

# Emotion, Age and Gender Extraction from Human Faces in Raspberry Pi based Video Surveillance with Low Cost Webcams.

Department of Information Technology,  
Indian Institute of Engineering Science and Technology, Shibpur

4th Semester Mini Project (IT2291)  
Under Supervision of Dr.Ruchira Naskar

---

Submitted by,  
Harshit Kumar(2022ITB011)  
Subhranil Nandy(2022ITB012)  
Soham Dalui(2022ITB039)

# Contents

- ❑ Introduction
- ❑ Use Cases
- ❑ Workflow of the system
- ❑ Raspberry Pi setup
- ❑ deepface
- ❑ Emotion, Gender and Age extraction from video stream
- ❑ References



# Objective

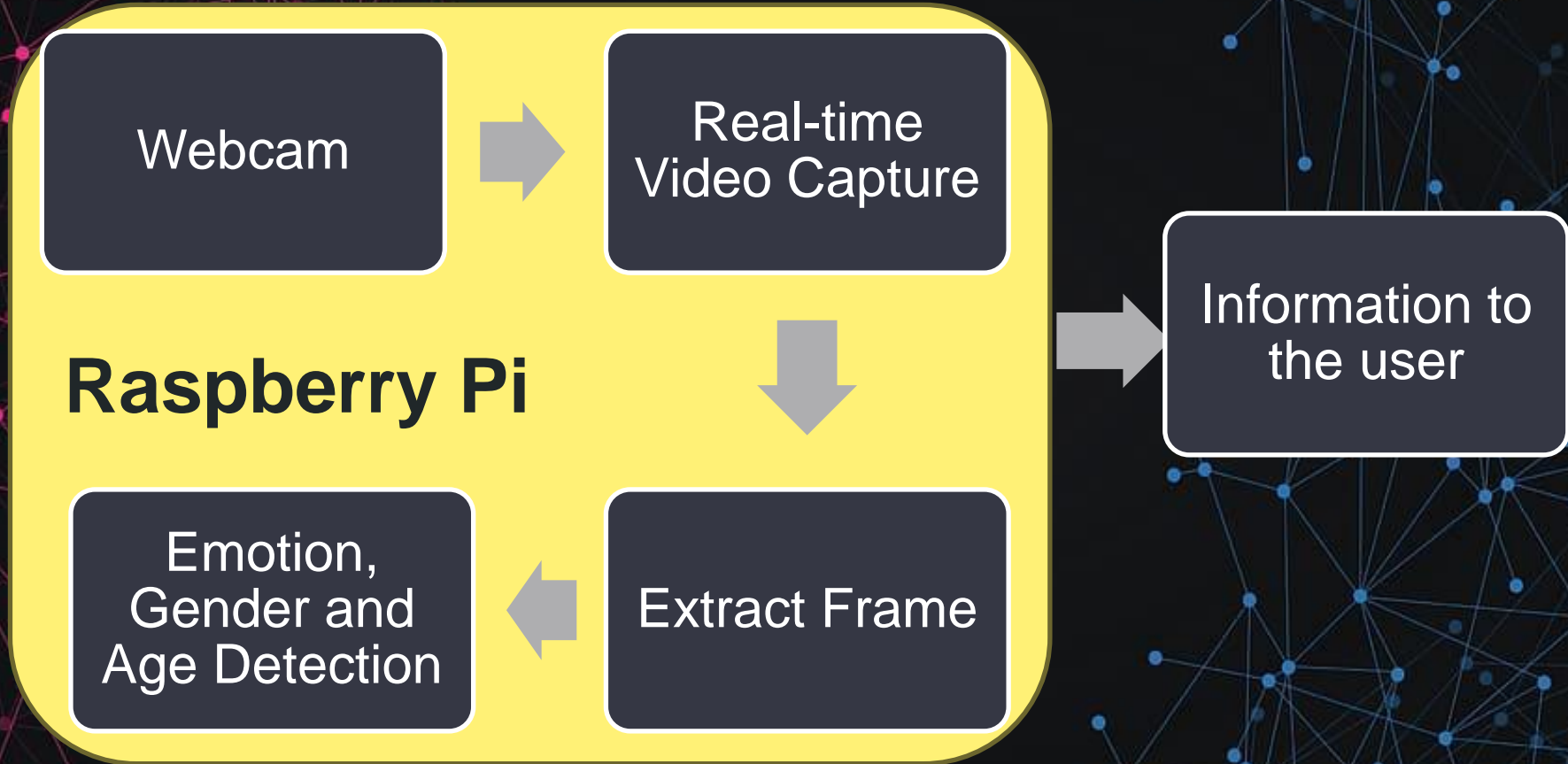
- ❑ To implement a real-time video surveillance using webcam/ CCTV, connected with a Raspberry pi, which forms an edge device.
- ❑ To extract the information like age, emotion and gender of the captured face using the light weight edge computation unit.

## Uses Cases:

- Use Case: Monitoring entry points (doors, windows) for unauthorized access.
- Scenario: Receive alerts when unknown faces are detected, along with their estimated age, gender, and emotional state.
- Customer sentiment analysis: Emotion detection can gauge customer satisfaction during interactions with staff, prompting interventions if negativity is detected.
- Targeted Advertising: By recognizing recurring customers and their demographics (age, gender), the system can display targeted promotions on nearby screens.



## Workflow of the system





[Emotion\_Detection]



18:55



Wastebasket

# Setup of Raspberry Pi 3B+

# Step-wise Setup Process

- Downloaded the latest Raspberry Pi OS : 64 bit Debian version: 12 (Bookworm)
- Used the Raspberry Pi Imager for writing the OS image to a SD card.
- Inserted the SD card into the Raspberry Pi and connected peripherals:
  - Keyboard, mouse, webcam in USB ports
  - Display cable in HDMI port
  - Provided power cord
- Powered up the Raspberry Pi and did some initial setup steps: language, username / password, connect to the network, etc.
- Performed a full system update & upgrade.
- Did the required steps for accessing Raspberry Pi via SSH (only terminal) and VNC (GUI with full access) for headless operation (accessing through laptop/ desktop w/o physically connecting Pi to it).

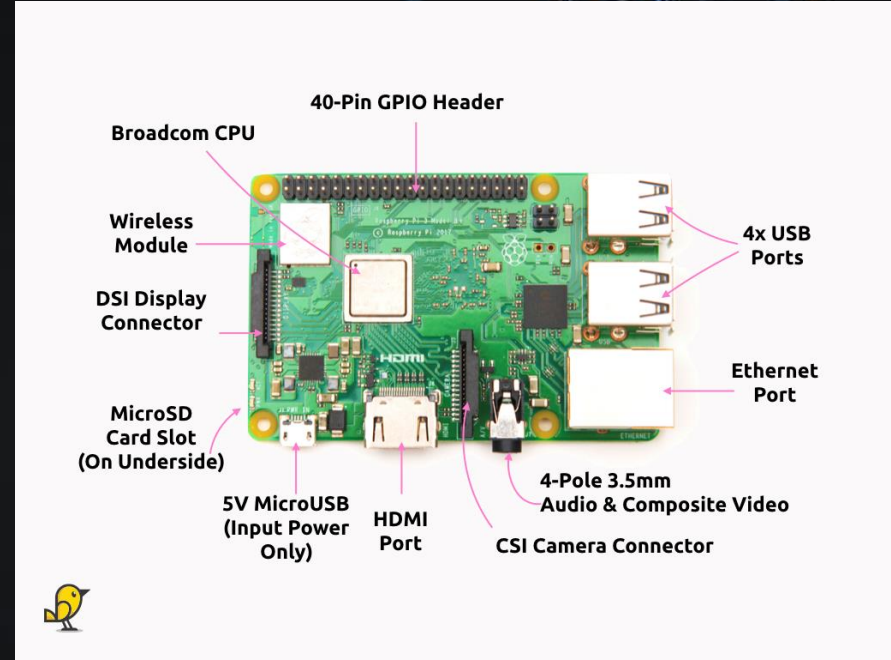


- Raspberry Pi 3B+
- 1GB RAM
- 128GB SD card storage





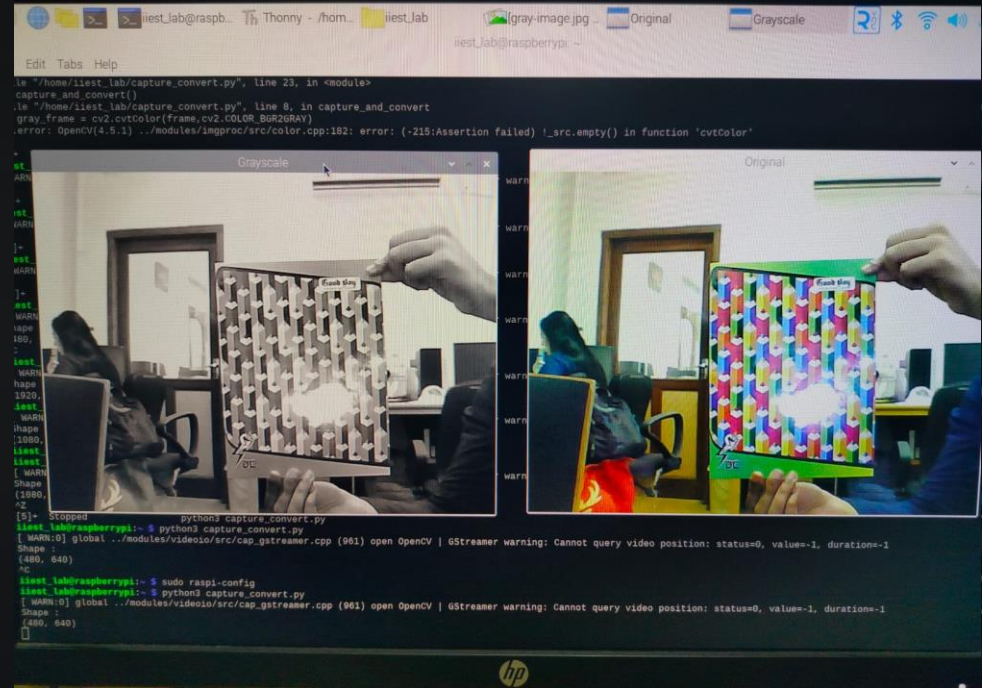
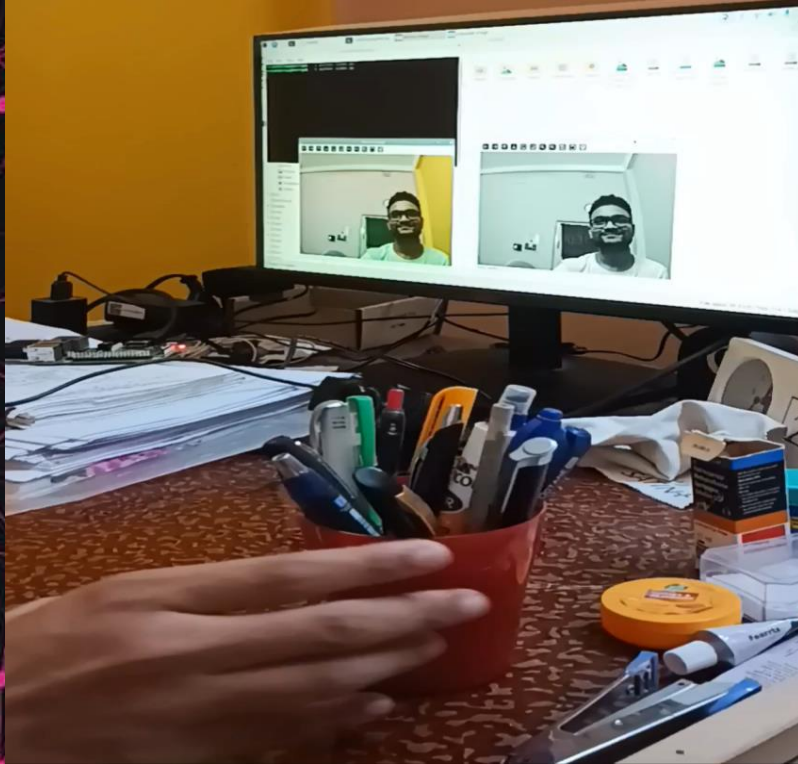
Raspberry Pi connected with USB webcam and connected to laptop wirelessly via VNC



Raspberry Pi 3B+



## Preliminary Performance Tests



# Realtime RGB to grayscale video conversion

# Emotion, Age and Gender Extraction from Human Faces at Edge



# Edge Computing

The background of the slide features a dark blue field with two distinct network-like patterns. On the left side, there is a cluster of red nodes connected by thin red lines. On the right side, there is a cluster of blue nodes connected by thin blue lines. These patterns suggest a distributed network or data flow.

Edge computing offers lower latency, faster data processing, reduced bandwidth usage, and improved security by processing data closer to the source or end-user device

# deepface and its Limitations in RPi

- We tried to run 'deepface' on Raspberry Pi using different models available in the framework but Raspberry Pi 3B+ (1GB RAM) couldn't support the real-time facial attribution computations.
- As a result, the RPi hanged every time we tested a model with the CPU chip becoming excessively heated up.
- Hence our mentor asked us to find a model which must be lightweight and could run within the resource constraints.

README MIT license

## deepface

pypi downloads 2M conda downloads 14k stars 10k license MIT Tests and Linting passing

blog sefiks.com youtube @sefiks follow @serengil patreon 31 patrons sponsors 2

DOI 10.1109/ASYU50717.2020.9259802 DOI 10.1109/ICEET53442.2021.9659697



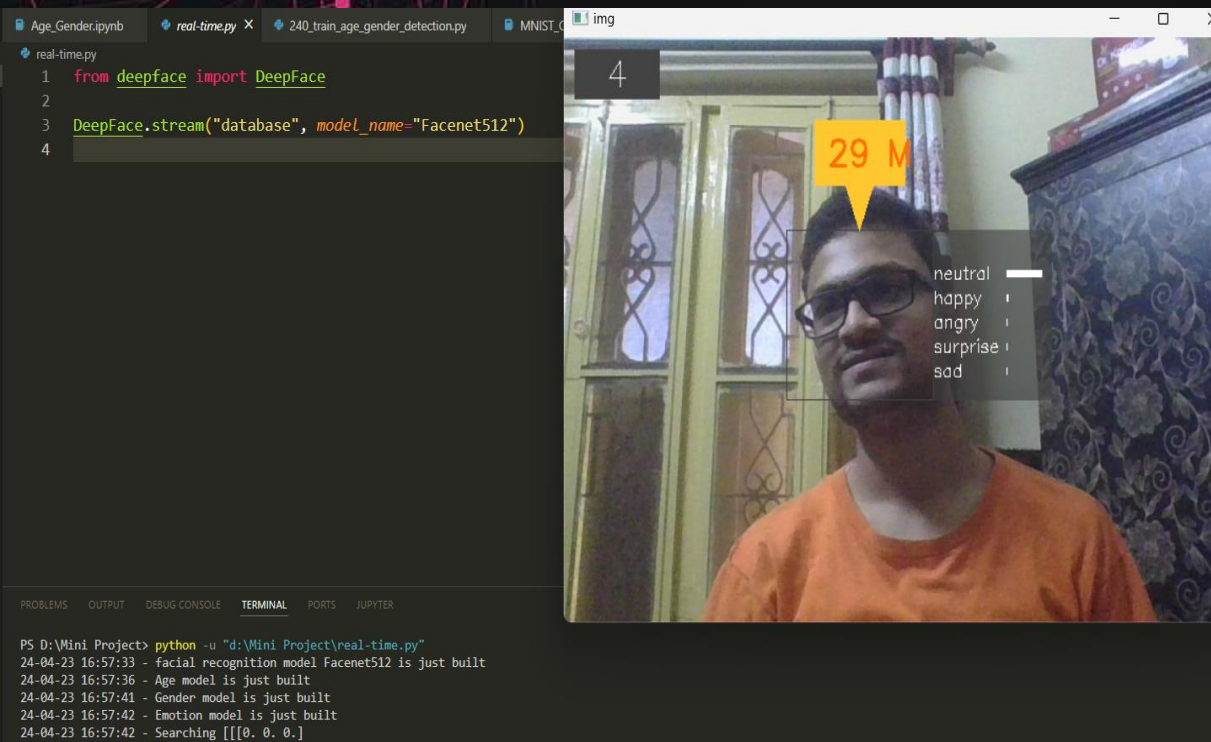
Deepface is a lightweight [face recognition](#) and facial attribute analysis ([age](#), [gender](#), [emotion](#) and [race](#)) framework for python. It is a hybrid face recognition framework wrapping **state-of-the-art** models: [VGG-Face](#), [FaceNet](#), [OpenFace](#), [DeepFace](#), [DeepID](#), [ArcFace](#), [Dlib](#), [SFace](#) and [GhostFaceNet](#).

Experiments show that human beings have 97.53% accuracy on facial recognition tasks whereas those models already reached and passed that accuracy level.

Installation pypi v0.0.90 conda-forge v0.0.89



# deepface running on Laptop with 8GB RAM



```
Age_Gender.ipynb | real-time.py x | 240_train_age_gender_detection.py | MNIST_G...
real-time.py
1 from deepface import DeepFace
2
3 DeepFace.stream("database", model_name="Facenet512")
4
```

4

29 M

neutral |  
happy |  
angry |  
surprise |  
sad |

PROBLEMS OUTPUT DEBUG CONSOLE **TERMINAL** PORTS JUPYTER

```
PS D:\Mini Project> python -u "d:\Mini Project\real-time.py"
24-04-23 16:57:33 - facial recognition model Facenet512 is just built
24-04-23 16:57:36 - Age model is just built
24-04-23 16:57:41 - Gender model is just built
24-04-23 16:57:42 - Emotion model is just built
24-04-23 16:57:42 - Searching [[[0. 0. 0.]
```

Model	Declared LFW Score
VGG-Face	98.9%
Facenet	99.2%
Facenet512	99.6%
OpenFace	92.9%
DeepID	97.4%
Dlib	99.3 %
SFace	99.5%
ArcFace	99.5%
GhostFaceNet	99.7%
<i>Human-beings</i>	<i>97.5%</i>

Models we tested along with their declared Labeled Faces in the Wild (LFW) scores

# Emotion Extraction from Human Faces

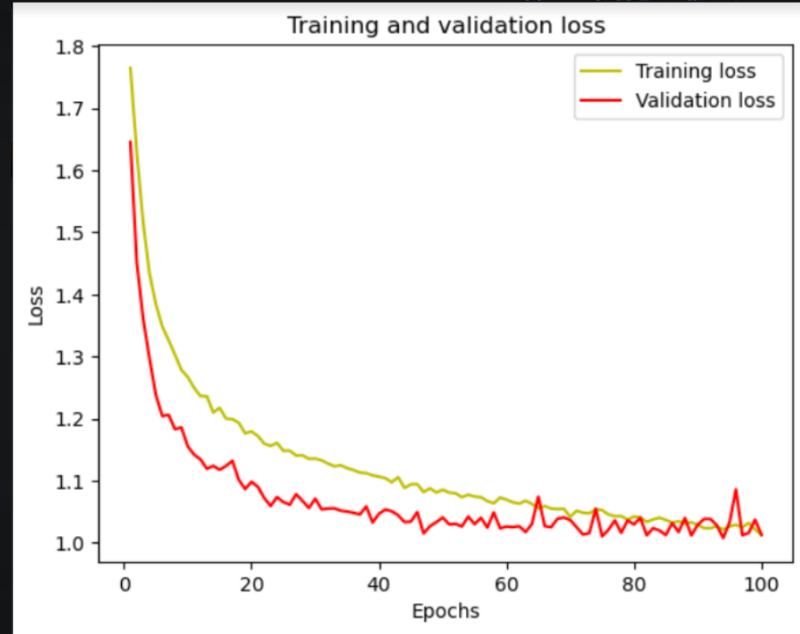
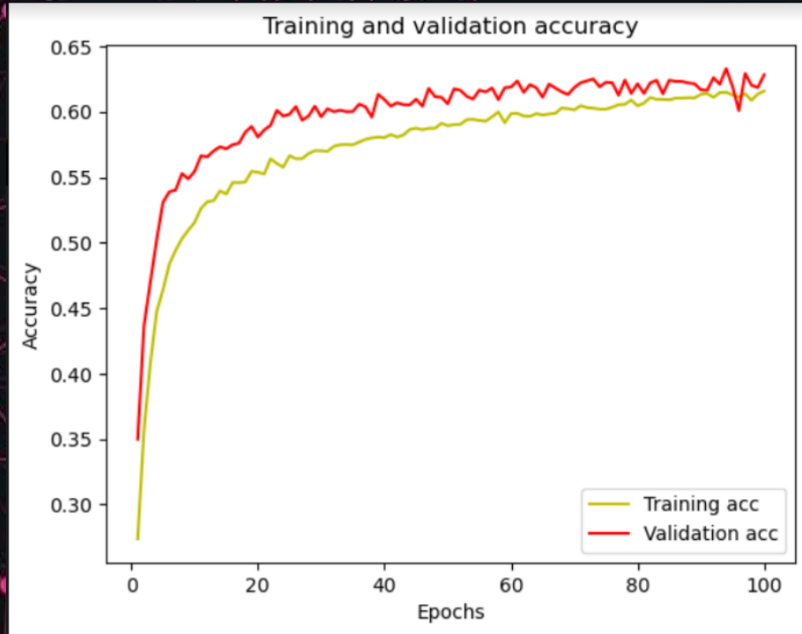
Model: "sequential"

## Model architecture

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 46, 46, 32)	320
conv2d_1 (Conv2D)	(None, 44, 44, 64)	18496
max_pooling2d (MaxPooling2D)	(None, 22, 22, 64)	0
dropout (Dropout)	(None, 22, 22, 64)	0
conv2d_2 (Conv2D)	(None, 20, 20, 128)	73856
max_pooling2d_1 (MaxPooling2D)	(None, 10, 10, 128)	0
dropout_1 (Dropout)	(None, 10, 10, 128)	0
conv2d_3 (Conv2D)	(None, 8, 8, 256)	295168
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 256)	0
dropout_2 (Dropout)	(None, 4, 4, 256)	0
flatten (Flatten)	(None, 4096)	0
dense (Dense)	(None, 512)	2097664
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 7)	3591
Total params: 2,489,095		
Trainable params: 2,489,095		
Non-trainable params: 0		

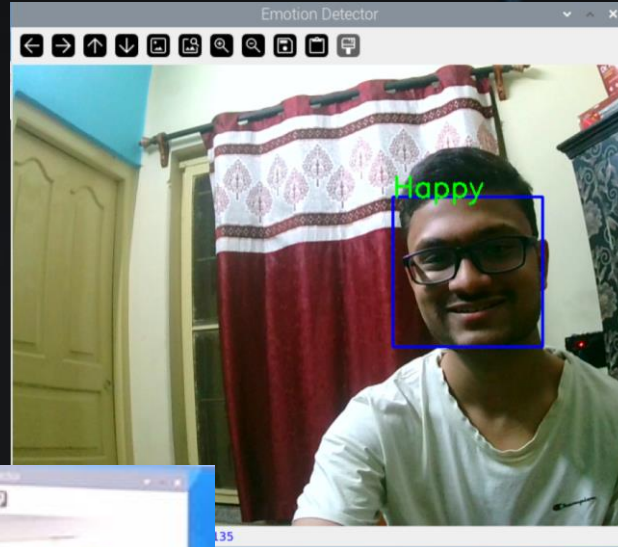
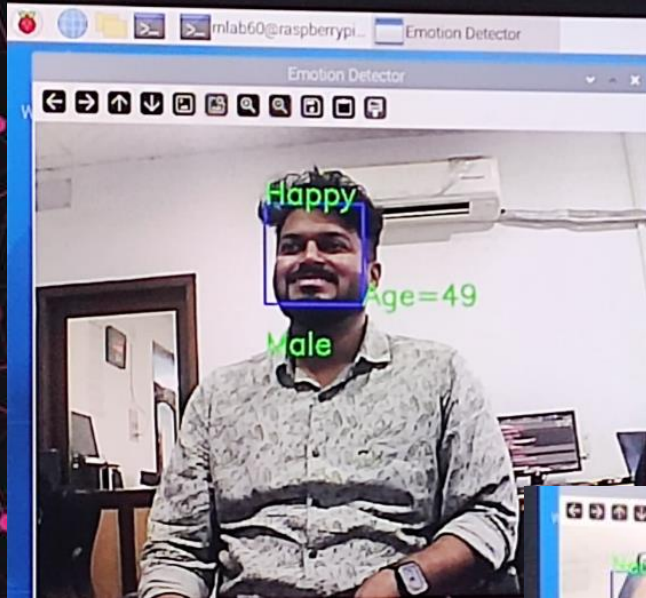
- The model uses Convolutional Neural Networks (CNNs) in order to predict the classes.
- Face datasets used: UTKFace
- Model is converted to TFLite (.tflite) format before being used in Raspberry Pi.

# Accuracy & Loss curves





# Real-time predictions



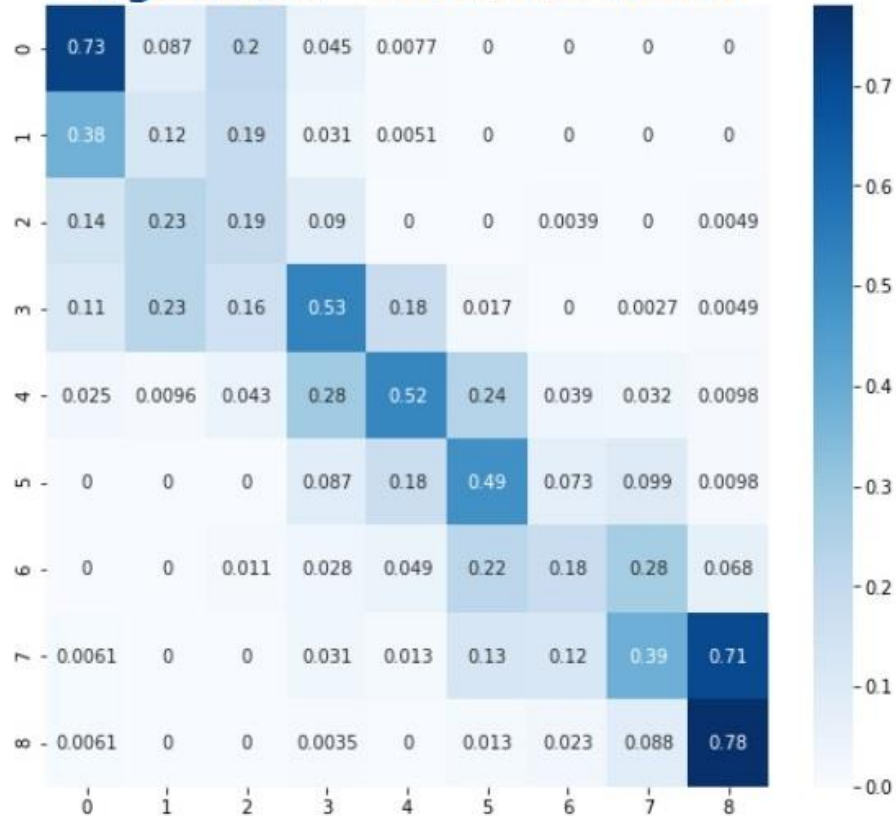


# Age and Gender Extraction from Human Faces

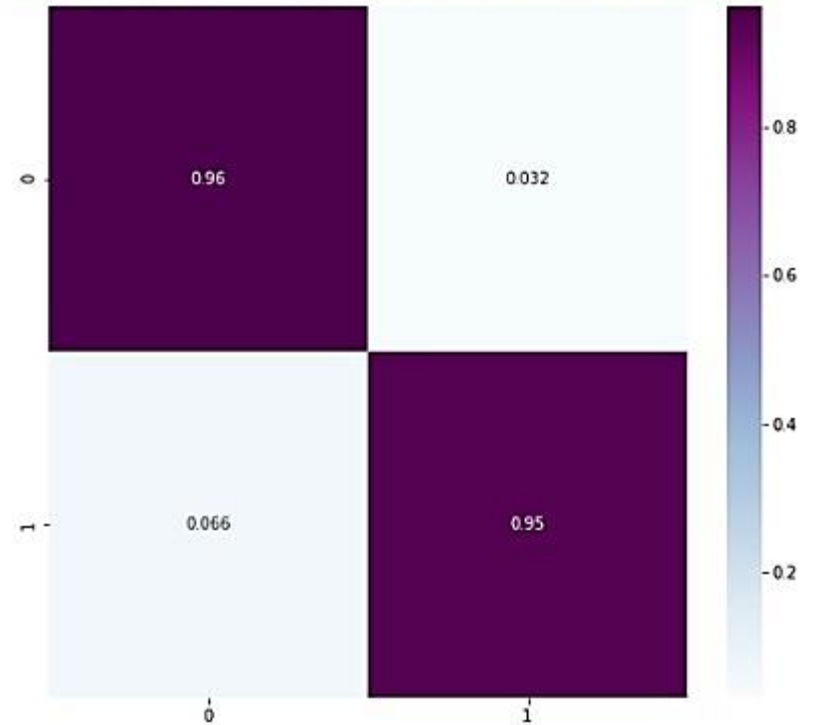
- The model uses Convolutional Neural Networks (CNNs) in order to predict the classes. The network used is MobileNet v1.
- Face datasets used: UTKFace combined with Appa-Real.
- Model is converted to TFLite (.tflite) format before being used in Raspberry Pi.

# Confusion Matrices

## Age model - Confusion Matrix

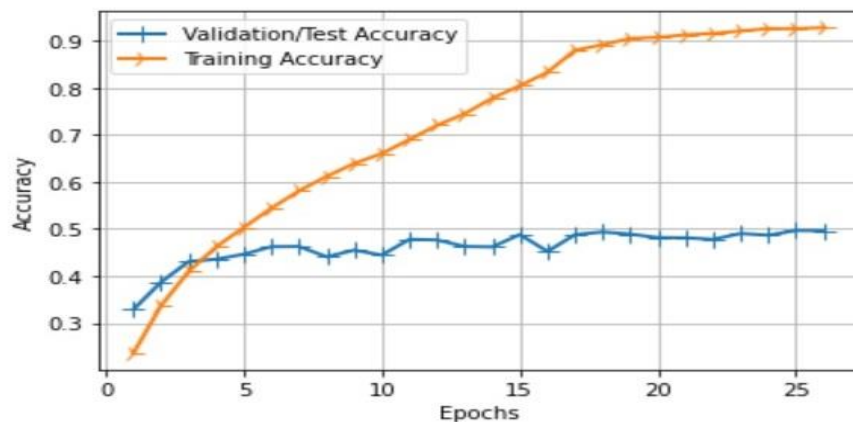
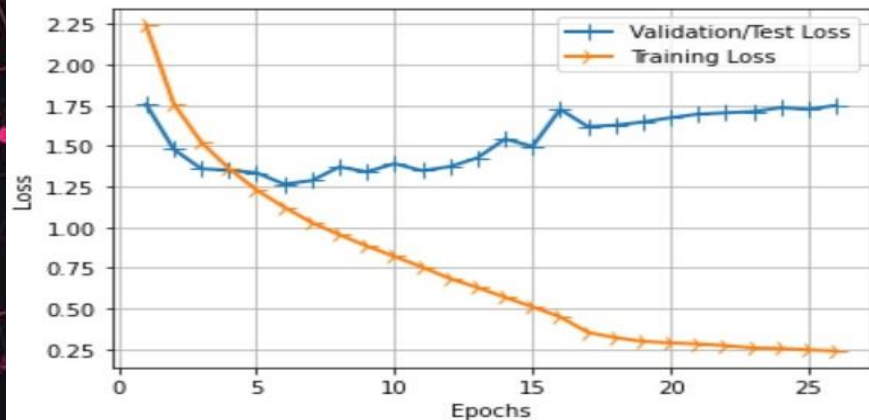


## Gender model - Confusion Matrix

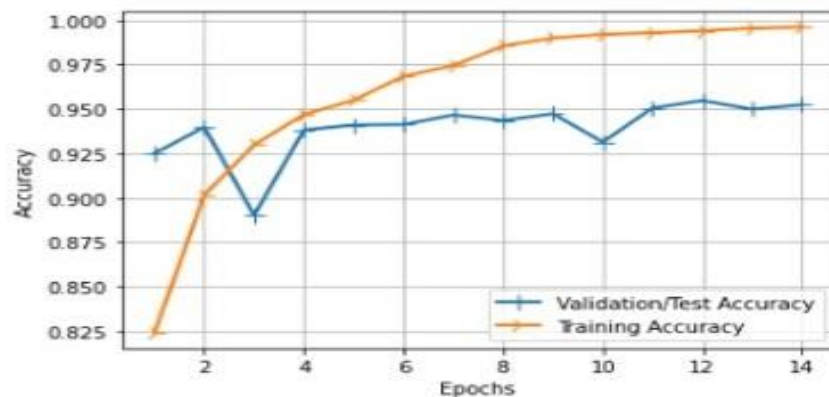
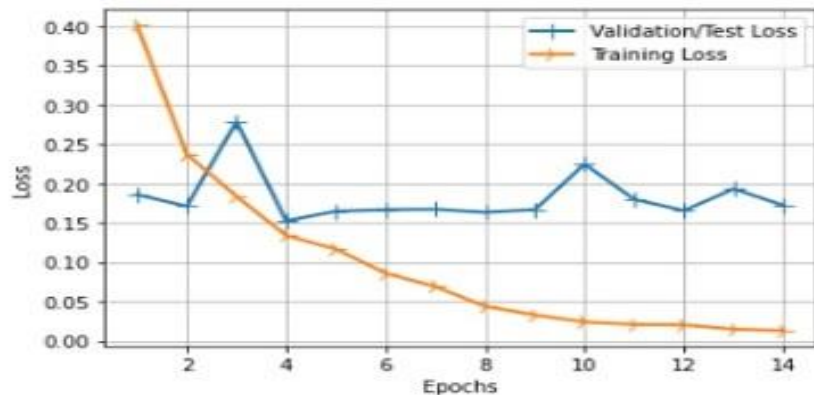


# Accuracy & Loss curves

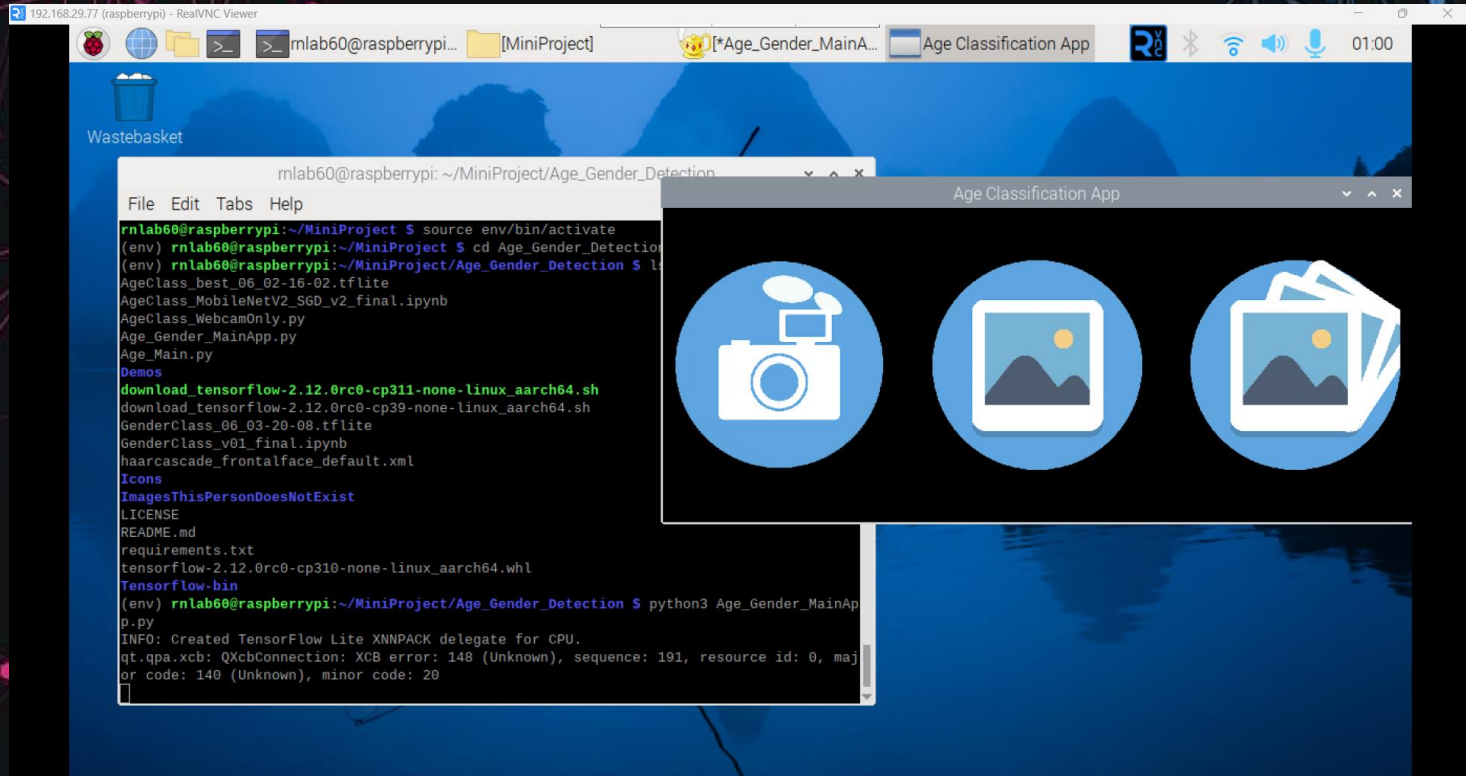
## Age model



## Gender model



# App Interface





# App Interface

192.168.29.77 (raspberrypi) - RealVNC Viewer

Wastebasket

mlab60@raspb... | [Age\_Gender\_De... | [Age\_Main.py - /... | Age Classificati... | Gender and Age ... | 01:22

mlab60@raspberrypi: ~/MiniProject/Age\_Gender\_Detect

File Edit Tabs Help

```
Total time: 450.53744316101074 ms.
Total time: 447.4518299102783 ms.
Total time: 448.09675216674805 ms.
Total time: 450.87122917175293 ms.
Total time: 448.8205909729004 ms.
Total time: 448.8856792449951 ms.
Total time: 448.3349323272705 ms.
Total time: 451.1752128601074 ms.
Total time: 459.40160751342773 ms.
Total time: 450.00720024108887 ms.
Total time: 448.86136054992676 ms.
Total time: 452.28004455566406 ms.
Total time: 449.77498054504395 ms.
Total time: 455.2583694458008 ms.
Total time: 493.65234375 ms.
Total time: 450.0772953033447 ms.
Total time: 453.95827293395996 ms.
Total time: 463.72461318969727 ms.
Total time: 450.09613037109375 ms.
Total time: 448.927640914917 ms.
Total time: 453.0012607574463 ms.
Total time: 449.59425926208496 ms.
Total time: 458.6191177368164 ms.
```

Gender and Age Classification: Live Webcam

Stop Camera

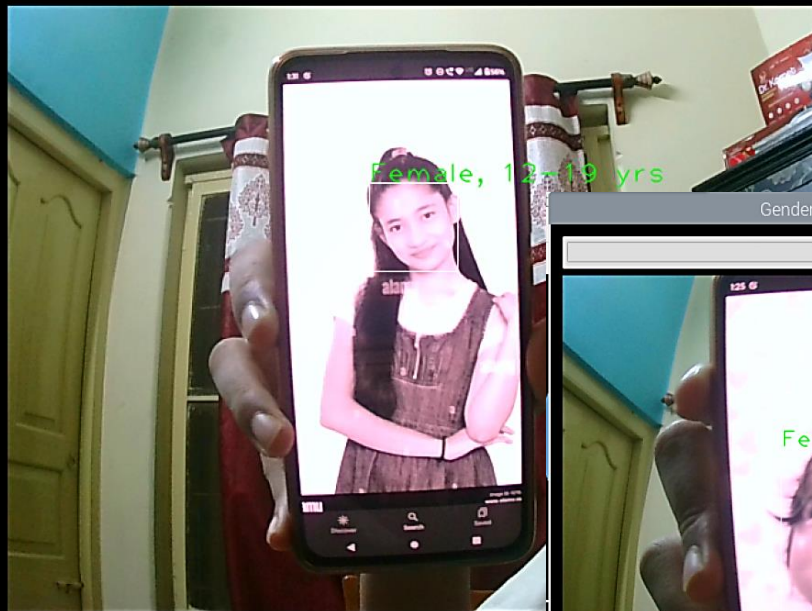
Male, 20-27 yrs

FPS: 1.58

# Predictions on random faces from the Internet

Gender and Age Classification: Live Webcam

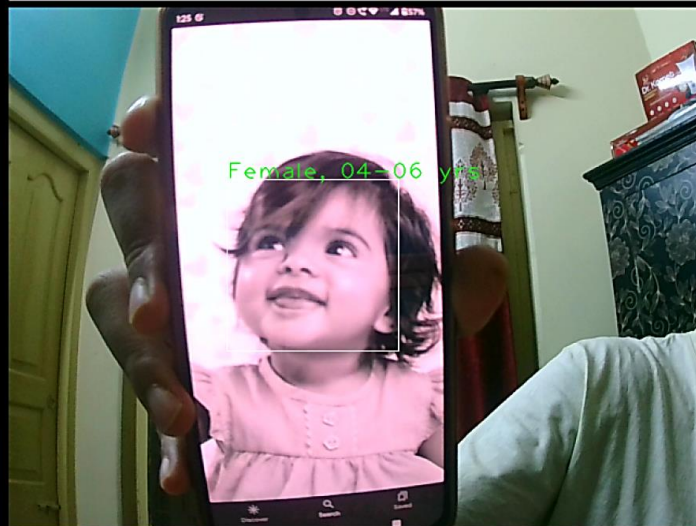
Stop Camera



FPS: 1.55

Gender and Age Classification: Live Webcam

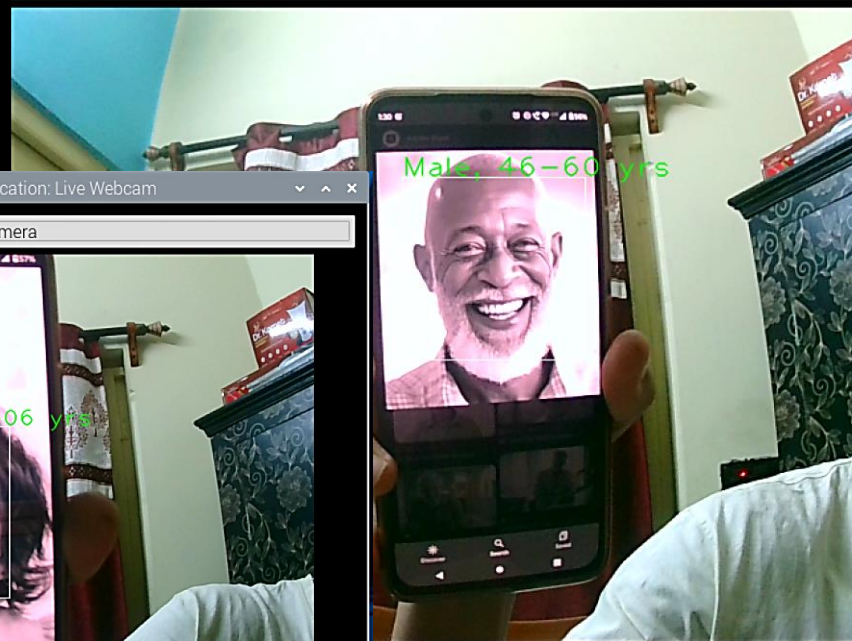
Stop Camera



FPS: 1.58

Gender and Age Classification: Live Webcam

Stop Camera





# Predictions on random faces from the Internet

Gender Classification: Live Webcam

Stop Camera



FPS: 2.32

Gender Classification: Live Webcam

Stop Camera



FPS: 2.33

# References

- ❑ **Bhattiprolu, S. (2023). *python\_for\_microscopists*. GitHub**  
**[https://github.com/bnsreenu/python\\_for\\_microscopists/blob/master/241\\_live\\_age\\_gender\\_emotion\\_detection/241\\_live\\_age\\_gender\\_emotion\\_detection\\_V2.0.py](https://github.com/bnsreenu/python_for_microscopists/blob/master/241_live_age_gender_emotion_detection/241_live_age_gender_emotion_detection_V2.0.py)**
- ❑ **Alexandru B, R. (2021). *Age-Gender-Classification-on-RaspberryPi4-with-TFLite-PyQt5***  
**<https://github.com/radualexandrub/Age-Gender-Classification-on-RaspberryPi4-with-TFLite-PyQt5>**



The background of the slide is a dark navy blue. It is decorated with two abstract network-like patterns. On the left side, there is a cluster of red nodes connected by thin red lines. On the right side, there is a cluster of light blue nodes connected by thin light blue lines. The nodes are small circles, and the lines are thin and slightly translucent. The overall effect is a modern, tech-oriented aesthetic.

# Thank You