

Submission folder contents

1. [Gender classification with Age estimation on UTK face](#) (Training Notebook 1)
2. [Age estimation Training Notebook – Second model](#) (Training Notebook 2)
3. [Inference Notebook](#) (for testing all models on UTK face, testing pipeline models, webcam live stream implementation on Colab)
4. Models **{All models are present to be used in Inference}**
5. [Finalised Pipeline Models](#) {based on Inference only models used in pipeline is given} size: **353MB**

Details Regarding Training Environment

The entire project was trained with Kaggle and Google Colab Notebook. **Please note that due to TensorFlow dependency issues, we had to use a older Runtime (Before Jan 30 2024) on Kaggle. [If you test the code on Kaggle we advise to do the same.](#)**

Session options ^

ACCELERATOR

GPU T4 ×2 ▼

Quota: 25:27 / 30 hrs

LANGUAGE

Python ▼

PERSISTENCE



No persistence ▼

ENVIRONMENT

Pin to original environment (2024-01-30) ▼

You won't get new packages, but your code is less likely to break. What is a notebook environment? ⓘ

Datasets

- ▶  merged-augmented-utk-faces-facial-age
- ▶  utkface-new

Merged and Augmented Dataset on UTK Face and Facial Age

This dataset contains 233K images with Age in their file name

<https://www.kaggle.com/datasets/skillcate/merged-augmented-utk-faces-facial-age-dataset/versions/2>

UTK Face new

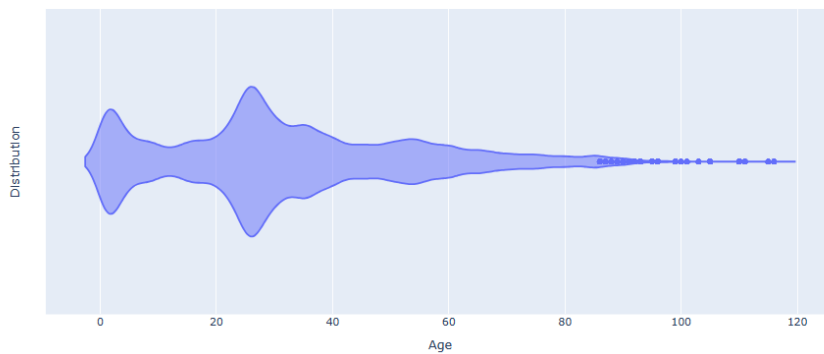
This contains 23k images with Age, Gender and Ethnicity encoded in the file name.

[UTKFace \(kaggle.com\)](https://www.kaggle.com/datasets/skillcate/merged-augmented-utk-faces-facial-age-dataset/versions/2)

Dataset Analysis

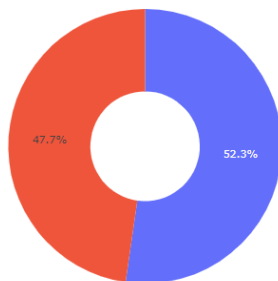
Merged and Augmented Dataset on UTK and Facial Age

Age Distribution

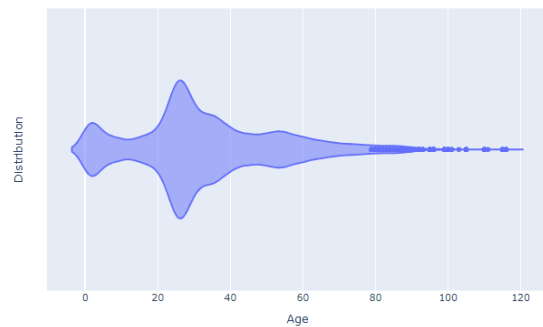


UTK Face Dataset

Gender Distribution (Donut Chart)



■ Male
■ Female

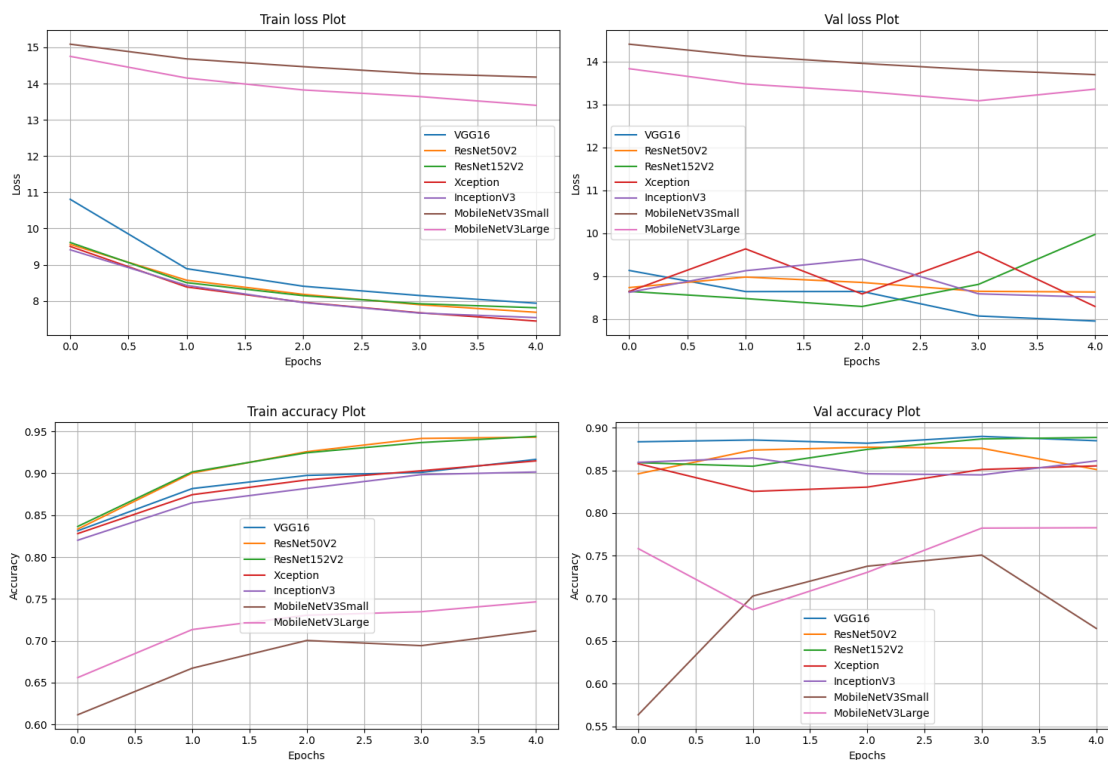


Reason for choosing two Datasets:

UTK Face as some issues like improper annotations, blurred or watermarked images etc., Also the Dataset is biased with age groups 25-35 yrs. To cater that the other dataset has more images below 10 yrs. Still, we expect the model to perform poorly on 10-20 yrs. and 60+ yrs. The male female ratio looks decent but that will be slightly biased to males. Though we expect good performances on this dataset, the overall real-world model performances might not be as accurate as expected with test dataset.

Exploratory Analysis on various pattern finding Model Backbones.

We went through Resnet50, Resnet150, Inception, Xception, Vgg16, Mobile-net Large and Small models with weights of Image-net for feature extraction from some research papers for classification and Regression.



The learning curve of Vgg16 looks better than Xception in Regression.

The Accuracy of Resnet 152V2 looks promising for Classification.

We ran few epochs on Resnet50 and Resnet152, VGG16, Xception for arriving to this conclusion.

We trained the models for 20 epochs with early stopping based and validation metrics and took the best model from last 5 epochs.

Training Notebooks:

- 1. Gender classification with Age Estimation on UTK Face**
- 2. Age Estimation on Merged and Augmented dataset**

If you train on Kaggle:

Use Runtime older than Jan 30 2024, Datasets are publicly available in it.

If you train on Colab or locally:

Download the UTK dataset and give the directory path in input as mentioned in the notebook

Overall Results:

VGG16 for age on UTK Face:

Training MAE: 6.1059

Validation MAE: 6.9152

Test MAE: 6.6553

Since the test set is only around 2K images we are getting data like this. We can say the average MAE is around 6.7. This motivated us for using the augmented dataset.

VGG 16 for age on Augmented Dataset on UTK and facial age:

Training MAE: 6.0149

Validation MAE: 5.9097

Test MAE: 5.9001

Here the results were very consistent and Test size is around 22K images which is fairly good enough benchmark.

Resnet 152V2 for Gender Classification on UTK Face:

Training accuracy: 0.8852

Validation accuracy: 0.8725

Test accuracy 0.87600

Here, the results were very consistent though the test size is small with 2K images.

Inference notebook

Inference notebook is expected to be run in **Google Colab** for using the webcam via Colab snippets.

Add all 3 Models in the **Models** to the notebook.

The models will be tested on 25% of UTK face dataset and results are provided.

Model metrics on 25% of randomized Entire UTK dataset in Inference notebook

Age VGG 16 UTK:	MSE: 74.6427	MAE: 5.8892
Age VGG 16 Augmented:	MSE: 67.2498	MAE: 5.6144
Gender Resnet 152V2 UTK:	Loss: 0.2610589	Accuracy: 0.88831

Finalizing the Models: *Entire pipeline designed model size: 353 MB*

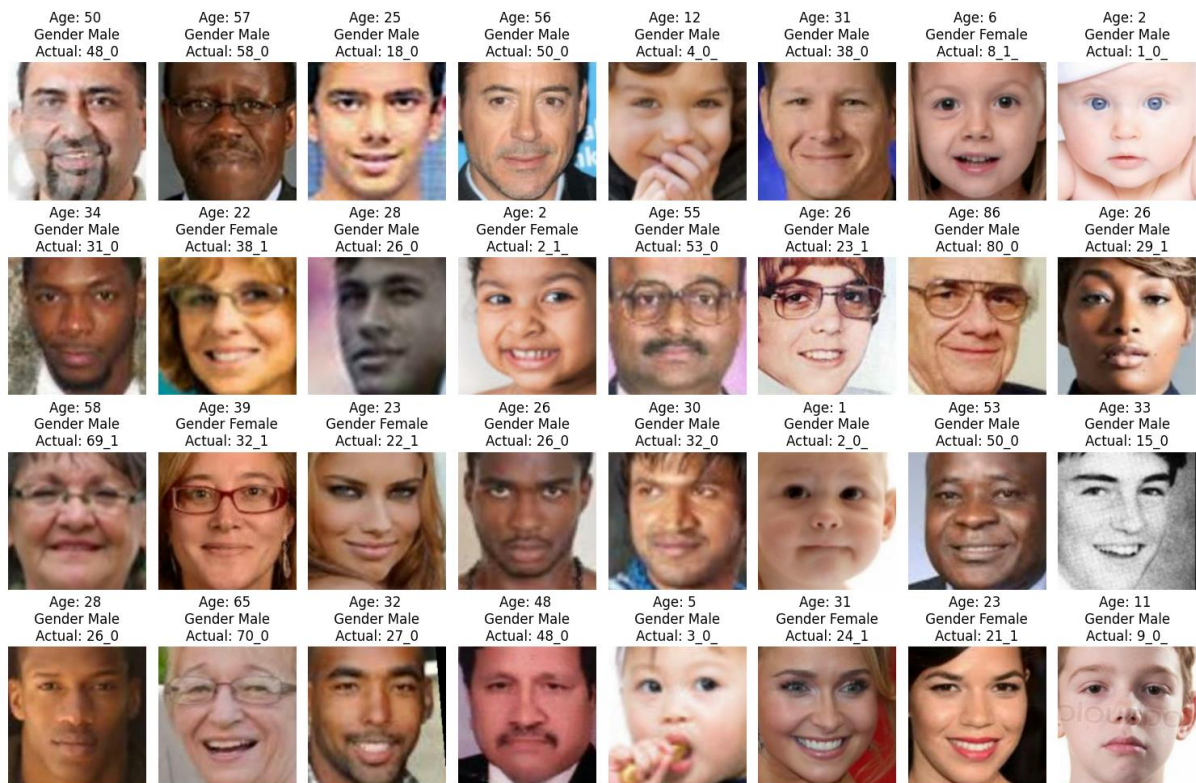
Age Estimation -> VGG 16 Backbone model trained on Augmented Dataset of UTK and Facial age (**129 MB**)

Gender Classification -> Resnet152V2 Backbone model trained on UTK Face original dataset (**223 MB**)

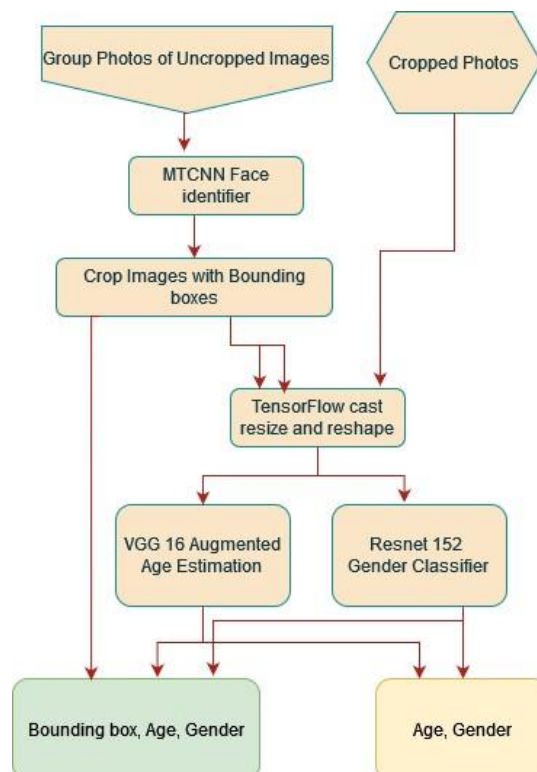
MTCNN -> In photo face detections we are using MTCNN for better accuracy

Harr Cascade -> MTCNN is sometimes slower, also produces logs which messes with the JS overlay used. So Harr Cascade is used for Webcam Video stream.

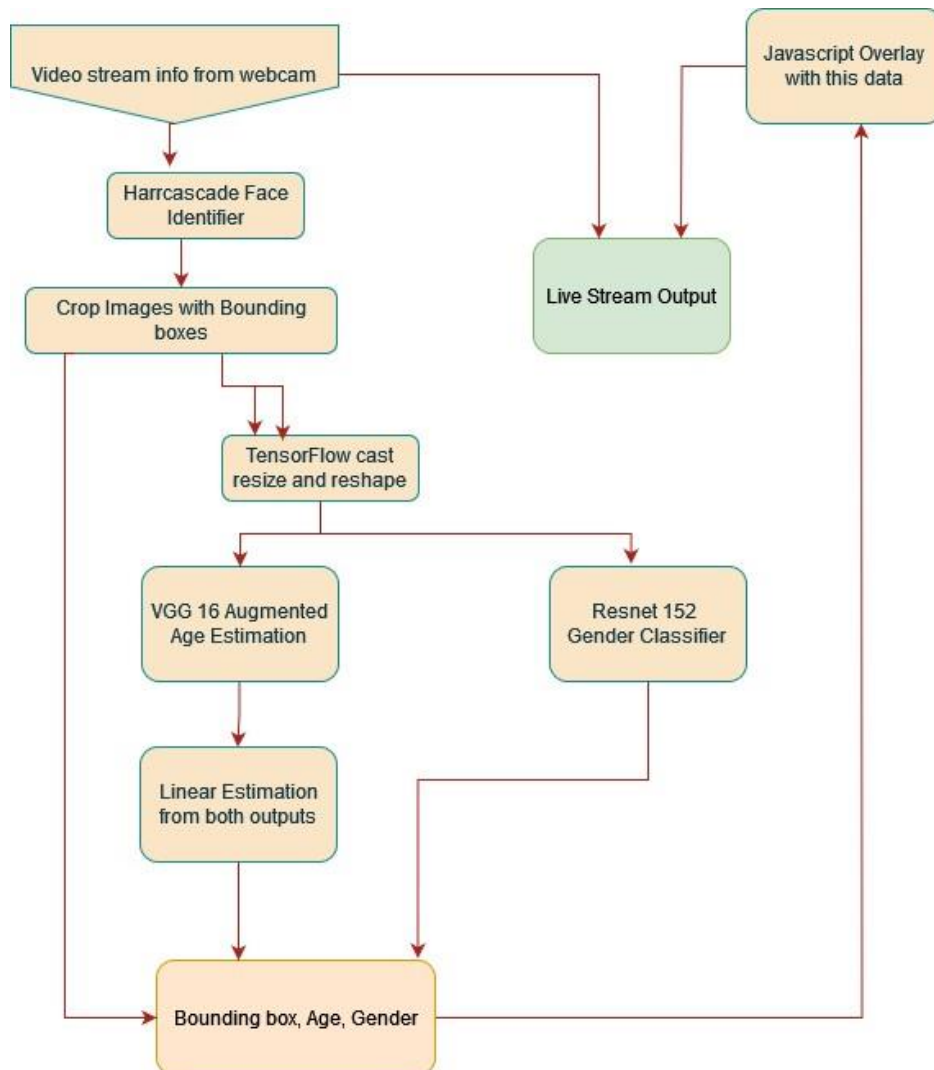
Visualizing Random sample images from the test set (Looks very promising)



Pipeline for Image Age and Gender Estimation



Pipeline for Live Stream Age and Gender Estimation



Functions provided for Evaluation:

- 1. find_age_gender_precropped(img_path)**
input: a string of the image file path
output: a tuple of predicted age (int), gender ["Male", "Female"]
- 2. find_age_gender_uncropped(img_path)**
input: a string of the image file path
output: a list (each element of the list is a list containing a bounding box list, age (int), gender ('Male' , 'Female'))
plot: it will also plot all the faces found with their age and gender in title

3. Video Stream cell

The last section contains code for Web cam stream.

Run the cells, it will start a web cam (give access if asked), the faces will be detected and bounding box with age, gender will be overlaid on the video which constantly updates. To exit you have to click on the video.

Guideline for getting good result in webcam:

- *If the person is far, his/her photo quality will be heavily degraded, since the prior output of webcam is already poor.*
- *Keeping the face in level with the webcam would give better results.*
- *Avoid bright light sources as much as possible.*
- *Avoid Bright light source in front of your face, which makes the model to be highly inaccurate.*
- *It can detect multiple faces also at same time, the results will be overlaid on the output video stream.*
- *Click on the video to stop the stream, otherwise the webcam will be on even if you cancel run of the cell.*

Final results summary:

Age model: **MSE:** 67.2498 **MAE:** 5.6144 **Size:** 129MB

Gender model: **Loss:** 0.2610589 **Accuracy:** 0.88831 **Size:** 223 MB

