# Gender and Age Classification

...

# Overview

- Introduction
- Dataset & Pre-processing
- CNN for Age and Gender Classification
- Block Diagram
- Procedure
- Challenges Faced
- Experiments & Results
- References

# INTRODUCTION

# Introduction

- The CNN we have implemented consists of three convolution layers and two fully connected layers and final output layer.
- SELU and ReLU are the activation functions that we have been using.
- Before feeding input the images are center cropped through preprocessing to make sure that the face region pixels are more concentrated.
- The implementation of the network is deep networks compared to the existing networks. So over fitting of the data can be prevented.
- The dataset used for training the network contains images which are not constrained i.e., there will be pose variation and motion blur in the image.
- Prediction of age will be done by distinguishing among EIGHT classes and gender among TWO classes.

# DATASET & PREPROCESSING

# DataSet :

- All the images are captured in the wild environment.
- There are 8 labels for the age as follows (0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60-)
- The Audience benchmark dataset comprises of about 29K photos.
- Gender labels are mentioned for the images along with the age labels.

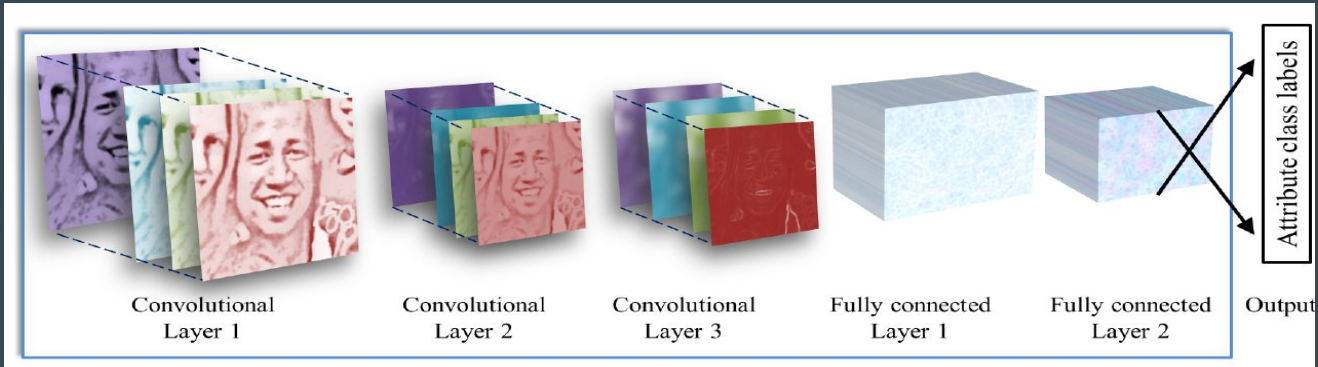
# Preprocessing :

- Data Augmentation which is strong combination of crop, flip, bright, contrast, scale, rotate.
- Data Rotation is in Specific Range.
- Scaling is in range of 0.8 to 1.2 along both directions of x and y.
- Brightness is under the constraints of max delta with 0.5.
- Upper bound and Lower bound values of contrast are under 0.5 & 1.5 respectively.

# Age & Gender classification using CNN

# DESCRIPTION of CNN:

- The network comprises of only three convolutional layers and two fully connected layers with a small number of neurons along with output layer for classification.
- We have used Max-Pooling Layer for downsampling, Local Response Normalization Layer for normalizing, Dropout Layer for reducing overfitting and Soft-Max Layer to perform non-maxima suppression.
- Below is the diagramatic representation of the network.



- Since Images are in 256 x 256 size, we have cropped into227 x 227 size.

- FIRST CONVOLUTIONAL LAYER:
  - In the first convolutional layer, the image is multiplied 96 times with 3 x 7 x 7 size filter. Later, it is passed through SeLU Activation Function, followed by a max pooling layer of 3 x 3 size with stride value as 2 and a lrn layer.

- SECOND CONVOLUTIONAL LAYER:
  - In the second convolutional layer, the previous layers output is 96 x 28 x 28 is multiplied 256 times with 96 x 5 x 5 size filter. Later, it is again passed through SeLU and a max pooling layer with same parameters as above along with local response normalization.

- THIRD CONVOLUTIONAL LAYER:
  - The output of previous layer is multiplied 384 times with 256 x 3 x 3 size filter. Later on, it is passed through SeLU Function and a max pooling layer.

# DENSE LAYERS :

- The first and foremost fully-connected layer which has 512 neurons gets input from third convolutional layer and later it is passed through SeLU and followed by a dropout layer.
- The second fully-connected layer undergoes changes same as first fully-connected layer.
- The third fully-connected layer classifies into the output. The output from third FC Layer is passed through Soft-Max Layer to perform Non-Maxima Suppression.

# Block Diagram

# Procedure

# TF Records :

- To feed the input for tensorflow, we use an input pipeline which takes a list of filenames (any supported format), shuffle them (optional), create a file queue, read, and decode the data.
- We have created 8 tf records for gender, age and age & gender individually. We have used 6 for training and 2 for testing.
- We have used inbuilt functions to read and decode the data.
- We can retrieve the label and image from those inbuilt functions which are nothing but training images and training labels respectively.
- For testing, we do same as above for testing image and testing labels.

# Initialization :

- Classes of age: (0,2) is 0, (4,6) is 1, (8,12) is 2, (15,20) is 3, (25,32) is 4, (38,43) is 5, (48,53) is 6 and (60,100) is 7.
- We have made some assumptions and rounded the classes for age.
- Classes for gender are labeled as shown here m is 0 and f is 1.
- The weights are initialized using pre-trained VGG Face network weight.

- In view of limiting the risk of the main challenge called Overfitting.
- We used dropout learning with a ratio of 0.5 (50% chance of setting a neurons output value to zero).
- Data Augmentation mentioned above along with cropping the image of 256 x 256 into 227 x 227.
- So that we can achieve multiple crop variations and mirror images in different Variations.
- Training itself is performed using Adam Optimizer with image batch size of 64 images.
- The initial learning rate is 0.005.

# ACTIVATION FUNCTIONS:

- The activation functions we have used are ReLU(Rectified Linear Unit) and SeLU(Scaled Exponential Linear Unit). RELU seems to be doing a much better job than SELU.
- But to our model, we have gained nearly 2% higher accuracy than ReLU. So, to propaganda our higher accuracy, we used to say SeLU as our Activation Function.
- To be genuine, it is possible that SELU is good in some configurations.

# CHALLENGES FACED

- Data-Preprocessing is an long winded task.
- Computational Time is higher for Training.
- Many Images don’t have gender and age labels.
- Some Images are not labeled as class, instead given as Integer.
- Age Detection is not easy considering many factors. So We have been classifying the image.
- We performed pre-processing steps in order to get multiple variations and pose variations.
- Images are center cropped for better feature learning so that only face pixels are used for processing.

# ACHIEVEMENTS

- Building a CNN model for classification of gender and age with higher accuracy than state-of-the-art models.
- We implemented the paper Age and Gender Classification using Convolutional Neural Networks by Gil Levi and Tal Hassner.
- We have registered the highest accuracy of 91% in gender and 55% in age using VGG pretrained weights and SeLU Activation function.
- Working of ReLU and SeLU.
- Using pre-trained weights has really helped us in reaching to the efficient output.

# EXPERIMENTS & RESULTS

| | 0-2 | 4-6 | 8-13 | 15-20 | 25-32 | 38-43 | 48-53 | 60 | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| MALE | 745 | 928 | 934 | 734 | 2308 | 1294 | 392 | 442 | 8162 |
| FEMALE | 682 | 1234 | 1360 | 919 | 2589 | 1056 | 433 | 427 | 9411 |
| BOTH | 1427 | 2162 | 2294 | 1653 | 4897 | 2350 | 825 | 869 | 19487 |

THE ADIENCEFACES BENCHMARK

| Architecture | Framework | Accuracy Observed |
|---|---|---|
| Proposed Architecture | Tensorflow | 50.2 +/- 1.2 |

AGE ESTIMATION RESULTS ON ADIENCE BENCHMARK

| Architecture | Framework | Accuracy Observed |
| --- | --- | --- |
| Proposed Architecture | Tensorflow | 86.1 +/- 4.7 |

GENDER ESTIMATION RESULTS ON ADIENCE BENCHMARK

| Framework | Hardware | Time For Training |
| --- | --- | --- |
| Tensorflow | Intel GPU | 10 hours |

TIME TAKEN FOR TRAINING

# References :


1. Gil Levi, Tal Hassner: Age and Gender Classification using Convolutional Neural Networks in CVPR, 2015
2. A. Krizhevsky, I. Sutskever, and G. E. Hinton: Imagenet classification with deep convolutional neural networks in Neural Inform. Process Syst., pages 10971105, 2012.
3. A. Dehghan, E. G. Ortiz, G. Shu, and S. Z. Masood. Dager: Deep age, gender and emotion recognition using convolutional neural network arXiv:1702.04280, 2017.

# The Team