International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 6, pp 193-200, June-2026

https://doi.org/10.47001/IRJIET/2026.106024

A Survey on Face Detection and Recognition Techniques

[1]Ashwini Bhamre, [2]Chetan K. Wani, [3]Shashikant K. More, [4]Aaradhya S. Sonar

[1]Assistant Professor, Department of Information Technology Engineering, PES Modern COE, Pune, Maharashtra, India
[2,3,4]Student, Department of Information Technology Engineering, PES Modern COE, Pune, Maharashtra, India

***Abstract* - The use of face detection and recognition has emerged as a theme of research in the field of computer vision due to their wide application in security systems, biometrics authentication, access control and intelligent video analysis. The rapid development of visual information produced by surveillance systems and digital imaging systems has raised the need to find automated, scaled up, and robust face recognition methods. This is a contribution that has made a lot of progress over the last ten years following the shift of traditional feature-based techniques involving handcrafted features to deep learning driven detection and embedding based recognition models.**

The paper is a thorough review of face detection methods, face recognition systems, strategies of similarity measurements and real-time surveillance systems. Classical techniques and the contemporary deep learning-based techniques are systematically reviewed and compared. The survey also considers the distributed and large-scale processing structures on one side, and system level issues like real-time performance, scalability, open set recognition and computational efficiency on the other. Judging by the comparative analysis, it is possible to determine the main research gaps and perspectives in the future and offer the systematized vision of the existing developments in face detection and recognition technologies.

Keywords: Face Detection, Face Recognition, Deep Learning, Surveillance Systems, RetinaFace, ArcFace, Facial Embeddings.

I. INTRODUCTION

Face detection and recognition has developed into a high profile area of research in computer vision and pattern recognition because of its broad range of application in security systems, biometric authentication, forensic pattern recognition, access control, and intelligent video surveillance The fast development of imaging equipment and surveillance infrastructure has created massive amounts of visual information that require automated and scalable face recognition algorithms, which can perform at various real world scenarios.

Early face recognition systems were mainly based on handwritten feature extraction schemes with classical classifiers Techniques Haar cascade classifiers [6], Principal Component Analysis (PCA) [7], and statistical appearance based models [8] were common owing to their computational simplicity and capability to run on a limited hardware platform but failed to perform well in an unconstrained settings, including variations in illumination, change in pose, conclusiveness, and low resolution imagery.

Convolutional Neural Networks (CNNs) The advent of deep learning has largely changed face analysis Multistage face detectors like MTCNN [4], single stage face detectors like RetinaFace [2], and embedding based recognition models like ArcFace [1] have both shown strong localization accuracy across scale and open set recognition Multistage detectors Multistage detectors, e.g. MTCNN [4], have demonstrated high localization accuracy across scales; single stage detectors Multistage detectors, e.g. RetinaFace [2], have demonstrated high localization accuracy in Recent work has, however, also been extended to scalable and real-time deployment models including distributed stream processing systems [12], CNN based real-time recognition systems [14], and hybrid edge cloud surveillance models [15] Nonetheless, the variety of existing models including traditional handcrafted models, deep embedding based models, similarity matching models, and system level architecture makes it hard to gain a coherent and systematic view of the discipline.

The purpose of the survey is to offer a critical overview of the face detecting and recognizing techniques The paper reviews critically the current methods of detecting faces with traditional and deep learning approaches, embedding recognition models and similarity strategies of recognition evaluation and the surveillance architecture in real-time By summarizing the current research studies and revealing the significant gaps in understanding the field, the paper will present a systematic overview of the current developments and suggest the important directions of the research.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

II. LITERATURE REVIEW

A. Face Detection Techniques

Face recognition Face detection is a basic element in surveillance-based systems of face recognition since inaccuracy at this point is directly related to the accuracy in extracting features and matching identities. In actual world surveillance setups, face detection will be required to work under harsh conditions like illumination variation, change of pose, occlusion, and scale variation. Consequently, there has been a lot of research done to apply more strength and reliability of face detection methods in unconstrained environment.

1. Traditional Face Detection Approaches

Earlier face detection algorithms were mostly based on manual features with the classical machine learning classifiers. The Haar Cascade classifier that uses Haar like features and an Adaboost based cascade learning algorithm was widely adopted, because of its low processing cost and real-time execution on the finite hardware platform. A number of missing person identification systems recognized Haar Cascade based detectors to detect facial parts in CCTV images due to their simplicity and simplicity to apply [6], [8].

Although these are the benefits, the traditional detectors are very sensitive to the changes in the environment like lighting conditions, background complexity, and facial position. Such restrictions tend to cause more false detections and missed faces especially in outdoor or large crowds in surveillance. To enhance stability in the detection, feature descriptors like Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) were proposed to obtain edge and texture information. Their handcrafted nature limited the capacity of these techniques to generalize even in different surveillance conditions although they provided incremental improvements. It has been reported that conventional detectors are not good at low resolution faces and partial occlusions that are usually found in long-range CCTV images [7], [8].

2. Deep Learning-Based Face Detection

The failure of the handcrafted feature-based methods prompted the use of convolutional neural network (CNN) based face detection models. CNN based detectors automatically discover hierarchical feature representations of data, which allow them to be more robust to changes in pose, scale, and illumination. Multistage detection systems like the Multitask Cascaded Convolutional Network (MTCNN) proposed a cascaded system with proposal, refinement, and output steps to refine and localize the face respectively, using the landmarks [4]. Although multistage detectors have good detection accuracy, they have a sequential processing nature that reduces further computation load. This can constrain their performance on large-scale, real-time surveillance systems that require low latency. However, several researchers have indicated that deep learning-based detectors can easily beat traditional effective systems in difficult setup, especially in case of dealing with occlusions, rotations of faces, and changes in resolution [7]. One of the most important challenges of surveillance usage is to have the optimal balance between accuracy of detection and the rate of processing.

3. Single-Stage Face Detection Using RetinaFace

Single-stage face detection models have come into the limelight to deal with the trade-off between accuracy and efficiency. The concept of face detection in RetinaFace is a dense prediction problem, so to do well in face detection, it uses Feature Pyramid Networks (FPN) to find facial features at different scales [1], [2]. In contrast to previous detectors, RetinaFace simultaneously predicts facial bounding boxes and five-point facial landmarks, making it possible to localize faces as well as allowing alignment in a single framework. Multiscale RetinaFace based techniques are shown to have high performances in small and partially covered faces which are usually overlooked by conventional and multistage detectors [1], [2]. Experiments with alternative backbone architectures, like ResNet and MobileNet variations, point to the flexibility of RetinaFace for both accuracy and real-time deployment specifications [1]. As per the literature, experimental results show that RetinaFace is highly accurate and has high recall on the benchmark datasets like WIDER FACE and hence is suitable in the surveillance-oriented applications [2].

4. Comparative Discussion and Observations

The surveyed literature indicates an evident shift from the handcrafted feature-based detectors to deep learning-based architectures. Although such traditional methods as Haar Cascade can also be used in controlled or low resources environments, they have limited robustness that limits its practical implementation in real world surveillance systems [6], [8]. CNN-based detectors are much better at enhancing the abilities of detection, but multistage pipelines might add latency, which can impact real-time performance [4]. The single-stage detectors like RetinaFace can offer a balance that is effective between accuracy and computational efficiency. The concurrent determination of facial landmarks also simplifies the downstream alignment and identification operations, which simplify the system and propagation of errors in the system [1], [2].

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

B. Face Recognition and Embedding-Based Methods

The most dominant part of biometric identification systems is face recognition, and the main task of this system is to depict the facial features of individuals in a small but very discriminative way. Recent face recognition systems mostly use the embedding models of deep learning, which encode facial images and convert them to feature vectors of fixed length. These embeddings allow the effective similarity-based comparison and scalable identity matching in large databases. Embedding-based methods have shown to be better than traditional methods of the classification-based approach, especially in unconstrained surveillance setups.

1. Traditional Face Recognition Approaches

Early face recognition methods were based on hand-written feature extraction methods (Eigenfaces, Fisherfaces, and Local Binary Pattern Histograms (LBPH)). The techniques were used to capture low-level statistical, or texture-based facial data and use distance measures or threshold-based decision rules as identity recognition techniques. LBPH-based methods have been embraced in several early missing person identification systems because of their simplicity and low computational cost [8]. Nevertheless, feature-based approaches that are handcrafted are rather weak when subjected to reality distortions like changes in illumination, facial pose, expression, and aging. It has been demonstrated that these methods do not work well on low-resolution surveillance images and in situations where partial occlusion is present, which are typical features of CCTV-based surveillance systems [6], [8].

2. Deep Learning-Based Face Recognition

The development of convolutional neural networks (CNNs) has been an innovative breakthrough in the field of face recognition because of their ability to learn hierarchical facial representations directly from the data automatically. The first CNN-based recognition systems modelled the problem as a closed-set classification problem, with each identity having a predefined class. Although useful with small datasets, these methods cannot be used in actual real-world surveillance situations where new identities are encountered all the time. To overcome this shortcoming, embedding-based learning frameworks were proposed. These models are trained with the objective of mapping a facial image to a high-dimensional feature space where samples of the same person are close and those of different persons are distant. This idea was popularized by FaceNet with the help of triplet loss, which directly sets relative distance constraints between anchor, positive, and negative samples. Later research has shown that embedding-based models are much better than classification-based systems in open-set recognition conditions, as is common in surveillance systems [7]. Besides embedding-based recognition models, several applied systems incorporate cloud-assisted facial recognition services to identify missing persons. Pawar et al. [9] introduce a practical framework which utilizes deep learning and facial encoding using a cloud recognition platform. Their system enables the uploading of facial images and matching them to a centralized database to rapidly identify and notify the authorities. These application-oriented designs indicate the possibility of implementing face recognition solutions in real world setting, but this depends on third party services to detect and expand.

3. ArcFace and Angular Margin–Based Loss Functions

ArcFace is a major progress of embedding-based face recognition by adding an additive angular margin on the loss function to boost the feature discrimination [1]. In contrast to traditional softmax loss, ArcFace normalizes feature vectors and class weights, which project embeddings on a hyperspherical surface. Angular distance therefore controls identity separation as opposed to vector magnitude, which makes the classes more separable. The surveyed studies always show that ArcFace-generated embeddings have high intra-class compactness and better inter class separation, which are very useful in large-scale face recognition applications [1], [3]. Compared to other margin-based loss functions, such as SphereFace and CosFace, ArcFace has been shown to be more accurate in recognition in adverse conditions like pose variation and low-resolution images [1], [2]. These features render ArcFace very suitable in recognition systems that are based on surveillance and demand a dependable identity discrimination.

4. Embedding Normalization and Similarity Based Recognition

Embedding-based recognition Embedding based recognition systems typically use L2 normalization of the feature vectors before matching so as to be scale invariant. Normalized embeddings enable the similarity to be computed consistently with the use of distance measures, including cosine similarity or Euclidean distance. Of these, the cosine similarity is the most popular since it is efficient in calculation as well as high-dimensional embedding space. Various implemented missing person identification systems based on surveys use the cosine similarity to match facial embeddings and identify identity matches [3], [5]. Similarity-based decision mechanism facilitates open set recognition wherein systems can reject an identity not found on the database through predefined similarity criteria. This is needed in a surveillance situation where most of the faces that are seen are not of registered faces.

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

5. Comparative Discussion and Observations

According to the literature that was surveyed, there is an obvious shift towards deep embedded-based models as opposed to handcrafted feature-based recognition methods. However, even though the conventional methods, i.e., LBPH and PCA are computationally lightweight, they are not so robust and this therefore limits their use in real world surveillance contexts [6], [8]. Conversely, embedding models developed by means of deep learning, in particular those that use angular margin losses such as ArcFace, are characterized by better discrimination ability, scalability, and resistance to environmental changes [1], [3].

C. Similarity Measures and Matching Strategies

Similarity measurement is one of the main elements of embedding-based face recognition systems because it defines the effectiveness of extracted facial representations in comparison to identify identity. In missing person identification by surveillance, similarity matching needs to be highly accurate and at the same time must be computationally efficient and open-set able to handle open set cases where unknown persons often occur. Consequently, extensive research has been dedicated to the development of the similarity metrics and decision mechanisms of the facial embedding comparison, which are reliable.

1. Distance-Based Similarity Metrics

In embedding-based recognition models, the surface images of facial identity are expressed as high-dimensional feature space fixed-length vectors. Identity comparison is widely done using distance-based similarity measures, with the most common ones being the Euclidean distance and the cosine similarity. Euclidean distance determines the absolute distance between two feature vectors in space, whereas cosine similarity determines the angular relationship by calculating the cosine of the angle between the two. Some of the missing person identification systems were surveyed, which used Euclidean distance with classifiers like k-Nearest Neighbours (k-NN) based on its simplicity and ease of implementation [5]. But Euclidean distance is susceptible to the difference in the magnitude of features, and this can adversely affect matching integrity in cases where embedding is not normalised.

2. Cosine Similarity for Embedding Comparison

Cosine similarity has since been the metric of similarity of choice in deep face recognition, especially normalized embedding-based systems. Cosine similarity can offer scale-invariant comparison and better performance in diverse acquisition conditions by solely comparing angular separation. In several of the studies surveyed, cosine similarity is used to match deep facial embeddings in a missing person identification and surveillance based application [3], [7]. Embeddings are explicitly L2 normalized and confined to hyperspherical manifold in ArcFace-based recognition schemes and so cosine similarity is a natural and productive identity comparison metric [1]. The empirical analysis conducted in the literature shows that cosine similarity has always shown a superior recognition accuracy and less false acceptance of falsehood compared to Euclidean distance in high dimensional embedding spaces [1], [3].

3. Threshold-Based Decision Strategies

There is no identity that is defined based on similarity computation without a proper decision mechanism. Application Threshold-based decision strategies are typically applied in open set recognition setups whereby the similarity score between a probe embedding and stored templates is compared against a predetermined threshold. When the score value is higher than the threshold, the identity is admitted as a match; otherwise, it is rejected as unknown. Several systems that are surveyed adopt set thresholds obtained by empirical validation [3], [7]. Although such thresholds are easy to apply, it is not always easy to generalize them in different settings, camera characteristics, or larger populations. The choice of threshold is a significant factor that can elevate the false positive or false negative rates to levels that are unacceptable, especially in the cases of missing person identification. Other systems used in the real world use the vendor-controlled similarity computation and decision systems instead of explicit distance modelling. In [9], the identity matching is carried out with the help of cloud-based face recognition APIs, where the similarity level and decision rule are internalized by the service provider. Although it makes system design and deployment simpler, this restricts transparency and control over corresponding behaviour, which can be paramount in sensitive systems such as missing person identification.

4. Open Set Recognition and Scalability Considerations

The face recognition systems that are based on surveillance by default are under open set conditions because most people who are viewed are not in the records. The similarity matching inherent in the embed-based similarity support opens set recognition through rejection based on similarity scores. however, Scalability is a significant issue with the increase in size of the identity database. According to studies surveyed, similarity that matches performance may reduce with large scale databases because of higher computational loads and similarity score distributions [5], [7]. The methods of embedding indexing, approximate nearest neighbour search and similarity score normalization have been investigated as methods of overcoming these challenges.

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

Nevertheless, practical and precise similarity computation on large-scale real-time surveillance systems has not yet been accomplished, as an open research problem.

5. Comparative Discussion and Observations

The literature reviewed shows that similarity-based matching strategies have considerable benefits over the classifier-based recognition strategies in open-set and large-scale setting. Although Euclidean distance and k-NN classifier are effective in cases of small data sets, they tend to be superficial with regard to scaling and less resilient in the real-world scenario [5], [6]. Conversely, cosine similarity and normalized deep embeddings are better in discrimination and cheaper to compute [1], [3]. Decision mechanisms using threshold allow practical control of both the false acceptance and rejection rates, but again close threshold control is the only way to ensure effective deployment. Cloud-assisted systems like [9] are fast to deploy and provide centralization of processing functions but create network availability, privacy of data, and vendor-specific recognition model dependencies. It is these constraints that drive the interest to study locally deployable and transparent embedding-based surveillance-driven identification frameworks.

D. Real-time Surveillance and Face Processing Frameworks

Surveillance enabled by real-time face processing is the core of the missing person detection, biometric surveillance, and video big data applications. As opposed to offline recognition systems, real-time systems need to handle real-time video streams with the highest latency requirements and be robust to pose variation, lighting variations, motion blur, and compression artifacts. Recent studies have thus been done not only on enhancing recognition accuracy but also on system architecture optimization, scalability of the system, and computational efficiency.

1. Real-time RGB Tracking and Dense Face Modelling

Dense monocular facial modelling and tracking is one of the important research directions in real-time face processing. Thies et al. [10] suggested a RGB pure real-time reenactment of the face that could have the capability of working at an interactive frame rate. It is a reconstruction framework that rebuilds the face identity based on a global non-rigid model-based bundling strategy and dense photometric tracking on a ground-based optimization pipeline executed on a GPU-accelerated system. A low-dimensional subspace deformation model is used to transfer the expression and is less complex to compute without sacrificing realism. The system has a frame rate of about 28 frames per second (FPS) at a high resolution, and this indicates that dense 3D facial tracking can be done in real-time with commodity hardware. Despite the main use being reenactment, the underlying architectural principles that are optimization-based, reduced parameter space modelling, and hardware acceleration can be directly applied to the surveillance-based real-time systems based on facial processing.

2. Real-time Recognition Under Operational Constraints

Understanding In a real-world surveillance, real-time face recognition systems need to meet high requirements on operation beyond crude precision. A real-time recognition model of illegal online soccer broadcasting was proposed by Correia et al. [11]. The system integrates the RetinaFace detection to process the reference image with Haar cascade detection to process real-time stream image to strike a balance between detection accuracy and computation speed. FaceNet128 and FaceNet512 embeddings are used to extract features, and verification is done by using cosine similarity. One of the design choices is to implement a zero False Positive Rate (FPR = 0), though at the cost of reduced recall. It is representative of real-world deployment needs in which erroneous identification may cause legal and functional repercussions. The system has a trade-off between precision-oriented threshold calibration and throughput performance with the system processing about 20 FPS. Through such constraint-aware architectures, there is recognition of threshold tuning, one shot recognition and operational reliability in real-time surveillance purposes.

3. Distributed and Large-Scale Stream Processing Architectures

Scalability is another user requirement of real-time surveillance systems, especially when using the system in a multi-camera or city-scale setting. Kazanskiy et al. [12] tested distributed face recognition systems on Apache Storm and IBM InfoSphere Streams stream processing systems. The experiment indicates that using distributed systems, it is possible to reach a real-time speed of around 24 FPS when implemented in a five-node cluster. The paper sheds light on the importance of stream processing models in processing large-scale continuous image streams by parallelization of OpenCV cascade-based detection and analysing latency trade-offs. In the tested configuration, Apache Storm was found to be more scalable and have better throughput than IBM Streams. In addition, fault tolerance and guaranteed message processing mechanisms are introduced so that the system will be robust in the presence of high data arrival rates. Besides distributed stream processing systems, hybrid edge cloud systems have been suggested to improve the scalability and responsiveness of smart surveillance systems. Kanagamalliga et al. [15] have suggested a real-time face recognition system

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

with Haar Cascade detection and Convolutional Neural Networks (HC CNN) to improve smart video surveillance. The system uses edge computing to detect faces initially and cloud-based processing to analyse them further, hence balancing both computational requirements and latency. Reports on experimental results showed a processing speed that is more than 30 FPS with better F1 scores than solution-only CNN methods. This paper presents the increased trend to edge-assisted intelligent surveillance systems that can support real-time operation as well as scalability.

4. Video Based Optimization Strategies for real-time Processing

In In addition to the architectural design, the optimization strategies at the algorithm level are also important in the realization of real-time performance in video-based recognition systems. The ESCO-based deep CNN architecture in [9] mitigates the computational cost in the continuous video recognition task, addressing the problem of computational strain by proposing a refined clarity-based frame selection mechanism and Enhanced Social Collie Optimization (ESCO) to handle hyperparameter optimization. Recent work by Pranav and Manikandan [14] provides the design and assessment of the real-time face recognition system that is founded on convolutional neural networks. The structure is proposed, which combines Viola Jones detection with a specialized CNN architecture that is optimized by systematic parameter optimization, such as convolution filter sizes, pooling strategies, and dropout settings. The AT&T data and real-time camera input experimental evaluation revealed recognition accuracy of 98.75% and 98.00%, respectively. Moreover, the analysis of the execution time showed that most of the latency is caused by image acquisition and image processing, and very little of the overall processing time is contributed by CNN classification. These results highlight the significance of optimization of architectural tuning and preprocessing in attaining real-time deployment efficiency.

5. System Level Challenges and Trade Offs

Regardless of the great improvements, the system of real-time surveillance and face processing still experiences various challenges. To start with, it is still hard to maintain high recognition accuracy under unconstrained conditions that include small-resolution imagery, occlusion, large variations in pose, and dynamically changing lighting.

Second, open set recognition on big databases brings about threshold calibration issues, particularly where there is a need to have erroneous false positive controls put in place. Also, scalability between more than two cameras would require distributed processing structures that can strike a balance between throughput and latency. This can be enhanced with increasing parallelism, at the cost of introducing overhead in synchronization and variation in latency. Likewise, precision-oriented systems tend to lose recall in order to guarantee the reliability of the operations of such systems, like in constraint-driven recognition systems. A second issue is to strike a balance between hardware acceleration and deployment possibility. Although frame rates can be achieved using the GPU-based optimization, large-scale deployment can need cost-effective distributed or lightweight solutions. Hence, it is necessary that real-time surveillance systems have a combination of discriminative embedding models and computationally efficient detection strategies, as well as scalable streaming infrastructures, to assure robustness and operational practicality. Future research proposals involve adaptive thresholding in dynamic settings, hybrid tracking recognition pipelines, efficient indexing in massive embedding databases, and combining small-form deep models to advantageous edge-assisted surveillance.

III. RESEARCH GAP AND FUTURE DIRECTIONS

Although major strides have been made in the field of face detectors, recognition, and real-time surveillance systems, there have been several research gaps that do not allow the implementation of effective missing persons identification systems. To begin with, most high-accuracy recognition models are tested using controlled benchmark data; surveillance conditions are brutal with problems such as low resolution, occlusion, motion blur, lighting variation, and extreme pose diversity. Even though embedding-based methods, like ArcFace, have proven to be good in terms of discriminative ability, their robustness on a large scale, unconstrained multi-camera systems are not well studied. Second, in real-time frameworks, computational efficiency and recognition precision are typically favoured at the expense of other frameworks that attempt to achieve both. False positive strict systems are known to trade off recall, whereas high recall systems are known to be unsafe. Strategies of adaptive threshold calibration of dynamically changing environments are open.

Third, current literature does not cover scalability in multi-camera and city-scale deployments in detail. Whereas the distributed stream processing framework can enhance throughput, there is a lack of effective indexing and quick similarity search systems on large embedding databases. There are further challenges associated with the open set recognition with ever-growing identity databases.

Fourth, tracking and recognition modules tend to be loosely integrated. There is little exploration on hybrid tracking recognition pipelines, which take advantage of temporal consistency to reinforce identity. The use of temporal

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

embeddings and motion-aware identity refinement can play a very important role in increasing real-time reliability.

Fifth, lightweight and edge-assisted deployments need to be researched. Most real-time systems require either GPU acceleration or cluster-based systems that are not always possible in resource-limited environments. It requires efficient compact model development, which preserves discriminative power with minimal computational overhead to be used in large-scale deployment.

Lastly, privacy-preserving and ethical issues of real-time surveillance systems seldom get incorporated into the system design. Safe embedding storage, encrypted similarity matching, and audit transparency mechanisms should be researched more to implement them responsibly. The resolution of these research gaps will be essential towards coming up with scalable, robust, and ethically sound real-time missing person identification systems.

REFERENCES

[1] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4690–4699, 2019.

[2] R. Sharma, A. Verma, and P. Kulkarni, "Multiscale Facial Detection Using RetinaFace Architecture," *International Journal of Advanced Computer Vision and Applications*, vol. 12, no. 3, pp. 45–52, 2025.

[3] P. S. Rao, K. S. Reddy, and M. V. Kumar, "AI Finder: An Intelligent Framework for Finding Missing Persons Using Face Match Algorithm," *International Journal of Innovative Research in Technology (IJIRT)*, vol. 11, no. 2, pp. 112–118, 2025.

[4] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[5] S. B. Sharath and H. Kumar, "Identifying Missing Individual Using Artificial Intelligence," *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, vol. 4, no. 1, pp. 98–103, 2024, doi: 10.48175/IJARSCT 19215.

[6] G. A. V. R. C. Rao, S. K. Reddy, and A. Kumar, "Face Detection and Tracking to Find Missing Person Using Haar cascade Algorithm," *International Journal for Research Trends and Innovation (IJRTI)*, vol. 8, no. 7, pp. 234–239, 2023.

[7] B. Vinavatani, R. Mehta, and A. Shah, "AI for Detection of Missing Person," *in Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC), IEEE*, 2022, pp. 1–6, doi: 10.1109/ICAAIC53929.2022.9792672.

[8] M. Bharathwaj, P. Naveen, and S. Ramesh, "Missing Child Identification Using Deep Learning," *International Journal of Creative Research Thoughts (IJCRT)*, vol. 9, no. 6, pp. 1021–1026, 2021.

[9] S. Pawar, A. Patil, and R. Kulkarni, "Find Missing Person Using Artificial Intelligence," *International Research Journal of Engineering and Technology (IRJET)*, vol. 8, no. 12, pp. 1564–1568, 2021.

[10] J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, and M. Nießner, "Face2Face: real-time face capture and reenactment of RGB videos," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2387–2395.

[11] H. A. Correia, D. Pontes, and J. H. Brito, "Face recognition in online soccer streaming for piracy detection," *Multimedia Tools and Applications*, vol. 84, pp. 30737–30756, 2025, doi: 10.1007/s11042 024 20389 3.

[12] N. L. Kazanskiy, V. Protsenko, and P. Serafimovich, "Performance analysis of real-time face detection system based on stream data mining frameworks," *Procedia Engineering*, vol. 201, pp. 806–816, 2017.

[13] J. C. Musale, A. K. Singh, and S. Shirke, "Effective face recognition from video using enhanced social collie optimization based deep convolutional neural network technique," *Journal of Visual Communication and Image Representation*, 2024.

[14] P.K.B.Pranav and J.Manikandan, "Design and evaluation of a real-time face recognition system using convolutional neural networks," *Procedia Computer Science*, vol. 171, pp. 1651–1659, 2020.

[15] K. Sa, A. Ra, B. B. S. Krishna, and V. Pa, "Advancements in real-time face recognition algorithms for enhanced smart video surveillance," *Procedia Computer Science*, vol. 230, pp. 486–492, 2023.

AUTHORS BIOGRAPHY

Mrs. Ashwini Bhamre, Assistant Professor, Dept. of Information Technology Engineering, PES Modern College of engineering, Pune, Maharashtra, India.

Chetan Wani, Student, Dept. of Information Technology Engineering, PES Modern College of engineering, Pune, Maharashtra, India.

ISSN (online): 2581-3048
Volume 10, Issue 6, pp 193-200, June-2026
https://doi.org/10.47001/IRJIET/2026.106024

Shashikant More, Student, Dept. of Information Technology Engineering, PES Modern College of engineering, Pune, Maharashtra, India.

Aaradhya Sonar, Student, Dept. of Information Technology Engineering, PES Modern College of engineering, Pune, Maharashtra, India.

Citation of this Article:

Ashwini Bhamre, Chetan K. Wani, Shashikant K. More, & Aaradhya S. Sonar. (2026). A Survey on Face Detection and Recognition Techniques. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(6), 193-200. Article DOI https://doi.org/10.47001/IRJIET/2026.106024

****