School

of

Electronics and Communication Engineering

Mini Project Report

on

ENHANCEMENT OF LOW LIGHT IMAGES

By:

1. **Chinmayee Mandi** 01FE19BEC188
2. **Amogh Joshi** 01FE19BEC199
3. **Sampada Malagi** 01FE19BEC244
4. **Palavi Kamat** 01FE19BEC032

Semester: V, 2020-2021

Under the Guidance of

Ujwala Patil

K.L.E SOCIETY'S
KLE Technological University,
HUBBALLI-580031
2020-2021

SCHOOL OF ELECTRONICS AND COMMUNICATION ENGINEERING

CERTIFICATE

This is to certify that project entitled **" Your Project Title "** is a bonafide work carried out by the student team of **"Chinmayee Mandi (01FE19BEC188) , Amogh Joshi (01FE19BEC199) , Sampada Malagi (01FE19BEC244) , Palavi Kamat (01FE19BEC032)"**. The project report has been approved as it satisfies the requirements with respect to the mini project work prescribed by the university curriculum for BE (V Semester) in School of Electronics and Communication Engineering of KLE Technological University for the academic year 2021-2022

Ujwala Patil **Guide**	**Nalini C. Iyer** **Head of School**	**N. H. Ayachit** **Registrar**

External Viva:

Name of Examiners **Signature with date**

1.

2.

ACKNOWLEDGMENT

The project Low light image enhancement becomes successful with the kind support and help of many individuals. We the team under Centre of Excellence in Visual Intelligence[CEVI] extend our sincere thanks to each one respectfully. The team expresses deepest gratitude to all the people who have assisted us in the process of completion of this project. All their contributions are deeply appreciated and acknowledged. The team would like to place on record our deep sense of respect to Nalini C Iyer, Head of the Department of School of Electronics and Communication for giving the opportunity to extend our skills in the directions of this project . The team expresses heartfelt gratitude to our guide Uma Mudenagudi, whose valuable insights proved to be vital in contributing the success of this project.

By:

Project Team

ABSTRACT

In this project, we present a method for the enhancement of images been captured in low-light conditions. The images captured in low-light conditions become uneasy for human eye visualization as human eye night vision is limited in comparison with other animals. In real-time application of a low-light image enhancement method, the basic expectations like fast computation, memory efficiency, visually appealing enhanced image, etc are to be met. The existing methods concentrate on the accuracy of enhancement at the cost of computational speed. Hence we propose a deep learning model which ensures that all the expected aspects are well-considered and met respectively. This model assures 30 percent speed-up without disturbing the enhancement quality where it is trained using raw low-light images and the results are ultra-high definition 4K resolution RGB images. This computation is achieved in just 1 second on a CPU and at 32 fps on a GPU. The model produced gives the best results without any fine-tuning done during the testing phase as well. We can have a generalized application for the model as object detection.

Contents

List of Figures

Chapter 1

Introduction

Images captured in outdoor scenes can be highly degraded due to poor lighting conditions. These images can have low dynamic ranges with high noise levels that can affect the overall performance of the computer vision algorithms. To make computer vision algorithms robust in low light conditions, we use image enhancement to improve the visibility of the images.

1.1 Motivation

Images captured in low-light conditions are unpleasing to human vision, the enhancement of such images helps the human society have a beautiful vision of images without any external light source requirement for capturing of images. This being the basic motivation, we urge to work on applications of this concept to help the needy at the root level and also where possible.

1.2 Objectives

1. Experiment on available learning based architectures towards Low light image enhancement.
2. Develop a learning based technique for enhancement of low light paired images of high ISO, low ISO and high exposure, low exposure.
3.Propose a learning based architecture to enhance the images captured in low light condition.
4.Demonstrate the results of the proposed methodology using the available data-set and compare the results with state-of-the-art methods

1.3 Literature survey towards low light image enhancement

1.Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement(CVPR-2020)[2]

A Deep Curve Estimation Network (DCE-Net) is devised to estimate a set of best-fitting Light-Enhancement curves (LE-curves) given an input image. The framework then maps all pixels of the input's RGB channels by applying the curves iteratively for obtaining the final enhanced image.

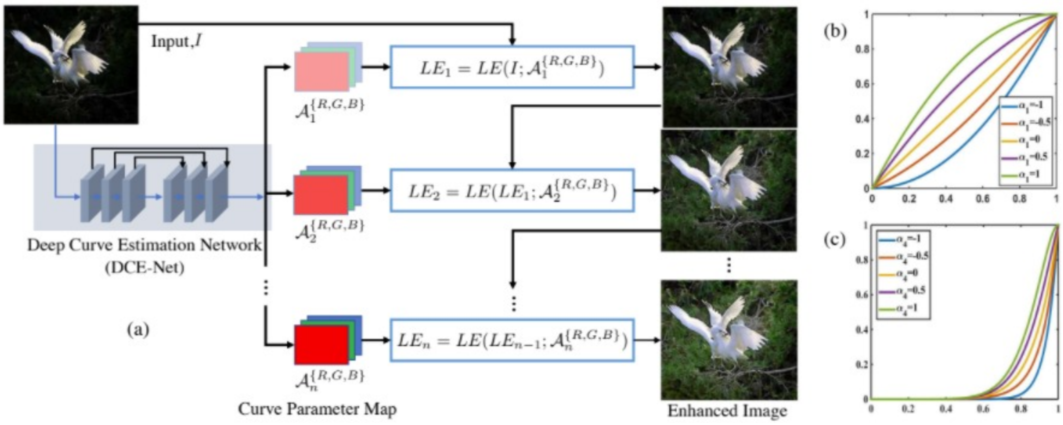


Figure 1.1: Block Diagram of Zero-Reference Deep Curve Estimation model

2.Learning to Restore Low-Light Images via Decomposition-and-Enhancement(CVPR-2020)[6]

It presents a novel network that first learns to recover image objects in the low-frequency layer and then enhances high-frequency details based on the recovered image objects. In addition, a new low-light image dataset is prepared with real noise to facilitate learning.

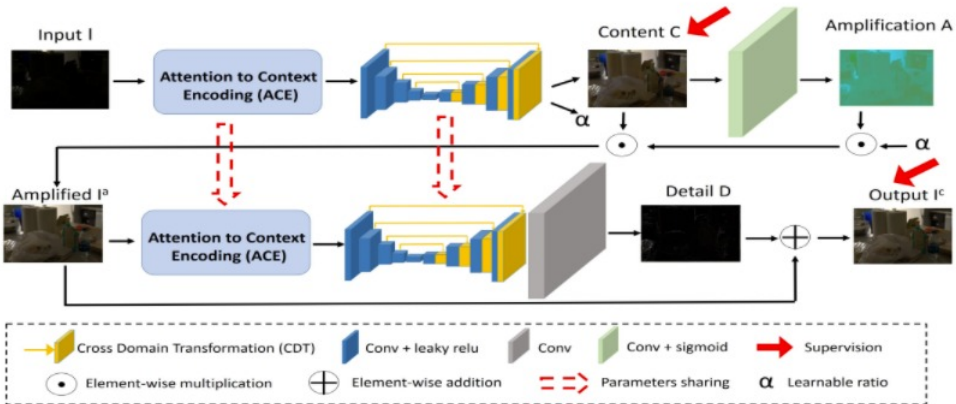


Figure 1.2: Block diagram of Learning to restore Low-light images via Decomposition and enhancement architecture.

3. From Fidelity to Perceptual Quality: A Semi-Supervised Approach for Low-Light Image Enhancement(CVPR-2020)[7]

The architecture here is divided into two parts: 1.recursive band learning : here the drbn is constructed to recover a normal light image based on low light input in recursive manner, 2.Band recomposition:the model here further learns to recompose the restored signals with perceptual quality guide

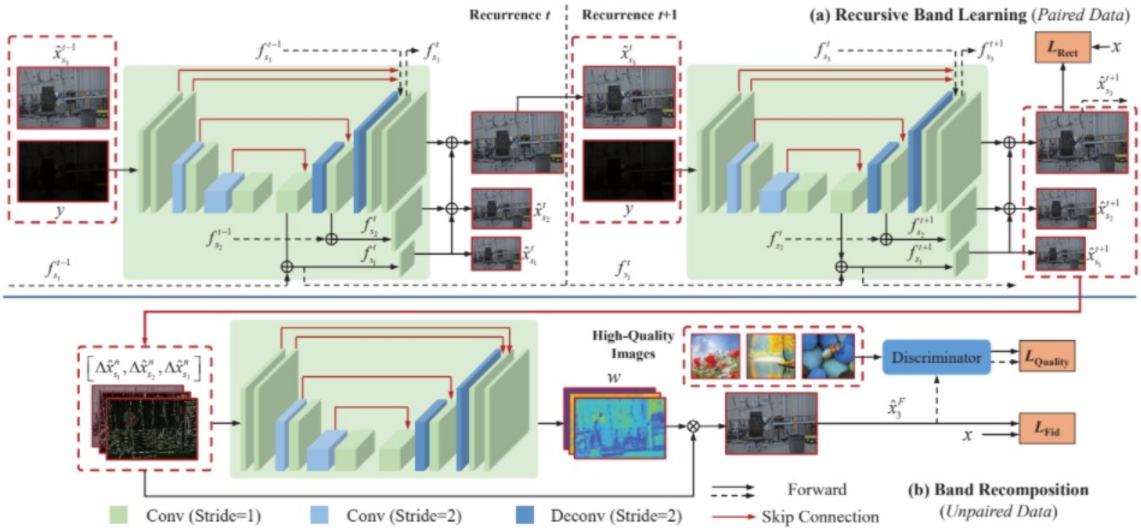


Figure 1.3: Block diagram of Fidelity to Perceptual Quality architecture.

4.A New Low-Light Image Enhancement Algorithm using Camera Response Model(CVPR-2020)[8]
A novel enhancement method using the response characteristics of cameras. First, we investigate the relationship between two images with different exposures to obtain an accurate camera response model. Then we borrow the illumination estimation techniques to estimate the exposure ratio map.

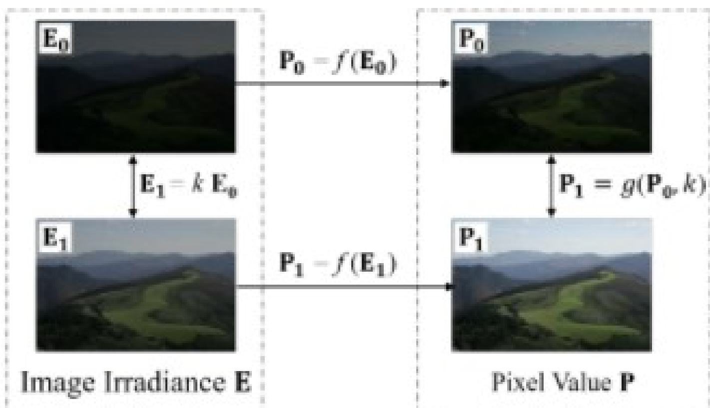


Figure 1.4: Block diagram of Camera response model architecture.

5.**Restoring Extremely Dark Images in Real Time(CVPR-2020)**[4]
A new deep learning architecture for extreme low-light single image restoration, despite of its fast lightweight inference, produces a restoration that is perceptually at par with state-of-the-art computationally intense models. To achieve this, we do most of the processing in the higher scale-spaces, skipping the intermediate-scales wherever possible.

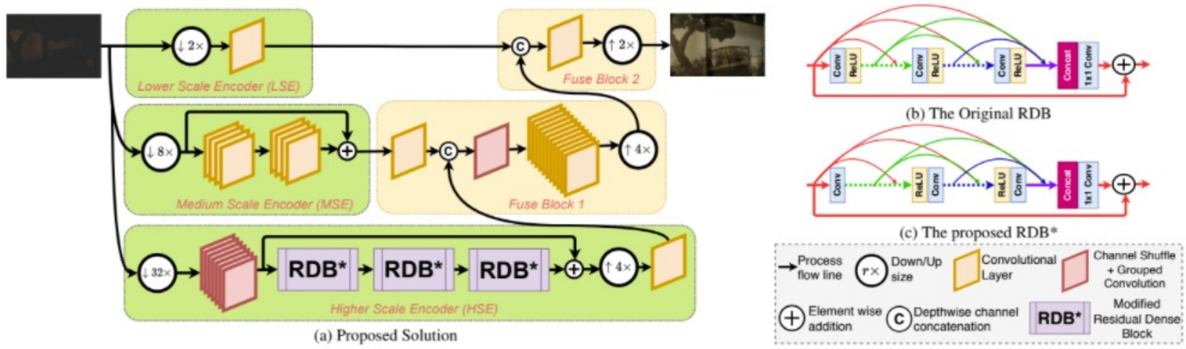


Figure 3. Architectural details of the proposed model. Depth of convolutional layer is roughly proportional to number of o/p features.

Figure 1.5: Block diagram of Restoring Extremely Dark Images in Real Time model architecture.

1.4 Problem statement

To propose a Learning based technique for Enhancement of images captured in low light conditions.

1.5 Application in Societal Context

Enhancement of images captured in low-light conditions, is one of the requirements in present societal conditions. If we can achieve more night vision capacity through such learning based models, it will be a tribute to may elements/roles of society.

Enhancement of images captured in low-light conditions can have novel societal applications as follows:

1.Night mode captures in mobile phones.
2.Analyzing low-light camera footage in forensics.
3.Help people with Nyctalopia.
4.Military applications:

a)Night vision equipment.

b)Espionage

c)Drones and combat planes in target acquisition during night missions.

d)Prevents mistargeting.

e)Counter insurgency.

5.It allows using inexpensive sensors to get quality images.

Chapter 2

System design

In this chapter, we will be looking towards the functional block diagram, the design alternatives and also about the final design which is being implemented.

2.1 Functional block diagram

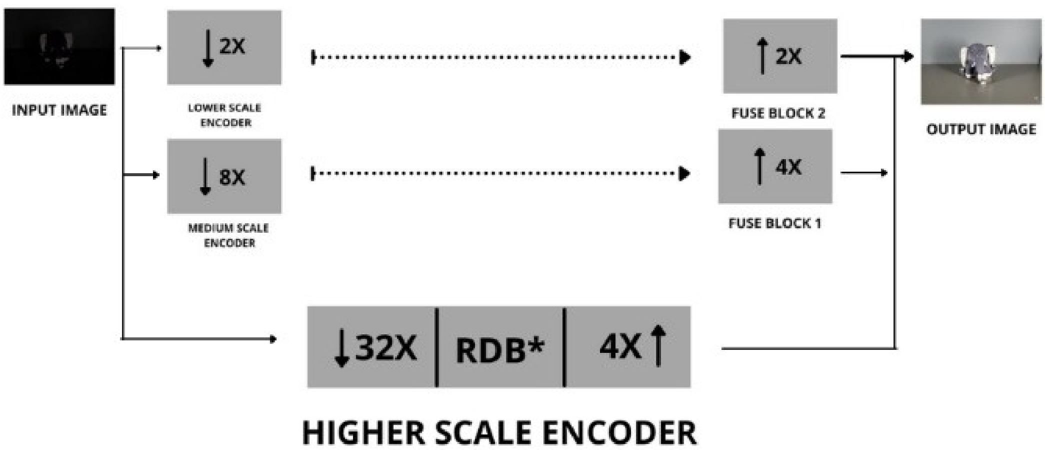


Figure 2.1: Functional block diagram of the proposed model.

Lower scale encoder[LSE]-Down-scales the input to 2 factor.
Medium scale encoder[MSE]-Down-scales the input to 8 factor.
Higher scale encoder[HSE]-Contains three different blocks:
i)Block 1: Down-scales the input to 32 factor.
ii)Block 2: Contains RDB* iii)Block 3: Up-scales the output of RDB* with the factor of 4.
Fuse block1: Up-scales and blends the output at the factor of 4.
Fuse block2: Up-scales the output with the factor of 2 and produces final output of the generator.

2.2 Design alternatives

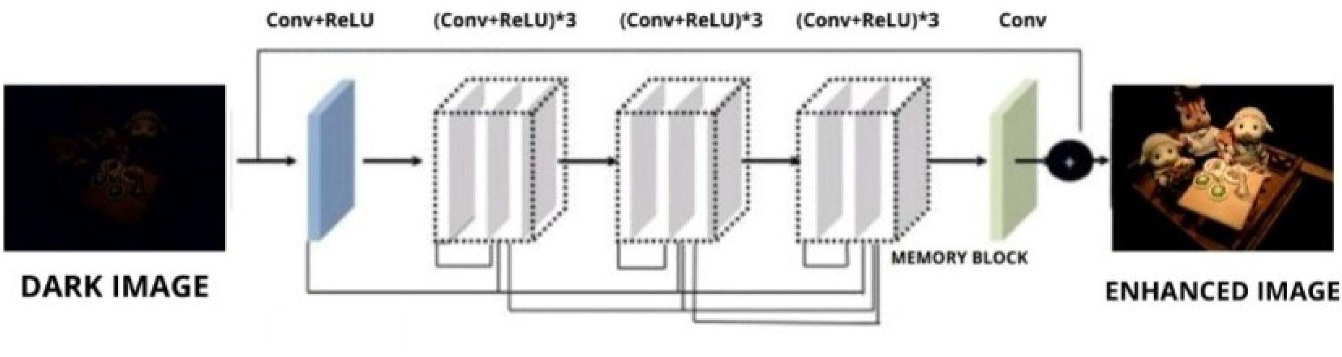


Figure 2.2: Color wise attention based model.

Color-wise attention network (CWAN) is proposed for low-light image enhancement based on convolutional neural networks. Motivated by the human visual system when looking at dark images, CWAN learns an end-to-end mapping between low-light and enhanced images while searching for any useful color cues in the low-light image to aid in the color enhancement process. Once these regions are identified, CWAN attention will be mainly focused to synthesize these local regions, as well as the global image. Both quantitative and qualitative experiments on challenging datasets demonstrate the advantages of our method in comparison with state-of-the-art methods.

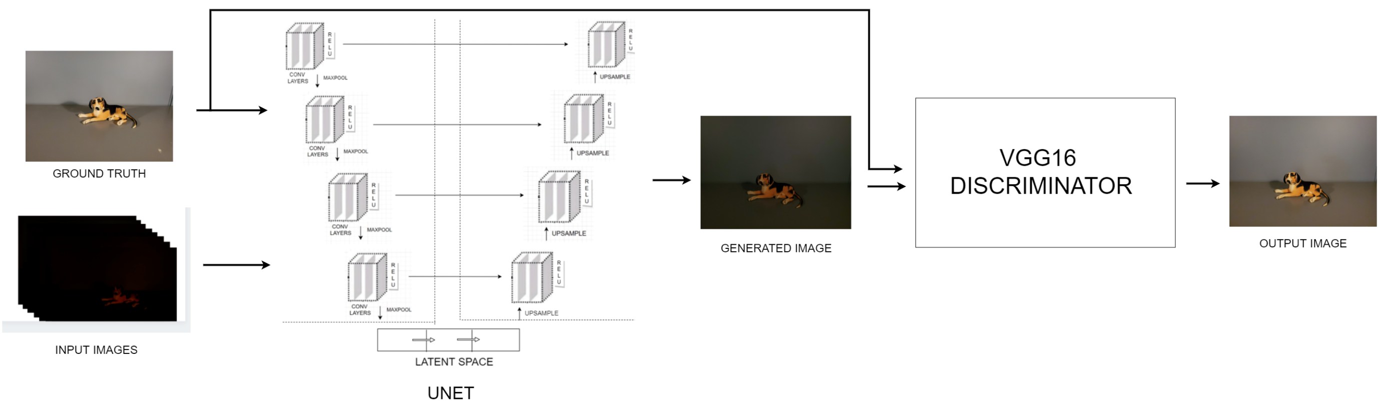


Figure 2.3: UNET based model

UNET based model proposed for lowlight enhancement looks upon the enhancement of light aspect of the image keeping the texture intact.It follows the 3 convolution layers

2.3 Final design

We select one of the optimal solutions based on its working, ease of implementation and results obtained.

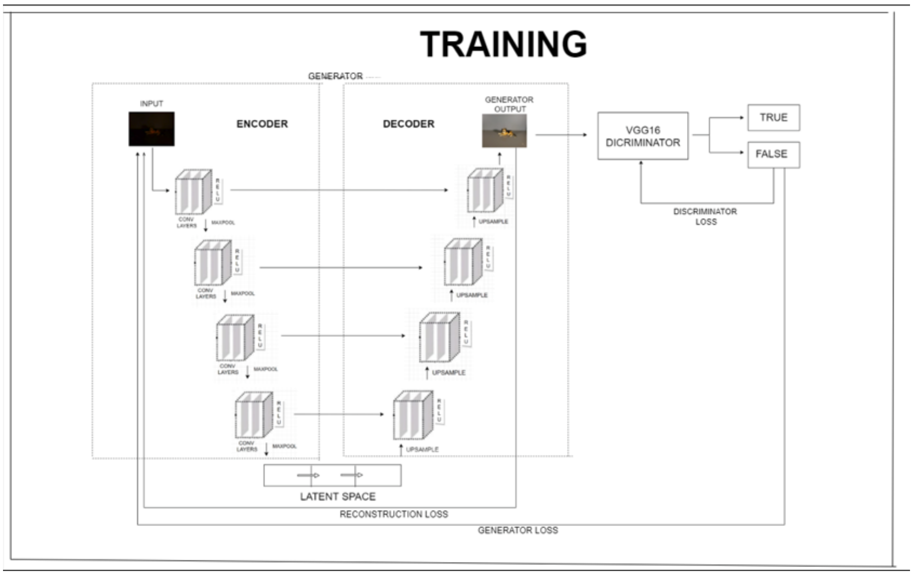


Figure 2.4: Training phase block daigram

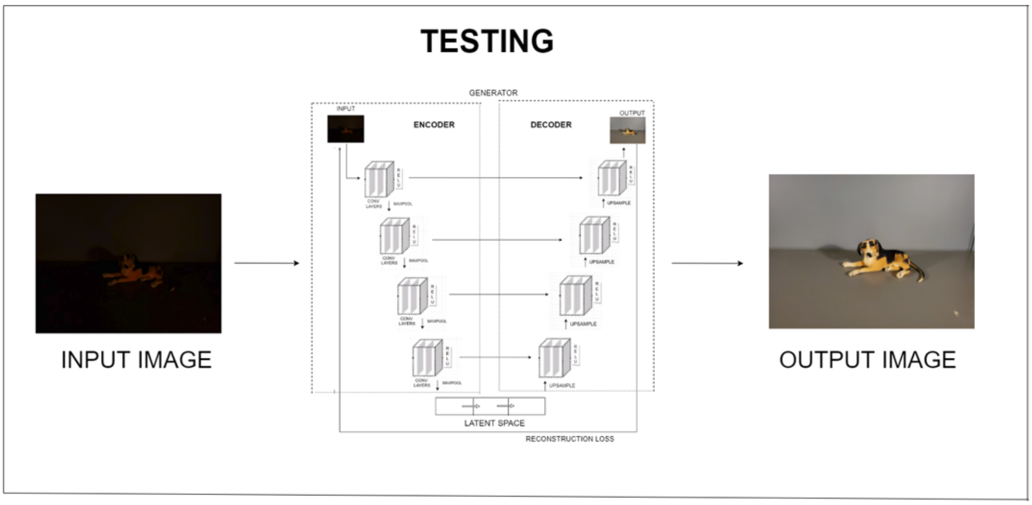


Figure 2.5: Testing phase block daigram.

Chapter 3

Implementation details

In this chapter, we will be looking into the specifications and the final system architecture, algorithm and the flowchart of the proposed architecture.

3.1 Specifications and final system architecture

As mentioned earlier this model is efficient n terms of both computational speed and the enhancement quality or the results.
To avoid computational overhead we have made required changes in the basic UNET architecture as shown in the below diagram.

As shown in the Figure3.1, lower scale encoder is at 1/2 resolution, medium scale encoder is at 1/8 resolution and the high scale encoder is at 1/32 resolution.

The functional block of high scale encoder is RDB* which is residual dense block helping us extract significant features from the input.

Here we provide raw images in the format '.ARW' as input to the network and generate restored RGB images saved in format '.jpeg' .

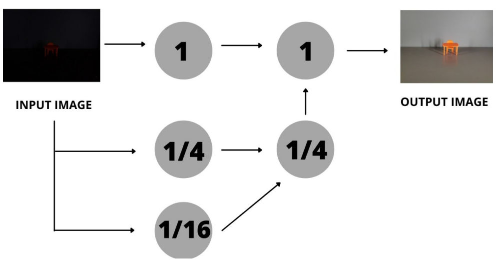


Figure 3.1: Proposed architecture.[1/r represents the scaled values.]

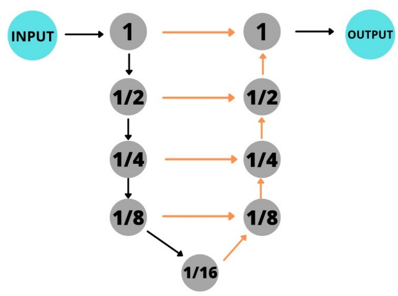


Figure 3.2: Basic Unet Architecture.[1/r represents the scaled values.]

The raw images are used in order to achieve accurate results without losing the required data, which usually occurs when captured images are processed through camera compression.

Fuse blocks are used to blend the low and medium frequency details.
The loss function used is MS-SSIM Loss.Activation function used is LeakyReLU refer Figure3.2

3.2 Algorithm

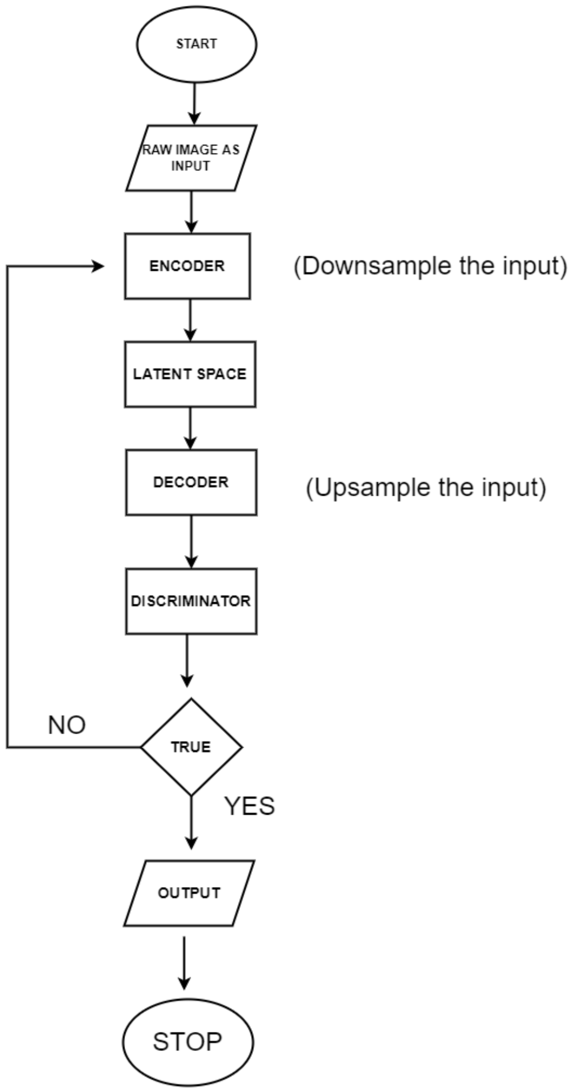
File Load:
train$_{f}iles = glob.glob("path")$
$gt_{f}iles = glob.glob("path")$
$test_{f}iles = glob.glob("path")$
$UNETAutoencoder:$
$UNET_{b}lock1:$
$x1 = train_{f}iles$
$x2 = CONV2D(x2)$
$x2 = CONV2D(x2)$
$x2 = CONV2D(x2)$
$x2 = RELU(x2)$
$x2 = MAXPOOL(X2)$
$UNET_{b}lock2:$
.
.
.
.
.
$UNET_{b}lock4:$

UNET$_{b}lock5:$
$x5 = CONV2D(x4)$
$x5 = CONV2D(x5)$
$x5 = CONV2D(x5)$
$x5 = RELU(x5)$
$x5 = MAXPOOL(X5)$
$UNET_{b}lock6:$
$x6 = CONV2D((x5, x4))$
$x6 = CONV2D(x6)$
$x6 = CONV2D(x6)$
$x6 = RELU(x6)$
$x6 = MAXPOOL(X6)$
$x7 = CONV2D(cat(x6, x3))$
$UNET_{b}lock7:$
$x7 = CONV2D(x7)$
$x7 = CONV2D(x7)$
$x7 = RELU(x7)$
$x7 = MAXPOOL(X7)$
$UNET_{b}lock8:$
$x8 = CONV2D(cat(x7x2))$
$x8 = CONV2D(x8)$
$x8 = CONV2D(x8)$
$x8 = RELU(x8)$

$x8 = MAXPOOL(X8)$
.
.
.
.
.
$UNET_block9:$
$output = x9$
$DISCRIMINATOR:$
$x = SELF.SLICE(X)$
$x = SELF.SLICE3(X)$
$RELU_2 = X \;\; x = SELF.SLICE3(X)$
$RELU_2 = X \; returnRELU_2$
return the model with required layers

3.3 Flowchart

Provides us the flow of entire process in this model.



Chapter 4

Results and discussions

In this chapter we discuss about the sony data-set used by lamba paper(citation) and our data-set.
With the understanding of above data-sets, we discuss and compare the results obtained by conducted experiments.

4.1 Datasets

We train the proposed architecture on two data-sets mainly, Sony data-set[1] where the image is captured in .ARW format (refer figure 4.1), in these the input images are the images captured using short exposure time of 0.1seconds and 0.4 seconds, and its corresponding ground truths captured in exposures time of 10s and 30s. Second, data captured using simulated environment with artificial lighting which is captured in .DNG format(refer images xx). Both these images are single channeled images. Images are captured using varying ISO keeping exposure constant.

Sony data-set

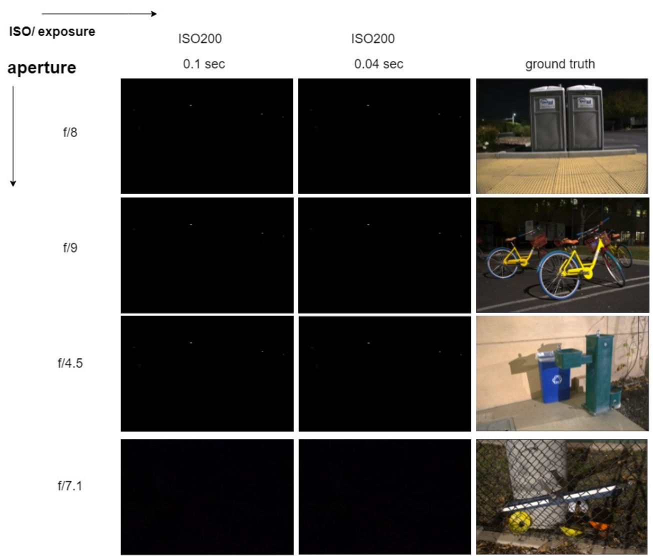


Figure 4.1: The above table represents the Sony dataset with varying exposure time and aperture.

Our data-set

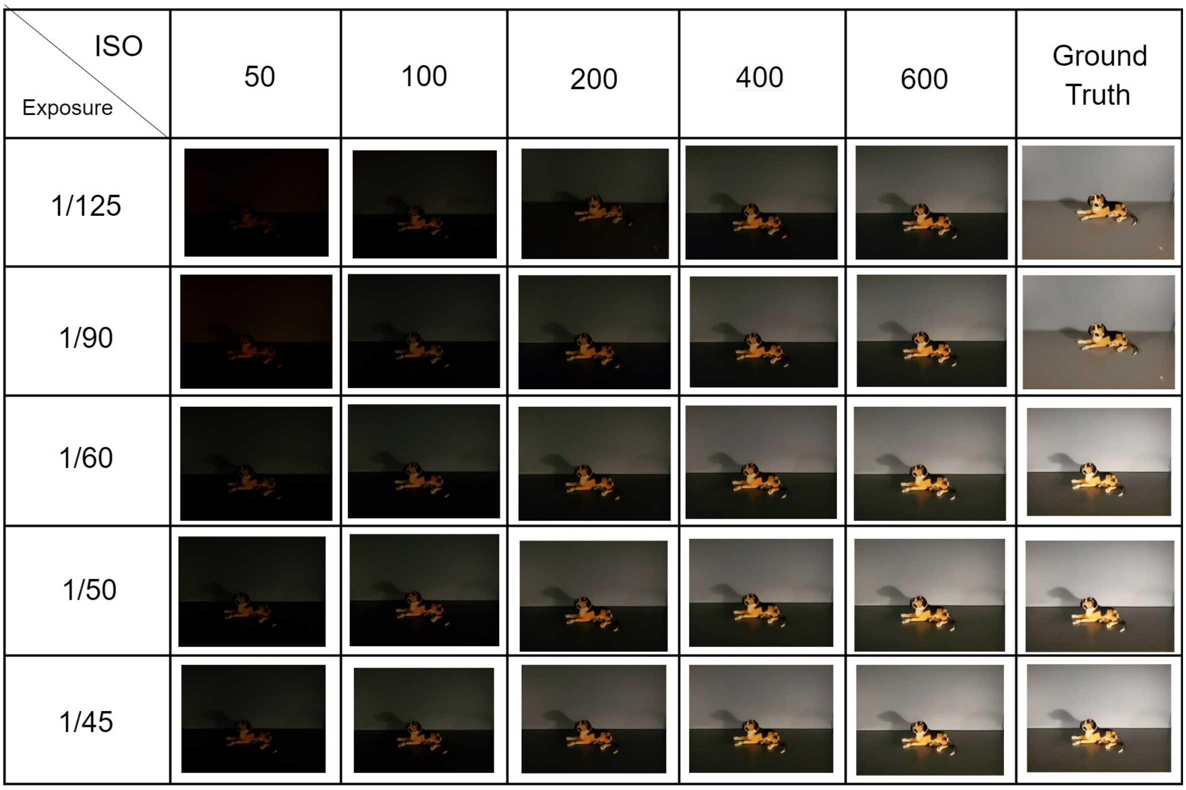


Figure 4.2: The above table represents the dataset captured in the simulated environment with varying exposure time and aperture.

4.2 Evaluation metrics

We evaluate the results generated by the proposed architecture using Peak Signal to Noise Ratio[3] (PSNR) and Structural Similarity Index Measure (SSIM)[5] evaluation metrics. **PSNR**[3] is the measure of power of corrupting noise that affects the fidelity of its representation and the ratio between the maximum possible power of signal, while **SSIM**[5] is a comprehensive metric that measures image quality loss caused by the processing such as data comprehension or data transfer loss, and is a complete reference metric which requires the ground truth and the processed image. PSNR[3] and SSIM are defined as:

$$PSNR[3] = 10log_{10}(\frac{P_{signal_{max}}}{MSE}) \tag{4.1}$$

where $p_{signal_{max}}$ is the maximum power of the signal

$$SSIM[5] = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{4.2}$$

where $\mu_x\mu_y\sigma_x\sigma_y and \sigma_{xy}$ are means, standard deviations and covariances for two images x and y

4.3 Experimental Setup

Experiment: 01

We have trained on Sony data-set using UNET auto-encoder with VGG16 discriminator having batch size 8, Adam optimizer, setting the learning rate at 1e-4. We trained the following model for one million iterations(125,000 epochs) using NVIDIA DGX-V100 with 32 GB VRAM.Evaluation of the following model is done using PSNR and SSIM metrics.

Experiment: 02

We have trained using the data-set captured in simulated environment having various ISO. We have used the same UNET auto-encoder with VGG16 discriminator having batch size 8 , Adam optimizer, setting the learning rate at 1e-4. We trained the following model for one million iterations(125,000 epochs) using NVIDIA DGX-V100 with 32 GB VRAM. We evaluate the following model using PSNR and SSIM metrics.

4.4 Experimental Results

We use two data-sets to train the model on:

- proposed architecture using Sony data-set
- proposed architecture using Custom data-set

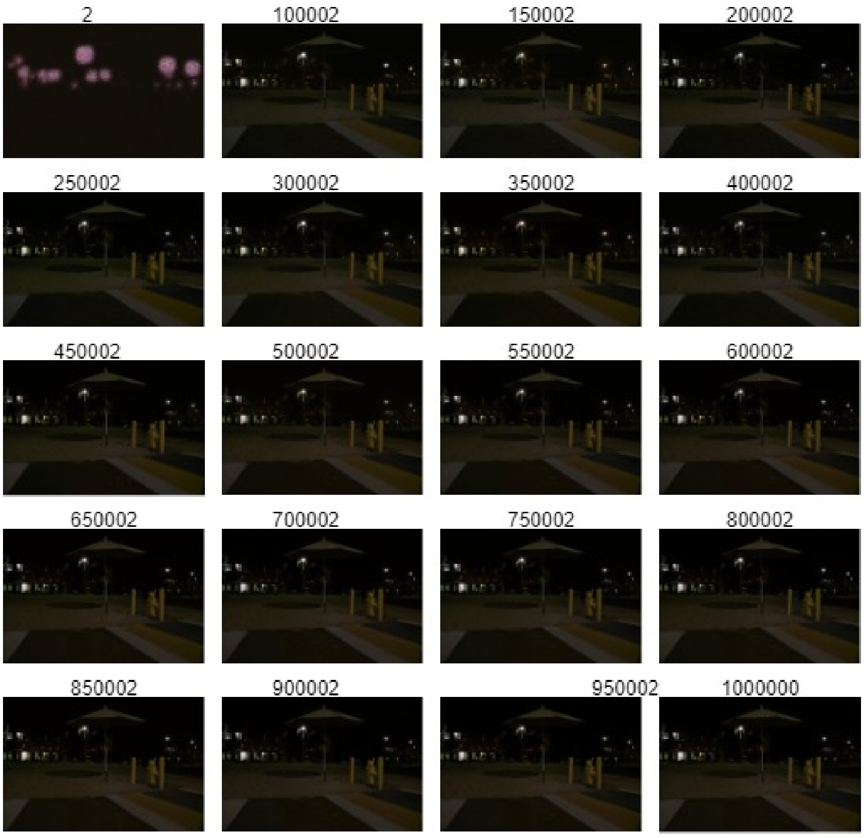


Figure 4.3: The above table represents the experimentation results using Sony data-set. The numbers above each image represents the iteration number.

Parameters	LDC	LAMBA	LAMBA(SONY)[4]†
PSNR	**29.56**	28.66	25.4994
SSIM	**0.799**	0.790	0.52511

The above table represents all the experimentation results.The best results are represented in bold and the second best is underlined.

Chapter 5

Conclusions and future scope

5.1 Conclusion

In conclusion it can be stated that low-light image enhancement project is elite and serves the society from root level to highest level of community.

5.2 Future scope

Enhancement of images captured in low-light conditions can have many future applications in military field for having an low light condition image detecting gadgets during highly classified missions and in a broader view of applications we can use this learning based model as a part of vision gadget to the needy community.

Bibliography

[1] Chen Chen et al. "Learning to see in the dark". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 3291–3300.

[2] Chunle Guo et al. "Zero-reference deep curve estimation for low-light image enhancement". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 1780–1789.

[3] Quan Huynh-Thu and Mohammed Ghanbari. "Scope of validity of PSNR in image/video quality assessment". In: *Electronics letters* 44.13 (2008), pp. 800–801.

[4] Mohit Lamba and Kaushik Mitra. "Restoring Extremely Dark Images in Real Time". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, pp. 3487–3497.

[5] Shiqi Wang et al. "SSIM-motivated rate-distortion optimization for video coding". In: *IEEE Transactions on Circuits and Systems for Video Technology* 22.4 (2011), pp. 516–529.

[6] Ke Xu et al. "Learning to restore low-light images via decomposition-and-enhancement". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 2281–2290.

[7] Wenhan Yang et al. "From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 3063–3072.

[8] Zhenqiang Ying et al. "A new low-light image enhancement algorithm using camera response model". In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2017, pp. 3015–3022.