



Missing Person Identification using Facial Recognition

Nithya Kalyani T^{1*}, Thejas Gowda², Vishnu Behera³, S Vineeth Sajjan⁴

¹ Assistant Professor, Department of Computer Science Engineering, Sri Sairam College of Engineering, Anekal, Bengaluru.

^{2,3,4} BE Students, Department of Computer Science and Engineering, Sri Sairam College of Engineering, Anekal, Bengaluru.

ABSTRACT: The great complexity and rising population in urban environs necessitate putting together advanced systems that can successfully detect and identify missing individuals. A paper which brings out an AI-based framework that primarily uses facial recognition and deep learning methodologies to address the issue of locating missing persons has been explored. Thus, the system so provided should rely on real-time CCTV recordings combined with a whole facial recognition system. This system will contain CNNs, YOLO for capturing faces, ResNet50 for improving feature extraction, and KNN for fast comparison in real time against a dedicated database of missing persons.

The system augments its scope by facilitating the secure submission of images and information pertaining to missing persons by community members, thus broadening the search capacity beyond conventional law enforcement assets. This engagement of the community enhances the prospects of prompt recovery and fosters improved cooperation between the public and governmental authorities. This framework offers a new approach toward enhancing public safety initiatives as well as shortening the response time to law enforcement through real-time processing, high accuracy, and scalability capabilities. The initiative tackles both technological and social dimensions of the search process, focusing on improved recovery rates, reduced search times, and therefore a safer community.

1. Introduction

Searching for missing persons is therefore one of the priorities that calls for prompt and applicable interventions, particularly in densely populated urban environments where traditional methods may not be effective. In the olden days, mainstream approaches to search for missing persons were very resource-intensive and time-consuming, such as sticking physical posters, broadcasting public notices through different media, and employing staff to manually patrol large areas in public places. Although they are crucial, these methods involve limitations in their scope of applicability, are labour-intensive, and usually introduce a delay in response that heavily impacts the potential success in data recovery. It thus calls for an approach that should be automated, scalable, monitorable, and detectable in real time with minimal usage of human resources. During the last few years,

enhancements in AI and facial recognition opened new avenues to bypass these hurdles. Using deep architectures such as CNN, this project attempts to automate the process of sifting through large spans of CCTV recordings in looking for potential matches to photographs in a missing people's registry with object detection schemes such as YOLO. The real-time object detection capability of YOLO, together with the strong feature extraction capability of ResNet50, allows for fast and accurate recognition, allowing the system to be effective in cluttered and complicated contexts. So as to get a high precision, KNN is implemented to match the faces detected within the live video feeds with entry in the missing persons data base while ensuring identification was valid. Alongside advancements in technology, this initiative also implements a community-focused strategy for the recovery of missing individuals. A secure digital platform enables members of the public to submit

photographs and details concerning missing persons, thereby broadening the search perimeter beyond areas under CCTV surveillance and promoting increased engagement from the community. By integrating automation powered by artificial intelligence with active community participation, the system significantly improves the scope and efficacy of missing person identification and simultaneously cultivates a sense of stronger cooperation between the general public and law enforcement entities. It is a building attempt for the speedy, precise and extensible missing person recovery framework, making a contribution to the improvement of public safety and paving the way for further active interaction in community welfare efforts.

2. Recent Works

Traditional methods used for finding missing people were the monitoring process, broadcasting the information through public media, and distributing leaflets. Though such methods existed, they were not so scalable or time efficient. This would often lead to a considerable delay in finding the missing persons. Automation brought in a key breakthrough; with techniques such as Viola-Jones and primitive feature-based models, this process was much more rapid. However, initial systems failed to handle real issues like occlusion, low resolution, and changes in lighting.

Most of these limitations have now been tackled by the recent developments in deep learning technologies. For instance, real-time object detection YOLO, as reviewed by Nadeem et al. (2022), has turned out to be very efficient in crowded urban scenarios for detecting faces in crowd surveillance videos. Similarly, the improved feature extraction property of ResNet50 proves efficient in facial recognition with challenging conditions like occlusion and angles as pointed out by Trupthi et al. (2024). Alam et al. (2024) used public input via safe interface achieving an accuracy level of 98% for missing persons identification. Despite these advancements, there are still differences. Most of the existing systems are too rigid to fit high-

density urban applications and do not balance accuracy with timeliness. More importantly, public data integration raises privacy issues because most systems have not specified protocols on how data could be managed with respect for privacy.

The proposed system integrates the most effective elements of modern technologies, addressing the identified weaknesses. The implementation of YOLO ensures reliable detection of crowds. Meanwhile, ResNet50 contributes to feature extraction under varying conditions, whereas KNN provides highly efficient fast matching against the overall database. To further enhance its effectiveness, the system incorporates safe community input. Additionally, it encompasses comprehensive protocols for very stringent privacy and security measures.

3. Proposed Work Explanation

3.1 CCTV Video Input:

The system collects video streams in real time from CCTV cameras. Such video feeds are continuous and mostly belong to cities with perpetual scenes like streets, shopping malls, or rail stations. The main challenge is to identify missing people from these constantly updating frames. The missing person can be captured in real time, directly from the footage captured. The video input can be from any conventional CCTV camera, and the system is expected to prove to be flexible with varying camera angles and settings.

3.2 Face Detection (YOLO):

YOLO (You Only Look Once) is an algorithm grounded in deep learning that functions to identify objects in real time. In this context, it specifically identifies faces in each individual frame of the video stream. YOLO is very efficient in crowded scenarios as it can detect and identify multiple faces in a single image in just one pass, making it ideal for public spaces like airports or train stations. YOLO operates by breaking down the input image into a grid, and making predictions about the bounding boxes along with the

probabilities of class in a single pass. Each cell predicts boxes and scores of confidences along with each box having a class probability attached to it.

YOLO utilizes one single CNN which, at the same time, makes predictions for multiple bounding boxes and class probabilities. The core part of YOLO includes convolutional layers that make it possible to extract higher-level features from the input image. To perform these tasks, it uses Darknet, a specially designed architecture. After features are extracted, YOLO further proceeds to use fully connected layers to predict the final output, such as the coordinate of bounding boxes, scores and class probabilities. The network's output is a tensor of dimensions $S \times S \times (B * 5 + C)$ where:

- $S \times S$ is the grid size.
- B is the number of bounding boxes each grid cell predicts.
- 5 corresponds to the bounding box parameters: 4 coordinates (centre, width, height) + 1 confidence score.
- C is the number of object classes (for facial recognition, $C = 1$) For each grid cell (i, j) , YOLO predicts the following parameters:

Bounding Box Coordinates: Each grid cell predicts B bounding boxes. Each bounding box is defined by 5 values:

- x_i, y_i : The centre of the box relative to the grid cell.
- w_i, h_i : Width and height of the box, normalized by the image width and height.
- C_i : Confidence score, representing the probability that the bounding box contains an object and how accurate the predicted box is.

These predictions are normalized to avoid large values when dealing with bounding boxes across different image sizes.

Class Probabilities: YOLO also predicts a vector of class probabilities $p(i, j)$ for each grid cell: $p(i, j) = P(\text{class}(j) | \text{object in cell})$ where j is the class index. In the case of facial detection, this would indicate the likelihood of the object being a face.

3.3 Feature Extraction (CNN):

After a face has been detected, the next step would come in to extract distinguishing facial features. These are done by Convolutional Neural Networks (CNNs), those types of deep learning models that have been particularly designed to work with images and to identify patterns. CNNs analyse the face to capture a very important number of specifics, ranging from eye distance to the nose form and the whole facial structure. These features are then converted into a feature vector, a numerical representation that can be used for comparison. The CNN is trained on datasets like VGGFace2, which includes various images of faces under different conditions (lighting, angle, etc.), ensuring the system can handle diverse real-world conditions. The feature vector is resistant to changes in illumination or indeed whenever the face is viewed from any angle. A typical CNN architecture for facial recognition consists of several key layers, each with a specific function in learning and extracting features from the input image. These layers apply a convolution operation on the input image or outputs of any previous layer. A convolution is an operation in which one uses filters or kernels in scanning the image to get patterns like edges, corners, and textures.

For the case of facial recognition, features like eyes, nose, mouth, and other relevant facial features will be identifiable. Mathematically, the convolution operation can be expressed as:

$$I' = (I * K) + b$$

Where:

- I is the input image.
- K is the kernel/filter.
- b is a bias term.

- * represents the convolution operation.

After the convolution, the Rectified Linear Unit (ReLU) activation function is typically applied to introduce non linearity into the model, allowing it to learn complex patterns in the data.

The ReLU function is mathematically defined as:

$$f(x) = \max(0, x)$$

Max pooling or average pooling is often used after convolutional layers to reduce the spatial dimensions of the feature maps while retaining the most important features. Mathematically, max pooling over a region of size 2x2 would look like:

$$\text{output} = \max(x_1, x_2, x_3, x_4)$$

Where x1, x2, x3, x4 are the values in the pooling window. The final classification step produces a feature vector, a compact numerical representation of the face by encoding key information about the face's features such as the distance between the eyes, nose shape, chin contours, etc.

3.4 Feature Refinement (ResNet50):

Then, the CNN-extracted raw features are passed on to the more complex deep learning architecture called ResNet50. These features are refined through a complex system that focuses on finer details, thereby improving accuracy and reducing errors. ResNet50 has been designed with specific considerations for the nuanced differences between faces and hence shows great performance even under partial or perspective views, like in profile. ResNet-50 is a 50-layer deep convolutional neural network that introduced the concept of residual learning. The residual learning framework allows ResNet-50 to train very deep networks effectively. Instead of learning a direct mapping $H(x)$, the model learns a residual mapping $F(x) = H(x) - x$.

Each residual block in ResNet-50 has the following form:

$$y = F(x, \{W_i\}) + x$$

Where:

- x: Input to the residual block.

- $F(x, \{W_i\})$: Residual function learned by the block using weights $\{W_i\}$.
- y: Output of the block (after adding the input x).

ResNet-50 is built using multiple residual blocks, organized into five stages. Each stage consists of a sequence of convolutional layers, batch normalization, ReLU activation functions, and skip connections. The 50 layers of its architecture is comprised of 48 convolutional layers, 1 max pooling layer, and 1 fully connected layers. ResNet-50 is pretrained on large-scale datasets like VGGFace2, which contain diverse images of faces with variations in lighting conditions, poses and angles, facial expressions, occlusions.

3.5 Matching (K-Nearest Neighbours-KNN):

After extracting and cleaning the facial features, the system conducts a comparative search of the current face against images in a database of missing people. In this work, the K Nearest Neighbours (KNN) matching algorithm is deployed. KNN is a non-parametric, instance-based learning algorithm that relies on distance metrics to find the closest points (neighbours) in a given dataset. In the context of facial recognition: KNN compares the feature vector of a detected face with all feature vectors stored in the Missing Persons Database.

It identifies the K closest matches based on a specified distance metric (e.g., Euclidean distance). Each detected face is represented as a feature vector x_i in R^n , where n is the dimensionality of the vector. The algorithm compares the query vector x_i which is of the detected face with all vectors y_i in the missing persons database. The similarity or dissimilarity between x_i and each database vector y_i is computed using a distance metric. The most commonly used metrics include:

Euclidean Distance:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

KNN compares the facial feature vector of the detected individual against the precomputed feature vectors in the missing persons database. Since KNN does not require a training phase, it is well-suited for real-time applications. New faces can be added to the database dynamically without retraining the model. KNN returns the K-closest matches, which can be used to shortlist potential identities for further verification.

3.6 Outcome (Notify Relevant Authorities):

This alerts and notifies the appropriate authority-law enforcement or security personnel-in case a

match is found in the face detected and any person in the missing persons database. The alert provides information regarding the person, including the individual's name, the last known location, and any other relevant details in connection with the lost person. In case no match is found, the system continues scanning through the video for other faces. Alerts are generally released in real-time so the concerned authorities can take quick steps for finding the whereabouts of the missing individual.

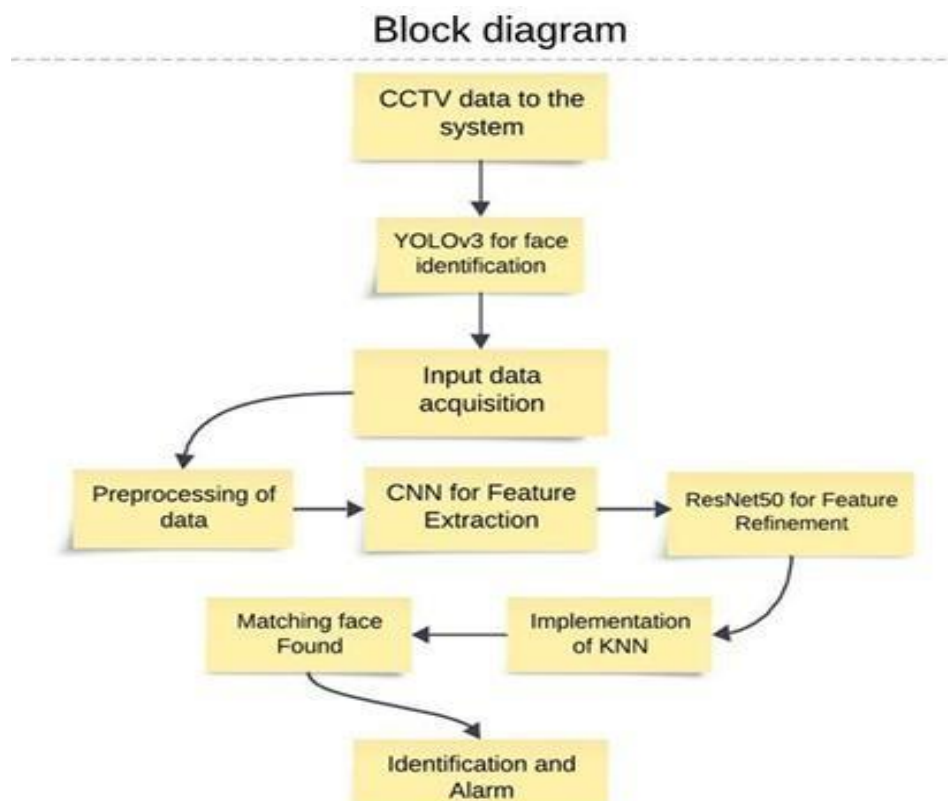


Figure 1: Block diagram

3.7 Datasets and Preprocessing: Training Dataset:

The basic training set implemented here is the VGGFace2 dataset, which contains a considerable number of facial images that expose a large degree of expressions, lighting, and angles. The CNN model trained in VGGFace2 yields generalization across various appearances and conditions. Missing Persons Collection: Contains photos of missing people together with associated metadata, including an age range and the last place the

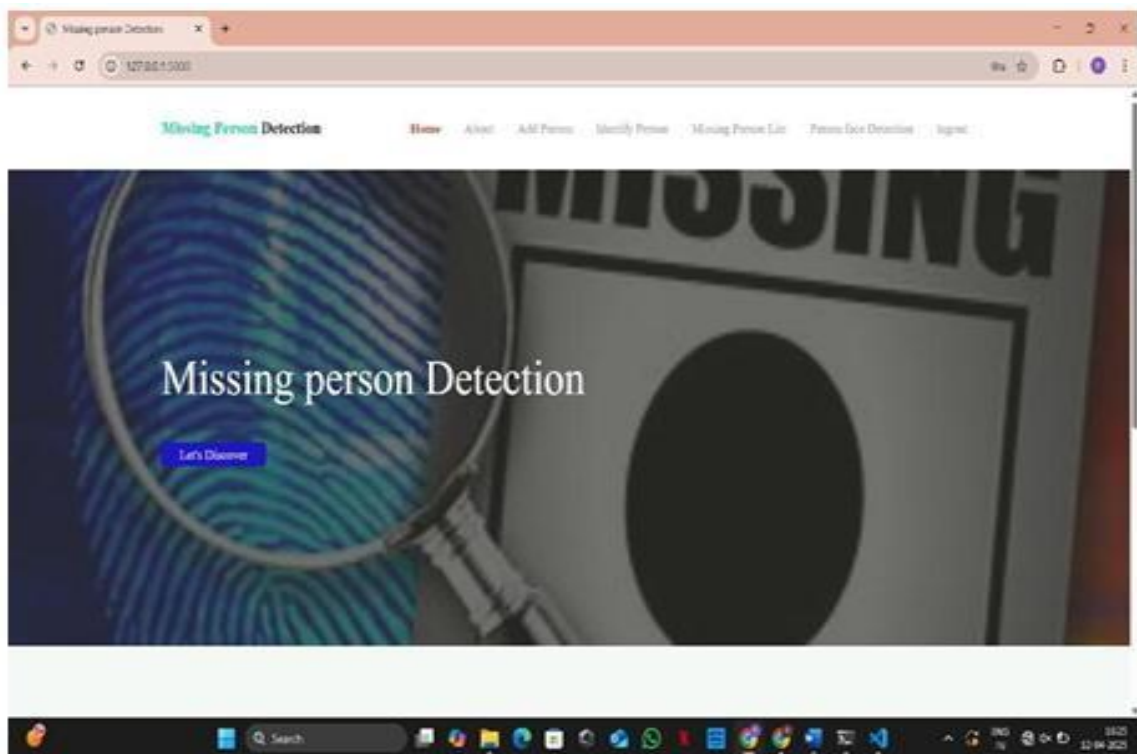
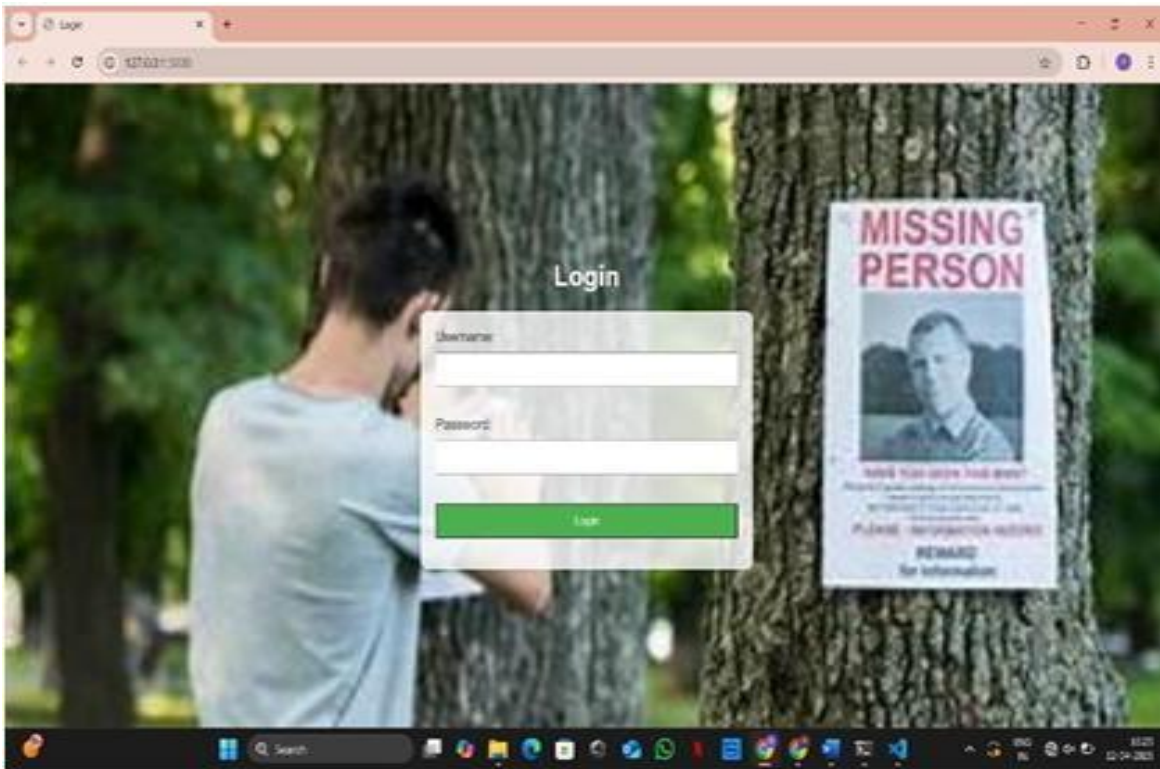
person was seen. Such information builds up the matching process by providing additional context and thus narrowing down the pool of potential matches.

4. Results and Discussion

We have successfully designed and implemented a website using flask framework and face recognition algorithm. Google Authentication techniques makes it easy for the users as they do not have to spend extra time logging in or any kind

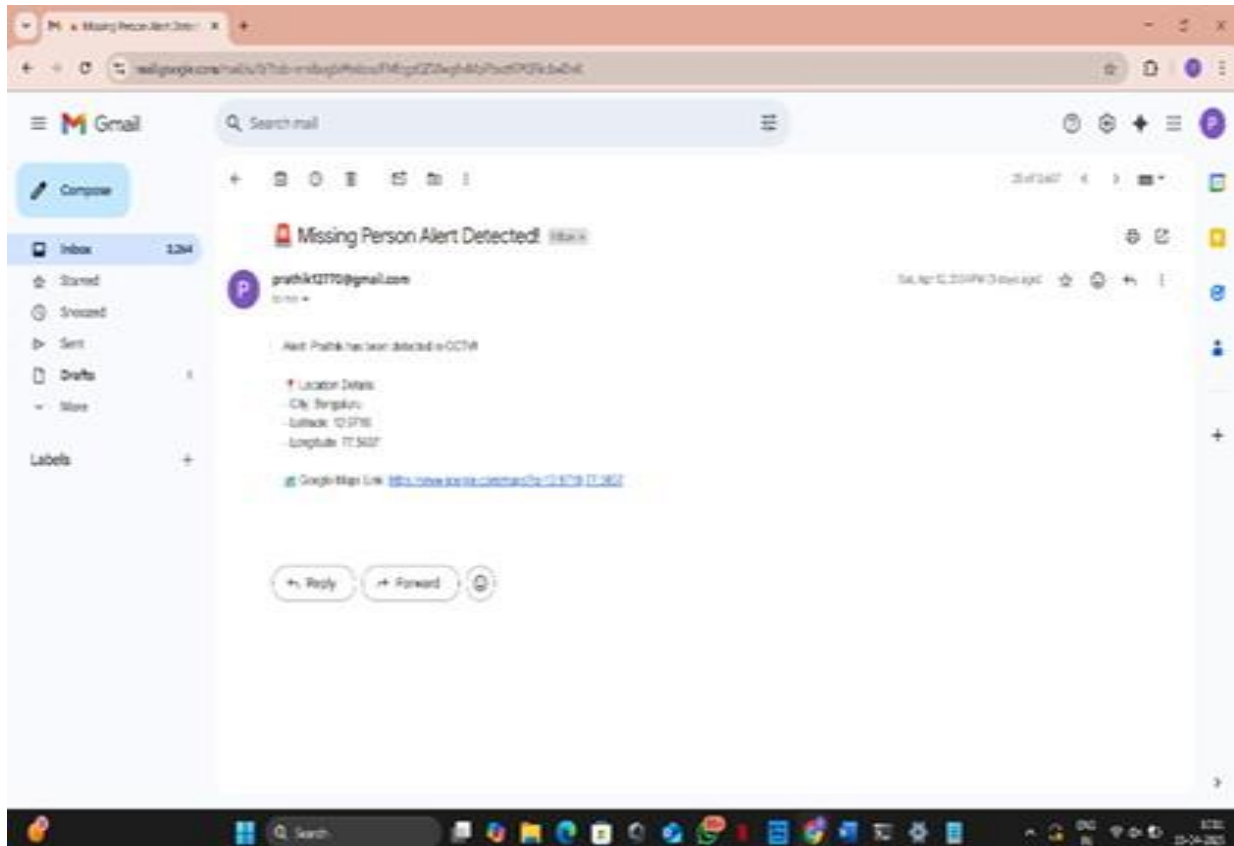
of registration fees. A high success rate has been observed in the process of determining people who

are missing from their locations using this model.



Missing Persons List

| # | Photo | Full Name | Date of Birth | Gender | Height | Weight | Details |
|---|-------|-----------|---------------|--------|--------|---------|------------------------------|
| 1 | | Portik | 2003-06-05 | Male | 178 cm | 56.0 kg | View Details |
| 2 | | Rishi | 2001-04-09 | Male | 180 cm | 56.0 kg | View Details |



5. Conclusions

This paper describes an artificial intelligence system for improving the process of missing person location by using advanced models such as YOLO for facial detection, CNN along with ResNet50 for extraction of features, and the KNN for matching. The integration of these technologically advanced components allows systems to provide real-time functional facial recognition and matching across databases, thus showing high effectiveness even in dynamic and expansive conditions. This will combine pre-existing CCTV infrastructures for economic efficiency, but adding a secure public upload portal promotes more cooperation between law enforcement and the community and increases the chances of locating missing persons considerably. Its ability to process live video streams in real time makes it especially beneficial in urban environments where prompt recognition can greatly improve the chances of recovery. Its ability to handle various situations such as low lighting, crowded scenes, and obstacles makes it resilient in many contexts. It is also scalable, thus suitable for

managing large datasets that may characterize regions with significant surveillance or population density.

Despite promising, it has challenges it faces, which include performance enhancement at low resolution and the reduction of false positives in dense environments. These are the challenges to be overcome and the computational power optimized for deployment at the edge to improve its applicability.

In addition, protecting data privacy in community engagement will be important to ensure that the public develops trust and becomes more involved. In summary, a practical application that bridges technology and what is seen in reality means an innovative approach for determining missing people. In the design and scope of the present embodiment, this system would allow it to become one critical public safety tool and certainly increase recoveries at hand while providing safer communities, too.

References

1. PD N. Harsha Sai; VV Sai Kiran; K. Rohith; D. RajeswaraRao, Year: 2022, "Identification of missing person using convolutional neural networks", In 2022 international conference on sustainable computing and data communication systems (ICSCDS), pp. 485 – 489.
2. Saravanan Alagarsamy; KV Sudheer Kumar; Polineni Vamsi; Devisetty Bhargava; B. Durga Hemanth, Year: 2022, "Identifying the Missing People using Deep Learning Method", In 2022 7th International Conference on Communication and Electronics Systems (ICCES), pp. 1104 – 1109.
3. Vishakha Shelke; Gunjan Mehta; Prajakta Gomase; Trisha Bangera, Year: 2021, "Searchious: Locating missing people using an optimised face recognition algorithm", In 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), pp. 1550 – 1555.
4. S. Ayyappan; S. Matilda, Year: 2020, "Criminals and missing children identification using face recognition and web scrapping", In 2020 International conference on system, computation, automation and networking (ICSCAN), pp. 1 – 5.
5. T. Smith, Year: 2022, A comparative analysis of best practices in a facial recognition policy for law enforcement agencies: A capstone in education with a concentration in educational leadership (Doctoral dissertation, DePaul University). DePaul University College of Education.
6. Abid Faisal Ayon; S. M. Alam, year: 2024, "Toward Digitalization: A Secure Approach to Find a Missing Person Using Facial Recognition Technology".