Attendance Monitoring System Using CNN

Akshita Hedaoo, Samyak Padwekar, Sakshi Aher
Vishwakarma Institute Of Technology, Pune
akshita.hedaoo16@vit.edu, samyak.padwekar16@vit.edu, sakshi.aher16@vit.edu,

***Abstract*— Conventionally, attendance records of students are taken manually by teachers by calling out the roll numbers in the class. The process is time-consuming , tedious and likely to get errors. Moreover, records of attendance are not easy to handle and store for the long- term. The research on attendance system have been going on for quite a long time, many solution have been proposed to make this system smart, but all those system have many drawbacks. In modern times, face recognition is the important aspect of computer vision. We proposed a system that uses the convolutional neural network (CNN) and recognition algorithms, which automatically detects the student when he/she sits in the classroom and marks the attendance by recognizing him/her. This system is developed by capturing real-time faces of students sitting in the class. Finally, the absentee lists are displayed in the CSV file. This System can be used in Government Schools and Colleges to monitor and validate attendance.The correctly recognized data gives us accuracy of approximately 74%.**

***Index Terms*—Class attendance, Face recognition, Deep learning(CNN).**

I. INTRODUCTION

Face recognition system is the area which more numbers of researches nowadays. Many methods have been discovered using different concepts for efficient face recognition. Face recognition is being highly used in areas like security systems. It can also be used for marking the attendance of students in a classroom. It is a monotonous and time consuming task to take the attendance manually everyday. Every classroom has almost 60 students and each lecture is of approximately 45 minutes. If we calculate average time taken for marking attendance manually in a single day is roughly around 1 hour. Many solutions have been already proposed to make this system smart yet efficient, but they have several drawbacks. One of the most popular method based on RFID is given in [1], but it has a limitation that it takes time in the same manner as students have to come one-by-one for marking attendance and it is very costly to implement. Smart attendance monitoring based on low energy Bluetooth device is proposed by the students of Delhi Technical University [2], but this system is intrusive as student can easily give proxy by bringing other student phone which is not an efficient way. Also students can give attendance from outside the class.

An approach based on scanning of fingerprint is proposed in [3]. It has a drawback that students have to stand for long time in a queue to mark their attendance and that takes ample amount of time for large number of students. Another solution which is based on barcode system is proposed in [5] where students have to show the barcode which is unique for each student on their ID card to mark attendance and that also is a very time consuming task. Several other approaches like speech based attendance system [6] and NFC (Near-Field Communication)based attendance system [7] are proposed but all of them have some kind of drawbacks. Smart attendance system based on face recognition is one of the non-intrusive method. It requires a camera to take a classroom picture and the image captured is given as input to our system and our system recognize each face from the input image and create a folder of the same. The camera is placed in such a way that every student is visible in the image captured. In the duration of 1 hour , it will take 4 images in the duration of 15 minutes and it will take the faces from each image accordingly. The process of recognizing the faces is divided into four steps. First, in our proposed system face detection is performed and stored in the folder. In the next step, the folder is passed to the CNN model and it recognizes the faces and returns the Gr No. The system is already trained with the database of students. Test image of the classroom is taken by the camera and is recognized by the system using haarcascade classifier and it creates an attendance sheet in CSV format. Then it marks the attendance of those students whose faces has been recognized .

II. RELATED WORKS

Nowadays, face recognition is one of the most deeply studied technologies in computer vision, with new approaches and encouraging results reported every year. So we try to improve attendance method by using the new technologies of face recognition. The acquisition of facial image is a very important part in the process of face recognition. Conventionally, researchers captured human facial image through shot pictures one by one, and then detected the human face from pictures to establish the face database. The [8] collect students' face data by taking pictures for each class, and then detecting the face in each pictures.It's time consuming, because of only a set of students' face data can be acquired in each class.In other words, if we need a large amount of face data, we have to take pictures in so many classes. However, in real scenes, the number of students is large and the class attendance always need to start from the first class. If we spend lots of time. Furthermore, the traditional methods are hard to guarantee the quality, and it is not realistic to complete the collection before the first class.In order to solve these problem, we propose a method of capturing face data through pre-class video capture,which can get a lot of face data quickly. The detail will be mentioned in Part III.With the development of face recognition, many different approaches are springing up every year. However, several classical methods are still used extensively. Such as Eigenface[9] , which be presented by

M. Turk and A. Pentland.The method of eigenface provides a simple and cheap way for face recognition, the training process is completely automatic and easy to code. Furthermore, it reduces statistical complexity in face image representation adequately. Once eigenfaces of the database are calculated, face recognition can be achieved in real time. However, the limitations of eigenface are also obvious, it is sensitive to lighting, scale and translation, and it requires a highly controlled environment. However, in actual scenes, the environment cannot be controlled, like light intensity, pose variation, facial expressions, image noise and etc. All of the change of external will affect the result of class attendance. In order to overcome the problem, we use the method of deep learning to finish face recognition by training the convolutional neural network which can learn the inner features automatically so that it can recognize faces, the detail will be mentioned in the following sections.

III. PROPOSED MODEL

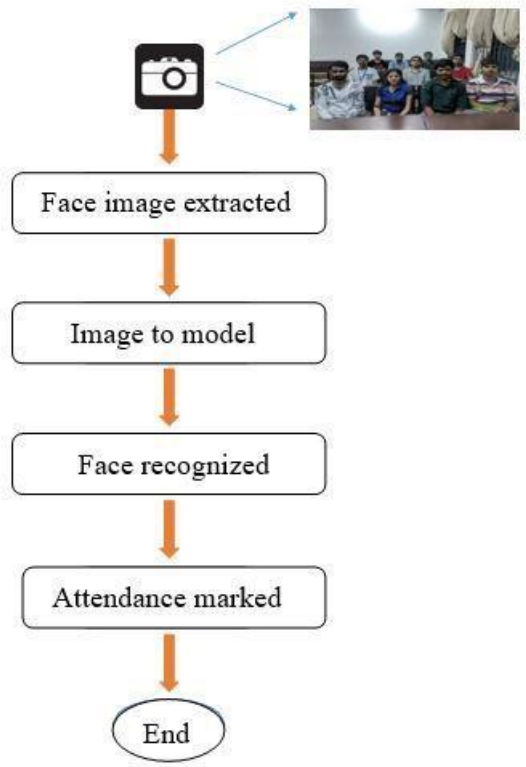

Figure 1: Model Flowchart

i. Database:
Videos of 72 students are captured which is then converted to images after various pre-processing like rotating the frames, taking snapshots at each second. Pre-processing also included cropping the image so that it could train on important features. The images available in the dataset contains the facial images from 0˚ to 180.˚

ii.Face Extraction
The image of classroom is then passed for face extraction. For extracting the face, we have used Harcascade Classifier A Haar Cascade is basically a classifier which is used to detect the object for which it has been trained for, from the source. The Haar Cascade is trained by superimposing the positive image over a set of negative images. The training is generally done on a server and on various stages.

iii. Storing images and testing:
The extracted images are stored in a folder as extracted and are passed to the pre-trained model along with two manually added convolutional layers. Transfer learning I done from pre-trained model to added layers.VGG-16 Model is a pre-trained model which gives the accuracy of the model. Conclusion can be made that CNN model's classifiers can classify images with more accuracy if transfer learning is used with pre-trained model.[11]

The ReLU layer
A ReLU layer does not change the size of the input, that is, x l and y share the same size. In fact, the Rectified Linear Unit (hence the name ReLU) can be regarded as a truncation performed individually for every element in the input:

$$y_{i,j,d} = \max\{0, x^{l}{}_{i,j,d}\}$$

with $0 \le i < H^l = H^{l+1}$, $0 \le j < W^l = W^{l+1}$, and $0 \le d < D^l = D^{l+1}$.
There is no parameter inside a ReLU layer, hence no need for parameter learning in this layer. Based on Equation 12, it is obvious that

$$\frac{dy_{i,j,d}}{dx^{l}_{i,j,d}} = [[x_{i,j,d} > 0]]$$

where [[.]] is the indicator function, being 1 if its argument is true, and 0 otherwise. Hence, we have

$$\left[\frac{\partial z}{\partial x^{l}}\right]_{i,j,d} = \begin{cases} \left[\frac{\partial z}{\partial y}\right]_{i,j,d}, & if\ x_{i,j,d} > 0 \\ 0, & otherwise \end{cases}$$

As soon as the network is trained and reached its expected goal,the training process stops and backpropagation starts to reduce the error.[17]

iv. Recognizing faces
Many algorithms can be used for face detection like using Viola-Jones[12] also know as ada boost algorithm[19],calculating image gradient in x and y detection[14] and also by consideing skin colour[15] but it also detect neck portion.So, HaarCascade Classifier is used. Haar-like features is used by Haar cascade classifier for human face detection. There are three formations of Haarlike features. The Algorithm looks for specific haar feature of a face. This detection takes the image and converts it into 24X24 window and smears each Haar feature to that window pixel by pixel. Initially, the algorithm requires a lot of positive images (images of faces) and negative images (images without faces) to train the classifier.
Then, these features are extracted. Features are numerical values determined from images that are used to distinguish

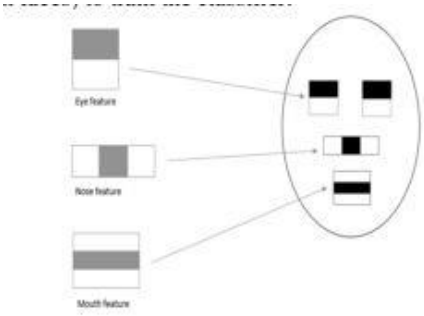

Figure 2 : Types of haar like features

one image from an0other each feature is a single value acquired by subtracting the sum of the pixels beneath the white rectangle from the sum of the pixels beneath the black rectangle .

Feature = Σ (pixels in black area) - Σ (pixels in white area)

The faces which were stored in folder is recognized and an array is formed where the students who were recognized is set as 1 and others as 0.

v. Marking the Attendance

After the recognition is done by the CNN module based on training set samples,the students which were recognized is marked as present in CSV file which was automatically generated using python.[16]

IV. RESULTS

1] Training and Validation Set:

Total of 72 classes are divided into training and validation with ratio of 8:2 that divides total images into 80% training set and 20% validation set. The total number of images available in the dataset is 71,009. There are two folders of which one folder contains the training set images which will be passed for training the model. The second folder is of validation set which will be used for the testing purpose.

2] Model Summary

Summary of CNN model with different layer types, output shapes and number of parameters is shown below. Total parameters and no. of images should be in proportion to avoid overfitting and underfitting.

3] Epochs and Training Accuracy:

Change in loss and accuracy with change in number of epochs is shown below. Accuracy increases as number of epochs increases as steps per epochs are same. The values of steps per epoch depends on the number of images each folder contains in the training set folders. Number of epoch depends on how much accuracy it gives. Mostly the number of epoch is 3 to 5.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d_1 (MaxPooling2	(None, 127, 127, 32)	0
conv2d_2 (Conv2D)	(None, 125, 125, 32)	9248
max_pooling2d_2 (MaxPooling2	(None, 62, 62, 32)	0
flatten_1 (Flatten)	(None, 123008)	0
dense_1 (Dense)	(None, 24)	2952216

Total params: 2,962,360
Trainable params: 2,962,360
Non-trainable params: 0

Figure 3 Model summary

4] Face recognition:

Face recognition is done by considering three different tables which are Table of Classes, Table of Files, Table of data values.

```
Epoch 1/3
600/600 [==============================] - 1926s 3s/step - loss: 0.3887 - accuracy:
0.8998
Epoch 2/3
600/600 [==============================] - 1969s 3s/step - loss: 0.0228 - accuracy:
0.9925
Epoch 3/3
600/600 [==============================] - 2152s 4s/step - loss: 0.0238 - accuracy:
0.9928
```

Figure 4.Epochs and accuracy

Key	Type	Size	
11810433	int	1	0
11810457	int	1	1
11810577	int	1	2
11810945	int	1	3
11811142	int	1	4
11812121	int	1	5
161054	int	1	6
161105	int	1	7
161164	int	1	8
161182	int	1	9
161211	int	1	10
161212	int	1	11
161248	int	1	12
161255	int	1	13
161290	int	1	14
161341	int	1	15
161401	int	1	16
161429	int	1	17
161470	int	1	18
161532	int	1	19
161551	int	1	20
1710112	int	1	21
1710592	int	1	22
1711034	int	1	23

Figure 5. Table of classes

Index	Type	Size	
0	list	0	[]
1	float32	(1, 24)	[[0. 1. 0. ... 0. 0. 0.]]
2	float32	(1, 24)	[[0. 0. 1. ... 0. 0. 0.]]
3	float32	(1, 24)	[[0. 0. 0. ... 0. 0. 0.]]
4	float32	(1, 24)	[[0. 0. 0. ... 0. 0. 0.]]
5	float32	(1, 24)	[[0. 0. 0. ... 1. 0. 0.]]
6	float32	(1, 24)	[[0. 0. 0. ... 0. 1. 0.]]
7	float32	(1, 24)	[[0. 0. 0. ... 0. 0. 0.]]

Figure 6 Table of data values

	A	B	C	D	E	F	G	H	I	J
1	Gr.No.	118104	'161164'	'161211'	'161212'	'161429'	'161290'	'161341'	'161401'	171034'
2	12/11/2019	0	0	0	0	1	0	0	0	1

Figure 7 Attendance marking in CSV

Table of Classes shows all 24 classes numbered from 0 to 23.Table of files shows 7 test images numbered from 0 to 6.Table of Data values gives the face recognition part.
From the results it can noted that the image with file no.0 and 2 in Table of files are incorrectly recognized and rest others are correctly recognized which gives us the accuracy of approximately 74%.

V.CONCLUSION

In this proposed paper, class attendance method is shown, which achieved through the Convolutional Neural Network (CNN). CNN needs a large number of data for training, so we design a strategy for facial data collected. The strategy of face collection solves problems of the traditional method which cannot collect facial data with high efficiency and high quality in the actual attendance. It can collect a large number of facial data by shooting video, which avoids collecting one by one and save a lot of time. At the same time, a large number of face images in the actual environment of complex and changeable are used to train the model of deep learning (CNN), the model can learn new features by training automatically, the new features can remove the inter-class difference as much as possible, such as light, noise, gestures and facial expressions, etc. In addition, the intra-class difference which can identify different humans is retained. The method using deep learning overcomes the effect of the change of external condition, the recognition rate in the actual situation is improved efficiently. With the improvement of recognition rate, the result of attendance will be more reliable and accurate. It also assists in overcoming the chances of proxies and fake attendance. In the modern world, a large number of systems using biometrics are available. However, the facial recognition turns out to be a viable option because of its high accuracy along with minimum human intervention. This system is aimed at providing a significant level of security. Hence, a highly pro-efficient attendance system for classroom attendance needs to be developed which can perform recognition on multiple faces at one instance. Also, there is no requirement of any special hardware for its implementation.

VI. REFERENCES

[1] M.K. Yeop Sabri, M.Z.A. Abdul Aziz, M.S.R. Mohd Shah, M.F Abd Kadir, "Smart Attendance System by suing RFID", Applied Electromagnetics 2007. APACE 2007. Asia -Pacific Conference on, 4– 6 Dec. 2007.

[2] R. Apoorv, P. Mathur, "Smart attendance management using Bluetooth Low Energy and Android", 2016 IEEE Region 10 Conference (TENCON), 2016.K.

[3] M.K. Yeop Sabri, M.Z.A. Abdul Aziz, M.S.R. Mohd Shah, M.F Abd Kadir, "Smart Attendance System by using RFID", Applied Electromagnetics 2007. APACE 2007. Asia -Pacific Conference on, 4– 6 Dec. 2007.

[4] N. I. Zainal, K. A. Sidek, T. S. Gunawan, H. Mansor, M. Kartiwi, "Design and Development of Portable Classroom Attendance System based on Arduino and Fingerprint Biometric", proc.of The IEEE 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M).

[5] K. Lakshmi Sudha, Shirish Shinde, Titus Thomas, "Barcode based student attendance system" International Journal of Computer Applications (0975 – 8887) Volume 119 – No.2, June 2015.

[6] S. Dey, S. Barman, R. K. Bhukya, R. K. Das, B C Haris, S. R. M. Prasanna, R. Sinha, "Speech biometric based attendance system", National Conference on Communications, 2014.

[7] Muhammet Baykar, Uğur Gürtürk, Ertuğrul Karakaya, "NFC based smart mobile attendance system", 2017 International Conference on Computer Science and Engineering (UBMK)

[8] Mehul K Dabhi, Bhavna K Pancholi, "Face Detection System Based on Viola - Jones Algorithm" International Journal of Science and Research (IJSR), Volume 5 Issue 4, April 2013.

[9] Turk M, Pentland A, "Eigenface based facial recognition", Journal of cognitive neuroscience, Volume: 3, Number: 1, 1991.

[10] Xiao-Yong Wei, Zhen-Qun Yang, Mining in-class social networks for large-scale pedagogical analysis, MM '12 Proceedings of the 20th ACM international conference on Multimedia,2012

[11] Abdullah,Mohammad S.Hasan, "An application of pre-trained CNN for image classificstion" 20th International Conference of Computer and Information Technology(ICCIT) 2017.

[12] Mehul K Dabhi, Bhavna K Pancholi, "Face Detection System Based on Viola - Jones Algorithm" International Journal of Science and Research (IJSR), Volume 5 Issue 4, April 2013.

[13] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, volume 2, p886-893, 2005

[14] Satyajit Kautkar, RahulKumar Koche,Tushar Keskar,Aniket Pande,Milind Rane, Gary Atkinson, "Face

recognition based on ridgelet transforms" Procedia Computer Science 2(2010) 35-43,ICEBT 2010.

[15] Rajatkumar Chauhan,Vivekanand Pandey,Lokanath M, "Smart Attendance System using CNN" International Journal of Pure and Applied Mathematics,Volume 119 No.15 2018.

[16] Y. Wen, K. Zhang, Z. Li, and Y. Qiao. A discriminative feature learning approach for deep face recognition. In European Conference on Computer Vision (ECCV).

[17] Meftah Ur Rahman "A comparative study on face recognition techniques and neural network" Department of Computer Science,George Mason University.

[18] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet : A unified embedding for face recognition and clustering.

[19] Wagh, J Priyanka, Agruti Chaudhari, Roshani Thakare, kare, Shweta Patil, "Attendance system based on face recognition using eigen face and PCA algorithms", Green Computing and Internet of Things (ICGCIoT) 2015 International Conference.