Northeastern University

EECE 5639 – Computer Vision

Professor: David Rosen

Project Report

April 24, 2024

Low-Light Image Enhancement

Palaniappan Yeagappan (NUID: 002846668)

Shiva Kumar Dhandapani (NUID: 002808148)

Balasubramaniam Renganathan (NUID: 002876644)

Gayathri Suresh (NUID: 002820711)

Low-Light Image Enhancement

Palaniappan Yeagappan[1] Shiva Kumar Dhandapani[2] Balasubramaniam Renganathan [3] Gayathri Suresh[4]
{yeagappan.p, dhandapani.s, renganathan.b, suresh.gay } @northeastern.edu
College of Engineering, Northeastern University, Boston, MA 02115, USA

***Abstract*—This study presents the implementation and analysis of a method termed Low-Light Image Enhancement. The method aims to enhance the transformative potential of end-to-end deep learning in low-light image noise reduction, drawing inspiration from recent advancements in the field. Utilizing convolutional neural networks, the focus is on leveraging single neural networks to contribute to the development of streamlined and efficient solutions in low-light image processing. This contribution seeks to advance technologies such as image noise reduction and face recognition in low-light environments. The primary objective of this project is to implement an algorithm capable of processing and enhancing images seamlessly, with a particular emphasis on real-time processing. The performance evaluation of the system involves the utilization of various publicly available datasets to ensure its adaptability across diverse low-light conditions by also improving image viewability. This report defines the structural composition of the system, analyzes the integration processes, discusses encountered challenges, and examines the implications of the findings for the progression of computer vision technology.**

Keywords: Signal-to-Noise Ratio (SNR), Stacked Sparse Denoising Autoencoders (SSDA), Trainable Nonlinear Reaction-Diffusion (TNRD), Multi-Layer Perceptrons (MLP), Deep Autoencoders (DA), Mean Square Error (MSE), Fully Convolutional Network (FCN), Peak Signal-to-Noise Ratio (PSNR), Convolutional Neural Networks (CNN), Ground Truth (GT), Original Image (OI), and Processed Image (PI)

I. INTRODUCTION

The low-light image enhancement technique is a significant achievement in the image enhancement field, with the potential to revolutionize image processing systems by merging many stages of the low-light image enhancement processes into a single neural network. Conventional methods for reducing noise typically follow a structured approach, which includes several steps like extracting features, identifying phonemes, and transcribing words. These procedures consist of various stages, like white balancing, demosaicing, denoising, and color correction.

One significant limitation of the conventional approach is that it amplifies noise during the image enhancement process. However, the introduction of end-to-end deep learning represents a significant improvement over this traditional strategy; the method employs convolutional neural networks, which enable the integration of the whole process into a single neural network model, simplifying the workflow. It provides a transformational approach, allowing for direct mapping of raw signals to image data, along with intermediary feature extraction stages, and also reducing the noise that is amplified in the pipeline process. The change is also very beneficial in other relevant fields, particularly in image noise reduction and face recognition. Furthermore, this technique has the potential for better noise reduction performance than the multi-step pipeline approach.

The primary goal of the project is to implement the Learning to See in the Dark [2] paper from scratch and integrate it with the traditional pipeline for CNN in an attempt to boost accuracy and minimize PSNR value. Previous studies have employed various methods, such as SSDA, TNRD, MLP, DA, and CNN. These methods have shown promising outcomes in image noise reduction, which intrigues us enough to learn more about this field.

The project aims to study and analyze a low-light image enhancement technique that is powerful enough to tackle various low-light environments, that challenge traditional image enhancement methods. The methodology involves a multi-faceted approach:

1) **Pre-Processing Pipeline:** This methodology uses a pre-processing pipeline to enhance image quality through steps like white balance correction, demosaicing, denoising, and sharpening, which enhance edges and high-frequency components for better clarity and detail.

2) **CNN Architecture:** The CNN enhances low-light images by interpreting high-to-low-exposure relationships. The processing pipeline in CNN includes Bayer format conversion, 4 channel packing, normalization, L1 loss function optimization, Adam optimizer refinement, data augmentation, and learning rate adjustments.

3) **Post-Processing and Training Refinement:** This image enhancement technique uses CNN-based transformations and advanced post-processing methods to enhance low-light images. It includes multidimensional color transformation, blind noise suppression, and sRGB color space standard adjustments for high-quality images.

II. RELATED WORK

In the field of computational photography and image processing, numerous techniques have been proposed to deal with the difficulties involved in improving low-light images. Chen et al. (2018) [2] presented a technique that uses CNN to directly improve low-light images from raw data. Significant improvements in noise reduction and color transformation were achieved by training a CNN on pairs of low-light raw photographs and equivalent well-exposed images, demonstrating the promise of deep learning for enhancing low-light images.

Zheng et al. (2021) [1] presented a hybrid learning framework for single image brightening in low-light conditions. Their approach efficiently preserves details in high-light sections while boosting low-light areas by combining model-driven and data-driven strategies. Refinement using a residual CNN produced better results than conventional image processing techniques, suggesting a viable way to improve the brightness and contrast of low-light images.

Retinex-Net is a deep learning system proposed by Wei et al. (2018) [3] that is particularly made for low-light improvement. The authors showed notable gains in low-light image quality by utilizing Decom-Net to break down images into reflectance and illumination components, and then Enhance-Net to adjust the illumination component. Their research demonstrates the advantages of combining data-driven and model-driven methods into a single, cohesive framework to tackle the challenges of low-light photography.

Tao et al. (2017) [4] proposed LLCNN, a CNN-based method designed for improving images in low light. Through the use of multiscale feature maps and the use of structural similarity (SSIM) loss, LLCNN was able to maintain textures and features during the enhancement process while effectively addressing gradient vanishing difficulties. Their approach performed better than earlier methods, offering important new information to the low-light image processing community.

To improve low-light images, Guo et al. (2017) [5] proposed LIME, which focuses on pixel-wise illumination map refinement. By estimating illumination maps and optimizing them with a structural prior, LIME outperformed previous techniques in terms of both quality and efficiency. This method has the potential to improve low-light images and boost the effectiveness of multimedia and computer vision algorithms later on.

These referenced works collectively represent significant contributions to the advancement of low-light image enhancement techniques. From deep learning-based approaches to hybrid frameworks that integrate model-driven and data-driven methods, each method offers unique insights and solutions to the challenges faced in computational photography and image processing. These contributions serve as valuable resources for further research and development in the field.

III. Project Outline

A. *Problem Statement*

Low light conditions still makes it difficult to take high-quality images, even with major advances in digital photographic technology. These challenges include compromised SNR and reduced visibility, which can hinder the interpretation of critical information in various applications such as autonomous driving, surveillance, and medical imaging. Conventional image processing pipelines and algorithms often struggle to effectively enhance low-light images while preserving important details and minimizing noise

B. *Objective*

The objective of this project is to implement and analyze the efficiency and accuracy of the low-light image enhancement technique and study how it improves visibility and perception in low-light conditions. The approach used aims to improve the performance of existing algorithms, such as the one proposed by Chen et al. [2], by leveraging a combination of model-driven and data-driven methods. The project seeks to address the limitations of conventional methodologies by employing end-to-end training of a FCN using raw short-exposure low-light photos and their long-exposure equivalents. By integrating spatial-temporal information and optimizing network architecture, the goal is to enhance the interpretability of complex low-light patterns and significantly improve image quality.

C. *Evaluation Metrics*

The performance of this low-light image enhancement technique will be evaluated using Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR). These metrics will quantify the overall image distortion, noise reduction, and preservation of image characteristics, respectively.

1) Mean Square Error (MSE):
Mean Square Error quantifies the average squared difference between the processed image and the ground truth, providing a measure of overall image distortion.

$$MSE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|^2 \tag{1}$$

where:

- N is the total number of pixels.
- y_i represents the pixel value in the ground truth image.
- $\hat{y}_i$ represents the pixel value in the processed image.

2) Peak Signal-to-Noise Ratio ($PSNR$):
Peak Signal-to-Noise Ratio measures the peak error between the processed image and the ground truth. Higher PSNR values indicate better image quality, especially in terms of noise reduction and signal fidelity.

$$PSNR = 10 \log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{2}$$

where:

- MAX is the maximum possible pixel value of the image (255 if represented in 8 bits).
- MSE is the mean squared error between the processed and ground truth images.

IV. METHODOLOGY

This methodology consists of a series of pre-processing pipelines, CNN design, and post-processing refinement phases in an effort to improve low-light photographs. Through the use of a combination of deep learning and conventional image processing techniques, each step is intended to maximize image quality.

A. *Pre-Processing Pipeline:*

The pre-processing phase of this methodology begins with the raw input data, which is transformed through a series of steps to improve image quality whilst making it prepared to undergo further analysis. This stage is built around traditional image processing modules like white balance correction, demosaicing, denoising, and sharpening, which use mathematical formulations to improve image quality.

Fig. 1: Tradiational Pipeline in the preprocessing pipeline

White balance correction is essential for accurate color representation since it neutralizes any color cast in the image. In mathematical terms, this method adjusts the intensities of different color channels to restore real color fidelity, as expressed by the formula:

$$I_{cor}(x, y) = I(x, y) * R \tag{3}$$

where $I_{cor}(x, y)$ is the corrected intensity at pixel (x, y), $I(x, y)$ is the original intensity, and R is the color correction factor.

Demosaicing is the process of interpolating color information from raw sensor data to recreate a full-color image. This approach involves evaluating nearby pixel intensities to fill in missing color values, as seen in the mathematical expression:

$$I_{dem}(x, y) = f(I_{raw}(x, y), I_{raw}(x \pm 1, y), I_{raw}(x, y \pm 1)) \tag{4}$$

where $I_{dem}(x, y)$ is demosaiced intensity at pixel (x, y), $I_{raw}(x, y)$ is the raw intensity values, and f is the demosaicing function.

Denoising techniques are employed to reduce noise in images acquired in low-light conditions while maintaining crucial features in the image. Gaussian filtering decreases noise by convolving the picture with a Gaussian kernel, as shown by the following formula:

$$I_{den}(x, y) = \frac{1}{\sum_{i=-k}^{k} \sum_{j=-k}^{k} w(i, j) * I(x + i, y + j)} \tag{5}$$

where $I_{den}(x, y)$ is denoised intensity at pixel (x, y), $w(i, j)$ is the Gaussian kernel weights, and K is the kernel size.

Sharpening modules enhance edges and high-frequency components to improve image clarity and detail. The Laplacian sharpening method, frequently paired with Gaussian smoothing, improves visual contrast and detail, as shown in the following mathematical expression:

$$I_{sh}(x, y) = I_{sm}(x, y) + k * (I(x, y) - I_{sm}(x, y)) \tag{6}$$

$I_{sh}(x, y)$ is the sharpened intensity at pixel (x, y), $I_{sm}(x, y)$ is the smoothed intensity, and k is the sharpening factor.

B. *Convolutional Neural Network (CNN) Architecture:*

The Convolutional Neural Network (CNN), a complex framework meticulously implemented to handle pre-processed low-light pictures and provide high-quality augmented outputs, is fundamental to our technique. The architecture of the CNN is designed to interpret complex mappings between high- and low-exposure pictures, which allows it to perform very well when boosting low-light imaging.

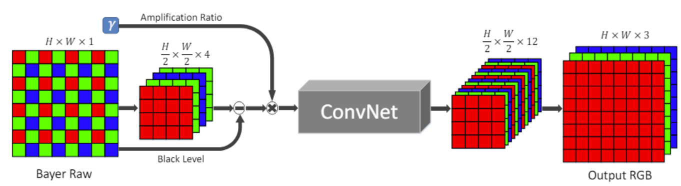


Fig. 2: Structure of the Convolutional Neural Network by Chen et al. [2]

During the training phase, the image data from the pre-processed images is loaded, and each image is converted into Bayer format and packed into 4 channels to facilitate processing. Normalization is also applied to ensure pixel values are within a standardized range (between 0 and 1). The L1 loss function, a fundamental component in the field of image processing, is used to rigorously optimize the CNN. This loss function assesses the absolute discrepancy between anticipated and ground truth pictures, causing the network to converge and iteratively refine its parameters. The CNN, with the Adam optimizer, navigates the complex landscape of image enhancement, refining its ability to detect small differences and improve picture quality.

Additionally, a variety of data augmentation techniques are used during training to improve CNN's robustness and flexibility in a variety of settings. These methods, which include rotation, flipping, and random cropping, provide variation to the training dataset and strengthen CNN's ability to generalize and perform well in practical settings. The CNN's training trajectory is carefully controlled by adjusting the learning rate, which guarantees optimal convergence and performance across training epochs.

C. *Post-Processing and Training Refinement:*

After undergoing CNN's transformational processing, the improved images go through one further round of refining to bring them up to normal sRGB color space standards. This post-processing work takes a multidimensional approach, including color transformation and blind noise suppression techniques, to further improve the color integrity and clarity of the images.

Post-processing takes the form of complex mathematical processes like sub-pixel layer adjustments that are aimed at restoring the images to their original resolution. This step refines pixel values and color channels by the application of statistical techniques and mathematical formulas, resulting in a smooth integration with the sRGB color space.

In Summary, this technique is a balanced combination of conventional knowledge and state-of-the-art techniques, designed to improve low-light images with good accuracy and improved quality.

V. Results and Observations

The implementation of Learning to See in the Dark [2] is studied, and the logic is re-implemented from scratch with the integration of the traditional pipeline into the pre-processing pipeline in an attempt to increase accuracy and reduce PSNR value.

A. Setup

The original implementation by Chen et al. [2] is done using an Intel i7 CPU and at least 64 GB of RAM, but due to the limitations, we have used an Intel i7 CPU and 24 GB of RAM. Because of this, we were not able to train the model with larger datasets; instead, we reduced the size to 1/4 of the original dataset and trained it for 100 epochs instead of 2000 epochs.

Various Python libraries like Rawpy 0.10, Scipy, Tensorflow, Seaborn, Scikit-Learn, etc. are used for implementing neural networks, reading raw images from digital cameras, and various numerical and scientific calculations.

The setup ensures that the program can efficiently utilize hardware resources, such as GPU acceleration, while also maintaining compatibility with the required software dependencies.

B. Observation

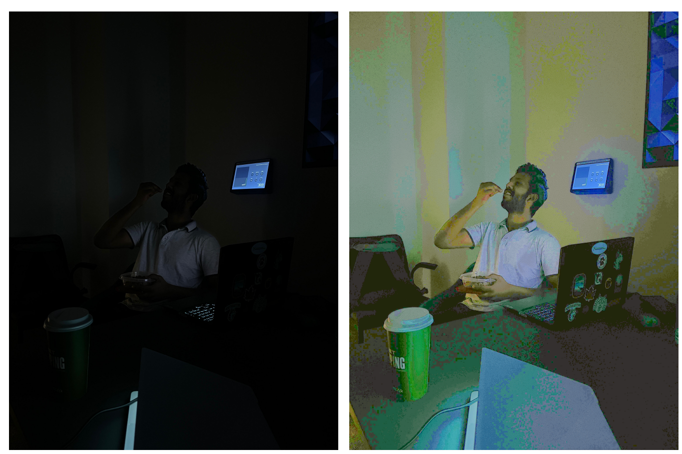
Upon comparing the images (with reference to the values of Ground truth, Original Image and Processed Image) reveals an illustration of this low-light image-enhancing algorithm's capabilities. The ground truth (GT) image shows a picture-perfect outdoor scene that is well-lit and over-exposed and captures a wide range of textures and colors with excellent clarity and fidelity. This GT image serves as a benchmark for evaluating how well the low-light improvement works.

With a compressed dynamic range and a severe loss of apparent detail, the original image (OI), which captures the essence of the difficulties associated with low-light photography, presents a scene that is almost undetectable. The scene's finer details are hidden by the dim illumination, which is underexposed, making the image unusable for either practical or artistic reasons.

When the enhancement algorithm is used, the processed image (PI) shows a remarkable change. The algorithm successfully recovers a significant quantity of information from the depths of darkness, restoring a visibility level that is close to the GT. The definition of the image has been greatly improved. Additionally, color rendition has a revival, with the PI displaying a refreshed palette better suited to the natural environment.

(a) Original Image (b) Processed Image (c) Ground Truth

Fig. 3: Enhanced Low-Light Image with Ground truth with Sony SID Dataset



(a) Original Image (b) Processed Image

Fig. 4: Image captured by the team using iPhone 13

The difference between processed images from a professional camera-equipped public dataset (*Fig. 3*) and those captured by our team using a smartphone (*Fig. 4*) is significant. Our analysis aimed to discern disparities in image quality, characteristics, and processing techniques. Professional camera images exhibit superior quality with high resolution, clarity, and color accuracy, while smartphone captures show reduced sharpness and color fidelity, along with exposure fluctuations and occasional blurring. The differences stem from inherent limitations in smartphone camera sensors and processing capabilities. Professional camera images undergo extensive post-processing for optimal enhancement, while smartphone images rely on automated algorithms for basic adjustments. Understanding these differences is crucial for informed decision-making in image selection and utilization across various fields. Future research could focus on bridging the quality gap between professional and smartphone captures, enhancing the applicability of smartphone photography in professional settings.

(a) Original Image (b) Processed Image (c) Ground Truth

Fig. 5: Output using Chen at el. algorithm [2]

The comparative analysis between Chen's algorithm and the implemented algorithm for low-light image enhancement unfolds a narrative of approches and their resultant imagery. The ground truth (GT) represents an impeccably lit scenario, serving as the baseline for natural color and detail.

Chen et al.'s [2] algorithm's processed image (PI) evidences a sophisticated approach to enhancement, with a combined effort to achieve luminance balance and color accuracy. It successfully reveals the obscured details of the original image (OI), reviving textures and hues that resonate with the GT. The execution of brightness correction maintains the scene's inherent contrast, avoiding over-saturation and retaining the natural ambiance. This suggests a precise calibration of enhancement parameters, resulting in a restoration of visibility that does not compromise the natural aesthetic of the environment.

On the other hand, our algorithm takes a more assertive stance in its enhancement. The images processed through this approach exhibit a marked increase in luminosity and a pronounced color saturation, diverging from the GT. Although visibility is significantly improved, the resulting images lean towards an over-enhanced spectrum, occasionally introducing noise and compromising the natural subtlety of the scene's original lighting conditions. This is because of training the model with fewer epochs and a small dataset. If the model was trained with the same amount of epochs and dataset, it would have revealed a perfect comparison between both algorithms.

The comparative PSNR evaluation reveals that Chen et al.'s algorithm outperforms the implemented algorithm with a notable margin, achieving a PSNR of 28.88 against 21.47. This quantitative measure suggests that Chen et al.'s approach yields an image that is closer to the original in terms of error minimization and noise reduction.

Conclusively, Chen's algorithm stands out for its ability to enhance underexposed photography to a level of refinement that aligns closely with human visual perception. The nuanced application ensures that the enhancement process meticulously balances the restoration of detail with the preservation of the scene's color integrity.

	PSNR
Chen et al.	28.88
Implemented Algorithm	21.47

TABLE I: Peak Signal to Noise Ratio comparison

VI. Conclusion

In this report, our study implemented and evaluated a low-light image enhancement model that can seamlessly improve the image quality while simultaneously trying to reduce the noise of the low-light images and assessing its performance in different low-light environments. Additionally, our work optimized the model's performance and also proved that the accuracy of pipeline processing can be improved by employing convolutional neural networks along with traditional low-light image enhancement pipelines. This integration significantly improves the robustness of the system.

Significant effort was dedicated to experimenting with diverse datasets and analyzing their performance. Comparisons with recent research provided insights into low-light image enhancement advancements. We acknowledged persistent challenges and highlighted potential future developments. The analysis conducted offers valuable insights and indicates future research directions for improving the model's effectiveness.

References

[1] C. Zheng, Z. Li, Y. Yang, and S. Wu, "Single image brightening via multi-scale exposure fusion with hybrid learning," IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 4, pp. 1425–1435, 2021

[2] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3291–3300, 2018

[3] . Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," 2018

[4] L. Tao, C. Zhu, G. Xiang, Y. Li, H. Jia, and X. Xie, "LLCNN: A Convolutional Neural Network for Low-Light image enhancement," in 2017 IEEE Visual Communications and Image Processing (VCIP), pp. 1–4, 2017

[5] Xiaojie Guo, Yu Li, and Haibin Ling. "LIME: Low-Light Image Enhancement via Illumination Map Estimation". In: Trans. Img. Proc. 26.2 (Feb. 2017), pp. 982–993. issn: 1057-7149. doi: 10 . 1109 / TIP.2016.2639450