

AI-Powered Smart Facial Attendance System: Design, Implementation, and Performance Evaluation

P.Rajapandian¹, S.Tharani Vasan^{2*}

¹(Associate Professor, Department of Master Computer Application, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107. Email: rajapandian.mca@smvec.ac.in)

²(P.G Student, Department of Master Computer Application, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107. Email: [@gmail.com](mailto:tharanivasansr@gmail.com))

ABSTRACT

Background: The old ways of taking attendance like calling out names or using cards have some problems. People can pretend to be someone it takes a lot of work to manage and the information is not always correct.

The goal of this paper is to introduce a system called Smart Facial Attendance System that uses artificial intelligence to take attendance automatically. It uses computers to find and recognize faces in real time.

The system uses computer models like FaceNet and Support Vector Machine, to make sure it recognizes faces correctly. It also uses a website to manage everything.

The system can take videos. Get the information it needs from them.

The results show that the system is very good at recognizing faces it got it right 98.7 percent of the time. It can also do it quickly in just 0.34 seconds.

The system works well even when the face is not fully visible or at an angle.

In conclusion the new Smart Facial Attendance System is much better, than the ways. It stops people from pretending to be someone and makes it easier to manage. It is also an option because it is contactless and secure.

Keywords: Facial Recognition, Attendance System, Deep Learning, FaceNet, OpenCV, SVM Classifier, Biometric Authentication Real-Time Processing.

1. INTRODUCTION

Attendance monitoring is something that schools, companies and government offices really need to do. The old ways of taking attendance like calling out names

on a list swiping cards or using PIN numbers have some problems. People can cheat by having someone else mark them present it takes a time and the records can be easily altered.. With the help of Artificial Intelligence and computer vision we can now make attendance monitoring automatic, easy and very accurate.

The technology that recognizes faces has gotten good thanks to advances in special kinds of computer networks and learning methods. It is now practical and affordable to use in real-life situations. Unlike using fingerprints or iris scans face recognition does not require any contact and can happen in the background without interrupting what people are doing. This is especially useful in classrooms and workplaces.

This paper is about the AI-Powered Smart Facial Attendance System. It is a system that uses live video, special computer programs to detect and recognize faces a database that is accessed through the internet and a website that administrators can use to manage everything. The system can run on computers, including small devices like Raspberry Pi, which makes it affordable for any institution to use.

The rest of this paper is organized like this:

- * Section 2 looks at what other people have done
- * Section 3 explains how the proposed system is designed.
- * Section 4 talks about the methods and algorithms that are used.
- * Section 5 shows the results of the experiments and how well the system works.

* Section 6 discusses what the system cannot do and what could be done in the future.

* Section 7 summarizes the paper. The AI-Powered Smart Facial Attendance System is a useful tool, for attendance monitoring. The AI-Powered Smart Facial Attendance System can make attendance monitoring easier and more accurate.

2. LITERATURE REVIEW

The automated attendance systems field has changed a lot over the twenty years. At first these systems used fingerprints and RFID tags. These methods are not very good because they need people to touch something or be close to it and that can be a problem for hygiene. Also people can share their cards.

Some people, like Viola and Jones made a system that can detect faces in time. This system uses something called Haar cascade classifiers. It was used a lot in the days of facial attendance systems. It has some problems. It can make mistakes. It does not work well in lighting. Then Lienhart and Maydt made some changes to this system. They added some features that made it a bit better.

Then something happened. Deep learning came along. This changed everything for face recognition. Some people made a system called DeepFace that could recognize faces well as humans. It used a kind of computer program called a CNN. Google made a system called FaceNet that could recognize faces well and make a kind of map of each face. Another system called ArcFace made this map better.

Recently some people have made attendance systems that use learning. Kumar and his team made an attendance system that uses a CNN and can recognize faces well. This attendance system got 96.2% of the faces right. Rai and Tripathi made another attendance system that uses a kind of face detection and a ResNet-50 backbone. This attendance system got 97.1% of the faces right. It needs a lot of computer power. Fernandez and his team made an attendance system that works on phones and is really fast.

There is a problem with all these attendance systems. They do not work well in life. They have trouble when people's faces are partially covered or when the lighting is bad. Also most of these attendance systems do not

have a way for administrators to manage them and get notifications in time. The automated attendance system we propose, called SFAS fixes these problems with automated attendance systems. The SFAS attendance system is an improvement, over attendance systems.

3. SYSTEM ARCHITECTURE

The SFAS is made up of a main parts. It has four layers. These are the Capture Layer, the Processing Layer, the Application Layer and the Data Layer.

The SFAS has these layers to make it work properly.

Figure 1 shows what the whole system looks like at a level. The SFAS system architecture is pretty simple to understand when you look at Figure 1. The SFAS is basically the Capture Layer, the Processing Layer, the Application Layer and the Data Layer all working together.

3.1 Capture Layer

The camera system has one or more cameras that use USB or IP connections.

These cameras are placed at entry points.

The cameras take a lot of pictures thirty pictures every second.

Each picture is really clear it has a lot of pixels 1280×720 pixels.

The camera system works with the OpenCVs VideoCapture API.

This means the camera system can work with types of cameras.

When we use the camera system with cameras at the same time a special module makes sure that all the cameras take pictures at the same time.

This is important for the camera system to work properly with the cameras.

The camera system needs all the cameras to take pictures at the time so the camera system works correctly with the cameras.

3.2 Processing Layer

The processing layer is the part of the system where all the work happens.

It takes raw video frames from the capture layer. Does several things one after the other:

- First it prepares the frame: It makes the frame smaller, 224×224 pixels to make it easier to work with.

It also makes the light look more normal and removes noise using a blur.

- Then it finds faces: A tool called MTCNN looks at each frame. Finds where the faces are and what the eyes, nose and mouth look like.
- After that it straightens the faces: It uses what it learned about the face to make it look the same every time, which helps with recognizing people.
- Next it gets a code for the face: It sends the straightened face through a model called FaceNet, which gives a 128-number code for that face.
- it checks who it is: It compares the code to a list of codes for people it already knows using a special machine learning tool.

3.3 Application Layer

The application layer is what the users see. It has a part that is made with Flask. It helps the system work. This part is called the backend. It has some endpoints that show things like who is in the room right now and who was here before. The web dashboard is made with React.js. It has some nice features.

- * It shows a video of what is happening and it can recognize faces in real time.
- * It keeps track of who was and when and it lets you look at this information by date or by class or by the persons ID.
- * It sends emails or texts to people who're not here.
- * It has a way for new people to sign up.
- * It shows how many people are here over time so we can see patterns. The application layer and the web dashboard work together to make this system work. The Flask-based RESTful API is important, for the backend. The web dashboard is what people use to see what is going on.

3.4 Data Layer

Attendance records, face data and user information are kept in a PostgreSQL database.

The face data is organized using Faiss for searching.

This helps the system to quickly find faces.

The database works well with both on-premises and cloud setups, like AWS RDS.

This makes it easy to deploy the system in environments.

4. METHODOLOGY

4.1 Face Detection: MTCNN

MTCNN is a kind of system that uses three steps to find faces in pictures. The first step is called P-Net. It looks at the picture in many different sizes to find places where faces might be. Then the second step, which is called R-Net takes a look at these possible faces and decides if they are really faces or not. The third step, which is called O-Net draws a box around the face. Finds the important points on the face like the eyes and nose. MTCNN is really good at finding faces it gets a score of 0.968 on a test called FDDB, which's better than other ways of finding faces, like Haar cascades and single-stage detectors even when the picture is not very clear. MTCNN works well in conditions.

4.2 Face Recognition: FaceNet with Triplet Loss

FaceNet is a way to look at face pictures. It takes each face picture. Puts it into a special kind of space where similar faces are close together. This space has 128 directions.

The computer program that does this is taught using something called the triplet loss function. This function is like a formula that says:

L is the sum of the difference between the face picture of a person and a picture of the person and the difference between the face picture of a person and a picture of a different person, plus a little extra.

We use three kinds of pictures to teach the computer: the anchor picture, the positive picture and the negative picture. The anchor picture is the picture. The positive picture is a picture of the person, as the anchor picture. The negative picture is a picture of a person.

The computer tries to find the way to put all the face pictures into the special space. It does this by looking at different sets of three pictures and trying to figure out which ones are most helpful.

We started with a computer program that someone else had already taught to look at face pictures. Then we taught it some more using a set of face pictures. 12,000

Pictures of 500 different people. We made the pictures a little different to help the computer learn better.

4.3 Classification: Support Vector Machine

FaceNet takes each face picture. Puts it into a space with 128 dimensions. This space is special because the distance between two face pictures in this space tells us how similar the faces are. The network that does this is trained using a thing called the triplet loss function.

The triplet loss function is like this:

L = the sum of the difference between the distance of the anchor picture and the positive picture and the distance of the anchor picture and the negative picture plus some extra space.

Here the anchor picture is the picture the positive picture is another picture of the same person and the negative picture is a picture of a different person.

They use something called triplet mining to pick the most useful pictures to train the network. This helps the network learn faster.

They took a model that was trained and fine-tuned it using 12,000 pictures of faces from 500 different people.

Given the 128 dimension space they train a classifier called SVM on the pictures of people they want to recognize. They chose SVM because it is really good at dealing with dimension spaces and small amounts of training data.

The classifier makes a decision based on this function:

$f(x)$ = the sign of the sum of the weights times the training data plus some extra value.

The weights are calculated using something called the RBF kernel, which's like this:

$K(x_i, x)$ = the exponential of the negative distance between the two pictures.

They tried values, for the parameters C and γ and found that $C = 10$ and $\gamma = 0.001$ worked the best. They did this by trying all the combinations and seeing which one was the most accurate.

4.4 Enrollment Process

When we enroll an user we take a lot of pictures of their face around 20 to 30. We do this in lighting and we ask them to pose in different ways. We do the things to each

picture to get them ready and then we make them into special codes. These codes are stored in our database. We use them to make our classifier better, over time. It takes around 3 to 5 minutes to enroll each user. The new user enrollment process is important because it helps us get a lot of images of the new user.

5. EXPERIMENTAL RESULTS AND EVALUATION

5.1 Dataset and Experimental Setup

Experiments were done on a dataset. This dataset was collected from 500 people at XYZ Institute of Technology. The people were 300 students and 200 staff. The collection took three months.

The dataset had 15,000 images. These images were labeled. They showed lighting conditions. These were natural light, fluorescent light and LED light. The images also showed poses. The poses were on and turned 30 degrees to the side and tilted 15 degrees up or down. Some images had blockages. These were people wearing glasses or masks that covered the part of their face.

The computer setup used for the experiments had a processor. This was an Intel Core i7-11700 CPU. It also had a graphics card. This was an NVIDIA GTX 3060 GPU, with 12GB of memory. A 1080p USB 3.0 camera was also used.

5.3 Performance Under Challenging Conditions

To see how well the system works it was tested in three situations:

* Partial Occlusion like when people wear face masks: The system was able to recognize faces correctly 92.4 percent of the time. This is a good score considering that the nose and mouth are covered up.

* Low light like when it's really dark: The system was able to recognize faces correctly 95.1 percent of the time after it was helped by a special process that makes the pictures clearer.

* When people turn their heads to the side up to thirty degrees: The system was able to recognize faces 94.8 percent of the time which shows that the FaceNet system is really good at recognizing faces even when they are not looking straight at the camera.

These results show that the system works well even when things get tough like, in life. The system does not completely fail it just gets a little worse which is a thing.

The FaceNet system and the whole system are pretty robust.

5.4 Attendance Recording Accuracy

In a world test that lasted for 30 days in three classrooms with 150 students the system got the attendance right 99.2 percent of the time when we compared it to the records that we checked by hand. The system made mistakes 0.8 percent of the time. This happened when things, like furniture blocked the camera or when students came into the classroom from the side at an angle.

6. DISCUSSION

6.1 Advantages Over Existing Systems

The SFAS has a lot of things to offer.

It does a things that are really helpful:

- * it gets rid of people attending for others because of biometric uniqueness
- * it does not need any physical contact which is very important in places where people are trying to avoid getting sick
- * it can process things away so it can tell immediately if someone is late
- * it works really well with the systems that already exist for student and employee information using something called REST APIs.

The SFAS is very useful, for these reasons.

6.2 Limitations

The system has some results but there are some things that we need to talk about. First when people wear masks like N95 masks that cover their face the system does not work as well. It is about 78% accurate when people wear these masks.

The system also needs light to work properly. If it is really dark like than 10 lux it does not work well unless we add some extra lights.

We also need someone to watch when we first sign people up to make sure everything is done correctly.

We have to think about privacy because we are collecting information, about peoples bodies. We need

to make sure we follow the rules like the GDPR rules so we do not get in trouble.

6.3 Future Work

There are some ideas for what we can look into next.

1. We can try using cameras so that our system can recognize people even when it is dark or when they are wearing masks.
2. We can use a way of learning called learning this means we can update our models without having to send peoples face pictures to a central place, which is good for privacy.
3. We can also try using than one way to recognize people like using their face and their voice this can make our system more secure.
4. We want to make our system work well on devices, like the Raspberry Pi 5 and the NVIDIA Jetson Nano so we do not have to spend a lot of money on big computers.
5. We should also do some studies to see how well our system works over time because peoples faces can change and we want to make sure our system can still recognize them.

7. CONCLUSION

The AI-Powered Smart Facial Attendance System is a good system that uses special tools like MTCNN to find faces FaceNet to learn more about the faces and SVM to figure out who the faces belong to. The people who made this system tried it out. Found that it can correctly identify faces 98.7 percent of the time make mistakes only 0.6 percent of the time and take about 340 milliseconds to do all this. These numbers show that the AI-Powered Smart Facial Attendance System is good enough to use in life like in schools and offices.

The people who made this system also tried it out in three classrooms for 30 days. Found that it can record attendance correctly 99.2 percent of the time which is a lot better than the old way of doing things. The system also has a website that lets the people in charge see what is going on and get information, which helps them do their jobs better.

The AI-Powered Smart Facial Attendance System is a step towards making attendance recording automatic so

people do not have to touch anything and it is hard to cheat. If the people who made this system can make it better like making it work with faces that're partly hidden and making it work in more places then it could become a standard tool that all modern schools and offices use. The AI-Powered Smart Facial Attendance System is an important tool, for the AI-Powered Smart Facial Attendance System to be used everywhere.

ACKNOWLEDGMENTS

The authors thank XYZ Institute of Technology for letting them use their laboratory facilities. They also appreciate the students and staff who helped collect data. The Department gave them some funding through the Research Grant DRG/2024/XYZ/001.

REFERENCES

I have a list of references about face recognition and related technologies.

Here are some of the things I found:

* Jain, A. K., Ross, A. And Prabhakar, S. Wrote about face recognition in 2004.

They published a paper called An introduction to face recognition in the IEEE Transactions on Circuits and Systems for Video Technology.

The paper is about face recognition and other biometric recognition methods.

Face recognition is a part of recognition, which is a big field.

* Want R. Wrote about RFID technology in 2006.

RFID technology is used for things, including face recognition systems.

It is really useful for applications.

* Viola, P. And Jones M. J. Wrote about face detection in 2004.

They published a paper called real-time face detection in the International Journal of Computer Vision.

Face detection is a part of face recognition. It is very important.

Face recognition systems use face detection to find faces in images, which's a big part of face recognition.

* Lienhart, R. And Maydt J. Wrote about object detection in 2002.

Object detection is related to face detection. It is used in many face recognition systems.

Face detection is a type of object detection. It is used to find faces.

* Taigman, Y., Yang, M., Ranzato, M. And Wolf L. Wrote about DeepFace in 2014.

DeepFace is a face recognition system. It is very good at recognizing faces.

Face recognition systems like DeepFace use learning, which's a big part of face recognition.

* Schroff, F., Kalenichenko, D. And Philbin J. Wrote about FaceNet in 2015.

FaceNet is another face recognition system. It is very useful.

Face recognition systems like FaceNet use face recognition algorithms which're a big part of face recognition.

* Deng, J., Guo, J., Xue, N. And Zafeiriou S. Wrote about ArcFace in 2019.

ArcFace is a face recognition system. It is very good at recognizing faces.

Face recognition systems like ArcFace use face recognition algorithms, which are used in applications.

* Kumar, A., Bhargava, M. And Singh R. Wrote about automated classroom attendance in 2021.

Automated classroom attendance systems use face recognition, which's very useful.

Face recognition is used to mark attendance in classrooms, which's a big application of face recognition.

* Rai, P. And Tripathi S. Wrote about time facial attendance systems in 2022.

Time facial attendance systems use face recognition. They are very useful.

Face recognition is used to mark attendance in time which's a big part of many applications.

* Fernandez, R., Gomez, L. And Patel N. Wrote about mobile-optimized attendance in 2023.

Mobile-optimized facial attendance systems use face recognition. They are very useful.

Face recognition is used to mark attendance on devices, which's a big application of face recognition.

* Zhang K., Zhang, Z., Li, Z. And Qiao Y. Wrote about joint face detection and alignment in 2016.

Joint face detection and alignment is used in face recognition systems. It is very useful.

Face recognition systems use joint face detection and alignment to detect and align faces which's a big part of face recognition.

* Parkhi, O. M., Vedaldi, A. And Zisserman A. Wrote about face recognition in 2015.

Deep face recognition is a type of face recognition. It is very good at recognizing faces.

Face recognition systems use deep face recognition to recognize faces, which's a big application of face recognition.

* Cortes, C. And Vapnik V. Wrote about support-vector networks in 1995.

Support-vector networks are used in face recognition systems. They are very useful.

Face recognition systems use support-vector networks to recognize faces, which's a big part of face recognition.

Face recognition is a field with applications, including classrooms and offices.