JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

“Fake Data Detection For Image Augmentation Using Generative Adversarial Network”

Aher Vrushali, Bhavsar Renuka, Deshmukh Nikita, Kumbhar Harshada, Prof.Tathe.S.G(Guide)

Department of Information Technology Prof. Rokade P.P. (H.O.D)
S.N.D College of Engineering & Research Center,Yeola Dr. Yadav D.M (Principal)

ABSTRACT

In colleges, universities, organizations, schools, and offices, taking attendance is one of themost important tasks that must be done on a daily basis. The majority of the time, it is donemanually, such as by calling by name or by roll number. The main goal of this project is tocreate a Face Recognition-based attendance system that will turn this manual process into an automated one. This project meets the requirements for bringing modernization to the way attendance is handled, as well as the criteria for time management. This device is installed in the classroom, where and student's information, such as name, roll number, class, sec, and photographs, is trained. The images are extracted using Open CV. Before the start of the corresponding class, the student can approach the machine, which will begin taking pictures and comparing them to the qualified dataset. Logitech C270 web camera and NVIDIA Jetson Nano Developer kit were used in this project as the camera and processing board. The image is processed as follows: first, faces are identified using a Haarcascade classifier, then faces are recognized using the LBPH (Local Binary Pattern Histogram) Algorithm, histogram data is checked against an established dataset, and the device automatically labels attendance. An Excel sheet is developed, and it is updated every hour with the information from the respective class instructor.

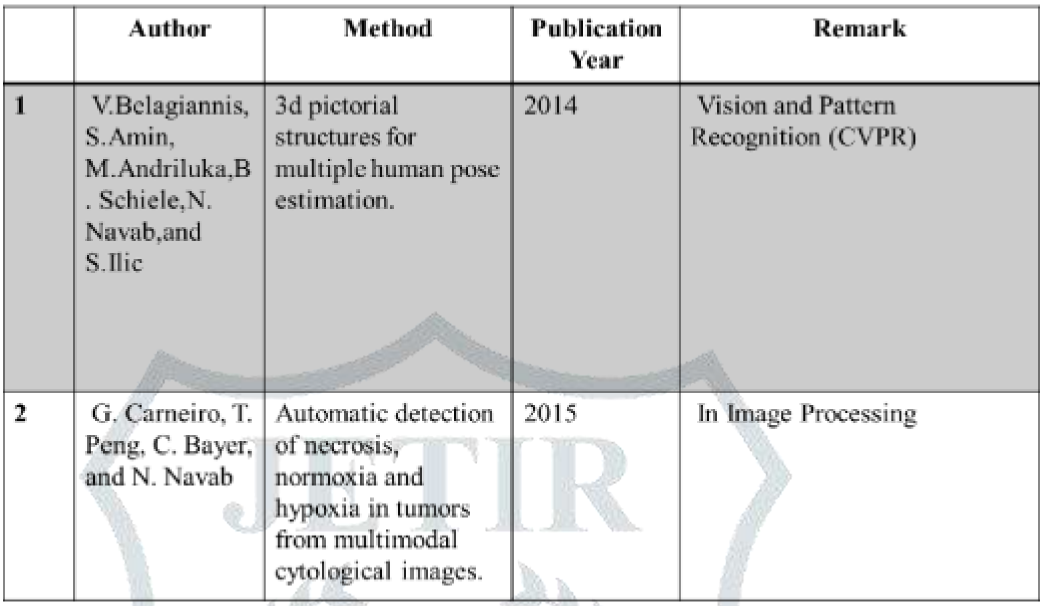
Keywords: Face Detection, Face Recognition, HaarCascade classifier, NVIDIA Jetson Nano

1.INTRODUCTION

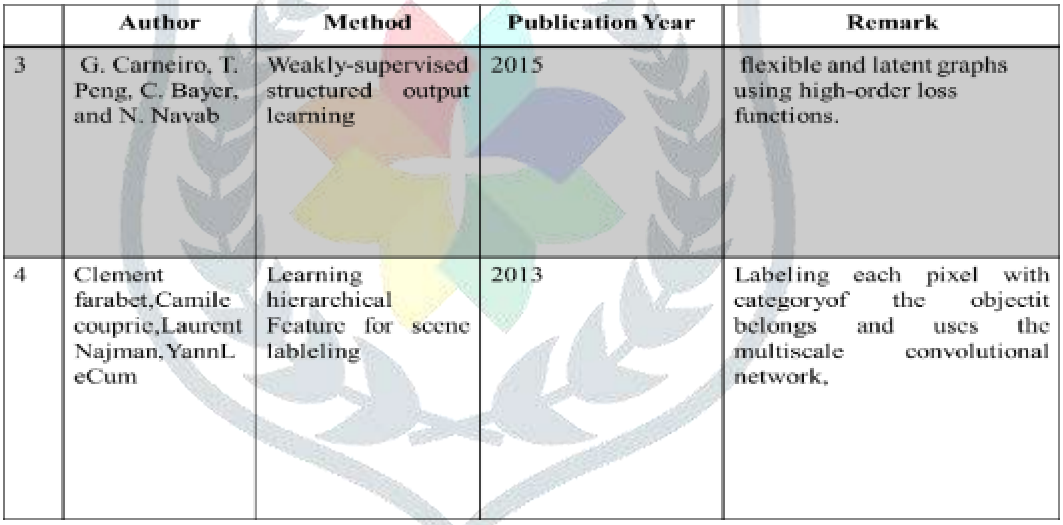
Face detection involves separating image windows into two classes; one containing faces (turning the background (clutter). It is difficult because although commonalities exist between faces, they can vary considerably in terms of age, skin color and facial expression. The problem is further complicated by differing lighting conditions, image qualities and geometries, as well as the possibility of partial occlusion and disguise. An ideal face detector would therefore be able to detect the presence of any face under any set of lighting conditions, upon any background. The face detection task can be broken down into two steps. The first step is a classification task that takes some arbitrary image as input and outputs abinary value of yes or no, indicating whether there are any faces present in the image. The second step is the face localization task that aims to take an image asinput and output the location of any face or faces within that image as somebounding box with (x, y, width, height).After taking the picture the system willcompare the equality of the pictures in its database and give the most related result.

2. LITERATURE SURVEY

Arun Katara et al. (2017) mentioned disadvantages of RFID (Radio Frequency Identification) card system, fingerprint system and iris recognition system. RFID card system is implemented due to its simplicity. However, the user tends to help their friends to check in as long as they have their friend’s ID card. The fingerprint system is indeed effective but not efficient because it takes time for the verification process so the user has to line up and perform the verification one by one. However for face recognition, the human face is always exposed and contain less information compared to iris. Iris recognition system which contains more detail might invade the privacy of the user. Voice recognition is available, but it is less accurate compared to other methods. Hence, face recognition system is suggested to be implemented in the student attendance system.



	Author	Method	Publication Year	Remark
1	V.Belagiannis, S.Amin, M.Andriluka,B. Schiele,N. Navab,and S.Ilic	3d pictorial structures for multiple human pose estimation.	2014	Vision and Pattern Recognition (CVPR)
2	G. Carneiro, T. Peng, C. Bayer, and N. Navab	Automatic detection of necrosis, normoxia and hypoxia in tumors from multimodal cytological images.	2015	In Image Processing



	Author	Method	Publication Year	Remark
3	G. Carneiro, T. Peng, C. Bayer, and N. Navab	Weakly-supervised structured output learning	2015	flexible and latent graphs using high-order loss functions.
4	Clement farabet,Camile couprie,Laurent Najman,YannL eCum	Learning hierarchical Feature for scene lableling	2013	Labeling each pixel with categoryof the objectit belongs and uses the multiscale convolutional network.

3. NEED OF PROJECT

This introduces the high possibility of human error in the detection process which necessitate an automated process.

Hence, In Our proposed system we are detection all kind of fake detection using image processing. we can use any normal user and also student use for study purpose.

4. PROPOSED SYSTEM

The proposed face recognition and fake detection is based on the deep learning technique of the face using the Local gray level matrix algo.

The accurate detection of deep fake image system consists of four phases such as

pre-processing,

dimensional reduction of image feature extraction.

Classification

5.EXISTING SYSTEM

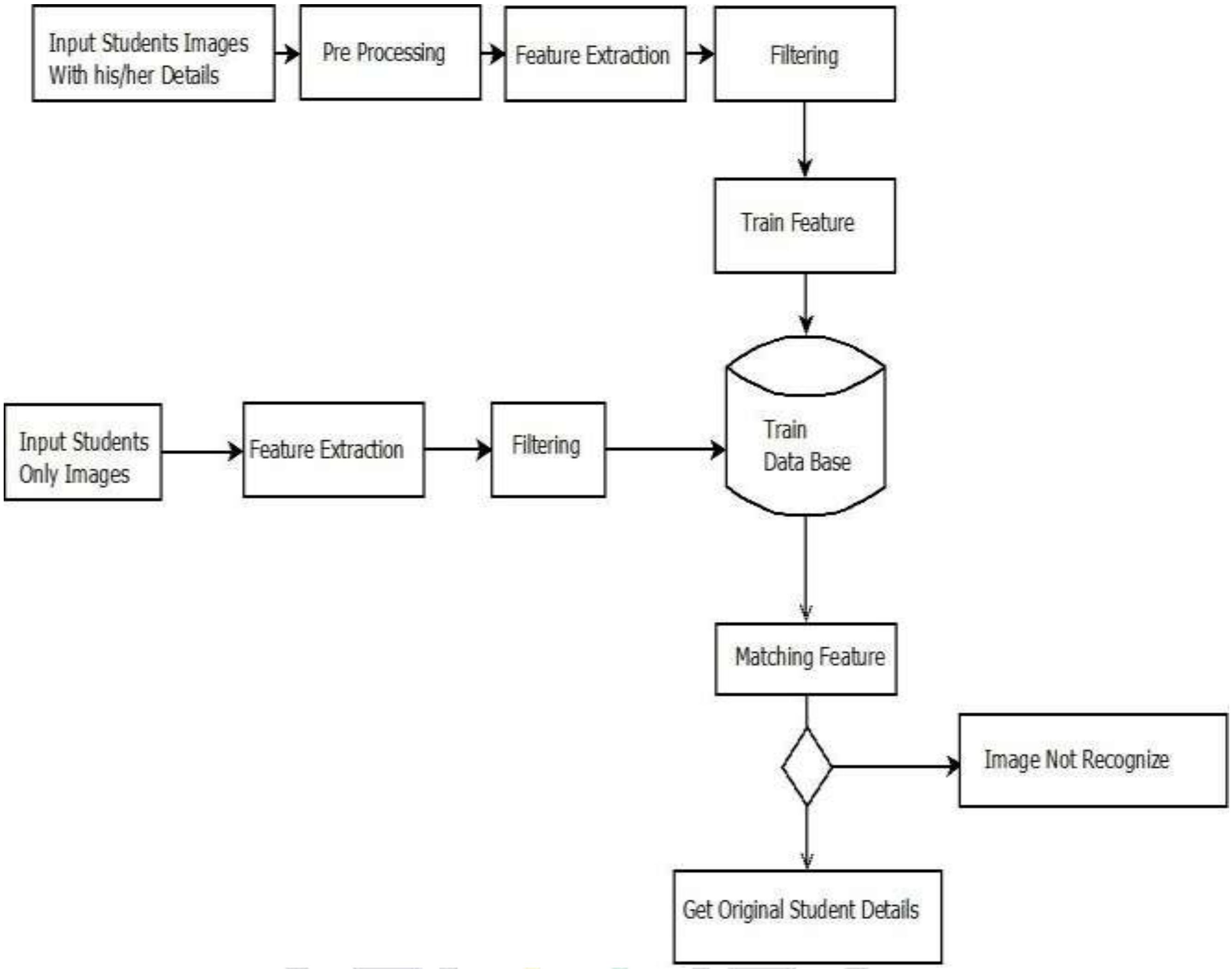
The main issues in the existing techniques are inaccurate, consumption time is high.

In this work we implement detecting of deep fake face image analysis using deep learning technique of fisherface using Local Binary Pattern Histogram (FF-LBPH).

Fisherface algorithm is used to recognize the face by reduction of the dimension in the face space using LBPH.

Deepfakes are being widely used as a malicious source of misinformation in court that seek to sway a court's decision. Because digital evidence is critical to the outcome of many legal cases, detecting deepfake media is extremely important and in high demand in digital forensics.

6.SYSTEM DESIGN



7.ALGORITHM

LBPH algorithm work step by step:

LBPH algorithm work in 5 steps.

1.Parameters: the LBPH uses 4 parameters:

- **Radius:** the radius is used to build the circular local binary pattern and represents the radius around the central pixel. It is usually setto1.
- **Neighbors:** the number of sample points to build the circular local binary pattern. Keep in mind: the more sample points you

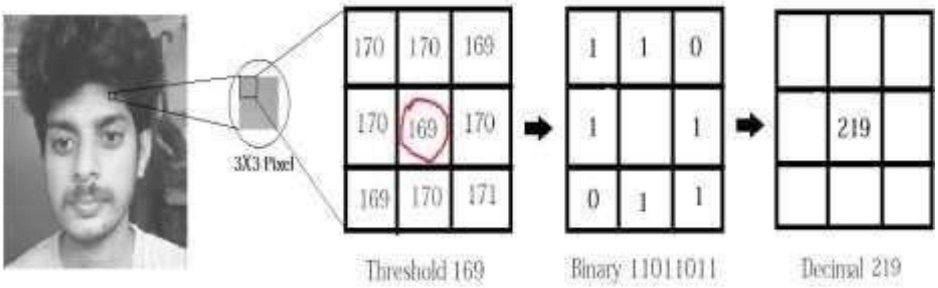
include, thehigher the computational cost. It is usually set to 8.

- **Grid X:** the number of cells in the horizontal direction. The more cells,the finer the grid, the higher the dimensionality of the resulting feature vector. It is usually set to 8
- **.Grid Y:** the number of cells in the vertical direction. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector. It is usually set to 8.

2. Training the Algorithm: First, we need to train the algorithm. To do so, we need to use a dataset with the facial images of the people we want to recognize. We need to also set an ID (it may be a number or the nameof the person) for each image, so the algorithm will use this informationto recognize an input image and give you an output. Images of the same person must have the same ID. With the training set already constructed,let's see the LBPH computational steps.

3. Applying the LBP operation: The first computational step of the LBPH is to create an intermediate image that describes the original image in a better way, by highlighting the facial characteristics. To do so, the algorithm usesa concept of a sliding window, based on the parameters radius and neighbors

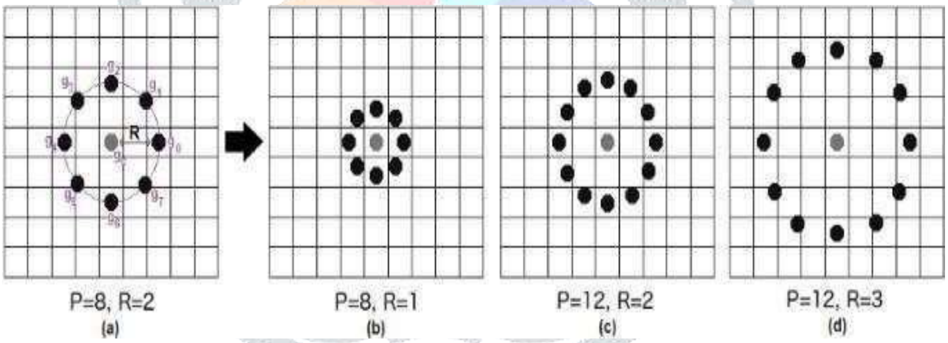
The image below shows this procedure:



Based on the image above, let's break it into several smallsteps so we can understand it easily:

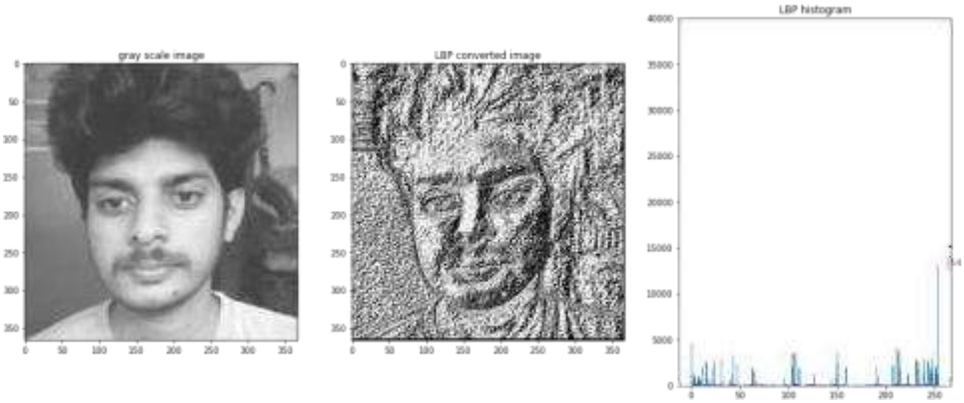
- Suppose we have a facial image in grayscale.
- We can get part of this image as a window of 3x3 pixels.
- It can also be represented as a 3x3 matrix containing the intensity of eachpixel (0~255).

- Then, we need to take the central value of the matrix to be used as thethreshold.
- This value will be used to define the new values from the 8 neighbors.
- For each neighbor of the central value (threshold), we set a new binary value. We set 1 for values equal or higher than the threshold and 0 for values lower than the threshold.
- Now, the matrix will contain only binary values (ignoring the central value). We need to concatenate each binary value from each position fromthe matrix line by line into a new binary value (e.g. 10001101). Note: some authors use other approaches to concatenate the binary values (e.g. clockwise direction), but the final result will be the same.
- Then, we convert this binary value to a decimal value and set it to the central value of the matrix, which is actually a pixel from the original image.
- At the end of this procedure (LBP procedure), we have a new image whichrepresents better the characteristics of the original image.



It can be done by using bilinear interpolation. If some data point is betweenthe pixels, it uses the values from the 4 nearest pixels (2x2) to estimate the value of the new data point.

4.Extracting the Histograms: Now, using the image generated in the laststep, we can use the Grid X and Grid Y parameters to divide the image into multiple grids shown.



Based on the image above, we can extract the histogram of each region as follows:

- As we have an image in grayscale, each histogram (from each grid)will contain only 256 positions (0~255) representing the occurrences of each pixel intensity.
- Then, we need to concatenate each histogram to create a new and biggerhistogram. Supposing we have 8x8 grids, we will have 8x8x256=16.384 positions in the final histogram. The final histogram represents the characteristics of the image original image.

5. Performing the face recognition: In this step, the algorithm isalready trained. Each histogram created is used to represent each imagefrom the training dataset. So, given an input image, we perform the steps again for this new image and creates a histogram which represents the image.

- So to find the image that matches the input image we just need to compare two histograms and return the image with the closest histogram.
- We can use various approaches to compare the histograms (calculate the distance between two

$$D = \sqrt{\sum^{n}(hist1_i - his}$$

histograms), for example: Euclidean distance, chi- square, absolute value, etc. In this example, we can use the **Euclidean distance** (which is quite known) based on the followingformula:

- So the algorithm output is the ID from the image with the closest histogram. The algorithm should also return the calculated distance, which can be used as a 'confidence' measurement.
- We can then use a threshold and the 'confidence' to automatically estimate if the algorithm has correctly recognized the image. We canassume that the algorithm has successfully recognized if the confidence is lower than the threshold defined.

8.SYSTEM REQUIREMENTS

SOFTWARE REQUIREMENTS

Operating System	: Windows 7
Front End	: Python
Back End	: My SQL
UML Design	: Rational Rose
Server	:XAMP

HARDWARE REQUIREMENTS

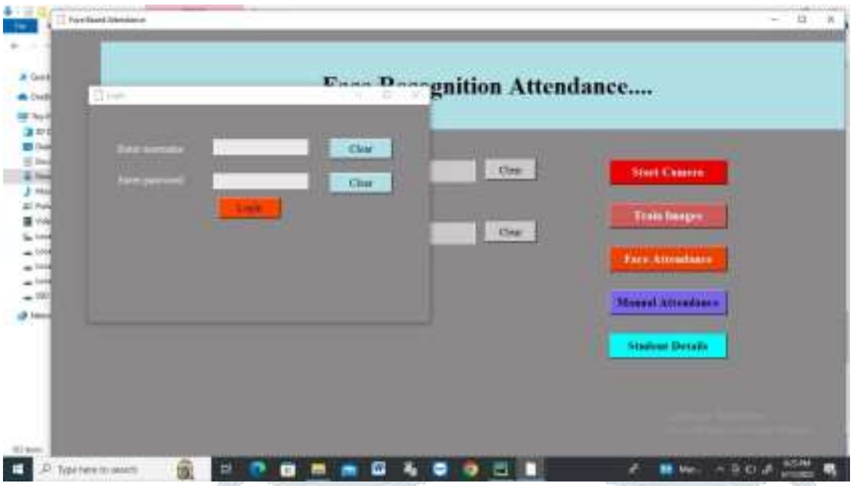
Hard Disk : 120 GB RAM : 512 MB

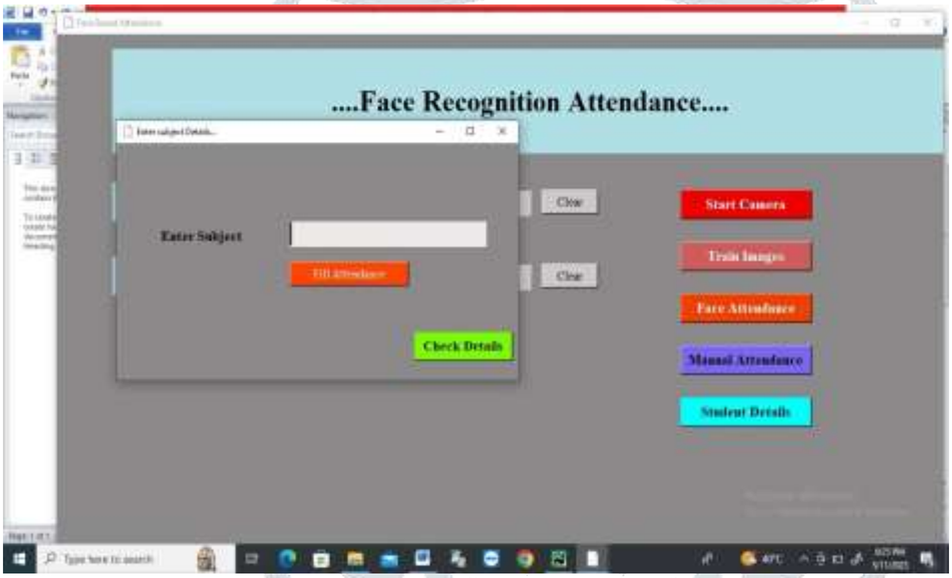
Processor : Pentium 4

Input device : Keyboard and Mouse Output device : Monitor

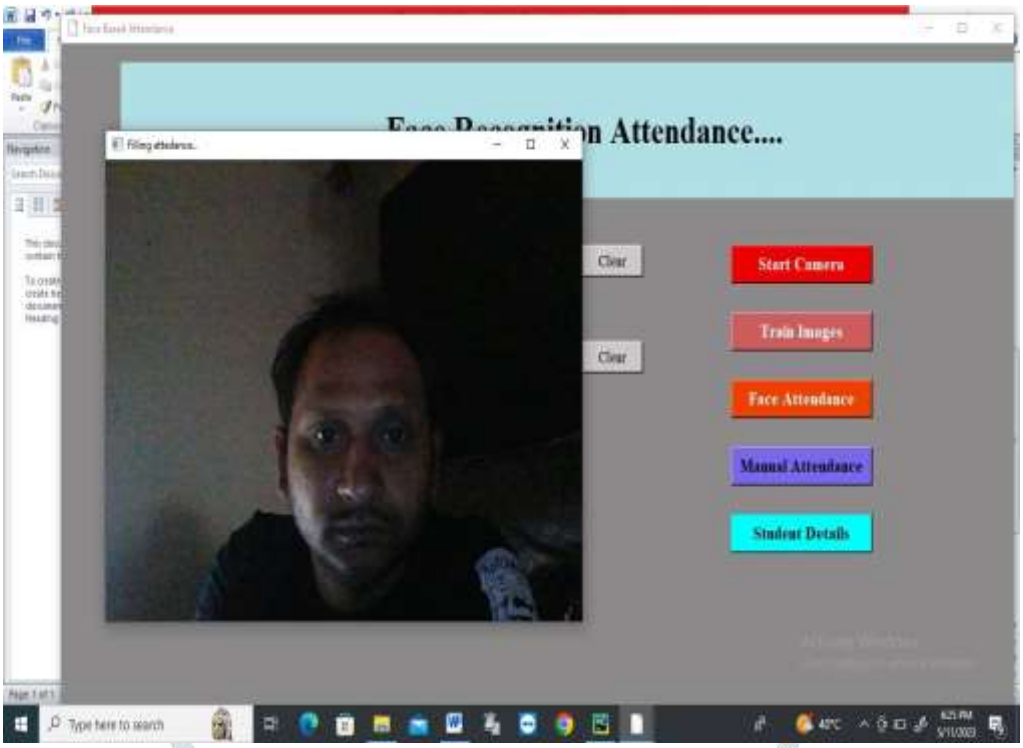
9.OUTOUT SCREENSHOTS

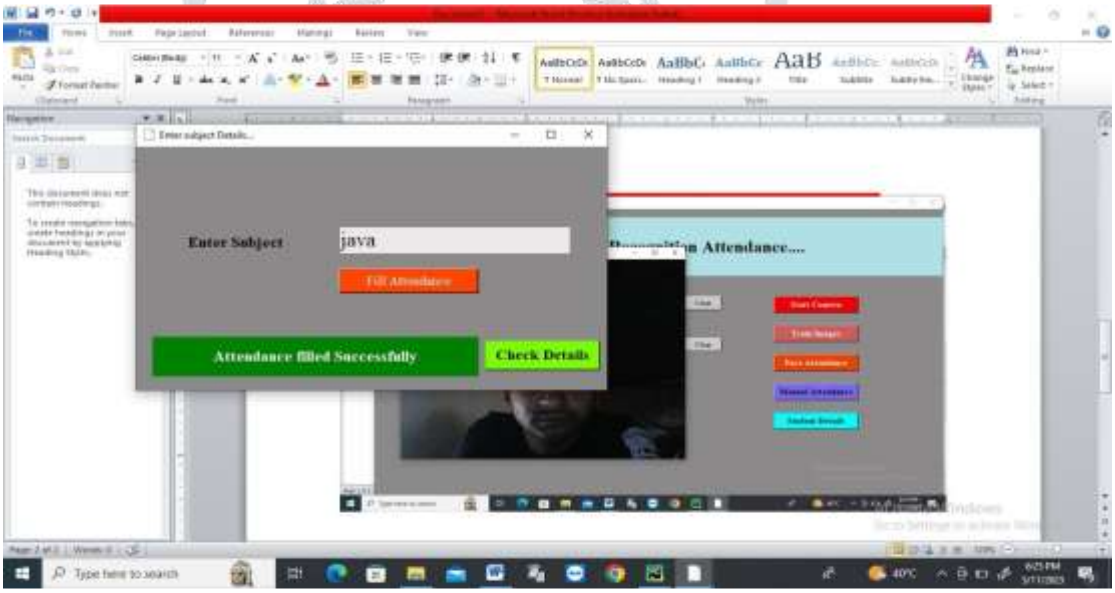
1.LOGIN MODULE





2.CAMERA MODULE

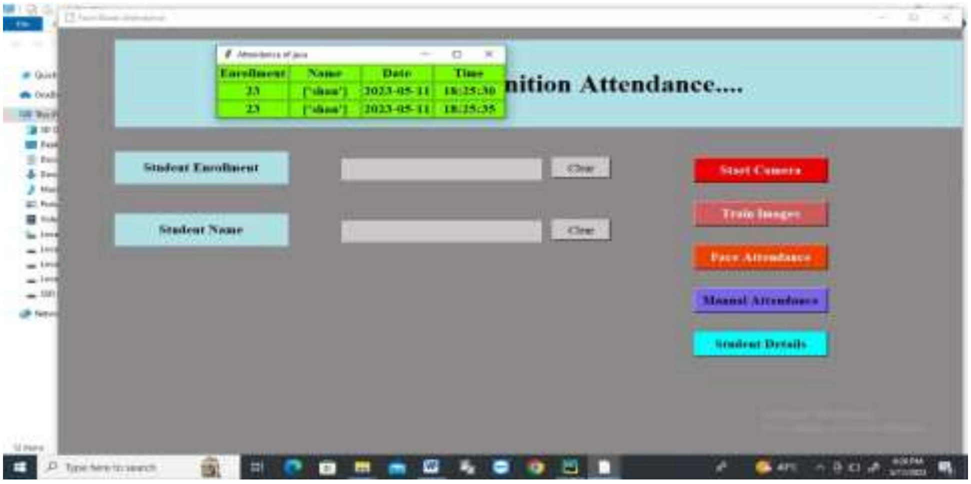




Enter Subject java

Attendance filled Successfully

3.DETECTION MODULE



Enrollment	Name	Date	Time
23	['shan']	2023-05-11	18:25:30
23	['shan']	2023-05-11	18:25:35

Student Enrollment

Student Name

10. CONCLUSION

Face recognition systems are part of facial image processing applications and their significance as a research area are increasing recently. Implementations of system are crime prevention, video surveillance, person verification, and similar security activities. The facerecognition system implementation can be part of Universities. Face Recognition Based Attendance System has been envisioned for the purpose of reducing the errors that occur in the traditional (manual) attendance taking system. The aim is to automate and make a system that is useful to the organization such as an institute. The efficient and accurate method of attendance in the office environment that can replace the old manual methods. This method is secure enough, reliable and available for use. Proposed algorithm is capable of detect multiple faces, and performance of system has acceptable good results.

11. REFERENCES

[1] D. G. Altman and J. M. Bland. Measurement in medicine: the analysis of method comparison studies. The statistician, pages 307–317, 1983.

[2] C. Bayer, K. Shi, S. T. Astner, C.-A. Maftei, and P. Vaupel. Acute versus chronic hypoxia: why a simplified classification is simply not enough. International Journal of Radiation Oncology* Biology* Physics, 80(4):965–968, 2011.

[3] V.Belagiannis, S.Amin, M.Andriluka,B. Schiele,N. Navab,and S.Ilic. 3d pictorial structures for multiple human pose estimation. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 1669–1676. IEEE, 2014.

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. Pattern Analysis and Machine Intelligence, IEEE

Transactions on, 23(11):1222–1239, 2001.

[5] L. Breiman. Random forests. Machine learning, 45(1):5–32, 2001.

[6] G. Carneiro, T. Peng, C. Bayer, and N. Navab. Automatic detection of necrosis, normoxia and hypoxia in tumors from multimodal cytological images. In Image Processing (ICIP), 2015 IEEE International Conference on, pages 2429–2433. IEEE, 2015.

12. FUTURE SCOPE

We are setting up to design a system comprising of two modules. The first module (face detector) is a mobile component, which is basically a camera application that captures student faces and stores them in a file using computer vision face detection algorithms and face extraction techniques. The second moduleis a desktop application that does face recognition of the captured images (faces) in the file, marks the students register and then stores the results in a database for future analysis.