ASISTENCIA – An AI Attendance Marking System

[1]V. Baby [2]K. Jhansi Lakshmi Bai [3]Aditya Cheruvu [4]Nagababu Veganti [5]Pavan Sesha Sai
[1]Associate Professor [2]Assistant Professor [3,4,5] UG (B.Tech Students)
[1]Department of CSE, VNR VJIET, Bachupally, Hyderabad, India.

Abstract **:** Face recognition is an important application of Computer-Vision branch of Machine Learning, owing to a wide range of uses in many fields. Identification of individuals in an organization for the purpose of attendance is one such application of face recognition. Maintenance and monitoring of attendance records plays a vital role in the analysis of performance of any organization. The proposal is to develop an AI Attendance Marking System to automate the traditional way of taking attendance. This system performs the daily activities of attendance marking and analysis with minimal human intervention, thus leading to saving of time in a class and make the attendance marking process much easier.

Indexed Terms **: Convolutional Neural Network (CNN), face recognition, computer vision.**

I. INTRODUCTION

As the educational needs of students at universities are increasing, and number of students joining universities are increasing, especially in Indian universities there exists a need to automate tasks like attendance taking and maintain a well-established system to track student attendance details. This increase both the amount of quality lecture time and also reduces the manual effort on the professor. Attendance of students is a very important and essential measure that a university relies on to ensure effective and quality education. Many Indian Universities have a mandatory rule to maintain the attendance of students above a percentage. Thus, it takes a few strict measures to enforce that in the real-world.

Several systems exist currently for the purpose of attendance in an organisation. In the traditional attendance marking system the professor has to call out all names and populate a list of students present in that class. After which, the lecturer marks attendance for all the students who attended that day in the attendance database system of the university. RFID tags is another popular solution to this problem. RFID tags method uses an RFID tag embedded in a card which is to be swiped by the individual of an organization to mark his/her entrance and exit times, thus recording the attendance details. The most recent solution to it is biometric systems. Biometric systems use fingerprint to uniquely identify a person and record their in-times and out-times.

However, there are several drawbacks in the existing systems. The traditional method is time consuming as all roll numbers need to be called out, and attendance details must later be manually posted into the attendance database which is a difficult task to do. In case of RFID and biometric methods, a long queue is formed after a class ends as all students leave the class at the same time. Additionally, RFID and vocal attendance methods can be easily violated, and these violations are unmonitored.

Thus, a simpler, and more effective way to carry out the task of attendance taking is to use modern image processing and face recognition algorithms, to identify and recognize multiple faces in images, captured by an automated image capturing system.

II. LITERATURE SURVEY

[1] "Attendance System based on Face Recognition" proposed a technique for face recognition using HOG (Histogram of Oriented gradients). However, there is another technique called CNN available for face recognition but CNN is more computationally expensive when compared to the HOG (Histogram of Oriented gradients).

"Real time Face Detection and Recognition in Video Surveillance" [2] proposed a system for face detection on video surveillance, but performing the face recognition on the live feed from multiple class rooms requires very high computational power and powerful systems to handle load. So, this prototype uses pictures taken by the camera in the class instead of video, which are easy to process when compared to video.

"Convolutional neural networks" [3] gave intuition about convolution neural nets and also about various types of convolutions like 1D, 2D, 3D convolutions and various types of available pooling layers and also explained the mathematics behind the convolutional and max pooling layers and loss functions used for training the neural nets.

III. PROPOSED SYSTEM

Proposed System involves usage of multiple-face detection and face recognition algorithms to identify all presentees in a picture taken by a photo capturing system.

3.1 Modules

- A photo capturing system to collect relevant group photos.
- A multiple Face-Recognition system using a class of computer-vision algorithms on the captured photos. Thus, attendance of many people in a photo is evaluated at a time, by saving a lot of time.
- An application (web and mobile) to manage the system and use the results. (tentative plan).

IV IMPLEMENTATION

4.1 Face Detection

Face detection is the first step in before performing the face recognition and it is used to identify the faces in the image.

4.1.1 HOG (Histogram of Oriented Gradients)

The HOG algorithm was first suggested for face detection by Shashuaet in the year 2005 It means Histogram of oriented Gradients. In this method the faces are recognized as the group of local histograms.

HOG is known as "quality descriptor". The task of a quality descriptor is to stereotype the object in such a way that same object produces as close as possible to the same characteristic descriptor when viewed in different circumstances.

These histograms engage the numbers at the area of image's local gradient directions that are head out. The realization of HOG is in order of the reception of gradient from image the set off histogram directions for specified locations. The normalization of histograms in the specified in location groups.

4.1.2 Face Detection using HOG method

The HOG method converts the given image into a black and white image, and computes gradient of the darkness of each pixel against the pixels around it.Then the computed gradients are represented in a matrix,Using pattern matching the obtained gradient matrix is compared with the training matrix., to find out faces in an image.

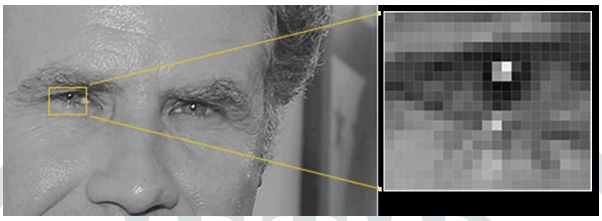

Fig 1 :Face Detection using HOG method

4.1.3 CNN

Earlier the HOG(Histogram of oriented gradients) and SVMs were widely used in face detection.
Both these algorithms do not perform well in situations when the images are little blurry and when the face in the image is tilted in some angle. Thus researchers worked constantly for better algorithm another algorithms.

CNN (Convolutional Neural Network), is such a powerful algorithm it widely used in computer vision algorithms like YOLO, ONESHOT Learning.

CNN consists of Convolution and pooling layers convolution layer is responsible for extracting features from the given image like straight edges, simple colours, and curves, this operation is performed used filters which has same depth as input image, The CNN starts by learning very low-level feature detectors, and as across the layers as its receptive field is expanded, learns to combine those low-level features into progressively higher-level features.

The next layer is polling layer, its function is to reduce the spatial size of the representation to reduce the number of parameters and computation in the network, and also to control overfitting. Pooling Layer operates independently on every slice of the input image and resizes it spatially, using MAX operation it is the most commonly used form of pooling. there are other types pooling available like Average pooling, Sum pooling these are according the problem requirement.

However, the computational cost for running the CNN is much higher when compared to the HOG face detector and CNN requires the powerful GPU for its working.

4.2 Face Recognition

Once the face is detected in the image the second task is to identify the person in that image, this is done by face recognition system.

ResNet is the one of architecture used to perform the face recognition.it contains convolution and pooling layers. When compared with other neural network architectures like Alex Net, Google net Resnet makes it possible to train up to hundreds or even thousands of layers and still achieves compelling performance.

There are various versions of ResNets exists based on number of convolution and pooling layers. ResNets are trained on 18,32,54,152 layers depending on the requirements. In this prototype uses the ResNet with 32 layers.

4.3 KNN: K-Nearest Neighbors Algorithm

KNN is the supervised Learning Algorithm in which the data is associated class labels, The Aim of this algorithm is to assign the new input data to the class that nearly matches it. Here K defines the number of nearest neighbors the algorithm should look for.

KNN can be used for both classification and regression predictive problems. But it is mostly used for classification purpose.

The encodings generated by the Resnet is passed to trained KNN-model on the saved images then new encoding is passed to KNN –model then it returns the class label of the pretrained encoding that nearly matches it.

4.4 System Architecture

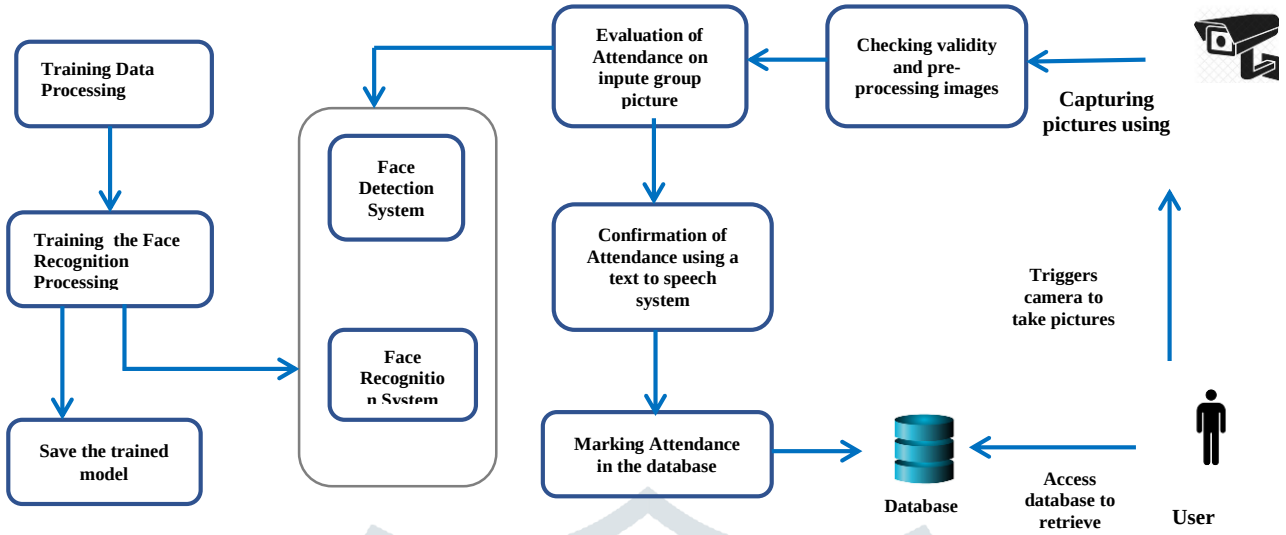

Fig 2: System Architecture

4.5 Hardware Requirements

Server Machine with the following minimum configuration:

- 8 GB of RAM, 16 GB recommended.
- Processor: Intel i7 6th gen and above is suggested.
- Nvidia GPU with at least 4 GB of VRAM.

Graphics Processing Unit must conform to Nvidia CUDA support and must be a high-performance GPU such as GTX 960m for mobile GPUs or GTX 1080 Ti for desktop variants with core count of at least a 1000. Raspberry pi 3B+ and above: Raspberry pi is a mini computer, which was initially designed and implemented to teach programming to school kids. However later realizing its potential as an IoT device, it has evolved into a great client system for performing small and simple tasks in an automated way, being placed at remote areas. Raspberry pi can be combined with various other electronic equipment such as cameras, infra-red sensors, sonar sensors for distance calculation, smoke detectors, etc. Raspberry pi camera module: The Raspberry Pi Camera Module v2 is a medium quality 8 megapixel Sony IMX219 image sensor custom designed with add-on board for Raspberry Pi, with a fixed focus lens. It's capable of producing 3280 x 2464 pixel static images, and also supports 1080p30, 720p60 and640x480p90 video recording. It is connected to Raspberry Pi by way of one of the small sockets on the board upper surface and uses the dedicated CSi interface, designed especially for interfacing to cameras.

- The face detection system is a CNN trained on millions of images to identify and locate multiple faces in a given image.
- The face recognition system is another CNN which outputs a vector of 128 facial features. Every person's face in the test set is encoded in this vector format and stored. The feature vector corresponding to a test image is evaluated and compared with test data to identify the face.
- The face recognition system is trained using the images of student faces against their roll numbers, and the training data is stored as a python pickle file.
- A raspberry pi with a camera module is used to collect and preprocess images and sent to a server for evaluation of attendance.
- The server processes the images and prepares a list of absentees, which is confirmed or modified by the professor (as a safety check), and the final attendance is posted in the database.
- The attendance is evaluated and raspberry pi reads out the absentees list aloud in the class to confirm attendance information.
- The user is given a chance to alter the attendance details via GUI, and the final attendance information is posted in the database.
- To user can use an end user application to operate the system and also retrieve attendance details.

V. RESULTS

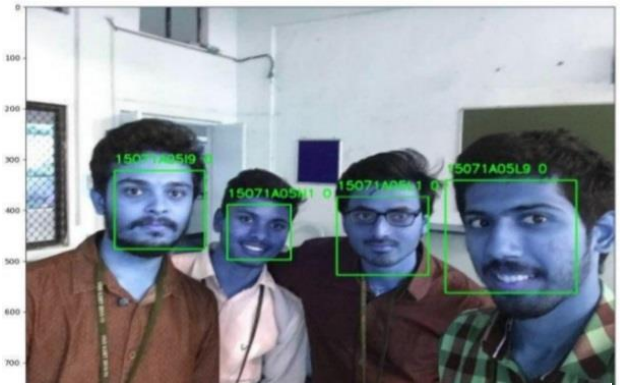

Fig. 3 Performance of the face detection and recognition algorithm on the trained faces.

Fig. 4 Performance of the face detection and recognition algorithm on the untrained faces

Performance of the face detection algorithm on large size inputs and empirical proof that it successfully classifies unknown faces.

VI. CONCLUSION

Based on the results obtained from the project, it may be considered as a prototype for a scalable and larger project to implement a full-scale attendance marking system for an institution.

The project involves various modules. The training module trains the model by generating 128-sided vector for each image which later clustered using clustering algorithm. This vector is saved for persistence as a pickle file. The face detection work on Hog or CNN algorithms to identify and recognize all known people in an image works on Resnet. Raspberry pie module with controls the camera for capturing the images and pre-process them and sends to server. The server performs the face detection and face recognition on the images and finds out all the people in the class.

REFERENCES

[1] K.Senthamil Selvin , P.Chitrakala , A.Antony Jenitha, " Face Recognition based Attendance Marking System".

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition".

[3] Navesh Sallawar, Shubham yende, Vaibhav padgilwar, Vishal kale , parag gorlewar, Gaurav Varma "Automatic attendance system by using face recognition".

[4] Berkant BAŞ, "Implementation of hog edge detection algorithm onfpga's".

[5] Navneet Dalal, Bill Triggs, "Histograms of Oriented Gradients for Human Detection".

[6] Database System Concepts: Textbook by Avi