Nakshatra-Drishti: A Supervised Learning Approach for Low Light Image Enhancement Using Convolutional Neural Networks

Nitesh Kumar
Faculty of Communication Engineering,
Military College of Telecommunication Engineering, Mhow,
Madhya Pradesh, India
neo.machine.race@gmail.com

Shailendra
Faculty of Communication Engineering,
Military College of Telecommunication Engineering, Mhow,
Madhya Pradesh, India
vermashailendra403@gmail.com

***Abstract*— Images captured under low-light conditions pose significant challenges for subsequent analysis due to degradation in quality, including noise, loss of scene content, inaccurate colour, and contrast information. In this paper, we propose a supervised learning-based convolutional neural network (CNN) model, Nakshatra-Drishti, specifically designed for enhancing low-light images, videos, and real-time camera feeds. The model is trained on paired datasets and extensively evaluated on various benchmarks, demonstrating remarkable results. We also introduce a user-friendly web-based software application that enhances image perception in poorly illuminated environments, facilitating more effective artificial intelligence analysis and decision-making processes.**

Keywords*— Low-light image enhancement, Convolutional neural networks, Supervised learning, Image perception, Artificial intelligence, Decision-making.

I. Introduction

Low-light or poorly illuminated environments often lead to degraded images, compromising their aesthetic quality and impairing viewer experience and interpretation. Consequently, conducting high-level computer vision operations and artificial intelligence analyses, such as object detection, change detection, and face recognition, becomes exceptionally challenging. Consequently, low light image enhancement (LLIE) techniques have emerged as a widely recognized field of ongoing research, with continuous developments occurring annually.

A. Background and Motivation

Algorithms ranging from histogram equalization to retinex model-based conventional methods have been developed to enhance low-light images [1],[2]. However, since 2017, there has been a surge in deep learning-based LLIE model developments, exhibiting superior accuracy, reliability, robustness, and speed compared to traditional methods [3],[4],[5]. These advancements have paved the way for more resilient computer vision applications in challenging lighting conditions.

B. Overview of Proposed Method and Testing

In this study, we introduce a novel deep learning algorithm based on Convolutional Neural Networks (CNNs) for low-light image enhancement. Our proposed Nakshatra-Drishti model effectively addresses diverse lighting conditions, including cloudy or foggy environments, and images captured at night or under non-uniform lighting. This model excels in handling high-dimensional image data and learns hierarchical feature representations, enabling significant enhancement in image quality. Trained on paired image data from the LOL dataset using supervised learning, our model has been extensively tested on various datasets and real-world images captured by digital cameras. It demonstrates remarkable performance improvements, including increased brightness, enhanced contrast, and noise reduction, even in complex low-light scenarios [6],[7],[8],[9]. Moreover, our model exhibits real-time processing capabilities, making it suitable for practical applications requiring rapid image enhancement and analysis.

Fig. 1. Images taken under sub-optimal lighting conditions

II. Literature Review

A. Survey of Low Light Image Enhancement Techniques

Low-light image enhancement has been an active field of research within the domain of image processing. Various traditional methods have been employed, including adaptive histogram equalization (AHE) [2], Retinex [1], and the multiscale Retinex model [10]. Recently, in order to strike a balance between detail and naturalness, a low-light image enhancement algorithm was proposed for non-uniform illumination images [11], utilizing a bi-log transformation method. Building upon logarithmic transformation, a weighted variational model was developed to estimate both the reflectance and illumination from an observed image, incorporating regularization terms [12]. Another approach to low-light image enhancement was presented in [13], which involved estimating the illumination of each pixel by identifying the maximum value in its RGB channels, followed by constructing an illumination map based on structural priors. Furthermore, a joint low-light image enhancement and denoising model was introduced in [14] through decomposition in a successive image sequence. Additionally, in [15], a Retinex model was proposed, integrating a noise map and comparing it with the conventional Retinex model for low-light image enhancement.

B. Analysis of Deep Learning Approaches in LLIE

Since 2017, the focus of low-light image enhancement has shifted towards the development of various deep learning models. These models can be categorized into supervised learning, reinforcement learning, unsupervised learning, zero-shot learning, and semi-supervised learning, depending on the learning strategy employed. Studies have shown that supervised learning strategies have been particularly successful, accounting for approximately 73% of usage. Supervised learning approaches typically involve paired training data, where low-light images and corresponding daylight image pairs are used for model training.

LL-Net, a stacked auto-encoder model, was introduced to learn joint denoising and low-light enhancement on the patch level [3]. Retinex-Net combined the Retinex theory with deep learning, providing an end-to-end framework [4]. HDR-Net incorporated deep learning network models with bilateral grid processing and local affine color transforms, along with pairwise supervision [16]. In the HDR domain, multi-frame low-light enhancement techniques have been developed [17], [18], [19]. A recent method proposed a learning approach to enhance low-light images directly from raw sensor data, focusing on avoiding amplified artifacts [20].

CNN-based solutions, which rely on paired data for supervised training, have been resource-intensive. LLNet, for instance, was trained on data simulated with random Gamma correction. Additionally, Generative Adversarial Networks (GANs) have been developed for image synthesis, translation, restoration, and enhancement, using paired training data [21], [22]. Unsupervised GANs, such as Enlighten GAN, have been proposed for learning inter-domain mappings without paired training data, thereby enhancing low-light images using unpaired low/normal light data [23]. Another approach, Zero-DCE, formulates light enhancement as image-specific curve estimation, producing high-order curves for pixel-wise adjustment on the dynamic range of input images [24]. An accelerated version, Zero-DCE++, was subsequently introduced [25].

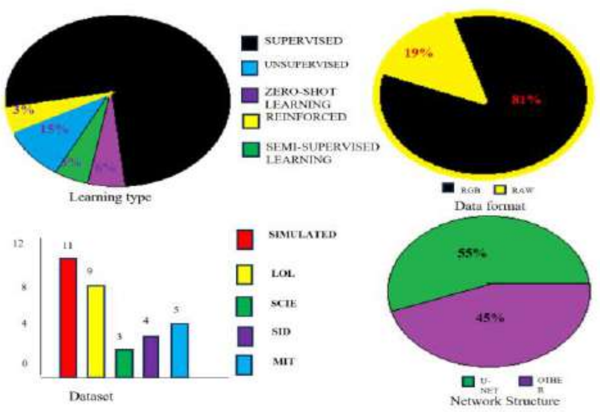

Fig. 2. Statistic analysis of deep learning-based LLIE Techniques

III. METHODOLOGY

A. Mathematical Modelling of Problem Statement

First give a common formulation of the deep learning based LLIE problem. For a low-light image

$I \in Y^{W \times H \times 3}$ of width W and height H, the process can be modelled as:

$$¥ = F(I; \phi)$$

where $¥ \in R^{W \times H \times 3}$ is the enhanced result and F represents the network with trainable parameters ϕ. The purpose of deep learning is to find optimal network parameters ø that minimizes the error.

$$ø = \operatorname{argmin} L(¥, Y),$$

The loss function L(¥,Y) drives the optimization of deep learning network. The loss functions used in our Nakshatra-Drishti model are Exposure loss, Colour constancy loss, Spatial constancy loss, illumination smoothness loss during training of deep learning model.

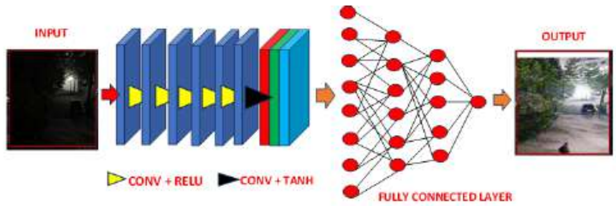

Fig. 3. Architecture Nakshatra Drishti deep learning Model

B. Architecture Design and Components

In the Architecture Design and Components section, we detail the structural layout and constituent elements of our Nakshatra-Drishti model, elucidating its network architecture, layers, and functional components. We highlight how these components synergistically contribute to the effective enhancement of low-light images.

1) Noise Filter: The denoise filter in our LLIE Model is a vital component designed to mitigate the adverse effects of noise commonly encountered in low-light images [28]. By employing a median-based denoising technique, it effectively suppresses Gaussian noise while preserving image details and edges. Additionally, it addresses other types of noise such as temporal noise, colour noise, and quantization noise, ensuring overall image quality enhancement.

2) Neighbouring Frame Utilization: Our Nakshatra-Drishti Model leverages information from neighbouring frames in video sequences to enhance performance and accelerate processing speed during video and real-time camera feed enhancement [29]. By exploiting temporal coherence, our algorithms better estimate scene characteristics and mitigate noise, ultimately resulting in improved image quality. This approach optimally utilizes computational resources, enabling real-time applications like surveillance and medical imaging.

3) Optimizer Function: Utilizing the Adam optimizer with a learning rate of 0.0001, our model refines parameters and minimizes loss to enhance accuracy [30]. By iteratively adjusting parameters based on the gradients of the loss function, Adam facilitates efficient convergence towards optimal solutions. The chosen learning rate balances exploration and exploitation, significantly improving the quality of enhanced images, videos, and real-time camera feeds.

4) Network Structure: Our Nakshatra-Drishti Model employs a U-Net type network structure, integrating multi-scale features and utilizing both low-level and high-level features [28]. This architecture is crucial for achieving optimized low-light enhancement by effectively integrating feature information.

5) Data Format: Trained using RGB data input images, our Nakshatra-Drishti Model caters to the prevalent RGB data format produced by various cameras. This approach ensures the recovery of clear details, vivid colour, noise reduction, and brightness enhancement in extremely low-light images, making it suitable for a wide range of applications.

In our Nakshatra-Drishti Model, we utilize a comprehensive set of loss functions:

- Exposure Loss: Adjusts image brightness by minimizing the difference between exposure levels of enhanced and ground truth images, enhancing visibility in low-light conditions.
- Colour Constancy Loss: Maintains colour consistency across images by penalizing deviations from true colours, ensuring realistic colour representation.
- Spatial Constancy Loss: Promotes spatial coherence and smoothness, reducing noise and blurriness for sharper, visually pleasing results.
- Illumination Loss: Enhances overall brightness and illumination quality, minimizing disparities between enhanced and desired illumination levels.
- Smoothness Loss: Encourages seamless transitions between pixels, reducing artifacts like noise and jagged edges for clearer, visually appealing images.

IV. Experimental Setup

A. Implementation and Training details

Our CNN-based Nakshatra-Drishti Model is trained using the LOL-dataset[26], incorporating paired low-light and over-exposed images for network training. We adopted a novel approach by developing a GUI-based integrated image enhancement, video enhancement, and live camera feed enhancement software application. To achieve this, we utilized a web-based framework created with the Flask library in Python, leveraging Visual Studio and Python for software development. During the execution of the deep learning model training module, Python codes are initiated, fetching the linked LOL-dataset. For model training, we employed an NVIDIA 2080Ti GPU with 512GB RAM, which completed training in approximately 40 minutes. We set the number of epochs to 100 and utilized Adam Optimizer with a learning rate of 0.0001 to adjust model parameters, minimize loss, and improve accuracy.

On a 12th Generation Intel® Core™ i3 processor CPU, the training time for the model is approximately 5 hours. Notably, our model demonstrates real-time processing capabilities, achieving a processing speed of approximately 500 frames per second for images sized 640×480×3 on GPU. The training/ work flow cycle is depicted in the Fig.4.

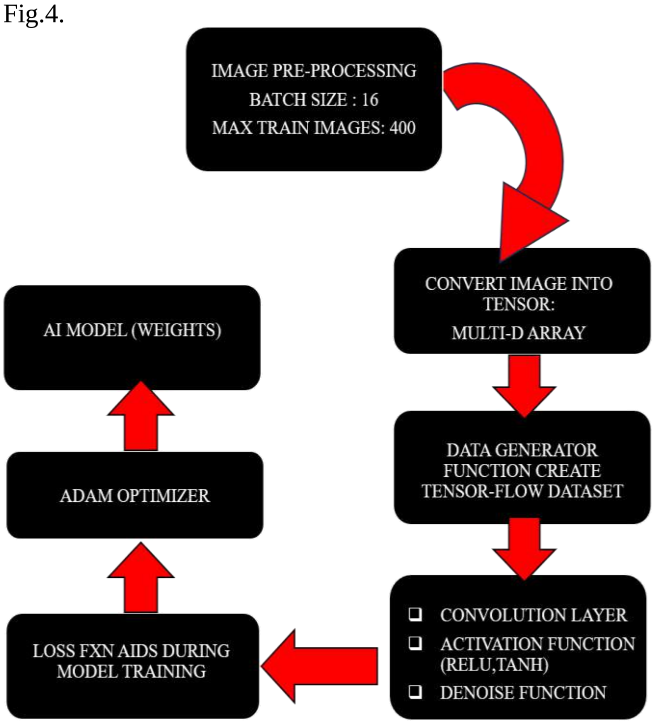

Fig. 4. Work Flow

B. Dataset Overview

For our Nakshatra-Drishti Model training, we relied on the LOL-dataset [26], an openly accessible dataset consisting of 500 pairs of images captured in varying lighting conditions. These pairs were divided into 485 training pairs and 15 testing pairs, with each image having a resolution of 400×600 pixels and being stored in RGB format. The LOL-dataset encompasses real-world captures as well as synthetic data, offering a diverse range of paired training datasets tailored for low-light image enhancement networks.

C. LLIE Model Training Procedure

Our LLIE model is primarily constructed on a CNN-based architecture, leveraging self-captured paired data from the LOL-dataset [26]. The dataset was partitioned randomly into 485 training pairs and 15 testing pairs, with images resized to dimensions of 400×600 pixels. Implementation was facilitated using TensorFlow and Keras libraries on an NVIDIA 2080Ti GPU, with a batch size of 16 employed during training. We initialized filter weights using a standard zero-mean Gaussian function with a standard deviation of 0.02, while biases were initialized as constants. ADAM optimizer with default parameters and a fixed learning rate of 0.0001 was utilized for network optimization. Activation functions such as ReLU and Tanh were incorporated into the model architecture. The execution of the code and the calculation of various parameters are illustrated in the accompanying screenshot depicted in Fig 5.

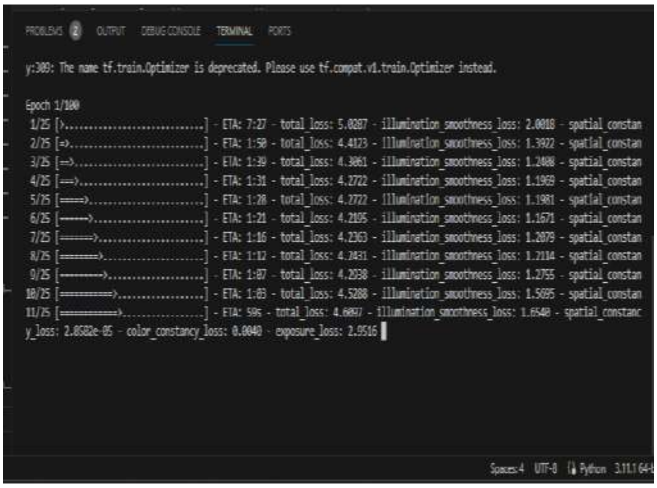

Fig. 5. Nakshatra-Drishti Model Training

D. Cost Function Visualisation

The integration of exposure loss, colour constancy loss, spatial constancy loss, illumination loss, and smoothness loss functions enhanced the model's performance, resulting in improved image enhancement results. The fluctuations of these loss functions over epochs are illustrated in the accompanying charts below:

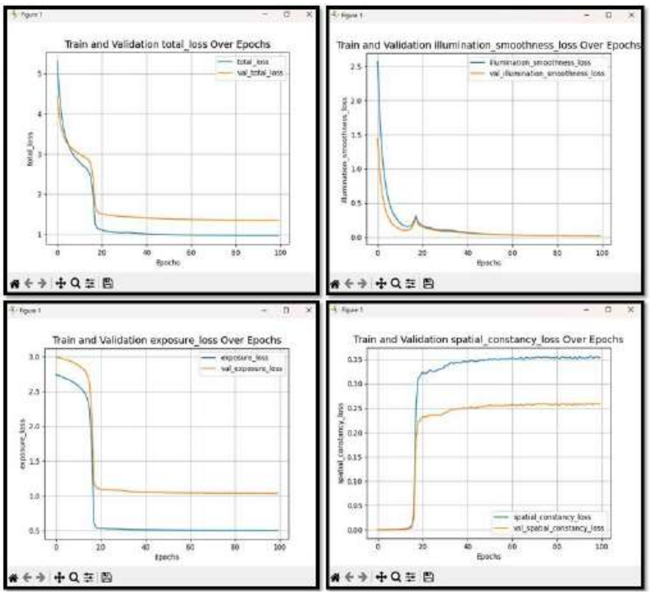

Fig. 6. Cost Function Visualisation

V. Results and Analysis

The web-based graphical user interface (GUI) of our low-light image enhancement model offers a user-friendly dashboard equipped with modules for image and video viewing, as well as live camera feed access. This interface simplifies user interaction by providing a centralized platform for uploading images, videos, and accessing live camera feeds, all conveniently accessible from a single location. With just a click, users can initiate the enhancement process, producing high-quality results effortlessly. Utilizing a Flask-based web interface on the frontend ensures smooth navigation between modules, delivering an intuitive and visually appealing experience for users.

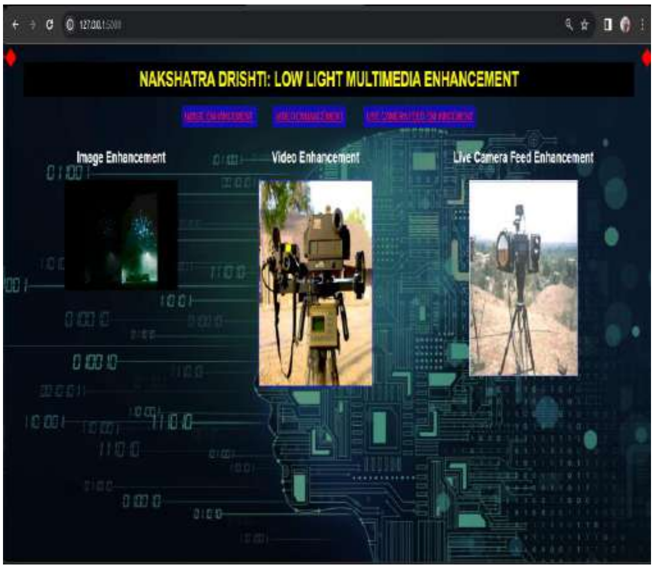

Fig. 7. GUI Dashboard of Nakshatra Drishti

A. Image Enhancement Module

The low-light image enhancement module serves as a fundamental component within our Nakshatra-Drishti model, aimed at enhancing the visual quality of images captured under low-light conditions. Utilizing the trained Nakshatra-Drishti model, this module enhances brightness, contrast, and overall clarity while mitigating noise and artifacts inherent in low-light photography. By analyzing input images and applying the learned weights from the deep learning model training process, this module unveils hidden details within shadows or darkness, thereby significantly improving image quality.

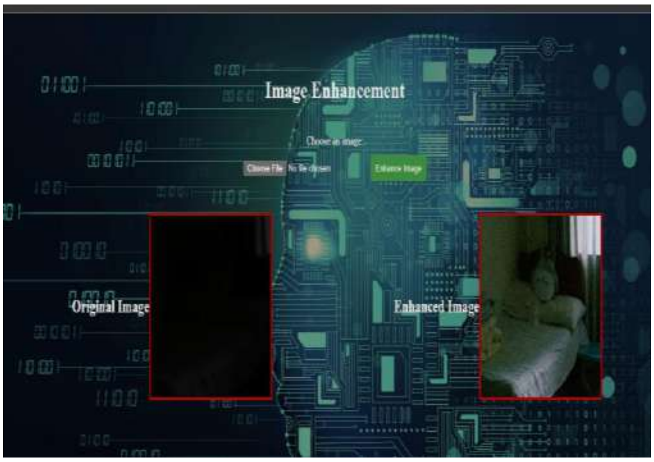

Fig. 8. Results of Image Enhancement Module

B. Low-Light Video Enhancement Module

The low-light video enhancement module is an extension of the image enhancement module, tailored to process videos captured in low-light environments. Leveraging the Nakshatra-Drishti model, this module enhances the visibility and clarity of video footage, ensuring that moving scenes captured in challenging lighting conditions appear clear and well-defined. Through consistent application of enhancement techniques across each frame of the video, this module delivers smooth and visually appealing results, enabling users to enjoy enhanced video playback without compromising quality or performance.

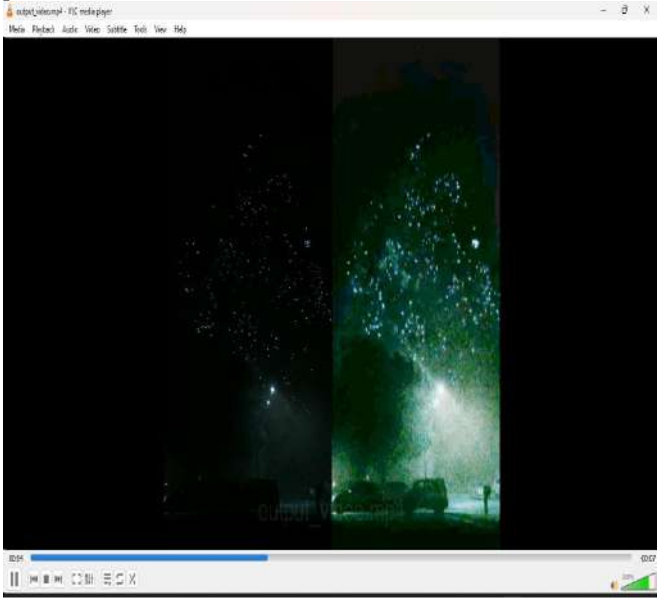

Fig. 9. Results of Low-Light Video Enhancement Module

C. Low-Light Real-Time Camera Feed Enhancement Module

The low-light real-time camera feed enhancement module offers real-time enhancement capabilities for live camera feeds, facilitating improved monitoring and visualization of scenes captured in low-light conditions. This module processes incoming video streams from connected cameras in real-time, applying enhancement algorithms to enhance visibility and quality on the fly. Through dynamic adjustment of enhancement parameters in response to changing lighting conditions, this module ensures that live camera feeds maintain clarity and detail, providing users with enhanced visibility for surveillance, monitoring, or other applications in low-light environments. Our Nakshatra-Drishti model has undergone testing with various cameras, including webcams, HD webcams, and IP cameras, demonstrating excellent results across diverse low-light scenarios.

Fig. 10. Results of Low-Light Real-Time Camera Feed Enhancement

VI. Analysis and Interpretation

A. Software Frameworks and Tools

The project is developed using open-source and platform-independent programming languages and software frameworks, including VSCodeUserSetup-x64-1.87.2, Python version 3.10.4, Flask web framework, and HTML. Minor configuration adjustments enable execution on LINUX environments as well.

B. Evaluation Metrics

In addition to subjective evaluations based on human perception, objective assessment of image quality is conducted using Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). These metrics are widely employed for Image Quality Assessment (IQA), with PSNR representing an infinite value and MSE approaching zero indicating superior image quality. The table below presents the experimental results on performance metrics used to evaluate the performance of our LLIE model on datasets, detailed below.

TABLE I. Quantitative comparisons on LOL-test and Brightening Train datasets in terms of MSE (×103), PSNR (in dB)

Learning	**Deep Learning Model**	**LOL-Dataset**		**Brightening-Train**	
		MSE	**PSNR**	**MSE**	**PSNR**
Supervised Learning	Nakshatra Drishti	1.18	18.85	1.39	17.95

C. Capabilities of the Nakshatra-Drishti Deep Learning Model

The Nakshatra-Drishti Deep Learning Model exhibits a wide array of capabilities, enabling it to proficiently address various tasks associated with enhancing low-light imagery. With its advanced architecture, the model can adeptly process low-light images, videos, and live camera feeds in real-time. This functionality is complemented by a user-friendly graphical user interface (GUI), which consolidates all enhancement solutions into a single, intuitive platform.

Users can effortlessly navigate the interface to access different features and functionalities. Additionally, the model provides insightful performance visualization, presenting graphical representations of performance parameters and metrics on a dedicated dashboard. These metrics, including epoch details acquired during model training and evaluation results on testing data, offer valuable insights into the model's performance. Moreover, the Nakshatra-Drishti model facilitates dynamic result storage, allowing users to store and retrieve results in a database for further analysis. Its flexibility is a notable asset, empowering security and control room operators to tailor model training according to their unique datasets and specific requirements.

VII. Conclusion

In conclusion, our research introduces the Nakshatra-Drishti deep learning model for enhancing low-light images, videos, and real-time camera feeds. The model demonstrates promising results across various datasets, surpassing existing techniques in terms of performance metrics and benchmarks. By integrating image, video, and live feed enhancement within a unified framework, accompanied by a user-friendly GUI, our approach offers a comprehensive solution to the challenges posed by low-light conditions.

Looking ahead, our study identifies several open issues for future exploration. Distinguishing semantic regions within low-light images remains a critical challenge, as existing methods often fail to consider the semantic information of different image regions. Similarly, the removal of unknown noises and artifacts poses significant challenges, necessitating further research efforts in these areas.

In summary, our work contributes to the advancement of low-light image enhancement methodologies, paving the way for more effective and robust solutions. By addressing current limitations and identifying future research directions, we aim to continually improve the quality of low-light image and video processing, thereby benefiting a wide range of applications and scenarios.

References

[1] E. H. Land, "The retinex theory of color vision," Scientific American, vol. 237, no. 6, pp. 108–129, 1977.

[2] S. M. Pizer et al., "Adaptive histogram equalization and its variations," Computer Vision, Graphics, and Image Processing, vol. 39, no. 3, pp. 355–368, 1987.

[3] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," Pattern Recognition, vol. 61, pp. 650–662, 2017.

[4] C. Wei et al., "Deep retinex decomposition for low-light enhancement," arXiv preprint arXiv:1808.04560, 2018.

[5] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1646–1654.

[6] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982–993, 2017.

[7] M. Zhu et al., "EEMEFN: Low-light image enhancement via edge-enhanced multi-exposure fusion network," in AAAI, 2020, pp. 13 106–13 113.

[8] C. Wei et al., "Deep retinex decomposition for low-light enhancement," arXiv preprint arXiv:1808.04560, 2018.

[9] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation," in Image Processing (ICIP), 2012 19th IEEE International Conference on, 2012, pp. 965–968.

[10] D. J. Jobson et al., "A multiscale retinex for bridging the gap between color images and the human observation of scenes," IEEE Transactions on Image Processing, vol. 6, no. 7, pp. 965–976, 1997.

[11] S. Wang et al., "Naturalness preserved enhancement algorithm for non-uniform illumination images," IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3538–3548, 2013.

[12] X. Fu et al., "A weighted variational model for simultaneous reflectance and illumination estimation," in CVPR, 2016, pp. 2782–2790.

[13] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982–993, 2017.

[14] X. Ren et al., "Joint enhancement and denoising method via sequential decomposition," in Circuits and Systems (ISCAS), 2018 IEEE International Symposium on, 2018, pp. 1–5.

[15] M. Li et al., "Structure-revealing low-light image enhancement via robust retinex model," IEEE Transactions on Image Processing, vol. 27, no. 6, pp. 2828–2841, 2018.

[16] M. Gharbi et al., "Deep bilateral learning for real-time image enhancement," ACM Transactions on Graphics (TOG), vol. 36, no. 4, p. 118, 2017.

[17] N. K. Kalantari and R. Ramamoorthi, "Deep high dynamic range imaging of dynamic scenes," ACM Trans. Graph, vol. 36, no. 4, p. 144, 2017.

[18] S. Wu et al., "Deep high dynamic range imaging with large foreground motions," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 117–132.

[19] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," IEEE Transactions on Image Processing, vol. 27, no. 4, pp. 2049–2062, 2018.

[20] C. Chen et al., "Learning to see in the dark," arXiv preprint arXiv:1805.01934, 2018.

[21] I. Goodfellow et al., "Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672–2680.

[22] X. Gong et al., "AutoGAN: Neural architecture search for generative adversarial networks," in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 3224–3234.

[23] Y. Jiang et al., "EnlightenGAN: Deep light enhancement without paired supervision," IEEE Transactions on Image Processing, vol. 30, pp. 2340–2349, 2021.

[24] C. Guo et al., "Zero-reference deep curve estimation for low-light image enhancement," in CVPR, 2020, pp. 1780–1789.

[25] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," TPAMI, 2021.

[26] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in Proceedings of the British Machine Vision Conference (BMVC), 2018.

[27] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[28] X. Zhang, X. Chen, X. Huang, and J. Gao, "Learning to See in the Dark via Wavelet Domain Attention Networks," in Proceedings of the European Conference on Computer Vision (ECCV), 2020.

[29] Y. Zhao, H. Guo, and J. Zhang, "Low-Light Image Enhancement Using a Conditional Generative Adversarial Network," IEEE Access, vol. 8, pp. 147744-147756, 2020.

[30] S. Li, Y. Huang, and Q. Tian, "Rethinking of Learning-based Low-light Image Enhancement: A Data-driven Perspective," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2020.