IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DEEP CONVOLUTIONAL NEURAL NETWORK BASED ATTENDANCE SYSTEM USING FACIAL RECOGNITION

[1]Dr. CN Sujatha, [2]B S Shreekar Goud, [3]K Neha Reddy,
[4]P Sri Harshavardhan
[1]Professor, [2]Student, [3]Student, [4]Student
[1]Electronics and Communication Engineering,
[1]Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad, India

Abstract: Attendance has been being considered a key factor by schools or colleges as a measure of participation of students in classwork. The traditional method is to mark the attendance of all the pupils manually which actually needs much time and effort. A smart alternative could be marking attendance by recognising the faces of the students. As online classes have become a part of our current education system, this method would go handy. Nowadays, almost everyone has access to gadgets like smart phones, tablets, personal computers or laptops. So the school/ college faculty can monitor the attentiveness of their students and obtain attendance report automatically. In order to achieve this, the information of students like their class, roll numbers and their faces is required primarily. The next step involves the training of the collected data. Finally, recognition of faces in order to mark attendance takes place.

Index Terms **- Facial Recognition, OpenCV, Haar Cascade, Deep Convolutional Nueral Networks, Local Biary Pattern Histogram.**

I. INTRODUCTION

The recent years have witnessed and undergone a great transition in the mode of communication. Due to covid-19, the whole world had to start online practices regardless of sector. Educational institutions are one of the foremost fields that have had an enormous impact and experienced compulsion to turn online. Our area of interest is attendance in online classes, more widely stating, online meetings where attendance plays a dominant role. Online mode of teaching faces several challenges that include ensuring active participation of students in the curriculum and their attentiveness to the lectures. This is where the regular method of noting attendance fails. Recognising the students present in the class and marking them present manually is both a time taking process and may give a scope for proxy; it does not promise the attentiveness of the students. The proposed method enables teachers to automatically get attendance reports of each class every day. It operates on the principle of Facial Recognition falling under the vast domain Image Processing.

Image Processing or DIP (Digital Image Processing) can be considered a process to manipulate an image by applying different transforms on it. The main motto of this method is either to get an enhanced version of the image or to extract a few useful or most important features from it. It comes under the vast domain of signal processing in which the input can be in the form of an image and the corresponding output is obtained in any of the forms: an image or characteristics or features, in other words, the information associated with the input image.

The paper will explain the application of the same in real-time where we collected the images of some students for testing the idea. They are used for training as the respective roll numbers are allotted. This whole part is done by coding in Python which is an open source language. A Python environment named PyCharm is made use of for the purpose; it enables the inclusion of all the required libraries like OpenCV and the algorithm employed for face detection is Haar cascade classifier. The features that are helpful in recognising the objects in an image are termed as Haar features. Cascading the classifiers enables the unimportant background to be set aside and focus on the prominent regions like face. Thus the system gets trained with the provided faces and the program assigns them the enrolment numbers of the students. Whenever the attendance of a class has to be taken, device cameras are turned on and if a face is detected, the corresponding roll number is displayed in the console window and the same gets marked present in the excel sheet created for attendance report.

1.1History

The techniques of digital image processing that are in use today were called as digital picture processing techniques. They are developed long back in 1960s at the Bell Laboratories, MIT (Massachusetts Institute of Technology) in the University of Maryland, and several other research facilities, with application to the satellite imagery, wire photo standards conversion, character recognition, medical imaging, photograph enhancement and videophone. In the beginning image processing held the purpose of improving the quality of an image. The aim of image processing was intended people in order to the improve people's visual effect.

The image which is input for processing is of low-quality in general and an improved quality image can be observed at the output. Usual image processing comprises of enhancement of image, image restoration, encoding and compression. American Jet Propulsion Laboratory (JPL) stands to be the first application to be successful. Image processing techniques like correction of geometry were used alongside gradation transformation, noise removal, etc. on number of lunar photos which were sent by the Space Detector Ranger 7 in the year 1964, considering the sun's position as well as moon's environment. It was a huge success since the mapping by the surface map of moon by the computer proved successful. Later, image processing got upgraded and more complex operations were performed on about 0.1 million photos the spacecraft sent back. Due to this maps which are used to indicate surface features of any piece of land, colour maps and panoramic mosaic of the moon are acquired, which accomplishes extraordinary results and laid a solid foundation for landing human on the moon's surface.

The price of processing changed into fairly high, however, with the computing gadget of that technology. That is modified in the 1970s, while digital picture processing proliferated as inexpensive computer systems and devoted hardware have become to be had. This brought about images being processed in actual-time, for a few dedicated problems inclusive of television requirements conversion. As well known-cause computers have become faster, they started to take over the position of devoted hardware for all however the maximum specialised and pc-extensive operations. With the fast computer systems and sign processors to be had in the 2000s, digital image processing has come to be the most common place shape of photo processing, and is typically used as it isn't only the most versatile approach but also the most inexpensive.

1.2Stages in Image Processing

There are different stages involved in digital image processing and are stated below.

Image Acquisition: In order to get into Image processing, it is essential to acquire an image at the primary stage. This is termed as Image Acquisition; an image fetched is converted into its digital form here. Pre-processing if any is done in this step and techniques include scaling.

Image Enhancement: It can be called a simple area of digital image processing where the features of interest or important parameters of an image get highlighted. These features can be the contrast in the image, image brightness, saturation etc.

Restoration of Image: Restoration refers to levelling up the way how an image appears that is improvement from a noisier version of the image taken as input.

Colour Image Processing: A mathematical model is employed to operate in the image. It is termed as colour space which is defined depending on the application. Intensity values are considered at this stage.

Wavelets and Multi-Resolution Processing: As the term suggests, the representation of the image is in terms of resolution and its degree. Wavelet transforms are used to detect the presence of noise in general.

Compression: It is often troublesome to store these images in their original qualities as space or memory constraints are considered. Hence it is important to reduce the storage size of the image. This process is achieved by encoding the image and the process is called image compression.

Morphological Processing: Basically, morphology refers to a set of tools or operations to deal with the shapes and structures that are present in an image. These components can be utilised in characterisation of the image. From this stage onwards, the outputs are image attributes rather than an image as in the previous stages.

Segmentation: Segmentation of the image into smaller segments or portions is done. It is a complex stage in digital image processing which is also time consuming but makes it easy for further processing steps.

Representation and Description: The next step is to represent the portions of an image in a form that could be processed by a computer system. Hence pixel representation can be observed at this stage. This data is initially in an unorganized form which is then described based on the image parameters such as regions, colours, textures, gradients etc.

Object recognition: The main purpose of this step is identification of the objects present in the image under processing. Many machine learning& deep learning algorithms are subjected to meet this purpose.

Knowledge Base: It is similar to database in any application; holds the whole information available about the image and its properties. Any stage of processing can fetch the required information from knowledge database. Its size or complexity is dependent on the resolution of receptor.

II. LITERATURE SURVEY

Looking back on the past years, it is evident form the research that students develop most of their cognitive skills in their educational phase of life; in other words, in their schooling. These skills developed at a tender age are what determine their progress in the future. Hence, it is essential to keep an eye on the attendance of students as it stands an important parameter [1] in the growth of students. Irregularity can result in a hindering effect on the same [2]. There are numerous methods for recording attendance of the students, one amongst them being through individual signatures. The procedure has various flaws, including the fact that it takes a considerable duration to check attendance; the attendance sheet might get lost, and the administration must make manual entries of attendance information into the database. To address this, the paper recommends a face-recognition-based [3] attendance system.

The fact is universally accepted that it is unfortunate for the world to have to witness the deadliest pandemic ever in recent years. The strike of COVID-19 stood challenging for educational institutions to keep going. Over ninety per cent students enrolled in different primary, secondary and professional courses were affected [4]. Almost all the institutions in the world had to adopt online mode of education programmes. In such a scenario, attendance becomes more prevalent and this system would ensure quality attentiveness of students in online classes. The framework comprises of a camera that records video of pupils in the class and communicates it to the management server. The picture of the individual is used as the input to dataset after which the image is pre-processed and the features of the face are retrieved using LBP (Local Binary Pattern) and Histogram [5]. 37 characteristics of shape and size are collected for each profile before it can be entered to the database. Much of this can be automated using our software. From a grey-scale image of the face, we discuss methods for extracting facial features such as the head contour, location of pupils, brows, and lips [6]. The approaches entail getting curves from detection algorithms first, then merging and generating them wherever needed.

In computer vision, face recognition is one of the problems with weight. Different algorithms are in use to serve the purpose and Haar cascade classifiers and LBPH algorithm which stand for Local Binary Pattern Histogram are discussed ahead. The LBPH approach outperforms existing Euclidean distance-based techniques like Eigen faces and Fisher-faces. Because of their resilience, we chose the Haar cascade for detection and the LBPH approach for recognition [7][8]. It withstands monotonic grey scale transforms with ease. Our system is evaluated utilising cases such as rate of facial recognition with and without a threshold in detecting faces.

III. OBJECTIVE

Attendance management system is a prototype i.e., a software developed to maintain the attendance of students or staff members regularly in the online meets like Zoom, WebEx, MS Teams etc. If a teacher takes attendance of every student in a class every period, it consumes almost 45 minutes of a day just for attendance. Even if we consider this way, we may not assure that this attendance is 100% prompt; there is a chance of proxy too. To overcome these issues, we came up with the solution to detect the faces of students using the camera embedded in their laptop screen and when their face is detected in an online class the attendance is automatically updated in the Excel sheet based on their class and roll number. This saves a lot of time and reduces the proxy issue for the teachers. We can confidently implement this technology and consider the attendance in any online meeting or class.

IV. METHODOLOGY

4.1Existing Methodology

The conventional method of taking attendance is confirming the presence of a student and marking his/ her attendance manually in a register. Faculty has to put a considerable time and effort in this task. It is almost the same in online classes; in addition, it does not ensure the attentiveness of students.

4.2Proposed Methodology

A lot of time can be conserved by introducing this technique in schools and colleges. The faculty just needs to be aware of handling and retrieving the attendance. It is a handy method in practice that anyone and everyone can get familiar to the operation. As soon as the camera is turn on, a known face is looked for. If one matches with any from the trained dataset of faces, the corresponding roll number or student is marked present for that class. Hence the generation of attendance report is automatic here as it gets uploaded to the excel sheet. This can be later accessed by the management for monitoring and other tasks. This way it proves to be better than the existing one due to its features such as the above stated.

V. IMPLEMETATION STEPS

The implementation is in four major steps:

- Creation of Dataset
- Training the Dataset
- Facial Recognition
- Excel-write

Figure 1, the flowchart depicts the sequential order of steps or operations to be performed in order to realise the model and test the same to obtain expected results in a flow.

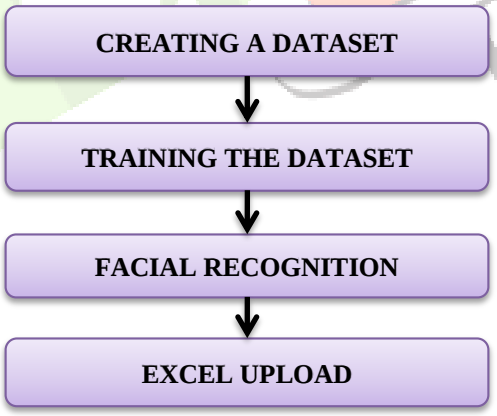


Figure 1 Implementation steps

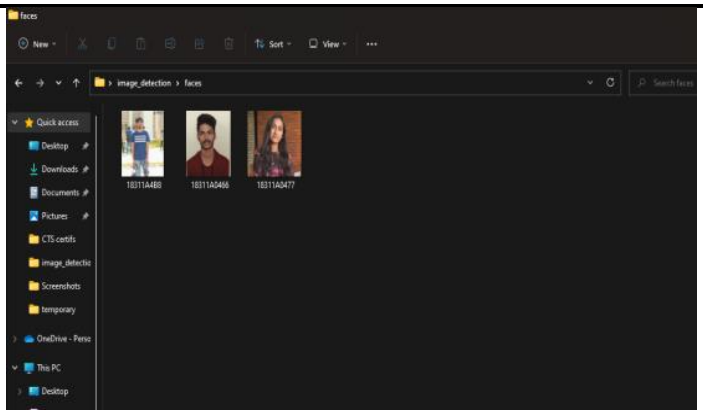


Figure 2 Dataset with images of students

5.1Creating a Dataset

Create the samples of users' faces using OpenCV by extracting the correct dimensions of the face. The following steps need to be followed while creating the dataset:

- Create a face detector system.
- Now, using Haar cascades, detect the faces within a particular frame.

Haar cascades are nothing but classifiers which classify the facial parts based on pixel values which then later converted into machine code language i.e. binary values. As OpenCV works with only grey scale images, it will store the samples in a grey scale manner. Extract the faces of around 30-40 samples for greater accuracy.

$$LBP(x_c y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c) \quad (1)$$

This conversion uses an algorithm; here, Local Binary Patterns Histogram- LBPH in short is used. Local Binary Operator is what does the operation in LBPH. An image must be in a square matrix, in other words, of a dimension MxM to be operated on by the above stated operator.

This feature looks at a picture element to its 8 nearest neighbours using the centre pixel intensity like an edge. If the neighbour's estimate is more considerable than or equal to the central value, it is made 1, otherwise it is made 0. As a result, we acquire a total of eight matched attributes from the eight neighbours. We acquire an 8-piece double number after uniting these attributes, which is a decimal number for our advantage is called pixel Local Binary Pattern. The pixel LBP cost is a decimal number in the range: 0 to 255. Monotonic grey scale transforms are not a problem for the LBP function. We must begin the training by acquiring the histogram statistics and loading the data of each face and facial characteristic into the computer file of YML format. The procedure for constructing a recognizer is as put below.

• Detect the faces in an image using the Haar cascades.
• Open the YML file generated in step 2 and read it.
• Provide the location of the YML file to be retrieved.

5.2Training

Training process includes a learning model. When this is trained, in other words, when it learns from a training data set, it is expected to make predictions about new data that is unknown. This dataset is referred to test data. A model can be any neural network. Deep CNN is what being used here. CNN stands for Convolutional Neural Network and the term 'deep' refers to the fact that there are at least 3 hidden layers in the network. DNNs (Deep neural networks) can be defined as the neural networks that comprise of many hidden layers. They are usually feed-forward in nature i.e. they form no loops.

Deep CNNs are known for their improved performance in Image processing and computer vision applications. It is attained due to their capability of extracting several features from the images. This is realizable since a deep CNN has 10 convolutional layers, 4 pooling. There is little requirement of pre-processing of input images. Any CNN in general has two major functions: Feature learning and classification. There are different layers responsible for each and are explained below.

Feature Learning: It includes the convolution and pooling layers and falls under training phase.

a) Convolution layer: It is the building block of a CNN and is responsible for the feature extraction from input images. This layer is a set of kernels or filters or feature detectors. Each filter is a group of learnable weights. An import Image is convoluted with a filter of dimensions MxM mathematically. The corresponding output is termed as feature map. Basic features in extraction are edges or lines, corners etc. As the number of layers increase, more complex features are captured.

ReLU activation function: Convolution layers are linear in nature and hence to deal with the moon-linearities in the input, non-linear activation function is used. It can be any of ReLU, Leaky ReLU, Sigmoid etc. ReLU (Rectified Linear Unit) function is in utility here. It can be represented as equation (2).

$$F(x) = \max(0, x) \quad (2)$$

Where x- parameter value

The maximum of input value and zero is the output. So the output for negative values of input is zero and when input value is a non-negative value, the same is reflected at the output as defined in equation (3).

$$\text{i.e. } f(x) = \begin{cases} 0, if\ x<0 \\ x, if\ x\geq \end{cases} \quad (3)$$

Pooling Layer: In general, it follows a convolutional layer. The more this entity is repetitive, in other words, the deeper the network, more complex features are extracted. It refers to the number of filters proportional to the volume depth. Receptive field also expands with depth.

The main motive of pooling layer is to reduce the dimensionality of the Feature map. It helps in minimising computational costs. Different polling techniques are in implementation based on the application. Two most popular techniques are max pooling and average pooling.

b) Classification: It is the application stage which is termed as inference.

Flattening: The pooled or downsized images in the form of matrix are converted into one dimensional array or a column vector to serve the purpose of classification.

Fully Connected layer: As the number of convolutional and pooling increases, abstract features get extracted which are usually not understandable by human intellect. FC (Fully Connected) layers come into picture in such situations to act as a translator between network language and humans. There can be multiple FCs in a CNN. An FC has multiple nodes; 128 here since in embedding, each face has 128 distinct measurements.

Softmax: Output of the ultimate FC layer is input to the Softmax output layer. All the values get normalised to fall between 0 and 1 at this layer. This behaviour makes it possible to append it to another network. Unlike Fully Connected layer, it has limited number of nodes based on the number of classes or categories the input images are to be classified into. For example, there are 30 students in a class; there will be only 30 nodes in the output layer. If any of these faces is recognised, the respective node indicates it.

5.3Facial Recognition

This system operates by locating and computing facial features from a given image in order to match a human face from a digital image or a video frame against a database of faces. It is commonly used to verify users through services such as ID verification. It functions on extracting features that are of use from an original image. Many such systems were first developed in 1960s starting as a type of computer program. These systems since then are in utility on smart gadgets as well as in other fields such as RPA which stands for Robotic Process Automation. FR (facial recognition) systems are classified as biometrics because computerised facial recognition requires the measurement of a person's physiological features. These techniques are less accurate than the biometric recognition technologies that deal iris and fingerprint, yet they are nonetheless useful and are in practice since they support no-contact processing.

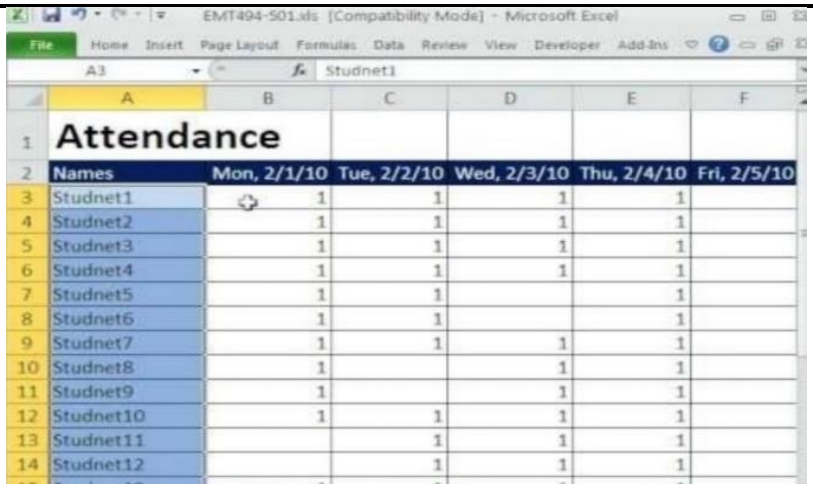
Modern human interface, security monitoring, and computer vision categorization have all used facial recognition systems. Public and private organisations use facial recognition technologies all around the globe nowadays. Their performance may differ; a few systems were abandoned in the past due their inefficiency. Face recognition systems have also sparked debate, with concerns that they infringe on residents' privacy concerns, frequently make inaccurate recognition, reinforce gender norms and racial bias, and fail to maintain crucial biometric information. Face recognition technologies have been banned in various locations across the US as a result of these allegations.

Meta revealed that it aims to close down Facebook's FR system, wiping the facial image information of over one billion active users, in response to rising societal issues. It would be one of the most significant revolutions in facial recognition deployment in the history of the technology. Face recognition software was first developed in the 1960s. Woody Bledsoe, Ch. Bisson and Helen Ch. Wolf collaborated on developing a computer that could detect human features. Because the positions of the facial characteristics in an image were to be defined manually before they could be utilized by the system for detection, their initial face detection and recognition program was called "man-machine".

User had to locate the coordinates of facial features such the eyes, mouth, nose, the widow peak in the hairline, in specific, pupil foci, the inner and outer corners of the eyes etc., on a graphic device. 20 lengths were measured making use of these coordinates.

In this way, a person might analyse around 40 photos every hour and construct a database of measured lengths. The distances for every image would then be automatically analysed, the difference computed, and the locked records returned as a probable match. Figure 3 shows the same. When a human face is considered, there exist certain points of interest which can be termed as facial feature points or landmarks. These are nothing but the crucial elements in a human face which differentiate one from another. This is why the process is also known as FFPD which stands for Facial Feature-Point Detection. Hence, such feature points are utilised by the system in learning the input faces and detecting or recognising them in the testing dataset or problem images.

Figure 3 Facial recognition



EMT494-501.xls [Compatibility Mode] - Microsoft Excel

File Home Insert Page Layout Formulas Data Review View Developer Add-Ins

A3 Studnet1

	A	B	C	D	E	F
1	Attendance					
2	Names	Mon, 2/1/10	Tue, 2/2/10	Wed, 2/3/10	Thu, 2/4/10	Fri, 2/5/10
3	Studnet1	1	1	1	1	
4	Studnet2	1	1	1	1	
5	Studnet3	1	1	1	1	
6	Studnet4	1	1	1	1	
7	Studnet5	1	1		1	
8	Studnet6	1	1		1	
9	Studnet7	1	1	1	1	
10	Studnet8	1		1	1	
11	Studnet9	1		1	1	
12	Studnet10	1	1	1	1	
13	Studnet11		1	1	1	
14	Studnet12		1	1	1	

Figure 4 Excel sheet for marking attendance

5.4Excel Upload

Excel is a spread sheet that supports computations and manipulation on the data entered in the sheet. It was introduced by Microsoft Corp. in the year 1985. It is an overlay of rows and columns. Intersection of two rows and two columns results in a box called cell. A cell is the basic unit of a sheet. By default, 3 sheets are available in an Excel document. It helps the project in following aspects.

- Providing a clear picture of attendance of a class as maintained in the registers conventionally.
- Maintain a report of daily attendance
- Enabling different computational tasks
- Assure effectiveness of online curriculum
- Analyse student productivity in the working hours

In the Fig.4, it can be observed that the first column is filled with the names of students whose attendance has to be marked; it covers the whole strength of the class. The succeeding columns hold the dates of working days on which online classes will be conducted. A cell in the excel sheet is marked '1' when a particular student's face is recognised on the corresponding day. It represents that he/ she is present on that day or period. Likewise, a blank cell represents the absence of the student.

VI. Test Results

Provided the training data, the model is then tested with another dataset known as testing data. With the knowledge of the faces of the students the model is trained with, it is expected to detect the face of a student belonging to the particular class and recognize him/ her with the corresponding enrolment number. When the code is run for testing, the following could be witnessed.

Running the program enables the integrated camera of the system or any camera module attached to it. When the camera is on, it looks for any face in its region of visibility. If a known face i.e. a face from the trained dataset is detected, the corresponding roll number will be displayed. In the fig. 5 shown, a student holding the roll number 18311A0466 happens to appear in front of the screen. Since his data is already trained, his face gets recognised and attendance is marked for the particular class. This is how our project works.

Figure 5 demonstrates the same. Results are obtained as expected when the model is tested with other students whose images are already learnt by the model. Each such face is recognised by the Convolutional Neural Network and the corresponding roll number is displayed below the face.

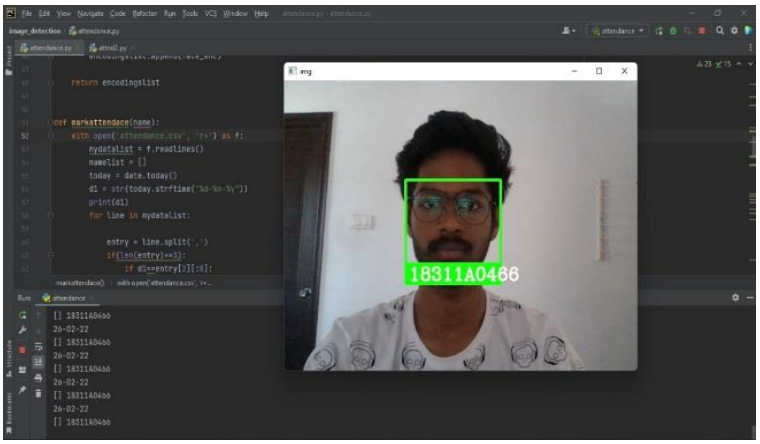


Figure 5 Test result: face of a student id recognized and roll number displayed

VII. CONCLUSION

As a result of testing under very robust conditions in this study, it is anticipated that real-world performance will be better. Models such as this one are quite good for feature detection since they use integral images and classifiers that make the detection process faster and ensure that no unnecessary objects are added to the frame when the machine learns. Haar Cascade with LBPH can recognize a face from any angle, regardless of whether the student is looking at the camera or tilting their head. The system can recognize faces even when the students are wearing glasses. It can also take attendance of multiple students at the same time. The system is pretty complex. This system deliberately gives high security to the user and it gives no space for any proxy attendance during the sessions which saves time and efforts of a teacher this can be utilized for the welfare of student. If we compare this model with the previous file system format it is way ahead in accessing data from any corner. A person without experience will not be able to mark attendance of multiple students at the same time. For a person with no experience, it won't be easy to mark attendance. The main disadvantage is that the dataset needs updating often since it's a facial recognition-based attendance system and it cannot operate without power. It is possible to improve the current model to exceed ninety percentage of efficiency. The project can address issues from image quality disparities and image size disparities. The current model operates at 85-90 per cent efficiency. This can be improved to exceed ninety per cent efficiency. An efficient model would recognise the person in the image even in low intensity light or low picture resolution. Also, handling the ambiguity in case of twins would be a great improvement.

REFERENCES

[1] D. D. Ready, "Socioeconomic disadvantage school attendance and early cognitive development: The differential effects of school exposure", Sociology of Education, vol. 83, no. 4, pp. 271-286, 2010.

[2] Smart attendance monitoring system (SAMS): a face recognition based attendance system for classroom environment Shubhobrata Bhattacharya, Gowtham Sandeep Nainala, Prosenjit Das, Aurobinda Routray

[3] Automatic attendance system for university student using face recognition based on deep learning Tata Sutabri, Ade Kurniawan Pamungkur, Raymond Erz Saragih International Journal of Machine Learning and Computing 9 (5), 668-674, 2019

[4] Post-Covid-19 education and education technology 'solutionism': A seller's market, Marko Teräs, Juha Suoranta, Hanna Teräs, Mark Curcher Postdigital, Science and Education 2 (3), 863-878, 2020

[5] An automatic attendance system using image processing Aziza Ahmedi, Suvarna Nandyal, The International Journal Of Engineering And Science (IJES) 4 (11), 1-8, 2015

[6] Craw, I., Ellis, H., and Lishman, J.R. (1987). Automatic extraction of face features. Pattern Recognition Letters, 5:183-187, February.

[7] A real-time face recognition system based on the improved LBPH algorithm XueMei Zhao;ChengBing Wei 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)

[8] Face Recognition based Attendance System using Haar Cascade and Local Binary Pattern Histogram Algorithm Bharath Tej Chinimilli;Anjali T.;Akhil Kotturi;Vihas Reddy Kaipu;Jathin Varma Mandapati, 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)

[9] de Haan, M., Johnson, M.H. and Maurer D. (1998) Recognition of individual faces and average face prototypes by 1- and 3-month-old infants. Centre for Brain and Cognitive.

[10] Digital Image Prcoessing-A Quick Review R. Ravikumar, Dr. V. Arulmozhi, International Journal of Intelligent Computing and Technology Vol. 2 January 2019 ISSN: 2457-0249

[11] Image parsing: Unifying segmention, detection and recognition Z Tu, X Chen, AL Yuille, SC Zhu- International Journal of computer vision, 2005- Springer