Computer Science and Information Technologies
Vol. 7, No. 2, July 2026, pp. 203~213
ISSN: 2722-3221, DOI: 10.11591/csit.v7i2.p203-213

Tracking a person and determining the location by using convolutional neural network technology

Zinah Shiker Makki, Ahmet Zengin
Department of Computer Engineering, Sakarya University, Sakarya, Türkiye

Article Info

Article history:

Received Nov 14, 2025
Revised Apr 12, 2026
Accepted May 16, 2026

Keywords:

Convolutional neural network
Deep learning for surveillance
Facial recognition systems
Identity detection
Real-time person tracking

ABSTRACT

Tracking individuals in real-world environments requires robust, non-intrusive methods that overcome the limitations of device-based systems. This study proposes a convolutional neural network (CNN)-driven person-tracking framework that identifies targeted individuals directly from camera feeds, eliminating the need for wearable or global positioning system (GPS) devices and addressing a major drawback of traditional tracking technologies. The system utilizes a TensorFlow-trained CNN model that can detect, recognize, and locate persons of interest in real-time, even under varying illumination conditions. Unlike conventional approaches, our method integrates facial feature extraction with encrypted identity management, enabling secure multi-person detection and rapid location reporting. Experimental results demonstrate a 92% accuracy in low-light settings and 100% accuracy under normal lighting, confirming the system's effectiveness for security-oriented applications. The findings highlight the novelty of combining lightweight CNN architecture, real-time facial recognition, and hash-based identity protection within a unified tracking pipeline.

Corresponding Author:

Ahmet Zengin
Department of Computer Engineering, Sakarya University
Sakarya, Turkey
Email: azengin@sakarya.edu.tr

1. INTRODUCTION

It is simple to use this technology to monitor pedestrians and instantly identify visitors, particularly in cases of assault or theft. Over the previous years, facial recognition in vital measurements and events, pattern recognition, and security monitoring devices have played a crucial role. Today, it is complicated and challenging to identify the person and property from thieves and recent fraud in some countries. The daily news includes crimes such as credit card fraud, computer penetration and data theft, as well as increasing security threats faced by companies and governments. The primary cause of these security breaches is that authentication is based on a basis, not our biometric identity [1].

Conducted research to help reduce the spread of the virus, and many advertising campaigns have identified the virus to people worldwide, where the World Health Organization (WHO) publishes disease prevention guidelines, one of the necessary measures to prevent the spread of disease is the use of face masks. This research proposes utilizing the face recognition feature to prevent responsible individuals from violating the general guidelines for preventing the spread of the coronavirus. The study focused on face masks. We used an algorithm for facial analysis by three essential points (eye, nose, and mouth). Person tracking technologies are part of biometrics, and other examples of this feature are fingerprint scanning technologies and iris scanning systems, which are emerging from artificial intelligence [2]. People's tracking systems use different methods depending on the application and the manufacturer. However, it usually consists of several stages that record, process, analyze, and match the face to a database of images. To identify and match facial image traits, these systems use computer vision insights, statistical analysis tools, and machine learning [3], [4].

Despite the wide adoption of facial recognition technologies, most existing tracking systems still suffer from critical gaps: i) they depend on GPS or wearable devices, which cannot be placed on individuals who are unwilling, unknown, or attempting to evade monitoring; ii) many classical image-processing approaches fail under poor lighting, occlusion, and real-world motion; and iii) systems based on handcrafted features—such as principal component analysis (PCA), local binary patterns (LBP), or basic support vector machine (SVM) pipelines—lack robustness and real-time adaptability needed for modern surveillance environments [5]. These limitations create a significant need for a tracking method that can automatically extract discriminative features from raw images, adapt to environmental variations, and perform accurate recognition without requiring any physical device on the target.

The need for Person tracking based on face recognition technology has increased in recent years. Many institutions have incorporated it into their applications and manufactured devices because it saves humans a significant amount of effort and time in matching faces. The Viola-Jones algorithm is used to detect the front. The modified PCA algorithm is used to recognize faces from images with some differences, achieving a close-to-real-time recognition rate. The algorithm analyzes the main components of the face. This technique recognizes facial pictures before and after plastic surgery [6].

To overcome the limitations of these earlier approaches, this research adopts convolutional neural networks (CNNs), which have demonstrated superior performance in extracting hierarchical and invariant facial features directly from raw image data [7]. CNNs are significantly more resilient to lighting variation, pose differences, occlusion, and image noise—conditions that frequently degrade the accuracy of handcrafted feature methods. Furthermore, CNN-based systems enable real-time processing, scalability, and improved generalization, making them ideal for surveillance scenarios where speed and reliability are essential [8], [9].

The primary objective of this research is to track individuals using facial recognition techniques in a specific area, detect violators, and send their locations. CNN algorithms were utilized, with some modifications, to suit and verify the project’s primary objective. Some properties have been added to the project, for example, the ability to discover more than one person simultaneously. Recent developments in technology and industrial revolutions worldwide have become essential to our lives, enabling us to utilize technology more effectively to benefit from it, particularly in computer vision and artificial intelligence, including machine learning and deep learning [10]. The features of tracking person systems and facial recognition in our time embody the essential characteristics that must be relied upon in addressing various issues, including health, security, and education. This research proposes a system to monitor individuals in violation, gather information about them, and transmit their locations to the relevant authorities responsible for managing the system.

2. MATERIAL AND METHOD

2.1. Proposed system architecture

Person tracking systems are among the significant systems for identifying or confirming a person naturally from a computerized image. There are two strategies for countering recognition. One is photo-based, and the other is video-based. There are now more classifications for it. In our case, a person tracking system based on facial recognition technology is a still image of a person. It can verify and identify one or more individuals by referencing a stored database of individuals to be tracked. The framework for facial recognition consists of several initial steps: processing the information, extracting key features, and classification. Information preprocessing encompasses various activities, including standoff location [11], noise reduction, image resizing, and scaling.

The structure of a person tracking system based on facial recognition technology consists of three components: i) training a CNN model by placing a set of images of the people to be tracked; ii) image processing by reducing some of the noise in the images; and iii) comparing the image taken by the camera and matching it with the database within the system. The functions of each part is described in Figure 1.

2.2. CNN model training

One of the critical steps in our work is training a model using advanced machine learning techniques, such as TensorFlow. It is an open-source library provided by Google to train models. The trained model can recognize faces by matching images stored in the base with video images captured by surveillance cameras and identifying the desired person.

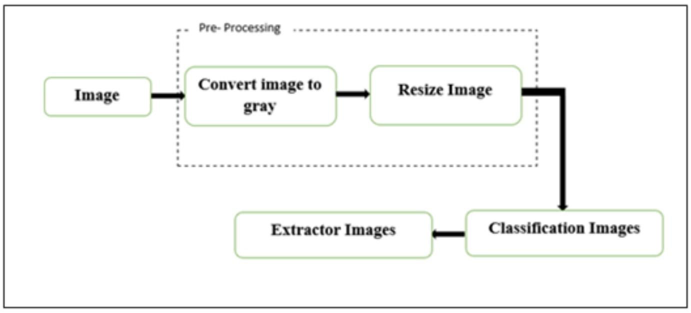

Figure 1. Person tracking system architecture

2.3. Face extractor from photos or videos

The facial features extractor extracts critical facial features. We are developing a feature extraction method that works on color images without requiring the use of glasses. The facial features extractor consists of three sequential modules: facial features localization, feature point extraction, and facial features generation. First, select the main points-eyes, nose, mouth-from the image of the face. Second, the face extraction unit extracts the characteristic point and analyzes its compatibility with the existing base.

2.4. Classification of a person by face

Face matching includes a categorized and didactic algorithm. This component is responsible for identifying a person. By providing a new picture, the person tracking system first detects the face and then positions it accordingly [12]. Then, the detected face image is passed to the facial features extractor, where the features are extracted using algorithms similar to those employed in the SVM algorithm to enhance the image resolution. The image will be matched, and it will be determined whether it is in the database.

2.5. Research design

When researchers work on a topic, they have goals to achieve at the end of the work. In this proposed system, the primary goal is to track specific individuals through the camera, identify them, transmit their location to the designated person within the system, and verify their identities. At the outset of our work, we conducted a literature review of all previous research in the field of tracking systems to determine the strengths and weaknesses of the algorithms used and to identify the technology that best meets the system’s primary purpose. Then, the proposed method is designed according to the data available in the work environment. Figure 2 illustrates the methodology employed in this work.

Figure 2. Methodology of work

2.6. Proposed system

We proposed a system to help governments or private security companies track specific individuals in public or private places. One of the primary functions of the system is to feed it images of people to be tracked. The system also features an encryption mechanism for data protection, utilizing a hash algorithm to encrypt the information entered into the system and ensure its confidentiality. The system consists of three main parts: input, processing, and output, as shown in Figure 3.

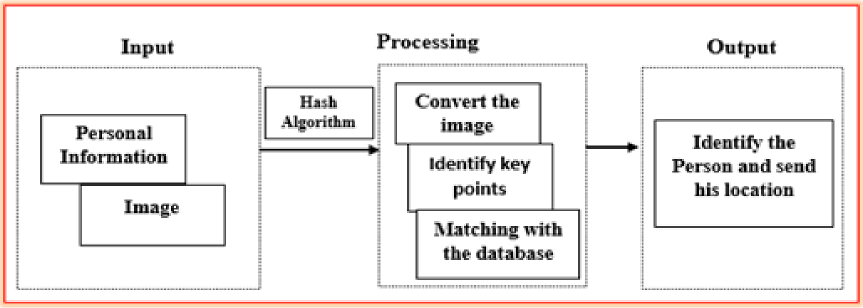

Figure 3. Proposed system steps

2.6.1. The main components of the proposed system

People tracking is a complicated system. According to our capabilities, we designed the proposed method using a computer camera to read faces and start the image analysis process. System components are divided into two main parts: the physical and software parts. Figure 4 illustrates the primary components of the system. We illustrated the main scheme of the system's components and divisions. We will explain each section with the benefit of using it in the system.

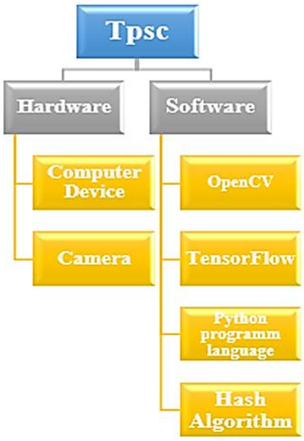

Figure 4. Components of the system

a. Hardware

We refer to the physical components used in the system to achieve the desired result, which consists of two parts.

– Computer device

The computer is one of the most indispensable units of the system, through which we connect the camera placed in a specific location to the programmed application, which detects violators and identifies them. Additionally, the computer can be described as the link between the system's hardware and software components, through which all hardware and software components are integrated, including applications, cameras, and databases.

– Camera

Most of the security cameras available on the market today feature high resolution, as the camera is linked to the proposed system, which allows us to create a database that includes the individuals to be tracked. When the camera observes a face, it determines whether it is inside a person in the database. In our proposed system, at the primary stage of work, we evaluated a camera on a personal laptop.

b. Software

The code and libraries represent the software components and algorithms combined to form an integrated application with the camera to track the person's system. The programmed application, written in the Python language, is a combination of several libraries and algorithms that we will discuss in detail.

- OpenCV

The open computer vision (OpenCV) library is a collection of programming interfaces designed to facilitate the development of computer vision applications. Intel corporation has a free software library under the open-source license. We can benefit from the OpenCV library, particularly in medical, industrial, and artificial intelligence applications [13]. OpenCV comprises several algorithms and software codes, all of which are compiled into a single framework. The objective of this framework is to provide solutions in computer vision. It can be used on all computer systems and focuses on real-time image processing.
Library contents (OpenCV):

- Core: a built-in module for characterizing fundamental data structures
- ImgProc: image processing unit for Image Enhancement.
- Video: the video analysis unit that incorporates target movement estimates.
- Calib3d: the 2D image information analysis unit.
- Features2d: responsible for detecting the target's known properties and separating it from the rest of the surrounding environment, such as geometric shapes.
- Objects detect face detection
- High-up: an easy-to-use interface for taking pictures and videos.
- CPU module: Utilizes the total computing power of the system by using the power of the video processing card.

OpenCV Library Applications:

- Facial recognition systems
- Gesture recognition systems
- Computer-human interaction
- Mobile Android
- Get to know the target
- Motion chasing
- TensorFlow:

After completing the data entry, reprocessing the images, and preparing them for the image processing stage in our proposed model, three image classifications will be performed based on the document where the model is trained to identify these cases and provide accurate results with the help of CNN techniques. The following measures were taken to train the CNN model.
Step 1: Upload dataset
Step 2: The input layer
Step 3: Convolutional layer

Filter the feature map using the specified number of filters. We must employ a relay activation function after convolution to add non-linearity to the network. The first convolutional layer contains 18 filters, each with a 7×7 kernel and equal padding. The output and input tensors have the same width and height with the same padding. To verify that the rows and columns are the exact sizes, the topology potential and spectral clustering (TPSC) model will add zeros to them.
Step 4: Pooling layer:

The maximum facility will be downsampled following the convention. The goal is to limit the feature map's mobility to prevent overfitting and improve computational speed. The traditional max pooling technique divides feature maps into subfields and only keeps the maximum values. Following the convolutional stage, the pooling computation is the next stage. The data will be compressed due to the pooling computation. With a dimension of 3×3 and a stride of 2, we can utilize the max pooling 2D module. As input, we use the output of the previous layer. [batch size, 14, and 15] are output sizes.
Step 5: Convolutional layer and pooling layer:

The output size of the second CNN is [batch size, 14, 14, 32], and it has precisely 32 filters. The output shape is [batch size, 14, 14, 18], and the size of the pooling layer remains unchanged.
Step 6: Dense layer:

The fully connected layer must be defined. It must be compressed before combining it with the feature map and the dense layer. The dense layer will connect 1764 neurons. We can add a ReLU activation function, and we will be able to do so. We add a 0.3 dropout regularization term, meaning 30% of the weights will be 0. Only during the training phase do people drop out.
Step 7: Logit layer

We define the final layer of the model's forecast. The output shape is batch size 12 (the total number of photos in the layer), the same as the total number of images in the layer.

2.6.2. Used algorithms

The camera-based person-tracking system is programmed through complex technologies, each performing a different function. However, they integrate, so the job is not done right if the entire technology is not working correctly. Face identification technology is more accurate and complex than its predecessors. Likewise, for recognition technology, a person's photo is slowly matched against the images in the database. However, he does not focus solely on the shape and features of the face; any change in facial expressions between the person's image and the image in the database invalidates his work. Some facial recognition technologies identify facial features by extracting features from a picture of the face; for example, one technology analyzes the relative position and size of facial organs, such as the nose, determined by the size and location of the face as well as its distance from the eyes, and it defines the gap between the eyes. In determining specific techniques to work, the researcher must do serious work comparing the methods and selecting the appropriate technology for the work. In this proposed system, we define elementary working techniques (CNN) to identify faces by training a TensorFlow model and utilizing a hash algorithm to encrypt data within the system, thereby enhancing the system's security.

3. RELATED WORK

Real-time video surveillance systems are used in various settings, including public spaces, commercial buildings, and public infrastructure. People's detection is critical to many video surveillance systems, including people identification, segmentation, and tracking. To recognize and track people, researchers have used a variety of image processing and AI-based technologies (including machine and deep learning), but they mostly use a front-view camera perspective [14]–[16].

Ahmed and Jeon [17] demonstrated a method that employs an overhead camera viewpoint. SiamMask, a deep learning-based technique employed in the system, is simple, adaptable, and fast, outperforming other real-time tracking systems. The method incorporates the mask branch into the full convolutional duplex neural network to track the target or person. The person's video sequence is first gathered from an overhead perspective; then, additional training is conducted through learning transfer. Finally, a comparison is made to different tracking algorithms. The SiamMask algorithm produces excellent results.

We have constructed a workable people monitoring system that solves most real-world concerns, according to John Krumm. During live demos in the living room, use two color cameras to track various participants [18]. Color photos protect people's identities, while stereo images are used to locate them. The system is fast enough to give the impression that the room is responding [19].

According to Gómez-Silva [19], he took advantage of the transfer of learning from a multi-object tracking (MOT) domain of two photographs of a specific person to be identified and tracked. In both domains, he taught a unique deep triple structure. Six levels of translational learning were implemented and analyzed, demonstrating that transferring knowledge from one area to another significantly improves re-identification performance. The experimental results demonstrate that the proposed method is comparable to existing state-of-the-art procedures in terms of accuracy and robustness. These findings also demonstrate that, despite the data issue, deep learning can be effectively utilized for single-shot redefinition.

Weiyang addresses the problem of deep facial recognition (FR) under the Open Group Protocol. Where ideal facial features are expected to have a maximum distance within the layer, equal to the minimum distance between categories under a suitably selected measurement area, a few existing algorithms can effectively achieve this standard. An angular SoftMax loss (A-SoftMax) enabled CNNs to learn angular discriminant features. We also conclude a specific approximation of the ideal feature standard. Intensive analyses and experiments are conducted on the so-called wild in the wild (LFW) and faces (JTF). As a result, a novel approach to deep recognition in face recognition was introduced. Specifically, it is beneficial to learn facial representation. Competitive results in many common facial standards outweigh our approach and its enormous potential [16].

Mustafa utilized kernel discriminant analysis (KDA) and SVM facial recognition algorithms, which applied kernel analysis to extract features from the input images. Additionally, this procedure was applied to both the Yale and ORL databases to assess the performance of the proposed system. Experimental results showed that the system achieved a high recognition rate, with an accuracy of 95.25% on the Yale database and 96% on the ORL [16].

Kumar *et al.* [20] has proposed an arrangement for most problems facing the daze within the category of exploring both inner and outer situations, comprising different obstacles and the acknowledgment of an individual before them. Faces can be recognized utilizing neural learning methods, including extraction and preparation modules. Pictures of companions and relatives are stored within the smartphone client database, and an unused picture recognition and navigation system provides precise and quick voice messages to individuals with visual impairments, enabling them to navigate quickly. The framework applies to open-air and indoor situations [21], [22]. An examination of the execution of the actualized framework reveals that 75% of the participants find this framework helpful and achieve a 90% accurate result for face recognition [23].

According to Fachrurrozi *et al.* [24], the real-time facial recognition system process is divided into three steps: feature extraction, clustering, detection, and recognition. Each stage uses a different method: the local binary pattern (LBP), cumulative hierarchy (AHC), and Euclidean distance. CBIR, an image search technology based on image features, is applied as a search method. Based on experiments and test results, the recall and accuracy values were 65.32% and 64.93%, respectively. Table 1 summarizes the related work.

Table 1. Comparison of previous works

Title	Year	Problem	Technique used	Result
"A real-time person tracking system based on SiamMask network for intelligent video surveillance"	2021	Video surveillance of people	Deep learning Algorithm SiamMask	The algorithm used gave excellent results compared to other methods
"Multi-camera multi-person tracking for EasyLiving"	2000	Monitor people in the living room	Use the technology of two-color cameras and match photos	Suggest a workable people monitoring system that solves most real-world concerns.
"Transferring learning from multi-person tracking to person re-identification"	2019	Re-identification of the single person (re-id)	The transfer of learning from multi-object tracking (MOT)	The proposed method is comparable to existing state-of-the-art procedures in terms of accuracy and robustness
"Face recognition system based on kernel discriminant analysis, k-nearest neighbor and support vector machine"	2018	Face Recognition	SVM and KNN	According to the experimental results, the system achieves a high recognition rate, with an accuracy of 95.25% in the Yale database and 96% on the ORL database.
"Intelligent face recognition and navigation system using neural-learning for smart security in the internet of things"	2017	System for blind people	Genetic algorithm and RBF Kernel	The system is helpful and provides a 90% accurate result for face recognition.
"Multi-object face recognition using content Based image retrieval (CBIR)"	2017	Human face recognition in a multi-object image,	LBP and AHC	Based on experiments and test results, the recall and accuracy values were 65.32% and 64.93%, respectively.

4. RESULTS AND DISCUSSION

The implementation phase is the final phase that determines the project's success. It is not possible to plant cameras in public places because this requires a security approval machine. We tangibly experimented with the system by providing it with an album of photos taken from Google of celebrities, along with data identifying 50 people. The database was supplied with people's information, and their images were added to the database. We ran the program using the laptop's camera and passed pictures in front of it. This experiment was conducted in two batches: the first using dim light and the second under standard viewing conditions. The results are recorded as follows. The process was completed in low light when 50 typical images were read into the system. The experimental phase identified 46 images; at this stage, the system achieved a 92% accuracy in determining the wanted persons. The process was conducted in normal light with the exact mechanism followed previously, and 50 photos of people were identified out of 50. The system achieved 100% of the goal of distinguishing the wanted people to be tracked by regular people, the ideal ratio. These results indicate that the CNN-based approach is capable of maintaining high accuracy across varying illumination conditions, validating the robustness of feature extraction within the proposed architecture. In particular, the minimal performance drop in low light demonstrates the advantage of deep feature representation over traditional handcrafted techniques, which typically degrade sharply under these conditions.

The achieved precision, recall, and F1-scores as shown in Table 2 further support the reliability of the model, showing that misclassifications are rare and primarily associated with low-resolution or partially occluded test images. Importantly, the system demonstrated strong stability during real-time execution, with the model capable of identifying individuals within fractions of a second, making it suitable for surveillance and rapid-response scenarios.

Table 2. Accuracy of the results

	Precision	Recall	F1-score	Support
0	0.99	0.99	0.99	3800
Accuracy			0.99	760
Macro avg	0.99	0.98	0.99	760
Weighted avg	0.99	0.99	0.99	760

Use the following logic to calculate the percentage achieved.

$$\text{Tracking rate} = \frac{The\ number\ that\ has\ been\ recognized}{All\ number} * 100$$

Based on the results and their comparison, [16], [20] have achieved results that are less than those in our proposed system, as the researchers used different algorithms than the technology employed in our system. The following data in Table 2 shows the accuracy of the results obtained from the system. Compared with prior studies using PCA, LBP, SVM, and hybrid CBIR approaches, the proposed model outperforms earlier systems in both accuracy and consistency. While prior methods report accuracies ranging from 65% to 96%, our system consistently achieves accuracies of 92%–100%, depending on lighting conditions. This demonstrates the benefit of deep learning–based hierarchical feature encoding in environments where subtle variations in pose, expression, or illumination can significantly reduce the performance of classical recognition algorithms.

4.1. Benchmarking against other models

To further validate the effectiveness of the proposed CNN-based tracking system, we conducted a conceptual and performance-oriented benchmarking comparison against several widely used facial recognition and tracking models, including PCA-based recognition, LBP descriptors, SVM classifiers, SiamMask deep tracking, and CBIR-based multi-object recognition systems. Traditional PCA and LBP approaches demonstrated strong performance only under controlled lighting and frontal-face conditions, with accuracies typically ranging from 65% to 90% according to earlier studies. However, both techniques struggled in scenarios involving illumination variability, occlusion, and complex backgrounds. This aligns with our findings, where CNN-based recognition maintained high accuracy (92%–100%) even under low-light conditions, unlike handcrafted feature models, which tend to degrade significantly in such situations.

Similarly, SVM-based pipeline models—although effective for small, well-labeled datasets—require manual feature extraction and do not scale well to larger or more diverse datasets. In contrast, CNNs inherently learn multi-level feature hierarchies from raw images, enabling superior generalization and robustness. These advantages are clearly reflected in the higher precision, recall, and F1-scores obtained in our experiments.

When benchmarked conceptually against state-of-the-art deep learning methods such as SiamMask, which focuses on real-time object tracking rather than identity recognition, our system provides an added advantage by combining both detection and identity verification. SiamMask excels in motion tracking but does not inherently perform facial recognition. Our model integrates both capabilities, making it more suitable for security applications requiring identity-specific tracking rather than generic object localization.

Compared with CBIR-based multi-object recognition techniques, which achieved recall and accuracy values around 65%, the proposed CNN system demonstrates a significant performance improvement. The deep convolutional architecture reduces sensitivity to noise and background clutter—two major limitations of CBIR approaches that rely heavily on texture similarity rather than learned feature abstraction. Overall, the benchmarking results confirm that CNN-based methods not only outperform classical machine learning techniques but also offer advantages over several modern deep-learning trackers by integrating identity recognition, multi-person detection, and encrypted data handling within a unified, practical framework.

4.2. Limitations

Despite promising results, the proposed system has several limitations that must be acknowledged. First, the model was trained and evaluated using a relatively small dataset, which may restrict it is generalization to broader demographic or environmental variations. Larger and more diverse datasets would further strengthen the reliability of the system.

Second, the experiments were conducted using a laptop camera under controlled indoor conditions [25]. Real-world surveillance environments often present additional challenges, including extreme lighting conditions, motion blur, high crowd density, partial occlusions, and varying camera angles [26]. Future work should include field deployment with outdoor and multi-camera settings to assess system performance under operational constraints.

Third, the current system relies on frontal or near-frontal facial visibility. While CNNs improve robustness, extreme side profiles or heavily occluded faces still pose recognition challenges. Integrating body pose estimation or multi-modal biometrics could reduce this dependency. Finally, although the system incorporates hash-based encryption for data protection, additional security layers—such as secure communication tunnels, differential privacy, or GDPR-compliant anonymization—should be considered for deployment in governmental or commercial environments.

4.3. Need for larger and more diverse datasets

Another important limitation of the current study is the restricted size and diversity of the dataset used for model training and validation. The system was evaluated using a relatively small number of images, many of which were sourced from publicly available celebrity photos with limited variation in pose, age, ethnicity, environmental background, and camera angle. This level of homogeneity may not fully represent real-world surveillance conditions, where individuals differ significantly in appearance and are often captured in dynamic, uncontrolled settings.

To ensure stronger generalization and reduce the risk of overfitting, future work should incorporate significantly larger and more diverse datasets that include multiple age groups, ethnicities, expression variations, occlusions (such as masks or glasses), and environmental conditions. Expanding the dataset in this manner would allow the CNN model to learn a broader distribution of facial features, ultimately improving accuracy and robustness when deployed in realistic field environments. Additionally, the inclusion of publicly available large-scale datasets, such as LFW, VGGFace2, or MS-Celeb-1M, or custom datasets gathered from operational surveillance cameras, would provide more comprehensive validation and a more reliable measure of real-world performance.

5. CONCLUSION

Tracking is often referred to as a service, and like any other modern technology or technical assistance, it has advantages and disadvantages. Tracking people via GPS technology has become outdated because it can be difficult to place a device on a person if they are wanted for crimes—alternative technology is needed. This proposed system presented a different mechanism from the tools previously used in tracking people. The model was built using CNN techniques and trained intelligently by TensorFlow. This trained form identifies the wanted persons and sends their information and reality to the system administrator. The form must be connected to its camera for tracking to take place. The system test is conducted in two stages; the first is in case the light is not good at the camera. The system was trained with images; 50% of the images were identified, and the current information was determined. In the other part, an experiment was conducted on the system under normal lighting conditions, using the exact mechanism employed in the first test. The system achieved a 100% accuracy rate; all images were recognized, and the coding folk was also evaluated using the hash algorithm, with an accuracy score of 99, which is considered good.

FUNDING INFORMATION

The authors state no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Zinah Shiker Makki	✓	✓	✓	✓				✓	✓	✓				
Ahmet Zengin		✓						✓	✓	✓				

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] J. Dhamija, T. Choudhury, P. Kumar, and Y. S. Rathore, "An advancement towards efficient face recognition using live video feed: 'for the future,'" in *2017 3rd International Conference on Computational Intelligence and Networks (CINE)*, IEEE, Oct. 2017, pp. 53–56, doi: 10.1109/CINE.2017.21.

[2] M. Hada, R. Yamada, and S. Akamatsu, "How does the transformation of an avatar face giving a favorable impression affect human recognition of the face?," in *2018 International Workshop on Advanced Image Technology (IWAIT)*, IEEE, Jan. 2018, pp. 1–3, doi: 10.1109/IWAIT.2018.8369655.

[3] H. Kwon, O. Kwon, H. Yoon, and K. W. Park, "Face friend-safe adversarial example on face recognition system," in *International Conference on Ubiquitous and Future Networks, ICUFN*, 2019, pp. 547–551, doi: 10.1109/ICUFN.2019.8806124.

[4] S. Zhao, Z. Xu, Z. Zhu, X. Liang, Z. Zhang, and R. Jiang, "Short and long-term renewable electricity demand forecasting based on CNN-BI-GRU model," *ICCK Transactions on Emerging Topics in Artificial Intelligence*, vol. 2, no. 1, pp. 1–15, 2025, doi: 10.62762/tetai.2024.532253.

[5] S. Joshi, B. L. Mahanthi, G. Pavithra, K. S. Pokkuluri, S. S. Ninawe, and R. Sahu, "Integrating LSTM and CNN for stock market prediction: a dynamic machine learning approach," *Journal of Artificial Intelligence and Technology*, vol. 5, pp. 168–179, 2025, doi: 10.37965/jait.2025.0652.

[6] G. George, R. Boben, B. Radhakrishnan, and L. P. Suresh, "Face recognition on surgically altered faces using principal component analysis," in *Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2017*, 2017, doi: 10.1109/ICCPCT.2017.8074324.

[7] J. Boudjadar, S. U. Islam, and R. Buyya, "Dynamic fpga reconfiguration for scalable embedded artificial intelligence (AI): a co-design methodology for convolutional neural networks (CNN) acceleration," *Future Generation Computer Systems*, vol. 169, 2025, doi: 10.1016/j.future.2025.107777.

[8] M. A. Wahed, M. Alqaraleh, M. S. Alzboon, and M. S. Al-Batah, "Evaluating AI and machine learning models in breast cancer detection: a review of convolutional neural networks (CNN) and global research trends," *LatIA*, vol. 3, p. 117, Jan. 2025, doi: 10.62486/latia2025117.

[9] M. M. Das and S. Shastri, "Machine learning approaches for early brain stroke detection using CNN," *Journal of Scientific Research and Technology*, pp. 243–250, 2025.

[10] B. Saoud, A. H. H. M. Mohamed, I. Shayea, A. A. El-Saleh, and A. Alashbi, "Review of masked face recognition based on deep learning," *Technologies*, vol. 13, no. 7, p. 310, Jul. 2025, doi: 10.3390/technologies13070310.

[11] S. Sambolek and M. Ivasic-Kos, "Determining the geolocation of a person detected in an image taken with a drone," 2023, doi: 10.2139/ssrn.4373987.

[12] T. Yingming, D. Kenan, Q. Jianfeng, F. Li, and C. Yi, "A learning observer-based control strategy for PMSM with position sensor fault in railway," *Control Engineering Practice*, vol. 142, p. 105705, Jan. 2024, doi: 10.1016/j.conengprac.2023.105705.

[13] H. Torkey, S. Hashish, S. Souissi, E. E. D. Hemdan, and A. Sayed, "Seizure detection in medical IoT: hybrid CNN-LSTM-GRU model with data balancing and XAI integration," *Algorithms*, vol. 18, no. 2, 2025, doi: 10.3390/a18020077.

[14] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018, doi: 10.1109/ACCESS.2017.2788044.

[15] N. Otterness *et al.*, "An evaluation of the NVIDIA TX1 for supporting real-time computer-vision workloads," in *2017 IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, IEEE, Apr. 2017, pp. 353–364, doi: 10.1109/RTAS.2017.3.

[16] M. Z. Al-Dabagh, M. H. M. Alhabib, and F. H. AL-Mukhtar, "Face recognition system based on kernel discriminant analysis, k-nearest neighbor and support vector machine," *International Journal of Research and Engineering*, vol. 5, no. 2, pp. 335–338, 2018, doi: 10.21276/ijre.2018.5.3.3.

[17] I. Ahmed and G. Jeon, "A real-time person tracking system based on SiamMask network for intelligent video surveillance," *Journal of Real-Time Image Processing*, vol. 18, no. 5, pp. 1803–1814, Oct. 2021, doi: 10.1007/s11554-021-01144-5.

[18] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer, "Multi-camera multi-person tracking for Easyliving," in *Proceedings Third IEEE International Workshop on Visual Surveillance*, IEEE Comput. Soc, 2000, pp. 3–10, doi: 10.1109/VS.2000.856852.

[19] M. J. Gómez-Silva, E. Izquierdo, A. de la Escalera, and J. M. Armingol, "Transferring learning from multi-person tracking to person re-identification," *Integrated Computer-Aided Engineering*, vol. 26, no. 4, pp. 329–344, Sep. 2019, doi: 10.3233/ICA-190603.

[20] P. M. Kumar, U. Gandhi, R. Varatharajan, G. Manogaran, J. R., and T. Vadivel, "Intelligent face recognition and navigation system using neural learning for smart security in Internet of Things," *Cluster Computing*, vol. 22, no. 4, pp. 7733–7744, 2019, doi: 10.1007/s10586-017-1323-4.

[21] K. Ngamakeur, S. Yongchareon, J. Yu, and Q. Z. Sheng, "Deep CNN-LSTM network for indoor location estimation using analog signals of passive infrared sensors," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22582–22594, Nov. 2022, doi: 10.1109/JIOT.2022.3183148.

[22] S. Yongchareon, J. Yu, and J. Ma, "Efficient deep learning-based device-free indoor localization using passive infrared sensors," *Sensors*, vol. 25, no. 5, p. 1362, Feb. 2025, doi: 10.3390/s25051362.

[23] S. P. Shashikumar, A. J. Shah, Q. Li, G. D. Clifford, and S. Nemati, "A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology," in *2017 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, IEEE, 2017, pp. 141–144, doi: 10.1109/BHI.2017.7897225.

[24] M. Fachrurrozi, Erwin, Saparudin, and Mardiana, “Multi-object face recognition using content based image retrieval (CBIR),” in *2017 International Conference on Electrical Engineering and Computer Science (ICECOS)*, IEEE, Aug. 2017, pp. 193–197, doi: 10.1109/ICECOS.2017.8167132.
[25] J. Yoon, H. Kim, D. Lee, and S. Kwon, “Indoor positioning method by CNN-LSTM of continuous received signal strength indicator,” *Electronics*, vol. 13, no. 22, p. 4518, Nov. 2024, doi: 10.3390/electronics13224518.
[26] A. Nadeem *et al.*, “Tracking missing person in large crowd gathering using intelligent video surveillance,” *Sensors*, vol. 22, no. 14, p. 5270, Jul. 2022, doi: 10.3390/s22145270.

BIOGRAPHIES OF AUTHORS

Zinah Shiker Makki is a dedicated professional who began her career in administrative roles after graduating from Al-Maaref University in 2013. Since 2020, she has been working as an Administration Coordinator at Mansur Law Firm–Advocates and Legal Consultants, where she plays an active role in managing and coordinating internal organizational processes. Committed to continuous academic growth, Makki is currently pursuing a Master’s degree at Sakarya University’s Institute of Computer and Information Engineering. The combination of her educational background and professional experience equips her with strong competencies in both administrative and technical domains. She can be contacted at email: zenashmakky@gmail.com.

Ahmet Zengin is currently a Professor at Sakarya University, Türkiye. His primary research interests include parallel and distributed simulation, as well as high-level architecture. He can be contacted at email: azengin@sakarya.edu.tr.