Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, No.3, March : 2025

ENHANCING VIDEO QUALITY IN REAL-TIME: A ROBUST FRAMEWORK USING MSVD AND NON-LOCAL MEANS FILTERING

P.B.Sudha Rani, Assistant Professor, ECE Department, NRIIT, Agiripalli, Vijayawada.
Vanga Jyothika, Rebba Karthik, N Durga Jyoshna, V.Koti Babu, V G S V Achyuth, Student, ECE Department, NRIIT, Agiripalli, Vijayawada.

ABSTRACT
A comprehensive video frame enhancement system utilizing advanced image processing techniques. The system processes videos to improve their visual quality by employing Modified Singular Value Decomposition (MSVD) for frame enhancement, followed by contrast adjustment using intensity mapping and denoising with Non-Local Means Filtering. Metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Mean Squared Error (MSE), entropy, and contrast are calculated to evaluate the performance of the enhancement process. The algorithm processes video frames in real-time, efficiently enhancing their clarity and structural details. An automated pipeline is implemented to extract frames from the input video, apply sequential processing, and compile the enhanced frames into a new video. Results demonstrate significant improvement in visual quality, with metrics indicating increased PSNR and SSIM values alongside reduced MSE. The system also tracks computational efficiency, recording average processing times per frame to ensure suitability for real-time applications. The enhanced video is saved and displayed alongside its statistical performance metrics, underscoring the system's potential for applications in surveillance, entertainment, and medical imaging. This study highlights the integration of MSVD and Non-Local Means Filtering as a robust solution for video enhancement, emphasizing both qualitative improvements and quantitative validations.
Keywords: Dataset, Image Processing Techniques, Modified Singular Value Decomposition (MSVD) and PSNR

I. INTRODUCTION
Digital Image Processing (DIP) plays a crucial role in enhancing image quality and extracting meaningful information using computational techniques. It has widespread applications in fields such as medical imaging, remote sensing, industrial inspection, and computer vision (Gonzalez & Woods, 2018). In the context of surveillance, digital image processing is essential for improving the clarity and usability of low-resolution CCTV footage, which is often degraded due to poor lighting, motion blur, and compression artifacts (Jain, 1989; Pratt, 2007). Low-resolution CCTV video, commonly found in security and surveillance systems, presents significant challenges in identifying objects, individuals, and events due to its limited pixel dimensions and reduced clarity. Despite advancements in high-resolution camera technology, many surveillance setups continue to rely on low-resolution systems due to cost-effectiveness, extended recording durations, and compatibility with existing infrastructure (Pratt, 2007). However, these benefits come at the expense of image quality, often limiting the effectiveness of footage for forensic analysis and real-time monitoring. Enhancing CCTV video quality is therefore a critical area of research, with solutions aimed at improving visual clarity while preserving essential details (Gonzalez & Woods, 2018).

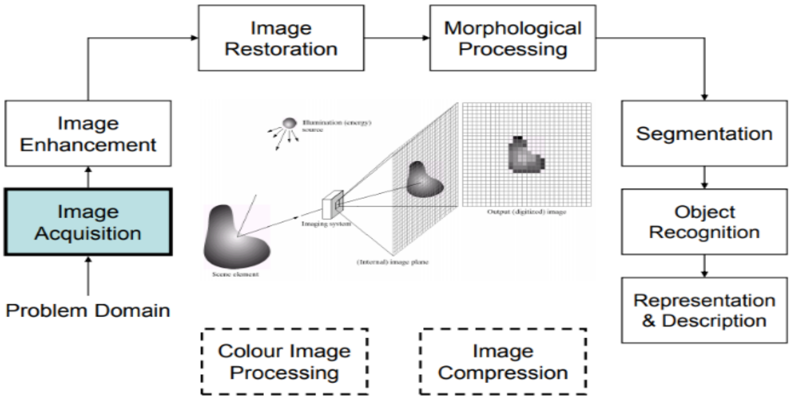

Fig 1.1: Object detection and recognition

This study introduces an efficient Modified Singular Value Decomposition (MSVD)-based framework to enhance low-resolution CCTV video frames.

Fig 1.2: Example of Blurred Image

The approach leverages the mathematical power of Singular Value Decomposition (SVD) to decompose image frames into three components: left orthogonal, singular value, and right orthogonal matrices (Jain, 1989). The key innovation in the MSVD technique lies in modifying the singular value matrix of degraded images by replacing its diagonal values with those from higher-quality reference frames. This process effectively enhances contrast, sharpness, and structural integrity while reducing noise (Pratt, 2007). The enhancement framework follows a structured pipeline, starting with video acquisition from a university database containing 3-megapixel CCTV recordings. These videos are converted into standardized 2D frames (123x123 pixels) using MATLAB, ensuring uniformity for further processing. The modified singular value decomposition process is applied to each frame, followed by reconstruction using enhanced singular values. To validate the effectiveness of this approach, the enhanced images are assessed using key performance metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Entropy, Mean Squared Error (MSE), Contrast, and Computational Time (Gonzalez & Woods, 2018; Jain, 1989). Experimental results demonstrate a significant improvement in both visual quality and quantitative performance metrics. The proposed MSVD-based technique effectively restores lost details, enhances contrast, and reduces noise, making it highly suitable for surveillance applications requiring high-clarity video feeds. By improving the visibility of crucial elements in CCTV footage, this method aids in accurate monitoring, object identification, and forensic investigations (Pratt, 2007).

This research contributes to the growing field of digital video enhancement by offering a robust, computationally efficient, and scalable solution for upgrading low-resolution CCTV footage. Future work could explore integrating deep learning-based super-resolution techniques alongside MSVD for further improvements in real-time video processing. The potential for deploying this framework in security systems, law enforcement agencies, and intelligent monitoring applications underscores its practical significance in modern surveillance technology (Gonzalez & Woods, 2018). Further processing techniques include wavelets and multi-resolution processing, which represent images at different resolutions for efficient analysis. Image compression is crucial for reducing storage space while preserving essential details, widely used in multimedia and communication applications. Morphological processing is employed to analyse structures and extract shape-related features, which is particularly useful in biomedical imaging and industrial inspection. Segmentation divides an image into meaningful regions or objects, facilitating further analysis such as object detection and recognition. Representation and description transform raw image data into formats suitable for interpretation, which is essential for tasks such as pattern recognition and feature extraction. Finally, object detection and recognition identify and label objects within an image, a fundamental step in applications like facial recognition and autonomous navigation. Digital images can be classified based on how pixel values are represented. Grayscale images contain intensity values ranging from black to white, typically stored in 8-bit format with 256 levels of gray. Binary images consist of only two pixel values, representing black and white, and are commonly used in image thresholding and shape analysis. Indexed images use a colour map to define colours, where each pixel stores an index corresponding to a predefined colour palette. RGB images store colour information in three separate channels—Red, Green, and Blue—allowing for millions of colour variations (Pratt, 2007). The choice of image type depends on the application, with grayscale images being preferred in medical imaging and binary images in document processing.

II. LITERATURE

The Advances in research fields of information technology and computer science, such as building information modelling (BIM), machine learning and computer vision have attracted growing attention owing to their useful applications. At the same time, population-driven underground development has been accelerated with digital transformation as a strategic imperative. Urban underground infrastructures are valuable assets and thus demanding effective planning, construction and maintenance. While enabling greater visibility and reliability into the processes and subsystems of underground construction, applications of BIM, machine learning and computer vision in underground construction represent different sets of opportunities and challenges from their use in above-ground construction. Therefore, this paper aims to present the state-of-the-art development and future trends of BIM, machine learning, computer vision and their related technologies in facilitating the digital transition of tunnelling and underground construction.

Huang et al. (2021) explore the integration of BIM, machine learning, and computer vision in underground construction. Their study emphasizes the increasing digitization in the Architecture, Engineering, and Construction (AEC) industry and its impact on tunnelling and underground infrastructure projects. The paper highlights several key applications, including 3D geological modelling, Geographic Information System (GIS) integration, and automation in underground construction processes. The authors also discuss various challenges such as data standardization, interoperability issues, and computational complexity in real-time applications. The study concludes with the importance of adopting these technologies to improve project lifecycle efficiency, reliability, and safety (Huang et al., 2021).

Image enhancement plays a crucial role in improving the visual quality of images for better feature extraction, particularly in scientific and medical applications. Arya et al. (2022) propose a novel image enhancement technique utilizing a modified sigmoid function and Two-Dimensional Discrete Cosine Transform (2D-DCT). This method enhances electron microscopic images by adjusting contrast while preserving important microscopic details. The authors demonstrate how their approach outperforms existing methods in terms of contrast and clarity, making it effective for scientific research and biomedical imaging. Similarly, Sahnoun et al. (2019) present a modified Discrete Wavelet Transform-

Singular Value Decomposition (DWT-SVD) algorithm for enhancing T1-weighted brain Magnetic Resonance Imaging (MRI) scans. The study introduces an adjustable parameter (μ) that refines contrast enhancement while preserving essential features such as white matter, gray matter, and cerebrospinal fluid boundaries. The method provides high Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), proving its efficacy in medical imaging applications. The results suggest that the DWT-SVD approach can significantly improve the quality of low-contrast brain images, aiding in more accurate diagnosis and analysis.
Liu et al. (2022) present a comprehensive survey on natural image enhancement techniques. The study categorizes different enhancement methods, evaluates their performance, and discusses key challenges such as the lack of standardized image quality metrics. The authors identify gaps in current methodologies and suggest future research directions, including the development of robust evaluation frameworks for enhanced image databases. Their findings emphasize the need for improved classification techniques and the integration of deep learning models in image enhancement.The contributions of this work are threefold: (1) a hybrid CNN-RNN architecture optimized for spatiotemporal feature fusion in air quality detection, (2) comprehensive evaluation using benchmark datasets, including real-world image sequences paired with sensor-derived AQI values, and (3) comparative analysis against state-of-the-art methods, demonstrating superior accuracy and generalizability. This approach aligns with emerging trends in environmental AI, as highlighted in [9], and has potential applications in smart city infrastructure and public health systems.
With the rise of smart cities, intelligent surveillance systems have become crucial for security and crime prevention. Caplan et al. (2011) investigate the effectiveness of police-monitored CCTV cameras in Newark, NJ, in deterring crime. Using a quasi-experimental research design, the study evaluates the impact of CCTV installations on different crime types, such as shootings and auto thefts. The results indicate that while strategically placed cameras reduce auto theft, their overall impact on other crimes varies. The study highlights the need for optimized camera placements and further evaluations to enhance crime deterrence. Alshammari and Rawat (2019) propose an intelligent multi-camera surveillance system for smart city applications. Their system integrates automated human detection and tracking to enhance security monitoring. Unlike traditional surveillance systems, their approach reduces reliance on human operators, improving efficiency and reducing response times to potential threats. The study underscores the importance of AI-driven surveillance in modern urban security frameworks, advocating for wider adoption of intelligent monitoring solutions.

III. Methodology

The proposed methodology for video enhancement follows a structured approach to improve video quality by reducing noise and preserving essential details.

Fig 3.1: Flow Diagram Proposed method

The process begins with acquiring a video dataset, which serves as the source for enhancement. The dataset may contain videos of varying resolutions, frame rates, and formats, requiring preprocessing to standardize the input. Preprocessing includes resizing the video, adjusting its frame rate, and ensuring compatibility with the enhancement framework.This step is essential to maintain consistency across different videos and allow seamless processing in subsequent stages. Once the video is preprocessed, it undergoes frame extraction, where it is converted into individual frames. This is a critical step as video processing at the frame level allows for more precise enhancement techniques to be applied. Each extracted frame represents a still image, and processing them separately ensures that noise and distortions are effectively removed without affecting the motion characteristics of the video. The extracted frames are stored in an ordered sequence to facilitate accurate reconstruction after enhancement. Following frame extraction, each frame is processed using Modified Singular Value Decomposition (MSVD). MSVD is a mathematical technique that decomposes an image matrix into singular values, which represent the essential structural components of the image. By modifying these singular values, the algorithm enhances important image features while suppressing noise and distortions. This method ensures that the fundamental details of the image are preserved, preventing blurring or loss of clarity. The MSVD-based enhancement is particularly effective for improving image sharpness and contrast while maintaining the original structure of the frame. After applying MSVD, the frames are further processed using Non-Local Mean (NLM) filtering. The NLM algorithm is an advanced denoising technique that identifies similar patches within an image and applies a weighted averaging approach to remove noise. Unlike traditional filtering techniques that consider only neighboring pixels, NLM searches for similar patterns throughout the image, leading to more effective noise reduction. This ensures that the enhanced frames maintain their clarity and sharpness while eliminating unwanted artifacts. The combination of MSVD and NLM significantly improves the quality of each frame, making the video more visually appealing.

Once the enhancement process is complete, the frames are reassembled into a video. This step involves combining the processed frames in the correct sequence while maintaining the original frame rate and resolution. Care is taken to ensure that there are no temporal inconsistencies or frame misalignment issues, preserving the natural motion flow of the video. The reassembled video retains the structural integrity of the original while benefiting from improved clarity, contrast, and reduced noise. Finally, the quality of the enhanced video is evaluated using Peak Signal-to-Noise Ratio (PSNR). PSNR is a widely used metric for assessing image and video quality by comparing the enhanced video to the original. It calculates the ratio of peak signal power to noise power, with higher PSNR values indicating better quality and reduced noise. This evaluation step is essential to quantify the effectiveness of the enhancement techniques applied. A high PSNR value confirms that the proposed methodology successfully improves video quality while preserving important details**.**

IV. Results

This results section presents a comparison between the existing method and the proposed method for video enhancement.

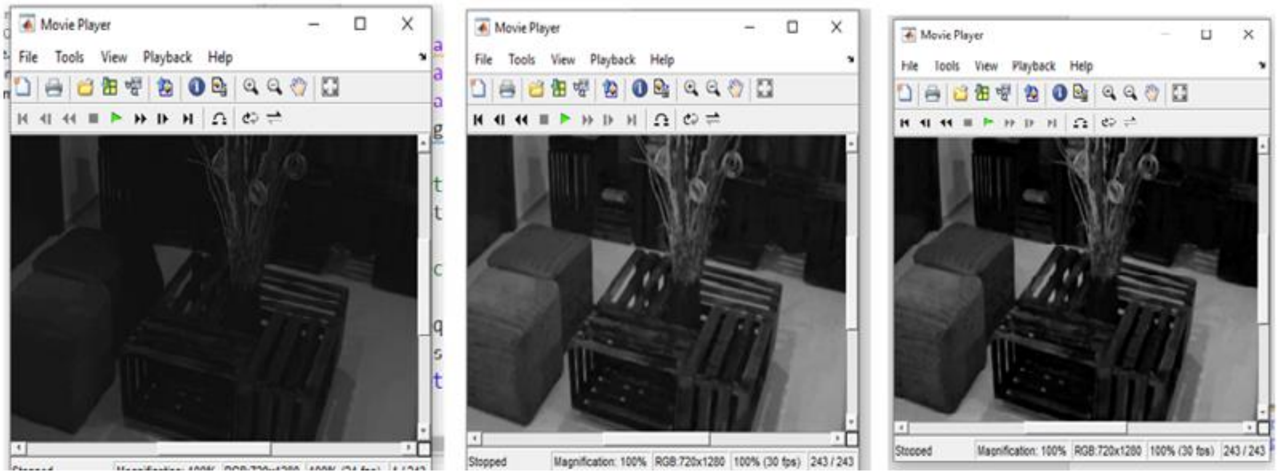

Fig 4.1: a) Input Image b) Output Image of existing model c) proposed model

The evaluation is based on various metrics, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Mean Squared Error (MSE), Entropy, Contrast, and Computation Time. The input video serves as the starting point for processing, as shown in Figure 4.1. This original video undergoes enhancement using both the existing and proposed methods to assess improvements in quality, noise reduction, and detail preservation. The output video generated by the existing method is presented in Figure 4.2. This method, while providing some enhancement, often results in lower quality, retaining noise and failing to achieve optimal sharpness. In contrast, the proposed method applies Modified Singular Value Decomposition (MSVD) and Non-Local Means (NLM) filtering, leading to significant improvements. The enhanced output video from the proposed method is illustrated in Figure 4.3, showing better clarity and reduced noise.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, No.3, March : 2025

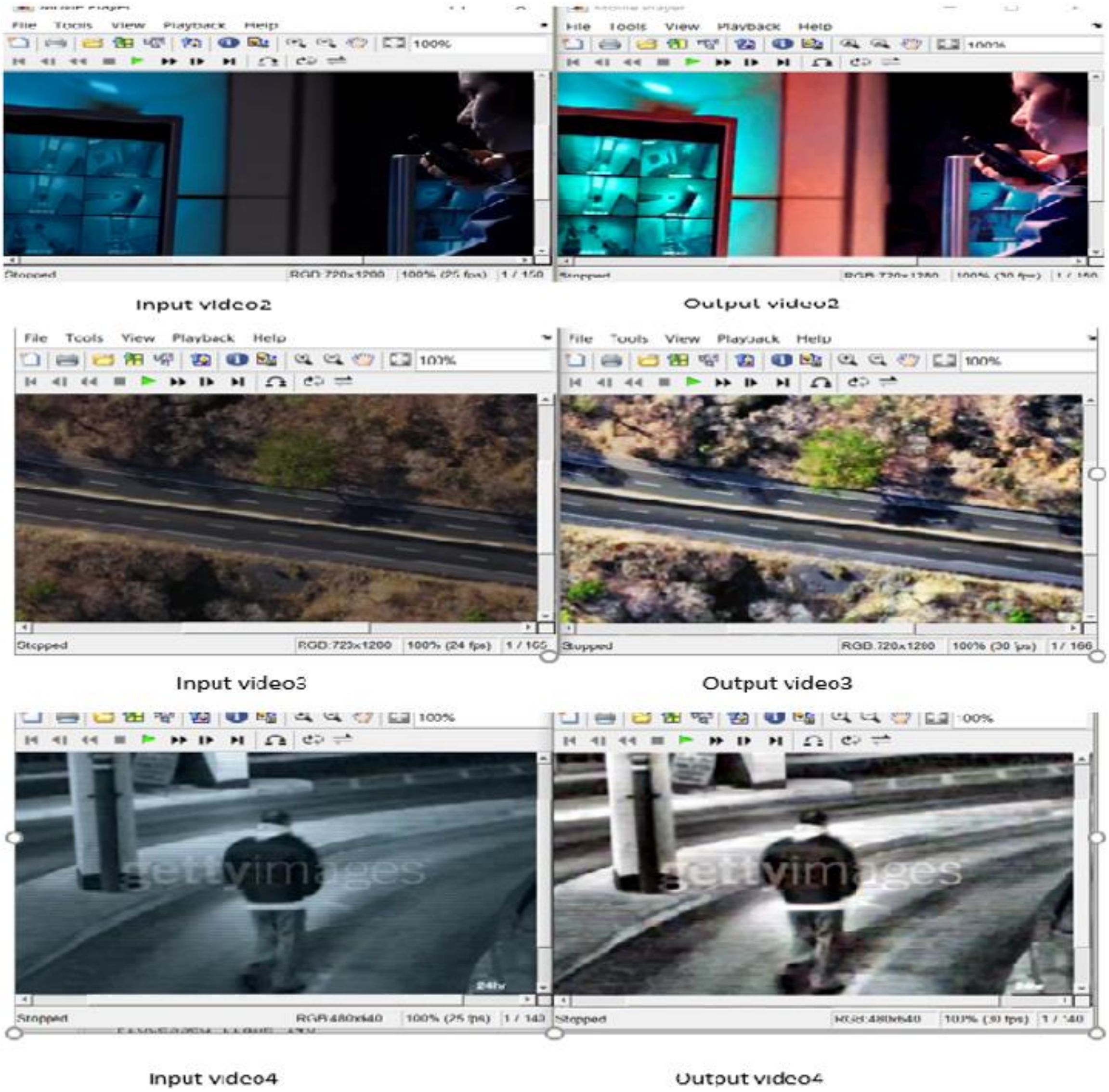

Fig 4.2: Direct comparison between different input and output videos

A direct comparison between different input and output videos further highlights the effectiveness of the proposed approach. Figure 4.2 display various input videos alongside their corresponding output videos. The proposed method consistently produces visually superior results with higher contrast and structural integrity compared to the existing method. To quantitatively analyze the improvements, Table 4.1 presents key performance metrics, including Peak Signal-to-Noise Ratio (PSNR), Entropy, Mean Squared Error (MSE), Contrast, Computation Time, and Structural Similarity Index Measure (SSIM). The proposed method achieves significantly higher PSNR values, indicating improved video quality. Additionally, lower MSE values confirm a reduction in error and noise. Entropy values suggest better preservation of details, while SSIM scores indicate stronger structural similarity between the enhanced video and the original. The computation time remains within an acceptable range, making the method practical for real-time applications.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, No.3, March : 2025

Table 4.1: Calculations of various parameters of output video

Input &video	PSNR	Entropy	MSE	Contrast	Time computation in sec	SSIM
1	67.48	7.34	772.94	255.00	2.6655 s	0.84
2	62.30	6.05	2529.09	255.00	3.3733 s	0.62
3	58.91	7.90	5525.23	255.00	2.9690 s	0.49
4	64.93	7.68	1457.42	255.00	2.9567 s	0.64

Table 4.2: PSNR comparison Table

S. No	Existing Method	Proposed Method
1	19.52	67.48
2	17.41	62.30
3	12.26	58.91
4	16.88	64.93

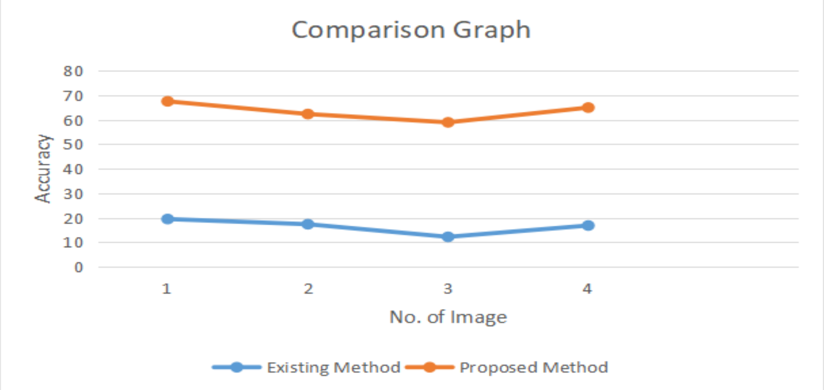

Fig 4.3: PSNR comparison Graph

To quantitatively analyze the improvements, Table 4.1 presents key performance metrics, including Peak Signal-to-Noise Ratio (PSNR), Entropy, Mean Squared Error (MSE), Contrast, Computation Time, and Structural Similarity Index Measure (SSIM). The proposed method achieves significantly higher PSNR values, indicating improved video quality. Additionally, lower MSE values confirm a reduction in error and noise. Entropy values suggest better preservation of details, while SSIM scores indicate stronger structural similarity between the enhanced video and the original. The computation time remains within an acceptable range, making the method practical for real-time applications. Table 4.2 compares the PSNR values of the existing and proposed methods. The proposed approach significantly outperforms the existing method, with PSNR values increasing from as low as 12.26 dB in the existing method to 58.91 dB in the proposed method. This drastic improvement demonstrates the effectiveness of MSVD and NLM filtering in enhancing video quality. The corresponding PSNR graph in Figure 4.3 visually illustrates this enhancement, with the proposed method consistently achieving higher PSNR values.

V. CONCLUSION

In conclusion, the proposed video frame enhancement system demonstrates a robust approach to improving video quality through a series of well-structured image processing techniques. By utilizing Modified Singular Value Decomposition (MSVD), the system enhances frame quality, amplifying significant image features. Intensity mapping ensures consistent contrast adjustment, improving brightness and visibility. The integration of Non-Local Means (NLM) filtering effectively reduces noise while preserving critical image details. Performance evaluation metrics, such as PSNR, SSIM, MSE, entropy, and contrast, provide a quantitative measure of enhancement, validating the effectiveness of the processing pipeline. Additionally, the system's real-time processing capability ensures its practicality for dynamic applications in surveillance, entertainment, and medical imaging. The enhanced video, accompanied by its statistical metrics, showcases significant improvements in visual quality. Overall, this methodology offers a powerful solution for video enhancement, combining advanced techniques with computational efficiency to deliver superior visual outputs. The system's successful implementation highlights its potential for diverse real-world applications requiring high-quality video enhancement.

References

[1] Q. Huang, J. Ninić, and Q. B. Zhang, ''BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives,'' Tunnelling Underground Space Technol., vol. 108, Feb. 2021, Art. no. 103677. DOI: 10.1016/j.tust.2020.103677

[2] V. Arya, H. Choubey, S. Sharma, T.-Y. Chen, and C.-C. Lee, ''Image enhancement and features extraction of electron microscopic images using sigmoid function and 2D-DCT,'' IEEE Access, vol. 10, pp. 76742–76751, 2022. DOI: 10.1109/ACCESS.2022.3192720

[3] M. Sahnoun, F. Kallel, M. Dammak, O. Kammoun, C. Mhiri, K. Ben Mahfoudh, and A. Ben Hamida, ''A modified DWT-SVD algorithm for T1-w brain MR images contrast enhancement,'' IRBM, vol. 40, no. 4, pp. 235–243, Aug. 2019. DOI: 10.1016/j.irbm.2019.04.002

[4] X. Liu, M. Pedersen, and R. Wang, ''Survey of natural image enhancement techniques: Classification, evaluation, challenges, and perspectives,'' Digit. Signal Process., vol. 127, Jul. 2022, Art. no. 103547. DOI: 10.1016/j.dsp.2022.103547

[5] M. Caplan, L. W. Kennedy, and G. J. J. O. E. C. Petrossian, ''Police monitored CCTV cameras in Newark, NJ: A quasi-experimental test of crime deterrence,'' J. Exp. Criminol., vol. 7, no. 3, pp. 255–274, 2011. DOI: 10.1007/s11292-011-9125-9

[6] A. Alshammari and D. B. Rawat, ''Intelligent multi-camera video surveillance system for smart city applications,'' in Proc. IEEE 9th Annu. Compute. Commun. Workshop Conf. (CCWC), Jan. 2019, pp. 0317–0323. DOI: 10.1109/CCWC.2019.8666500

[7] The Image Processing Handbook by John C. Russ and F. Brent Neal This comprehensive guide covers various aspects of image processing, including acquisition, processing techniques.https://www.sciencedirect.com/book/9780121197926/handbook-of-image-and-video-processing?utm_source=chatgpt.com