Automatic Attendance Monitoring System Using Deep Learning

Shanmuhappriya M[1], Dr. Sudha Sadhasivam G[2]
{20mz33@psgtech.ac.in[1], gss.cse@psgtech.ac.in[2]}

UG scholar, Dept. of Computer Science and Engineering, PSG College of technology, Coimbatore, India[1], Professor and Head, Dept. of Computer Science and Engineering, PSG College of technology, Coimbatore, India[2]

Abstract. The manual method of attendance management is time consuming and difficult to maintain. So, the process of management of attendance in schools and universities should be automated. Many biometric systems can be used to record attendance. Face recognition is used frequently. The proposed work aims to record attendance without human interference. In this method the camera is fixed in the classroom.Faces are detected from the captured image and then recognized using trained model. The attendance is then marked and verified by class teacher. The attendanceis marked and absenteeism is sent to the parents. The proposed tooluses Deep Convolution Neural Network(DCNN) in which Max-Margin Object Detection(MMOD) and Histogram of Orientation Gradient(HOG) in Dlib used for face detection. Facenet algorithm responsible for extracting high quality features from the given faces is used for face recognition. The model is trained using K-Nearest Neighbour(KNN) and the framework is loaded in Jetson nano for testing.

Keywords: Automatic attendance, Face recognition, FaceNet, Deep learning, Jetson nano.

1 Introduction

Attendance monitoring and management in schools and universities are done using conventional methods. The traditional way of marking attendance is to call the student's names one-by-one in chronological order and mark the attendance in the attendance book. This attendance book is maintained for every class.This attendance is used to calculate the attendance percentage of the students for the final examination. It is a day-to-day routine work. The process of taking attendance and calculating the attendance percentage is done by the class teacher. This older approach of marking attendance during class is time consuming tiring task and prone to error. There is a possibility of misplacing the attendance book. Also, there exist higher chance for the students to cheat by giving proxy while answering the attendance.As, calling a person by name for marking attendance takes lots of time, then comes several techniques to reduce the time taken for attendance marking. One of such methods is card swipe system. This system allows the students to use swipe cards which is compatible with magnetic card reader. The attendance is stored in the database and the absenteeism is marked as well. The Radio Frequency Identification(RFID) based attendance system is used to cut off the time taken in marking the attendance. The RFID based attendance system uses three components to do this process. The tags which is the RFID cards, the RFID reader, and

the host computer. The RFID tag has an integrated antenna and memory. The RFID tag's memory contains all the information written on it. The rectangular plastic card that encloses the tag is very similar to the bank's ATM card or credit card. This RFID based attendance system has a reader which reads the data of the RFID tag. The reader emits radio waves. The tag as a response to this radio waves sends the information present in it. The host computer in the system stores the attendance of the students. It then evaluates the attendance information of the students which is obtained through the RFID reader and tags. Though this process takes only less amount of time to mark attendance, the chances of cheating is high. Students can use the cards of another students to mark the attendance even if he/she is not present. To overcome this disadvantage, a new technology of marking attendance has been employed widely, which is attendance marking using fingerprint. It is one of the biometric approach which is now used to mark attendance very widely. This approach is not only followed in schools and colleges but also in work place to ensure where the person is truly present or not. Another biometric authentication technology of marking attendance is using iris. The photographs of the student's eye is taken as the input. The iris part of an eye in the frame is clipped out. Then it is divided in to blocks. The blocks in turn is converted in to feature values. Then matching is performed to recognize the person. This method is widely used in immigration control, national ID and crime investigation. All these above mentioned conventional approaches of marking attendance is either time consuming or prone to proxies which allows the students to give a fake attendance for a person who is absent for the class. To overcome the difficulties faced by the aforementioned techniques for attendance management, a new biometric approach for automatic attendance management task is proposed in this paper. The proposed work usesface recognition to mark the attendance. Images are taken from the classroom by the high definition cameras mounted on the walls. From group photos only the face portion of the images which is the Region of Interest (RoI) of the faces are clipped and all the unwanted portions of the frames are removed. It is done by automatically locating the faces from the frames and drawing the bounding boxes around the faces only. With this image the class attendance of the students is marked in the database and the absenteeism is sent to the respective parents. To speed up the process of attendance marking the entire system is loaded in to the Jetson nano. Jetson nano is the Graphics Processing Unit(GPU) enabled platform used for video processing and live streaming, which is suitable for everyday use.

The paper is organized in an accompanying way. Section – II explains the work related to the paper. Section – III gives the detailed description of the proposed system architecture which explain the jetson nano implementation, the algorithms and the libraries used for face recognition.Section – IV shows the experimental results and finally the conclusion and the future scope has found in section – V.

2 Literature Survey

Varadharajan et al.[1] suggested an automatic attendance management system by using face detection approach in this paper. Eigen faces is set of Eigen vectors which are used in computer vision problem of face recognition. For a given image the background subtraction is applied to cut all the unnecessary parts of the image and take only the faces of the people. All detection techniques normally uses the background subtraction to detect the faces. The background can be easily subtracted from the frames of the captured video. As only the face of the person moves by variation in their face position and the other things in the background

remains unchanged, the static parts of the frames are subtracted. Background subtraction is done. Gray scale image and binary image are subjected to background subtraction. But most commonly first the image is converted to gray scale and then the background subtraction takes place. Background subtraction improves the accuracy of the detected faces. The detected face is then cropped. The cropped image of each face is taken for the comparison of images in database. The proposed work uses face recognition to identify the detected faces. Eigen value approach is used in this paper to recognize the detected faces from the frames. The proposed system has accuracy of face detection is 93% and the accuracy of face recognition obtained is 87%. Face recognition is the difficult task in this research and could be improved further. The demerit of the proposed system is even though eigen face approach of recognizing faces is simple, speedy and also has learning ability, the face recognition for large volume dataset is quite a tedious process. It work well for frontal face images with sufficient light. For the face images with variations in head positions the accuracy of the result is less.

Omar et al. [2] suggested an approach for class attendance management system by using face recognitionmethod is used. The proposed work uses Haar-featured based cascade algorithm for face detection and Local Binary pattern(LBP) is used for face recognition. The faces of the students sitting inside the classroom are captured using the cameras. Raspberry Pi runs Raspberry Linux Operating System which is installed in micro SD card. The raspberry pi is connected with the camera and the 5-inch screen for the purpose of displaying. The dataset holds images of eleven people or students. By operating with the aforementioned tools and techniques the system achieved the accuracy rate of 95%. The demerit of the system is LBP is prone to light sensitivity. When the light is on the value of the image pixels increases and when the light is off the value of the image pixels gets decreased. Therefore, it increases the distance between the trained data and the newly captured image of the same person, making the system difficult to recognize.

Serign et al. [3] suggested an improved face recognition algorithm and its application in attendance management system. This approach combines Local Binary Pattern algorithm with advanced image processing techniques such as adjusting contrast levels, bilateral filter, histogram equalization, image blending to overcome some of the issues while recognizing the face images from the frames. These adjustments helps in improving the existing Local Binary Pattern algorithm. The accuracy of face recognition is compared by using several methods. Local Binary Pattern algorithm along with Support Vector Machine and Particle Swarm Optimization(PSO) gives the accuracy result of 96.54%. When the original LBP alone is used only 89.3% of accuracy is obtained. Local Binary Pattern along with Support Vector Machine and Dual Cross Platform(DCP) gives the accuracy of about 97.50% and the proposed method gives the accuracy of about 99.0%. The drawback of the research is it does not address the issue of occlusion and mask faces in facial recognition.

Clyde et al. [4] suggested an approach for automatic attendance management system using face recognition approach for the proposed work. The application called Droid Cam is used. This application runs on laptop which is linked with the android mobile phones. In this approach, the android device can capture the face images of the student for marking attendance. In the proposed work,the Haarcascade algorithm is used for detecting the faces and the Local Binary Pattern Histogram(LBPH) is used for recognizing the faces from the given image frames. The android smart phone captures the live video. Frames are obtained from the video. Haar Cascade is used to detect the Region of Interest(ROI) for each face. LBPH calculates the histogram values of each face. At last, the attendance is marked in the system. It is done only if the captured image matches with the trained image dataset stored in

the database. The demerit of the system is Haar feature is sensitive to pose variations, quality of the image captured and illumination changes.

Smitha et al.[5] suggested a face recognition based attendance management system for the proposed work. In the suggested method,Haar-Cascade classifier and Local Binary Pattern Histogram algorithm are used for face detection and recognition process respectively. The camera takes the live video of the classroom, through which the faces of the students are taken. The faces from the image frames are alone cropped to obtain the Region Of Interest(ROI) which is done by using Haar Cascade Classifier with OpenCV. Then the image frames converted from RGB format to Gray scale format images. Face is recognized using Local Binary Pattern(LBP). These LBPs are then converted into decimal value format and then histograms of all those decimal values are made. At the end, one histogram will be formed for each face images in the training data. The histogram of an input face image of the frames which is to be identified is calculated. Then it is compared with the pre-computed histogram values. If the histogram values matches the corresponding labels are taken as a result to identify the particular person/student. Webcam is used to detect faces. In low resolution, the system obtains 90% of accuracy, where in high resolution, the system achieves up to 94%. The demerit of the proposed system is that it is difficult for the system to identify faces when there exists pose variations.

Kumar et al.[6] suggested challenges in face detection and recognition techniques. HaarFeature Algorithm is used for face detectionand Local Binary Pattern Histogram(LBPH) for face recognition is used in the proposed model. OpenCV is used to perform operation. LBPH is a binary based best texture operator. The accuracy level of LBPH is 89%.

Patil et al. [7] suggested an approach for automatic students attendance marking system using image processing and machine learning. Face detection is done by Viola-Jones algorithm. For feature selection Linear Discriminant Analysis(LDA) is applied. K-Nearest Neighbor(KNN) and the Support Vector Machine(SVM) is used for face recognition. In the proposed work the face images of the students are acquired through camera. By using histogram equalization the image processing is handled. Then the features of the student's faces from the frames are extracted. Then the vector dimensionality is reduced by Linear Discriminant Analysis(LDA) algorithm. After the feature extraction and dimensionality reduction, The face recognition is done by Linear Discriminant Analysis, K-Nearest Neighbor, Support Vector Machine to classify faces. Then the attendance taken by the systems is logged in the excel sheet. The accuracy of the LDA along with the KNN is 97% and the accuracy of the LDA along with the SVM is about 95%. The demerit of the proposed system is the time complexity. Also the dataset used is very small. To reduce the time complexity parallelism can be applied.

Samridhi Dev et al. [8] proposed a student attendance system for schools and colleges using face recognition approach. The proposed system makes use of Haar classifier, K-Nearest Neighbor, Support Vector Machine, Convolution Neural Networks, Generative adversarial networks, gabor filters. Students never stay at one position they will be moving around or changing the head position all the time. So the image captured by the camera need have to be a crystal clear one all the time. The images taken could be a blurry image also. To overcome this blurriness of the images and to upgrade those images as a good one to make the process efficient generative adversarial networks is used. After the process of generative adversarial networks an image is obtained which is ameliorated. This image is now used to detect the face location and for feature extraction with Haar classifier and gabor filters. After detecting the face locations, the face is recognized using three algorithms which is K-Nearest Neighbor, Support Vector Machine and the Convolution Neural Networks. After completing the face

recognition process the proposed system generates the name of the person and identification number of the students whose faces are identified on the input image frames. The proposed system achieves the high precision and less computational complexity. Accuracy of the system has been improved by using Gabor filters. Among the three algorithms (KNN, CNNN, SVM) K-Nearest Neighbor algorithm proved to have the highest accuracy of 99.27%. CNN has less computational time complexity and the SVM is very less efficient.

Nandhini et al.[9] suggested a Student smart attendance management system through face recognition using machine learning algorithm approach. Haar Cascade algorithm and Multi-Task Cascaded Convolutional Neural Network(MTCNN) model is used to detect faces from the frame. Decision Tree Machine Learning algorithm is used in this paper for recognizing the faces from the frames.Two or more frames are taken from the video. In this work, the Region of Interest(ROI) is marked using Haar cascade approach and MTCNN. After this step, the faces are detected for each frame of the video. Decision Tree Machine learning algorithm is employed in this proposed system for comparing with pre-trained images. The accuracy is calculated for three different datasets. For AT & T dataset the accuracy calculated is 92.89%. For Yale dataset, 93.12% accuracy is obtained. 92.87% is obtained for Facepix dataset. This work can also be further extended to mark attendance for similar faces such as identical twins where a very small and minute change can be noted from the faces of the identical people. The faces with different angles are hard to be identified.

Xue Lv et al.[10] proposed an Application of Face Recognition Method Under Deep Learning Algorithm in Embedded Systems. An optimized Multi-task Cascaded Convolutional Network (OMTCNN) algorithm for detecting the faces from the frames is used in this paper, and lightweight face recognition algorithm based on CNN(LCNN) for face recognition. The fundamental function of the system includes the registration and recognition of the face. The system first loads the face detection part and the recognition part. The face images are then collected. Then, the face detector determines the location of the face from the given image, and the face image feature is extracted through the face recognition method. At last, the result is obtained by comparing the feature with the face image captured with the database. For the proposed model Labelled Faces in the Wild (LFW) public test dataset is used. With OMTCNN 95.78% accuracy is reached. Without OMTCNN only 84.63% is achieved. The proposed multi-core embedded face recognition system has good real-time performance and a very low false detection rate for face images.

Mishra et al.[11] propose the multiscale parallel deep CNN (mpdCNN) architecture for the real low-resolution face recognition. The mpdCNN architecture uses different activation functions such as Rectified Linear Unit(ReLU), tanh, Sigmoid, Leaky ReLU and Exponential Linear Unit(ELU). Softmax activation function is used as the output layer of the network in all the cases. The SCface Database is used by the model. The accuracy obtained for low resolution is 88.6% and for high resolution the accuracy obtained is 99%. Multiscale parallel deep Convolution Neural Networks architecture is efficient, robust and lightweight for face recognition purpose on both real-time low-resolution and high-resolution images.

Tamilselvi et al.[12] suggested a face recognition system based on Hybrid Robust Point Set Matching Convolutional Neural Network (HRPSM_CNN) under unrestrained environmental condition. For pre-processing Laplacian of Gaussian (LoG) filter is used to overcomeirregular noise sensitivity. To pull out noise a smoothening process is done, wherein the low resolute images of Gaussian type are generated and imposed on image. The Viola Jones algorithm extracts features by Haar like structure. Then they are classified by cascade classifier along with Ada Boost algorithm.As the features of the face images are used, it reduced the time and increases the efficiency. HRPSM_CNN is used to recognize faces under

unpredicted and unusual situations. The system accepts image as an input and audio as an output. Face is detected from the input image and by applying HRPSM-CNN, the features are extracted and the face image is classified. The recognized face corresponds to the name which is obtained from the database. The text to audio conversion module connected to the visually challenged person helps to identify which person is standing nearby. Accuracy rate of 97% is achieved and under difficult situations like various lightings and weather conditions it shows better recognition of 95%. The demerit of this method is the system fails when the situation is uncontrolled.

Saleem et al.[13] suggested face recognition using facial features. the authors have tried to analyze the aptness of two face recognition algorithms namely facenet and facial landmarks. The features are extracted, pre-processed and converted into an array and the ROI is matched to identify the person. When the image is given as an input the frame resolution (640, 480) is verified. If not the next frame is taken. Pre-processing is done by BGR to RGB illumination correction. Then the location of the face is extracted and encoding is done. The Feature extraction is done by calculating the distance between eyes and nose and length of the face done by Histogram of Oriented Gradient(HOG). Now the new image for testing is given as input and the pre-processing, encoding, and feature extraction is done. Feature comparison is done with the training model and the classification is made. The algorithm is trained on 96 images with different persons and it takes around 2 seconds per image to train the algorithm on the image and once trained, the system takes 2.50 seconds for classification. The accuracy of the system is 82.432%. The future scope the system can be a hardware deployment of the system that will be cost-effective and user-friendly. The time taken for training the algorithm can further be reduced if the algorithms embedded with You Only Look Once(YOLO) which detects and identifies various objects in the picture in real-time is employed. The performance of the proposed work can be improved if the training dataset of face images, size is increased further.

Wang et al[14] suggested a violence detection and face recognition system based on deep learning. Brute force face detection and face recognition models are established. This model is best suited for real-time public security defense system. The Conditional Neural Process(CNP) model based on SPP is proposed in this system. The Spatial Pyramid Pooling(SPP) based model belongs to the improved "cross space method". TheSPP-based Convolution Neural Networks eliminates the image pre-processing step. Thereby the training process in the proposed system is done with ease and also it is fast.This is considered to be a very critical task in many practical applications. The technique of face recognitionachieves 92% accuracy in crowd dataset and 97.6% accuracy in hockey dataset. ConvNet requires large datasets to process and train Neural Network. By using only the smaller dataset the system will face overfitting problem. Convolution Neural Network is slow due to the operations involved in the pooling layers. Training process will take more time, if the system does not have good GPU.

Farayola et al.[15] proposed aframework which isface recognition with deep learning approach.The proposed model uses Nvidia GeForce GTX 1060 of 6GB RAM. TensorFlow framework is incorporated for its flexible architecture. Convolution Neural Networks classifier is used to successfully achieve Face detection.Thereby making an approximate contrast on the instance. In initial part convolution operation is performed to pre-process the given input face images and pooling operation is performed to reduce the dimensionality or features. It is the layerin which the image features of the face images are extracted and then detected. In convolution layer, filter is applied to the given data and provides the activation result. Pooling layer is used for reducing the size of the input data. Max Pooling down samples an input data,

reducing the input dimensionality. At last the input is given to the fully connected layer of the Convolution Neural Network which resides at the final stage of the system. Every input neuron has a full connection to every activation unit. This process is flattened to make single vector of values such that each of it will represent a certain feature belonging to a label. The errors obtained is minimized by back propagation. The model is used forpre-trained VGG Face for face recognitionfrom the set of faces tracked in video or image capture the accuracy achieved by the system is about 97%. This approach can be applied to any camera footage system and can be used by security sectors to fight against crimes. The system has the ability to learn and train on very large datasets.

Pranav KB et al.[16] suggested a Design and Evaluation of a Real-Time Face Recognition System using CNN in his proposed work. The proposed system is Convolution Neural Network (CNN-DL) combined task of feature extraction and classification. The proposed CNN architecture is designed using Open Source Neural Network library called Keras running on top of TensorFlow. The hardware used is Lenovo make laptop with intel I5-7200 processor operating at 2.5GHz with NVIDIA GeForce 940MX GPU and 8GB RAM. The convolution layer includes (conv + ReLU). The real-time image is taken by using the cameras which is mounted on the walls.Then the Viola-Jones algorithm is applied for face detection. Then the detected face is converted in to grayscale and resized to 120x120 pixels. The result is fed to Conv layer 1 which has 32 filters of size 3x3 pixel. Then the next Conv layer 2 has different set of 32 filters of size 3x3 pixel. The output of conv 2 is fed to MaxPool function using a window size of 4x4 pixel. In dropout layer, it is observed that during evaluation that a drop rate of 0.5 yielded maximum accuracy for proposed application. Again, the output is passed to a set of conv3, pool2 and dropout layers. Then passed to flatten layer. Finally, the fully connected layer takes part for classification. Then face id is recognized. The recognition accuracy of the system can be obtained by optimizing the number and window size of the convolution filter layers and pooling layers.

Lizze D'Cruz et al. [17] suggested a contactless attendance system using Siamese neural network based face recognition. The contactless attendance system comprises of face recognition using one-shot learning on Siamese network, MySQL database is used for storing the attendance and Gmail is used for messaging the parents about the absenteeism. In the proposed model the system has two Convolution Neural Networks.The two CNN are symmetrical with two input images. One image is pre-computed and another image is a newly captured image from the webcam for recognizing the person. The closeness i.e. the similarity and the dissimilarity of the images inputted are measured by calculating the Euclidean distance. Triplet loss is used for calculating the loss function, which resides at the last layer of the convolution neural network. In this model, the input images are given with filter size and padding. Pooling layers of the convolution networks are responsible for dimension reduction, thereby reducing the computation time. The computational time is reduced by down sampling of images. The fully connected layer of the CNN is used for predicting the best label to describe the face image. The softmax classification is used for mapping. Back propagation is used for minimizing the errors. If the similarity score i.e. the distance between two face images compared is large / high, the image is unknown. If the similarity score is small, the face of the person is recognized. More convolution layers in the network results in more number of parameters. Hence a large dataset is required for this approach to prevent overfitting problem.

Dhanush Gowda et al. [18] proposed a face recognition based attendance system in his paper. The system uses live video stream for marking the attendance. OpenCV is used to extract the frames from the live video which is captured by the camera fixed inside the

classroom walls. The system uses Dlib for face detection and face recognition. For each face in the given frame, facial landmarks are detected and a normalized wrapped patch face is detected. Using the detected facial landmarks, the face descriptor is computed which is the 128 Dimensional vectors. The labels with face descriptor is saved in the disk and enrolled. For a new input image or the query image, the face descriptor is calculated by detecting the facial landmarks and the Euclidean distance is calculated between the input/query image and the already enrolled image. For the same person the distance between the face images i.e. the input image and the stored image is less than 0.6. The method recognizes multiple faces from the given frame in less time with high accuracy.

Sati et al.[19] suggested Face Detection and Recognition, Face Emotion Recognition through NVIDIA Jetson Nano. NVIDIA Jetson Nano, ideally suited as an IoT edge device because of its small size and connectivity options. Face detection is implemented using Deep Neural Network(DNN) face detector. For face recognition is one of the deep metric learning approach is used in this paper. A list of 128 real-valued numbers is the output feature vector used for quantifying the face in the Dlib facial recognition network that is trained using triplets. Nano is booted with the Operating System of Ubuntu with libraries like OpenCV, NumPy and Keras installed in it. For face detection, 99.96% accuracy is obtained for the face closer to the webcam. 99.31% for the detected face that is detected a bit farther away. 89.99% accuracy is obtained when lighting is dim. The face is detected with an accuracy of 72.87% where the face is not at a straight angle. For face recognition, accuracy of 79.85% and 86.41% in different scenarios, different backgrounds.

3 Proposed System

The above mentioned methods for face recognition is interesting but the efficiency is not up to the mark. Sometimes the system fails to recognize person. Failure occurs due to pose variations, illumination, occlusion and facial expressions which makes the face recognition process less accurate.The model discussed above takes more time to train the dataset and provides less accuracy results while processing large datasets. Smaller datasets lead to the overfitting problem and showed wrong results for same faces.The proposed system aims to recognize the faces of the students with better accuracy with minimal time complexity. Jetson nano is used to speed up the processing time and facenet algorithm is employed as the face recognition technique to directly map the features of the faces in to vector space. This section discuss the architecture of the proposed system.

4 System Architecture:

The hardware used in the proposed work is based on Nvidia's Jetson nano. The Jetson nano is chosen for the proposed application because of the compact size of the device and it supports parallelization to speed up the process. The jetson nano enables Graphics Processing Unit(GPU) for parallel processing, with 128-core Maxwell, CPU quad core ARM A57 1.43 GHz with memory $GB 64-bit LPDDR4 25.6 GB/s which has USB 4x USB 3.0, USB 2.0 Micro-B. For displaying purpose jetson has Display HDMI and a Display port. Several communication protocols for the device are GPIO, I2C, I2S, SPI, UART. It has the board power rating of 5V 2A, 5V 4A. Jetson nano has Solid State Drive(SSD) of storage capacity

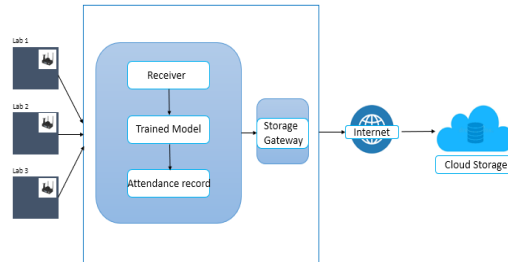64GB. Jetson nano provides high performance and manages the workload in a very small device.

Figure 1 Jetson nano for automatic attendance monitoring system

The proposed architecture diagram shown above illustrates the working of Jetson nano in attendance automation. Streaming is done through the camera. We use Pi-cam to record or to capture the real-time video of the students inside the classroom. The camera module is connected with the Jetson nano device for transmitting image and video. This point-to-point data transmission is handled by MIPI-Camera Serial Interface(CSI). The video data captured by the camera is transmitted through the Jetson nano over wireless communication network Wi-Fi to the host computer (Fig.1) . The host computer is already waiting for the data to receive from the Jetson nano for further processing. Once the data is received by the host computer, the videos are converted in to frames. The frames containing the students' face images are now passed to the training model. The trained model is capable of recognizing the students' faces by the pre-trained dataset. The model identifies the faces with corresponding labels and the attendance records are stored in the disk. The attendance data files residing on the disk can be simply copied to the cloud storage. This process is suitable for schools which is a small environment. But for universities and for large enterprise environment, a more scalable method of storing data is mandatory, as the environment tends to grow every year. The better solution for copying the attendance data files from the local disk to the cloud storage can be done through the storage gateway. The storage gateway acts as a bridge between the local host and the cloud storage. An internet connection is mandatory to send the data from the storage gateway to the cloud storage.
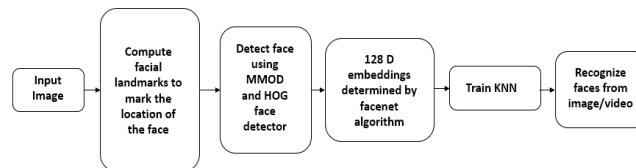

Figure 2 Block diagram for proposed system

The various modules in the proposed learning system includes

Face Detection : The system performs deep learning based face detection using Dlib model. The algorithm used for detecting the faces from the captured frames is Max-Margin Object Detection(MMOD) and Histogram of Orientation Gradient(HOG). MMOD is chosen for the system as the face detection model can run on Graphics Processing Unit(GPU) and make it super-fast. Max-Margin Object Detection is used to improve the face detection

methods like Histogram of Orientation Gradient. For more robust face detection this model is used. The model reads the input image which is provided. The Face detection algorithm localizes the face from the given image frames. The face locations are marked by eliminating all the unwanted static backgrounds and locate just the face alone. Then the face part is alone cropped or clipped for further operations like face recognition. Both the face detection model works in a similar fashion except the rectangle object returned by the MMOD detector.

Face Recognition : The proposed model uses FaceNet algorithm for quantifying the face. FaceNet algorithm is a Deep Convolution Neural Network(DCNN) which maps each input face image to 128 Dimensional vector space which are called face embeddings. FaceNet has the capacity to map face images from the frames to the Euclidean space. So that the distances corresponds to the similarity measure. The distance or the similarity measure between two face image of the same person is very small and for distinct faces it is large.

Training model : It is the process of training the classifier model to classify or to identify the face images belongs to which label which is already pre-trained and stored in the database. Comparison of the input images and the trained model takes place by the differences in the similarity measures among them. The classifier model used by the proposed method is K-Nearest Neighbor(KNN). The similarity is measured by comparing the vectors, which is done by distance metric. KNN uses Euclidean distance metric to do this process of finding the similarity measure between the datapoints.
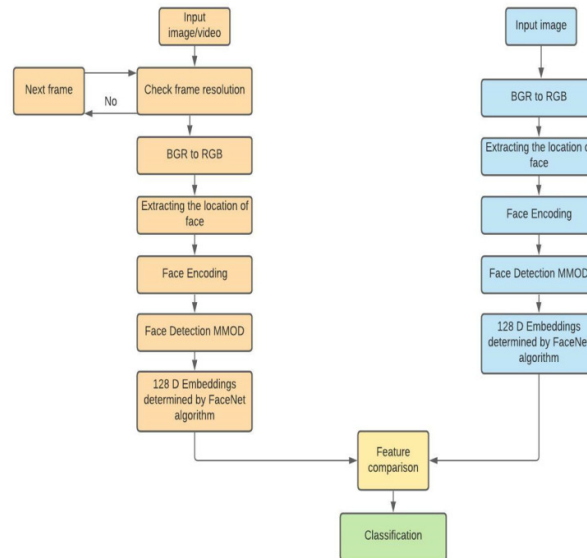

Figure 3 Process flow diagram

Image captured by the camera is given as an input to the system. The images are then converted from BGR format to RGB format for performing further operations on face recognition. The face location is marked and encoded in to an array further used to extract features from the image. Then then faces from the frames are detected by using the obtained facial landmarks. The features of the faces from the frames are detected using Max-Margin Object Detection(MMOD) and Histogram of Orientation Gradient(HOG) which is one of the Dlib's model for face detection. Then the 128 Dimension Embeddings are determined by the

facenet algorithm. The facenet model / algorithm is used to create a face embeddings which is the 128 D vectors for each face which is detected in the previous step. Facenet model / algorithm directly learns the mappings from the face images to the Euclidean space. The distances in the space provides the similarity measure of the faces. Then K-Nearest Neighbor is developed as a classifier model for predicting the identity of the given face from the inputted frames. Comparison is done between both training dataset and test dataset. Based on the level of the similarity the image is classified.

Libraries used

Nvidia Software Development Kit will be used to flash the Jetson nano to install the libraries and Application Program Interface(APIs). Additionally some of the python libraries which are required to install are Dlib, Numpy, Face_recognition, Python Imaging Library(PIL), Scikit-learn (Sklearn), pickle, Time.

For building the proposed system the libraries used are Keras OpenFace which is the face_recognition library is for using the facenet algorithm, Imutils is used for preprocessing the given input face images. Dlib library is used for detecting the faces from the image frames. It detects the face using landmark detection, which is done by predicting the position of 68 coordinates that maps the features of the face. The Python Imaging Library(PIL) is used to load the dataset images. Numpy is used for encoding the pixel values in array format. Computer reads all the image frames by pixel values which is in number format. Sklearn library is used to develop the KNN classifier, pickle library is used to save the model in to the disk. To predict the time taken to process the model for training and testing the data, we use Time library.

5 Experimental Results

For the proposed work, students face images are taken as the dataset. 11 students faces are taken each with 15 images for training the data and 5 images per person to test the data. The entire code is written in Jupyter notebook. The libraries mentioned in section – III are installed by using the anaconda prompt. The training data and the test data are stored in separate folders. The face_recognition.load_image() function is used to load the training data images. Through each training image the face locations are extracted by face_recognition.face_locations() function and the bounding box is drawn around the face. If the training data contains exactly one face in the image, the face_recognition.Face_encoding() function is used to calculate the face encodings of the loaded face images. Then the face encodings of each image corresponding to the label which is the student's name is added to the training data. Knn_clf.fit() function is used to train the classifier model.The model is then saved in the disk using pickle.dumb() function. While testing the model for students' face images, the model identified the faces accurately for all the students. The time taken to train the model is 5minutes 35seconds. To test the model, the time taken is 1minute 16seconds. To calculate the time taken we use time() function from Time library. The accuracy obtained by our model is 100% for the student's face dataset. The accuracy is calculated by the function accuracy_score() from metrics of scikit-learn library.

6 Conclusion

In this work, the automatic attendance management system uses face recognition to mark the attendance for which a Deep Convolution Neural Network approach is used. Max-Margin Object Detection(MMOD) and Histogram of Orientation Gradient(HOG) used to detect faces from the loaded images. MMOD is robust. MMOD while implemented on NVIDIA GPU Jetson nano makes the process super-fast. For quantifying faces the facenet algorithm is used. Here each input image is mapped to 128 Dimensional vector space. Facenet is a Deep Convolutional Neural Network which learns the face embeddings of the given faces through the training dataset. Then the identity of the given face is predicted by using the K-Nearest Neighbor algorithm. KNN algorithm is employed for searching process. Using the vector values for the given face images the similarity is measured and the vectors are compared using the distance metric. KNN uses Euclidean distance metric to find the distance between two face vectors or data points. The proposed work is implemented in NVIDIA's Jetson nano which performs the process in very less time. The workload is efficiently balanced by the small device. The work may be improved in future for faces wearing mask. Location dependent attendance management system can also be tried. Data can be maintained in blockchain for privacy.

References

[1] Varadharajan, E.; Dharani, R.; Jeevitha, S.; Kavinmathi, B.; Hemalatha, S. (2016). [IEEE 2016 Online International Conference on Green Engineering and Technologies (IC-GET) - Coimbatore, India (2016.11.19-2016.11.19)] 2016 Online International Conference on Green Engineering and Technologies (IC-GET) - Automatic attendance management system using face detection. , (), 1–3. doi:10.1109/GET.2016.7916753 . IEEE

[2] Salim, Omar Abdul Rhman; Olanrewaju, Rashidah Funke; Balogun, Wasiu Adebayo (2018). [IEEE 2018 7th International Conference on Computer and Communication Engineering (ICCCE) - Kuala Lumpur, Malaysia (2018.9.19-2018.9.20)] 2018 7th International Conference on Computer and Communication Engineering (ICCCE) - Class Attendance Management System Using Face Recognition. , (), 93–98. doi:10.1109/ICCCE.2018.8539274 . IEEE.

[3] Bah, Serign Modou; Ming, Fang (2019). An improved face recognition algorithm and its application in attendance management system. Array, (), 100014– . doi:10.1016/j.array.2019.100014. Elsevier

[4] Gomes, Clyde; Chanchal, Sagar; Desai, Tanmay; Jadhav, Dipti; Patil, M.D.; Vyawahare, V.A. (2020). Class Attendance Management System using Facial Recognition. ITM Web of Conferences, 32(), 02001–. doi:10.1051/itmconf/20203202001

[5] Smitha, Pavithra S Hegde, Afshin (2020). Face Recognition based Attendance Management System. International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org ISSN: 2278-0181 IJERTV9IS050861, Vol. 9 Issue 05, May-2020

[6] Ravendra Kumar, Rahul Sajwan, Praveen Kumar Jha, Sachin Sharma(2020). Challenges in face detection and recognition techniques. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-8, June 2020

[7] Patil, V., Narayan, A., Ausekar, V., & Dinesh, A. (2020). "Automatic Students Attendance Marking System Using Image Processing And Machine Learning". 2020 International Conference on Smart Electronics and Communication (ICOSEC). doi:10.1109/icosec49089.2020.9215. IEEE.

[8] Dev, Samridhi; Patnaik, Tushar (2020). [IEEE 2020 International Conference on Smart Electronics and Communication (ICOSEC) - Trichy, India (2020.9.10-2020.9.12)] 2020 International Conference on Smart Electronics and Communication (ICOSEC) - Student Attendance System using Face Recognition. , (), 90–96. doi:10.1109/ICOSEC49089.2020.9215441 . IEEE

[9] Nandhini R, Kumar P (2020) Student Smart Attendance Through Face Recognition using Machine Learning Algorithm International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-9 Issue-1, May 2020

[10] Lv, X., Su, M., & Wang, Z. (2021). Application of Face Recognition Method Under Deep Learning Algorithm in Embedded Systems. Microprocessors and Microsystems, 104034. doi:10.1016/j.micpro.2021.104034 . Elsevier

[11] Mishra, N. K., Dutta, M., & Singh, S. K. (2021). "Multiscale parallel deep CNN (mpdCNN) architecture for the real low-resolution face recognition for surveillance". Image and Vision Computing, 104290. doi:10.1016/j.imavis.2021.104290. Elsevier.

[12] M. Tamilselvi, S. Karthikeyan, "An ingenious face recognition system based on HRPSM_CNN under unrestrained environmental condition", Alexandria Eng. J. (2021), https://doi.org/10.1016/j.aej.2021.09.043. Elsevier

[13] Saleem, S., Shiney, J., Priestly Shan, B., & Kumar Mishra, V. (2021). "Face recognition using facial features". Materials Today: Proceedings. doi:10.1016/j.matpr.2021.07.402. Elsevier

[14] Wang, Pin; Wang, Peng; Fan, En (2021). Violence detection and face recognition based on deep learning. Pattern Recognition Letters, 142(), 20–24. doi:10.1016/j.patrec.2020.11.018 . Elsevier

[15] Michael Farayola, Aman Dureja(2020) Face Recognition With Deep Learning. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 07, JULY 2020 ISSN 2277-8616

[16] K B, Pranav; J, Manikandan (2020). Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks. Procedia Computer Science, 171(), 1651–1659. doi:10.1016/j.procs.2020.04.177 . Elsevier.

[17] D'Cruz, Lizzie; Harirajkumar, J. (2020). Contactless attendance system using Siamese neural network based face recognition. Materials Today: Proceedings, (), S2214785320380834–. doi:10.1016/j.matpr.2020.10.462 . Elsevier

[18] Dhanush Gowda H.L., K Vishal, Keertiraj B. R, Neha Kumari Dubey, Pooja M. R.(2020). Face Recognition based Attendance System. International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-018, Vol. 9 Issue 06, June-2020

[19] Vishwani Sati, Sergio M´arquez S´anchez, Niloufar Shoeibi, Ashish Arora, and Juan M. Corchado(2020). Face Detection and Recognition, Face Emotion Recognition Through NVIDIA Jetson Nano. Novais, Paulo; Vercelli, Gianni; Larriba-Pey, Josep L.; Herrera, Francisco; Chamoso, Pablo (2021). [Advances in Intelligent Systems and Computing] Ambient Intelligence â " Software and Applications Volume 1239 (11th International Symposium on Ambient Intelligence) || . , 10.1007/978-3-030-58356-9(), –. doi:10.1007/978-3-030-58356-9

[20] D. S. Vijayan, A. Leema Rose, S. Arvindan, J. Revathy, C. Amuthadevi, "Automation systems in smart buildings: a review", Journal of Ambient Intelligence and Humanized Computing https://doi.org/10.1007/s12652-020-02666-9