MONOCHROME AUGMENTED LOW-LIGHTIMAGE ENHANCEMENT

Arnav Shah
SRM Institute of Science & Technology
Kattankulathur, Tamil Nadu, India
ak0339@srmist.edu.in

Yash Kulkarni
SRM Institute of Science & Technology
Kattankulathur, Tamil Nadu, India
ys2400@srmist.edu.in

Abheelash Mishra
SRM Institute of Science & Technology
Kattankulathur, Tamil Nadu, India
ap2907@srmist.edu.in

Vaishnavi
SRM Institute of Science & Technology
Kattankulathur, Tamil Nadu, India
vv9426@srmist.edu.in

Sharat S Embrandiri
Samsung Research Institute
Bengaluru, India
sharat.e@samsung.com

Subin K B
Samsung Research Institute
Bengaluru, India
subin.kb@samsung.com

Ankur Jain
Samsung Research Institute
Bengaluru, India
jain.ankur@samsung.com

Dr. S. Babu
Associate Professor
SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu,India
Babus@srmist.edu.in

Dr. R. Rajmohan
Assistant Professor
SRM Institute of Science & Technology
Kattankulathur, Tamil Nadu, India
rajmohar@srmist.edu.in

ABSTRACT

Low light short exposure photography is challenging, but an important factor in capturing images in temporarily dynamic scenes avoiding unwanted effects such as ghosting, motion blur, camera shakes, image artifacts, etc. Monochrome augmented low-light image enhancement aims to get improved low-light short-exposure images by using an additional monochrome sensor and its data. Monochrome images typically possess a higher SNR (Signal-to-Noise Ratio) and better luma information, since it avoids the attenuation by the Bayer Filter. The objective here is to develop a deep learning based approach to enhance low-light short exposure images from the main sensor by using an additional low resolution monochrome sensor.

1. Introduction

In recent years, the surge in digital imaging devices has completely transformed how we capture and hold onto life's moments, regardless of lighting conditions. Yet, the persistent challenge of low-light situations continues to cast a shadow over the quality of our captured images, dulling their clarity and vibrancy. When faced with these less-than-ideal conditions, traditional methods of enhancing images often stumble, prompting us to seek out innovative approaches to overcome these limitations.

Low-light images suffer from inherent limitations that compromise their visual quality and appeal. The reduced availability of light in low-light environments results in diminished visibility and clarity, obscuring fine details and textures within the scene. Moreover, the heightened levels of noise and graininess in low-light images detract from their sharpness and crispness, introducing distracting artifacts that degrade overall image quality. Additionally, the limited color fidelity inherent in low-light conditions leads to inaccuracies or desaturation in captured colors, diminishing the vibrancy and realism of the scene. Challenges in handling the wide dynamic range present in low-light scenarios further exacerbate these issues, often resulting in loss of detail in both highlights and shadows. Furthermore, the prolonged exposure times required in low-light settings increase the risk of motion blur, further compromising image sharpness and clarity. Collectively, these factors contribute to the perceived inferiority of low-light images compared to those captured in well-lit conditions, highlighting the importance of effective low-light image enhancement techniques.

This research aims to tackle the shortcomings that plague low-light image quality head-on. Our drive stems from recognizing that conventional enhancement techniques struggle when it comes to mitigating the adverse effects of challenging lighting conditions. These struggles manifest in noticeable reductions in image clarity, color fidelity, and an unwelcome increase in distracting noise.

While existing literature acknowledges the importance of improving low-light image quality, current methodologies fall short of providing a comprehensive solution. We're missing a solution that can effectively balance the preservation of essential visual details while also minimizing the distractions caused by noise in low-light environments.

This study takes a departure from traditional methods by tapping into the inherent advantages of monochrome information to enhance images. Our approach is all about boosting the overall quality and perceptual fidelity of low-light images by seamlessly integrating monochrome features with low light images. We're banking on the idea that monochrome channels harbor valuable visual information even in the face of challenging lighting conditions, giving us a unique opportunity to elevate the overall visual appeal of the resulting images.

Throughout this paper, we'll walk you through how we conceptualized, designed, and implemented our project. We'll conduct a thorough evaluation of its effectiveness using a variety of datasets that cover various low-light scenarios. We'll compare our results with those of prevailing state-of-the-art methodologies to showcase the unique contributions of our approach. Ultimately, the findings of this research not only add to the growing body of knowledge in image processing but also provide a practical solution to enhance the visual quality of images captured in demanding lighting conditions.

2. Related Work

Low light image enhancement has been addressed in several works. One approach is the use of convolutional neural networks (CNNs) for image enhancement. DEANet, proposed by Wen Yu et al., combines frequency and content information of images using three subnetworks: decomposition, enhancement, and adjustment networks **[1]**. Another CNN-based method is proposed by Xueyan Zhou et al., which decomposes the illumination of the input image into high-frequency and low-frequency components and adjusts the illumination iteratively until the final distribution is uniform **[2]**. Additionally, a novel architecture called MIRNet, presented by the authors in Context_3, maintains spatially precise high-resolution representations while capturing strong contextual information from low-resolution representations. These works demonstrate improved performance in enhancing low light images compared to existing methods **[3]**.

An interesting approach proposed by Zi-ang Li et al. **[4]**, addresses the challenge by introducing a novel approach utilizing a teacher-student framework for efficient knowledge transfer between networks, leveraging distillation loss based on attention maps. Additionally, it presents a gradient-guided low-light image enhancement network, comprising enhancement and gradient branches that collaboratively preserve structural information. Experimental validation demonstrates the method's superiority over existing techniques, highlighting its potential for practical low-light image enhancement applications. Sergei et al. addresses the pertinent issue of enhancing low-contrast monochrome images in an automatic and computationally efficient manner **[5]**. The central focus is on the adaptive enhancement of integral contrast in complex monochrome images through the application of nonlinear statistical non-inertial transformations. This approach aims to improve the overall quality of images characterized by low contrast, small-sized objects, and non-uniform illumination. A notable contribution of the paper is the exploration of the effectiveness of histogram-based methods in enhancing the integral contrast of intricate images. Particularly, the focus is on images with low-contrast details and challenging lighting conditions. To evaluate the proposed techniques, a comparative analysis is conducted, which involves the assessment of no reference methods for evaluating generalized image contrast. This evaluation leverages histogram-based metrics and expert assessments to gauge the method's efficacy, advancing the field of adaptive contrast enhancement for complex monochrome images, particularly in automated settings.

Numerous studies have aimed to enhance the efficiency of low-light image enhancement in the raw domain. In pursuit of a computationally expedient low-light enhancement system, researchers in **[6]** introduced a lightweight architecture (RED) tailored for extreme low-light image restoration. Additionally, they devised an amplifier module to estimate the amplification factor derived from the input raw image.

Fangjin Liu et al. **[7]** used datasets such as the Vasileios Vonikakis (VV), Multi-exposure image fusion (MEF), DICM and low-light dataset (LOL) to employ a method based on recursive networks for low-light image enhancement. Recursive networks are a type of deep learning architecture that can learn hierarchical features and patterns from images, making them suitable for image enhancement tasks. The research paper proposed by Feifan Lv et al. **[8]** explores the topic of enhancing low-light images using attention-guided methods for low-light image enhancement. The method described in the statement employs a multi-branch convolutional neural network along with attention maps and a reinforcement network for enhancing low-light images. To train this method, a synthetic dataset is created with carefully designed low-light simulations, making it larger and more diverse than existing datasets. The approach uses two attention maps: one to identify underexposed areas and well-lit regions and another to differentiate noise from actual textures. The multi-branch enhancement network adapts to the input image, and a reinforcement network further enhances color and contrast. Extensive experiments on various datasets demonstrate that this method produces high-quality results for low-light image enhancement, surpassing current state-of-the-art methods in both quantitative and visual evaluations.

3. Methodology

In conducting this research, our overarching approach centered on addressing the challenges inherent in low-light image enhancement through the

Figure 1: Monochrome and low-light images captured

development of this project. The primary objective of our study was to propose a novel methodology that leverages monochrome information to augment the enhancement process, thereby improving the visual quality and perceptual fidelity of low-light images.

To achieve this objective, our methodology was designed to encompass several key steps, each tailored to contribute to the overall research goal. The methodology was crafted to facilitate data collection, model architecture design, training, and evaluation processes, thereby enabling a systematic and comprehensive exploration of the proposed approach.

By structuring our methodology around these essential components, we aimed to ensure a rigorous and methodical approach to addressing the research problem. Each component in this project was meticulously crafted to support these objectives, from the preprocessing of the dataset to the training and evaluation of the proposed model.

3.1. Data Collection

For this study, data collection involved gathering three types of images: monochrome images, low-light images, and corresponding ground truth image. The dataset consisted of a curated selection of images representing a range of low-light conditions. Monochrome images, which served as the primary input for the proposed enhancer, and low-light images were either sourced from image datasets available in the public domain or were custom-made. Ground truth images, depicting the same scenes as the low-light images but captured under normal lighting conditions, were also included to facilitate evaluation and comparison.

To acquire the necessary data for our project, we captured over 440+ sets of images. We also utilized open-source datasets such as the LOL (LOw-Light) dataset to increase the amount of data we can train the model on.

3.2. Data Preprocessing

Data preprocessing is a critical phase in our research endeavor, aimed at optimizing the input data to facilitate effective model training and evaluation. In this section, we detail the preprocessing steps applied to the raw input images and ground truth images, which are essential for enhancing the quality and alignment of the dataset.

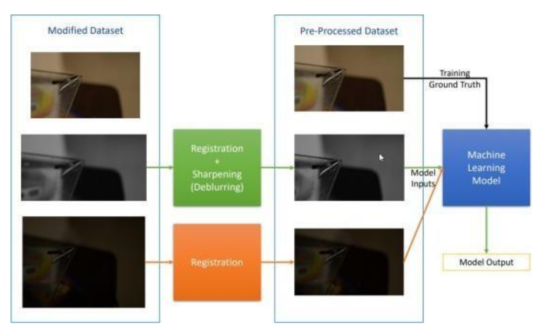

Figure 2: Data preprocessing pipeline architecture

The initial preprocessing step involves feature registration, a pivotal process to ensure accurate alignment between the monochrome input images and their corresponding ground truth color images. To achieve this, we employ the Scale-Invariant Feature Transform (SIFT) algorithm.

This algorithm identifies key points and descriptors in both images, enabling the computation of a transformation matrix using the RANdom SAmple Consensus (RANSAC) algorithm.

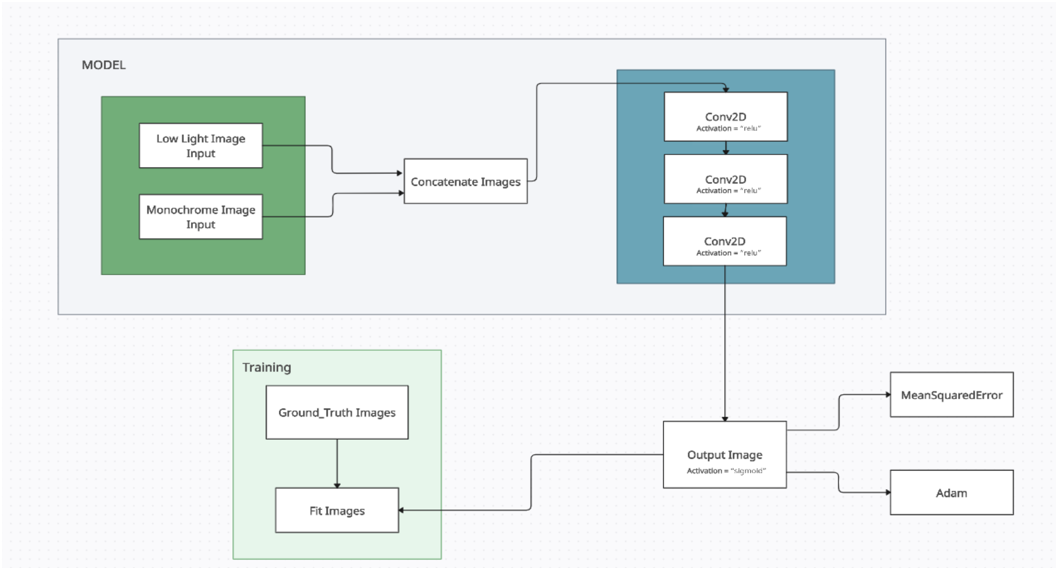

Figure 3: Architecture of the model

RANSAC is a robust estimation method used to fit models to data points in the presence of outliers. It is particularly useful in computer vision, image processing, and geometric modeling tasks where the data may be corrupted by noise or outliers. The basic principle of RANSAC involves randomly selecting a subset of data points, fitting a model to these points, and then testing this model against the remaining data points. Data points that are consistent with the model (inliers) are identified, and the model is refined using only these inliers. This process is repeated for a specified number of iterations, with the final model being the one that has the largest number of inliers. Subsequently, this transformation matrix is utilized to register the features between the monochrome input image and the ground truth color image, ensuring precise alignment for effective model training. Following the feature registration process, additional enhancement operations are performed on the monochrome input images to augment their quality before feeding them into the model. These enhancement techniques include applying Gaussian blur to mitigate noise and improve overall image quality. Furthermore, a weighted blending operation is employed to enhance contrast and emphasize details in the monochrome images. These preprocessing steps aim to enrich the visual information available to the model during training, thereby enhancing its ability to accurately enhance low-light images while preserving essential details and minimizing noise artifacts.

Throughout the preprocessing pipeline, robust error handling mechanisms are implemented to address any potential issues that may arise during image registration or enhancement operations. These mechanisms ensure the smooth execution of the preprocessing pipeline, even in the presence of unforeseen challenges, thereby maintaining the integrity and reliability of the dataset. By meticulously implementing these preprocessing steps, we ensure that the input data fed into our model is appropriately aligned, enhanced, and optimized for training. This not only enhances the robustness and effectiveness of the model but also improves its ability to accurately enhance low-light images, ultimately contributing to the advancement of low-light image enhancement techniques.

3.3. Model Training

Model training represents the cornerstone of our research efforts, where we utilize convolutional neural networks (CNNs) to learn the intricate mappings between low-light input images and their corresponding ground truth representations.

This section delves into the nuanced methodologies and intricacies underlying the model training process, elucidating the steps taken to ensure the robustness and efficacy of our low-light image enhancement model.

3.3.1. Checkpoint Callback Configuration

The essence of model training lies not only in the optimization of neural network parameters but also in monitoring the model's progress and preserving the best-performing configurations. To this end, we configure a 'ModelCheckpoint' callback, a versatile tool provided by the TensorFlow framework. This callback serves a dual purpose:

- ➢ First, it periodically saves the model weights during training, enabling us to capture snapshots of the model's evolving state.
- ➢ Second, it facilitates the selection of the best-performing model weights based on a specified monitoring metric.

By designating the loss function as the metric to monitor, we aim to capture the model configurations that minimize the Mean Squared Error (MSE) between predicted and ground truth images. Additionally, we specify the frequency of saving checkpoints (after each epoch) to ensure a comprehensive record of the model's trajectory throughout training.

3.3.2. Model Training Procedure

The crux of model training lies in the iterative optimization of the model's parameters to minimize the disparity between predicted and ground truth images. Leveraging TensorFlow's high-level API, we invoke the 'fit()' method on the constructed CNN model, kickstarting the training process. During each training iteration, the model ingests batches of training data comprising monochrome images, their corresponding low-light counterparts, and ground truth images.

These input-output pairs serve as the basis for updating the model's weights via backpropagation, guided by the Adam optimizer. Notably, we opt for a batch size of 1 to facilitate fine-grained adjustments to the model parameters, thereby enhancing convergence and mitigating potential overfitting. The training unfolds over a predetermined number of epochs, with each epoch representing a complete pass through the entire training dataset. Furthermore, to gauge the model's generalization capabilities and guard against overfitting, we allocate a portion of the training data (20%) for validation purposes. This enables real-time evaluation of the model's performance on unseen data, providing valuable insights into its efficacy and generalization prowess.

Upon the completion of the training process, the trained model encapsulates the acquired feature representations and optimized parameters. This trained model is capable of seamlessly enhancing low-light images with newfound clarity and fidelity.

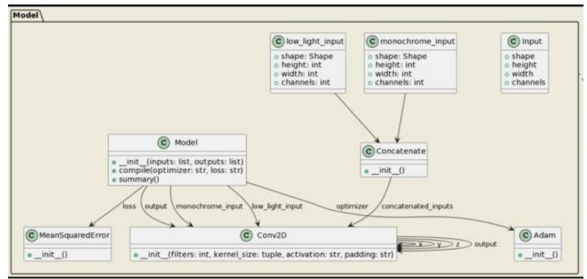

Figure 4(a): Class diagram of the model

Figure 4(b): Class Diagram of Training

4. Results

The images generated through the proposed approach undergo a rigorous training process, leveraging monochrome input images, low-light counterparts, and ground truth images. In order to prevent overfitting and guarantee robustness, the model's parameters are fine-tuned during training. The trained model is capable of seamlessly enhancing low-light images

When evaluating the enhancement of low-light images, the proposed method exhibits superior performance when compared to alternative models. The study makes use of established methodologies, including DEANet, recursive networks, and attention-guided methods, to illustrate the distinctive advantages of the proposed method. The model's efficacy in enhancing the visual quality and perceptual fidelity of low-light images is bolstered by the employing of monochrome information and the rigorous training procedure. Furthermore, the study underscores the significance of confronting the inherent difficulties associated with improving images in low light and offers a pragmatic resolution for enhancing images captured under harsh lighting circumstances.

In comparison to established methodologies, the proposed approach yields outputs that demonstrate enhanced low-light image quality, structural similarity, and minimal pixel-wise variations. These results demonstrate the efficacy of the deep learning-based approach in this regard.

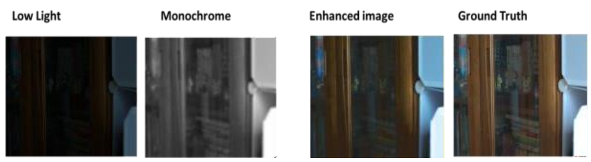

Figure 5(a): Input, Output & Expected output comparison

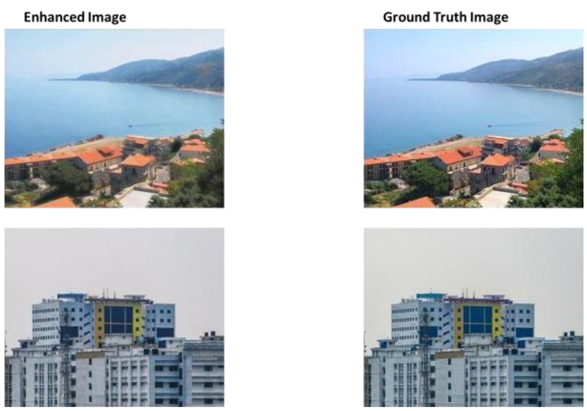

Figure 5(b): Output & Expected output comparison

Three metrics are used to assess the quality of an image: Mean Squared Error (MSE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR).

In our case, the MSE is 0.0090, which is relatively low. This suggests that the pixel values between the ground truth and the output image are quite similar, as MSE measures the average squared difference between corresponding pixels of these two images.

The SSIM, a metric that considers luminance, contrast, and structure for a more comprehensive measure of image quality, is 0.7890. This value is close to 1, indicating a high level of structural similarity between the ground truth and the output image.

Lastly, the PSNR, which measures the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation, is 74.88. This is a relatively high value, suggesting good image quality.

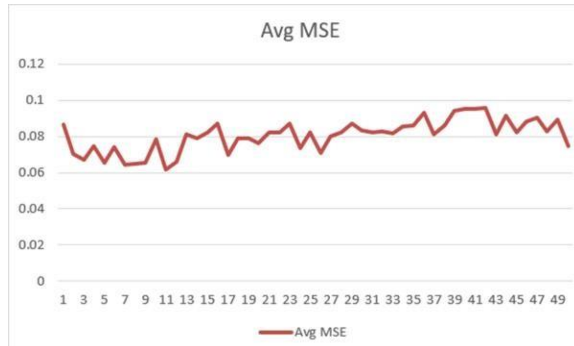

Figure 5(c): Average MSE graph

Figure 5(d): Average SSIM graph

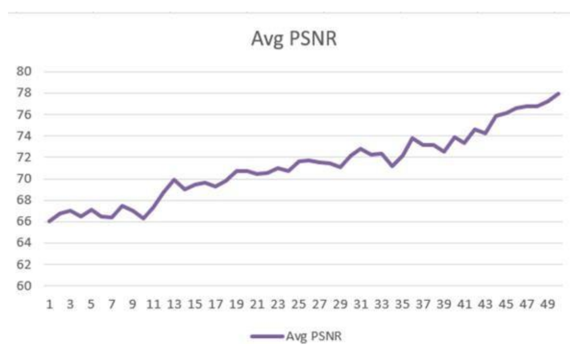

Figure 5(e): Average PSNR graph

In summary, the low MSE indicates low pixel-wise differences, the high SSIM indicates high structural similarity, and the high PSNR suggests good image quality.

5. Conclusion

The suggested approach, which uses monochrome information to improve low-light images, showed encouraging outcomes in terms of enhancing perceptual fidelity and visual quality. The study offers a workable answer to the problems associated with low-light image enhancement and adds to the expanding corpus of knowledge in image processing. The approach aims to balance the preservation of important visual details while minimizing noise-induced distractions in low-light environments by integrating monochrome features with low-light images. The study employs a comprehensive research methodology that includes data collection, preprocessing, model training, and evaluation. The use of robust error handling mechanisms and open-source datasets is emphasized. Metrics like Mean Squared Error (MSE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR) show that the results have low pixel-wise differences, high structural similarity, and good overall image quality, demonstrating the efficacy of the suggested method in enhancing low-light images.

The study has certain limitations, such as the need for more investigation and improvement of the suggested methodology to deal with particular issues like motion blur and color fidelity in dimly lit environments. Further areas for improvement could be the model's scalability to accommodate larger datasets and the approach's generalization to a wider range of low-light scenarios.

In order to improve the robustness and efficacy of the low-light image enhancement model, future work may entail the development of more sophisticated preprocessing techniques to further optimize the input data for model training as well as the investigation of additional deep learning architectures and algorithms. Moreover, the incorporation of real-time image processing functionalities and the modification of the approach to suit various imaging devices and environments may represent beneficial avenues for further investigation.

References

[1] Zhihao Wang, Yifan Zhang, and Xiaoyu Zhang. DEANet: Decomposition Enhancement and Adjustment Network for Low-Light Image Enhancement.

[2] Wen Yu, Liquan Zhao, Tie Zhong. Unsupervised Low-Light Image Enhancement Based on Generative Adversarial Network.

[3] MD Sana Meharaj, Esampelli Malathi, Chittumalla Keerthana, Poosala Harshitha and Dasari Mahesh Kumar. Low Light Image Enhancement using MIRNet.

[4] Zi-ang Li, Yuehuan Wang and Jinpu Zhang. Low-light image enhancement with knowledge distillation.

[5] Sergei Yelmanov and Yuriy Romanyshyn. Automatic Enhancement of Low-Contrast Monochrome Images.

[6] Mohit Lamba and Kaushik Mitra. Restoring extremely dark images in real time.

[7] Fangjin Liu, Zhen Hua, Jinjiang Li and Linwei Fan. Low-Light Image Enhancement Network Based on Recursive Network.

[8] Feifan Lv, Yu Li and Feng Lu. Attention Guided Low-Light Image Enhancement with a Large Scale Low-Light Simulation Dataset.