Underage Driving Detection - Age Recognition Using Face Detection

Y. Angeline Christobel[a], and K. Keerthiga[b]

[a]
Dean, School of Computational Studies, Hindustan College of Arts & Science
 Padur, Chennai
[b]Hindustan College of Arts & Science, Padur, Chennai

Abstract: Facial recognition with age variation is one of the challenging tasks. Artificial Intelligence age predictions can be used in many fields such as smart human-machine interface growth, health, electronic commerce. Age may be a soft biometric trait which aids enforcement in identification of several crime and victims of underage group. Prediction of people's age accurately from their facial images is an ongoing active problem. A statistical pattern recognition approach for solving this problem is proposed in this paper. Convolutional Neural Network (CNN), a deep learning algorithm, is used as extractor of features. CNN requires less processing than other classification algorithm. In this paper, face images of individuals have been trained with convolutional neural network and age with high rate of success has been predicted. The images cover wide range of poses, facial expression, lighting, occlusion and resolution. In recent years, the causality of traffic accidents caused by minor aged people driving have been gradually increasing. There are several serious injuries and damages due to increase in major accidents. Therefore, in this study, age detection using deep learning was developed to alert and prevent these large-scale disasters using facial recognition technology.

1. Introduction

According to recent survey, many lives are shattered by underage drivers going for a joyride in the vehicles of their parents or relatives. People are aware of many unlicensed drivers going for the short drive – but often overlook the nature of the risks and the consequences of these youngsters. Teen drivers aged 16-19 years are fourfold more likely than older drivers to crash. The aim is to predict the age of individuals using a face image.

Facial analysis gained recognition in the computer vision community in the past. Human's face contains features that determine identity, age, gender, emotions, and the ethnicity of people. Age predictions from an unfiltered real-life face are yet to meet the requirements of real-world applications. Regardless of the progress computer vision community keeps making continuous improvement of the new techniques that improves the state of the art, but still has room for improvement. Different methods have been proposed to solve this classification problem. Here, the age classifications task is formulated as a classification problem in which the CNN model learns to predict the age from a face image. Motor vehicle accidents remain the leading cause of death for teenagers, accounting for nearly 41% of fatalities among the age groups of young people ranging from 13 to 19 years. Despite of the deadly toll results and with the lack of driving experience, it still reflects the fact that the teenage brain remains a piece ongoing. Teenagers are prone to risk taking, impulsive behaviour and sensation seeking all of which causes trouble. To overcome and prevent this, the goal is to develop an Artificial Intelligence and machine learning technology to help us keep teens safe by predicting their age by their faces and preventing them from driving. The aim of this paper is to propose and experimentally evaluate an automated system that would be able to predict the age of an individual from facial images.

This paper is organized as follows: Section 2, the studies conducted by other researchers will be presented and Section 3, methodology to detect age will be explained. Section 4 presents experimental results and discussion. The conclusion will be stated in section 5.

2. Literature Review

In [1], S. U. Rehman et al. introduced a new method for automatic age estimation of face images. Initially they introduced a Gabor wavelet transform for age estimation to achieve real-time and fully-automatic aging feature extraction. In their work, CNN handles multitasking of Facial detection and emotional classification.

The prediction of age, gender and ethnicity on the East Asian Population using a Convolutional Neural Network (CNN) was explored by N. Srinivas et al. [2].

N. J ain et al. [3] introduced a hybrid deep CNN and RNN (Recurrent Neural Network) model. This model improves the overall result of face detection. The dataset used in this model are MI Facial Expression and JAFFE.

A convolutional network architecture was proposed by G. Levi et al. [4] that classifies the age with small amounts of data. The Audience Benchmark was used to train the model.

S. Turabzadeh et al. [5] proposed a system in which a real time automatic facial expression was designed. The proposed work was implemented and tested on an embedded device which could be the first step for a specific facial expression recognition chip for a social robot.

A. Dehghan et al. [6] presented an automated recognition system for age, gender and emotion that was trained using deep neural network.

A. Krizhevskyetal. [7] presented a work that segregates 1.2 million images into 1000 different categories with the help of a deep convolutional neural network. The results obtained proved that supervised learning can deliver exceptional accuracies.

In [8], Y. H. Kwon and N. da Vitoria Lobo proposed methods for age estimation based on calculating ratios between different measurements of facial features (eyes, nose, mouth, chin) that are localized and ratios between them are calculated by their sizes and distance measures. Using hand-crafted rules, face is classified into different age categories.

3. Methodology

Today's existing system provides less support for preventing underaged driving. This is due to human efficiency of inspecting every single face every minute is not possible. Also, it is difficult for a human to predict ages from just the face. But it is not the case for a machine, a machine which is trained with better algorithms inspects every face, extract features, apply neural network to predict ages and highlights underaged people.

Age detection is the process of automatically discerning the age of the person from a photo of their face or from live video. Here age detection is implemented in two stages:

Stage 1: Detect faces in input image/video stream

Stage 2: Extract the face Region of Interest and apply age detector algorithm to predict the age of the person

For Stage 1:

Any face detector that is capable of producing bounding boxes in an image can be used. In this study, DNN Face Detector is used. It is supported Single-Shot-Multibox detector and uses ResNet-10 Architecture as the backbone. The model is trained using images available from the web. OpenCV provides 2 models for this face detector.

1. Floating point 16 version of the original caffe implementation.
2. 8-bit quantized version using TensorFlow.

Caffe model implementation is used in this study.

The method has the following merits:

- Most accurate compared to other face detector methods
- Runs at real-time on CPU
- Works for different face orientations-up, down, left, right, side-face, etc.
- Detects faces across various scales

For stage 2:

Once face detector has produced the bounding box coordinates of the face in the image/video stream, move to stage 2.

Stage 2: Identifies the age of the person. For detecting the age, caffe model is used. When using OpenCV's deep neural network module with Caffe models, there are two sets of files need to be used:

1. The .prototxt file which defines the model architecture.
2. The .caffemodel file which contains the weights for the actual layers
3. Dataset used is Adience. This dataset serves as a benchmark for face photos and is inclusive of various real world imaging conditions like lighting, pose, and appearance. It has a total of 26,580 photos of 2,284 subjects in eight age ranges and it is about 1GB in size. The models used have been trained on this dataset. Figure 1 below shows the methodology diagram.
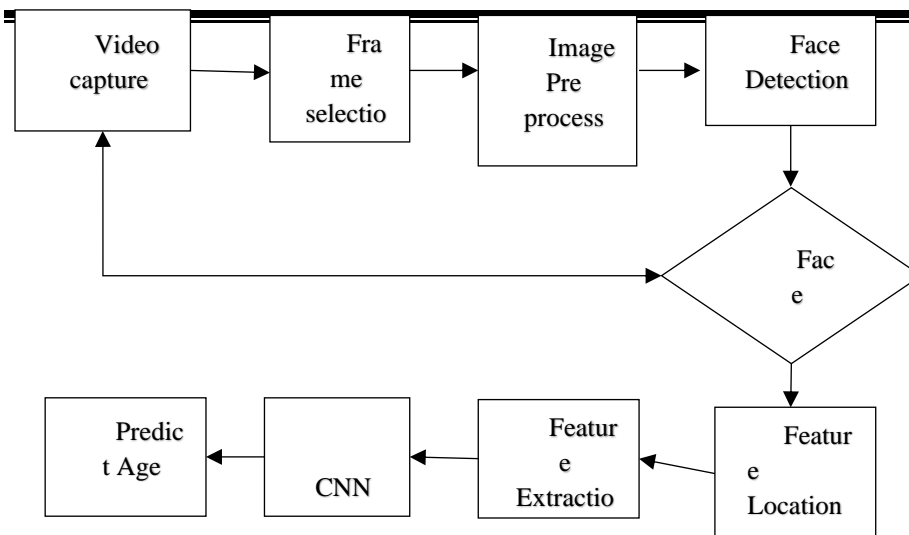
Figure 1: Methodology Diagram

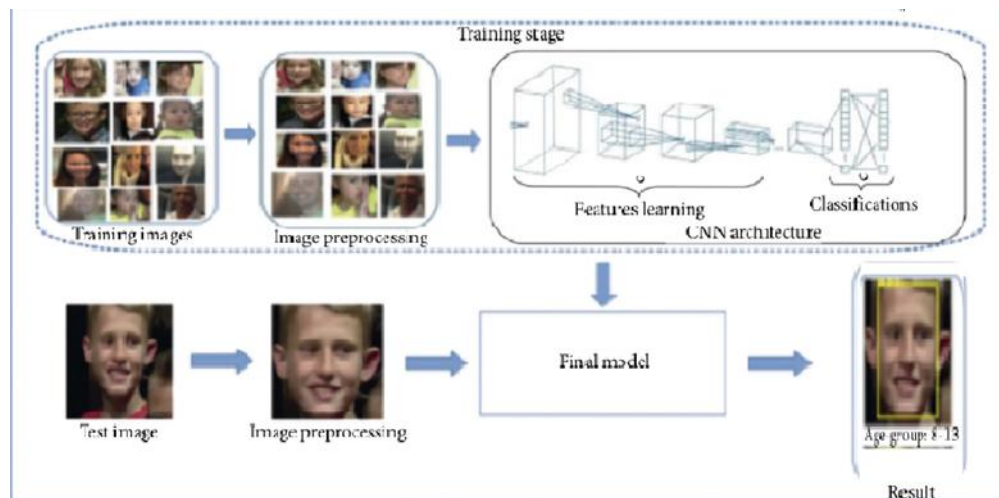The figure 2 shows the process diagram.

Figure 2: Process Diagram

3.1 Algorithms

3.1.1 Face Detection: OpenCV DNN Module

DNN-Deep Neural Network helps to detect face from an image or video streams. It is a Caffe model which is predicated on the Shot-Multibox Detector (SSD) and uses ResNet-10 architecture as its backbone. It introduced post OpenCV 3.3 in its deep neural network module.

3.1.2 Age Recognition: CNN Model

In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the main categories to do image recognition, image classifications, object detections, face recognition, etc., are some of the areas where CNNs are widely used. CNN image classification takes an input image, processes it and classifies it under certain categories. Computer sees an input image as array of pixels depending on the image resolution. Based on the image resolution, h x w x d (h = Height, w = Width, d = Dimension) will be seen.

In deep learning CNN models are trained and tested in such a way that when each input image will pass through a series of convolution layers with filters (Kernals), pooling, fully connected layers (FC) and finally applies SoftMax function to classify an object with probabilistic values between 0 and 1.
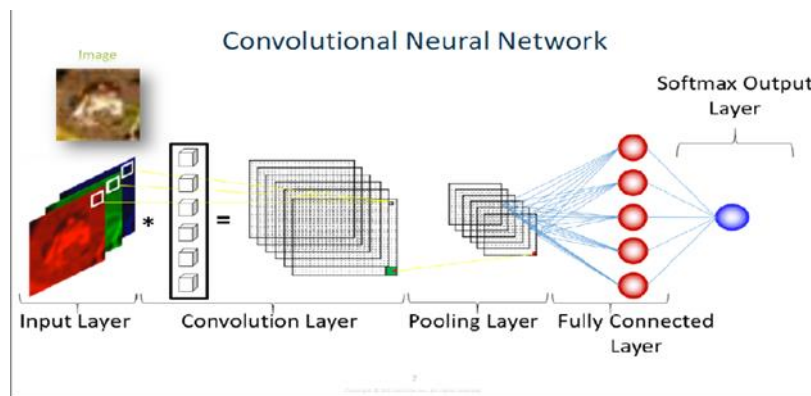
Figure 3: CNN Diagram

Input Layer

Input layer stores the raw pixel values of the image. The input layer has three colour channels (R,G,B).

Convolution layer

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel. The dot product is computed by sliding the filter over the image at a certain stride. Stride is the number of pixels, slide over each time when the filter is moved. The dot product creates a new matrix called the convolved feature or feature map.

Non-Linearity (ReLU)

ReLU stands for Rectified Linear Unit for a non-linear operation. The output is $f(x)=\max(0, x)$.

ReLU's purpose is to introduce non-linearity in our convNet.

Pooling Layers

Pooling layers are commonly place between successive convolution layers to simplify the output from a convolutional layer. A pooling layer down samples to reduce the dimensions.

Types of Pooling Layers:

Max Pooling

Average Pooling

Sum Pooling

Max pooling takes the largest element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements in the feature map call as sum pooling.

Output Layer

The Output Layer is a fully connected layer that is associated with a particular loss function that computes the prediction error. Being a fully connected layer, the neurons in the output layer connect to all of the activations in the previous layer.

The SoftMax function takes a vector of scores and transforms it to a vector of values between 0 and 1 that sum to 1. Thus, applying SoftMax as the activation function ensures that the sum of output probabilities from the fully connected layer is 1.

Age recognition using OpenCV's is quite popular. In this paper OpenCV's DNN module is used for face detection and CNN model for age group recognition.

3.2 Dataset

The accuracy of CNN design is tested using the Adience dataset [9], designed for age and gender classification. The Adience set consists of images automatically uploaded to Flickr from smart- hone devices. Because these images were uploaded without prior manual filtering, as is typically the case on media webpages (e.g., images from the LFW collection [10]) or social websites (the Group Photos set [11]), viewing conditions in these images are highly unconstrained, reflecting many of the real-world challenges of faces appearing in Internet images. Adience images capture variations in head pose, lightning conditions quality, and more. It has a total of 26580 photos of 2,248 subjects in eight age ranges.

Figure 4: Data Set Model

3.3   DATA PRE-PROCESSING

Facial image pre-processing may not be necessary if the source is similar to a standard passport photograph. However, facial images in-the-wild may have characteristics such as various pitch/roll/yaw angles, multiple subjects per image, background noise, varying image size and quality, etc. Such photos require image pre-processing and normalisation to align and remove unnecessary features. Before passing an image in the network, pre-processing is done by resizing the image to the size it was trained. After resizing the image, normalization is done.

Figure 5: Data Pre-Processing

To Detect the face, first loaded a pretrained model to the code. The detecting devices like web cam or any other camera device will look for face and check with model then display the age group of a person.

4. Experimental Results and Discussion

Here are few test cases of hypothesis of what problems that the technology could possibly encounter. Given below are few images with different perceptive of faces views.

Test Case1: Front view

In figure 6, the front view of a person is shown. As seen in the figure, the age of the individual is predicted in the age group range of 25-32 years with 96% accuracy.
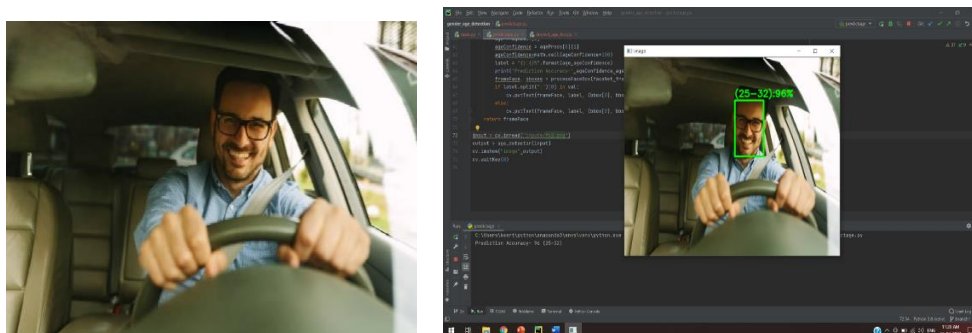

Figure 6: Front view

Test case2: Side view

In test case 2, the side view of a young individual is given to predict the age. The age is predicted in the age group range of 8 to 12 years. The side view gives the accuracy of 41%. The accuracy is less because in the audience dataset, the age range below 8 to 12 is not trained and this image falls below 8 years of age.
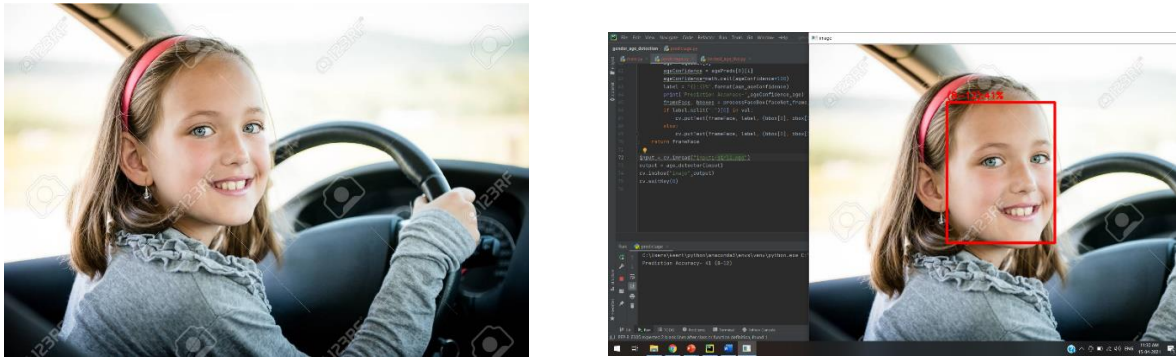
Figure 7: Test case 2 – side view

Test Case 3: High-resolution image

In test case 3, the goal is to check the age of a high-resolution image. The high-resolution image gives the accuracy of 82% and age is predicted in the age group range of 8 to 12 years.
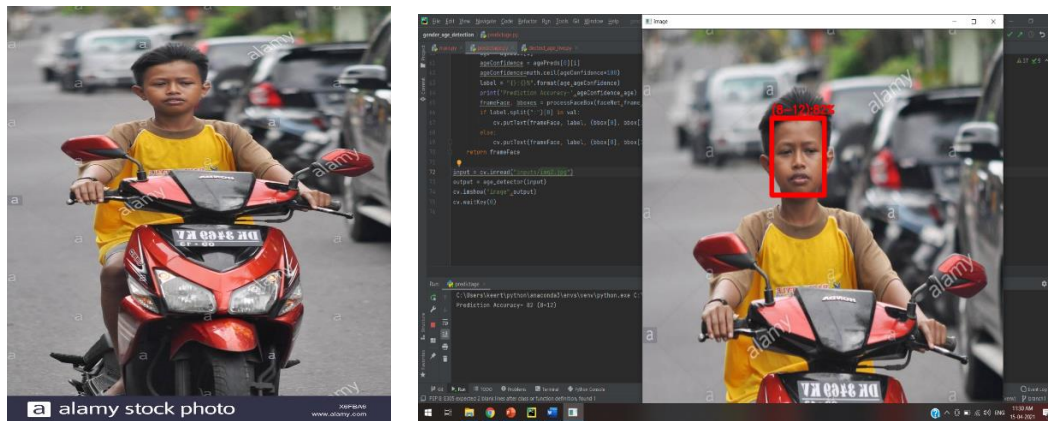

Figure 8: Test Case 3-High resolution image

The testing with different views gives more accurate prediction. To detect the underage driving persons, this methodology will be helpful and accidents can be reduced.

5. Conclusion

Predicting the age of an individual with just the face is a challenging process as there are number of factors that determine how old anyone appears, also including their lifestyle, job, habits and importantly genetics. Generally, people like to hide their age. So, when a human struggle to predict someone's age, even a machine learning model will also struggle as it follows algorithms built by man. Therefore, it is necessary to asses all age prediction results in terms of perceived age rather than actual age. This study is designed to detect and prevent underaged drivers from driving. It is developed and tested with test cases on the design. Overall, the accuracy of the model is decent but can be improved by using more data, and better network architectures.

References

1.   S. U. Rehman, S. Tu, Y. Huang, and Z. Yang, Face recognition: A Novel Un-supervised Convolutional Neural Network Method, IEEE International Conference of Online Analysis and Computing Science (ICOACS), 2016.

2.   N. Srinivas, H. Atwal, D. C. Rose, G. Mahalingam, K. Ricanek, and D. S. Bolme, Age, Gender, and Fine-Grained Ethnicity Prediction Using Convolutional Neural Networks for the East Asian Face Dataset, 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), 2017.

3.   N. Jain, S. Kumar, A. Kumar, P. Shamsolmoali, and M. Zareapoor, Hybrid Deep Neural Networks for Face Emotion recognition, Pattern Recognition Letters, 2018.

4.   G. Levi, and T. Hassner, Age and Gender Classification Using Convolutional Neural Networks,IEEE Workshop on Analysis and Modeling of Faces and Gestures (AMFG), IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Boston, 2015.

5.   S. Turabzadeh, H. Meng, R. M. Swash, M. Pleva, and J. Juhar, Realtime Emotional State Detection From Facial Expression On Embedded Devices, Seventh International Conference on Innovative Computing Technology (INTECH), 2017.

6.   A. Dehghan, E. G. Ortiz, G. Shu, and S. Z. Masood, Dager: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Network, arXiv preprint arXiv: 1702.04280, 2017.

7.   A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet classification with deep convolutional neural networks, Communications of the ACM, vol. 60, no. 6, pp. 8490, 2017

8.   Y. H. Kwon and N. da Vitoria Lobo. Age classification from facial images. In Proc. Conf. Comput. Vision Pattern Recognition, pages 762–767. IEEE, 1994.

9.   Dataset downloaded from Kaggle website:- https://www.kaggle.com/age-groupclassification- with-cnn

10.   G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face

recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts,

Amherst, 2007.

11.   A. C. Gallagher and T. Chen. Understanding images ofgroups of people. In Proc. Conf. Comput. Vision PatternRecognition, pages 256–263. IEEE, 2009