Low Light Image Enhancement Techniques

Ketki Patankar
University of California, San Diego
A-69032361
kpatankar@ucsd.edu

Varun Sankar Moparthi
University of California, San Diego
A-69030810
vmoparthi@ucsd.edu

Abstract

Research in low-light image enhancement has gained significant momentum in recent years due to the challenges posed by images with poor illumination, such as those affected by intense shadows, uneven lighting, backlighting, or taken at night. These images often suffer from issues like underexposure, noisy artifacts, low contrast and color distortion. Initially, traditional methods, including histogram equalization and gamma correction, were commonly used to address a few of these challenges. However, recent advancements have shifted towards deep learning techniques, that utilize Convolutional Neural Networks (CNNs) or Transformers to mitigate these drawbacks. In this paper, we first explore several state-of-the-art low-light image enhancement techniques, implementing and comparing them using Peak Signal-to-Noise Ratio (PSNR) as the evaluation metric. This analysis leads us to identify Retinexformer as the best-performing model. Next, we thoroughly examine the output of Retinexformer and uncover the presence of halo artifacts, particularly around sharp edges. The focus of this work then shifts towards investigating and implementing strategies to mitigate these halo effects, aiming to improve the overall quality of enhanced images.

1. Introduction

Low Light Imaging Enhancement (LLIE) is one of the popular and yet a challenging task in the area of Image processing and Computer vision, where it plays a significant role in many real time applications like; Night-time surveillance, Medical imaging and Autonomous driving. Images captured in low-light conditions typically suffer from several challenges. The light signal is minimal, resulting in poor visibility, while noise levels are elevated due to the low signal-to-noise ratio. Additionally, quantization errors introduced by the image sensor further degrade image quality. As a result, the image does not align with human perception. The Low light image enhancement techniques aim to brighten images captured in poorly lit environments, enhance the contrast, mitigate image degradation and remove image noise aiming to restore the original image. The LLIE methods proposed in some earlier work try to recover the low light images to match the human perception. However, these methods often rely on high-quality image capture techniques with minimal distortion which is not the case in real time images.

The Standard LLIE methods for image enhancement like Histogram Equalization and Gamma Correction, while improving the image contrast may amplify the noise. In the new era, limelight of Deep Learning(DL) has proposed plethora of solutions for low light image improvements using deep convolutional neural network (CNN). CNN-based low-light image enhancement methods can be broadly classified into two categories. The first category focuses on learning a direct mapping from low-light degraded images to well-lit ground truth images, without explicitly considering human perception. The second category, inspired by Retinex theory, utilizes multiple CNNs to separately process different image properties.

To overcome the challenges of multi-stage training observed in case of CNN based Retinex implementation, the RetinexFormer paper [4] proposes a single end-to-end model capable of capturing long-range dependencies and non-local similarities in an image. This addresses the limitations of CNNs in identifying these crucial patterns. A key example of this approach is the Transformer model, which has gained prominence for its ability to capture long-range dependencies in data using self-attention mechanisms. By combining the Retinex Theory and Transformer model the new Retinexformer model has been introduced [4]. This model consists of two parts. The first part, the One-Stage Retinex-Based Framework (ORF), incorporates Retinex theory and performs two tasks: image decomposition into reflectance and illumination, followed by noise suppression and correction of under/overexposure and color distortion. The second part uses an Illumination-Guided Transformer (IGT), which captures long-range dependencies.

1.1. Problem Statement

The goal of this project is to explore and analyze various state-of-the-art low-light image enhancement techniques. These methods aim to improve the contrast and denoise low-light images, which are typically characterized by poor visibility, low contrast, and significant noise, particularly in the darker regions. The enhancement technique should, therefore, not only brighten the image but also preserve important details while reducing noise, resulting in a cleaner, more informative output that maintains a natural and realistic appearance.

General Mathematical Objective:
$\arg\min_{I_e} \lambda_1 L_{contrast}(I_e, I_{true}) + \lambda_2 L_{denoise}(I_e, I_{true})$

Here I_e refers to enhanced output image, I_{true} refers to ground truth image, λ_1, λ_2 are weights given to loss functions of contrasts and de-noising.

At the outset, the primary objective was to identify the most effective existing model for enhancing low-light images. After conducting an extensive comparison of several models, the best-performing model, Retinexformer, was selected as the baseline for further analysis. Although Retinexformer delivered promising results, certain artifacts, particularly the halo effect around sharp edges, were observed in the enhanced images. Consequently, the focus shifted to investigating the causes of these halo artifacts and exploring potential strategies for mitigating them.

2. Previous Methods

When addressing low-contrast problems, traditional techniques such as histogram equalization and gamma correction are commonly used for image enhancement [1]. Gamma correction, a non-linear gray-level transformation, adjusts an image's brightness and contrast by modifying the mapping between input and output pixel values through a gamma parameter. Histogram equalization, on the other hand improves global contrast by redistributing pixel intensity values evenly across the entire image. However, this can lead to over-enhancement in bright regions and a loss of detail in darker areas.

To mitigate these issues, methods like Adaptive Histogram Equalization (AHE) and Contrast Limited Adaptive Histogram Equalization (CLAHE) have been developed [2]. These techniques apply local histogram equalization on smaller image regions, or tiles. CLAHE improves upon AHE by adding a clipping threshold to prevent over-amplification, limiting contrast enhancement to avoid excessive brightening. While CLAHE outperforms AHE, it can still struggle with highly non-uniform lighting. Additionally, both methods may introduce color artifacts and fail to recover pure black areas, requiring sophisticated approaches to handle extreme lighting or restore true blacks.

The Retinex theory offers a more robust solution for handling such non-uniform lighting. This model is capable of recognizing color and brightness independently of lighting variations. The core idea behind Retinex is that an image can be decomposed into two components: reflectance and illumination. While reflectance (the inherent color and texture of objects) remains relatively constant, illumination (the lighting affecting the scene) varies smoothly across the image. Retinex models the illumination as a smooth, spatially varying component, often approximated using a Gaussian filter. By estimating the illumination across the image and then normalizing the reflectance, Retinex allows for consistent color perception, even under changing or uneven lighting conditions.

Inspired by the idea of image decomposition and enhancement, a low-frequency-based approach for low-light image enhancement [3] is proposed, driven by two key observations. First, the low-frequency component of an image contains more structural information and is less susceptible to noise compared to the high-frequency component. This makes enhancing the low-frequency layer of a noisy low-light image easier than enhancing the entire image. Second, image primitives such as edges, lines, and corners have low dimensionality and can be represented with smaller examples. By leveraging low-frequency information, networks can more easily reconstruct the high-frequency details. Based on this, the proposed model consists of two main parts: an Attention to Context Encoding Module (ACE), which adaptively decomposes and enhances the low and high-frequency layers, followed by a Cross-Domain Transformation Block (CDT), which suppresses noise and enhances image details.

A more advanced image enhancement technique, DiffLight [5], uses a dual-branch pipeline to improve low-light images. One branch focuses on denoising the image using a Diffusion model and enhancing brightness using LEDNet framework. The Diffusion model works by gradually adding noise to an image and training a model to reverse this process, and hence effectively cleaning up the image. The second branch utilizes a UNet transformer with full self-attention mechanism, which helps capture both global and local dependencies within the image and map distinct regions in image. Additionally, to address the blocky artifacts often seen in patch-based methods, DiffLight incorporates Progressive Patch Fusion (PPF), which smoothly blends patch boundaries, ensuring a more seamless and natural-looking result.

In this paper, we implemented four state-of-the-art methods for preliminary results: Contrast Limited Adaptive Histogram Equalization (CLAHE), the ACE and CDT based Frequency Decomposition method, LLFormer unit

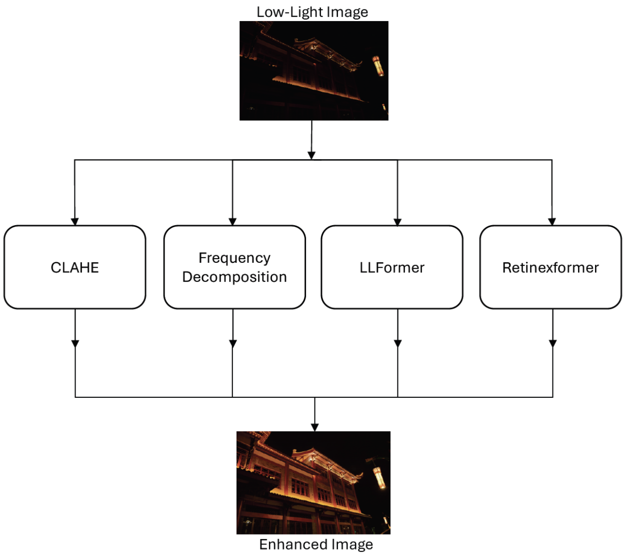

Figure 1. Preliminary Implemented models

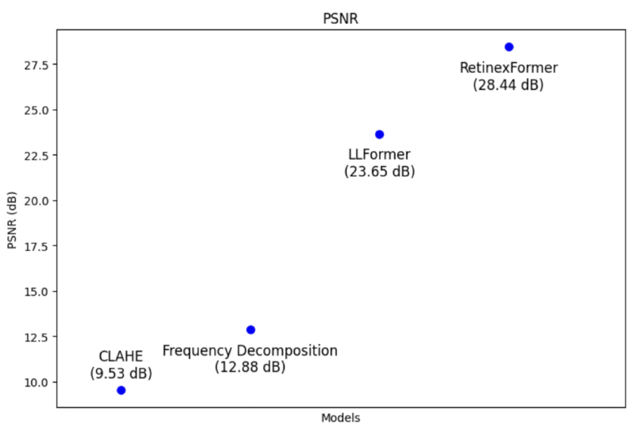

Figure 2. Model Performance (PSNR)

of DiffLight and Retinexformer as shown in Figure 1

Each model was tested on 15 images, and their performance was evaluated using the PSNR metric. The figure 2 above shows the performance of each model, with Retinexformer yielding the best results. Based on its superior performance, Retinexformer was selected as the baseline for further analysis and enhancement.

2.1. Theory

2.1.1. Retinex Theory

It aims to explain and model how humans perceive the color and light under varying illumination conditions. The main idea is to dissolve the image into two components, Reflectance(R) and Illumination(L).

$$I(x, y) = R(x, y)L(x, y)$$

$$log(I(x, y)) = log(R(x, y)) + log(L(x, y))$$

The primary goal is to estimate the reflectance (R) by compensating for the illumination (L). Since illumination changes smoothly while reflectance remains constant, a spatial filtering operation (such as a Gaussian or low-pass filter) is applied to remove high-frequency components from the image. This results in an estimate of the illumination (L), which is then used to recover the reflectance (R').

$$L(x, y) \approx I(x, y) \circledast G(x, y))$$

$$R^{'}(x, y) = exp^{log(I(x,y)-log(I(x,y)\circledast G(x,y))}$$

It has different variants like Single-Scale Retinex (SSR) which is the simplest form of original algorithm, Processes the image at a single scale by using a logarithmic function to enhance local contrast and simulate color constancy. It subtracts a low-pass filter from the original image to model lighting variations; Multi-Scale Retinex (MSR) which is improved upon SSR. It applies the Retinex model at multiple scales (different resolutions), combining the results to improve both fine details and overall contrast, which is effective for complex images with varying light intensities and details. Multi-Scale Retinex with Color Restoration (MSRCR) extends MSR by adding a color restoration step that corrects color shifts during image enhancement, preserving natural color tones, applied vivdly where both color fidelity and contrast enhancement are critical. Deep Retinex(DeepLearning based Retinex) uses deep neural networks to automatically learn optimal parameters for color constancy and lighting correction, improving image quality under complex conditions, usually applied in low-light photo enhancement, and complex visual scenes.

2.1.2. Transformer

The Transformer deep learning model leverages a novel mechanism called self-attention, which allows it to process sequences of data in parallel, enabling much more efficient training and handling long-range dependencies.

The self-attention mechanism allows each token to interact with every other token in the input image. Unlike traditional sequence models, which process tokens sequentially, self-attention computes relationships between all pairs of tokens simultaneously by giving attention scores. It is calculated as following query (Q), key (K), and value (V)

$$Attention\_weight(Q, K, V) = softmax(\frac{QK^T}{\sqrt{dk}})V$$

In practice, multiple attentions are used to allow the

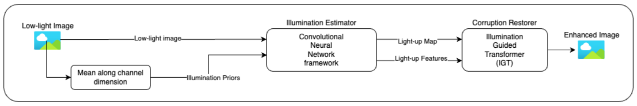

Figure 3. Block Diagram of Retinexformer

model to focus on different parts of the input image simultaneously, using multiple heads, the model is able to capture different aspects of the input image in parallel, providing richer representations.

Transformer has the encoder-decoder architecture, where encoder part has L layers of combination of Multi-head attention and feed forward networks and decoder part has L layers but has an additional multi-head attention sub-layer to attend to the encoder's output. Combined architecture will look like this,

$$Enc\_Out = Multihead(Self\_Attention(X), X)$$

$$Dec\_Out = Multihead(Self\_Attention(Y), Enc\_Out)$$

2.1.3. Retinexformer

The Retinexformer framework is a single-stage Retinex-based approach integrated with an Illuminance-guided transformer designed to effectively remove noise/corruption from images.

While it is clear from the below equation that Retinex theory assumes that the image I(x,y) is corruption free and perfectly represent the scene, practically, low-light images often contain noise due to the low signal-to-noise ratio (SNR), and color distortions caused by uneven lighting or sensor limitations. RetinexFormer addresses this by incorporating perturbation terms into the low-light image model as shown below.

$$I(x,y) = (R(x,y) + R^{'}(x,y))\,(L(x,y) + L^{'}(x,y))$$

Additonally, the conventional retinex theory focuses on removing noise from the Reflectance and completely overlooks the estimation errors that could come up in the illumination map. All of these corruptions are modelled well in the Retinexformer model as in Figure 3.

Furthermore, by incorporating a Transformer, the model leverages attention mechanisms, which compare each pixel with every other pixel and assign attention scores. This determines how much focus a pixel should have on other pixels when generating its representation, specifically the illuminance, since in this we are using illumination-guided Transformer model. This facilitates us to interpret relationship between far away regions within the image that may influence each other. Similarly we are able to model the similarity in structure or textures that appear at distant locations in the image.

Figure 4. Low Light Image of buildings

Figure 5. Retinexformer generated output

2.2. Preliminary Results

Analysis of low-light images and their output generated by RetinexTransformer model reveals the superiority of the algorithm. It is clear from Figures 4 and 5 that this method is capable of improving image contrast without over-saturating the brighter regions in the image.

On further analysis of the model output as shown in figure 6, 7, we were able to identify certain recurrent issues like the appearance of Halo artifacts at multiple edges in the image, color bleeding, and local object blurring especially visible around small objects like leaves. These issues are highlighted in the images given below.

2.3. Halo Artifact

Halo's in images, project themselves as a collection of successively decreasing intensities that appear around sharp edges. Essentially, these halo's appear in two different forms:

CVPR
#Group-10

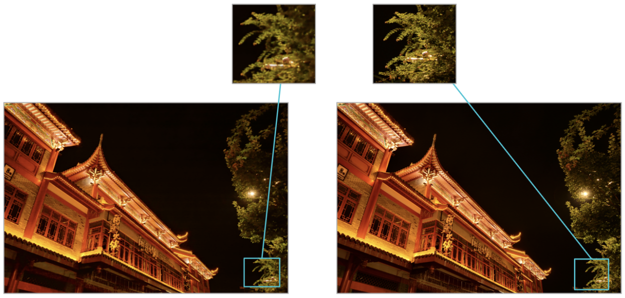

Figure 6. Object blurring in the model generated output

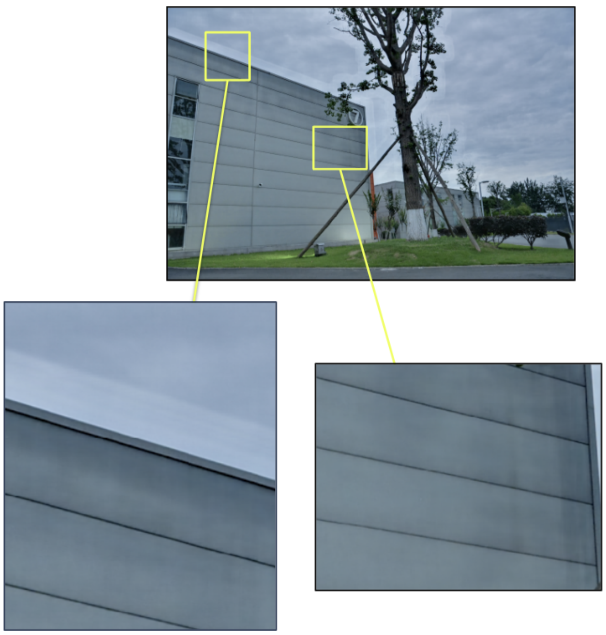

Figure 7. Halo combined with color bleeding

- Positive Halo: Appears as a bright band of progressively changing intensities at edges between high and low-contrast regions.
- Negative Halo: Appears as a darker band of progressively changing intensities at edges between high and low-contrast regions.

2.3.1. Cause of Halo Effect

A few potential reasons for the generation of Halo artifacts are listed below:

- In context to the case of Retinex framework, this artifact could be a result of an issue with Illumination and Reflectance separation.
- Inappropriate use of attention mechanism where there is an excessive focus on boundaries
- Due to local contrast adjustment in retinex where the contrast between neighbouring pixels is increased and causes over-exaggeration of transitions.

Figure 8. Sub-bands from Wavelet Decomposition

2.4. Exploratory Analysis(Cons of our approach)

2.4.1. Halo artifact not detected as edges

Since halo projects itself as a collection of successively reducing intensity, it should appear in the frequency domain in the form of high-frequency components. In order to analyze this, we plotted the four sub-bands of the wavelet decomposition of our input image. As a result, we observed that not all the halo's visible in the image were captured as high-frequency components. This can be confirmed in the figure 8.

Due to this we limited ourselves to non-frequency decomposition based method for halo detection and removal.

2.4.2. Clouds captured in Halo mask

Our approach used manual thresholding to detect halo regions. This selected intensity range was also shared by a large area of the cloud region in the image, due to which the halo mask extended into the cloud regions. Whatever filtering operation was carried out on the halo, also affected this cloud area equivalently. To avoid this, we tried to differentiate the cloud and halo region using their texture. Ideally, the clouds would have a smoother texture when compared to the sharp edges that appear in case of halo. For this, we used the texture descriptor called Local Binary Patterns. The algorithm works by comparing each pixel with its neighbouring pixels that are defined by a grid of a selected size (default 3X3). If the neighbouring pixel has a greater intensity than the center pixel, it is assigned a value 1. In figure 9 we can see that there are certain regions in the clouds that also have sharp transitions and share the same texture as that of the halo. Hence, we could not go ahead with this method on differentiating the cloud from our halo mask. This is the reason why our final model output suffers with certain amount of data loss in the cloud regions.

Figure 9. Local Binary Patterns - Texture Descriptor

Figure 10. Halo Mask

3. Proposed Methods

3.1. Approach 1: Halo mask creation and neighbourhood-based filtering

After a process of trial and error, we successfully determined the optimal upper and lower thresholds for detecting the halo in the image. Using these thresholds, we generated a binary mask that highlights the halo region, as shown in the figure 10.

We then perform Neighbourhood-based filtering in this halo region, and update each pixel in it, with its neighbourhood mean. In our case, the neighbourhood size chosen is 70. For each pixel we take 70 pixels above, below and on its left and right and average over their intensity values. An important point to note here is that, we have also set a lower limit on which intensities present in this grid will be picked up. This is done to avoid the issue of occurrence of a dark black border at certain edges where our central pixel is surrounded by dark regions like trees.

Figure 11. Mask for removing white patches

3.2. Approach 2: Combination of approach one and anisotropic diffusion

As an extension to approach one, we performed Anisotropic diffusion on its filtered output. Anisotropic diffusion, also called perona-malik diffusion, is originally used for preserving edges while performing the diffusion on low-gradient regions. It works by determining the gradient of the image in the x and y directions to detect the horizontal and vertical edges, respectively. The edge strength is then computed in the form of a gradient magnitude. This is used to determine the diffusion coefficient, which is a Gaussian function applied to the gradient magnitude of the image, scaled by a parameter kappa. The pixel values are then updated according to the gradient values, with the update weights determined by the diffusion coefficient.

In our approach, we have performed 50 iterations of this diffusion process which helps gradually smoothen the areas around the edges. There are two parameters that play an important role in the diffusion process. These are the kappa and gamma values. Kappa defines how sensitive our edges will be to diffusion. In our case, we have set a high kappa value of 40. Gamma parameter controls the speed of diffusion, it tells us how much each pixel changes in each iteration. We have set gamma to a low value of 0.05 so that the diffusion is slow and more controlled.

In the output of this approach, we observed certain white patches in the image. To mitigate this issue, we performed a second round of manual thresholding and neighbourhood-based filtering. The mask we used for this is shown in Figure: 11.

4. Results

The results of both the approaches are highlighted below. We can clearly see the reduction in the halo around the edges as we go from one approach to another.

Figure 12. Gray scale image of model output with halo

Figure 13. Output image for approach 1

Figure 14. Model output for approach 2 - Diffused image

Figure 15. Approach 2 output after patch removal

5. Summary

In this study, four state-of-the-art models for low-light image enhancement were implemented. Among them, RetinexFormer exhibited the best PSNR and delivered superior qualitative results compared to other decomposition-based methods. However, some issues, such as halo artifacts, were observed in the enhanced images produced by the model. To address this, a post-processing method was proposed for detecting and removing these halo artifacts. The best results were achieved by combining Halo Mask-based Neighborhood Filtering with the Anisotropic Diffusion technique, effectively reducing the presence of halo artifacts in the enhanced outputs.

5.1. Future Scope

The enhancement of low light image quality through deep learning models encounters challenges like halo artifacts and color bleeding. By directly incorporating techniques to handle these issues within the model's architecture, the system can become more efficient and robust.

5.1.1. Incorporating halo reduction techniques in the model

The halo reduction techniques discussed are currently applied as post-processing steps to the model output. However, these techniques or ideas could be integrated directly into the model's architecture, making it more robust and generalized to handle halo artifacts during the Retinexformer image enhancement process itself.

5.1.2. Addressing other issues in deep learning models

Other issues observed in some images were Color Bleeding and Local Object Blurring, which could be effectively addressed within the deep learning framework by incorporating techniques to mitigate them directly into the model.

6. Contributions

The tasks, like the literature review, implementation of state-of-the-art methods, development and implementation of the proposed method, and miscellaneous activities, were equally divided among the team members. Each individual made significant contributions toward achieving the project goals.

Name	Contributed Percentage
Varun Sankar Moparthi	50%
Ketki Patankar	50%

References

[1] P. P. Anoop and R. Deivanathan, “Advancements in low light image enhancement techniques and recent applications,” Journal of Visual Communication and Image Representation, vol. 103, p. 104223, Aug. 2024, doi: 10.1016/j.jvcir.2024.104223.

[2] Y. Li, Y. Luo, Y. Zheng, G. Liu, and J. Gong, “Research on Target Image Classification in Low-Light Night Vision,” Entropy, vol. 26, no. 10, Art. no. 10, Oct. 2024, doi: 10.3390/e26100882.

[3] K. Xu, X. Yang, B. Yin, and R. W. H. Lau, “Learning to Restore Low-Light Images via Decomposition-and-Enhancement,” in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA: IEEE, Jun. 2020, pp. 2278–2287. doi: 10.1109/CVPR42600.2020.00235.

[4] 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 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France: IEEE, Oct. 2023, pp. 12470–12479. doi: 10.1109/ICCV51070.2023.01149.

[5] Y. Feng et al., “DiffLight: Integrating Content and Detail for Low-light Image Enhancement,” in 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA: IEEE, Jun. 2024, pp. 6143–6152. doi: 10.1109/CVPRW63382.2024.00619.

[6] P. Perona and J. Malik, ”Scale-space and edge detection using anisotropic diffusion”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629-639, 1990.

[7] Digital Image Processing by Rafael C. Gonzalez and Richard E. Woods, 4thedition, Pearson.