VISITOR CHECKING SYSTEM

BONGU KARTHIK
Computer Science Engineering
ABV-IIITM
Gwalior, India

GUMIDELLI CHANDRAHAS
Computer Science Engineering
ABV-IIITM
Gwalior, India

MARAM VIGNESH
Computer Science Engineering
ABV-IIITM
Gwalior, India

Prof.SHASHIKALA TAPASWI
Computer Science Engineering
ABV-IIITM
Gwalior, India

***Abstract*—The visitor checking system is a critical component of security and access control in various environments, like educational institutions. Conventional approaches often rely on manual identification methods, such as photo identification cards or manual verification by security personnel, leading to potential inaccuracies and security breaches. In this project, we propose a novel visitor checking system leveraging Siamese networks to enhance visitor identification and authentication. Siamese networks are a class of deep learning architectures designed for measuring similarity between two inputs. In our system, we utilize Siamese networks to learn feature representations of visitors based on their facial images.The network we used has been trained on a large dataset, the system can then effectively compare incoming visitors' facial features with those in the database for verification and identification. The proposed visitor checking system consists of two main stages: enrollment and verification. During the enrollment stage, visitors provide their identification details and have their facial images captured by the system. The Siamese network then processes these images and generates a compact and discriminative feature representation, creating an embedded vector for each visitor**

***Index Terms*—SiameseNetwork,HaarCascade,FaceNet,Machine Learning.**

I. Introduction

In recent times, educational institutes have become increasingly conscious of the need to maintain a safe and secure environment for students, staff, and visitors alike. As they strive to strike a balance between maintaining an open and welcoming atmosphere while ensuring the safety of their community, modern technologies offer innovative solutions to address these challenges effectively. One such technology is the implementation of a Visitor Checking System, a comprehensive approach that combines security and efficiency to manage the entry and exit of visitors on campus.

Traditional methods, such as manual sign-ins and ID verification, can be cumbersome, time-consuming, and prone to errors. However, by harnessing the power of Siamese Neural Networks, educational institutes can significantly improve the visitor management process. Siamese Neural Networks are a specialized class of artificial neural networks designed to compare and recognize similarities or differences between two input data points. This unique architecture makes them particularly well-suited for facial recognition tasks, as they can identify and authenticate individuals with remarkable accuracy and speed.

A. Problem/Motivation

In this project, we explore the application of Siamese Neural Networks in educational institutes to create a robust and efficient Visitor Checking System. By leveraging the power of deep learning algorithms, the system will be capable of quickly and accurately verifying visitors' identities, reducing the risk of unauthorized access, and enhancing the overall security posture of the institute. Throughout this project, we will delve into the technical aspects of Siamese Neural Networks and their application in facial recognition. We will also explore the ethical considerations and privacy concerns associated with implementing such technology on an educational campus.

By presenting a well-rounded examination of the Visitor Checking System based on Siamese Neural Networks, this project aims to provide educational institutes with invaluable insights into how cutting-edge technology can be employed to create a safer and more efficient environment, fostering an optimal learning and working atmosphere for all stakeholders.

B. Objectives

The objective is to develop a Machine Learning model to register the visitors who want to enter the campus. This model helps the campus authorities to identify the visitors in case of any emergency situation occurs. The model is trained on different faces using a neural network architecture known as siamese networks. First is the Visitor Registration to create a system that efficiently and properly gathers visitor information, decreasing the need for manual data entry. The system should be able to process visitor information such as name and contact details and second the verification is to verify the identity of a visitor in case of any emergency and Reduction of manual processes by automating the task of registering each visitor entering the campus, the number of mistakes that can occur the whole process is reduced. The time taken to register each visitor is also greatly reduced.

II. Literature Review

A. Background of the Project

Maintaining a secure environment has risen to the top of the priority list for institutions in today's fast-paced world. A visitor checking system is one of the steps done to enhance security in institutions. A visitor checking system is a technological tool created to control and observe the flow of guests coming into and leaving a location.A visitor checking system's main goal is to increase security by precisely identifying people and making sure that only authorized visitors are allowed entry. Visitors must check in using the designated system when they arrive, which records their identity and the reason for their visit.

III. Related Works

A. Face Detection

Face detection is a computer vision task that involves locating and identifying human faces within images or video frames. The primary goal of face detection is to determine whether there are any faces present in the given input data and, if so, to identify their bounding boxes (rectangular regions that encompass the faces). Several machine learning models have been developed to perform face detection. Here are a few well-known ones:

Haar Cascades: Haar Cascades is a machine learning-based face detection technique proposed by Viola and Jones in 2001. It uses a series of simple image features called Haar-like features along with a machine learning algorithm (often AdaBoost) to classify regions of an image as either containing a face or not. Despite its age, Haar Cascades remains widely used due to its efficiency and effectiveness for real-time face detection applications.

You Only Look Once (YOLO): YOLO is another widely used object detection framework that can be applied to face detection tasks. It divides the image into a grid and predicts bounding boxes and class probabilities for each grid cell. YOLO is known for its real-time performance and can detect faces efficiently.

B. Face Verification

In recent years, several face recognition techniques have been developed, each with its unique architecture and approach. One notable algorithm is DeepFace, created by Facebook in 2014. It employs a CNN-based architecture with six convolutional layers and a loss function for face verification. Google introduced FaceNet in 2015, which uses a 22-layer CNN architecture combining Zeiler and Fergus, and GoogLeNet. FaceNet applies the triplet loss function for face recognition. Another significant development came from the Visual Geometry Group (VGG) at Oxford University in 2015, called VGGFace. This architecture is based on VGG and consists of 18 convolutional layers along with a triplet loss function for enhanced performance. Carnegie Mellon University researchers proposed another approach using GoogLeNet architecture, followed by the triplet loss function. This model is known as OpenFace, and it optimizes the parameters to enable training with smaller datasets. These algorithms were evaluated using the Labeled Faces in the Wild (LFW) dataset. DeepFace achieved an accuracy of 97.35% ,aceNet achieved 99.63% accuracy, VGGFace achieved 98.95% accuracy, and OpenFace achieved an accuracy of 92.92%.

IV. Methodology

Face detection is a computer vision task that involves locating and identifying human faces within images or video frames. The primary goal of face detection is to determine whether there are any faces present in the given input data and, if so, to identify their bounding boxes (rectangular regions that encompass the faces).Several machine learning models have been developed to perform face detection.

A. Siamese Networks

Siamese Neural Networks are a specialized class of artificial neural networks designed to compare and recognize similarities or differences between two input data points. The term "Siamese" comes from their architectural resemblance to Siamese twins, as they share the same parameters and architecture while processing two distinct inputs simultaneously. These networks have found significant applications in tasks such as facial recognition, signature verification, one-shot learning, and similarity-based clustering. The primary goal of Siamese Neural Networks is to learn a meaningful representation of the input data such that similar data points are mapped closer together in the learned feature space, while dissimilar data points are pushed further apart. This makes them highly effective in tasks where measuring similarity or dissimilarity between data samples is essential. The key components of a Siamese Neural Network include:

- **Shared Architecture** The Siamese network consists of two or more identical subnetworks, sharing the same set of parameters. These sub-networks process their respective input data independently but with the same learned weights and biases.
- **Feature Extraction** Each sub-network in the Siamese architecture extracts relevant features from its input data. In the context of facial recognition, for example,these sub-networks analyze facial images and transform them into a lower-dimensional representation capturing unique facial features.
- **Distance Metric** The learned feature representations from the shared sub-networks are then compared using a distance metric, typically Euclidean distance or cosine similarity. The distance metric quantifies the similarity between two input samples, with smaller distances indicating higher similarity.
- **Loss Function** The Siamese Neural Network is trained using a specific loss function that encourages similar samples to have small distances and dissimilar samples to have large distances. The contrastive loss function and triplet loss function are commonly used in this context.

The training process of Siamese Neural Networks usually involves providing pairs of samples, each labeled as either similar or dissimilar. During training, the network learns to adjust its parameters to minimize the distance between similar samples and maximize the distance between dissimilar ones, effectively learning to distinguish between them. Siamese Neural Networks have shown great promise in various applications, particularly in scenarios where labeled training data is limited, as they enable effective oneshot learning. Furthermore, their ability to learn meaningful feature representations has made them a popular choice for similarity-based tasks, contributing to enhanced performance and accuracy in numerous domains.

B. Haar Cascades

Haar cascades, also known as Haar classifiers or Haar features, are an object detection algorithm used for identifying objects or specific patterns within digital images. Haar cascades are particularly well-suited for detecting faces in images or video streams. They work by applying a series of simple rectangular Haar (fog)-like features to different sub-regions of an image. These features are calculated by subtracting the sum of pixel values in white rectangle regions from the sum in black rectangle regions.

A Haar cascade is a collection of trained Haar classifiers, organized in a specific structure. Each classifier is trained to identify a specific visual pattern or feature, such as an edge, a line, or a specific texture. The cascading aspect of the algorithm refers to the arrangement of classifiers in a cascade, where each subsequent classifier becomes more complex and specific than the previous one.

During detection, the Haar cascade scans an image or video frame using a sliding window technique. At each position, the cascade applies each classifier in a cascade sequentially, and if a classifier fails to detect the pattern, the detection process is stopped for that particular window. This approach allows for efficient computation, as it quickly rejects regions that are unlikely to contain the object of interest.

We used HaarCascade to identify faces in the images taken by the camera and crop the image to only have the face of the person. This makes it easy for the faceNet to easily convert face images into face embeddings without any errors which could be caused due to the background, thus increasing the accuracy of our project.

C. FaceNet

A facial recognition system called FaceNet was created by Google researchers using deep learning. It is intended to produce high-dimensional embeddings, sometimes referred to as face vectors or face embeddings, which represent the particular traits and features of a face. Then, these embeddings can be applied to tasks involving faces, such as face verification, grouping, and identification. Convolutional neural networks (CNNs) are an architecture that FaceNet uses to extract facial information from input photos. In order to learn a mapping from the input image space to a high-dimensional feature space where faces of the same person are close together and faces of different people are far away, the network is trained on a large dataset of face photos. FaceNet's main goal is to discover an embedding space in which the Euclidean distance between face embeddings can accurately gauge how similar two faces are. FaceNet strives to produce a face representation that is independent of elements like position, lighting, and expression while preserving the fundamental facial features by encoding faces into a high-dimensional space.

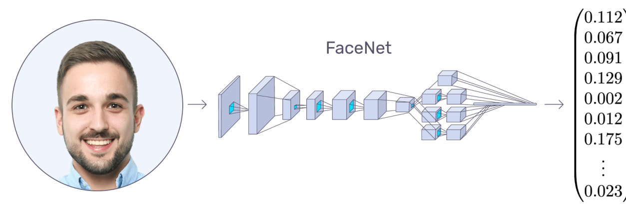

Fig. 1. FaceNet

V. Results

A. Working

The two main functions of the visaitor checking system is registratioin and identification of visitors. The visitor checking system utilizes Haar cascades and FaceNet for face identification and registration. When a visitor arrives, the system captures their image using a camera and applies the Haar cascade classifier to detect their face. The detected face is then cropped. Next, FaceNet generates a unique fixed-length vector (embedding) for the face image. This embedding, along with the visitor's information, such as name is saved in a database, completing the registration process. During identification, when a visitor approaches the system again, their image is captured and processed similarly. The FaceNet model generates an embedding for the detected face, which is then compared with the stored embeddings of registered visitors. If the similarity between the embeddings exceeds a certain threshold, the system identifies the visitor with a known identity. Based on this identification, the system can perform specific actions, such as granting access to certain areas or providing personalized information to the visitor.

B. Website

Backend of this model is created using FLASK framework in which the API end points like GET and POST are built to get the data and fetch the data and then display the output in the frontend. The frontend application is made using HTML, CSS, bootstrap where user can choose whether to verify, add, delete a visitor. The face captured in the website goes through our model to give the results required by the user.We have deployed the ML model through Flask. It helps us in running the model on localhost. The server can be further deployed to heroku so it can be accessed anytime instead of limiting it to local system. We can also view the visitors visiting the

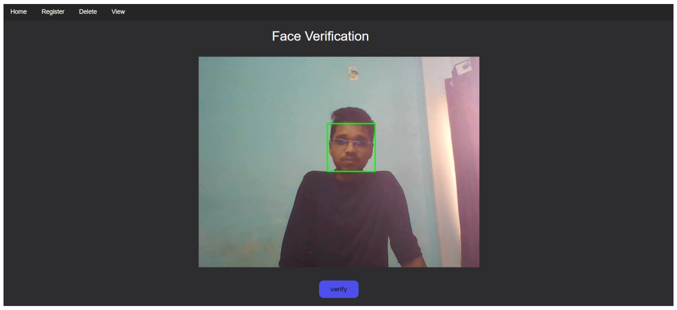

Fig. 2. Before Verification

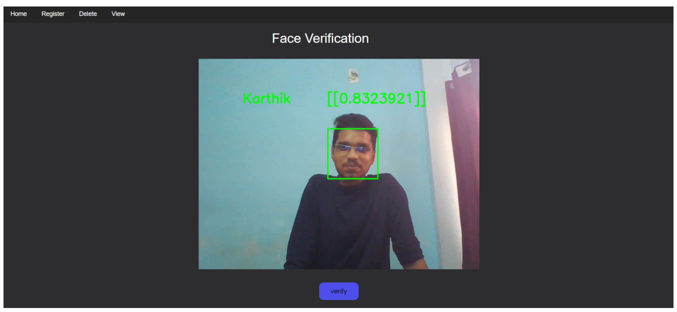

Fig. 3. After Verification

campus

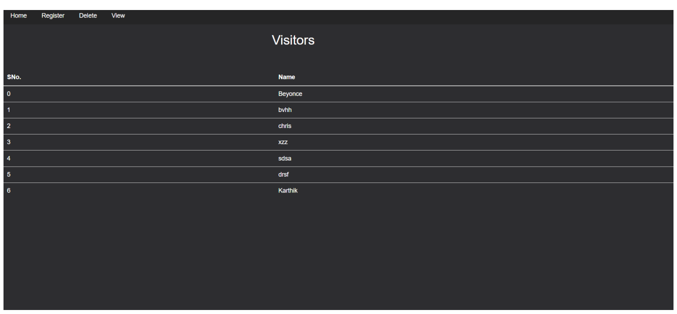

Fig. 4. Visitors List

VI. Conclusion

A. Future Scope

- **Integration with Access Control Systems** To automate the entry procedure, integrate the facial recognition system with access control systems. On the basis of known visits, this can entail unlocking doors or providing access to particular locations.
- **Database and Scalability** Especially when working with huge volumes of data, optimise the face recognition database for quicker and more effective searches. To ensure scalability and performance, take into account using advanced database administration approaches.
- **Multi-modal Biometrics** To improve the overall security and precision of visitor identification, think about combining different biometric modalities, such as facial recognition with fingerprint or iris recognition.

B. Limitations

- **Limited Pose and Viewpoint Tolerance**Haar Cascade and FaceNet might not handle extreme pose variations or non-frontal views well. If the visitor's face is significantly tilted or viewed from an unusual angle, the system's accuracy may decrease.
- **Dependency on Lighting Conditions** The performance of Haar Cascade and FaceNet can be sensitive to lighting conditions. Illumination changes, shadows, or variations in ambient lighting can affect the accuracy of the face detection and recognition processes.
- **Single Modality** The visitor checking system using Haar Cascade and FaceNet relies solely on facial features for identification. In cases where the face is obscured, damaged, or not visible due to accessories (e.g., masks) or medical conditions, the system may not be able to identify visitors accurately.

References

[1] Paper: "FaceNet: A Unified Embedding for Face Recognition and Clustering" Authors: Florian Schroff, Dmitry Kalenichenko, James Philbin Link: https://arxiv.org/abs/1503.03832 Applications Conference. IEEE, 2004.

[2] Paper: "Siamese Neural Networks for One-Shot Image Recognition" Authors: Koch, G., Zemel, R., Salakhutdinov, R. Link: https://www.cs.cmu.edu/ rsalakhu/papers/oneshot1.pdf

[3] Paper: "Rapid Object Detection using a Boosted Cascade of Simple Features" Authors: Viola, P., Jones, M. Link: https://www.cs.cmu.edu/ efros/courses/LBMV07/Papers/viola-cvpr-01.pdf

[4] Paper: VISITX: Face Recognition Visitor Management System Authors: Mr. Rikshit Makwana, Mr. Romil Nandwana, Mr. Jayshil Jain, Mr. Tejas Laxmeshwar, Mr. Shirish Sabnis Link:https://www.irjet.net/archives/V6/i4/IRJET-V6I4220.pdf

[5] Flask Documentation Link: https://flask.palletsprojects.com/en/2.3.x/