AUTOMATIC ATTENDANCE MONITORING

[1]Abhishek Singh Rawat,[2]Akash Yadav,[3]Gopal Tibrewal,[4]Sayuj Koul,[5]Prof. Priyanka Paygude, [6]Prof Snehal Chaudhary
[1,2,3,4]Research Scholar, [5,6]Professor
[1,2,3,4,5,6]Department of Information Technology
[1,2,3,4,5,6]Bharati Vidyapeeth(Deemed to be) University College of Engineering, Pune, Maharashtra, India

Abstract : Attendance is essential part of education assessment system. Almost all schools and universities use conventional attendance sheet for attendance management which is laborious, time consuming and less efficient. Few of Universities have moved onto finger based biometrics for attendance monitoring but still it is not widely used due to expenses and setup system which requires hardware and software installation on the other hand facial recognition based attendance monitoring can be done with a camera and software no additional hardware is required. Moreover the cheap cost of camera and wide availability also makes attendance based on facial recognition more suitable and easily applicable. The performance and accurate nature of attendance management based on image recognition makes it more promising and reliable for such application.

We selected model based on deep neural network which was found to have better performance than other algorithms like support vector machine and classifier cascades.

IndexTerms- **Attendance, biometrics, face detection, face recognition, eye blinking, MTCNN, Inception-resnet**

1. INTRODUCTION

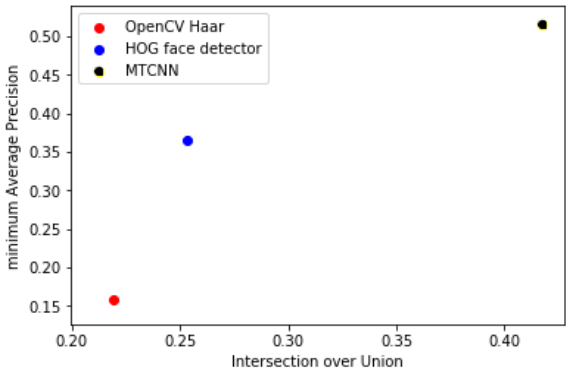
Biometrics is a way of using a person's physical or biological features to substantiate one's identity. Its purpose is to increase accuracy of identification. The oldest way is fingerprints for identification which is based on the fact that everyone has unique set of fingerprints. Biometrics also include other verifying methods like Iris scan [1], Retinal Scan [2], Voice Recognition [3] and Facial Scan [4].

Face recognition [4] is most efficient and accurate of all biometric techniques [5]. Like other techniques it also requires a hardware but comparing to hardware required by others like retinal scanner, iris scanner or fingerprint scanner hardware used in face recognition i.e. camera which is easily available and accessible. Furthermore, it is one of the most flexible way of authentication [6] as it also works even when subject is unaware. Face recognition works by scanning different facial indexes and mapping them with data in database.

1.1. PREVIOUS WORKS

In Indian Universities either fingerprint biometric system or attendance sheet are used for attendance monitoring and updating. The fingerprint system needs extra and costly hardware. The disadvantages of attendance sheet usage are wide known. It faces problems like fake attendance, sheet tearing, attendance cutting and many more. Similarly fingerprint system faces problems [6] with students having dry skin problem and skin conditions worsen in winters. Furthermore, they both are time-consuming, less efficient and their reliability is also debatable.

We used MTCNN (Multi Task Cascade Convolution Neural Network) [7] in our approach after comparing it with widely used models in real time applications. Viola Jones [8] detection algorithm is one of the earliest and still in use method for face detection. It is still used with real time applications. We compared IoU(Intersection over Union) and mAP(minimum Average Precision) scores of Haar Cascade Face Detector [8], HOG face detector and MTCNN in Graph 1 and Table 1.



Graph 1 mAP- IoU score

MODEL	IOU	mAP
OpenCV Haar Cascade Face Detector[8]	0.219	0.159
DLib HOG Face Detector	0.253	0.365
Tensorflow MTCNN Face Detector[7]	0.417	0.517

Table 1 IoU and mAP score of face detection model

Network	Top-1 Error	Top-5 Error
BN-Inception[11]	25.2%	7.8%
Inception-v3 [11]	21.2%	5.6%
Inception-ResNet-v1[11]	21.3%	5.5%
Inception-v4 [10]	20.0%	5.0%
Inception-ResNet-v2 [11]	19.9%	4.9%

Table 2 Single crop - single model experimental results.

Similarly for face recognition after observing the comparison (table 2) with various other available state of art networks we chose to use Inception-Resnet v2.

2. METHODOLOGY

The attendance monitoring is achieved by using MTCNN [7] (multi-task cascade convolution neural network) to detect face and Inception-resnet is used to recognize the face from the list of students. It was chosen due to better performance [8] over other models. The MTCNN is also trained to detect eyes. After eyes are detected eye blinking is tested to ensure that the detected face is not image. The eye-blinking is added as a security measure that guarantee's student's presence at scene and to abstain mischievous practices to spoof the attendance system. The Fig. 1 shows the steps followed in our approach.

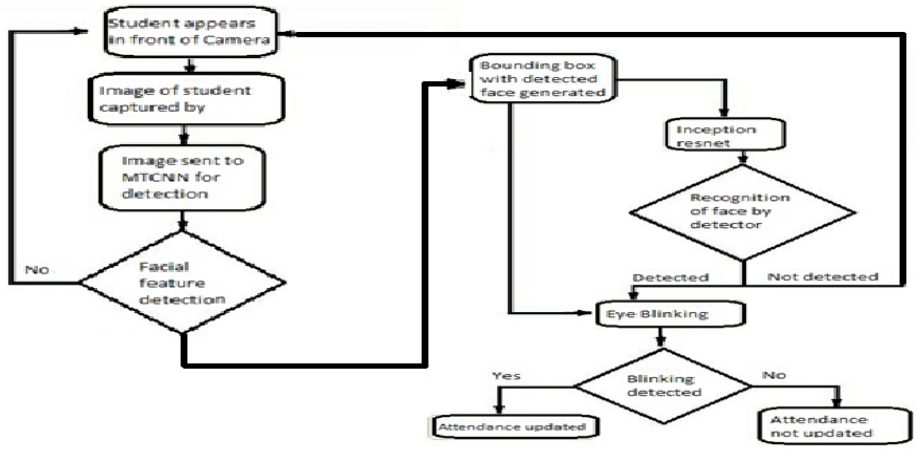


Fig 1 Process of Automatic attendance monitoring

The attendance is maintained in csv (comma separated value) file daily. Students who are present mark their attendance by appearing in front of camera which automatically detect and recognize their face. It also detects their presence by observing eye blinking which prevents the circumvention and validates the presence of student, next to which csv file is updated for recognized student.

2.1. FACE DETECTION

Face detection [9] is diversely used technique used in various industries and sectors for different applications depending on the need. It is used in countless ways with suitable variation. In our project we used MTCNN (Multi-Task Cascade Convolution Neural Network) to detect faces of students.

The MTCNN model consists of three networks [7] which are 1.) Proposal Network (P-Net), 2.) Refine Network (R-Net), 3.) Output Network (O-Net). The fig 1 shows the pipeline of MTCNN. The difference b/w P-net, R-net and O-net is on structure level. R-net consists of more Full-Connection Layer than P-net and O-net consist of more Convolution Layer than R-net. These networks are named after their operations. These networks are used in three phases in sequential manner. MTCNN phases:

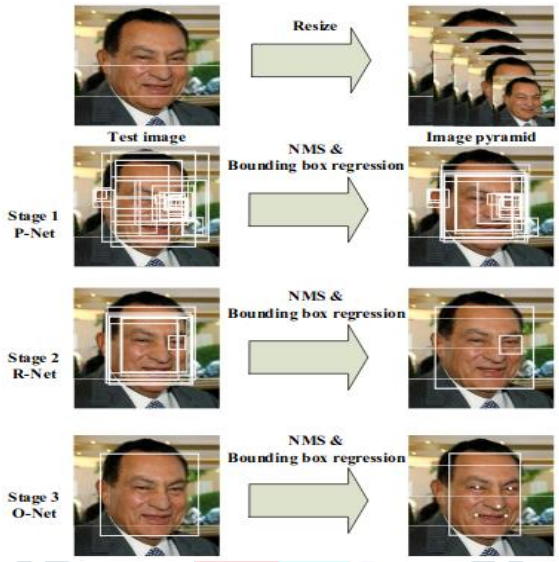


Fig 2 Working of MTCNN

1.) Phase–I: Image is passed and multiple copies of various scales are produced which are fed to P-net which is responsible for producing output with bounding boxes and a list containing confidence level of associated bounding box which is used to remove low confidence boxes. Following this, coordinates are converted to unscaled coordinates and using NMS (Non-Maximum Suppression) final bounding box is produced whose coordinates are then converted to real image coordinates and these are passed to second phase.
2.) Phase–II: This phase is similar to previous phase but here the false positive areas are removed following the similar process as that of Phase-I. Output of first phase is taken as input and padding is done and image is passed to R-net which produces result similar to P-net and after going through NMS and updating coordinates and bounding box, passes them to third phase.
3.) Phase-III: Working of this phase is similar to previous ones. The difference comes in the output of O-net which produces three results: 1.) Coordinates of the bounding box, 2.) Coordinates of the 5 facial landmarks, and 3.) Confidence level of each box. After going through NMS the final result is delivered.

The delivered result is a dictionary consisting of three keys: 1.) Box (Bounding box coordinates), 2.) Confidence (Box Confidence level) and 3.) Key points (Face features coordinates).

2.2 FACE RECOGNITION

The necessity of computer vision came when we faced problem which needed computer to detect objects. Further down the road the need to identify these objects was also felt. The face detection and recognition are application of object detection. The face recognition is part which comes after face detection. In many areas the next step after face detection is needed only detection isn't enough. Like in cases of authentication or monitoring it is not enough only to detect face. In our monitoring system we have used Inception-Resnet v2 model [10]. The Inception-Resnet is hybrid model of Inception model [11] and Resnet [12] model. These both models were introduced to solve the problems faced in two different axes of layer addition. The Resnet model is a deep residual network [12] which was made to solve the problem of vanishing gradients as layers are added and net is made deeper while the Inception model focuses on making network wider rather than deeper. It promotes addition of multiple filters at same level [11]. Following this principle inception module with dimension reduction was introduced.

The Inception-Resnet model gives advantages of both model and provides better accuracy for face recognition [10].

2.3 EYE BLINKING

While eye blinking can be used to determine human behavior and condition it also indicates motion in

	A	B	C
1	Abhishek Singh Rawat	10	
2	Abhishek Tiwari	8	
3	Akash Gupta	9	
4	Akash Yadav	7	
5	Aniket Choudhary	9	
6	Divya Diksha Kumari	5	
7	Garima Tripathi	6	
8	Gopal Tibrewal	10	
9	Kartik Shrivastav	9	
10	Mahak Satija	8	
11	Manan Taneja	7	
12	Nandini Kalra	5	

	A	B	C
13	Omveer Singh	6	
14	Pawan Jha	10	
15	Priya Ranjan	9	
16	Rahul Koul	9	
17	Sahibdeep Singh	10	
18	Sahil Jain	3	
19	Sayuj Koul	10	
20	Shikhar Goswami	7	
21	Shouvik Bagchi	8	
22	Shraddha Agarwal	7	
23	Yash Kesharwani	9	
24	Yashasvi Bhat	6	

Fig 3 Attendance csv

frame. We used this to differentiate between real and fake face. In our monitoring system we trained a network for detecting eye blinking by observing the frame for opening and closing of eyes. We used our net to first detect the eyes then if it detected opening and closing in successions then blinking is detected, thus face is not image and student is present in front of camera at that moment.

3. RESULT

A. TRAINING

In our model we first collected data of students in form of 10sec clip in which student had to move his face from left to right. This was done in order to provide different frames having face image from different angles. The data first passed through MTCNN which provided bounding box and also facial landmarks coordinates which were later used. The data then passed through Inception-Resnet v2 which stored the face embeddings in form of 128d vector. This vector was stored in dictionary which acts as a database for storing students data.

B. TESTING

The model was tested against numerous students consisting of those who had their image data trained, those who didn't and model was also tested against persons having facial similarity to students with trained image data. The model worked excellently providing success rate of more than 99% given that favorable conditions were met and even in case where there was lightening problem the model worked greatly with accuracy over 95.68%. The model's threshold was set 90% to make it able to distinguish between similar faces and because with the usage of HD camera problems like detecting face from low quality image was solved.

4. CONCLUSION

We used transfer learning and used state of art networks [7], [10] in our approach to meet the goal. We compared our model to the prevailing attendance monitoring system in Indian Universities [13] and found that usage of camera for this purpose not only ease the process but also don't face problems like the conventional methods [14]. We selected the algorithms for our model based on comparison done on their performance [9], [10], [11] on various parameters like precision, error rate etc. Using this approach is very easy, viable and achievable. There is no need for expensive and cumbersome installation in application of our method. It is highly collectable and non-intrusive way to monitor attendance..

REFRENCES

[1] S. Mann and L. M. Mann, "System and method for aircraft passenger check-in and boarding using iris recognition," ed: Google Patents, 2000.

[2] Condurache, Alexandru Paul, Johannes Kotzerke, and Alfred Mertins. "Robust retina-based person authentication using the sparse classifier." *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012.

[3] Sweetlin, J. Dhalia, V. Aswini, and R. Dhanusha. "Speech based attendance application register." *Recent Trends in Information Technology (ICRTIT), 2016 International Conference on*. IEEE, 2016.

[4] Hjelmås, Erik, and Boon Kee Low. "Face detection: A survey." *Computer vision and image understanding* 83.3 (2001): 236-274.

[5] Ibrahim, Dyala R., Abdelfatah A. Tamimi, and Ayman M. Abdalla. "Performance analysis of biometric recognition modalities." *Information Technology (ICIT), 2017 8th International Conference on*. IEEE, 2017.

[6] Masupha, Lerato, et al. "Face recognition techniques, their advantages, disadvantages and performance evaluation." *Computing, Communication and Security (ICCCS), 2015 International Conference on*. IEEE, 2015.

[7] Zhang, Kaipeng, et al. "Joint face detection and alignment using multitask cascaded convolutional networks." *IEEE Signal Processing Letters* 23.10 (2016): 1499-1503.

[8]Raya, I. Gusti Ngurah Made Kris, Agung Nugroho Jati, and Randy Erfa Saputra. "Analysis realization of Viola-Jones method for face detection on CCTV camera based on embedded system." *Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics), 2017 International Conference on*. IEEE, 2017.

[9] Dang, Kirti, and Shanu Sharma. "Review and comparison of face detection algorithms." *Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on*. IEEE, 2017.

[10] Szegedy, Christian, et al. "Inception-v4, inception-resnet and the impact of residual connections on learning." *AAAI*. Vol. 4. 2017.

[11] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

[12] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[13] Ahamed, Oli, and Md Rasedujjaman. "Biometric smart attendance kit with fingerprint scanner by using microcontroller." *Electrical & Electronic Engineering (ICEEE), 2015 International Conference on*. IEEE, 2015.

[14] Ghosh, Swamendu, et al. "Smart Attendance System." *2018 International Conference on Smart City and Emerging Technology (ICSCET)*. IEEE, 2018.