As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 03, pages 293-297.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue03.html
DOI: https://doi.org/10.5281/zenodo.19014331

Smart Motion Based Video Enhancement for Low Light Environments using Multi-Modality

N.M.Ramalingeswararao[1], A.Chandu[2], K.Sai Durga Aravind[2], D.Murali[2], Y.Shanmukhasairam[2]

Department of ECE, Godavari Global University, Rajamahendravaram, INDIA.
Department of ECE, Godavari Institute of Engineering & Technology (A), Rajamahendravaram, INDIA.

To Cite this Article
N.M.Ramalingeswararao, A.Chandu, K.Sai Durga Aravind, D.Murali & Y.Shanmukhasairam (2026). Smart Motion Based Video Enhancement for Low Light Environments using Multi-Modality, 12(03), 293-297. https://doi.org/10.5281/zenodo.19014331

Article Info
Received: 06 February 2026; Revised: 03 March 2026; Accepted: 08 March 2026.

KEYWORDS

Low-light video enhancement, motion-based processing, pseudo-infrared representation, temporal consistency, noise suppression, adaptive brightness, real-time video processing

ABSTRACT

Videos captured in low-light conditions often exhibit poor visibility caused by inadequate illumination, sensor noise, low contrast, and motion-related artifacts. Such degradations significantly affect both visual perception and the performance of vision-based systems operating in dark environments. This paper proposes a smart motion-based video enhancement framework that follows a multimodal strategy tailored for low-light scenarios. The method combines RGB video frames with a pseudo-infrared (pseudo-IR) illumination representation generated from the same input sequence. The pseudo-IR component offers robust global illumination cues for brightness enhancement, while the RGB modality helps retain natural color information and fine structural details. To improve temporal stability, motion information between consecutive frames is utilized to enforce consistency and suppress flickering effects commonly observed in frame-by-frame enhancement approaches. In addition, illumination-guided reflectance recovery, adaptive brightness control, and a two-stage noise suppression strategy are applied to avoid noise amplification in dark regions. Experimental results indicate that the proposed framework produces visually clearer and temporally stable videos compared to existing methods, while maintaining low computational complexity suitable for real-time applications such as surveillance and autonomous navigation in low-light environments.

INTRODUCTION

Videos recorded under poor lighting conditions typically exhibit degraded visual quality due to weak illumination, sensor noise, low contrast, and motion-related artifacts. These challenges make it difficult to perceive important scene details and can significantly impact the performance of vision-based systems used in applications such as night-time surveillance, traffic monitoring, and autonomous navigation [11]. Despite extensive research in low-light enhancement, achieving visually stable and reliable results in dynamic scenes remains a challenging problem.

Many existing approaches enhance low-light videos by processing each frame independently. While such methods can improve brightness and contrast, they often introduce temporal artifacts such as flickering and inconsistent illumination across frames, particularly in the presence of motion [1], [7]. In practical scenarios, even small frame-to-frame inconsistencies can noticeably degrade viewing quality. As a result, recent studies have highlighted the importance of exploiting temporal information to ensure smooth and consistent enhancement across video sequences.

Illumination modeling plays a critical role in low-light enhancement. Retinex-based methods decompose an image into illumination and reflectance components, enabling controlled brightness improvement while preserving structural details [4], [10]. However, approaches relying solely on RGB information often struggle in extremely dark environments, where illumination cues are weak and noise becomes dominant [11]. To overcome these limitations, multimodal enhancement strategies have attracted increasing attention. By introducing additional illumination-related information, such as near-infrared cues, these methods provide stronger guidance for brightness restoration while maintaining natural color appearance [8], [14]. Since dedicated infrared sensors are not always available, generating pseudo-infrared representations directly from RGB frames has emerged as a practical and cost-effective alternative. Motivated by these observations, this work proposes a motion-aware multimodal video enhancement framework that combines RGB information with pseudo-infrared illumination guidance to improve visual quality while remaining computationally efficient.

LITERATURE SURVEY:

Zhang et al. [1] proposed a learning-based strategy that propagates enhancement information from earlier frames to subsequent frames to reduce temporal inconsistencies. While this approach improves visual smoothness, its performance depends strongly on the effectiveness of the underlying single-frame enhancement model and does not explicitly incorporate illumination modeling.

Lv et al. [2] introduced an unsupervised low-light video enhancement framework based on spatial–temporal co-attention transformers. Although the model effectively captures inter-frame dependencies, the complexity of the transformer architecture limits its suitability for real-time deployment.

Wu et al. [3] presented STARNet, which aggregates information from neighboring frames to enforce temporal consistency. Despite achieving stable brightness across frames, the method operates solely on RGB inputs and lacks explicit illumination guidance, which restricts its performance in severely dark scenarios.

Retinex-based approaches have also been widely investigated for low-light enhancement. Lee and Yang [4] employed illumination–reflectance decomposition to enhance brightness while preserving details. Xu et al. [5] further extended this concept by incorporating spatial–temporal consistency constraints, resulting in smoother video outputs, though their approach remains limited to RGB-based processing.

Deep learning models for temporal consistency have further advanced video enhancement. Zhu et al. [6] presented a spatial-temporal learning strategy to produce stable frame enhancements, and Rota et al. [7] applied recurrent neural networks to refine frames over time, reducing flickering. Both approaches, however, depend on the quality of initial frame enhancement.

Multimodal approaches bring in additional illumination data to guide brightness improvements. Niu et al. [8] used near-infrared data along with RGB frames to restore brightness more effectively but required specialized sensors. To avoid extra hardware, generating pseudo-infrared signals from RGB frames has become a practical alternative [8].

Other studies, such as Ye et al. [9], focus on propagating enhancement information across frames to preserve temporal coherence, though they do not emphasize illumination modeling. Li and Anantrasirichai [10] proposed an unsupervised Retinex-based method to handle varying lighting conditions but did not use multimodal fusion.

Surveying the literature shows that low-light video enhancement faces common challenges: noise amplification, flickering, and high computational costs

[11]. These challenges motivate the need for efficient, illumination-aware, and motion-consistent enhancement methods. The approach proposed in this work addresses these issues by combining pseudo-infrared guidance with motion-aware temporal processing, achieving stable and visually improved results without requiring additional sensors.

PROPOSED METHOD:

Enhancing videos captured in low-light conditions is challenging due to poor visibility, motion blur, noise amplification, and low contrast. To address these issues, we propose a smart motion-based video enhancement framework that integrates both RGB video frames and a pseudo-infrared (pseudo-IR) illumination representation. In addition, motion information between consecutive frames is explicitly considered to ensure smooth and temporally consistent enhancement.

The pseudo-IR representation is designed to capture the global illumination characteristics of the scene, which provides effective guidance for brightness enhancement without over-amplifying noise. At the same time, the RGB frames preserve natural color information and fine texture details. Unlike conventional frame-wise enhancement approaches that often lead to flickering and unstable brightness, the proposed method leverages motion-aware temporal information to propagate enhancement results across frames. This helps maintain visual consistency, particularly in regions containing moving objects.

To further improve visual quality, noise suppression is applied at both spatial and temporal levels. Spatial denoising reduces sensor noise within individual frames while preserving edges, whereas temporal denoising exploits redundancy across neighboring frames to suppress flickering noise. The overall framework is computationally efficient and does not require additional hardware sensors, making it suitable for real-time applications such as surveillance systems, autonomous vehicles, and intelligent transportation systems operating under low-light conditions.

METHODOLOGY:

Our proposed video enhancement framework enhances low-light videos in several carefully designed steps, combining RGB frames, pseudo-infrared illumination, motion information, and noise suppression.

1. Multi-Modal Decomposition:

Each RGB frame is split into two components:

i. Pseudo-infrared illumination (L) – captures the overall lighting in the frame. It is generated by converting the RGB frame to grayscale and applying a Gaussian smoothing filter:

$$L = G\sigma\ (Gray(I))$$

where, I is the input RGB frame, and $G\sigma$ is the Gaussian smoothing with standard deviation σ.

ii. Reflectance (R) – preserves details and textures of the scene. The RGB frame can be approximately reconstructed as:

$$I = R \cdot L$$

This decomposition allows us to enhance brightness without losing fine details or colors.

2. Motion-Aware Temporal Consistency:

Enhancing each frame separately often causes flickering. To prevent this, we consider the motion between consecutive frames using optical flow or frame difference techniques. The enhancement of the previous frame is propagated to the current frame according to motion:

$$It_{enhanced} = \alpha \ .\ It_{current} + (1 - \alpha)\ .\ It\text{-}1_{warped}$$

Here:

- Itcurrent is the currently enhanced frame,
- It-1warped is the previous frame adjusted according to motion,
- α ($0 < \alpha \leq 1$)controls the blending of current and previous frames.

This ensures smooth brightness transitions and consistent edges in moving objects.

3. Illumination Modeling and Reflectance Recovery:

The pseudo-IR illumination L guides how brightness is adjusted in each frame. We perform adaptive brightness control using a gamma-based adjustment:

$$L_{enhanced} = L\gamma$$

where γ is a parameter ($0 < \gamma \leq 1$) controlling the brightness scaling.

The final enhanced frame is reconstructed by combining the reflectance with the enhanced illumination:

$$I_{enhanced} = R\ .\ L_{enhanced}$$

This preserves textures and colors while improving overall brightness and contrast.

4. Noise Suppression:

Brightening low-light videos can make noise more noticeable, especially in dark areas. To fix this, our method reduces noise both within each frame and across consecutive frames, keeping edges and details sharp. This results in cleaner, smoother, and more stable video output, we apply a two-stage noise suppression:

i. Spatial denoising: smooths flat regions while preserving edges.

ii. Temporal denoising: uses information from neighboring frames to remove flickering noise.

The combination of these steps results in a cleaner, more stable video.

BLOCK DIAGRAM:

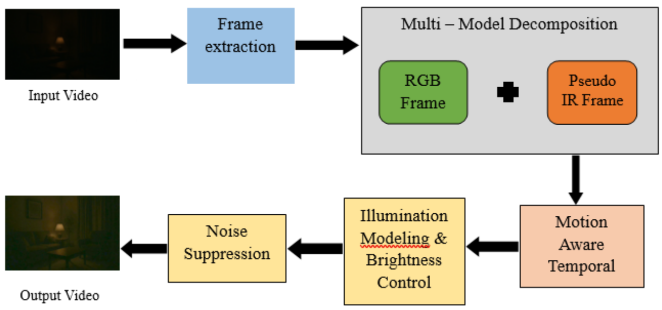

Fig 1: Block Diagram

PERFORMANCE COMPARISON :

Comparison of the performance of the proposed Multi-Modality enhancement with existing methods is presented below. This show the improvement in the proposed method PSNR and SSIM Values. The table is shown below:

Table 1: Performance Comparison Table

Methods	PSNR Value	SSIM Value
Zero-DCE Curve Estimation **[12]**	21.5 dB	0.78
Improved Retinex Model **[4]**	21.8 dB	0.79
EnlightenGAN **[13]**	22.2 dB	0.80
Temporal Consistency Network (TCN) **[1]**	22.5 dB	0.82
TempRetinex **[10]**	22.8 dB	0.83
Recurrent Temporal Network **[7]**	22.9 dB	0.83
ST Propagation & Reconstruction **[9]**	23.0 dB	0.84
ST Co-Attention Transformer **[2]**	23.1 dB	0.84
Multi-Modality (Proposed Method)	**23.45 dB**	**0.85**

GRAPH REPRESENTATION:

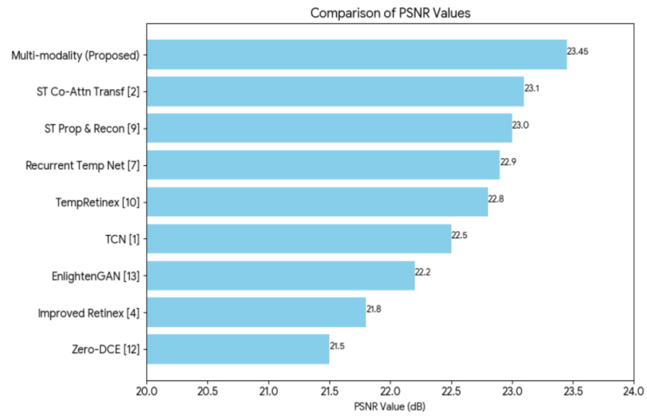

Fig 2: PSNR Comparison

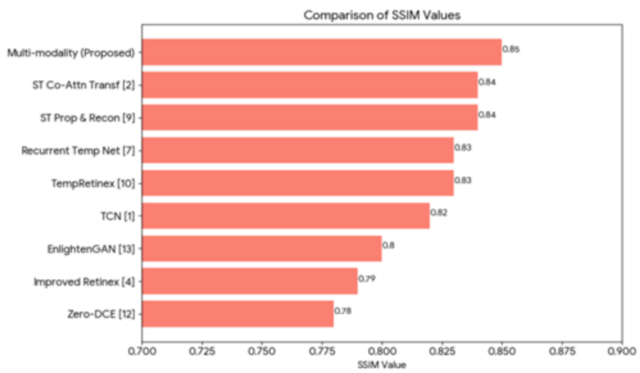

Fig 3: SSIM Comparison

CONCLUSION:

In this work, a smart motion-based video enhancement framework for low-light environments was presented using a multimodal approach. By integrating RGB information with a pseudo-infrared illumination representation and incorporating motion-aware temporal processing, the proposed method effectively enhances visual clarity while maintaining consistent brightness across video frames. The use of pseudo-infrared guidance enables controlled illumination enhancement, whereas motion-based propagation helps reduce flickering and preserve structural continuity in dynamic scenes.

Experimental evaluation shows that the proposed approach achieves improved quantitative and qualitative performance, obtaining a PSNR of 23.45 dB and an SSIM of 0.85, and outperforming several existing low-light video enhancement techniques. The framework is computationally efficient and does not require additional sensing hardware, making it suitable for practical deployment in real-world scenarios. Future work will focus on improving robustness under extremely low illumination and further optimizing the method for implementation on resource-constrained platforms.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Fan Zhang, Yang Li, Shuowei You, and Yejin (Y.) Fu, "Learning Temporal Consistency for Low-Light Video Enhancement from Single Images," in Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, Jun. 2021.https://openaccess.thecvf.com/content/CVPR2021/papers/Zhang_Learning_Temporal_Consistency_for_Low_Light_Video_Enhancement_From_Single_CVPR_2021_paper.pdf.

[2] X. Lv, S. Zhang, C. Wang, W. Zhang, H. Yao and Q. Huang, "Unsupervised Low-Light Video Enhancement With Spatial-Temporal Co-Attention Transformer," in IEEE Transactions on Image Processing, vol. 32, pp. 4701-4715, 2023.https://pubmed.ncbi.nlm.nih.gov/37549080/.

[3] Zhe Wu, Zehua Sheng, Xue Zhang, Si-Yuan Cao, Runmin Zhang, Beinan Yu, Chenghao Zhang, Bailin Yang, and Hui-Liang Shen, STARNet: Low-Light Video Enhancement Using Spatio-Temporal Consistency Aggregation, Pattern Recognition, Volume 160, April 2025. https://doi.org/10.1016/j.patcog.2024.111180.

[4] Sung-Ling Lee and Shih-Hsuan Yang, "Efficient Low-Light Video Enhancement Based on Improved Retinex Algorithms," in Proceedings of the 2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW 2023), Brisbane, Australia, Jul. 2023.

[5] Xiaogang Xu, Kun Zhou, Tao Hu, Jiafei Wu, Ruixing Wang, Hao Peng, and Bei Yu, "Low-Light Video Enhancement via Spatial-Temporal Consistent Decomposition," in the Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI 2025),2025.https://www.ijcai.org/proceedings/2025/238.

[6] Lingyu Zhu, Wenhan Yang, Baoliang Chen, Hanwei Zhu, Shiqi Wang, and Xiandong Meng, "Temporally Consistent Enhancement of Low-Light Videos via Spatial-Temporal Compatible Learning," International Journal of Computer Vision, vol. 132, pp. 4703–4723, 2024. https://link.springer.com/article/10.1007/s11263-024-02084-w.

[7] Claudio Rota, Marco Buzzelli, Simone Bianco, and Raimondo Schettini, "A Recurrent Neural Network for Temporal Consistency in Low-Light Videos Enhanced by Single-Frame Methods," vol. 31, pp. 2795–2799, 2024. https://ieeexplore.ieee.org/document/10709345/.

[8] Muyao Niu, Zhihang Zhong, and Yinqiang Zheng, "NIR-Assisted Video Enhancement via Unpaired 24-Hour Data," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp.10778–10788.

[9] Jing Ye, Changzhen Qiu, and Zhiyong Zhang, "Spatio-Temporal Propagation and Reconstruction for Low-Light Video Enhancement," Digital Signal Processing, vol. 139, 2023.https://www.sciencedirect.com/science/article/abs/pii/S1051200423001665.

[10] Yini Li and Nantheera Anantrasirichai, "TempRetinex: Retinex-Based Unsupervised Enhancement for Low-Light Video Under Diverse Lighting Conditions", arXiv preprint, 2025.https://arxiv.org/abs/2511.09609.

[11] Chongyi Li, Chunle Guo, Linghao Han, Jun Jiang, Ming-Ming Cheng, Jinwei Gu, and Chen Change Loy,"Low-Light Image and Video Enhancement Using Deep Learning: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 12, pp. 9396–9416, Dec. 2022. https://ieeexplore.ieee.org/document/9609683/.

[12] Guo, Chunle, Li, Chongyi, Guo, Jichang, Loy, Chen Change, Hou, Junhui, Kwong, Sam, and Cong, Runmin. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 1780–1789.

[13] Jiang, Yifan, Gong, Xinyu, Liu, Ding, Cheng, Yu, Fang, Chen, Shen, Xiaohui, Yang, Jianchao, Zhou, Pan, and Wang, Zhangyang. EnlightenGAN: Deep Light Enhancement Without Paired Supervision. IEEE Transactions on Image Processing, vol. 30, pp. 2340–2349,2021. EnlightenGAN: Deep Light Enhancement Without Paired Supervision | IEEE Journals&Magazine|IEEEXplore.

[14] Zhong, Mingchen, Lu, Xin, Li, Dong, Xu, Senyan, Jiang, Ruixuan, Fu, Xueyang, and Yin, Baocai. CompEvent: Complex-valued Event-RGB Fusion for Low-light Video Enhancement and Deblurring. arXiv preprint arXiv:2511.14469 (2025). https://arxiv.org/abs/2511.14469.