International Journal of Technology and Empirical Research

GNIT RESEARCH MANTRA Vol: 1, Issue: 1, September 2021

A Haar Cascade Classifier Based Approach for Detecting Objects

Srabani Kundu
Dept. of CSE
Guru Nanak Institute of Technology Kolkata, West Bengal, India
E-mail: srabani.kundu@gnit.ac.in

Rupak Chakraborty
Dept. of CSE
Guru Nanak Institute of Technology Kolkata, West Bengal, India
E-mail: rupak.chakraborty@gnit.ac.in

Sangeeta Bhattacharya
Dept. of CSE
Guru Nanak Institute of Technology Kolkata, West Bengal, India
E-mail: sangeeta.bhattacharya@gnit.ac.in

published online at https://gnitresearchmantra.in/

Abstract: Face detection and identification is now a days a problem of research interest. Detecting faces and identifying it can be done easily using computer vision. Machine learning based classifiers are useful for doing this. The application area of this can be in various places like content-based image retrieval, video coding, video conferencing, crowd surveillance, and intelligent human–computer interfaces, objects identification of visually impaired people etc. In this paper, OpenCV module is considered along with Haar Cascade classifier to detect and identify human faces. Further it is extended to detect gender and ages of human being so that it can help visually impaired people for identifying. The approach also provides solution for detecting any kind of objects with highest accuracy.

Keywords—Image Classification; Object Detection; Classifiers; OpenCV

I. INTRODUCTION

Millions of people live in this world with incapacities of understanding the environment due to visual impairment. Although they can develop alternative approaches to deal with daily routines, they also suffer from certain navigation difficulties as well as social awkwardness. For example, it is very difficult for them to recognize a particular known person, address any unknown person etc. in an unfamiliar environment such as school, workplace, etc. A blind and visually impaired people find it difficult to know whether a person is talking to them or someone else during a conversation. Computer vision technologies, especially the deep convolution neural network, have been rapidly developed in recent years. It is promising to use the state-of-art computer vision techniques to help people with vision loss. Not only it is useful for vision loss people it has wider application areas like to prevent retail crime, unlock phones, find missing reasons, protect law enforcement and many more. For performing these tasks, the crucial step is to identify the faces properly. For doing those various kinds of approaches based on machine learning classifiers are proposed since decades. A brief survey of those approaches are carried out in next section. Section III illustrates objective along with the proposed model whereas Section IV details modules and its descriptions. Experimental results are shown in Section V and the final Section draws the conclusion of the paper.

II. LITERATURE REVIEW

A survey of Face recognition model was proposed by W. Zhao et al. [8], and E. Hjelams et al. [13] in the year of 2001 and 2003 respectively. They surveyed various classification, edge-based approaches to detect faces effectively. Detecting faces from complexed backgrounds like colored background etc. become much more difficult. G. Yang et al. proposed one automated face detection system based on the idea of a hierarchical knowledge-based method. The method was basically based on the idea of an improved edge-detection approach [15]. Another technique was proposed by R.L. Hsu et al. [14] to identify faces from colored images. The idea was based on the combination of a novel lightning compensation technique and a non-linear color transformation. It is also possible to know the ages of human beings from the identified faces. By training the classifiers and then predicting the ages were successfully achieved by different researchers. The survey of those techniques can be found in the literature [9]. In the year of 2008, Y. Fu et. al. proposed a linear regression-based approach to estimate human age after applying manifold method of analysis on face images [10]. Our survey continued and it was found that one innovative approach based on the Fuzzy LDA method and Gabor feature of images was effectively proposed [12]. This hybrid classifier generated higher precision values when the technique was applied to the group of consumer images. Identifying objects properly from the group of persons is also a challenging task. One efficient solution was proposed by A. C. Gallegher et al. in the form their research paper [11]. They introduced contextual features which encapsulated the group structure locally and globally both for identifying the group people properly. Survey continued and found more recently developed classifications of face images based on AdaBoost, Boosted Cascade, l_2 techniques [5-7]. Like detection of faces, identifying other objects along with faces also holds current research interest. This method specially helps to the visually impaired people a lot. Object detection techniques are surveyed by P. Panchal et al. in their research work in detail [4]. Two very recent flexible guide mechanism was proposed by S. Durgadevi et. al. and S. Udgirkar et. al. to help blind people. First and foremost images were captured by camera, then classification was done by machine learning classifier and finally those images were converted to audio signal for assisting blind people [1-2]. Multiple objects recognition in a single image becomes very difficult because of lightning, rotation, positioning, mirroring, occlusion, scaling etc. To solve the effects of these, various approaches are proposed since decades. Survey of those approaches are noted by K. Khurana et. al. [3] in their paper. The surveyed techniques are template-matching, color-based, active-passive, shape-based, local-global features etc. Corresponding classifiers which were used with those approaches are can also be found there. Inspired by the above-sated literature, we are also proposing models that will demonstrate three models namely face recognition, age-gender detection and object detection.

III. OBJCECTIVE AND SYSTEM MODEL

A. Objective

The objective of the proposed approach is to build a system that converts visual input into audio signals which may lead to a practical product to help the blind or visually impaired navigate.

B. Proposed Model

Identification of people is a major challenge faced by the visually impaired. The increase in computation capability of mobile devices gives motivation to develop applications that can assist visually impaired persons. The proposed system is designed to take advantage of the portability of mobile devices and provide a simple user interface that makes use of the system easy for the visually impaired. Key design requirements for a portable system include small device size and low weight. To achieve this goal, a wearable camera fitted with the glass, mobile device and earpiece (for getting audio output) are used to form a compact and lightweight system.

The proposed model is based on the Haar Cascade Classifier and OpenCV. Brief description of these two are given below:

i. Haar Cascade classifier

Haar Cascade classifier is an effective face detection approach which was proposed by Paul Viola and Michael Jones in 2001. This is basically a machine learning based approach where a cascade function is trained from a lot of images both positive and negative. Based on the training it is then used to detect the objects in the other images.

ii. OpenCV_

OpenCV (Open Source Computer Vision) is a popular computer vision library started by Intel in 1999. The cross-platform library sets its focus on real-time image processing and includes patent-free implementations of the latest computer vision algorithms. In 2008 Willow Garage took over support and OpenCV 2.3.1 now comes with a programming interface to C, C++, Python and Android. OpenCV is released

Fig.1: Block Diagram of Face Detection & Recognition

under a BSD license so it is used in academic projects and commercial products alike. OpenCV 2.4 now comes with the very new FaceRecognizer class for face recognition.

The currently available algorithms are:

- Eigenfaces
- Fisherfaces
- Local Binary Patterns Histograms.

Features of OpenCV:

There are several features for the OpenCV library, some of the general features are as follows:

- Open-source
- Fast speed
- Easy to integrate
- Ease of coding
- Fast prototyping

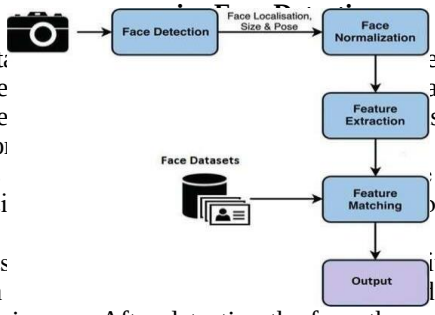
To solve the problem, haarcascade_frontalface_default.xml algorithm from Haar cascade classifier and Local Binary Patterns Histograms(LBPH) algorithm for face recognition from OpenCV are used here. Block diagram of face detection, age & gender analysis and object detection approaches are shown in the Fig. 1, 2 and 3 respectively.

Fig.2: Block Diagram of Age & Gender Analysis

Fig.3: Block Diagram of Object Detection

IV. MODULES AND ITS DESCRIPTIONS

A. Face Detection and Recognition Module



First st em using Haar-cascade ating new Haar-cascade scades that was used fo caused random objects is incorporated to obtai bject is created using through the cv2.Cas ive XML file. A camera leoCapture() to capture images. After detecting the face, the captured image is cropped and resized and converted into grayscale for further operations.

ii. Training the classifier

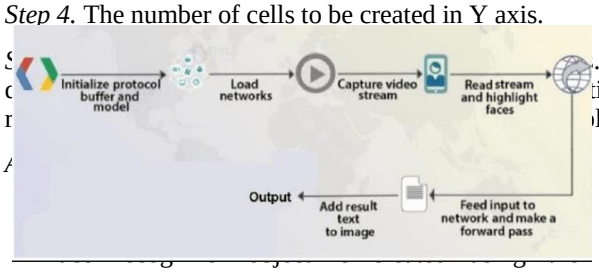
OpenCV enables the creation of XML files to store features extracted from datasets using the FaceRecognizer class. The stored images are imported, converted to grayscale and saved with IDs in two listswith same indexes. FaceRecognizer objects are created using face recogniser class like this cv2.face. createLBPHFaceRecognizer(). Steps of this algorithm is as follows:

Step 1: The radius from the centre pixel to build the local binary pattern.

Step 2. The Number of sample points to build the pattern. Having a considerable number will slow down the computer.

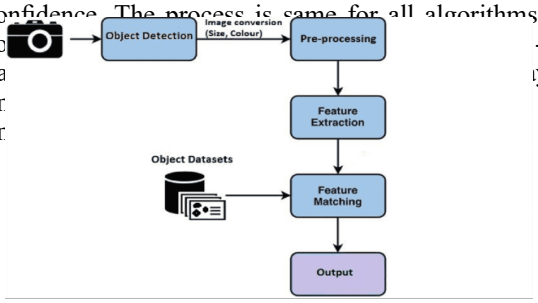
Step 3. The Number of Cells to be created in X axis.

Step 4. The number of cells to be created in Y axis.



. If the tion will le.

desired parameters. Face detector is used to detect faces in the image, cropped and transferred to be recognized. This is done using the same technique used for the image capture application. For each face detected, a prediction is made using FaceRecognizer.predict() which return the ID of the class and confidence. The process is same for all algorithms and if the co -1. Finally, na y the name or will print ur



B. Age and Gender Detection

To address the unknown person for a blind people we have to detect the age and gender of an unknown person. It is very difficult to detect the exact age of a person. So we make our Age and Gender Detection module as a classification problem. This module have been trained on the Adience dataset which are kept on our public domain and datasets are present also on our module holding folder.To implement this Age and Gender Detection module we need some packages which are OpenCV(cv2), math, argparse etc. cv2 will be used for capturing the image from the frame and perform other image related functions. argparse is the recommended command line parsingpackages or module in the python standard library.

This module works at three phases:

- Detect faces
- Data processing from dataset & classify age & gender
- Put results on image and display the results which are age & gender.

For face detection, we have a .pb file which is a protobuf file (protocol buffer) and it holds the trained weights of the model. We can use this to run the trained model. While a.pb file holds the protobuf in binary format, one with the .pbtxt extension holds it in text format. For age and gender, the .prototxt files describe the network configuration and the .caffemodel file defines the internal states of the parameters of the layers. For face, age, and gender the module initialize protocol buffer and model and initialize the mean values for the model and the lists of age ranges and genders to classify. Now we use the readNet() method to load the networks. The first parameter holds trained weights and the second carries network configuration. Now we capture video stream by cv2.VideoCapture() function to classify on a webcam's stream. we set padding to 20. Now until any key is pressed,the module read the stream and store the content into the names hasFrame and frame. If it isn't a video, it must wait, and so module call up waitKey() from cv2, then break. The module make a call to the highlightFace() function with the faceNet and frame parameters, and what this returns, we will store in the names resultImg and faceBoxes. And if we got 0 faceBoxes, it means there was no face to detect. Here, net is faceNet and this model

is the DNN Face Detector. The detailed working of the module is given in the following algorithm.

Step 1: This module create a shallow copy of frame and get its height and width.

Step 2: The module create a blob from the shallow copy.

Step 3: This module set the input and make a forward pass to the network.

Step 4: faceBoxes is an empty list now. for each value in 0 to 127, define the confidence (between 0 and 1). Wherever we find the confidence greater than the confidence threshold, which is 0.7, we get the x1, y1, x2, and y2 coordinates and F

Step 5: Then, we put up rectangles on the image for each such list of coordinates and return two things: the shallow copy and the list of faceBoxes.

Algorithm 2: **The working process of DNN Face Detector**

But if there are indeed faceBoxes, for each of those, we define the face, create a 4-dimensional blob from the image. In doing this, we scale it, resize it, and pass in the mean values. We feed the input and give the network a forward pass to get the confidence of the two class. Whichever is higher, that is the gender of the person in the picture. To classify the gender we give male and female. Then, we do the same thing for age. Atlast we'll add the gender and age texts to the resulting image and display it with imshow().

C. Object Detection

1.In order to detect any object we are using Mobile net SSD because it is one of the best methods which has a good balance between accuracy and speed, We will be able to run with our CPU almost real time and it will be able to detect lot of objects with a good amount of accuracy.

2.We have to install open cv python to acess all the libraries.

3.We have to import our image so that we can detect any object.

4.So, we have to import our files within that we will import configuration file.

5.Have to create our model. To create model we have to add weightpath and configuration path.

6.Now we need to send our image to our model then it will give us the predictions.

7.We have to define the confidence threshold so that we can define that particular point that which point we want to detect. For defining threshold we create the bounding box on our objects.

8.We have to subtract the value and have to write the origin points. We have to decide the font size and have to put scale and color, thickness.

9.We have to change this so that we can run It with an actual webcam.

10.We have to define certain parameters on how big our image is.

11.We have to define our image.

12. Now, we have to run it and test in up.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The dataset is created by our own. Then we are applying our technique to this dataset to get the results. Fig. 4 & Fig. 8 are showing the datasets whereas Fig. 5, 6, 7, 9, 11 are displaying the outcomes of the approach. Fig. 6 & 7 are detecting faces, Fig. 9 is identifying multiple objects in a single image, Fig. 10 and 11 are showing that non-human being objects are also traced by the algorithm.

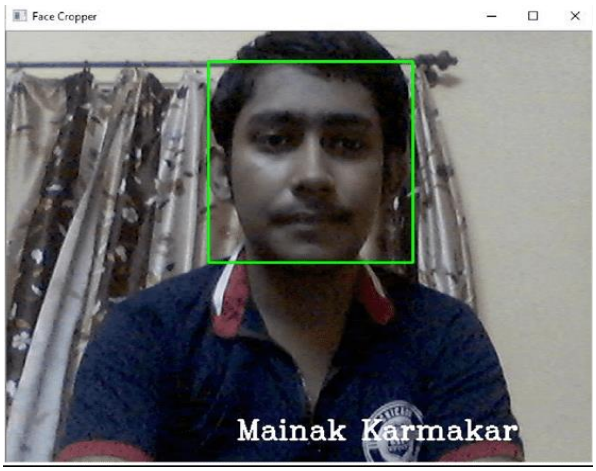
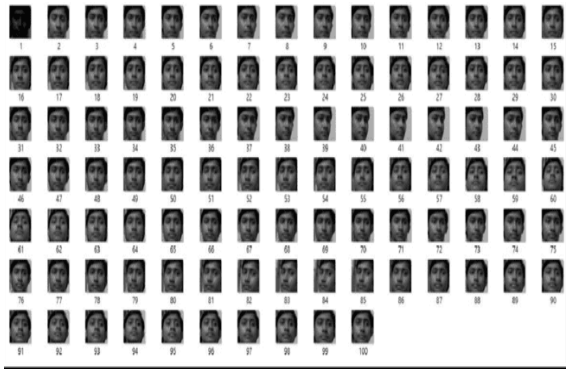


Fig.4: Dataset of a person for face recognition



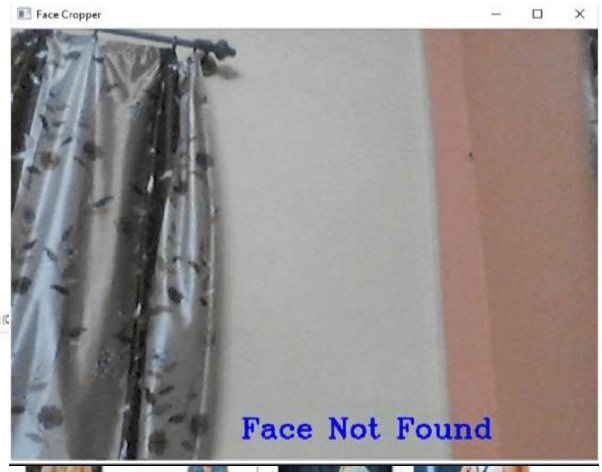


Fig.5: no face is present in front of the camera

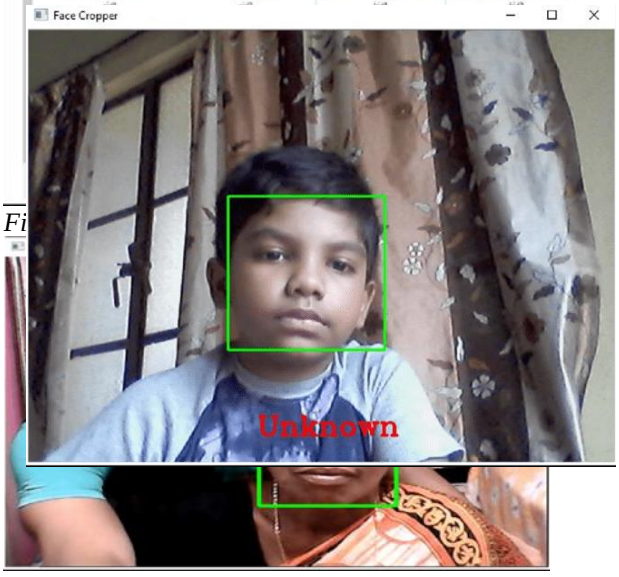


Fig.9: Experimental results for age & gender detection

Fig.6: Face is detected but not recognized

Fig.10: Experimental results for object detection

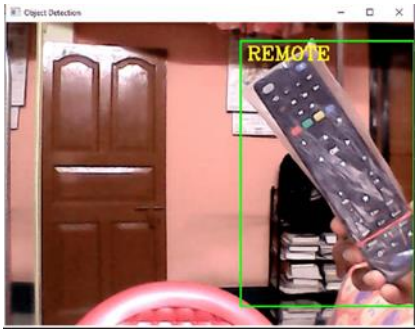


Fig.11: Experimental results for object detection

VI. CONCLUSIONS AND FUTURE WORK

The proposed approach is applied on created dataset and the result that was taken, showing that it is providing prominent outcomes with higher accuracy. The approach is not only applied to detect faces but also it is helpful to analysis gender and age when they are trained properly. It can also be seen that the approach is quite capable of identifying objects from the group of objects. The approach can easily be applied to make objects identify to visually impaired people. Only conversion of identified image to audio signal is required and designing of its corresponding device is required. This approach will be compared with some deep learning-based approaches as well in future to show its effectiveness.

VII. ACKNOWLEDGMENT

This work is the collaborative work of all the authors to propose an idea of detecting objects along with its age and genders. No copyrighting is associated with this work.

REFERENCES

[1] S. Udgirkar, S Sarokar, S Gore, D Kakuste, and S Chaskar, " Object Detection System for Blind People," International Journal of Innovative Research in Computer and Communication Engineering, vol. 7, no. 6, June, 2020.

[2] G. Eason, B. Noble, and I.N. Sneddon, "Smart Machine Learning System for Blind Assistance," IEEE International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 10-11 Dec, 2020.

[3] K Khurana and R Awasthi, "Techniques for Object Recognition in Images and Multi-Object Detection," International Journal of Advanced Research in Computer Engineering & Technology, vol. 2, no.4, pp.1383-1388, April 2013.

[4] P Panchal, G Prajapati, S Patel, and H Shah, "A Review on Object Detection and Tracking Methods," International Journal For Research In Emerging Science And Technology, vol. 2, no.1, pp. 7-12, January 2015.

[5] Q. Shi,, A. Eriksson, , A V D Hengel, and C. Shen, "Is face recognition really a compressive sensing problem?," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 20-25 June, 2011.

[6] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2,8-14 December, 2001.

[7] J. Yoon, and D. Kim, "Frontal face classifier using AdaBoost with MCT features," IEEE International Conference on Control, Automation, Robotics and Vision (ICARCV), 7-10 December, 2010, DOI:10.1109/ICARCV.2010.5707954

[8] W. Zhao, ., R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," Acm Computing Surveys, vol.35, no.4, pp.339-458, December 2003.

[9] Y. Fu, G. Guo, and T. S. Huang, " Age synthesis and estimation via faces: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.32, no.11, pp. 1955-1976, November, 2010.

[10] Y. Fu and T. S. Huang, "Human age estimation with regression on discriminative aging manifold," IEEE Transactions on Multiledia, vol. 10, no.4, pp.578-584, May 16, 2008.

[11] A. C. Gallagher and T. Chen, "Understanding images of groups of people," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 20-25 June, 2009

[12] F. Gao and H. Ai, " Face age classification on consumer images with gabor feature and fuzzy lda method," Internattional Springer Conference on Advances in Biometrics, pp.132-141, 2-5 June, 2009, DOI: 10.1007/978-3-642-01793-3_14

[13] E. Hjelams, B. K. Low, " Face Detection: A Survey ," Computer Vision and Image Understanding, vol. 83 no.3, pp.236-274, September, 2001.

[14] R. L. Hsu, M.L. Mottaleb, A.K. Jain, "Face Detection in Color Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.5, pp. 696-706, August, 2002.

[15] G. Yang, T.S. Huang, "Human Face Detection in a Complex Background," Pattern Recognition, vol. 27, no.1, pp.53-63, January, 1994.