98.30

- **Research Void:** A more sophisticated pre-processing method like by applying dynamic size blocks for feature extraction in input images to enhance the pose variation can be applied to improve the accuracies.

3.2 A Lightweight Convolutional Neural Network for Real and Apparent Age Estimation in Unconstrained Face Images

O. Agbo-Ajala and S. Viriri, in IEEE Access September,2020

- **Problem Statement:** Designing a lightweight CNN model of fewer layers to estimate the real and apparent age of individuals from unconstrained real-time face images that can be deployed on mobile devices.
- **Methodology:** Proposed method for the age estimation task are as follows:
  - ✓ Face Detection and Alignment
    - Face Detection
    - Face Alignment
  - ✓ CNN Lightweight Architecture
- Mean Absolute Error (MAE) is the evaluation metric used in this paper.

Layer type	Output size	Filter size/Stride
Input	$224 \times 224 \times 3$	–
Conv1	$56 \times 56 \times 96$	$7 \times 7/4 \times 4$
Stage2	$28 \times 28 \times 256$	5×5
Stage3	$14 \times 14 \times 384$	3×3
Conv4	$7 \times 7 \times 384$	3×3
HPooling		
Softmax		

$$MAE = \sum_{i=1}^{N} \frac{|l_i - l_i^*|}{N}$$

- **Results/Conclusion:** Model achieved the bestperformance with fewer parameters and a reductionin computational time which satisfies the real-time requirement of the model for its applicability on mobile devices.
- **Research Void:** There is need to consider a more robust and quality image pre-processing algorithm that detect the unfiltered images faster for real-time estimation of the images.

3.3 Deep Learning based approach to detect Customer Age, Gender and Expression in Surveillance Video

E. P. Ijjina, G. Kanahasabai and A. S. Joshi (11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020)

- **Problem Statement: Designing a Deep learning model to detect customer age, gender and expression in surveillance video.**
- **Methodology:The surveillance videos from CCTV cameras in stores is given as input to the proposed framework to detect faces and estimate age, gender and expression of the person. Processes involved are:**
  - ✓ **Face detection using Haar Cascade object detection model**
  - ✓ **Age and gender detection using Wide Residual Network architecture**
  - ✓ **Expression recognition using mini Xception model**
- **Results/Conclusion: 82.9% accuracy rate for gender detection and 70.8% accuracy for age-range.**
- **Research Void: The model has difficulties in detecting the side facial profile of customers not facing the camera, thereby affecting its accuracy. Model is optimized for detecting the age and gender traits of customers belonging to Indian origin.**

3.4 Age and Gender Prediction using Deep Convolutional Neural Networks

Rafique, A. Hamid, S. Naseer, M. Asad, M. Awais and T. Yasir (International Conference on Innovative Computing (ICIC), 2019)

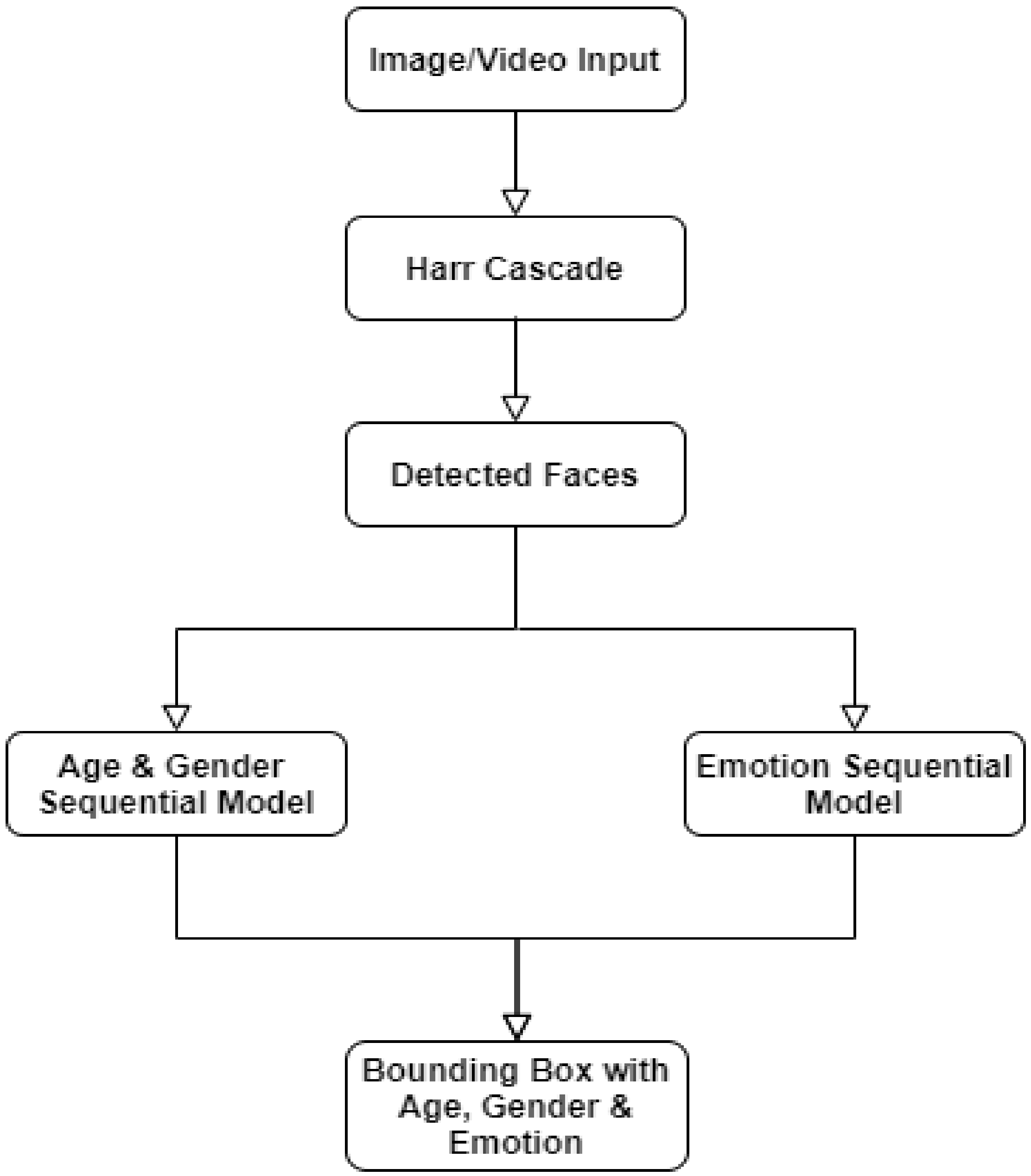
- **Problem Statement:** Designing a deep CNN model to predict age and gender of a person.
- **Methodology:** Deep CNN is used to improve age and gender prediction using HAAR Feature-based Cascade classifiers. Gender classification is done by using neural networks trained on the small set of frontal faces image. It uses 3D structure of head and image intensities for classifying age. The proposed network consists of

three convolutional layers in which two of them are fully connected with small number of neurons. Processes involved are:

- ✓ HAAR Feature-based Cascade Classifiers
- ✓ Local Binary Pattern (LSB) descriptor
- ✓ Support Vector Machine (SVM) classifier

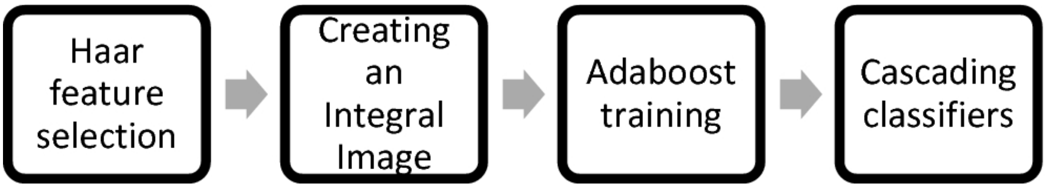
- **Results/Conclusion:** Using deep CNN, the accuracy for age and gender prediction is 79%.
- **Research Void:** Over fitting is intensified when considering deep convolutional neural networks due to large parameters, accuracy can be improved by using efficient algorithms and more precise CNN architecture.

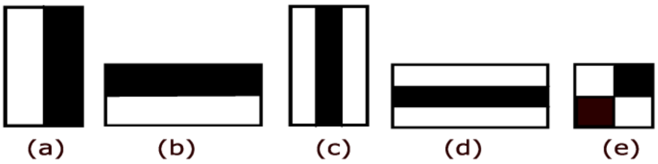
Chapter -4
Block Diagram



Chapter -5
HAAR Feature Selection

- It is an Object Detection Algorithm used to identify faces in an image or a real time video.





- Haar Feature Selection:
  - (a)&(b) Edge Feature
  - (c)&(d) Line Feature
  - (e) Four Rectangle Feature
- Creating an Integral Image: To reduce computational time
- AdaBoost Training: Selects which haar feature is relevant
- Cascading Classifiers: Stage-wise process

Chapter -6

Age & Gender Detection Model

- Convolutional Neural Network (CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.
- Model Type: Sequential Model

Convolution Layer No.	**No. Of Nodes**	**Kernel Size**
1	128	3
2	128	3
3	256	3
4	512	3

Table 1: Age Model

Convolution Layer No.	**No. Of Nodes**	**Kernel Size**
1	36	3
2	64	3
3	128	3
4	256	3
5	512	3

Table 2: Gender Model

The model type which we have used is Sequential model. Sequential is the easiest way to build a model in keras. It allows us to build model layer by layer using add() function. For age detection we have used Regression which is a supervised machine learning technique to predict continuous values. For gender detection we have used binary classification which is used to predict a label.

For age model we have used four convolutional layers and for gender model we have five convolutional layers. After each convolutional layer we have used maxpool2D() which down samples the input along its spatial dimension. For the models the kernel size is 3, which basically means the size of the filter matrix is 3x3. The activation function which we have is ReLU which stands for Rectified Linear Unit. 'Dense' is the layer type which we will be using for the output layer. The output layer will contain only 1 node. .

6.1 Training and Testing Results

- **Dataset: 1523 images from UTKFace**
- **Number of epochs: 50**
- **Accuracy: 79.52 %**
- The shape of each input is 200x200 pixels. The number of epochs which we have used is 50. Epoch basically means the number of complete passes of the entire training datasets

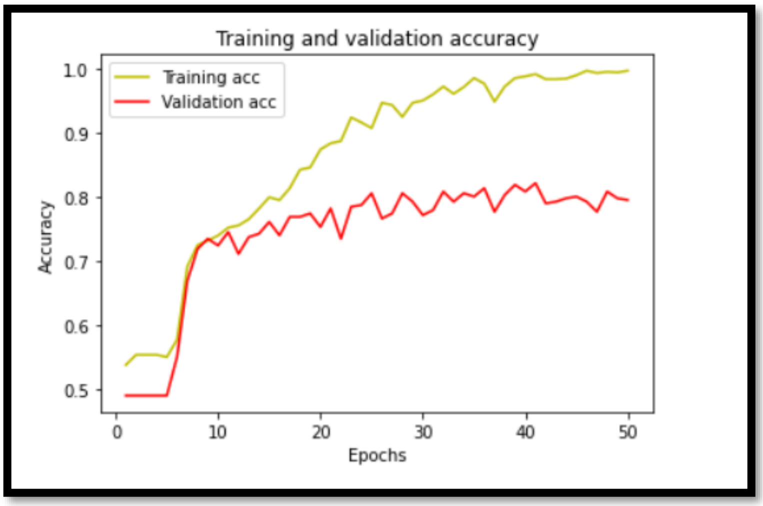


Fig 1: Training and Validation Accuracy

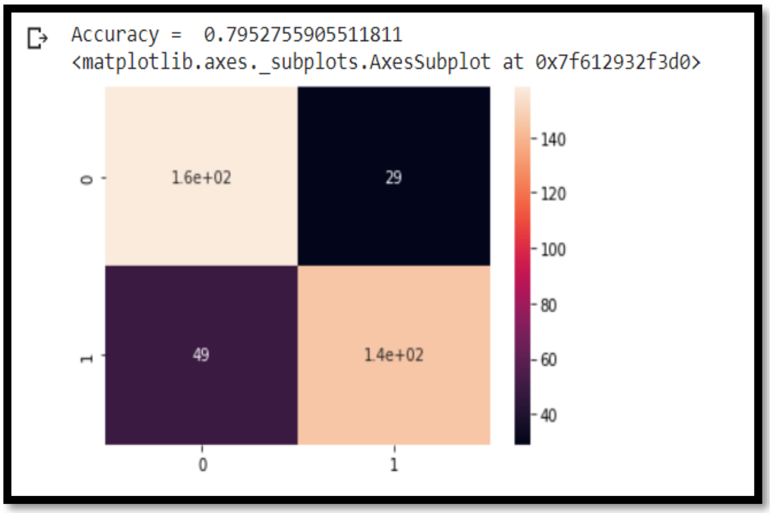
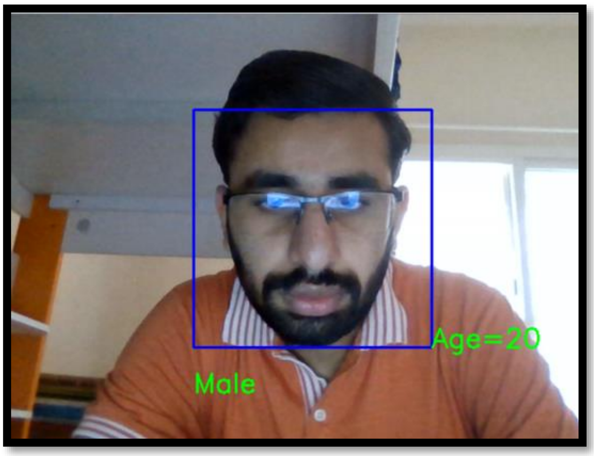


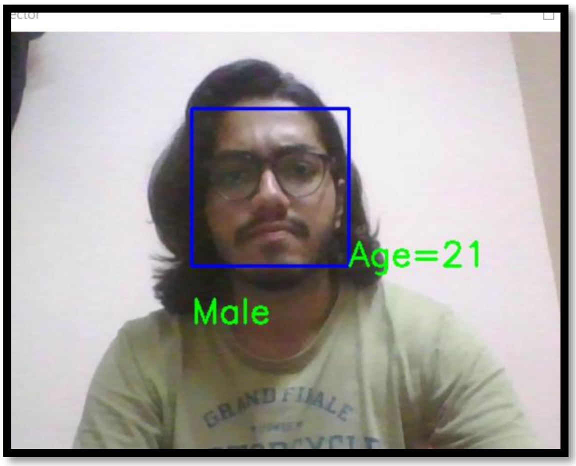
Fig 2: Confusion Matrix

The following figures are the results obtained from running the code in google colab. The first figure shows the plot of Accuracy vs Epochs. The yellow curve indicates the Training Accuracy and red curve indicates the Validation Accuracy. As we increase the number of epochs, we see that the training and validation accuracy increases. For age and gender detection we are getting an accuracy of 79.52%.

The second figure is a plot of confusion matrix. Confusion matrix is a technique which summarizes the performance of classification algorithm. In the plot, 0 stands for Male and 1 stands for Female. For male gender, 160 images are correctly classified and 29 images are misclassified. For female gender, 140 images are correctly classified and 49 images are misclassified.

6.2 Outputs Obtained





These are the results obtained, Gender and Age has been correctly identified.

Chapter 7

Emotion Detection Model

Four convolution layers have been used:

Convolution Layer No.	No. Of Nodes	Kernel Size
1	32	3
2	64	3
3	128	3
4	256	3

Table 3: Emotion Model

- Activation function used are ReLU and softmax.The softmax function is a function that turns a vector of K real values into a vector of K real values that sum to 1.
- Loss type used is ‘categorial_crossentropy’. Categorical crossentropy is basically a loss function that is used in multi-class classification tasks.
- Optimizer used is ‘adam’.
- Emotion labels: ['Angry’, ‘Disgust', 'Fear', 'Happy’, 'Neutral’, 'Sad’, 'Surprise’]
- Dataset: 4176 images from Kaggle
- Test/Train split = 75:25

7.1 Training and Testing Results

- Number of epochs = 50
- Accuracy = 78.12%

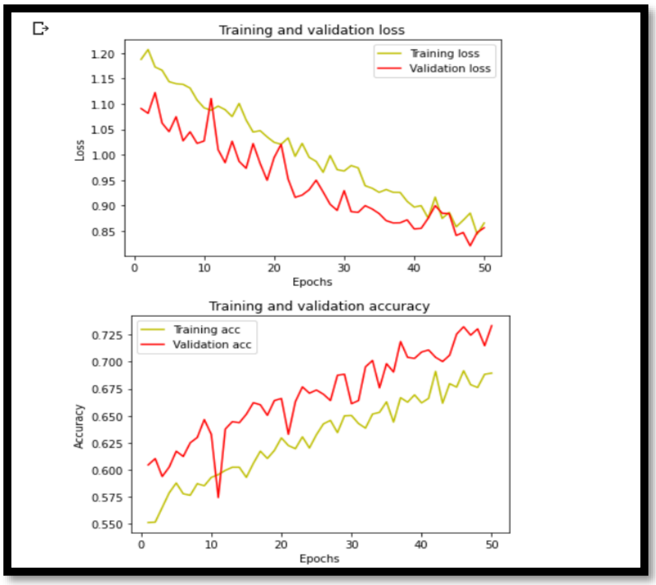


Fig 3: Training and validation plot

The above figure is a plot of Training and Validation loss and accuracy for each epoch. From the plot we see that as we increase the number of epochs the Training and Validation loss decreases and Training and Validation accuracy.

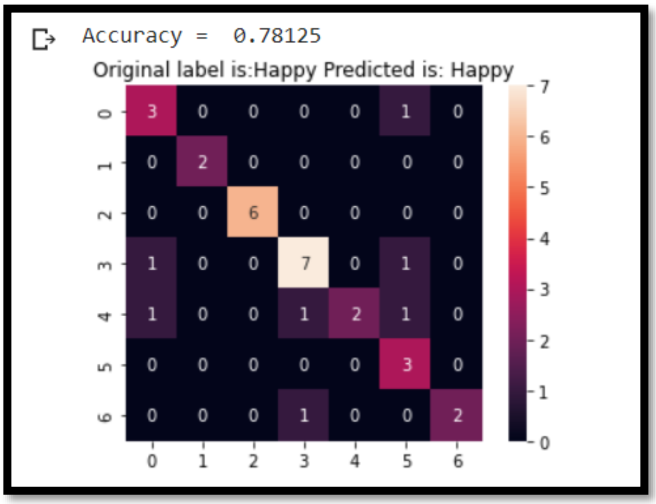
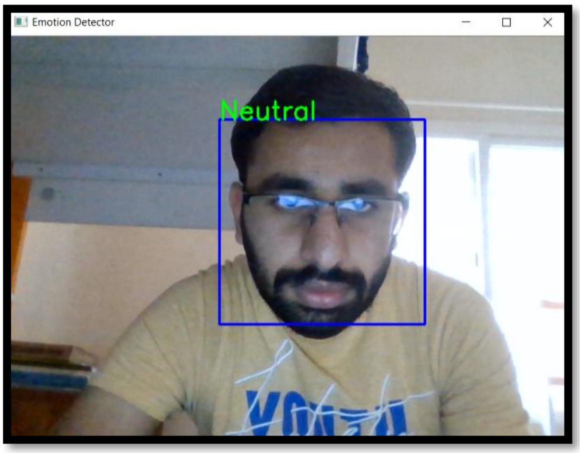
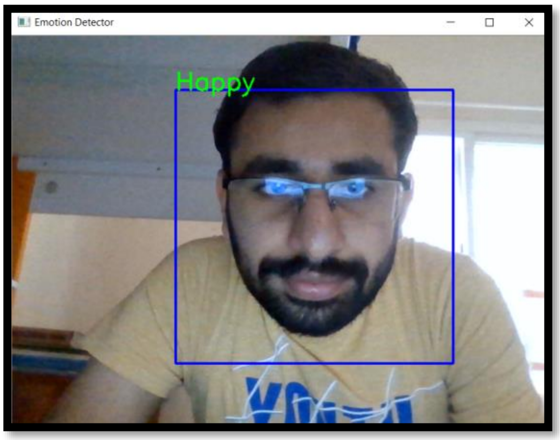


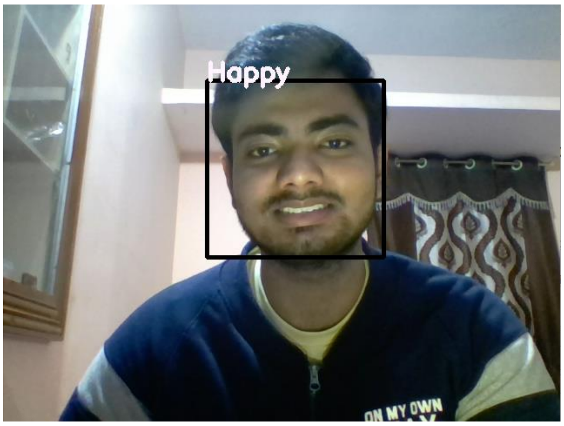
Fig 4: Confusion Matrix

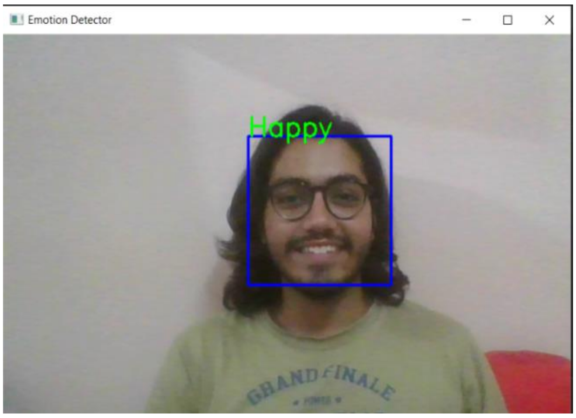
The above figure is plot of confusion matrix. In the figure the numbers labelled from 0 to 6 stands for different labels of emotion. For example, 0 stands ‘Angry’ and 1 stands for ‘Disgust’.In the figure 3 images of ‘Angry’ are correctly classified and 1 image is misclassified. Similarly, we see that, 7 images of 'Happy' are correctly classified and 2 images are misclassified

7.2 Outputs Obtained



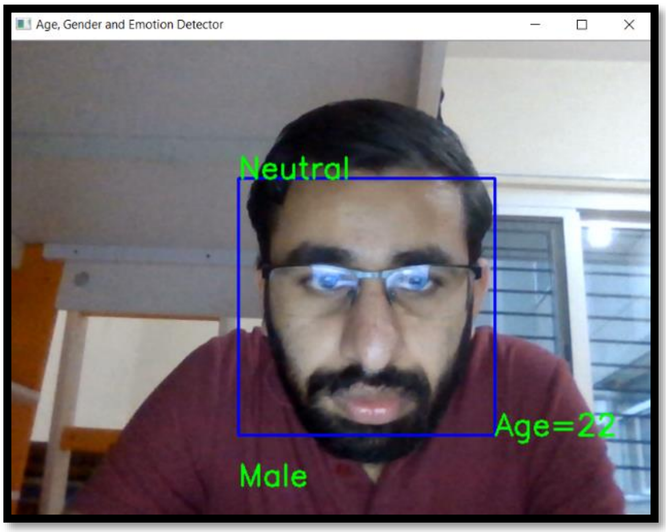






Chapter -8

Age, Gender & Emotion Detection Results



The above figures are the results of real time Age, Gender and Emotion Detection. We have loaded all the three trained models in one code, that is Age, Gender and Emotion. We have used single for loop to show all the three labels that is Age, Gender and Emotion. We have used cv2.VideoCapture(0) function to operate the web-cam. The labels are displaced using cv2.putText(). The title that is 'Age, Gender and Emotion Detector is displayed using cv2.imshow(). We have also used detectMultiScale() function, so whenever we move face front or back, the face is still detected and even multiple faces can be detected.

Chapter -9

Conclusion

- By using deep CNN model, we are getting accuracy of 79.52% for Age and Gender detection and accuracy of 78.12% for Emotion Detection.
- The accuracy can be further increased by using larger datasets.
- Future work would include improving prediction under low lighting conditions and accurate prediction when the person is situated far away from the camera.

Chapter -10

References

- E. P. Ijjina, G. Kanahasabai and A. S. Joshi, "Deep Learning based approach to detect Customer Age, Gender and Expression in Surveillance Video," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225459.

- I. Rafique, A. Hamid, S. Naseer, M. Asad, M. Awais and T. Yasir, "Age and Gender Prediction using Deep Convolutional Neural Networks," 2019 International Conference on Innovative Computing (ICIC), 2019, pp. 1-6, doi: 10.1109/ICIC48496.2019.8966704.

- O. Agbo-Ajala and S. Viriri, "A Lightweight Convolutional Neural Network for Real and Apparent Age Estimation in Unconstrained Face Images," in IEEE Access, vol. 8, pp. 162800-162808, 2020, doi: 10.1109/ACCESS.2020.3022039.

- Agbo-Ajala, Olatunbosun&Viriri, Serestina. (2020). Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces. The Scientific World Journal. 2020.

- Tawhid, Md. Nurul Ahad & Dey, Emon. (2018). A Gender Recognition System from Facial Image. International Journal of Computer Applications. 180. 5-14. 10.5120/ijca2018915852.