Automated Attendance System and Voice Assistance using Face Recognition

P.Bhavani\textsuperscript{a}, C.Saraswathi\textsuperscript{b}, K.Soba\textsuperscript{b}, and R.Preethika\textsuperscript{b}

\textsuperscript{a}Associate Professor, Department of Computer Science & Engineering, Sri Manakula Vinayagar Engineering College, Puducherry.
\textsuperscript{b}UG Student, Department of Computer Science & Engineering, Sri Manakula Vinayagar Engineering College, Puducherry.

\textbf{Article History}: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

\textbf{Abstract}: In this work we propose a Live Attendance Marking System for institutional purpose. This system will enable the department to mark the attendance of students automatically by recognizing their faces. The system is based on the face detection and recognition algorithms and automatically recognizes a student whenever he/she comes across the camera module. Further after the recognition, it automatically updates his/her attendance in the Cloud web server database. The basic architecture and all the algorithms used in the model are described in the paper elaborately. Moreover in order to evaluate and enhance the performance of the system many of the filter functions are used which provide the capability to capture and recognize the images even in dull and low light places. This project also suggests the technique to ensure that a student can mark the attendance only once in a day. The live attendance system is much more efficient to traditional attendance systems both in saving time and in maintaining the database.

\textbf{Keywords}: Viola Jones, Region Based Convolution Neural Network (RCNN), Histogram of Oriented Gradients (HOG)

1. INTRODUCTION

1.1 Image Processing

The term digital image refers to processing of a two-dimensional picture by a digital computer. In a broader context, it implies digital processing of any two-dimensional data. A digital image is an array of real or complex numbers represented by a finite number of bits. An image given in the form of a transparency, slide, photograph or an X-ray is first digitized and stored as a matrix of binary digits in computer memory. This digitized image can then be processed and/or displayed on a high-resolution television monitor. For display, the image is stored in a rapid-access buffer memory, which refreshes the monitor at a rate of 25 frames per second to produce a visually continuous display. An image processor does the functions of image acquisition, storage, pre-processing, segmentation, representation, recognition and interpretation and finally displays or records the resulting image. A digitizer converts an image into a numerical representation suitable for input into a digital computer. Some common digitizers are Microdestiometre, Image dissector, Videocon camera, Photo sensitive solid state array, Flyingpoint scanner.

The following block diagram gives the fundamental sequence involved in an image processing system.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{block_diagram.png}
\caption{BLOCK DIAGRAM OF FUNDAMENTAL SEQUENCE INVOLVED IN AN IMAGE PROCESSING SYSTEM}
\end{figure}

1.2 History of image processing
The Field of image processing is continually evolving. During the past five years, there has been a significant increase in the level of interest in image morphology, neural networks, full-colour image processing, image data compression, image recognition, and knowledge-based image analysis systems. Image processing methods stems from two principal application areas: improvement of pictorial information for human interpretation, and processing of scene data for autonomous machine perception. Image is better than any other information form for our human being to perceive. Vision allows humans to perceive and understand the world surrounding us. Image understanding, image analysis, and computer vision aim to duplicate the effect of human vision by electronically. Basically image enhancement means improvement in appearance of image. Image restoration means to restore an image as it is after applying any operations. Image compression is to reduce the amount of data of an image. Processing of digital images involves procedures that are usually expressed in algorithmic form. Most image processing functions can be implemented in software. The only reason for specialized image processing hardware is the need for speed in some applications or to overcome some fundamental computer limitations.

1.3 Categories of image processing

Most of the common image processing functions available in image analysis systems can be categorized into the following four categories: Preprocessing, Image Enhancement, Image Transformation, Image Classification and Analysis

1.3.1 Preprocessing

Preprocessing functions involve those operations that are normally required prior to the main data analysis and extraction of information, and are generally grouped as radiometric or geometric corrections. Radiometric corrections include correcting the data for sensor irregularities and unwanted sensor or atmospheric noise, and converting the data so they accurately represent the reflected or emitted radiation measured by the sensor.

1.3.2 Image Enhancement

Image enhancement, is solely to improve the appearance of the imagery to assist in visual interpretation and analysis. Examples of enhancement functions include contrast stretching to increase the tonal distinction between various features in a scene, and spatial filtering to enhance (or suppress) specific spatial patterns in an image.

1.3.3 Image Transformation

Image transformations are operations similar in concept to those for image enhancement. However, unlike image enhancement operations which are normally applied only to a single channel of data at a time, image transformations usually involve combined processing of data from multiple spectral bands.

1.3.4 Image Classification and Analysis

Image classification and analysis operations are used to digitally identify and classify pixels in the data. Classification is usually performed on multi-channel datasets (A) and this process assigns each pixel in an image to a particular class or theme (B) based on statistical characteristics of the pixel brightness values. There are a variety of approaches taken to perform digital classification. We will briefly describe the two generic approaches which are used most often namely supervised and unsupervised classification.

1.4 Image Processing Architecture

Image processing architecture is depicted in Fig. 1. Image processing architectures can be classified based upon processing power, data bandwidth and interim storage.

1.4.1 Characterization of Processing: Power Processing power is the capacity of an architecture for performing manipulations of image data per unit of time. Since different architectures have different sets of manipulations in their instruction set a common denominator of operations is required for use as a scale of measurement of such capacity.

1.4.2 Characterization of Data Bandwidth: As like other real time data processing and communication application areas, the guiding and movement of data among operating nodes may impact system responsiveness or throughput as much or more than the operation timing itself. The specific parameters of such
data bandwidth which affect overall behavior of an architecture may include N, the number of independent data path, B the bandwidth of each path, M the degree of match between the data output rate and format requirements of transmitting nodes, and the input rate and format requirements of receiving nodes. The third parameter accounts for degradation of potential parallelism due to partial data unavailability. An example of such loss is when one node output is serial (such as that of a typical raster-scan image sensor) and it is connected to another node which requires an entire image’s data to be present before it can commence processing.

1.4.3 Characterization of Interim Storage: The previously discussed two attributes of an architecture provide insight into the raw power to manipulate and move data at a relatively elementary level of image operators. Another attribute of an architecture is the ability to save intermediate or “scratch” image data as a complex vision algorithm proceeds through several phases of image conditioning, segmentation, analysis, feature detection, etc. The critical features of such storage are its capacity and the directness of its accessibility to the operating nodes which eventually need to process the intermediate data. In other words, direct access by a node to just the desired pixels of a required intermediate image will support more system capability than will indirect.

1.4.4 Other attributes are:
- Size of images processable directly by the architecture
- Multiple versions of available systems
- Input/output bandwidth
- Program loading and/or distribution overhead
- Dynamic control of program flow
- Dynamic control of processing resolution

![Image Processing Architecture](image)

**Figure 1** Image processing architecture

1.5 Components of Image Processing

Image processing is the combination of the different elements involved in the digital image processing. Digital image processing is the processing of an image by means of a digital computer. Digital image processing uses different computer algorithms to perform image processing on the digital images. It consists of following components:

- Image Sensors: Image sensors senses the intensity, amplitude, co-ordinates and other features of the images and passes the result to the image processing hardware. It includes the problem domain.
- Image Processing Hardware: Image processing hardware is the dedicated hardware that is used to process the instructions obtained from the image sensors. It passes the result to general purpose computer.
- Computer: Computer used in the image processing system is the general purpose computer that is used by us in our daily life.
- Image Processing Software: Image processing software is the software that includes all the mechanisms and algorithms that are used in image processing system.
- Mass Storage: Mass storage stores the pixels of the images during the processing.
- Hard Copy Device: Once the image is processed then it is stored in the hard copy device. It can be a pen drive or any external ROM device.
• Image Display: It includes the monitor or display screen that displays the processed images.
• Network: Network is the connection of all the above elements of the image processing system.

1.5.1 Applications

Some of the applications of Image processing are
• Intelligent Transportation Systems
• Remote Sensing
• Machine/Robot vision
• Colour processing
• Video processing
• Biomedical Imaging techniques
• Defence surveillance
• Microscopic Imaging
• Automatic Visual Inspection System

2. RELATED WORKS

From the various researches, observed that there are some relevant works which similar to the current proposal whose ideas, works and drawbacks are listed below:

(Samridhi Dev,Tushar Patnaik,2020)The proposed system makes the use of Haar classifiers, KNN, CNN, SVM, Generative adversarial networks, and Gabor filters. Efficient and robust device for taking attendance in a classroom without any time consumption and manual work. Three algorithms have been used which are K-nearest neighbor, convolutional neural networks, and support vector machine and the method SVM algorithm proved to be less efficient Convolutional neural networks evinced to have low computational complexity

(E.Varadharajan, R.Dharani, S.Jeevitha, B.Kavinmathi, S.Hemalatha, 2018) says the camera is fixed in the classroom and it will capture the image, the faces are detected and then it is recognized with the database and finally the attendance is marked. In this paper they used eigen value method which is the most suitable method. This method is more suitable because of its speed. Hence here we are going to eigen value method to recognize the faces. Hence the background is subtracted only once in a set of image. For the purpose of accurate face detection we go for background subtraction. Then this is checked for images under different conditions. This method requires only simple hardware for installation. The students who attends a specific lecture, laboratory, sectionor exam at its specified time duration, thereby saving a lots of effort and time and decreasing interruptions and disruption

(Muhammad Ayat Hidayat, Holong Marisi Simalango, 2018) In this paper, attendance is still done manually using paper, where the paper will be signed by students later. This student attendance system is done by conducting data collection, system analysis, system design, and system implementation. In this case, the application is using the methods or procedures that exist in the API Usage and the REST API is based on REST architectural style, which is using text-based JSON message over HTTP transport to communicate each other. In this process the student opens the attendance application, after opening, the application will ask the user to activate Bluetooth. After Bluetooth is activated the application will check the existence of IBeacon. IBeacon functions as an identity for each classroom in the lecture building.

(KaneezLaila Bhatti , Laraib Mughal, FaheemYarKhuhawar ,Sheeraz Ahmed Memon,2019) The main theme of this system based on face recognition to maintain the attendance record of students. The daily attendance of students is recorded subjectwise which is stored already by the administrator. The system automatically starts taking snaps and then apply face detection and recognition technique to the given image and thererecognize students are marked as present and their attendance update with corresponding time and subject id. This Methodprovides of Deep learning techniques. To develop this system, histogram of oriented gradient method is used to detect faces in images and deep learning method is used to compute and compare feature 
facial of students to recognize them.

(Pooja G.R , Poornima M , Palakshi S , Bhanu Prakash Varma M , Krishnna N,2014)In this paper they proposed an automated attendance management system. This system is based on face detection and recognition
algorithms. In this Method, by using Webcam the student is automatically detected when he enters the class room and marks the attendance by recognizing him. proposed special techniques in order to handle the threats like spoofing. When compared to traditional attendance marking this system saves the time and also helps to monitor the students. Training is slow, but detection is very fast. The task of face detection in Viola Jones algorithm uses a 24x24 window as the base window size to start evaluating these features in any given image.

(Benfano Soewito, ford Lumban Gaol,Echo Simanjuntak,FergyantoE.Gunawan,2016) The proposed attendance system uses smartphone to verify the employee. The system provides two options for doing verification, such as fingerprint and voice recognition. This Attendance system using fingerprint and voice as a method of identifying and recording the coordinates GPS to ensure the user’s location when do attendance. This system includes of smartphone or mobile device. This framework for novel image compression based method for multispectral fingerprint biometric system MFCT consists of two phases Enrollment and Matching. Both phases include the application of corresponding transformation processes to get the new compressed images. During the enrolment process, the multispectral fingerprint is taken and wavelet decomposition is applied to obtain the four different band of the image and features are extracted.

3. EXISTING SYSTEM AND ISSUES

The Automated Attendance System and Voice Assistance is one among the image processing systems. According to base paper the algorithm of Convolutional neural networks evinced to have low computational complexity and SVM algorithm proved to be less efficient. The difficult task in this system is face recognition and also the accuracy rate is low. A facial feature can be used in the different computer vision algorithms like face detection, expression detection and many video surveillance applications. Recently, face recognition systems are attracting researchers toward it. This approach has three different methods such as SVM, CNN and KNN and it has been presented. For SVM and MLP based approach, the features are extracted using PCA and LDA feature extraction algorithms. In KNN based approach, the images were directly feed to the KNN module as a feature vector. The proposed approach shows the good recognition accuracy for KNN based approach. The SVM, CNN and KNN achieves the testing accuracy around 88%, 95% and 97% on self-generated database respectively. By this service the ambulance may not require the third person to indicate the address on their own. This approach is explained clearly with the help of a block diagram mentioned below as in Fig.2.

```
| Input Image | Face Detection | Face Segmentation | Feature Extraction | Matching with trained Image | Report Generation |
```

**Figure 2. Operating process of Existing System**

The above block diagram states firstly, two or more cameras depend on the need, and the size of the classroom has to be installed on the ceiling of the classroom from where it covers the entire area. Image captured from these cameras will be considered as an input to the system. There may be a possibility of getting image blurred due to movements of students, for better efficacy image can be upgraded using Generative Adversarial
The proposed system is to deliver an effective attendance system more efficient and make it a successful one. There exists an attendance system facility which identifies the student face by collecting the dataset and trained the image through many algorithms and generating the attendance report but the algorithm used posses the low accuracy to train the student dataset and the generated attendance report not maintained for future works. To overcome the issues, we propose an attendance system using face recognition with help of voice assistance, which is the best foot forward to reduce the loopholes encountered by the system. The Attendance system using face recognition consists of two phases, face detection and face recognition. The performance of detection algorithms such as Viola Jones and deep learning-based detection was compared and deep learning-based detection was preferred. On the recognition front, deep learning-based face recognition was used and optimum results were obtained. It was found that dlib using Convolutional Neural Networks (Fast-RCNN) yielded better results than dlib using Histogram Oriented Gradients (HOG) because HOG only detects a frontal image, while CNN detects faces from all angles and ensures no discrepancies during face detection and recognition. Also, it was observed that the speed of training datasets was higher in RCNN as it uses GPU as compared to HOG, which uses CPU for training the dataset as well as recognition. Fig.3 shows the Flow work of the system.

3.1 Challenges in Existing System

Though the existing system is good at being tested on these conditions KNN proved to be better by achieving the overall accuracy of 97% when tested on conditions listed above CNN achieved the overall accuracy of 95% and SVM achieved an accuracy of 88%. Viewing the aspect of time complexity, CNN is exposed to have low time complexity. It takes more time to train the collected student dataset.

4. PROPOSED SYSTEM

The idea behind the proposed system is to deliver an effective attendance system more efficient and make it a successful one. There exists an attendance system facility which identifies the student face by collecting the dataset and trained the image through many algorithms and generating the attendance report but the algorithm used possesses the low accuracy to train the student dataset and the generated attendance report not maintained for future works. To overcome the issues, we propose an attendance system using face recognition with help of voice assistance, which is the best foot forward to reduce the loopholes encountered by these systems. The Attendance system using face recognition consists of two phases, face detection and face recognition. The performance of detection algorithms such as Viola Jones and deep learning-based detection was compared and deep learning-based detection was preferred. On the recognition front, deep learning-based face recognition was used and optimum results were obtained. It was found that dlib using Convolutional Neural Networks (Fast-RCNN) yielded better results than dlib using Histogram Oriented Gradients (HOG) because HOG only detects a frontal image, while CNN detects faces from all angles and ensures no discrepancies during face detection and recognition. Also, it was observed that the speed of training datasets was higher in RCNN as it uses GPU as compared to HOG, which uses CPU for training the dataset as well as recognition. Fig.3 shows the Flow work of the system.
A student Dataset will be created by detecting their face in different Angles. Also the created Dataset for each student is mentioned with their Unique “id”. In this method, the student is asked for an unique ID by voice assistance. The capturing of image is done by OpenCV (Opensource Computer Vision Library) which is used to translate the Digital Image into the computer Understandable Text. Once the “id” is taken, the face of student is captured through webcam. With the help of dataset already created, the captured image is compared. The training of captured image is done in framework of TensorFlow and using Region based Convolutional NeuralNetworks (Fast-RCNN) yielded better results than dlib using Histogram Oriented Gradients (HOG) because HOG only detects a frontal image, while CNN detects faces from all angles and ensures no discrepancies during face detection and recognition. The matching is done by the captured image compared with the student created Dataset.

Once the comparison is finished, the Excel Sheet is automatically opened to generate the Attendance Report. The Generated student report is update on cloud server for future use. After the report is generated, the student is informed by the Voice Assistant that “Your Attendance is Marked”. Additionally the Voice Assistance have been included to give the Student verification while entering id.

Fig.4 below shows the architectural working the proposed system. It provides the accurate Student Attendance Marking System. It improves the accuracy rate and utilize less time to Train the dataset and used of Viola-Jones face detection algorithm to detect the image of the student. By using Viola Jones algorithm it brings the advantages of feature selection that is sophisticated and an invariant detector that locates scales. We can scale the features instead of scaling the image itself. Also, it was observed that the speed of training datasets was higher in RCNN as it uses GPU as compared to HOG, which uses CPU for training the dataset as well as recognition.
4.1 Dataset Creation

First step in every biometric system is the enrolment of a person using general data like their name and their unique biometric features as templates. A dataset of student is created with their names and unique id. The dataset have trained and stored in newly created folder with their required names. Image is captured from the camera and then student face is detected. After the face is detected, that face is cropped and then it is enhanced using histogram equalization and noise filtering so that exact features can be extracted. These unique features are then stored in the face database with certain id of that person.

4.2 Capturing the Images of Student

Before enabling the webcam the student are asked to enter the unique id then using Web camera attached that is continuously capturing images of students, detect the faces in images and compare the detected faces with the dataset and mark the attendance. In this technique we used viola jones algorithm for detecting the face of student. Using this technique enhance the efficiency and accuracy of the detection process.
4.3 Face Recognition & Marking Attendance

In this method, after the Face is detected, the capturing of image is done by OpenCV (Opensource Computer Vision Library) which is used to translate the Digital Image into the computer Understandable Text. With the help of dataset already created, the captured image is compared. The matching is done by the captured image compared with the student created Dataset. In this model we used Region based Convolutional Neural Networks (Fast-RCNN) yielded better results than dlib using Histogram Oriented Gradients (HOG) because HOG only detects a frontal image, while CNN detects faces from all angles and ensures no discrepancies during face detection and recognition. Then after the comparison is finished, the Excel Sheet is automatically opened to generate the Attendance Report. After the report is generated, the student is informed by the Voice Assistant that “Your Attendance is Marked”. Here the Voice Assistance have been included to get the Student unique id is entered correctly.

4.4 Update on Cloud Server

After the testing phase is completed the admin or the class advisor can upload the excel sheet of report in cloud server for future purpose. This cloud server is created very highly confidentiality because it can access with required login id and password. By this we provide a secure Attendance Report and generating the report in cloud server datas of the students will never lost and it can used for later retrieval.
4.5 Steps Involved In Our Methodology

1. Dataset Creation
2. Train the image and Pre-processing of the image (RGB to gray image, contrast adjustment, adaptive threshold)
3. Detection of face
4. Verification and Validation
6. Upload on Cloud Server

5. CONCLUSION

Human beings perform face recognition automatically every day and practically with no effort. Although it sounds like a very simple task for us, it has proven to be a complex task for a computer, as it has many variables that can impair the accuracy of the methods, for example: illumination variation, low resolution, occlusion, amongst other. In face recognition is basically the task of recognizing a person based on its facial image. It has become very popular in the last two decades, mainly because of the new methods developed and the high quality of the current videos/cameras. Our Proposed system can be more helpful and saves more time for the teachers instead of taking student attendance manually. Hence with the help of automated attendance system we were able to get our desired results and our lot of productive time is saved. We have also dismissed the probabilities of marking the proxies which occurs due to the manual attendance system. For such a complex process, we don’t require any peculiar hardware, only camera and database servers are required. Using this method we can replace all the old methods. Efficient and automatic attendance management is introduced in this research. The management of attendance in this method is more simple and the attendance is taken more accurately.

REFERENCES


16. Samridhi Dev, Tushar Patnaik “Student Attendance System using Face Recognition” The International Conference on Smart Electronics and Communication (ICOSEC 2020)


