Attention-Driven Multi-Scale GAN with Complementary Learning Sub-Network for Low-Light Image Enhancement

Pansilu Wijesiri
Department of Computer Science and Engineering
University of Westminster
London, UK
pansiluwijesiri@gmail.com

Guhanathan Poravi
School of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
guhanathan.p@iit.ac.lk

***Abstract*—Low-light image enhancement (LLIE) is a vital aspect of downstream vision applications like surveillance, photography, and autonomous systems, where inferior lighting conditions result in degraded image quality and impaired performance of such vision tasks. This research presents AMGAN, an Attention-Driven Multi-Scale GAN, which integrates a Complementary Learning Sub-Network (CLSN) to address the challenges demonstrated by the non-uniform lighting. The CLSN generates the inverse grey maps as the attention maps, thereby guiding the enhancement process by emphasizing the under-illuminated areas while preserving well-illuminated regions. The model employs a U-Net-based generator with multi-scale feature extraction from both the spatial and frequency domains, thereby attending to texture preservation and natural color consistency. A dual Markov discriminator evaluates both local (patch-based) and global consistency, ensuring the visual coherency of the generated results. The model is trained and evaluated in quantitative and qualitative aspects on multiple benchmark datasets, including LSRW, LOLv1, LOLv2-real, and LOLv2-synthetic. Results demonstrate the potential of the proposed method towards state-of-the-art methods with averages of 18.64 dB of PSNR and 0.468 of SSIM values while maintaining perceptual quality. Ablation studies demonstrate the significance of CLSN and the frequency feature integration. The proposed model is produced as a solution for enhancing low-light images, with potential application across various domain-specific visual data in challenging lighting conditions.**

***Index Terms*—Low-Light Image Enhancement, GAN, Attention Mechanisms, Frequency Domain Processing, CNN.**

I. Introduction

Image enhancement is a crucial preprocessing step in computer vision, which aims to improve the visual perception and overall quality of the images captured in challenging environmental conditions [1]. Images that are acquired in low-light environments often suffer from low levels of brightness, contrast with color distortions, and noise [2]. The significance of low-light image enhancement lies in the ability to improve the performance of various high-level computer vision tasks [3]. High-quality images retain more information, which is essential for tasks such as object detection, image recognition, medical imaging, semantic segmentation, autonomous driving, surveillance, and face recognition [1], [4].

Low-light image enhancement (LLIE) techniques are targeted to enhance the visual quality by adjusting the illumination distribution and removing noise [2]. Images captured in inadequately illuminated environments can lead to loss of detail and lower contrast levels, which negatively affect human visual perception and the subsequent computer vision tasks that utilize the degraded image [4]. The goal of LLIE is to restore images to normal brightness, retain and enhance intricate details, and facilitate processing for subsequent downstream vision tasks [4].

Deep Learning techniques, especially Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), have been known to revolutionize image enhancement through the capture of spatial features with better detail and noise resistance. CNN-based methods are well-known in supervised learning through reliable local feature learning and end-to-end processing [2]. GAN-based methods are known for the generation of realistic images, addressing the generalization compartment [4]. Despite their advancements, CNNs and GANs alone leave room for improvement [2]. Therefore, combining the strengths of both CNNs and GANs presents a promising avenue for performance enhancement for a reliable solution [4]. Furthermore, the integration of additional feature retention from other candidate domains, such as the frequency domain, and the utilization of attention-mechanism-based enhancement for brightness holds significant potential [5]. To address such challenges and identified potential, this research proposes the Attention-Driven Multi-Scale GAN with Complementary Learning Sub-Network, such as AMGAN. The model aims to enhance the low-light images by leveraging both spatial and frequency domain features while adaptively adjusting the brightness factor through an inverse-gray-map-based attention mechanism. The key contributions of this work include:

- **Novel CLSN Module:** Introduces an inverse gray-map-based Complementary Learning Sub-Network to handle the non-uniform lighting adaptively by driving attention toward the areas based on the level of enhancement

required.

- **Multi-Scale Feature Extraction:** Utilizes both the spatial domain and the frequency domain information to preserve image detail, texture, and naturalness of the image.
- **Comprehensive Evaluation:** The author extensively benchmarks the performance on LSRW, LOLv1, LOLv2-real, and LOLv2-synthetic datasets, thereby demonstrating competitive quantitative and qualitative performance compared to the state-of-the-art methods.

The rest of the paper is organized as follows: Section II presents a review of related work in low-light image enhancement. Section III details the proposed AMGAN architecture, highlighting the CLSN and GAN integration. Section IV discussed the experimental setup and datasets used. Section V presents the evaluation results and comparisons with benchmark methods. Finally, Section VI concludes the paper and outlines the directions for future research.

II. Related Work

A. Traditional LLIE Approaches

Histogram Equalization (HE) is used to enhance the image contrast by spreading out the pixel intensity distribution [5]. However, it struggles with over-enhancement and noise amplification, more specifically in scenarios with non-uniform lighting conditions [1]. Retinex theory models the image as illumination and reflectance components [6]. Yet the Retinex theory itself struggles with color distortion and is unable to handle diverse and complex lighting variations [2]. While traditional methods offered early insights into the domain of LLIE, they led to struggle with noise, non-uniform illumination, detail, and texture loss, thereby prompting the researchers to explore deep learning techniques [1].

B. Deep Learning-Based LLIE

In the deep learning era, Zero-DCE [7] enhances brightness adaptively to the curve estimation strategy, yet the performance of the method in extremely dark regions diminishes. It dynamically adjusts the pixel and higher-order curves, yet may still lead to under-enhancement or over-exposure [6]. EnlightenGAN [8] employs unsupervised generative adversarial training for LLIE, yet it introduces artifacts and texture loss, more specifically when exposed to complex lighting environments [2]. Although it can learn feature mapping between the low-light image and the normal-light image without paired matching images, it tends to retain or increase noise in extremely dark areas and cause artifacts [9]. RetinexNet [10] utilizes Retinex decomposition within a deep learning framework but is unable to accompany adaptive mechanisms for non-uniform illumination [2]. While it enhances the images through the decoupling of illumination and reflectance components based on Retinex theory, it can produce results with color shifts [2], [3]. KinD [11] combines Retinex theory with a convolutional neural network, which divides the network into illumination and reflectance estimation. However, the enhanced results can have inadequate brightness overall, with noticeable artifacts and color deviations in enhanced areas [2]. SCI [12] addresses local darkness yet leads to uneven enhancement [2]. LIME [13] generates images with excessive contrast, non-uniform colors, and dazzling effects in local regions [14]. JED [15] has a serious enhancement problem due to information loss, with images that are not clear enough while texture details are blurred [14]. While these deep learning approaches improve the enhancement quality of traditional techniques, most lead to inadequate levels of enhancement in terms of non-uniform lighting, intricate texture, and spatial consistency factors.

C. Attention Mechanisms in LLIE Models

Attention mechanisms are frequently involved in LLIE models to drive focus toward salient features, yet they still have limitations [5], [16]. Although these methods show good performance, they all capture spatial information within the same scale [16]. Spatial attention mechanisms are integrated into models such as KinD [11] and MBLLEN [17]. For example, a bright attention mechanism can predict the illumination distribution spread in different regions of the image [5]. Some approaches utilize the channel and spatial attention modules to sequentially learn the regions to drive attention and in consideration of the spatial and channel domains, respectively [18]. The hybrid attention mechanism combines both the channel attention mechanism and the spatial attention mechanism to make up for the shortcomings of each domain [19]. Existing methods predominantly utilize the attention mechanisms in the spatial domain, yet the detail and texture refinement achievable through the frequency-domain utilization is neglected [20]. Many methods primarily focus on spatial domain information for addressing image degradation but often overlook the frequency domain information [20]. Methods based on CNNs may not be enough to address noise and color distortions, reigning from the inability to establish long-range feature dependencies [21]. Most methods employ dense dot-product self-attention as the core of the Transformers, and it may introduce potential noise by allowing redundant features that interfere with the attention map [20].

Despite these improvements, these attention-based models overlook the importance of utilizing frequency-driven detail, texture enhancement, and non-uniform lighting handling [6].

III. Methodology

The AMGAN system consists of an integration of a Complementary Learning Sub-Network through the inverse gray map produced and input to the generator network of the GAN, in which the spatial and frequency domain features are extracted for fine-grained detail and texture retention. The CLSN, inspired by the reinforcement branch technique demonstrated by LightingNet method [22], generates the inverse gray map with the illumination constraints for targeted brightness enhancement. Proceeding forward, the GAN, accompanied by the U-Net-based generator, retrieves the intricate image details in both the spatial and frequency domains of the degraded image. A dual-Markov discriminator is further established

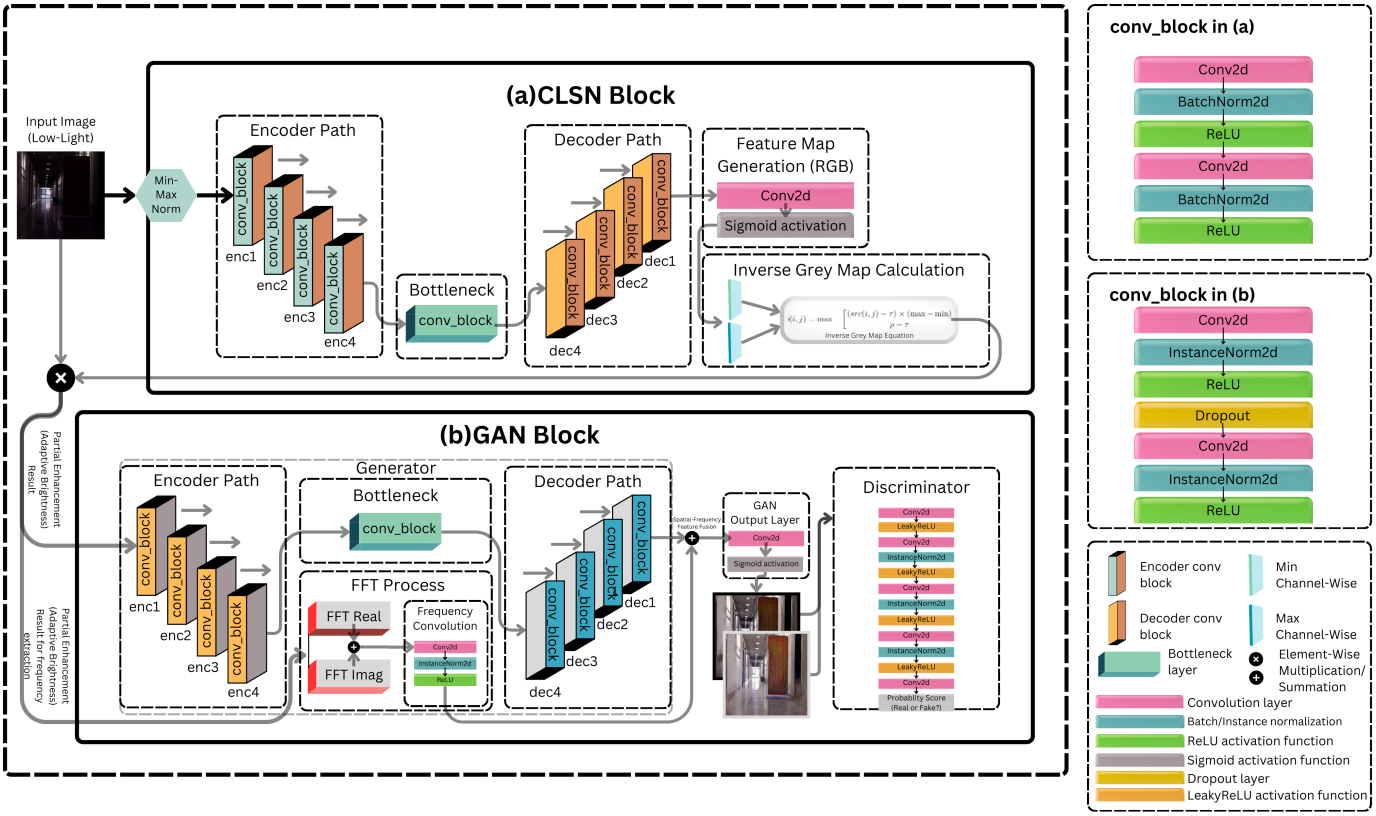

Fig. 1. The overall high-level architecture of the proposed AMGAN framework, illustrating the CLSN and GAN blocks along with frequency and spatial feature fusion.

to enable patch-based discrimination along with global (full-image) discrimination, thereby ensuring the realism of the enhancement.

A. Complementary Learning Sub-Network (CLSN)

1) Illumination constrain-based inverse gray map: The CLSN model generates an inverse gray map based on the illumination constraint equation inspired by MLLEN-IC [19], thereby driving attention to the regions that require greater attention due to inadequate illumination and leaving the rest with minimal enhancement. The resulting gray map is applied over the original degraded image, which selectively enhances the image regions based on the generated weights in the attention map.

2) The Inverse Gray Map as the Attention Map: The illumination constraint $\Phi(i,j)$ is considered for the RGB color channels, as inspired by the [19].

$$\Phi(i,j) = \phi_R(i,j) \otimes \phi_G(i,j) \otimes \phi_B(i,j) \tag{1}$$

Where $\phi_R(i,j)$, $\phi_G(i,j)$, and $\phi_B(i,j)$ are the illumination constraints for red, green, and blue channels, separately, which contributes to accurate attention calculation than generating the illumination constraint for the entire image, as a single channel, which leads to the loss of channel-specific image details. And, the $\otimes$ denotes the element-wise multiplication. Here each illumination constraint for each channel (R, G, B) is depicted as,

$$\phi(i,j) = \max - \left[\frac{(src(i,j) - \tau) \times (\max - \min)}{\rho - \tau}\right] \tag{2}$$

Where the pixel intensities at the coordinate positions (i,j) are formulated as $src(i,j)$. The max and min are the maximum and minimum channel-wise intensity values for the region. And τ and ρ are the maximum and minimum values within the channel. The generated inverse gray map is then element-wise multiplied by the original degraded image, thereby enhancing the brightness factor selectively. The CLSN, which is the CNN, is accompanied by the U-Net encoder-decoder architecture.

B. GAN-Based Image Refinement Phase

1) Fast Fourier Transform (FFT) for Frequency Domain Processing: As the U-Net-based generator network accompanies frequency features, the high-frequency details are extracted through a technique called Fast Fourier Transform (FFT), thereby extracting the Fourier features alongside the spatial feature. Here, the 2D Discrete Fourier Transform (DFT) technique is accompanied as followed by [3],

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)} \tag{3}$$

TABLE I
COMPARISON OF PSNR AND SSIM VALUES OF DIFFERENT LOW-LIGHT IMAGE ENHANCEMENT METHODS ON MULTIPLE DATASETS.

Method	LSRW		LOLv1		LOLv2-real		LOLv2-synthetic	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
RetinexNet [10]	15.609	0.414	16.77	0.560	15.47	0.567	17.13	0.798
DRBN [23]	16.734	0.507	20.13	0.830	20.29	0.831	23.22	0.927
RUAS [24]	14.271	0.461	18.23	0.720	18.37	0.723	16.55	0.652
MIRNet [25]	16.470	0.477	24.14	0.830	20.02	0.820	21.94	0.876
SNR-Net [26]	16.499	0.505	24.61	0.842	21.48	0.849	24.14	0.928
RetinexMamba [27]	19.536	0.576	24.03	0.827	22.00	0.849	25.89	0.943
RetinexFormer [28]	19.570	0.578	25.16	0.845	22.80	0.840	25.67	0.930
DPEC [29]	19.643	0.576	24.80	0.855	22.89	0.863	26.19	0.939
AMGAN (Ours)	**18.953**	**0.571**	**22.851**	**0.801**	**21.252**	**0.811**	**23.553**	**0.891**

Where the spatial domain image for color profile, contrast, and further brightness enhancement is as *f(x,y)*. The remaining detail and texture components of the image, which are quintessential for the visual and structural coherency, are taken over by the frequency feature maps, $F(u, v)$, as in the equation (3), which is gated at a varying range to manage the frequency feature utilization across enhancement. Here M and N are the image dimensions, while $e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$ is the complex exponential basis function.

As to reconstruct the frequency feature maps generated by DFT to the enhanced representation, the Inverse Fourier Transform (IFT) is used, as:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{4}$$

The GAN adversarial loss is utilized by both the generator and discriminator. As depicted by the equation (4), the generator accompanies generator loss to fool the discriminator with the pseudoimage produced.

$$L_{\text{GAN}}(G) = \mathbb{E}_{z\sim p_z} \left[\log(1 - D(G(z)))\right] \tag{5}$$

Where $D(z)$ is the pseudoimage produced by the generator, and $D(G(z))$ is the probability that the generated pseudo image is real, which is determined by the discriminator. As such, the discriminator accompanied the discriminator loss to maximize the values to improve the classification as real or fake for the generator-produced images.

$$L_{\text{GAN}}(D) = \mathbb{E}_{x\sim p_{\text{data}}} \left[\log D(x)\right] + \mathbb{E}_{z\sim p_z} \left[\log(1 - D(G(z)))\right] \tag{6}$$

As in the equation (6), if $D(z)$ is closer to 1 and $D(G(z))$ is closer to 0, then the discriminator is performing up to a reliable standard. Due to the critical importance of determining the perceptual similarity between the enhancement and the ground truth image, the VGG16-based perceptual loss is utilized in the generator component.

$$L_{\text{VGG}} = \sum_{i=1}^{L} \frac{1}{N_i} \left\|\phi_i(G(z)) - \phi_i(x)\right\|^2 \tag{7}$$

Where, $\phi_i(x)$ are the feature maps from the *i-th* VGG16 layer, and N_i is the number of elements in the accompanied layer, thereby assuring the perceptual similarity values for the enhancement put forward. In the final enhancement produced, the enhancement is a bi-product of adaptive brightness enhancement through inverse gray maps, refined detail and texture through Fourier techniques, and natural color profiles through spatial domain features, all undergoing discrimination at patch-level for a reliable and fine-grained enhancement.

IV. EXPERIMENTAL SETUP

A. Datasets

To evaluate the performance of the proposed AMGAN method, several benchmark datasets were utilized, ensuring comprehensive testing across diverse lighting conditions. The **LSRW** dataset comprises 5,650 low-light image pairs representing a diverse range of real-world low-light scenarios, with a split of 5,000 pairs for training, 600 pairs for validation, and 50 pairs for testing, thereby enhancing generalization across environments. The **LOLv1** dataset includes 485 image pairs for training, 100 pairs for validation, and 15 pairs for testing, which includes natural indoor and outdoor low-light scenes. The **LOLv2-real** dataset contains 689 pairs of image samples captured in real-world conditions spanning across various dark environments. Meanwhile, the **LOLv2-synthetic** dataset consists of 1,000 synthesized image pairs, divided into 800 for training, 100 for validation, and 100 for testing. These synthetic samples provide controlled benchmarking to evaluate the robustness of the model.

B. Implementation Details

The AMGAN method was implemented using PyTorch and Python 3.10 in Google Colab by leveraging the GPU capabilities with an NVIDIA Tesla T4 for accelerated computations with CUDA. The training process was conducted in two stages:

1) The CLSN was trained independently to generate the inverse grey map as the attention map for non-uniform illumination correction, with a learning rate of 5×10^{-4}, decay rate of 5×10^{-5} over 600 epochs.
2) The GAN was subsequently trained with the CLSN-produced image as the input to the generator, further while preserving the texture and naturalness, using cosine annealing with a learning rate of 2×10^{-3}, decay similarly over 600 epochs.

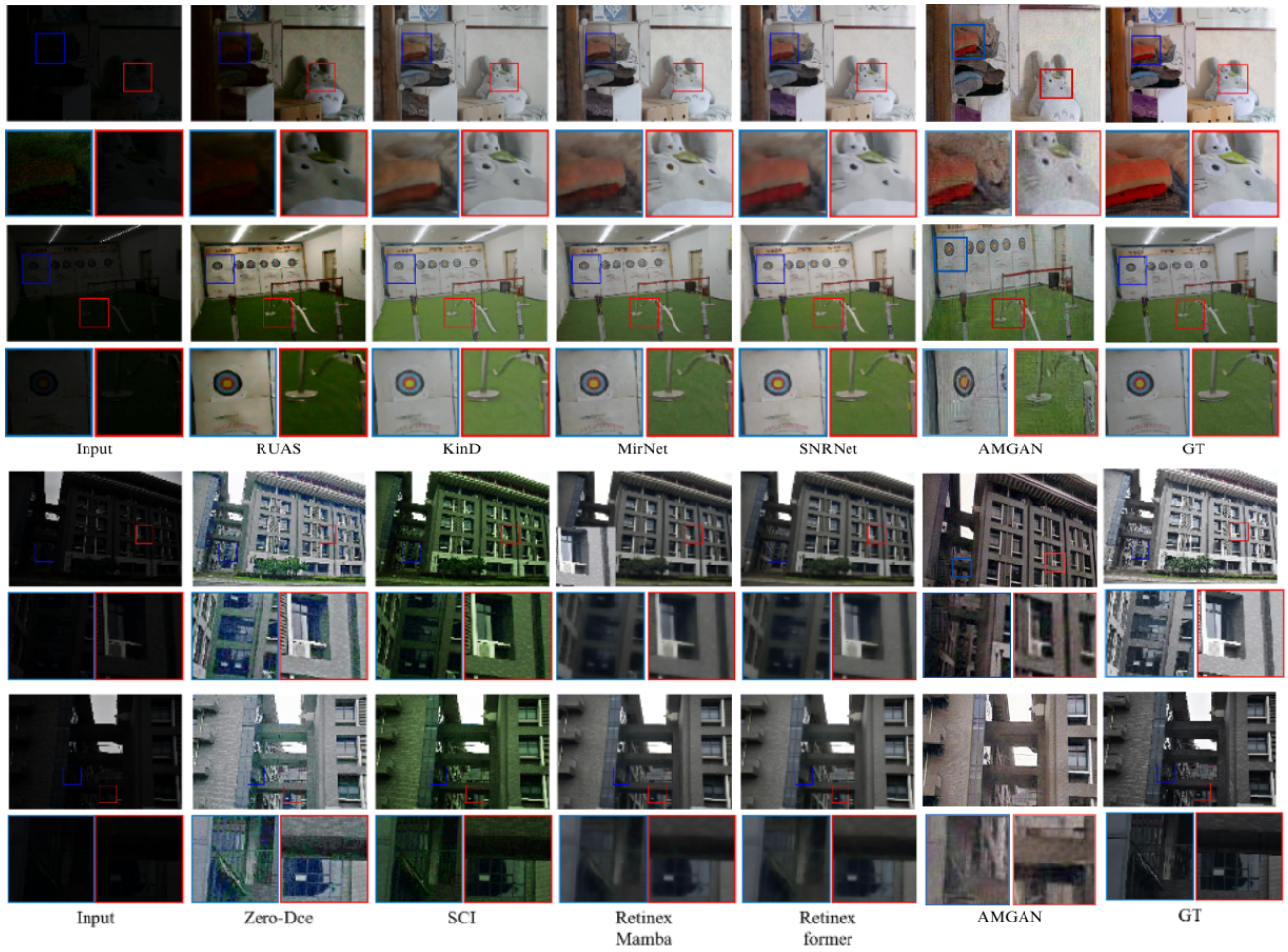

Fig. 2. Qualitative comparison of AMGAN against existing LLIE methods on LOLv1(top) and LSRW (bottom) datasets. The results are compared across different low-light scenarios, showcasing the enhanced brightness, texture, and contrast provided by AMGAN. Zoomed-in patches highlight detail recovery.

The Adam optimizer was utilized for both stages, with $\beta_1 = 0.5$ and $\beta_2 = 0.999$ for stable gradient updates.

V. Results

A. Quantitative Results

As in table I, proposed AMGAN method significantly outperforms traditional Retinex-based models while remaining highly competitive with state-of-the-art methods like DPEC and RetinexFormer, balancing computational efficiency and enhancement quality. On the LSRW dataset, AMGAN achieves 18.953 dB PSNR and 0.571 SSIM, significantly improving over RetinexNet (15.609 dB, 0.414 SSIM) and DRBN (16.734 dB, 0.507 SSIM) by reducing noise amplification and enhancing illumination consistency. Similarly, on LOLv1, AMGAN with scores of 22.851 dB PSNR and 0.801 SSIM surpasses RetinexNet (16.77 dB, 0.560 SSIM) and DRBN (20.13 dB, 0.830 SSIM), providing better brightness correction while minimizing artifacts. AMGAN competes closely with transformer-based models as RetinexFormer (25.16 dB, 0.854 SSIM) while maintaining lower computational overhead by avoiding self-attention mechanisms that are capable yet with higher complexity and inference time. On LOLv2-real, AMGAN (21.252 dB, 0.811 SSIM) performs comparably to RetinexMamba and SNR-Net, excelling in natural texture preservation and enhancement, where these methods struggle with color consistency.

B. Qualitative Evaluations

Fig. 2 compares AMGAN with state-of-the-art LLIE methods, including RUAS [24], KinD [11], MirNet [25], SNRNet, Zero-DCE [7], SCI [12], RetinexMamba [27], and RetinexFormer [28]. The results highlight differences in brightness correction, texture retention, noise suppression, and structural consistency. RUAS and RetinexMamba often introduce underexposure and loss of detail, while Zero-DCE and SCI are unable to meet color consistency. KinD and MirNet enhance brightness but cause a haze effect. RetinexFormer and RetinexMamba reduce noise but compromise the brightness factor. RUAS experiences under-enhancement and Zero-DCE distort object boundaries. KinD and SNRNet soften edges excessively, causing detail loss. AMGAN maintains well-defined contours without overprocessing. While Retinex-

Former achieve slightly higher scores, their high computational cost limits usage in real-time applications. AMGAN balances enhancement quality and efficiency, making it a practical choice for real-world application.

VI. Conclusion

This paper proposed AMGAN, an attention-driven multi-scale GAN reinforced with a complementary learning sub-network, for low-light image enhancement. By leveraging an inverse grey-map-based attention mechanism and multi-scale feature extraction utilizing frequency domain, AMGAN enables enhancement of brightness, preservation of fine textures, handles noise, and maintains structural integrity. Quantitative results demonstrate that AMGAN outperforms traditional Retinex-based methods while remaining competitive with state-of-the-art transformer-based approaches. Additionally, AMGAN achieves a balance between enhancement quality and computational efficiency, making it suitable for real-world applications. Qualitative evaluations confirm that AMGAN provides more natural brightness correction, increased texture retention, and robust noise balancing compared to existing LLIE methods. In conclusion, AMGAN offers a scalable and efficient solution without excessive processing overhead. Future work could focus on optimizing the architecture of the AMGAN method for further computational efficiency and exploring transformer-based hybrid enhancements, and also extending the framework to handle domain-specific tasks presents promising research directions.

References

[1] N. Yu, J. Li, and Z. Hua, "FLA-Net: multi-stage modular network for low-light image enhancement," *Vis. Comput.*, vol. 39, no. 4, pp. 1251–1270, 2023, doi: 10.1007/s00371-022-02402-8.

[2] Z. Tian et al., "A Survey of Deep Learning-Based Low-Light Image Enhancement," *Sensors*, vol. 23, no. 18, p. 7763, 2023, doi: 10.3390/s23187763.

[3] K. Shang, M. Shao, Y. Qiao, and H. Liu, "Frequency-aware network for low-light image enhancement," *Comput. Graph.*, vol. 118, pp. 210–219, 2024, doi: 10.1016/j.cag.2023.12.014.

[4] S. Hu and L. Li, "An attention mechanism and GAN based low-light image enhancement method," in A. Bhattacharjya and X. Feng, Eds., *SPIE*, Hangzhou, China, May 2023, p. 143. doi: 10.1117/12.2681113.

[5] J. Fu et al., "Low-light image enhancement base on brightness attention mechanism generative adversarial networks," *Multimed. Tools Appl.*, vol. 83, no. 4, pp. 10341–10365, 2024, doi: 10.1007/s11042-023-15815-x.

[6] X. Fang et al., "A non-uniform low-light image enhancement method with multi-scale attention transformer and luminance consistency loss," *Vis. Comput.*, 2024, doi: 10.1007/s00371-024-03452-w.

[7] C. Guo et al., "Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 1777–1786, doi: 10.1109/CVPR42600.2020.00185.

[8] Y. Jiang et al., "EnlightenGAN: Deep Light Enhancement Without Paired Supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021, doi: 10.1109/TIP.2021.3051462.

[9] G. Han et al., "Low-light images enhancement and denoising network based on unsupervised learning multi-stream feature modeling," *J. Vis. Commun. Image Represent.*, vol. 96, p. 103932, 2023, doi: 10.1016/j.jvcir.2023.103932.

[10] C. Wei et al., "Deep Retinex Decomposition for Low-Light Enhancement," *ArXiv*, 2018. [Online]. Available: https://www.semanticscholar.org/paper/Deep-Retinex-Decomposition-for-Low-Light-Wei-Wang/c7798bbbcd2e85593a2b2d37b92dabac63420cd8.

[11] Y. Zhang, J. Zhang, and X. Guo, "Kindling the Darkness: A Practical Low-light Image Enhancer," *IEEE Xplore*, 2019. [Online]. Available: http://arxiv.org/abs/1905.04161.

[12] L. Ma et al., "Toward fast, flexible, and robust low-light image enhancement," Jul. 2024. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2022/papers/Ma_Toward_Fast_Flexible_and_Robust_Low-Light_Image_Enhancement_CVPR_2022_paper.pdf.

[13] X. Guo, Y. Li, and H. Ling, "LIME: Low-Light Image Enhancement via Illumination Map Estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, 2017, doi: 10.1109/TIP.2016.2639450.

[14] X. Chen, J. Li, and Z. Hua, "Retinex low-light image enhancement network based on attention mechanism," *Multimed. Tools Appl.*, vol. 82, no. 3, pp. 4235–4255, 2023, doi: 10.1007/s11042-022-13411-z.

[15] X. Ren, M. Li, W.-H. Cheng, and J. Liu, "Joint Enhancement and Denoising Method via Sequential Decomposition," Apr. 28, 2018, *arXiv*, arXiv: 10.48550/arXiv.1804.08468.

[16] S. Hu, J. Yan, and D. Deng, "Contextual Information Aided Generative Adversarial Network for Low-Light Image Enhancement," *Electronics*, vol. 11, no. 1, p. 32, 2022, doi: 10.3390/electronics11010032.

[17] F. Lv, F. Lu, J. Wu, and C. Lim, "MBLLEN: Low-Light Image/Video Enhancement Using CNNs," presented at the *British Machine Vision Conference*, 2018. [Online]. Available: https://www.semanticscholar.org/paper/MBLLEN%3A-Low-Light-Image-Video-Enhancement-Using-Lv-Lu/70cb4bdd05cccc1f99cf690582e66b7637b81da7.

[18] Y. Wang and Z. Zhang, "Global attention retinex network for low light image enhancement," *J. Vis. Commun. Image Represent.*, vol. 92, no. 103795, p. 103795, 2023, doi: 10.1016/j.jvcir.2023.103795.

[19] G.-D. Fan, B. Fan, M. Gan, G.-Y. Chen, and C. L. P. Chen, "Multiscale low-light image enhancement network with illumination constraint," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 11, pp. 7403–7417, 2022, doi: 10.1109/tcsvt.2022.3186880.

[20] S. Yuan, J. Li, L. Ren, and Z. Chen, "Multi-Frequency Field Perception and Sparse Progressive Network for low-light image enhancement," *J. Vis. Commun. Image Represent.*, vol. 100, p. 104133, 2024, doi: 10.1016/j.jvcir.2024.104133.

[21] S. Huang, H. Dong, Y. Yang, Y. Wei, M. Ren, and S. Wang, "IATN: illumination-aware two-stage network for low-light image enhancement," *Signal Image Video Process.*, vol. 18, no. 4, pp. 3565–3575, 2024, doi: 10.1007/s11760-024-03021-7.

[22] S. Yang, D. Zhou, J. Cao, and Y. Guo, "LightingNet: An integrated learning method for low-light image enhancement," *IEEE Trans. Comput. Imaging*, vol. 9, pp. 29–42, 2023, doi: 10.1109/tci.2023.3240087.

[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 Trans. Image Process.*, vol. 30, pp. 3461–3473, 2021, doi: 10.1109/TIP.2021.3062184.

[24] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, "Retinex-Inspired Unrolling With Cooperative Prior Architecture Search for Low-Light Image Enhancement," presented at the *IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2021, pp. 10561–10570. Accessed: Feb. 02, 2025. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2021/html/Liu_Retinex-Inspired_Unrolling_With_Cooperative_Prior_Architecture_Search_for_Low-Light_Image_CVPR_2021_paper.html.

[25] S. W. Zamir et al., "Learning Enriched Features for Real Image Restoration and Enhancement," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds., Cham: Springer International Publishing, 2020, pp. 492–511. doi: 10.1007/978-3-030-58595-2$_3$0.

[26] X. Xu, R. Wang, C.-W. Fu, and J. Jia, "SNR-Aware Low-Light Image Enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 17714–17724. Accessed: Feb. 02, 2025.

[27] J. Bai, Y. Yin, Q. He, Y. Li, and X. Zhang, "Retinexmamba: Retinex-based Mamba for Low-light Image Enhancement," May 20, 2024, *arXiv:2405.03349*. doi: 10.48550/arXiv.2405.03349.

[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," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, Oct. 2023, pp. 12504–12513. [Online]. Available: http://arxiv.org/abs/2303.06705.

[29] S. Wang, Q. Lu, B. Peng, Y. Nie, and Q. Tao, "DPEC: Dual-Path Error Compensation Method for Enhanced Low-Light Image Clarity," Nov. 01, 2024, *arXiv:2407.09553*. doi: 10.48550/arXiv.2407.09553.