Lost Person Recognition using MTCNN & FaceNet

Shubham Golwal
Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology
Andheri, Mumbai

Rhutik Parab
Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology
Andheri, Mumbai

Dilip Patel
Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology
Andheri, Mumbai

Reeta Koshy
Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology
Andheri, Mumbai

ABSTRACT

The National Crime Records Bureau (NCRB) reports that between 2016 and 2021, an average of 3.4 lakh people in India were reported missing, indicating that the long-standing problem of missing persons has reached concerning proportions. This is equivalent to 930 people per day or 39 people per hour. At the same time, big cities like New Delhi, Chennai, and Mumbai have seen a sharp increase in their surveillance infrastructure; New Delhi leads the world in this regard, with an astounding 1,826.58 cameras per square mile.

To address this, we implemented a system that makes use of the latest developments in computer vision, particularly the FaceNet and MultiTask Cascaded Convolutional Neural Network (MTCNN) algorithms. Our mission is to use the vast potential of surveillance technology and minimize the increasing number of missing person cases. Our approach uses behavioral analysis and facial recognition to build a robust and effective system that can identify missing people from photos or videos, even when there are multiple people involved. We see our suggested solution as a revolutionary means of addressing the pressing social issue of missing people, guaranteeing a more prompt and precise response enabled by cutting-edge technology.

Keywords

FaceNet, Face Recognition, MTCNN, Missing Person Detection, Convolutional Neural Network, Facial Embeddings.

1. INTRODUCTION

The prevalent issue of missing persons remains a profound societal concern, with thousands of individuals reported missing globally each year. Addressing this challenge necessitates innovative solutions leveraging cutting-edge technologies to expedite the detection and recovery of lost individuals. In response to this exigency, our research introduces a holistic approach to lost person detection by harnessing the capabilities of advanced algorithms like FaceNet, which is grounded on the principles of the InceptionResNet v1 architecture, and employing Locality-Sensitive Hashing (LSH) to efficiently retrieve embeddings or vector data. Our system boasts a user-friendly interface enabling individuals to upload images of missing persons, facilitating a collaborative effort to locate and reunite loved ones. By transforming these images into numerical embeddings and maintaining a well-structured MongoDB database, the system streamlines the matching process, enhancing data management efficiency.

As per the latest statistics from the National Crime Information Center (NCIC), over 609,000 missing persons cases were reported in the United States alone in 2021. These staggering figures underscore the critical need for effective lost person detection systems. In our paper, we present a comprehensive analysis of our approach, detailing the datasets used, conducting an exhaustive literature survey on existing methodologies, thoroughly examining the FaceNet model grounded on InceptionResNet v1, elucidating the significance of Locality-Sensitive Hashing in retrieving embeddings, and showcasing the results derived from our experimental evaluations. Notably, our system provides real-time notifications upon positive matches, significantly reducing response times and bolstering the chances of reuniting families with their missing loved ones.

Our research delves into the technical intricacies of our system, elucidating the significance of the FaceNet model based on InceptionResNet v1 architecture. This model employs deep neural networks to convert facial features into high-dimensional embeddings, facilitating efficient data organization and retrieval. Moreover, Locality-Sensitive Hashing plays a pivotal role in swiftly retrieving embeddings or vector data, enabling rapid and accurate processing of images. Through rigorous evaluation, we demonstrate the practicability and efficiency of our approach in addressing the critical issue of lost person detection. We strongly believe that our research holds immense potential to substantially contribute to ongoing efforts to locate missing individuals and offer much-needed solace to their families.

2. Literature Review

The development of facial recognition technologies has attracted a lot of interest lately. Notably, the Attention Mesh lightweight neural network was introduced by Ivan Grishchenko et.al [1]. This neural network focuses specifically on important facial features like the lips and eyes when predicting 3D face meshes. Its unique real-time processing speed of 50 frames per second and its single-network design make training easier. However, the use of 2D error-based accuracy measurements and its reliance ona large dataset with precise landmarks indicate possible difficulties in real-world application.

In order to improve discriminative power, Jiankang Deng [2] et.al presents a novel method that adds an angular margin to the softmax loss and offers a geometric explanation. While the consistency it adds to the training process is admirable, evaluating its practical considerations is made more difficult by the lack of precise implementation specifics. ArcFace, however, is a notable attempt to advance deep face recognition techniques. In addition, the importance of digital image processing increating a face recognition system is emphasized in the publication "Face Detection and Recognition System using Digital Image Processing (2020)" by Gurlove Singh [3] et.al. Here, the main goal is to verify a person's identity using only their distinctive face traits. The use of digital image processing methods is emphasized as a way to improve dependability and accuracy. To illustrate the complexity involved in the face detection process, the report does address issues related to lighting, potential obstacles, and environmental factors. Age, skin tone, and facial expression variations highlight the complex issues that must be taken into account while developing reliable face recognition algorithms.

2.1 Comparative Analysis

Comparing the recent and notable works in the Facial Recognition field, Ivan Grishchenko et al. [1] introduced the Attention Mesh lightweight neural network, which is notable for its 50 frames per second real-time processing performance and unique network architecture that makes training it simple. On the other hand, its practical application in real-world circumstances is called into question due to its dependence on 2D error-based precision assessments and the requirement for a large dataset with accurate landmarks. Jiankang Deng et al. [2] provide ArcFace, a novel technique that improves discriminative power by adding an angular margin to the softmax loss and offering a geometric interpretation. Its consistent training method is admirable, but evaluating its practical implications is made more difficult by the absence of concrete implementation details. The importance of digital image processing in face identification is also emphasized by Gurlove Singh et al. [3], who seek to authenticate people using distinguishing features on their faces. The research sheds light on the intricacies involved in the face detection process by highlighting the difficulties caused by lighting conditions, potential obstructions, and environmental factors. Age, skin tone, and facial expression variations highlight even more the complex factors that must be taken into account in order to construct trustworthy face recognition algorithms.

3. Methodology

The Methodology for proposed work consists of following aspects

3.1 Common Datasets used for face detection and recognition

The Labeled Faces in Wild dataset and the Facescrub dataset were the two main datasets used in the study. The information below outlines how to access and preprocess these datasets.

3.1.1 Labeled Faces in the Wild Dataset

Labeled Faces in the Wild (LFW) [4] is a widely used dataset for face verification and recognition tasks. It was created by researchers at the University of Massachusetts, Amherst. It has more than 13,000 labeled images of faces collected from the wild, meaning the images are taken in uncontrolled, real-world conditions. There are over 5,000 individuals represented in the dataset. It is known for its challenging nature due to variations in lighting, pose, age, and facial expressions which makes it a good benchmark for evaluating the robustness of face recognition algorithms.

$$\text{Euclidean Distance} = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} \quad (1)$$

3.1.2 Facescrub Dataset

The Facescrub dataset [5] comprises a total of 106,863 face images of male and female 530 celebrities, with about 200 images per person.The images were sourced from the internet, capturing faces in authentic, uncontrolled environments. The dataset includes annotations for both names and genders associated with the depicted faces.

3.2 Multi-task Cascaded Convolutional Network Model

Our lost person detection system relies on the Multi-task Cascaded Convolutional Networks (MTCNN) algorithm for accurate face detection and preprocessing. MTCNN, a robust front-end component with three cascaded networks, identifies faces in images of various poses and orientations. First, the P-Net generates candidate face regions as a proposal network. The second stage, the R-Net, refines these proposals and eliminates false positives to improve face detection. The third and final stage, the O-Net, refines face landmarks and highlights key facial features.

The MTCNN algorithm's complex design helps it handle real-world images, making it essential to our lost person detection system. Its cascading structure captures and refines facial features at multiple levels, enabling face detection in complex images. Fig.. 1 shows the MTCNN architecture schematic to demonstrate its operation. Fig.. 1 shows how the P-Net generates candidate face regions, the R-Net refines them, and the O-Net refines facial landmarks for accurate and reliable face detection. This diagram illustrates the MTCNN model's multi-stage cascaded nature and role in our lost person detection system.

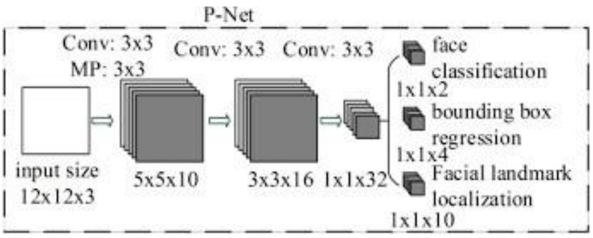


Fig 1. MTCNN Architecture Diagram

Fig. 1 shows the MTCNN model's multi-task cascaded structure, including P-Net, R-Net, and O-Net and shows how the algorithm generates candidate face regions, refines them, and detects faces accurately, a crucial step in our lost person detection system

3.3 FaceNet Model

Our lost person detection system relies on the FaceNet model to embed facial features in high-dimensional space. FaceNet uses deep neural networks to extract features and create a unique and consistent representation for each face, making facial data storage and retrieval efficient. FaceNet's ability to generate embeddings with reduced intra-class distance and increased inter-class distance makes it ideal for accurate face matching. For facial matching the FaceNet uses Euclidean distance (L2 distance) between two *n*-dimensional vectors *a* and *b* can be calculated using the Eq. 1:
In the Eq. 1:

- a_i and b_i represent the components of vectors *a* and *b* respectively.
- The summation (Σ) is performed over alldimensions (i = 1 to n).
- The square of the difference between corresponding components is calculated for each dimension.
- The square root of the sum of these squared

differences gives the Euclidean distance between the two vectors.

The FaceNet model's architecture is diagrammed in Fig. 2 to explain its inner workings. The deep neural network that embeds facial features is shown in this diagram. Facial images are processed by the model to extract critical features and create a high-dimensional vector. Lost person detection is efficient and accurate when these embeddings are stored in the database. This diagram illustrates FaceNet's complex but crucial role in our system.

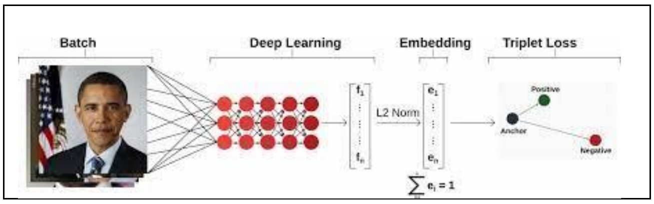


Fig 2. FaceNet Model Architecture Diagram

The FaceNet model's deep neural network structure converts facial features into high-dimensional embeddings, as shown in this diagram. The model's ability to create unique and consistent embeddings is essential for accurate face matching in our lost person detection system.

4. Objectives of Proposed Methodology

The primary objectives of proposed work are:

1. To develop a facial recognition system using FaceNet and MTCNN algorithms that facilitates the efficient andaccurate identification of missing persons.
2. Automation of the comparison process by implementing advanced computer vision techniques, ensuring a systematic and reliable approach to matching facial features.
3. Enhancing the robustness of the system to handle scenarios with limited facial embeddings, ensuring accurate results even in challenging conditions.
4. Extending the system's capability to process a variety ofinputs, including video clips and images, to provide a versatile solution for different data sources.
5. Implementing a user-friendly interface for easy management and maintenance of a comprehensive database of individuals of interest.
6. Exploring and integrating additional deep learning architectures to further improve facial embedding comparison capabilities.
7. Fine-tuning algorithms and improving feature extractionmechanisms to detect subtle differences and similarities in facial embeddings, thereby enhancing the overall accuracy of the identification process.

5. PROPOSED METHODOLOGY

Lost person detection requires cutting-edge technology to streamline the process. Our system combines the advanced FaceNet and MTCNN algorithms into a cohesive and effective lost person detection framework. The user-friendly system makes it easy for anyone to help reunite loved ones.

5.1 System Workflow

- Upload User Image: Users can upload missing person images to the system. This user-friendly feature aids group lost person detection.
- MongoDB Integration and Embedding Conversion: After uploading an image, our system uses FaceNet to create high-dimensional facial embeddings. The missing persons' data is organized in a MongoDB database using these embeddings.
- Embedding Average: The system takes a smart approach when multiple images of the same person are available. It averages individual embeddings to reduce redundancy and improve matching accuracy.
- Real-time Matching: The system searches the database for suspects and lost people using their images. It thoroughly searches all stored images for embeddings that match the provided image. This thorough search improves identification odds.
- Enhancement of Facial Features: MTCNN is used by the system to precisely detect and align facial features. By doing this, the accuracy of the facial embeddings is improved and the extracted facial features are guaranteed to be precisely aligned and normalized.
- Frontend Notification: If a match is found, the system immediately notifies the frontend. Real-time notifications speed up action and reunions.
- Constant Learning and Improvement: To update and enhance its facial recognition skills over time, the system has a feedback loop. It uses system performance data and user feedback to optimize the database and match algorithms for improved identification results over time.

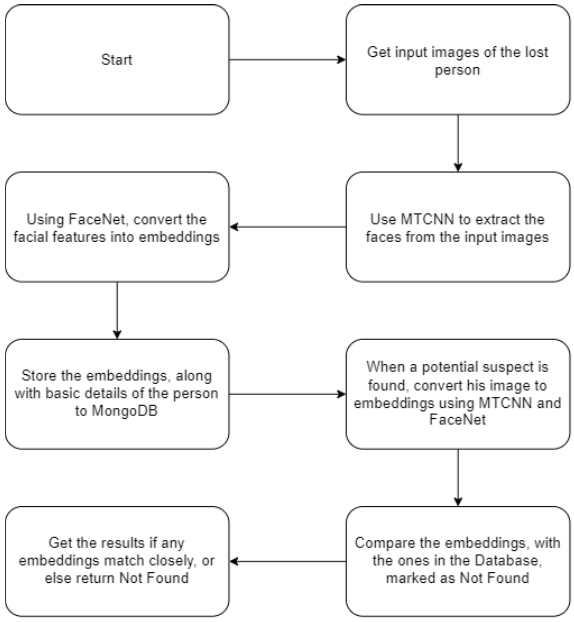


Fig 3. Data Flow Diagram

The workflow of the suggested system is shown in this Fig. 3, which shows the sequential steps involved in lost person detection. First, the lost person's input is acquired, and faces are extracted from the input image using the MTCNN model. The Facenet model is then used to turn the extracted facial features into embeddings. These embeddings are then kept forfuture use in the MongoDB database along with the individual's basic information. In the event that a possible suspect is found during the search, MTCNN and Facenet are used to turn their image into an embedding.

The newly created embeddings and theembeddings kept in the database are contrasted. When a close match is found, the individual is labeled as found and the relevant information is obtained from the database. On the other hand, the system returns a "not found" status if no match is found. This all-encompassing strategy guarantees a dependable and effective framework for missing person identification, improving search and rescue efforts.

5.2 Limitations and BoundaryConditions

1. Limited Embeddings in FaceNet: The FaceNet model utilized in the system generates 128-dimensional embeddings. While effective, this dimensionality limitation could impact the system's ability to capture intricate facial features, potentially leading to challenges in distinguishing individuals with very similar facial characteristics.
2. Challenges in Long-Distance Surveillance: The practicality of live camera feeds for long-distance surveillance poses a challenge. Beyond a certain distance, facial details become less discernible, impacting the accuracy of facial recognition. This limitation is exacerbated in scenarios where the subject's face is not adequately illuminated.
3. Quality of Camera Feed: The success of the facial recognition system is contingent on the quality of the camera feed. Low-quality or pixelated images may hinder the accurate detection of facial features, especially in real-world surveillance settings with varying environmental conditions.
4. Sensitivity to Lighting Conditions: The system's performance is susceptible to variations in lighting conditions. Suboptimal lighting may result in the degradation of facial visibility, affecting the precision of the recognition algorithm.
5. MTCNN Face Detection Probability: The MTCNN face detection algorithm relies on probability thresholds to identify faces in an image. While effective, this approach might lead to a bias towards more prominent facial features, potentially missing faces with lower probabilities, particularly in scenarios involving crowdedor complex scenes.
6. Similarity in Facial Features: In situations where individuals share similar facial characteristics, the system's ability to accurately distinguish between them may be compromised. This limitation is inherent in facial recognition systems and necessitates additional measuresfor robust identification.
7. Height and Perspective Considerations: The assumption that cameras are mounted at an average height of 8-7 meters introduces a boundary condition. Faces above this height might be captured at medium resolution, impacting the system's ability to recognize individuals based on facial details, especially in crowded or dynamicenvironments.
8. Visibility in Low Resolution: Faces not visible in low-resolution feeds pose a significant challenge. The system's effectiveness is contingent on the visibility of facial features, and low-resolution images may result in incomplete or insufficient information for accurate identification.
9. Considering these limitations and boundary conditions is essential for a comprehensive understanding of the system's capabilities and its applicability in diverse real-world scenarios

6. IMPLEMENTATION

During the process of building our facial recognition system, a comprehensive Flask application was developed. This application was designed to identify people who went through our system as well as people whose images were provided by the administrator in a seamless manner. The user interface gives the administrator the ability to decide in real time whether to allow passage or not, which opens the door to the possibility of automating the process.

6.1 Face Detection

Our implementation employs the Multi-task Cascaded Convolutional Networks (MTCNN) algorithm for accurate face detection. MTCNN is a robust front-end component with three cascaded networks, enabling precise identification of faces in images with multiple individuals.

6.2 Face Landmark Estimation

To enhance facial feature extraction, we utilize facial landmark estimation. This step involves identifying and marking key facial features, such as eyes, nose, and mouth, with high precision. This process is crucial for aligning and normalizing facial features.

6.3 Machine Learning

A deep Convolutional Neural Network (CNN) is implemented using FaceNet, a powerful deep learning model. This CNN is trained to recognize and extract 128 different facial measurements, creating a unique representation for each face.

6.4 Face Recognition

Our face recognition functionality is powered by a Support Vector Machine (SVM) classifier trained on the facial measurements obtained from the FaceNet model. This trained SVM classifier efficiently matches faces, allowing for accurate identification even in scenarios involving multiple people in images or video clips.

This integrated approach ensures a comprehensive andefficient system for face detection, landmark estimation, machine learning-based feature extraction, and accurate face recognition from images and video clips. The utilization of FaceNet and MTCNN contributes to the robustness and effectiveness of our proposed solution.

6.5 User Interface

1. Reporter: An admin can upload the images of the reported lost person, and add their details, to upload to the Database. All the lost persons in the Database are visible to the admin, along with their Found/Not Found status
2. Founder: The Founder can upload image or a video of the suspected person, to check if he exists in the Database, and if he/she matches someone, their name is displayed on the screen Application, while Fig. 6. and Fig.. 7 shows the Founder section. A Reporter, uploads the image/images of the person who is reported to be lost (Fig.. 4), and also adds some basic details of the lost person, like name, age, location, etc.

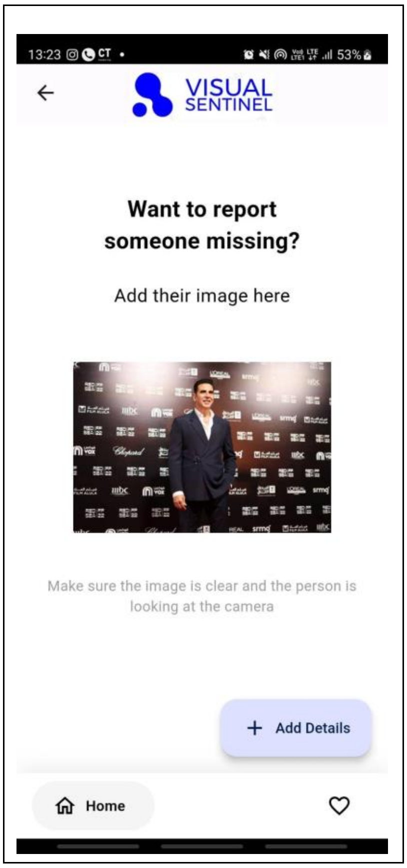


Fig 4. Lost Person Image selected

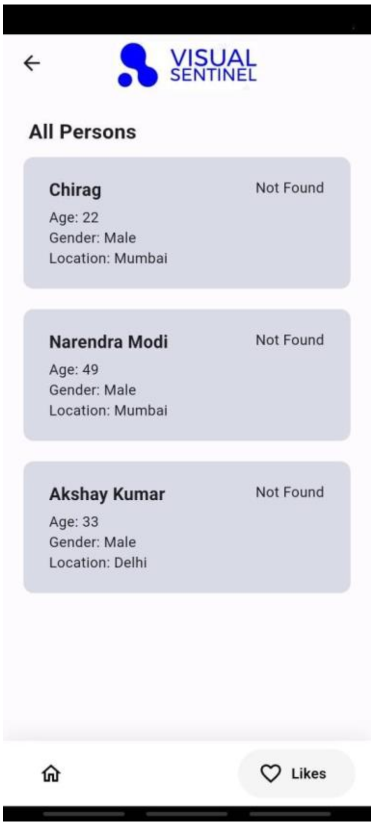


Fig 6. Suspected Person Image selected

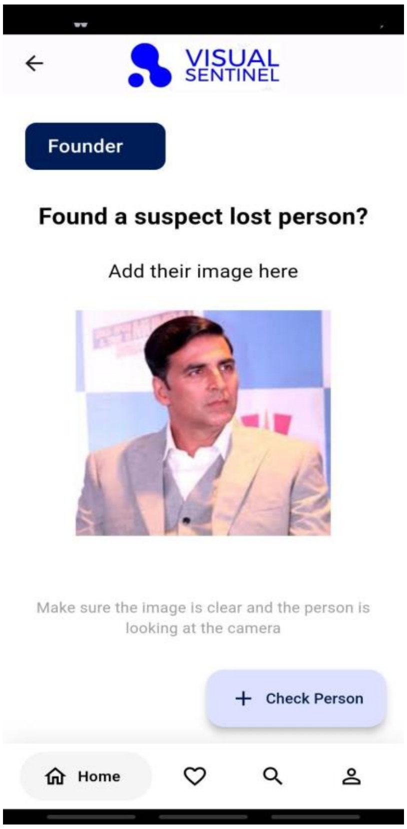


Fig 5. All Persons Details Screen

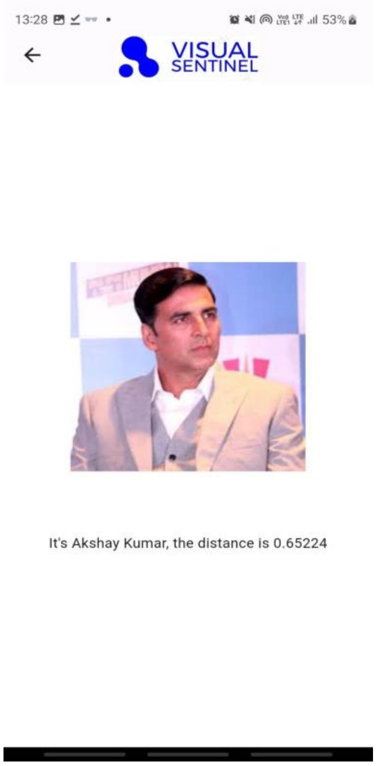


Fig 7. Result when a match is found

Fig.. 4 and Fig.. 5 shows the Reporter (Admin) section ofthe The images are then converted to embeddings and are stored to the Database along with the lost person details.The Reporter can view all the persons in the Database along with their Found/ Not Found status (Fig.. 5). Now, any normal user, which is a

Founder, can upload the image of a person who is suspected to be lost (Fig.. 6), and if it matches the embeddings with anyone from the Database, then their name is displayed to the user (Fig.. 7)

7. SYSTEM ARCHITECTURE

A Reporter (Admin) and several founders work together to create a comprehensive framework in our implemented application that addresses the crucial problem of missing persons location. The Reporter takes on the duty of augmenting the database with the images of individuals who are absent, as well as vital particulars like name, age, gender, and location. The system uses an advanced procedure to turn these images into embeddings, and then it stores the embeddings along with related information in a MongoDB database. Founders can also upload videos or pictures of people they believe are missing at the same time, which starts a database search. To find possible matches and help reunite missing people with their families, the application compares these uploaded files with the stored embeddings using a matching algorithm. This cooperative strategy makes use of the synergy between Reporters' data input and Founders' real-time queries, improving the system's efficiency and efficacy in finding missing people. For founders attempting to solve cases of missing persons, the integration of MongoDB guarantees a scalable and organized storage solution for embeddings and related data, enabling easy retrieval and comparison procedures.

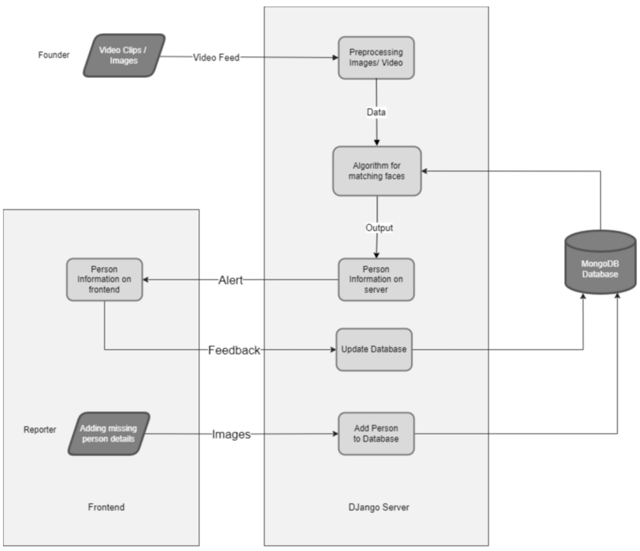


Fig 8. System Architecture diagram

Fig. 8: System Architecture Diagram provides a detailed representation of the proposed lost person detection system, showing how a strong framework can easily integrate different parts for effective operation. By uploading photos via the user interface, the user starts the process and sets off a series of actions. First, the system converts the uploaded images into embeddings using the FaceNet algorithm; this is an essential step for effective data representation and subsequent matching. The MongoDB database then stores these embeddings along with relevant information like name, age, gender, and location, offering an organized and scalable storage solution. An embedding averaging mechanism is incorporated to improve matching accuracy by enabling the system to produce a greater number of representative embeddings for comparison. The system's real-time matching features allow it to quickly match uploaded images with the database, making it easier to find possible matches quickly. Frontend notifications are the end product of this process, informing users of any matches discovered and offering amore efficient user experience. The overall efficacy of the lost person detection system is enhanced by this architectural design, which guarantees a smooth and effective flow of information encompassing user interaction, data processing, storage, and real-time feedback

8. EXPERIMENTS RESULTS

When we conducted an investigation into the efficacy of our strategy for reuniting missing individuals with their families, we obtained findings that are not only satisfying but alsoinstructive. We used the Facescrub dataset, which includes 2,400 photographs of humans in a range of postures and perspectives, to test the capabilities of our lost person recognition engine. The dataset [5] allowed us to put ourengine through its paces. These photographs were gathered from a wide variety of different sources, and each one depicts a real-life instance of a person going missing somewhere in the world. We have made a concerted effort to ensure that the dataset contains a diverse collection of facial expressions, lighting configurations, and orientations. This is done to ensure that our system is able to successfully handle realistic situations involving the detection of missing persons.

The photos' noise, resolution, and quality were improved, demonstrating the difficulties of processing user-uploaded images from a variety of equipment and situations. Noise, resolution, quality, postures, and angles were altered in the photos. We provided so many examples to show our system's versatility and dependability in processing a wide range of image inputs. FaceNet successfully integrated these photographs, demonstrating its versatility in assessing facefeatures.

A multi-stage cascaded MTCNN model was what made it possible to achieve the astounding accuracy of face detection. This model resolved issues relating to the orientation andposture of the dataset.

The quantitative measures showed that the system had an accuracy level of 99.43% when it came to identifying lost people from the dataset. This was shown by the fact that the system was able to do this. The Proposed System architecture is shown in Fig. 8.The fact that our approach proved successful in real-world scenarios despite the fact that theillumination and image quality changed served as evidence that it is applicable to a wide variety of contexts.

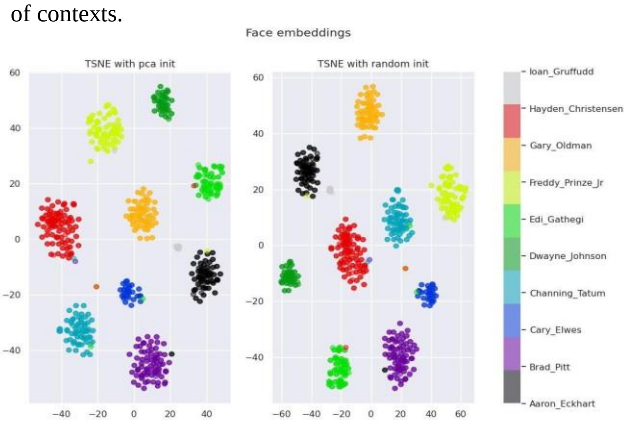


Fig 9. Plot of facial embeddings from testing dataset

.These findings not only represent a significant improvement for the field of facial recognition-based applications, but they also provide more evidence that our system for locating lost persons is dependable and efficient..

Fig. 9 shows the clustering of facial embeddings for a sample of test dataset. Referring to this visualization we can see that the facial embeddings are calculated accurately in most cases.

Fig. 10 and 11 show the heatmaps for the euclidean distance between the facial embeddings obtained from the training and testing datasets, respectively. We sampled 2,000 unaugmented testing image embeddings for visualization in both figures

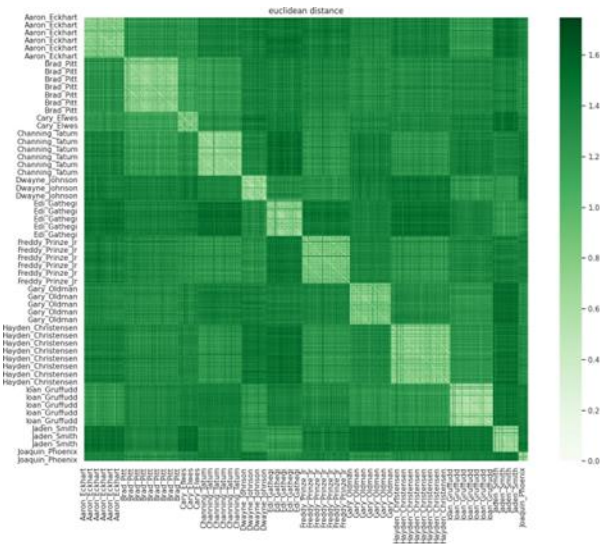
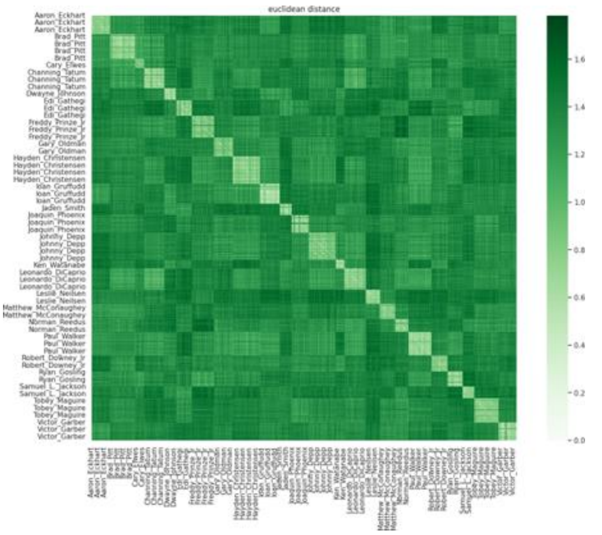


Fig 10. Heatmap for distance matrix of training results



. **Fig 11.** Heatmap for distance matrix of testing results

From Fig. 10 and 11 we can observe that the facial embeddings for the same people have a maximum euclidean distance of 0.8 units. This indicates that the model is able to learn robust and consistent representations of individuals, even across different images.

8. CONCLUSION

Our research illustrates the substantial influence of cutting edge facial recognition technologies on lost person detection systems using the MTCNN and FaceNet models. Through the efficient use of FaceNet for high-dimensional facial embeddings and MTCNN for facial feature detection, the system has demonstrated a significant increase in both the efficiency and accuracy of missing person identification. These models' smooth integration into the system's architecture has improved retrieval capabilities, which make it easier for families of the missing to quickly and reliably reunite with their loved ones. This method demonstrates the potential of advanced deep learning techniques to address practical societal issues, especially in the area of recovery and identification of missing individuals.

9. FUTURE WORK

Future updates to the system could include the ability to use facial recognition to identify suspicious activity in addition to lost person detection. The system can be strengthened to identify suspicious activity and possible security threats by integrating sophisticated surveillance techniques and intelligent monitoring algorithms. State-of-the-art deep learning architectures and methodologies could be explored and implemented to further improve the system's facial embedding comparison capabilities. The system can detect subtle differences and similarities between facial embeddings much better by fine-tuning the algorithms and improving the feature extraction mechanisms, which will result in a more accurate and dependable identification process. These developments could improve the system's overall performance and create a safer, more effective environment for security monitoring and missing person detection.

10. REFERENCES

[1] Ivan Grishchenko, Artsiom Ablavatski, Yury Kartynnik, Karthik Raveen-dran, and Matthias Grundmann. Attention mesh: High-fidelity face mesh prediction inreal-time, 2020.

[2] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. ArcFace: Additive angular margin loss for deep face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10):5962–5979, Oct 2022.

[3] Gurlove Singh and Amit Kumar Goel. Face detection and recognition system using Digital Image Processing. In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), pages 348–352, 2020.

[4] G. B. Huang, T. Berg, and E. Learned. LFW Face Database. https://vis-www.cs.umass.edu/lfw/, n.d. Computer Vision Lab.

[5] S. Winkler. resources. vintage - resources. https://vintage.winklerbros.net/facescrub.html, 2021, March 23. Retrieved November 14, 2023.

[6] Chunming Wu and Ying Zhang. 2021. MTCNN and FACENET Based Access Control System for Face Detection and Recognition. Autom. Control Comput. Sci.55, 1 (Jan 2021), 102–112. https://doi.org/10.3103/S0146411621010090

[7] Khan, Soumya & Sengupta, Diganta & Ghosh, Anupam & Chaudhuri, Atal. (2023). MTCNN++: A CNN-based face detection algorithm inspired by MTCNN. The Visual Computer. 1-19. 10.1007/s00371-023-02822-0.

[8] N. Singh and R. M. Brisilla, "Comparison Analysis of Different Face Detecting Techniques," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, 2021, pp. 1-6, doi: 10.1109/i-PACT52855.2021.9696583.

[9] Raju, Anand & Saravanan, Thirukkumaran & Arul, Jawahar. (2022). AI based face recognition system using FaceNet deep learning architecture. AIP Conference Proceedings. 2640. 020031. 10.1063/5.0118073.

[10] Li H, Hu J, Yu J, Yu N, Wu Q. UFaceNet: Research on Multi-Task Face Recognition Algorithm Based on CNN. Algorithms. 2021; 14(9):268. https://doi.org/10.3390/a14090268

[11] Qi, S., Zuo, X., Feng, W. and Naveen, I.G., 2022, December. Face Recognition Model Based On MTCNN And Facenet. In 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC) (pp. 1-5). IEEE.

[12] M. Alansari, O. A. Hay, S. Javed, A. Shoufan, Y. Zweiri, and N. Werghi. Ghostfacenets: Lightweight face recognition model from cheap operations. IEEE Access, 11:35429–35446, 2023

[13] Pham, T.N., Nguyen, N.P., Dinh, N.M.Q. and Le, T., 2022, November. Tracking Student Attendance in Virtual Classes Based on MTCNN and FaceNet. In Asian Conference on Intelligent Information and Database Systems (pp. 382-394). Cham: Springer Nature Switzerland.

[14] Xu, X., Du, M., Guo, H., Chang, J. and Zhao, X., 2021. Lightweight facenet based on mobilenet

[15] Gu, M., Liu, X. and Feng, J., 2022. Classroom face detection algorithm based on improved MTCNN. Signal, Image and Video Processing, 16(5), pp.1355-1362