

**Deep Learning**

***Facial Emotion, Age and Gender Detection from Images***

***Team number: 8***

*Siva Yogitha Mokkapati*

*Venkat Nihaal Akula*

**School of Graduate Professional Studies**

Data Analytics

DAAN 897 – Deep Learning

Spring II, 2020

**Document Control**

**Work carried out by:**

|  |  |  |
| --- | --- | --- |
| **Name** | **Email Address** | **Exhaustive list of Tasks** |
| Siva Yogitha Mokkapati | Sqm6303@psu.edu | Data collection, preprocessing, transfer learning and model building for emotion detection |
| Venkat Nihaal Akula | Vxa5088@psu.edu | Data collection, preprocessing, transfer learning and model building for age and gender detecttion |
|  |  |  |

**Revision Sheet**

|  |  |
| --- | --- |
| **Date** | **Revision Description** |
| 4/25/2020 | Preparing the data |
| 4/30/2020 | Building models |
| 5/5/2020 | Changing models as per required |
| 5/10/2020 | Prediction file |
| 5/20/2020 | Apply transfer learning |
| 5/25/2020 | Compare models |

**TABLE OF CONTENTS**

[1 Introduction 3](#_Toc33129402)

[2 Problem Statement 3](#_Toc33129403)

[2.1 Challenges 3](#_Toc33129404)

[2.2 Related Works 3](#_Toc33129405)

[2.3 Importance 3](#_Toc33129406)

[3 Data Collection 3](#_Toc33129407)

[4 Data Preprocessing 3](#_Toc33129408)

[5 Methodology 4](#_Toc33129409)

[6 Results 4](#_Toc33129410)

[7 Discussion of Results 4](#_Toc33129411)

[8 References 5](#_Toc33129412)

# Introduction

Humans have the innate ability to distinguish faces, facial emotions, gender and to estimate age. Today because of the availability of GPU’s, hard disk and technologies we can make computers do a lot of work that humans do. This opens tons of applications. Face, emotion, gender and age recognition are often helpful in improving the security, criminal identification and in several facial applications like Snapchat and Instagram. Of these the most important one is the criminal identification where estimating the age and gender of a person helps to identify the criminal fast. Emotion detection in crime scene helps us know the criminal easily. Not always will humans be able to identify the gender and age correctly. Sometimes, humans can be manipulated by appearances, but computers cannot. Hence, the need for gender, age and emotion detection using computers.

# Problem Statement

To identify the emotion, gender and age from facial images.

# Challenges

* To construct an optimal model for Problem statement.
* To apply transfer learning and check if it is able solve the problem of emotion, gender and age with a good accuracy.
* To determine different class boundaries and its implications for age detection.
* Use of different datasets for emotion and age, gender detection.

## RELATED WORKS

We used two datasets for our project. Both the datasets were used in other projects. However, a combination of those two datasets are not used in any project because one contains array pixels to grey scale images while other contains RGB images. For the emotion detection dataset almost of the projects that used the dataset only performed emotion detection as there are only labels for emotions while we did not limit ourselves to just emotion detection. [1] DAGER published in 2017 presented how they were able to classify the gender age and emotion using CNN. In this paper they combined three datasets to train their model of which one is the FER dataset that we used for emotion detection. However, their methodology is different from ours as we they estimated an age instead of age group. Also, they did not combine the outputs like we did. [2] Another paper published in 2019 focused only on emotion detection. In [3] SVM classifier is used for classifying the emotions after deep feature extraction. The uniqueness of our projects is the ability to use different datasets for emotion and gender, age detection and used their combination for predicting other images.

## IMPORTANCE AND IMPACTS

One of the important reasons is the various applications of this problem that range from security, to person identification, to human-computer interaction. The need for detecting emotion of a person in a crime scene can help us know who the criminal is easily. Also, emotion detection can help us know the Psychological state of the person. Gender and age detection can improve the Revenue of Snapchat and Instagram like applications by providing more facial options. There is need to build a model that can outperform human’s ability in recognizing and distinguishing human attributes because of these variety of applications.

# Data Collection

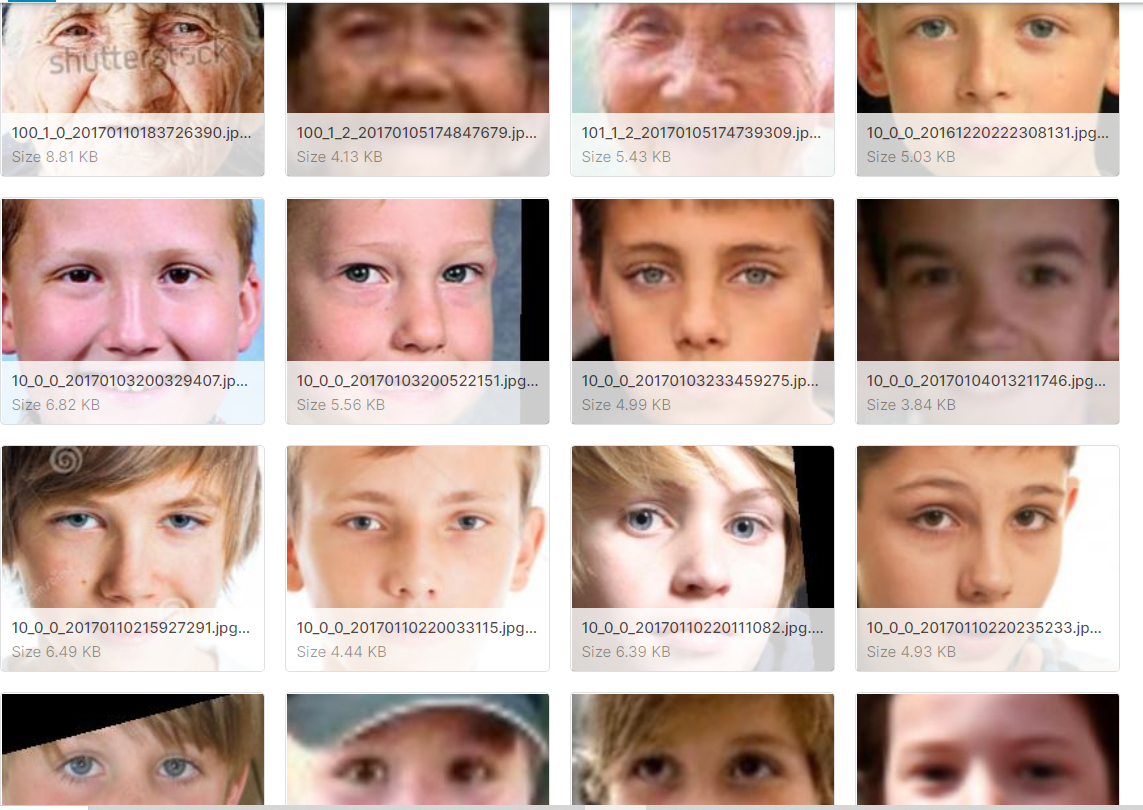
The dataset for emotion classification is taken from FER Challenge posted in Kaggle. The dataset has approximately 35000 records, with attributes image pixel array and emotion value.

<https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data>. The pixels converted to images looks like:



The dataset for age and gender classification is taken from:

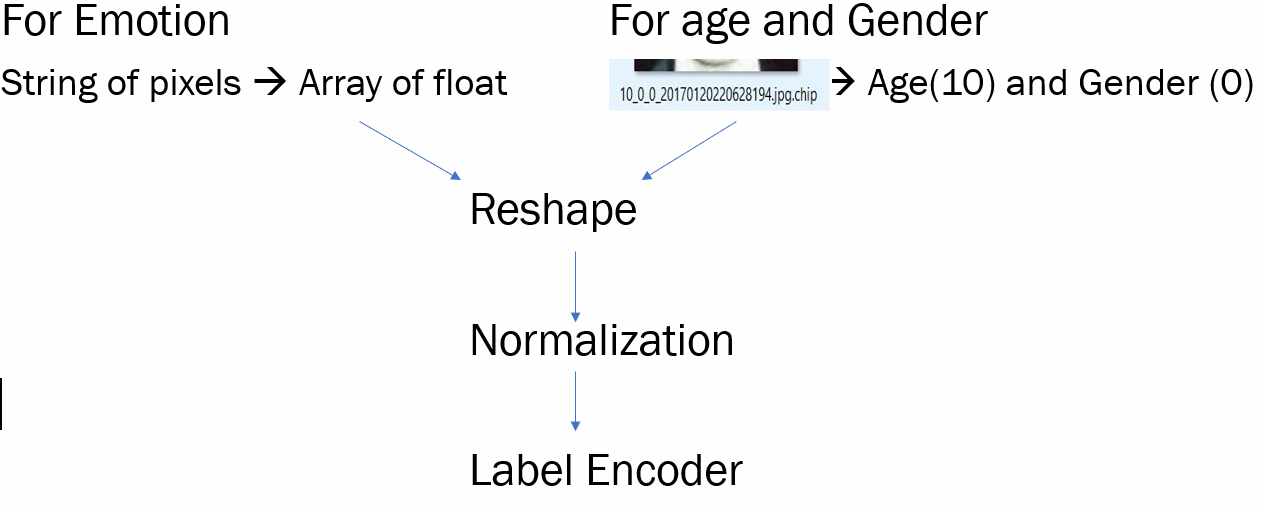
<https://susanqq.github.io/UTKFace/>. It has approximately 23000 images. The image looks like:



These facial images are used to train the model for identifying the facial emotions, gender and age.

# Data Preprocessing

* For emotion detection, the dataset contains image pixels stored in string format in a data frame. So, the pixels should be separated and converted to float format. For gender and age dataset the image name contains the age and gender. Hence, they should be extracted from the image name. Later, for both the datasets the arrays are reshaped as per required. The image arrays are flattened and normalized (dividing by 255). The labels are converted to categorical variables using label encoder. Arrays are augmented while fitting the model like width shift, zoom, height sift, rotation, flip etc.



# Methodology

* We have initially reshaped the images dataset into required format of (48,48,1) and (48,48,3) for emotion classification model and age and gender classification model respectively.
* For emotion classification model we have built a basic dense layer neural network model initially and improvised it by adding more complex convolution layers and additional parameters until an optimal model is obtained.
* For application of transfer learning methodology, we have used VGG model which resulted in a less accurate model.
* For Gender and Age classification we have first used a simple CNN model with 32,64 and 128 kernel sizes , later we have added different parameters and different kernel sizes to find an optimal model for Age and Gender prediction respectively.
* Choice of model for Emotion classification:
  + Why do you choose this model?

We have built a complex Convolution Neural Network model as its best approach for retrieving information stored in images and to make accurate class prediction.

* + Type of Neural Network (CNN)
    - Cost functions: Categorical Cross entropy
    - Activation functions: Elu , Softmax(output)
    - Pooling Layers: Max Pooling(2D)
    - Optimizer: Adam
    - Regularization: Drop out, Batch Normalization
    - Kernal initializer: he\_normal
* Choice of model for Age classification:
  + Why do you choose this model?

We have built a complex Convolution Neural Network model as its best approach for retrieving information stored in images and to make accurate class prediction.

* + Type of Neural Network (CNN)
    - Cost functions: Categorical Cross entropy
    - Activation functions: Relu , Softmax(output)
    - Pooling Layers: Max Pooling(2D)
    - Optimizer: Adam
    - Regularization: Drop out, Batch Normalization
    - Kernal initializer: he\_normal
* Choice of model for Gender classification:
  + Why do you choose this model?

We have built a complex Convolution Neural Network model as its best approach for retrieving information stored in images and to make accurate class prediction.

* + Type of Neural Network (CNN)
    - Cost functions: Categorical Cross entropy
    - Activation functions: Relu , Softmax(output)
    - Pooling Layers: Max Pooling(2D)
    - Optimizer: Adam
    - Regularization: Drop out, Batch Normalization
    - Kernal initializer: he\_normal
* Strategy to validate models
  + We have divided the entire data set into train and validation in 80:20 ratio, accuracy and loss values are calculated and plotted in between train and validation data set for each epoch of the model. This process was used for all the models built for Age, Gender and Emotion classification.
  + Transfer learning VGG model was built for emotion classification as additional metric for model comparison.
  + The only limitation we faced was to have imbalanced dataset for different class label which was handled using ImageGenerator for Emotion classification.
* How good are they in solving your problem?
* All the three optimal models that were selected have an individual accuracy of 70%.
* All the models also have an overall average precision, recall and f1 score of 0.6 per class individually
* Our models Vs traditional machine learning
* We can conclude that the CNN models we have built can process information from both low resolution and high-resolution images more accurately when compared to traditional models.
* Further, Our CNN models can only get better with only time by training with more and more varieties of images thus further enhancing accuracy from 70%.
* Assumption/constraints considered while applying the method on data.
* The constraint that are faced while building emotion classification model was the input was trained on greyscale image, thus when an RGB image is fed to predict module it must be rescaled. Thus, has a chance of losing some details.
* The constraint that are faced while building Age classification model are determination of different class boundaries.
* Software/[Database] used
* Anaconda Navigator.
* Jupyter Notebook.
* Anaconda Prompt.
* VS code

# Results and INterpretation

* Accuracy and Loss values are plotted for all the model generated for emotion, age and gender classification
* Plots of Models for Emotion classification:
* Model 1: Basic Dense layer model.



* The above model is an underfitting model.
* Model 2 : Simple CNN model



* The above model is an overfitting model.
* Model 3 : Complex CNN model (**OPTIMAL MODEL**)



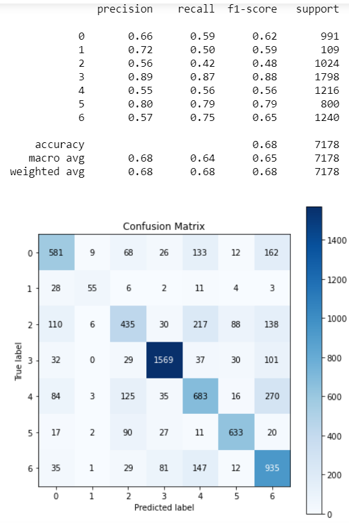
* This is our optimal model with training accuracy and validation accuracy of 73% and 68% respectively.
* Model 4 : Complex CNN model with only top 3 emotions



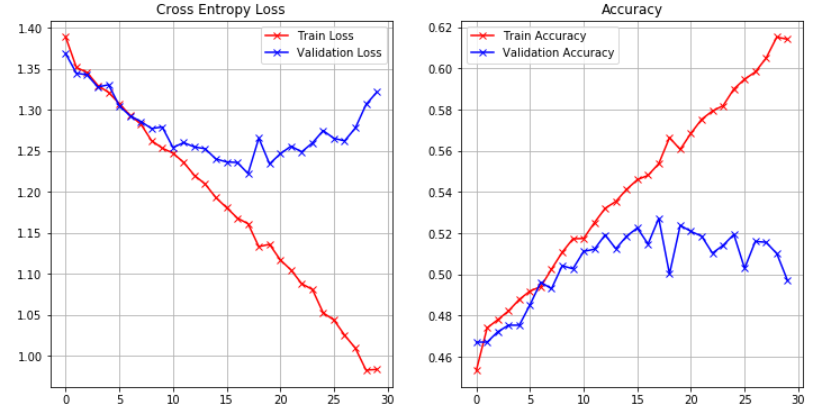
* An optimal model was built for only top 3 emotions.
* Model 5 : VGG model



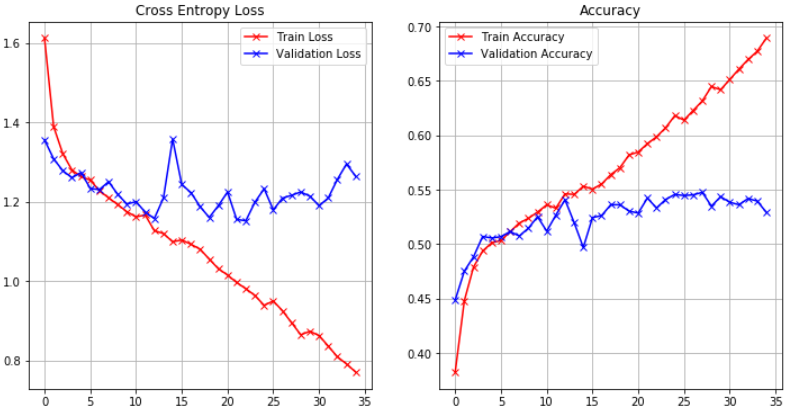
* Transfer learning was applied to construct a VGG model but an underfitting model was built.
* Confusion Matrix of our optimal model:



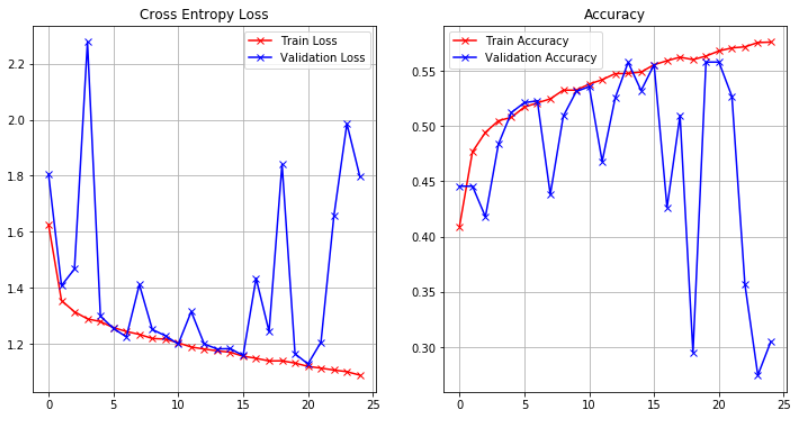
* The above confusion matrix shows the precision, recall and f1 score values.
* Plots of Models for Age classification:
* Model 1: Basic CNN model.



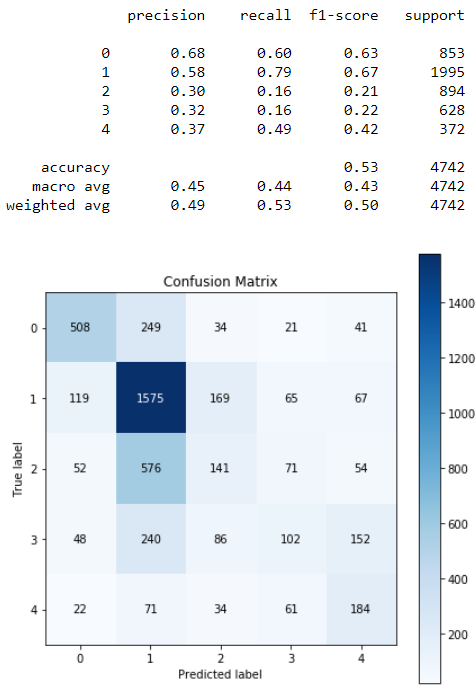
* The above model is an overfitting model
* Model 2: A complex CNN model was built (**Optimal Model**)



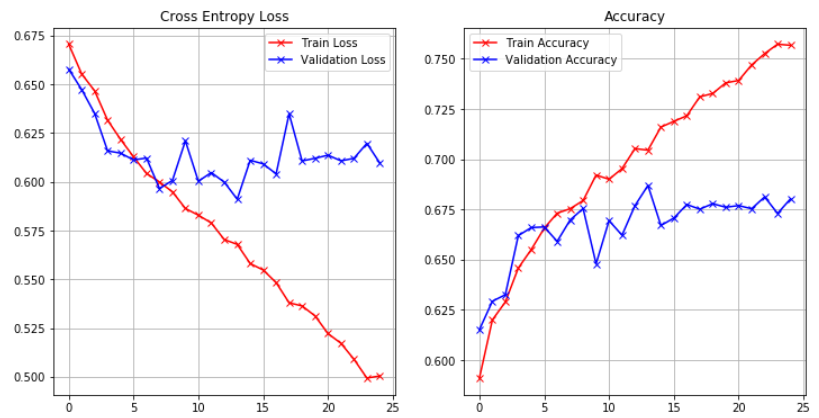
* This is our optimal model with training accuracy and validation accuracy of 63% and 54% respectively.
* Model 3: A complex CNN model is built with ’Elu’ as activation function



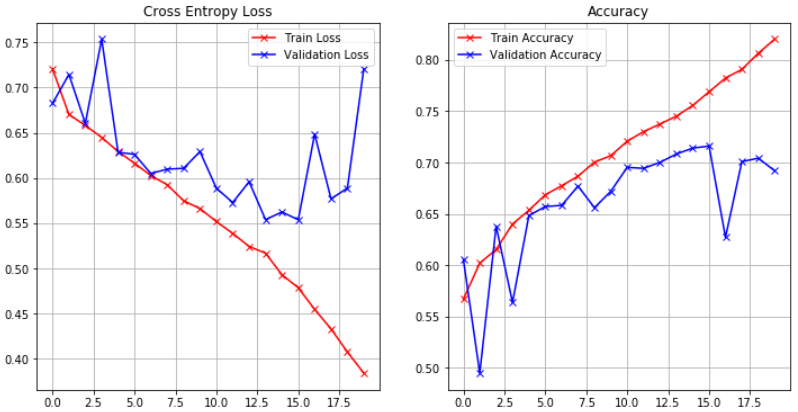
* The above model is an underfitting model.
* Confusion Matrix of our optimal model:



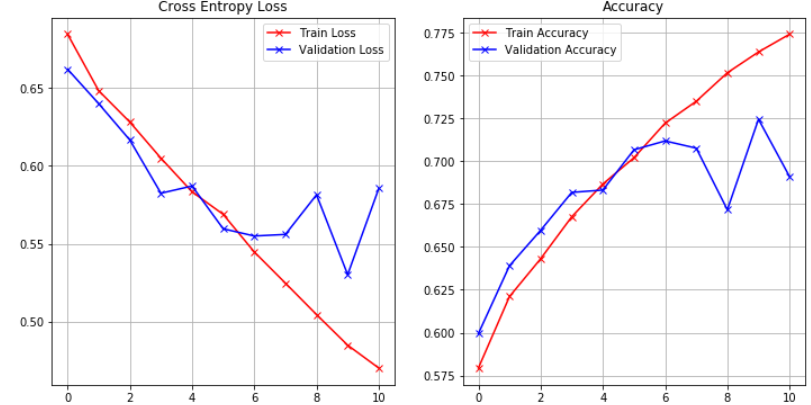
* The above confusion matrix shows the precision, recall and f1 score values.
* Plots of Models for Gender classification:
* Model 1: A basic CNN model with more convolution layers



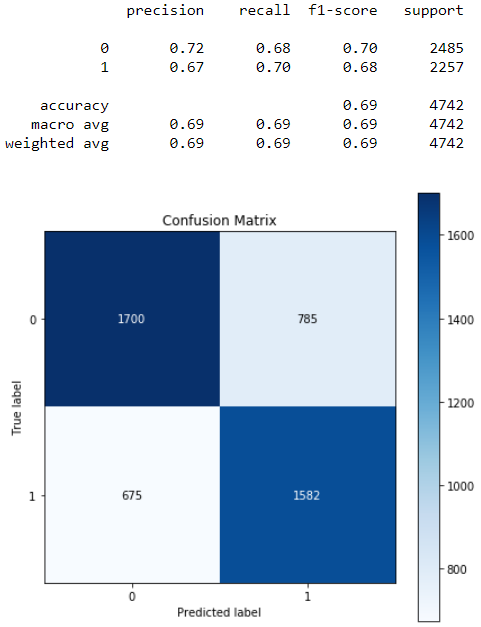
* This is an underfitting model.
* Model 2: A simple CNN model with bigger kernels of (96, 256, 384) (**Optimal Model**)



* This is our optimal model with training accuracy and validation accuracy of 83% and 70% respectively.
* Model 3: A basic CNN model with less convolution layers



* This is an over fitting model and attained accuracy of 83% at 10th epoch.
* Confusion Matrix of our optimal model:



* The above confusion matrix shows the precision, recall and f1 score values.

# Discussion of Results

* **Outputs:**

|  |  |
| --- | --- |
|  | A person wearing a suit and tie smiling at the camera  Description automatically generated |
|  | A person smiling for the camera  Description automatically generated |
|  | A person smiling for the camera  Description automatically generated |
|  | A person wearing a suit and tie  Description automatically generated |
|  | A person in a blue shirt  Description automatically generated |
|  | A person wearing a suit and tie smiling at the camera  Description automatically generated |
|  | A person in glasses looking at the camera  Description automatically generated |
|  | A person wearing a suit and tie  Description automatically generated |

* The combined optimized model built can be used to identify age, gender and emotion from a given image. We can get an accuracy of 73 % approximately for the 3 attributes combined. However, there are some images being misclassified in any one of the attributes. Because VGG model accepts only RGB images, we had to broadcast the array to convert gray scale to RGB image leading to the poor performance of transfer learning. We need to build more complex models in order to achieve high performance. The future scope of this project includes the use of more pre trained models, consider more datasets of different sizes, colors and to build a graph model using them.

# Your Feedback

This project helped us understand many deep learning concepts. The practical part helped us better understand how to use pre trained models (transfer learning), how to build our own models, how to tune the parameters etc. Also, there were more resources we could have tried many other models and compare them.

# References

1. Dehghan, Afshin & Ortiz, Enrique & Shu, Guang & Masood, Syed. (2017). DAGER: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Network.
2. Özdemir, Mehmet & Elagoz, Berkay & Alaybeyoglu, Aysegul & Sadighzadeh, Reza & Akan, Aydin. (2019). Real Time Emotion Recognition from Facial Expressions Using CNN Architecture. 10.1109/TIPTEKNO.2019.8895215.
3. Bargal, Sarah & Barsoum, Emad & Ferrer, Cristian & Zhang, Cha. (2016). Emotion recognition in the wild from videos using images. 433-436. 10.1145/2993148.2997627.
4. Levi, Gil & Hassncer, Tal. (2015). Age and gender classification using convolutional neural networks. 34-42. 10.1109/CVPRW.2015.7301352.
5. Tal Hassner, Shai Harel, Eran Paz, Roee Enbar. Effective Face Frontalization in Unconstrained Image.
6. Slides and notebooks uploaded in canvas.
7. <https://riptutorial.com/keras/example/32608/transfer-learning-using-keras-and-vgg>
8. <https://towardsdatascience.com/predict-age-and-gender-using-convolutional-neural-network-and-opencv-fd90390e3ce6>
9. <https://www.learnopencv.com/age-gender-classification-using-opencv-deep-learning-c-python/>
10. <https://www.researchgate.net/publication/225173459_Classification_of_Face_Images_for_Gender_Age_Facial_Expression_and_Identity>
11. <https://www.geeksforgeeks.org/image-classifier-using-cnn/>