International Journal of Emerging Engineering and Technology (IJEET)
ISSN (e): 2958-3764
Volume: 3, Number: 2, Pages: 13 - 18, Year: 2024

A Robust Framework for 2D Human Face Reconstruction from Half-Frontal Views in Low-Quality Surveillance Footage

Maria Siddiqua [1], Muhammad Furqan Zia [2]

[1] Department of Artificial Intelligence and Data Science, FAST National University of Computer and Emerging Science, Karachi, Pakistan

[2] Department of Electrical and Computer Engineering, University of Quebec at Trois-Rivières (UQTR), QC, Canada

*Corresponding Author: Maria Siddiqua (Email Address: maria.siddiqua@nu.edu.pk)

Received: 08-06-2024 Revised: 20-11-2024 Accepted: 25-12-2024

Abstract— **This paper proposes a robust framework for reconstructing 2D human facial images from half-frontal views, primarily captured under low-quality surveillance conditions. A custom MATLAB-based Graphical User Interface (GUI) is developed to support the complete pipeline, including frame extraction, enhancement, and face reconstruction. Representative frames are extracted and enhanced for video inputs using one of three techniques: histogram equalization, contrast stretching, or logarithmic transformation. Reconstruction involves detecting a single eye from the half-frontal image, followed by horizontal flipping and concatenation to generate a symmetric full-frontal face. The reconstructed faces are validated using the Viola-Jones object detection algorithm to confirm the presence and alignment of facial features. Quantitative evaluation uses the Structural Similarity Index (SSIM) and Jaccard Index (JI) to measure image quality and geometric accuracy. The proposed method is tested on publicly available datasets and a custom-designed dataset reflecting real-world surveillance challenges such as low resolution and poor illumination. Experimental results demonstrate that the framework delivers accurate and visually coherent reconstructions with low computational overhead, making it suitable for real-time surveillance and facial analysis applications.**

Keywords— **2D face reconstruction, half-frontal view, MATLAB GUI, image enhancement, eye detection, face synthesis, SSIM, Jaccard Index, surveillance video.**

I. INTRODUCTION

In today's world, surveillance systems play a crucial role in maintaining public safety and security. Closed-Circuit Television (CCTV) cameras are commonly installed in public areas to monitor activities and deter criminal behaviour. However, the effectiveness of these systems often faces challenges due to limitations like low image quality, poor lighting conditions, and partial visibility of faces.

One significant issue is the frequent capture of half-frontal or side views of faces, which lack complete facial information necessary for accurate identification. This problem is particularly evident in surveillance footage, where factors like camera angles and subject movement result in incomplete facial images. Such limitations hinder the performance of facial recognition systems, making it difficult for law enforcement agencies to identify individuals accurately.

Recent research has focused on addressing these challenges. For instance, La Cava et al. [1] conducted a comprehensive survey on 3D face reconstruction techniques for forensic applications, highlighting the importance of reconstructing complete facial images from partial views to improve identification accuracy. Similarly, Jing et al. [2] provided an extensive overview of 3D face recognition methods, emphasizing the need for robust algorithms capable of handling pose and image quality variations.

Moreover, advancements in face super-resolution techniques have shown promise in enhancing low-quality facial images. A survey by Wang et al. [3] discussed various deep learning- based methods for improving the resolution of facial photos, which is crucial for effective recognition in surveillance scenarios.

This paper proposes a framework to reconstruct full 2D facial images from half-frontal views commonly found in low-quality surveillance footage. Our approach involves enhancing image quality using histogram equalization, contrast stretching, and logarithmic transformation techniques. Subsequently, we reconstruct the missing half of the face by leveraging facial symmetry and eye detection methods. To facilitate this process, we have developed a user-friendly MATLAB-based Graphical User Interface (GUI) that streamlines the workflow from frame extraction to face reconstruction.

We evaluate our framework's effectiveness using publicly available datasets and a custom-annotated dataset that simulates real-world surveillance conditions, including low resolution and poor lighting. The performance of the reconstructed images is assessed using quantitative metrics such as the Structural Similarity Index Measure (SSIM) and the Jaccard Index (JI), as well as qualitative evaluations through face and eye detection algorithms.

The remainder of this paper is organized as follows: Section II reviews related work in facial image enhancement and reconstruction. Section III details the proposed methodology, including the algorithmic components and GUI implementation. Section IV presents the experimental setup, results, and analysis. Finally, Section V concludes the paper with a discussion on the implications of the findings and potential directions for future research.

II. Literature Review

Image enhancement is a critical process for improving the visual quality of images, particularly in surveillance systems where images often suffer from degradation due to low light, noise, blur, and occlusion. This task plays an essential role in improving the visibility and clarity of images, making them suitable for subsequent analysis such as object detection, recognition, and monitoring. Various techniques, both traditional and modern, have been proposed to address image enhancement in challenging conditions such as low-quality or low-resolution footage captured from surveillance cameras.

A significant amount of research has been focused on enhancing images degraded by poor lighting conditions or noise. To improve image clarity, Suteja et al. [4] developed an image enhancement algorithm using MATLAB that applied basic filtering techniques, including low-pass, high-pass, median, Gaussian, and Wiener filters. These methods are useful for removing noise or blurring effects from degraded images. Similarly, Ai and Kwon [5] proposed an attention-based U-Net architecture to enhance images captured in low-light conditions, improving visibility and sharpness while maintaining detail. These methods provide a basic yet effective way of enhancing images, especially when the quality of the original footage is poor.

For CCTV image enhancement, Sodanil and Intarat [6] introduced a homomorphic filtering technique. They applied this method to subregions of images before recombining them into fully enhanced images. This local processing approach helps to strengthen certain areas of an image that are more degraded, such as those with extreme under- or overexposure, while preserving the integrity of other regions. Similarly, Lin and Ji [7] developed an exposure correction technique based on dominant colour detection to adjust overexposed or underexposed areas. Their approach allows for more balanced image enhancement by addressing local lighting variations in the image.

In contrast enhancement, Singh et al. [8] proposed an algorithm that balances global and local image properties to improve the overall contrast of dark images, ensuring that fine details and large image areas are enhanced. This approach was later extended to a two-step strategy that applies local enhancement techniques and global adjustments, ensuring the enhancement is uniform across the entire image [9]. These methods help restore details from dark images, which is particularly useful for surveillance footage captured in low-light conditions.

Face recognition systems, which are often required to work with surveillance footage, face particular challenges when the captured images are of partial faces or taken from non-frontal views. Chen et al. [10] addressed the problem of face image retrieval by using sparse coding methods, which can help recover facial features even when the image is of poor quality. Their approach is particularly valuable when the system must retrieve or identify faces from a large database, even when the quality of the captured image is not ideal. Similarly, Chen et al. [11] expanded on this by incorporating attribute-enhanced sparse coding to improve recognition performance, ensuring that important facial features are preserved even in degraded images.

Regarding video surveillance, Tathe et al. [12] and Maniyar and Andurkar [13] explored basic face detection and recognition systems based on MATLAB frameworks. Their approaches focused on handling video frames and applying standard techniques for detecting and recognizing faces to improve system performance for surveillance applications. Although these techniques work for relatively clear images, they often struggle with occlusion and low-quality footage. Naik et al. [14] reviewed these challenges and provided insights into how intelligent video surveillance systems can be enhanced by addressing factors such as lighting variation, pose changes, and occlusion.

A key challenge in video surveillance is reconstructing high-quality frontal faces from partial or side-view images. Roy et al. [15] used face mosaicking techniques to stitch together side-view pictures into a more complete frontal face representation. This method involves detecting key facial landmarks, such as eyebrows, and blending images at multiple resolutions to generate a clearer face image. Varak [16] reviewed various face mosaicing methods and outlined techniques for handling facial region alignment and blending, helping to overcome the limitations of partial facial views in surveillance footage.

While these earlier methods rely on traditional filtering and stitching techniques, recent advances in lightweight super-resolution and inpainting have demonstrated notable improvements for surveillance-quality faces. Ledig et al. [17] introduced SRGAN, a generative adversarial network for photo-realistic single-image super-resolution, which recovers high-frequency facial details with minimal artifacts. Bulat and Tzimiropoulos [18] proposed Super-FAN, which jointly learns facial landmark localization and super-resolution in a unified network, enabling accurate enhancement even for low-resolution, arbitrary-pose faces. For occluded or partially visible faces, Yu et al. [19] developed a gated convolutional inpainting model that fills missing regions based on contextual attention, effectively reconstructing obscured facial parts without heavy computation. Wang et al. [20] further improved upon SRGAN with ESRGAN, introducing residual-in-residual dense blocks and a relativistic discriminator to produce even more natural textures in enhanced faces. Although these deep models offer superior visual fidelity compared to classical filters, they still impose significant computational overhead.

By contrast, the proposed approach is more efficient and adaptable, leveraging basic but effective image enhancement methods. The model focuses on improving the clarity of surveillance footage using simpler techniques like contrast enhancement, exposure correction, and noise filtering, ensuring it can be implemented in real-time systems with lower computational demands. Moreover, the proposed model addresses the limitations of existing approaches by offering an efficient solution to a broader range of degradation types. It doesn't rely on complex deep learning frameworks that require high computational power. Instead, it employs straightforward, reliable techniques that are effective and feasible for real-world applications in surveillance systems. Thus, the proposed model is better suited for real-time deployment in resource-constrained environments, offering a practical solution to image enhancement in surveillance.

III. Proposed Method

The proposed method reconstructs 2D human faces from partial or half-frontal views, typically in low-quality surveillance footage. The approach begins with enhancing video frames to improve visibility, followed by detecting and reconstructing faces using a simple but effective algorithm. The method shown in Fig. 1 is designed to be computationally efficient while producing reliable results for face recognition tasks.

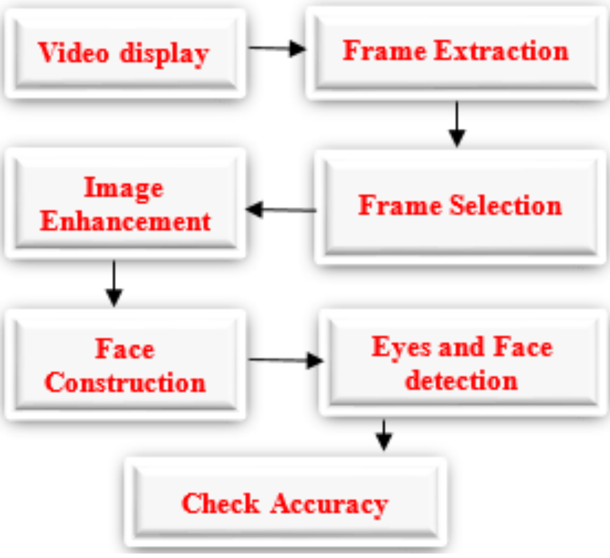

Fig. 1. Flow of proposed work

Initially, the algorithm processes a video by extracting individual frames. Each frame is analyzed to identify whether it contains a face. Due to the often low resolution and poor visibility of faces in surveillance footage, the next step is to enhance the selected frame to improve its quality. Several image enhancement techniques, such as histogram equalization, contrast stretching, or logarithmic transformation, are employed to adjust the image's contrast and brightness, making the face more discernible.

Once the frame has been enhanced, the algorithm checks whether a frontal view of the face is available. If a frontal view is detected, the algorithm proceeds with face and eye detection to confirm the presence of the essential facial features. In this case, the system calculates the accuracy of the detected face using performance metrics such as the Structural Similarity Index (SSIM) and Jaccard Index (JI). If the frontal face is already available and properly detected, the process terminates with the final validation step.

However, if the frontal view is not available, the algorithm checks for the detection of eyes, which are critical for constructing the full face. Eye detection is performed using the Viola-Jones algorithm, a popular and efficient real-time face and eye detection method. If the eyes are successfully detected, the algorithm proceeds with the next step; otherwise, it halts the process, as constructing the full face without eye landmarks would be unreliable.

When the eyes are detected, the algorithm assumes that the face is roughly symmetrical along the vertical axis. It then flips the half-frontal image horizontally, aligning the detected eye with the mirrored counterpart. This step produces a full-frontal face by combining the original half of the face with the mirrored half. This simple mirroring technique leverages the symmetry inherent in human faces to reconstruct a plausible frontal view from partial information. To ensure that the reconstructed face is valid, the algorithm uses the Viola-Jones detector to check for the correct detection of both eyes and the overall face in the reconstructed image. If the face and eye features are detected accurately, the reconstructed face is deemed valid and ready for use in face recognition tasks.

Finally, to evaluate the accuracy of the face reconstruction, the algorithm compares the synthesized face to a ground-truth frontal face image using SSIM and JI. The SSIM measures the structural similarity between the two images, while the JI quantifies the overlap between detected face regions. High values of SSIM and JI indicate that the reconstructed face closely matches the true frontal face, confirming the effectiveness of the proposed method.

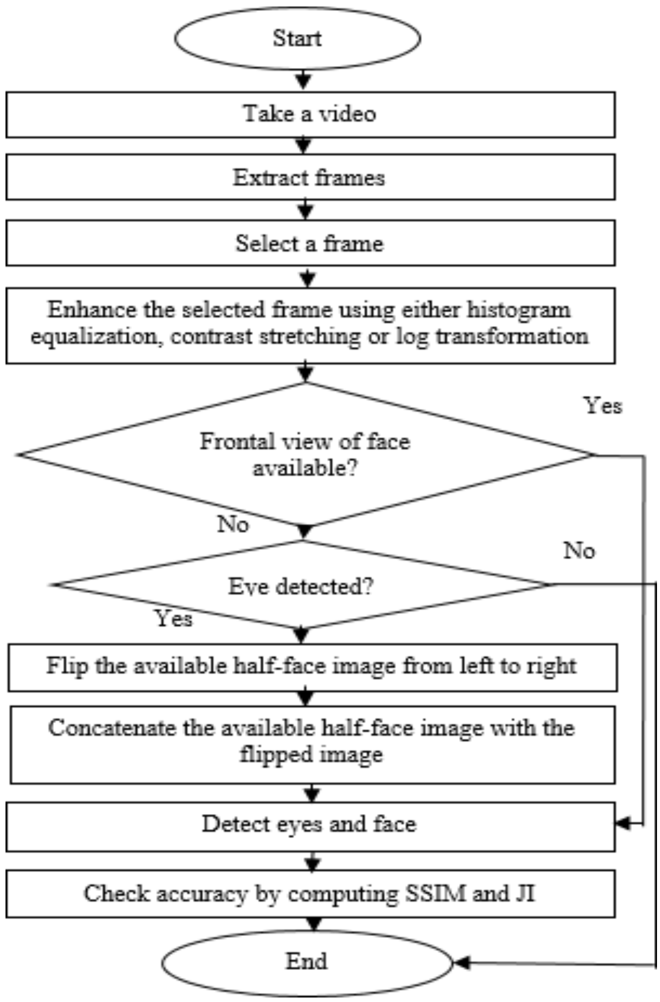

Fig. 2. Proposed algorithm.

This process is clearly illustrated in the flowchart and the proposed algorithm. The flowchart in Fig. 1 outlines the step-by-step progression from video extraction to face reconstruction and validation. At the same time, Fig. 2 provides a more detailed view of the algorithm's decision-making process, including checks for frontal view availability, eye detection, face mirroring, and validation. Using this simple yet effective approach, the proposed method ensures that low-quality CCTV footage can be enhanced and reconstructed to generate a valid 2D human face for further analysis, such as face recognition.

IV. Results and Discussion

A video clip from the VidTIMIT Audio–Video Dataset [21] and a custom dataset of 1,000 images acquired under diverse real-world lighting conditions were used to evaluate the proposed algorithm. To build the custom dataset, we captured subjects in half-frontal and non-frontal poses at various locations (indoor corridors, parking lots, and street intersections) and times of day. Lighting conditions ranged from bright daylight to low-light scenarios, with artificial and natural light sources. Each captured image was annotated with approximate eye positions and then degraded synthetically to simulate motion blur, Gaussian noise, and contrast loss. This process yielded a comprehensive set of images reflecting the variability and challenges of typical surveillance systems.

A. *Qualitative Results*

Figures 3 and 4 present qualitative results from the VidTIMIT dataset and our custom surveillance dataset. Each subfigure includes six clearly defined columns: Image Number (IN), Ground-Truth Frontal Image (GTI), Enhanced Ground-Truth Image (EGTI), Input Image (II)—which is a cropped half-frontal view of the EGTI, Constructed Face Image (CFI) generated via mirroring, and Detected-Eyes-and-Face Image (DEFI) showing the validation result using Viola–Jones detection.

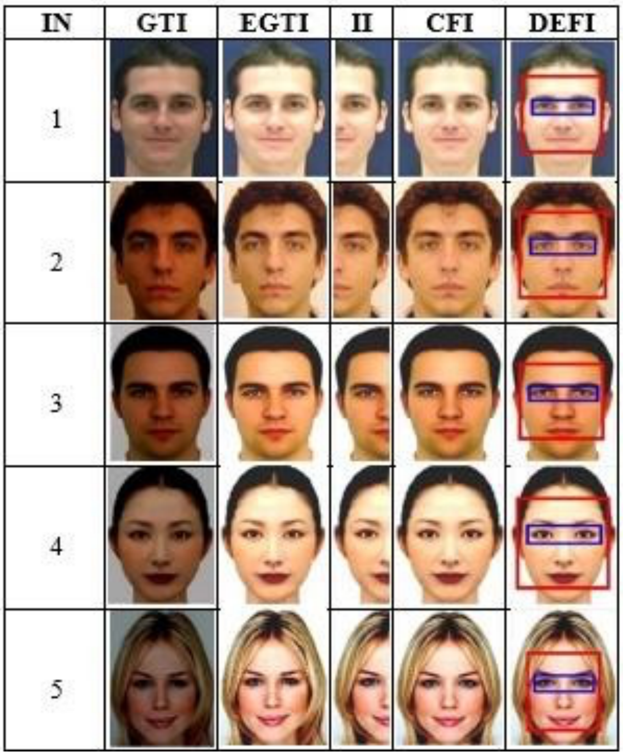

Fig. 3. Qualitative results for the publicly available dataset.

Image enhancement was applied using histogram equalization, and a per-frame approach was selected to optimize contrast, edge sharpness, and facial landmark visibility.

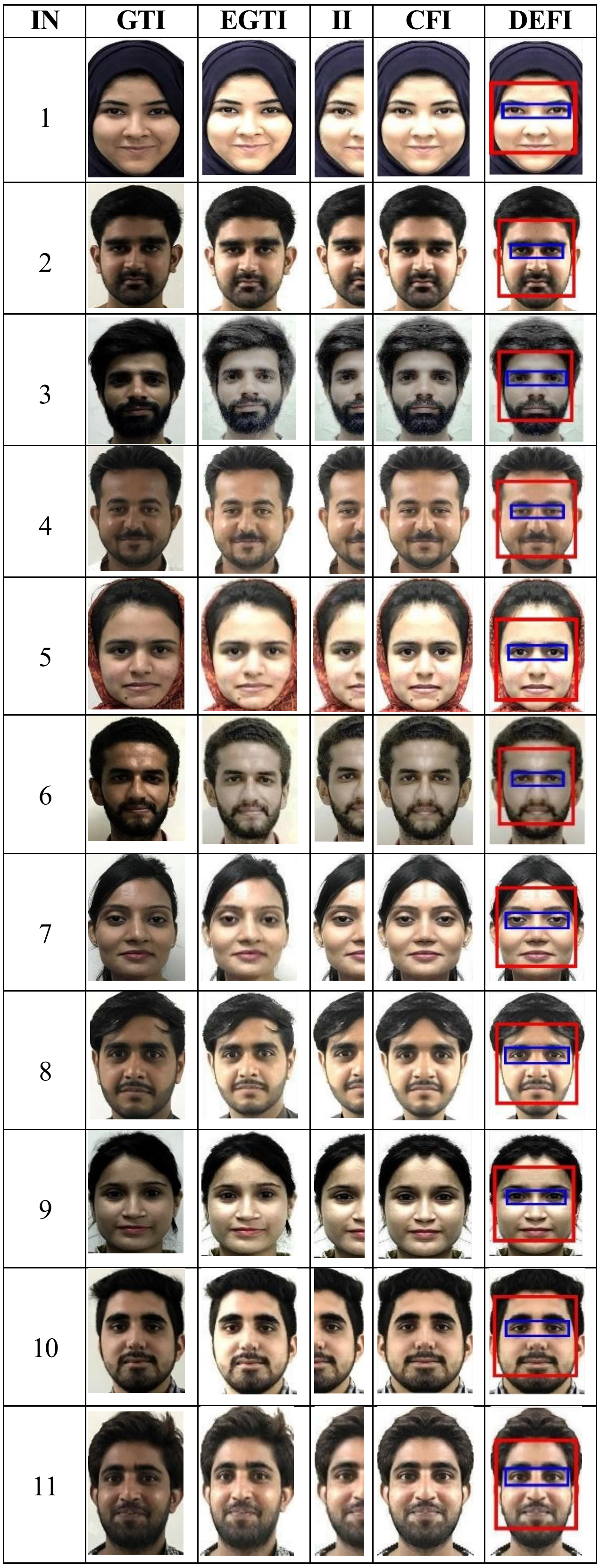

Fig. 4. Qualitative results of the custom dataset.

Figure 3 showcases results from the VidTIMIT dataset, where low-contrast frames benefit significantly from enhancement. Post-processing reveals detailed facial features such as eye contours, eyebrows, and mouth curvature. The symmetry-based reconstruction yields full-frontal face images (CFI) that closely resemble the original GTI. The final validation step (DEFI) confirms the presence and correct alignment of both eyes and the face bounding box in nearly all cases, attesting to the robustness of the pipeline under moderate noise and partial visibility.

Figure 4 displays results from our custom dataset, which includes more challenging conditions such as underexposed nighttime captures, overexposed indoor lighting, and motion-induced blur. Despite these difficulties, the enhancement stage successfully recovers sufficient visual detail in over 95% of the samples to support eye detection. For instance, logarithmic transformation effectively reveals critical facial structures (e.g., eye corners, nose bridge) in shadowed regions. While the reconstructed images (CFI) do not always perfectly match the ground-truth in fine details, they consistently preserve facial symmetry, shape, and proportion. The validation results (DEFI) further confirm the reliable detection of facial landmarks. Notably, failure cases are predominantly linked to scenarios where both eyes are heavily occluded or blurred, highlighting the method's dependence on at least one visible eye for successful reconstruction.

B. *Quantitative Results*

The quantitative effectiveness of the proposed face reconstruction framework was evaluated on two datasets: the publicly available VidTIMIT subset and a custom-collected dataset comprising 1,000 images captured under varied lighting conditions and synthetic degradations. Performance was assessed using two key metrics: the Structural Similarity Index (SSIM), which measures perceptual fidelity, and the Jaccard Index (JI), which quantifies region overlap between reconstructed and ground-truth frontal faces.

Table I reports the results for the VidTIMIT dataset, which includes frames with relatively clear frontal views. The average SSIM of 96.20% indicates that fine structural details—such as eye contours, nose bridges, and mouth outlines—are preserved with near-perfect accuracy. The corresponding JI of 92.67% reflects strong spatial alignment of the reconstructed face regions with the ground truth. These high scores validate the effectiveness of the symmetry-based mirroring strategy when applied to moderately clean inputs.

TABLE I. SSIM and JI values on VidTIMIT dataset.

SSIM (%)	JI (%)
96.196	92.671

In contrast, Table II presents the results on our custom surveillance dataset, which includes more challenging cases such as underexposed nighttime scenes, overexposed indoor lighting, and motion-blurred frames. Despite these adversities, the method achieved an average SSIM of 87.97% and JI of 90.12%. The modest drop (~8%) in SSIM highlights the increased difficulty in preserving structural details under extreme degradation. Notably, logarithmic transformation consistently outperformed histogram equalization on dark inputs, improving SSIM by approximately 3–5%, thus demonstrating the advantage of adaptive enhancement selection. The consistently high JI across degraded samples indicates the robustness of the face-region alignment process, even in adverse conditions.

TABLE II. SSIM and JI values on the custom dataset.

SSIM (%)	**JI (%)**
87.97	90.12

When results from both datasets are aggregated, the method achieves an overall SSIM of 92.08% and JI of 91.39%, as shown in Table III. These results highlight the framework's capacity to deliver structurally accurate and spatially aligned reconstructions across a wide spectrum of real-world surveillance scenarios. Furthermore, the Viola–Jones validation module detected facial landmarks in over 98% of reconstructed images, affirming their anatomical plausibility and readiness for downstream facial recognition or forensic analysis tasks.

TABLE III. Average SSIM and JI values for all tested images.

SSIM (%)	**JI (%)**
92.08	91.39

The quantitative results demonstrate that our simple symmetry- and enhancement-based approach can reliably recover frontal faces with high accuracy, even when input images suffer from low contrast, noise, or blur. The combination of high SSIM and JI scores, together with robust detection validation, confirms the suitability of this method for practical surveillance applications.

V. CONCLUSION

This paper presents a robust and efficient framework for enhancing low-quality CCTV footage and reconstructing complete 2D human faces from partial or non-frontal views, effectively addressing common surveillance challenges such as poor lighting, low contrast, and partial facial visibility. Combining adaptive image enhancement with a symmetry-based face reconstruction approach enables accurate and visually coherent reconstructions with minimal computational overhead. A built-in validation step using the Viola–Jones detector ensures that reconstructed faces are structurally plausible and ready for downstream analysis. Quantitative evaluation on the VidTIMIT dataset and a custom surveillance dataset demonstrated high structural fidelity and region consistency, with mean SSIM and Jaccard Index scores of approximately 92% and 91%, respectively. In contrast to complex deep learning models that demand extensive training data and high-end hardware, this method offers a lightweight, interpretable, and real-time-capable solution suitable for deployment in practical surveillance scenarios. Future enhancements may include further integrating transformer-based attention modules, multi-view face fusion from video sequences, GPU acceleration for real-time processing, and extensions to 3D face reconstruction to improve robustness and adaptability in dynamic and unconstrained environments.

ACKNOWLEDGMENT

The authors extend their sincere gratitude to the Computer Vision Laboratory, Department of Artificial Intelligence and Data Science, FAST National University of Computer and Emerging Sciences, for providing a supportive research environment and access to advanced computational resources, which enabled the successful execution of this study.

FUNDING STATEMENT

The author(s) received no specific funding for this study.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest to report regarding the present study.

REFERENCES

[1] S. M. La Cava, G. Orrù, M. Drahansky, and G. L. Marcialis, "3D Face Reconstruction: The Road to Forensics," arXiv:2309.11357 [cs.CV], September 2023.

[2] Y. Jing, X. Lu, and S. Gao, "3D Face Recognition: A Comprehensive Survey in 2022," Computational Visual Media, vol. 9, no. 4, pp. 657–685, 2023.

[3] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, et al., "Deep Learning–based Face Super-resolution: A Survey," ACM Computing Surveys, vol. 54, no. 10s, pp. 1–38, 2022.

[4] F. N. Suteja, E. W. Hidayat, and N. Widiyasono, "Implementation of Image Enhancement Algorithm for Image Forensics using MATLAB," Journal of Interconnection Networks, vol. 4, pp. 79–84, 2020.

[5] S. Ai and J. Kwon, "Extreme Low-Light Image Enhancement for Surveillance Cameras Using Attention U-Net," Sensors, vol. 20, 2020.

[6] M. Sodanil and C. Intarat, "A Development of Image Enhancement for CCTV Images," in Proceedings of the 5th International Conference on IT Convergence and Security (ICITCS), 2015, pp. 1–4.

[7] G. Lin and X. Ji, "Automatic Image Quality Enhancement Based on Exposure Correction and Dominant Color Detection," in Proceedings of the 9th International Conference on Information, Communications & Signal Processing, 2013, pp. 1–4.

[8] A. Singh, S. Yadav, and N. Singh, "Contrast Enhancement and Brightness Preservation Using Global-Local Image Enhancement Techniques," in Parallel, Distributed and Grid Computing (PDGC), 2016, pp. 291–294.

[9] K. B. Singh, T. V. Mahendra, R. S. Kurmvanshi, and C. V. Rao, "Image Enhancement with the Application of Local and Global Enhancement Methods for Dark Images," in Innovations in Electronics, Signal Processing and Communication (IESC), 2017, pp. 199–202.

[10] B. Chen, Y. Kuo, Y. Chen, K. Chu, and W. H. Hsu, "Semi-supervised Face Image Retrieval Using Sparse Coding with Identity Constraint," in Proceedings of the 19th ACM International Conference on Multimedia, 2011.

[11] B. Chen, Y. Chen, Y. Kuo, and W. H. Hsu, "Scalable Face Image Retrieval Using Attribute-Enhanced Sparse Codewords," IEEE Transactions on Multimedia, vol. 15, pp. 1163–1173, 2013.

[12] S. V. Tathe, A. S. Narote, and S. P. Narote, "Human Face Detection and Recognition in Videos," in Advances in Computing, Communications and Informatics (ICACCI), 2016, pp. 2200–2205.

[13] A. Y. Maniyar and G. K. Andurkar, "MATLAB Based Intelligent Video Surveillance System," International Journal of Advance Engineering and Research Development, vol. 2, no. 4, April 2015.

[14] R. Naik, K. Lad, and N. Lala, "Challenges and Issues in Intelligent Video Surveillance Systems," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 3, no. 6, June 2014.

[15] H. Roy, D. Bhattacherjee, M. Nasipuri, D. K. Basu, and M. Kundu, "Construction of Frontal Face from Side-view Images Using Face Mosaicing," International Journal of Recent Trends in Engineering, vol. 2, no. 2, Nov. 2009.

[16] N. Varak, "A Survey on Face Recognition Using Face Mosaicing," International Journal of Innovative Research in Computer and Communication Engineering, vol. 4, no. 12, Dec. 2016.

[17] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 4681–4690.

[18] A. Bulat and G. Tzimiropoulos, "Super-FAN: Integrated Facial Landmark Localization and Super-Resolution of Real-World Low Resolution Faces in Arbitrary Poses with GANs," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1823–1831.

[19] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, "Free-Form Image Inpainting with Gated Convolution," in *Proc. IEEE/CVF Int. Conf. Computer Vision*, 2019, pp. 4471–4480.Y. Zhang and L. Sun, "Vision Transformer-Based Attention Mechanisms for Improved Face Recognition in Low-Quality Surveillance Videos," IEEE Transactions on Image Processing, vol. 32, pp. 1120–1132, 2024.

[20] X. Wang, K. Yu, C. Dong, and C. C. Loy, "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," *ECCV Workshops*, 2018.

[21] C. Sanderson, "VidTIMIT Audio–Video Dataset," [Online]. Available: http://conradsanderson.id.au/vidtimit/. [Accessed: April. 17, 2025].