International Journal on Emerging Research Areas (IJERA) *ISSN: 2230-9993*

DOI: 10.5281/zenodo.14650941

A Review of Methodologies for Detecting Missing and Wanted People Using Machine Learning and Video Surveillance

George P Kurias
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
Email: georgepkurias2025@cs.ajce.in

Gokul Krishna AU
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
Email: gokulkrishnaau2025@cs.ajce.in

Jifith Joseph
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
Email: jifithjoseph2025@cs.ajce.in

Sharunmon R
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
Email: sharunmonr2025@cs.ajce.in

Linsa Mathew
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
Email:linsamathew@amaljyothi.ac.in

***Abstract*—This literature review explores recent advancements in machine learning and image processing techniques for the detection of missing and wanted individuals through video surveillance. Sixteen studies were analyzed, focusing on various methodologies such as convolutional neural networks (CNNs), gait analysis, Bayesian networks, and 2D-3D facial recognition frameworks. The review highlights the applications, strengths, and limitations of these approaches in terms of accuracy, real-time applicability, and robustness in challenging conditions such as low-resolution footage and occlusion. This work aims to provide insights into the current research landscape and identify potential areas for future exploration, with a focus on improving the efficiency and scalability of surveillance-based identification systems.**

***Index Terms*—CNN, Facial Recognition, Gait Analysis, Bayesian Networks, Surveillance, Missing Persons Detection.**

I. Introduction

The problem of identifying missing or wanted individuals through video surveillance has gained increasing attention in law enforcement and public safety. The current methods rely on manual review and comparison, which are time-consuming and inefficient. This project leverages machine learning models to automate this process, using methodologies such as CNNs for facial recognition and Bayesian networks for predicting likely locations of suspects.

Several existing approaches, reviewed in this paper, address different aspects of human identification in forensic and real-world scenarios. This work integrates methodologies from these papers to create a robust system suitable for real-time video analysis, enabling faster identification and alerting.

II Machine Learning Techniques for Detecting Missing and Wanted People

A. *Convolutional Neural Networks (CNNs) for Gait and Face Recognition*

CNNs are widely used for person identification, specifically for facial and gait recognition. These networks excel in learning spatial hierarchies of features directly from image data. In the case of facial recognition, CNNs can handle variations in lighting, pose, and facial expressions, making them suitable for surveillance footage where such conditions are not ideal. For gait recognition, CNNs analyze walking patterns, which are unique to individuals and can serve as a robust identifier when faces are not visible. This technique is particularly useful in low-resolution or crowded scenes where facial recognition struggles. [13], [9].

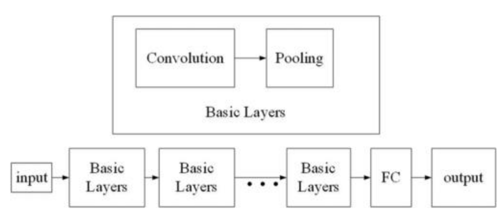

Fig. 1. CNN-based Gait and Face Recognition [13].

B. *2D to 3D Facial Image Analysis*

Significant advancement in facial recognition involves converting 2D images into 3D models to improve recognition accuracy. By detecting key facial landmarks and constructing a 3D facial mesh, these systems can overcome issues of pose

DOI: 10.5281/zenodo.14650941

variation, occlusion, and lighting inconsistencies. This method is especially beneficial in surveillance environments where the

DOI: 10.5281/zenodo.14650941

subject may not always face the camera directly, allowing for multi-angle identification. [1], [11]

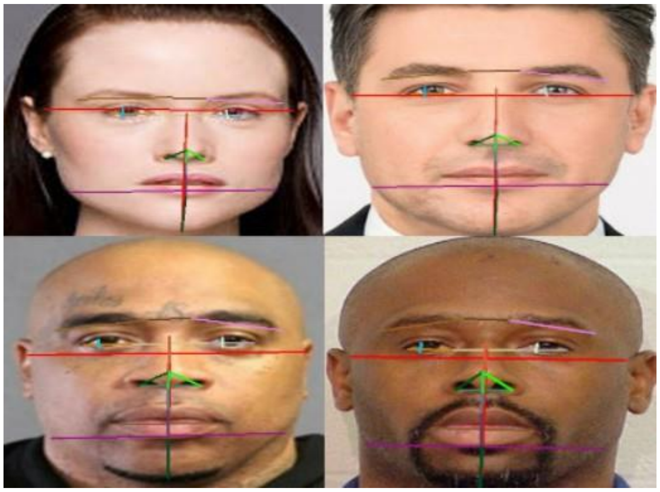

Fig. 2. Feature Extraction using Euclidean Distances in 2D Facial Images [1].

C. Siamese Neural Networks for Face Verification

Siamese networks have proven effective in scenarios where the goal is to compare two images and verify whether they belong to the same individual. These networks consist of twin subnetworks that learn to output similar embeddings for identical images and dissimilar embeddings for different ones. This technique is robust against changes in lighting, pose, and image quality, making it an excellent tool for facial verification in dynamic surveillance conditions. It is especially useful for recognizing partially occluded faces or in cases where subjects appear at different angles. [10]

D. Mask R-CNN for Person Detection and Segmentation

A deep learning model used for object detection and instance segmentation. In the context of surveillance, it is used to accurately detect and segment individuals from video frames. The segmentation mask allows the system to isolate a person from the background, improving the accuracy of subsequent facial or gait recognition. This technique is particularly effective in complex scenes where multiple individuals are present, ensuring the system focuses only on the relevant subjects. [14]

E. Bayesian Networks for Predictive Movement and Location Modeling

In dynamic environments like video surveillance, predicting an individual's next location is valuable for narrowing down search areas or tracking movements over time. Bayesian networks provide a probabilistic approach to modeling uncertainties and dependencies between different variables, making them highly suitable for predicting the movement or location of suspects or missing individuals. In a surveillance system,

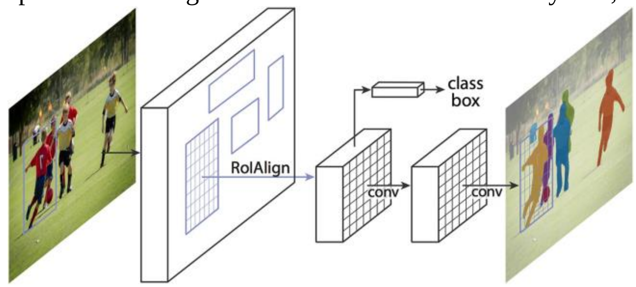

Fig. 3. Detection and Segmentation using Mask R-CNN [**?**].

Bayesian networks can process historical data, including previous movements, location context, and behavior patterns, to estimate where a person is likely to appear next. These predictions can then guide real-time video analysis, focusing resources on areas with higher probabilities of finding the person of interest. Bayesian networks also adapt as new information becomes available, refining predictions dynamically and improving the system's overall efficiency in large surveillance networks. [6]

F. Semantic Part Detection for Facial Recognition

One key challenge in facial recognition is accurately identifying individuals when parts of the face are obscured. Semantic part detection focuses on extracting and analyzing specific facial regions, such as the eyes, nose, mouth, and jawline, to improve recognition accuracy. By isolating the most informative parts of the face, models can achieve higher precision, even when other parts of the face are partially hidden by masks, glasses, or headwear. This approach is particularly useful in low-resolution footage or in cases where the face is not fully visible. Semantic part detection is also valuable in forensic settings, where a partial facial image might be the only available evidence. [5], [7]

G. Deep Learning-Based Image Quality Assessment

Surveillance footage often suffers from inconsistencies in quality, such as blur, low contrast, or noise, which can de- grade the performance of identification models. Image quality assessment using deep learning techniques helps to evaluate and filter out low-quality images that might affect recognition accuracy. These systems assess key quality indicators like sharpness, brightness, and contrast at the pixel level, ensuring that only high-quality images are processed for recognition. This pre-processing step enhances the reliability of identification models, reducing false positives and improving the overall robustness of surveillance systems, especially in environments where camera quality may vary or deteriorate over time. [7]

H. Generative Models for Synthetic Data Creation

International Journal on Emerging Research Areas (IJERA) *ISSN: 2230-9993*

DOI: 10.5281/zenodo.14650941

Generative models, particularly Generative Adversarial Networks (GANs), are becoming increasingly important in training identification models for surveillance. These models generate synthetic facial or gait data under various conditions, such as different lighting, poses, and occlusions, augmenting the available training data. GANs can create realistic facial images that simulate the variations typically encountered in real-world surveillance, such as changes in angle or facial expressions. This helps to train more resilient models that can perform well even when real-world data is limited or biased. Additionally, generative models can simulate rare or extreme conditions, improving the model's robustness in identifying individuals in challenging situations. [3]

I. Multi-Task Learning for Combined Identification

Multi-task learning (MTL) is a deep learning paradigm where a single model is trained to perform multiple related tasks simultaneously, such as face and gait recognition. By sharing information between tasks, MTL improves generalization and reduces the need for task-specific data. In surveillance, MTL enables a single model to handle both facial and non-facial identification methods, switching between them as needed depending on the video quality or the visibility of the individual. This flexibility enhances the system's ability to identify individuals even when one method (e.g., facial recognition) fails due to occlusion or poor lighting. MTL also improves efficiency by enabling the system to process different identification tasks concurrently, making it suitable for real-time applications. [15]

II. Data Processing Steps

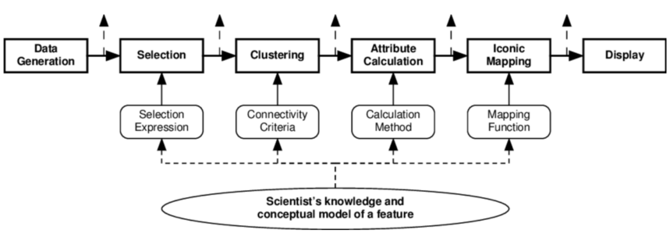

Fig. 4. Data Processing Pipeline.

A. Frame Extraction

Individual frames from video footage are to be extracted. Video is essentially a sequence of still images, and each frame can be analyzed independently. Extracting frames allows the system to handle the temporal aspects of video data while focusing on individual images for detection and identification. The frame rate may be adjusted depending on the nature of the video.

B. Object Detection and Segmentation

Detection and segmentation of individuals from the background using object detection techniques, such as Mask R-CNN. The model draws bounding boxes around detected people and segments them to isolate each person from the rest of the scene. This step is crucial for reducing noise in the data and ensuring that the model focuses only on relevant parts of the image (i.e., the person being tracked).

C. Face and Body Detection

After individual isolation, the system applies face detection algorithms to locate the face within the bounding box. For cases where the face is not visible, the system can apply gait detection to analyze walking patterns. This step helps differ- entiate between different individuals in the frame, especially when they are moving.

D. Landmark Detection

For facial recognition tasks, landmark detection is performed to identify key points on the face (e.g., eyes, nose, mouth). These landmarks are essential for building 2D or 3D models, which are used to improve recognition accuracy. Landmark detection allows the model to track the relative positions of facial features and ensure alignment, even when the face is partially obscured or rotated.

E. Data Augmentation

Augmentation techniques modify the original images to simulate different conditions that may be encountered in real-world surveillance, such as variations in lighting, resolution, and occlusion. This step generates additional training data by rotating, scaling, flipping, or altering the brightness of the images, which helps the models generalize better across diverse environments.

F. Image Quality Assessment

Poor-quality frames negatively affect the accuracy of the identification process. An image quality assessment step is applied to filter out blurry, noisy, or low-contrast images. Only frames that meet a certain quality threshold are passed to the next processing steps. This ensures that the machine learning models receive high-quality input, improving the reliability of the system.

G. Feature Extraction

For facial recognition, CNNs extract hierarchical features from the facial regions, such as eye shape, nose structure, and mouth contours. For gait recognition, features related to body movements, posture, and leg motion are extracted. The extracted features are then passed to classification models for identification.

III. Challenges and Future Directions

Despite the advancements, several challenges remain.

A. Challenges

1. Low-Quality Video Footage: Many surveillance cameras

DOI: 10.5281/zenodo.14650941

produce low-resolution video, which makes facial recognition difficult. Enhancing video quality or optimizing the system for low-resolution inputs will be crucial.

2. Occlusion and Crowded Scenes: Identifying individuals in crowds or when partially occluded poses a significant challenge. Future work will include developing more advanced segmentation techniques to isolate individuals more accurately.

3. Real-Time Processing: Achieving real-time processing speed without sacrificing accuracy is difficult, especially when running multiple layers of detection and identification. Optimization of the model to run efficiently on lower-powered hardware is a key future direction.

4. Data Privacy and Ethics: Implementing such a system in public places raises concerns about data privacy and ethics. Future work will involve addressing these concerns and ensuring that the system complies with local laws and regulations.

B. Future Research Opportunities

Future research could focus on developing lightweight models optimized for real-time video processing on edge devices and surveillance cameras. Reducing the computational load of facial and gait recognition models would enable quicker identification, making these systems more feasible for deployment in resource-constrained environments like public surveillance or mobile law enforcement units.

Another promising direction is the integration of multimodal data, combining visual inputs with other data streams such as audio, thermal imaging, or infrared. This approach could improve the accuracy of detection systems in low-light or occluded environments, where traditional video-based methods struggle.

Further exploration of hybrid models that combine facial and gait recognition, along with Bayesian networks for movement prediction, may enhance the system's robustness, especially in dynamic or crowded scenes. The use of ensemble learning techniques to fuse predictions from multiple models could also increase the system's reliability by reducing false positives and handling varying conditions.

In addition, research into real-time data collection paired with cloud-based processing could improve the scalability of these systems for larger surveillance networks. By offloading complex computations to the cloud, the system could process multiple video feeds simultaneously without overloading local devices, while maintaining responsiveness and speed.

Lastly, privacy-preserving techniques such as differential privacy and federated learning should be explored further. These methods would allow for secure, decentralized processing of surveillance data, ensuring that individual privacy is respected while still achieving accurate identification of missing or wanted individuals.

IV. CONCLUSION

The reviewed studies demonstrate various machine learning methodologies and techniques for detecting and aiding missing and wanted individuals in video footage. Techniques such as facial recognition, gait analysis, and Bayesian networks provide complementary solutions that, when integrated, offer a comprehensive approach to identifying individuals in dynamic environments. The reviewed methodologies lay the groundwork for developing more advanced systems, but challenges such as data quality, occlusion, and real-time processing must still be addressed.

REFERENCES

[1] G. Sanil, K. Prakash, S. Prabhu, V. C. Nayak, and S. Sengupta, "2D- 3D Facial Image Analysis for Identification of Facial Features Using Machine Learning Algorithms With Hyper-Parameter Optimization for Forensics Applications," 9 August 2023.

[2] X. Li, S. Lai, and X. Qian, "DBCFace: Towards Pure Convolutional Neural Network Face Detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 4, April 2022.

[3] C. Fu, X. Wu, Y. Hu, H. Huang, and R. He, "DVG-Face: Dual Variational Generation for Heterogeneous Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 6, June 2022.

[4] X. Wei, Y. Song, O. Mac Aodha, J. Wu, Y. Peng, J. Tang, J. Yang, and S. Belongie, "Fine-Grained Image Analysis With Deep Learning: A Survey," December 2022.

[5] J. Sun, S. Xia, Z. Sun, and S. Lu, "Cross-Model Deep Feature Fusion for Face Detection," September 2020.

[6] D. Reilly, M. Taylor, P. Fergus, C. Chalmers, and S. Thompson, "Misper-Bayes: A Bayesian Network Model for Missing Person Investigations," April 6, 2021.

[7] P. Terhörst, M. Huber, N. Damer, F. Kirchbuchner, K. Raja, and A. Kuijper, "Pixel-Level Face Image Quality Assessment for Explainable Face Recognition," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 5, no. 2, April 2023.

[8] W. Wang, X. Wang, W. Yang, and J. Liu, "Unsupervised Face Detection in the Dark," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, January 2023.

[9] M. Joseph and K. Elleithy, "Beyond Frontal Face Recognition," 22 March 2023.

[10] C. Song and S. Yong, "Face Recognition Method Based on Siamese Networks Under Non-Restricted Conditions," 20 April 2022.

[11] M. Jabberi, A. Wali, B. Neji, T. Beyrouthy, and A. M. Alimi, "Face ShapeNets for 3D Face Recognition," 17 May 2023.

[12] S. Yang, Y. Wen, L. He, and M. Zhou, "Sparse Common Feature Representation for Undersampled Face Recognition," *IEEE Internet of Things Journal*, vol. 8, no. 7, April 1, 2021.

[13] A. M. Saleh and T. Hamoud, "Analysis and Best Parameters Selection for Person Recognition Based on Gait Model Using CNN Algorithm and Image Augmentation," 2023.

[14] P. C. Neto, J. R. Pinto, F. Boutros, N. Damer, A. F. Sequeira, and J. S. Cardoso, "Beyond Masks: On the Generalization of Masked Face Recognition Models to Occluded Face Recognition," 22 August 2022.

[15] H.-B. Kim, N. Choi, H.-J. Kwon, and H. Kim, "Surveillance System for Real-Time High-Precision Recognition of Criminal Faces From Wild Videos," 9 June 2023.

[16] J. Zhao, S. Yan, and J. Feng, "Towards Age-Invariant Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 1, January 2022.

[17] B. Varghese, "Review Paper on Biometrics Authentication based on Liveness Detection Methods," *International Journal on Emerging Research Areas*, vol. 03, no. 01, 2023, doi: 10.5281/zenodo.8019348.