Real-Time Age, Sex and Emotion Detection (Using Deep Learning)

Ritesh Kumar Dubey
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Shivam Kripashankar Singh
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Tanmay Gupta
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Prakhar Kumar
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Project Guide: Dr. Pawan
*Associate Professor, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

ABSTRACT

In this paper, we describe a real-time default system that estimates age and gender using a set of face sequences from the video camera. Age and gender equation system consists of four steps: i) detection and removal of facial expressions in the input video; ii) selection of images of the front face from the extracted face areas using the shape of the head; iii) duplicate face detection and facial removal; and iv) age and gender equality using statistics facial features. Here, LBP features with AdaBoost components are used to locate the surface area video frame, and front-facing images are selected using the 3D pose simulation method. In addition, Particle-based tracking system is used to remove duplicate faces and to improve condition census accuracy, and Gabor-LBP features are used to measure age and gender using SVM line and AdaBoost categories. In tests, a large number of facial data sets are used for training and to evaluate the proposed method, and the maximum performance is achieved by age and gender average: 72.53% of age and 98.90% gender.

Keyword: - *Facial Characteristics, Age, Emotion and Gender, Face Recognition, CNN, AdaBoost, Gabor-LBP, real-time, Artificial Intelligence, Machine Learning, Deep Learning.*

1. INTRODUCTION

Facial features play an important role in understanding and recognizing emotions. An important factor in human interaction normalization of facial features and body language. In 1971, Friesen and Ekman pointed out that facial features are related worldwide with specific emotions. Even animals do the same types of muscle movements as human beings belong to a certain attitude, in spite of their racial background, education, birth, etc. Therefore, when properly modelled, this is often the case can serve as a very useful feature in the human machine interaction. In this paper, we propose a real-time default system for age and sex ratio pictures. To achieve this, face detection methods and posture measurement methods are used to detect priorities facial images. Then, LBP and Gabor filters

are used to extract facial features, and supervised reading. The Adaboost method is used to measure age and gender. Ensuring the operation of proposed method, tests are performed using a measurement website.
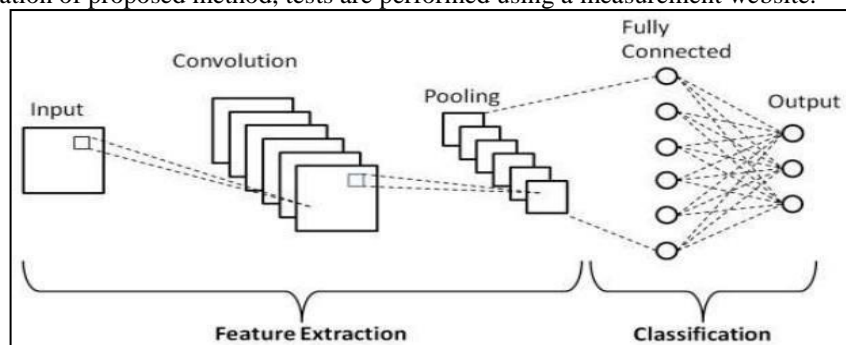
Fig -1: CNN Architecture

The function of distinguishing emotions, feelings on a person's face is divided into seven classes, namely: "anger, disgust, fear, happy, sad, surprised and neutral". Finally, we try to predict either Male or Female among them according to gender categories task. All these tasks can be performed on a single surface like many faces in one frame. Our full real time facial recognition, emotional isolation and sexuality separation takes less than 0.5 seconds. FER2013 and IMDB are two data sets we used the function of the separation of emotions and sex in order. The FER 2013 data set contains information about 35,887 grayscales, 48x48 7-dimensional face pictures, labelled as: "Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutrality" with 28709 and 7178 training photos confirmation photos namely 80/20 split. The IMDB database is too large Face data containing data for various artists. This the data set contains approximately 470,000 images. This database provides mat file as a metadata containing several features such as these face school, second face school, gender and age for all picture. Images with only one front face have more face points, while the image that contains the most faces has one small facial points. The second face points show that for sure the image contains a second face. Otherwise we take only those pictures with one face and too many faces before. To find out, we select only the images they contain face points ≥ 3. Finally, the IMDB is divided into an 80/20 ratio of training and validation set.
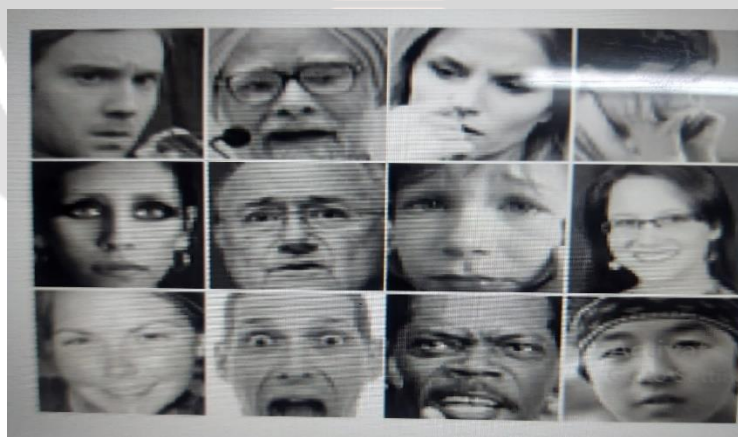
Fig -2: Various Gray images

1.1 Gender and age segregation

Gender and age segregation are facial features which play a vital role in social interaction. Age and sex segregation is important function in numerous application with more information, for example access controls, low authorizations, visual effects and many more.

1.2 Facial Expression detection

Emotion or expression detection on face images can communicate and impact individual's feelings. It is intriguing and to basic comprehend the feelings passed on by the images and how the visual content of the image implies expression.

2. LITERATURE SURVEY

Md jashim uddin Dr. paresh Chandra Barman, khandaker Takdir Ahmed et al. [1] proposed a detection system using CNN model which can achieved 95% accuracy rate in age, gender detection with IMDB-WIKI dataset and 66% accuracy achieved in emotion detection with FER dataset.

Thakshila R.kalansuriya and Anuuja T. Dharmaratne et al.[2] proposed an age gender detection system Using artificial neural network which achieved 70.5% accuracy rate. where FERET and FGNET datasets are used.

M.R Dileepa and Ajit Dantib et al. [3] proposed an age gender prediction system using Neural network and sigma control limit which achieved 95% accuracy rate.

Sepidehsadat et al. [4] suggested that the utilization of Gabor filter will make it simpler for the network to focus on the face, because the output direction is perfectly matched with facial wrinkles, and wrinkles will become the input of CNN. The network focuses on providing useful features with 7% age accuracy and 2% gender accuracy.

Imane et al. [5] proposed a face detection system using HAAR cascade, CNN based on FER13(KNN) for normalization and emotion detection and Gabor binary pattern histogram sequence is used for pattern scanning. The model use for different machine learning technique (SVM, KNN, random forest and classification, regression tree) and uses KNN and SVM algorithm to obtain a high accuracy rate which reaches 70%.

Rajesh et al [6] Here proposed real- time emotion detection used CNN with 9 layers for training and categorizing 7 different types of emotions which gives an accuracy of around 90.

The emotion and gender detection system using CNN proposed by Md Abdullah-al-Imran et al. [7] which achieved for emotion detection 66% with FER emotion recognition dataset and with IMDB-WIKI dataset achieved 95% for gender detection.

3. METHODOLOGY

We propose a collective architectural model of two CNN models, one is a Mini-Xception and the other is a simple CNN model for 4 layers. We combine these two models using "Measurement". The Mini-Xception architecture was developed by Francois Chollet, creator of the Keras library. This is a deep CNN architecture that contains depth differentiate convolutions. A centralized global integration algorithm is used in this model so that it does not become completely dependent a connected layer of training parameters. Depth wise Separable Convolutions consists of two types:
    (1) Deep Transformation (2) Direct Transformation.
The main purpose of using these layers is to distinguish between the cross-linking of the channel and the local relation. This is the first layer uses D $\times$ D on each M input channel and beyond N $1 \times 1 \times$ M conversion filter to combine M numbers of input channels and N numbers of output channels. This the layer reduces the total calculation compared to normal convolutions by $1 / N + 1 / D2$ Why Merge? The collection is machine learning a model that combines the predictions of two or more models. Predictions made by different models can be combined uses statistics, such as mode or description, or other complexities methods. There are two main reasons for using the collection model over one model is as follows:
    1. Functionality
    2. Strength
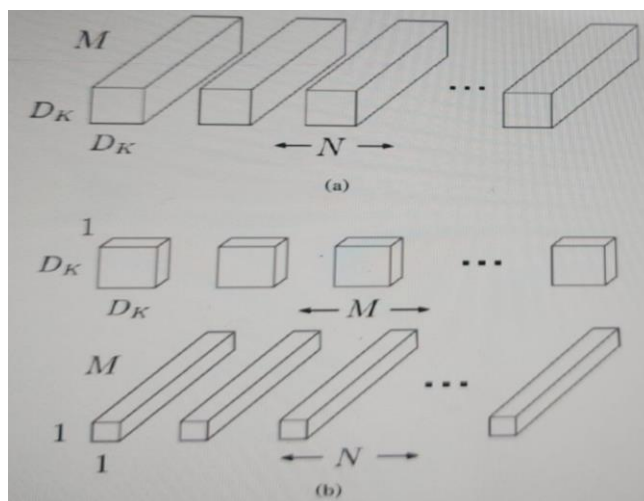
Fig -3: (a)Simple Convolutions (b) Depth wise Separable Convolutions

3.1 Functionality

An integrated model can make better pre dictions and result in better performance than any other donor model.

3.2 Strength

Collection reduces the spread of predictions and model performance. Why Rate? The middle layer reduces the contrast factor in the final model of the neural network causing a decrease in the spread of model performance to find out more confidence in predicting model outcomes. Through the work of separating emotions, we trained our Ensemble model 100 times and gender partition function we have trained only the mini-Xception model (part of our integration model) 100 times. We used Adam optimizer as an optimization algorithm and cross section Loss of entropy as a function of loss. We tested models with appropriate verification conditions for our emotional and gender-segregation function in FER-2013 and the IMDB data set respectively. Because for training purpose we used CPU Intel i5 7th generation @ 2.50GHz and 8GB NVIDIA GEFORCE 1050Ti GPU.

4. EXPERIMENTAL RESULTS

Consequences of our real-time feelings and gender segregation hidden facial functions are detected. All real-time pipeline involves: facial recognition, emotional separation and one-step sex classification. This implementation can work in both groups and single images and live input provided by webcam. We look at our results in Table I and compare them with it other previous methods. We can see the ensemble model provides better performance. Model Architecture a framework for combining models to reduce variability in the final model of the neural network which reduces I spread to model performance to gain confidence in it model prediction. Figure represents general confusion the matrix of our Integrated model of emotional separation function. Figure shows Precision, Recall, F1-points and support metrics for each emotion class in the Fer-2013 database. Because to calculate model accuracy in FER-2013 and IMDB Database, Using accurate metrics. Feature Recognition the work is trained using the Integrated model and the accuracy achieve 68% in test images in the FER-2013 database. Gender the split function uses the Mini-Xception model and is achieved 95% accuracy of test images in IMDB database.
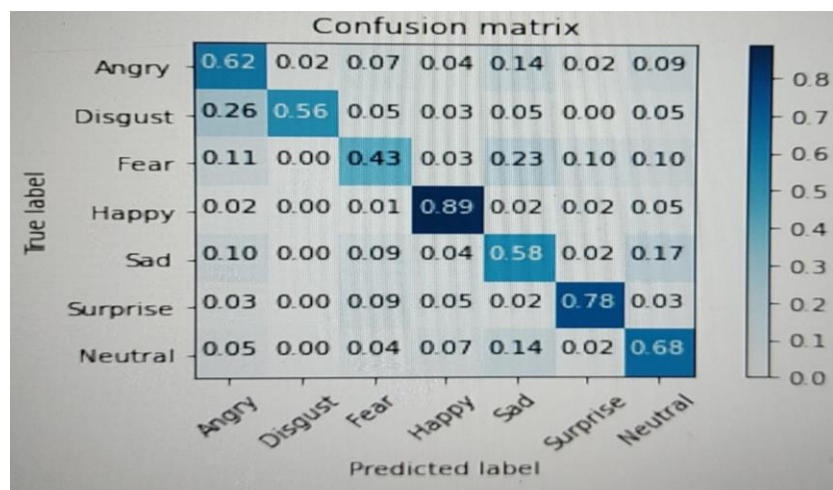
Chart -1: Normalized Confusion Matrix of our Ensemble model for Emotion classification task on FER-2013 dataset.

Emotions	Precision	Recall	F1-score	Support
Angry	0.62	0.62	0.62	958
Disgust	0.72	0.56	0.63	111
Fear	0.57	0.43	0.49	1024
Happy	0.86	0.89	0.87	1774
Sad	0.55	0.58	0.56	1247
Surprise	0.78	0.78	0.78	831
Neutral	0.62	0.68	0.65	1233
Accuracy			0.68	7178

**Table -1:** Results Matrix of our Ensemble model for Emotion classification task on FER-2013 Dataset

4.1 Emotion Classification

The proposed model can be used for standard partitions purpose while keeping the thoughts in real time. Eventually a real-time system was created using: facial recognition, emotions segregation and gender segregation into a single module.
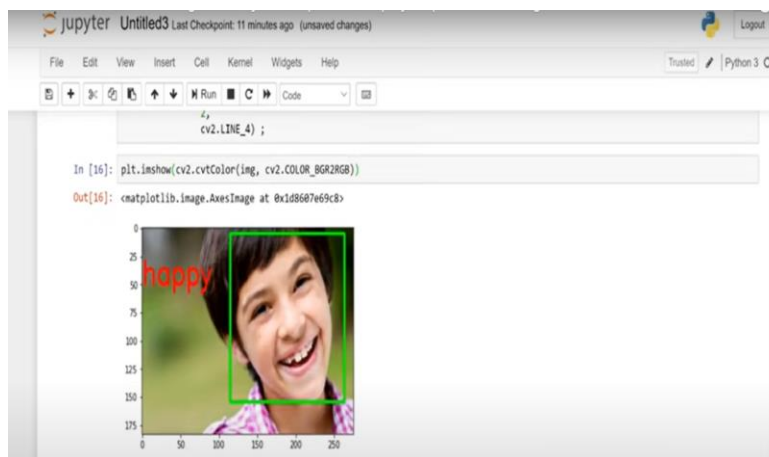
Fig -4: Single face emotion classification when image is passed as an input.
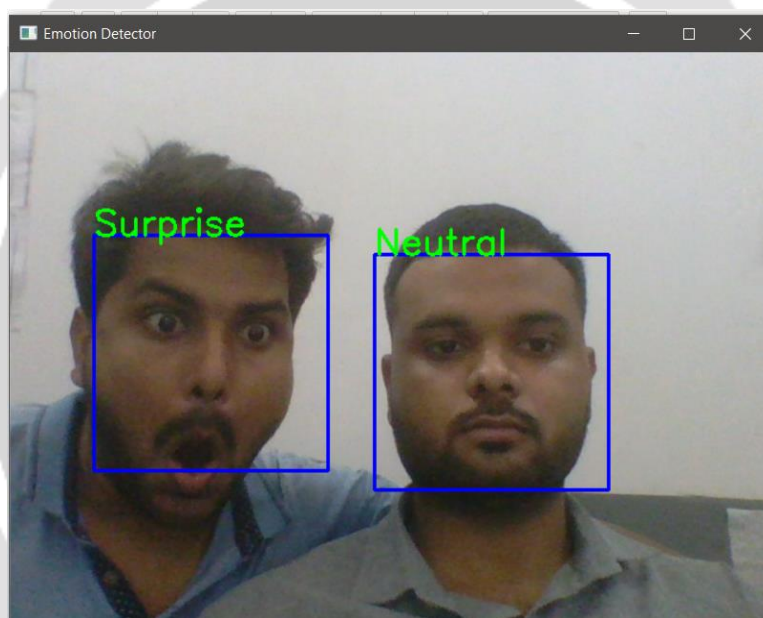
Fig -5: Multiple face emotion classification using Video feed as an input.

This model can detect emotions through analysis the facial expressions of the person shown on the webcam as input to model. Emotions on a person's face is divided into one of seven categories: Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutrality.

4.2 Age and Gender Estimation

After face and eye detection, a face image can be customized as an image of a fixed size local eye areas. Then, the Gabor-LBP histogram framework (Gao et al., 2008) is used extract the representative strength features of a local histogram (Lee et al., 2012). Use Filters from other Gaussian, the size of the facial features can be reduced by 60% compared with Gabor wavelet filters. Filters based on first and second Gaussian orders. Filter images from Gaussian sources are numerous translated by combining standard face images and filters from the Gaussian alternatives. Then, one by one The Gaussian output filter image is converted to an LBP map image using an LBP operator. The LBP patterns in each pixel can be summarized by summarizing the boundary values. measured by two forces, reflecting the local structure of the landscape image texture. Finally, each image of the LBP map is divided into separate regions with the previously defined drum size, and a total of 36 barrels are used for each LBP map image. Historical LBP histograms for sub-regions combined to make the final sequence of histogram as a representation of the face. In general, age estimation can be seen as a problem of multiple stages or setbacks, as well as gender equilibrium is

defined as a two-stage problem. Therefore, although we can use AdaBoost directly on to measure gender, using Adaboost in age estimates is not straightforward. For example, suppose we have an eight-year course. Then, training data can be divided into two subgroups repeatedly, and Adaboost can be trained as such that the two subgroups are well discriminated against. As shown in the picture, a total of seven divisions exist required to train age measurement data.

5. CONCLUSIONS

The proposed model can be used for standard partitions purpose while keeping the thoughts in real time. Eventually a real-time system was created using: facial recognition, age prediction, emotion and gender segregation into a single module.
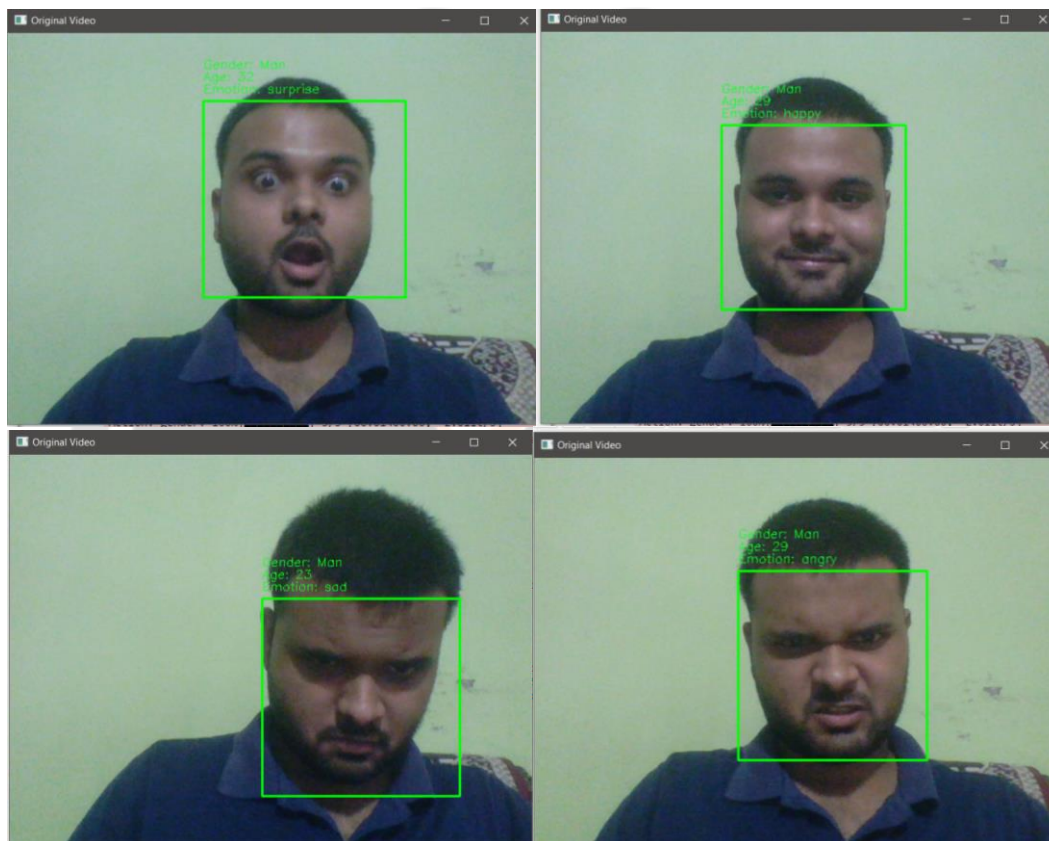
Fig -6: **Top Left:** gender-man, age- 32, emotion- surprise
Top Right: gender-man, age-29, emotion- happy
Bottom Left: gender-man, age-23, emotion- sad
Bottom Right: gender-man, age-29, emotion- angry

5. REFERENCES

[1]. Md. Jashim Uddin, Dr. Paresh Chandra Barman, Khandaker Takdir Ahmed S.M. Abdur Rahim, Abu Rumman Refat, Md Abdullah-Al- Imran6 "A Convolutional Neural Network for Real-time Face Detection and Emotion & Gender Classification" IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)

[2]. Thakshila R. Kalansuriya and Anuja T. Dharmaratne,"Neural Network based Age and Gender Classification for Facial Images" International Journal on Advances in ICT for Emerging Regions

[3]. M. R. Dileepa and Ajit Dantib "Human Age and Gender Prediction Based on Neural Networks and Three Sigma Control Limits" ISSN: 0883-9514 (Print) 1087- 6545 (Online) Journal homepage: http://www.tandfonline.com/loi/uaai203

[4]. 2018-Sepidehsadat Hosseini, Seok Hee Lee, Hyuk Jin Kwon, Hyung Il Koo Nam Ik Cho, "Age and Gender Classification Using Wide Convolutional Neural Network and Gabor Filter", IEEE2018.

[5] Imane Lasri, Anouar Riad Solh Mourad E Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network", IEEE- 2019.

[6] Rajesh Kumar G A, Ravi Kant Kumar Goutam Sanyal, "Facial Emotion Analysis using Deep Convolutional Neural Network",2017 International Conference on Signal Processing and Communication (ICSPC). doi: 10.1109/ cspc.2017.8305872, Pg.No.- 369 to374.

[7] Md Abdullah-Al-Imran "A Convolutional Neural Network for Real-time Face Detection and Emotion & Gender Classification'' e-ISSN: 2278-2834, p- ISSN: 2278-8735. Volume15, Issue 3, Ser. I (May - June2020), PP 37-46.