International Journal of Communication, Science and Technology (eISSN: 3049-0650)
vol. 1. no. 3 (June-July, 2025), pp.16-23

RetinexGAN: A Hybrid Approach for Low-Light Image Enhancement

Vrashabh Patil
Department of CSE
KLE Dr. MSSCET
Belagavi, Karnataka, India

Savita Bagewadi
Department of CSE
KLE Dr. MSSCET
Belagavi, Karnataka, India

Uma Mudenagudi
Department of CSE
BVBCET
Hubbali, Karnataka, India

***Abstract*—Low-light image enhancement is a significant challenge in computational imaging, necessitating innovative solutions that maintain visual information under suboptimal illumination. This research presents a novel hybrid methodology that integrates RetinexNet decomposition with a Generative Adversarial Network (GAN)-based enhancement model, effectively addressing key limitations in existing low-light image processing techniques. By decomposing images into reflectance and illumination components using an advanced deep learning framework, our approach develops a comprehensive strategy for reconstructing high-quality images. The method emphasizes the effective separation and enhancement of image components while preserving structural integrity and perceptual consistency. This two-stage enhancement process merges Retinex theory with sophisticated neural network architectures, exploring innovative strategies to transform images captured in challenging lighting conditions. Our methodology aims to overcome traditional challenges in low-light imaging, including noise amplification, color distortion, and loss of structural details.The findings demonstrate the robustness of this approach across various low-light conditions, ensuring adaptability beyond specific domains such as surveillance, autonomous driving, and medical imaging. The model is evaluated on diverse low-light environments, highlighting its potential for broader real-world applicability.**

***Index Terms*—Low-light image enhancement, Retinex theory, Generative Adversarial Network (GAN), Image processing , Reflectance ,illumination**

‘

I. Introduction

Low-light imaging presents significant challenges in computational imaging, as quality levels of images are often found to be deteriorated on account of reduced visibility, diminished contrast, and heightened noise levels. These issues degrade visual quality and impair the performance of critical tasks such as object detection, recognition, and tracking, which are essential in applications like autonomous driving, medical imaging, security surveillance, and remote sensing. Therefore, innovative techniques for low-light image enhancement are vital to ensure the reliability of these systems under challenging lighting conditions.

This research introduces a novel hybrid methodology that integrates RetinexNet decomposition [1] with a Generative Adversarial Network (GAN) [2]-based enhancement model. This approach effectively addresses key limitations in existing low-light image processing techniques by decomposing images into reflectance and illumination components using an advanced deep learning framework. Our method emphasizes the effective separation and enhancement of these components while preserving structural integrity and perceptual consistency.

The Retinex theory [3], developed by Land and McCann in the 1970s, serves as a foundational framework for this study. It models human vision's ability to perceive consistent colors under varying lighting conditions by decomposing an image into two components: reflectance (representing intrinsic properties of the scene) and illumination (which accounts for varying lighting conditions). By estimating and correcting the illumination component while preserving reflectance, Retinex-based methods have shown significant potential in improving image visibility in low-light environments.

Traditional methods, such as histogram equalization and gamma correction often fail to address the complexities of low-light scenarios effectively in spite of providing straightforward enhancements. Recent advances in deep learning have introduced novel architectures. They are designed with focus on low-light enhancement; however, many still struggle with issues such as over-enhancement, color distortion, and visual artifacts.

The key contributions of this research are as follows:

- Proposed a GAN [2]-based Retinex model: This new method merges Retinex theory with GANs to improve low-light images. It also tackles usual issues like over-enhancement and artifacts.
- Strength across different uses: The model shows strength in important areas like security, self-driving cars, and healthcare imaging, where good visual rebuilding is very important.
- Performance review: This study provides a detailed look at how well the suggested method works, comparing it to current techniques with numbers like PSNR, SSIM, and more, as well as visual checks for quality.

The rest of this paper is set up like this Section II gives background on main ideas of Retinex theory. Section III looks at previous work in low-light image enhancement and points out different methods that exist. Section IV explains the methodology, focusing on the suggested GAN-based Retinex approach. Section V shows the results, including a detailed analysis with both numbers and visual assessments. Lastly, Section VI wraps up the paper by summarizing key points, contributions, and possible future research paths in low-light image enhancement.

II. Background

This section provides a foundational overview of Retinex theory. This theory models the ability of human visions in terms of perceiving consistent colors under varying lighting conditions. Retinex- based approaches serve as the basis for many low-light image enhancement techniques by estimating and correcting the illumination component.

A. Retinex Theory

Retinex theory [3], introduced by Land and McCann, explains how human vision perceives colors consistently across different lighting conditions. The theory suggests that an observed image can be decomposed into two components:

$$I(x, y) = R(x, y) \cdot L(x, y) \tag{1}$$

where:

- $I(x, y)$ is the observed image,
- $R(x, y)$ is the **reflectance** component, representing the intrinsic properties of the scene, and
- $L(x, y)$ is the **illumination** component, representing the varying lighting conditions.

Retinex-based image enhancement techniques aim to recover the **reflectance** while appropriately estimating and adjusting the **illumination**, improving visibility in low-light conditions.

1) Variants of Retinex Theory: Several computational models have been developed based on Retinex theory:

- **Single-Scale Retinex (SSR)** – Enhances contrast by applying Retinex processing at a fixed scale but struggles with over-enhancement or loss of details.
- **Multi-Scale Retinex (MSR) [4]** – Improves upon SSR by applying Retinex processing at multiple scales, balancing detail enhancement and color consistency.

Retinex theory serves as a foundation for modern low-light image enhancement techniques, inspiring methods that combine Retinex with deep learning to achieve improved performance.

III. Related Work

This section reviews existing methods for low-light image enhancement, focusing on their strengths, weaknesses, and limitations. We categorize these approaches into Retinex-based methods, deep learning models, and GAN-based techniques.

A. Retinex-Based Methods

Conventional Retinex-based methods leverage the theory's ability to decompose images (as described in Section II.A) into reflectance and illumination components. For example, [1] uses guided image filtering to refine the illumination map. However, a key limitation of many traditional Retinex-based methods is their assumption that low-light images are noiseless, leading to significant noise amplification. Furthermore, they often rely on handcrafted priors that require extensive parameter tuning, limiting their adaptability to diverse scenarios.

B. Deep Learning Approaches

Deep learning has brought about noteworthy advancement in low-light image enhancement. Zero-Reference Deep Curve Es- timation [5] reformulates the enhancement task as pixel-wise curve estimation guided by non-reference loss functions. Wei et al. [1] combined Retinex decomposition with deep learning. However, these methods often necessitate complex multi- stage training pipelines that can hinder practical application. URetinex-Net [6] incorporates deep unfolding optimization with Retinex theory. While effective, these models can struggle with generalization across varying conditions and complex illumination scenarios.

C. GAN-Based Methods

Generative Adversarial Networks (GANs) enhance low-light images by learning illumination mappings while preserving details. EnlightenGAN [7] enables unsupervised enhancement, while RetinexGAN [8] integrates Retinex decomposition for illumination correction. SID-GAN [9] improves extreme low-light images. These approaches leverage adversarial training to enhance visibility while reducing noise.

Building on these advancements, our work introduces a novel GAN-based Retinex model that integrates GANs with Retinex decomposition for adaptive illumination enhancement. This approach aims to retain structural details while mitigating artifacts commonly found in previous methods. By addressing the shortcomings of existing techniques in terms of contrast improvement, illumination correction, and noise removal, our methodology offers a more robust solution for low-light image enhancement.

IV. Methodology

This section details our proposed RetinexGAN framework for low-light image enhancement. RetinexGAN leverages the strengths of both RetinexNet for robust reflectance-illumination decomposition and a custom-designed Generative Adversarial Network (GAN) to refine the illumination component and enhance overall image quality. While RetinexNet [1] excels at separating reflectance and illumination, it can sometimes produce results with residual noise and artifacts. To address these limitations, we introduce a GAN-based refinement stage that learns to generate visually appealing and artifact-free enhanced images. RetinexGAN builds upon the foundation of previous low-light enhancement techniques by combining the interpretability of Retinex theory with the generative power of GANs. This framework integrates the RetinexNet [1] model with a custom GAN-based enhancement model. The combination of these techniques aims to exploit the strengths of both methods to overcome the limitations of traditional low-light image enhancement approaches, which often fail to maintain visual fidelity or preserve important structural details. The methodology is composed of two main stages: the first involves decomposing the input image using the Retinex theory to separate illumination from reflectance, and the second involves enhancing these decomposed components using a Generative Adversarial Network (GAN) [2]. This two-tier process effectively addresses both illumination degradation and noise while preserving the intrinsic scene features.

A. RetinexNet: Decomposing Illumination and Reflectance

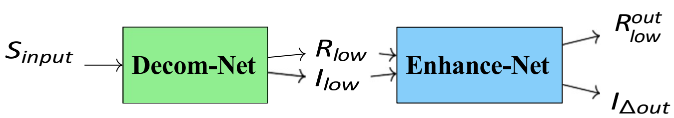

Fig. 1. Decomposition and Enhancement Pipeline by RetinexNet.

At the core of our methodology is **RetinexNet**, a deep learning-based model that applies the Retinex theory to decompose low-light images into two distinct components: illumination and reflectance. The Retinex theory [3] posits that an image can be modeled as the product of these two factors. The reflectance map captures the intrinsic properties of the scene (e.g., textures and shapes), while the illumination map reflects the environmental lighting conditions (e.g., brightness and shadows).

- **Reflectance Map (**R**)**: This map contains the intrinsic details of the scene and remains unaffected by lighting conditions. It preserves crucial features such as texture, edges, and structural elements necessary for object recognition and scene understanding.
- **Illumination Map (**L**)**: The illumination map models the lighting effects responsible for low-light conditions, including shadows, highlights, and ambient light. These effects can obscure fine details in low-light images.

RetinexNet [1]is specifically designed to handle this decomposition process by leveraging deep convolutional neural networks to accurately separate these components through two main modules:

1) **Decom-Net**: This module extracts the reflectance and illumination maps using multiple convolutional layers to learn features that help isolate reflectance from illumination. The reflectance map represents true scene information while the illumination map models varying lighting effects.
2) **Enhance-Net**: This module refines the illumination map by adjusting its brightness, contrast, and clarity to correct low-light conditions without altering the reflectance map. By enhancing the illumination map, we restore brightness and visibility, making images appear more natural and easier to interpret.

The outputs from RetinexNet are:

- **Enhanced Illumination Map (**I_{Δ}**)**: This map represents an improved version of the illumination where lighting effects have been corrected for better contrast and brightness.
- **Reflectance Map (**R_{low}**)**: This map retains the structural integrity of the original image with no alterations made to the reflectance information.

Once these components are generated, they are passed on to the GAN-based enhancement model for further refinement. This stage ensures that both the enhanced illumination map is optimized further and overall image quality is improved.

B. GAN-based Enhancement Model: Refining Illumination and Reflectance Maps

Following the decomposition of the illumination map using RetinexNet, the enhanced illumination map and the original reflectance map are fed into the **GAN-based enhancement model** for further refinement. This GAN architecture enhances the image by improving the visual quality of both the illumination and reflectance maps while ensuring that the structural details of the scene are preserved.

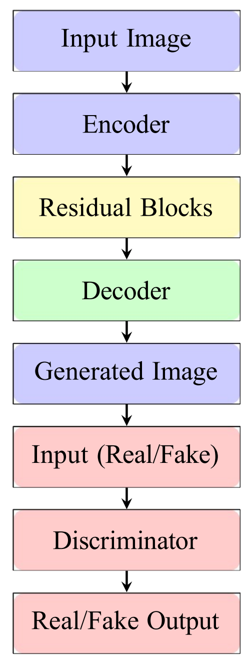

Fig. 2. RetinexGAN Architecture

The GAN [2] model, depicted in the diagram, consists of two primary components: the **Generator** and the **Discriminator**, which are trained adversarially.

- The **Generator** takes the input image and processes it through several stages: first, it passes through the **Encoder** [10], which extracts features, then through **Residual Blocks** [11], which enhance the features and finally, the **Decoder** [10], which generates the output image. The **Generated Image** is the final output produced by the Generator.
- The **Discriminator**, placed below the Generator in the diagram, receives the **Input (Real/Fake)** images, which include both real images and the generated ones. It

processes these inputs through a **Discriminator** block and outputs a **Real/Fake Output**, indicating whether the input image is real or generated.

The Generator and Discriminator are connected such that the **Generated Image** from the Generator is passed to the **Discriminator**, which then classifies it. The adversarial training ensures that as the Generator improves in producing more realistic images, the Discriminator becomes better at distinguishing between real and fake images. This constant interplay between the two components leads to the Generator yielding increasingly high-quality results.

1) Generator Architecture: The Generator network is designed to refine both the illumination and reflectance components output by RetinexNet. It learns to generate a more visually appealing image by enhancing brightness, contrast, and structure while preserving scene details. The Generator follows a deep neural network architecture with the following layers:

1) **Encoder** [10]: The encoder extracts high-level features from the input components (the enhanced illumination map and reflectance map). It consists of several convolutional layers that reduce spatial dimensions while capturing important features at multiple scales. The architecture starts by applying a convolution operation, followed by batch normalization and ReLU activation to ensure efficient feature extraction and transformation. The residual blocks, located at the end of the encoder, enhance the learning process by enabling the model to retain and process important information from earlier layers.

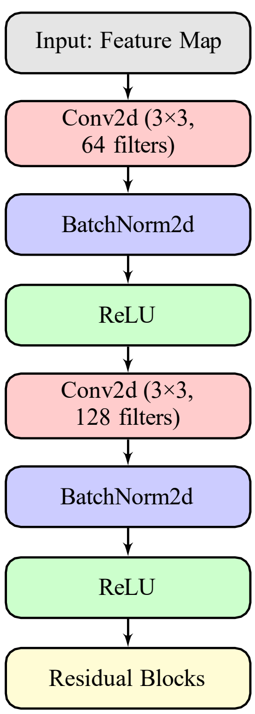

Fig. 3. Encoder Architecture.

The encoder effectively captures important features by progressively refining the feature map through convolutional layers. The use of batch normalization ensures stable training by normalizing the output of each layer, while the ReLU activations introduce non-linearity. The residual blocks further enhance the learning capability by allowing the network to retain critical information from the initial layers.

2) **Residual Blocks** [11]: These blocks refine the image by learning residual features. Each block contains two convolutional layers with skip connections, allowing the network to learn additional fine details that contribute to overall enhancement. The residual blocks are designed to retain crucial information from the input and combine it with learned features to produce more detailed output. This helps improve the network's performance by avoiding vanishing gradients and enabling it to learn deeper features more effectively.

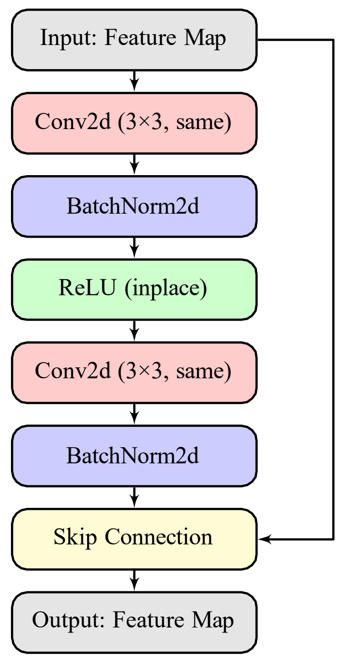

Fig. 4. Residual Block.

The residual block shown in the diagram includes two convolutional layers, each followed by batch normalization and ReLU activation. The output of the second convolution is then added to the original input via a skip connection. This skip connection helps retain information from earlier layers, allowing the network to learn both fine-grained features and overall patterns. By incorporating these residual connections, the block facilitates the learning of more complex features while improving training stability and performance.

3) **Decoder** [10]: The decoder restores spatial resolution using transposed convolution layers. It reconstructs the image from the feature maps and outputs the enhanced image. The decoder is responsible for mapping the compressed, high-level features back into the image space,

which is crucial for generating a full-resolution image. The final layer applies a Tanh activation function, which helps scale the pixel values between $[-1, 1]$, making the output suitable for visual representation.

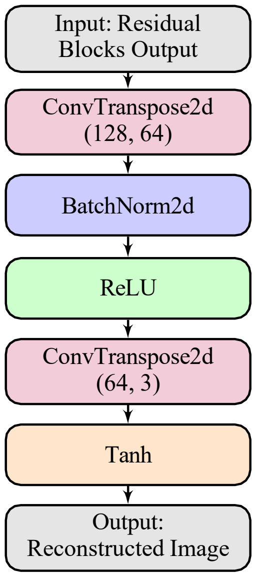

Fig. 5. Decoder Architecture.

The decoder architecture shown in the diagram restores the spatial resolution of the image. The input to the decoder is the output of the residual blocks, which contains refined feature maps. It first applies a transposed convolution layer to upsample the feature map. Batch normalization and ReLU activations help stabilize training and introduce non-linearity. Finally, the decoder uses another transposed convolution layer to produce the image with three channels (RGB) and applies the Tanh activation function to ensure the pixel values are appropriately scaled for output.

2) Discriminator Architecture: The Discriminator network's primary role is to distinguish between real and generated images using a patch-based approach. The input image is divided into non-overlapping patches of size $N \times N$ (e.g., 64x64 pixels), and each patch is classified independently as either real or fake. This patch-based approach encourages the Generator to produce high-frequency details and realistic textures at a local level. The Discriminator consists of four convolutional layers, each followed by a Leaky ReLU activation and batch normalization. The final layer outputs a single value representing the probability that a given patch is real.

1) **Convolutional Layers**: The Discriminator uses four convolutional layers to extract features from input images at multiple scales.
2) **Leaky ReLU Activation**: Leaky ReLU activations allow gradients to flow through negative values during training, preventing issues associated with "dying ReLU," which can hinder learning, especially in deep networks.
3) **Batch Normalization**: Batch normalization is applied after each convolutional layer (except the final layer) to stabilize training and reduce internal covariate shift, allowing for higher learning rates and faster convergence.
4) **Sigmoid Output**: The final convolutional layer outputs a sigmoid value representing the probability that an input image patch is real. This output is then used to calculate the adversarial loss, guiding the Generator to produce more realistic patches.

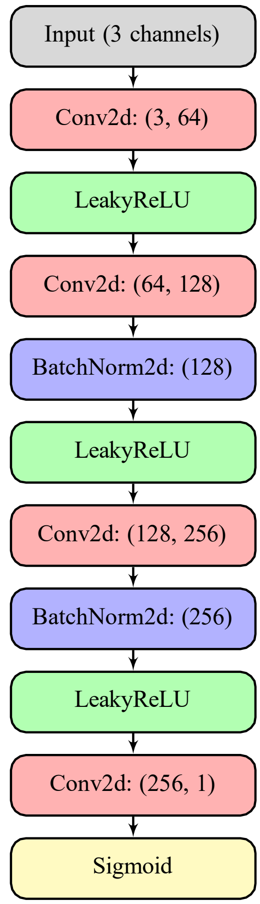

Fig. 6. Architecture of the Patch Discriminator Network used for distinguishing real and generated patches.

In summary, our GAN-based enhancement model effectively refines low-light images by leveraging both illumination and reflectance maps while maintaining structural integrity through adversarial training.

3) Loss Functions: To optimize the performance of the Generator and Discriminator, we use a combination of adversarial loss, pixel-wise loss, perceptual loss, and image quality losses (PSNR and SSIM) [12]. These losses guide the Generator to produce visually realistic images while maintaining high fidelity to the original scene structure. The total loss for the Generator is a weighted sum of the individual loss components:

$$L_G = L_{adv} + \lambda_1 L_{pixelwise} + \lambda_2 L_{perceptual} + \lambda_3 L_{PSNR} + \lambda_4 L_{SSIM}$$

where L_{adv} is the adversarial loss, $L_{pixelwise}$ is the pixel-wise loss, and $L_{perceptual}$ is the perceptual loss, with additional terms for PSNR and SSIM [12]. The weighting factors (λ_1, λ_2, λ_3, λ_4) control the relative importance of each loss term. The specific form of each loss function and the values of the weighting factors were determined experimentally.

- **Adversarial Loss** [2]: The Generator aims to minimize the adversarial loss by trying to deceive the Discriminator. The Discriminator, in turn, works to correctly classify images as real or generated. This adversarial setup improves the realism of the generated images.
- **Pixel-wise Loss**: The pixel-wise loss ensures that the generated image is close to the ground truth in terms of pixel values. This helps reduce artifacts and ensures better image quality.
- **Perceptual Loss**: Perceptual loss compares the feature representations of the generated and ground truth images, encouraging the Generator to learn high-level features that align with the target.
- **PSNR and SSIM Losses** [12]: These image quality metrics help preserve the overall image structure and reduce the distortion caused by low-light enhancement.

C. *Computational Requirements*

The proposed method was trained on a dual Tesla T4 GPU setup (16GB VRAM each) in the Kaggle environment, leveraging multi-GPU parallelism to optimize computational efficiency. The inference time per image is 0.12s, with RetinexNet decomposition accounting for 40% and the GAN-based enhancement stage contributing 60% of the total processing time. Given the complexity of the model, GPU acceleration is essential, as CPU-based execution significantly increases latency. The multi-stage framework produces better low-light image quality but future optimizations including model pruning and quantization and knowledge distillation will enhance computational efficiency for real-time applications in resource-constrained environments.

D. *Overall Approach*

In summary, the proposed method combines the RetinexNet [1] decomposition for low-light image enhancement with a GAN-based refinement model to produce high-quality, detailed images. The two-stage pipeline allows for effective separation of illumination and reflectance components, followed by advanced enhancement techniques using GANs to refine the image. The use of multiple loss functions, including perceptual and image quality-based losses, ensures that the enhanced images maintain structural integrity while appearing natural and clear.

V. Results and Evaluation

In this section, we present the results of the proposed low-light image enhancement method. We evaluate the performance of our approach in terms of both qualitative and quantitative metrics. The following experiments were conducted to validate the effectiveness of our method.

A. *Qualitative Evaluation*

To showcase the improvements brought about by our proposed method, we present a selection of low-light images from the LOL [1] dataset alongside their enhanced versions. These results highlight the effectiveness of our Retinex-based decomposition and GAN-based enhancement technique in significantly boosting image brightness, contrast, and detail preservation.

1) Overview of the Enhancement Process: Our low-light image enhancement model follows an intuitive input → encoder → residual → decoder architecture. In this section, we will walk through each stage of the processing pipeline, illustrating how each contributes to the overall enhancement of the images.

Fig. 7. Input Illumination Map.

The original image, displayed in Figure 7, represents the illumination map generated by RetinexNet. This serves as the starting point for our model and reflects the challenging lighting conditions present in the scene. As you can see, this image requires significant enhancement to improve clarity and detail representation.

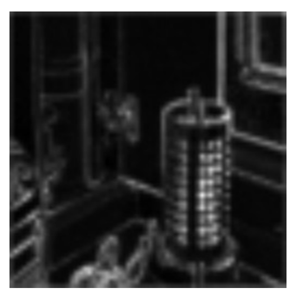

Fig. 8. Encoder architecture output.

After passing through the encoder, the image is transformed into a feature-extracted representation, as shown in Figure 8. The encoder plays a vital role in isolating key features—like edges and textures—while filtering out unnecessary data. This simplification helps the model to place emphasis on the important vi- sual elements needed for effective enhancement. The encoded image presents a distilled version of the original, capturing essential details that will be refined in later steps.

The residual component, illustrated in Figure 9, captures the differences between the original illumination map and its enhanced counterpart. This image highlights areas where

Fig. 9. Residual architecture output.

significant corrections are applied, particularly in terms of brightness, contrast, and sharpness. The residual image is crucial for ensuring that enhancements do not distort important details, allowing us to selectively enhance underexposed regions while maintaining edge sharpness.

Fig. 10. Final enhanced output.

The final output, shown in Figure 10, is produced after passing through the decoder, which reconstructs the image from both encoded features and residual enhancements. This stage restores visual details, resulting in an image characterized by improved brightness and contrast without introducing artifacts or overexposure. The output showcases a remarkable enhancement in perceptual quality, where finer details become clearer and natural colors are preserved.

The visual transition from the input image to the encoded, residual, and decoded images demonstrates how effectively our model enhances low-light images. By employing an encoder-decoder framework with residual enhancements, our approach successfully addresses challenges associated with low-light conditions—such as noise, poor contrast, and detail loss. In comparison to traditional enhancement methods, our model achieves a more balanced improvement while preserving fine details and avoiding issues like over-brightening or noise amplification.

This qualitative evaluation showcases how effectively our proposed method enhances low-light images, bringing out hidden details and improving overall clarity. The enhanced images demonstrate the model's ability to preserve essential features while making the scene more visually appealing.

To further illustrate our findings, we provide examples of low-light images alongside their high-lighted enhanced versions:

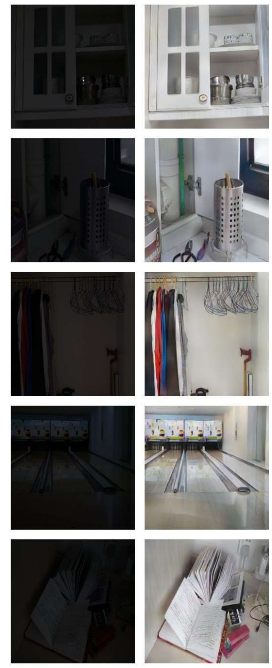

Fig. 11. Blended low-light and high-light image generated by our model.

B. Quantitative Evaluation

We used the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for ensuring that the performance of our method is quantitatively evaluated. PSNR is a widely used metric to assess the quality of images, with higher values indicating better image quality. SSIM contributes to the measurement of the structural similarity between images, with values closer to 1 indicating higher similarity. We compared our results with state-of-the-art low- light enhancement methods on the LOL [1] dataset to assess the improvements in terms of PSNR and SSIM.

Method	PSNR (dB)	SSIM
SID [13]	14.35	0.436
3DLUT [14]	14.35	0.445
DeepUPE [15]	14.38	0.446
RF [16]	15.23	0.452
DeepLPF [17]	15.28	0.473
IPT [18]	16.27	0.504
UFormer [19]	16.36	0.771
RetinexNet [1]	16.77	0.560
Sparse [20]	17.20	0.640
EnGAN [7]	17.48	0.650
RUAS [21]	18.23	0.720
FIDE [22]	18.27	0.665
DRBN [23]	20.13	0.830
KinD [24]	20.86	0.790
Restormer [25]	22.43	0.823
MIRNet [26]	24.14	0.830
SNR-Net [27]	24.61	0.842
Retinexformer [28]	25.16	**0.845**
Proposed Method	**26.32**	0.804

TABLE I
COMPARISON OF PSNR AND SSIM FOR VARIOUS LOW-LIGHT IMAGE ENHANCEMENT METHODS ON LOLV1.

C. Discussion

The results show a noteworthy improvement in PSNR for our method compared to traditional and Retinex-based methods. This significant improvement in PSNR can be attributed to the GAN-based refinement stage, which effectively reduces noise and enhances finer details. The proposed method achieves a PSNR [12] of **26.32 dB**, which represents a substantial enhancement in image quality, as higher PSNR values cor- respond to clearer and more visually accurate images. The enhanced images demonstrate better visibility, reduced noise, and preserved structural details, confirming the effectiveness of combining RetinexNet and a GAN-based enhancement approach.

This significant improvement in PSNR highlights the potential of our method in real-world applications requiring high-quality low-light image enhancement.

VI. Conclusion

In this paper, we presented a hyrid approach for low-light image enhancement method based on RetinexNet and Generative Adversarial Networks (GANs). Our approach combines the advantages of Retinex-based decomposition and GAN-driven enhancement to significantly improve image visibility, brightness, and detail preservation. Through both qualitative and quantitative evaluations, we demonstrated the superiority of our method over existing low-light enhancement techniques. The results show notable improvements in PSNR, indicating enhanced image quality, with better noise reduction and preservation of structural details. Our method outperforms traditional approaches and state-of-the-art methods such as RetinexNet and EliganentGAN , making it a promising solution for real-world low-light image enhancement tasks. Future work will focus on further optimizing the model for real-time processing and extending it to more challenging scenarios, such as dynamic low-light conditions.

References

[1] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," 2018. [Online]. Available: https://arxiv.org/abs/1808.04560

[2] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014. [Online]. Available: https://arxiv.org/abs/1406.2661

[3] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.

[4] Z.-u. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Proceedings of 3rd IEEE international conference on image processing*, vol. 3. IEEE, 1996, pp. 1003–1006.

[5] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," 2020. [Online]. Available: https://arxiv.org/abs/2001.06826

[6] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang, and J. Jiang, "Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 5891–5900.

[7] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," 2021. [Online]. Available: https://arxiv.org/abs/1906.06972

[8] T. Ma, M. Guo, Z. Yu, Y. Chen, X. Ren, R. Xi, Y. Li, and X. Zhou, "Retinexgan:unsupervised low-light enhancement with two-layer convolutional decomposition networks," *IEEE Access*, vol. 9, pp. 56 539–56 550, 2021.

[9] Y. Zhang, X. Di, B. Zhang, Q. Li, S. Yan, and C. Wang, "Self-supervised low light image enhancement and denoising," *arXiv preprint arXiv:2103.00832*, 2021.

[10] K. Aitken, V. V. Ramasesh, Y. Cao, and N. Maheswaranathan, "Understanding how encoder-decoder architectures attend," 2021. [Online]. Available: https://arxiv.org/abs/2110.15253

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015. [Online]. Available: https://arxiv.org/abs/1512.03385

[12] A. Hore´ and D. Ziou, "Image quality metrics: Psnr vs. ssim," in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 2366–2369.

[13] C. Chen, Q. Chen, M. Do, and V. Koltun, "Seeing motion in the dark," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 3184–3193.

[14] H. Zeng, J. Cai, L. Li, Z. Cao, and L. Zhang, "Learning image-adaptive 3d lookup tables for high performance photo enhancement in real-time," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 4, pp. 2058–2073, 2022.

[15] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 6842–6850.

[16] S. Kosugi and T. Yamasaki, "Unpaired image enhancement featuring reinforcement-learning-controlled image editing software," 2019. [Online]. Available: https://arxiv.org/abs/1912.07833

[17] S. Moran, P. Marza, S. McDonagh, S. Parisot, and G. Slabaugh, "Deeplpf: Deep local parametric filters for image enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 12 823–12 832.

[18] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, C. Xu, and W. Gao, "Pre-trained image processing transformer," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021, pp. 12 299–12 310.

[19] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general u-shaped transformer for image restoration," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 17 662–17 672.

[20] W. Yang, W. Wang, H. Huang, S. Wang, and J. Liu, "Sparse gradient regularized deep retinex network for robust low-light image enhancement," *IEEE Transactions on Image Processing*, vol. 30, pp. 2072–2086, 2021.

[21] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, "Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 10 556–10 565.

[22] K. Xu, X. Yang, B. Yin, and R. W. Lau, "Learning to restore low-light images via decomposition-and-enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 2278–2287.

[23] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "Band representation-based semi-supervised low-light image enhancement: Bridging the gap between signal fidelity and perceptual quality," *IEEE Transactions on Image Processing*, vol. 30, pp. 3461–3473, 2021.

[24] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," 2019. [Online]. Available: https://arxiv.org/abs/1905.04161

[25] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," 2022. [Online]. Available: https://arxiv.org/abs/2111.09881

[26] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao, "Learning enriched features for real image restoration and enhancement," 2020. [Online]. Available: https://arxiv.org/abs/2003.06792

[27] X. Xu, R. Wang, C.-W. Fu, and J. Jia, "Snr-aware low-light image enhancement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 17 714–17 724.

[28] Y. Cai, H. Bian, J. Lin, H. Wang, R. Timofte, and Y. Zhang, "Retinexformer: One-stage retinex-based transformer for low-light image enhancement," 2023. [Online]. Available: https://arxiv.org/abs/2303.06705