

An Improved Image Enhancement Technique for Low Light Images Using Deep Learning Approach

¹Rajesh Gopakumar and ²Karunakar A Kotegar

¹Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India.

²Department of Data Science and Computer Applications, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India.

¹rajesh.g@manipal.edu, ²karunakar.ak@manipal.edu

Correspondence should be addressed to Rajesh Gopakumar : rajesh.g@manipal.edu

Article Info

Journal of Machine and Computing (<http://anapub.co.ke/journals/jmc/jmc.html>)

Doi: <https://doi.org/10.53759/7669/jmc202404060>

Received 02 January 2024; Revised from 04 April 2024; Accepted 06 June 2024

Available online 05 July 2024.

©2024 The Authors. Published by AnaPub Publications.

This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Abstract - Image enhancement in facial detection is a critical component of facial recognition systems. Face identification in an uncontrolled environment is affected by a multitude of difficulties such as poor light levels, low-resolution cameras, occlusions from surrounding objects, and tiny faces in distant photographs. Low signal-to-noise ratio, low brightness, and noise in low-light photographs lead to issues such as color distortion and poor visibility, which makes it challenging to identify faces. Many techniques to enhance low-light images have been developed, improving the face detection system's accuracy. This will improve the picture at the expense of higher running costs and lower model robustness. The proposed technique, DCE-Net, uses performance-intensive deep learning and light-enhanced image properties. A non-referential deep learning technique was employed to acquire and modify the image attributes. A set of loss functions designed to perform without ground-truth images is the foundation of the deep network learning employed. Compared to the current referential methods, straightforward non-referential light estimation curve mapping minimizes the computational demand for low-light image improvement. Several experiments conducted on standard datasets demonstrated the efficacy and reduced computational requirements of the approach. The effectiveness of this method is supported by both the qualitative and quantitative outcomes. The PSNR and SSIM computation for paired images shows promising results using the proposed image enhancement technique.

Keywords - Face Detection, Image Enhancement, Deep Learning, Convolutional Neural Networks, Face Recognition.

I. INTRODUCTION

Image enhancement has become a very important and ongoing research topic for a long time now. One of the necessary phases of a face recognition system is face detection. Many factors present in an unconstrained environment affect face recognition, like poor resolution cameras, low-light environment, occlusions due to other objects in the scene, the small-sized faces in images taken from a far distance etc. It is challenging to distinguish faces in lowlight images due to problems such as color distortion and poor vision caused by low signal-to-noise ratio, low contrast, and noise [1]. Many techniques have been developed over time to improve lowlight images, advancing the accuracy of the face detection system. These methods improve the picture at the expense of higher running costs or lower model robustness. Figure 1 displays a sample low-light photograph along with an improved version. Technological advancements in the field of computing devices, such as GPUs and TPUs, have encouraged the use of convolutional neural networks for large deep learning models [1], [2]. Training large deep learning models requires huge computing power, leading to a high impact on the environment. This impact can be reduced by using computationally efficient algorithms and avoiding training the model from scratch if it is not necessary.

In this work, we have considered enhancing low-light images by improving the light parameters of images without paired supervision [2] using an improved deep learning algorithm. This addresses the development of a face detector for low-light images taken under various lighting conditions. Four different datasets, discussed in Section 3, are used for various experiments. The DCE-Net-based image enhancement technique yielded good quantitative and qualitative results for all the four datasets.

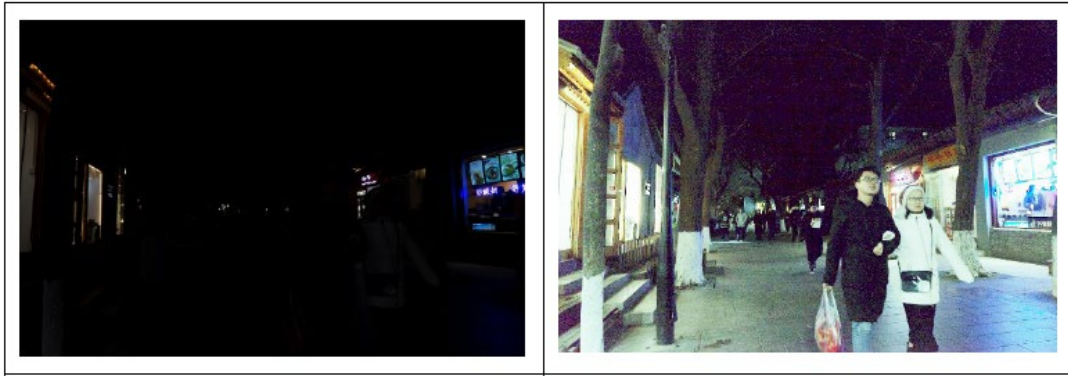


Fig 1. Typical lowlight image and corresponding improved image

II. RELATED WORKS

Liang et al. [3] proposed a deep-learning approach to effectively denoise and enhance images. Image enhancement was achieved by utilizing a neural network to forecast model coefficients in conjunction with feature extraction supported by illumination and noise maps. The proposed method is superior to traditional methods. Bhattacharya et al. [4] used image enhancement without paired supervision, which differs from conventional paired supervision methods. The proposed method uses various consistencies in the geometry and illumination to enhance the image. Benchmark datasets, which provided competitive results, were used for extensive experiments.

With a Retinex-based CNN, Hai et al. [5] deconstructed, denoised, and increased the contrast by utilizing the spatial and frequency features of the low-light image. The LSRW, a real-world paired dataset, was employed in the learning and training processes. The effectiveness of the proposed strategy was demonstrated by the extensive use of public datasets in the tests. By avoiding uneven lighting and noise problems, Liang et al.'s [6] recurrent quasi-exposure generation fused with multi-exposure efficiently improved face detection. The proposed method outperformed other algorithms on available datasets when tested on the DarkFace benchmark low-light dataset. Using face detectors, Wang et al. [7], [8] transformed normal-light images to low-light images without the need for low-light annotations. Utilizing bidirectional pixel-level translation, self-supervised learning, and reciprocal curve-based illumination adjustment, a high-low adaptation technique was applied. The test results demonstrated that the approach performed exceptionally well, even in the absence of low-light annotations.

A auto-balancing lighting model for enhancing images was introduced by Ma et al. [9]. The system uses a stacked illumination learning strategy with a weight distribution that converges between the level's outcomes to reduce processing costs. Hai et al. [10] used a fully convolutional network to split images into light and reflectance, and subsequently reduced reflection noise to enhance the clarity of the image. Retinex image decomposition and enhancement algorithms were used by Gasparyan et al. [11] to enhance the prominence and precision of photographs taken under low-light conditions. Using benchmark datasets, the proposed method was tested against several existing methods and yielded impressive results. Lu et al. [12] suppressed and eliminated the low-light effect generated by low-light environments using a modified GAN with light processing and Gaussian filtering. The experimental approach to extracting light-invariant characteristics from facial photos is effective. To establish the levels of enhancement, Guo et al. [2] suggested a picture enhancement technique based on light curve estimation that uses deep neural networks and loss functions. The recommended method is based on learning unpaired or non-paired data. Tests conducted on the reference datasets demonstrated the statistical and qualitative efficacy of the technique in object recognition.

III. PROPOSED METHODOLOGY

The major constraint in low-light face detection is the difficulty in identifying the facial features. The primary solution for this is to enhance the image and prepare the image for locating facial points. A technique for improving low-light images was devised using CNN. The system learns the illumination properties of the image input using a Deep Curve Estimation Network (DCE-Net) [2], and the light curve is modified to obtain an improved version. Various loss functions have been devised to make necessary adjustments using light enhancement curves without affecting the image. The model was tested and analyzed quantitatively and qualitatively to study the impact of the enhancement on low-light images using the proposed method.

The most common trend observed in image enhancement methods is the use of reference images for learning and testing on low-light images. This becomes impractical and difficult for real-life facial applications because they generate only low-light images. In such situations, it is always better to develop a non-reference method that requires only low-light images [4]. In this study, we used a non-reference method for image enhancement. Several experiments were conducted to confirm the improved efficacy of these images. To evaluate the enhanced image quality, we computed non-reference loss functions, as explained in Section 3.2. The peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were computed using the enhanced and low-light images to measure the decline in image quality. The following sections discuss the DCE non-reference framework for picture enhancement, and the loss functions are shown after each section. The various loss functions, SSIM, and PSNR calculations, as well as an MTCNN-based face detection framework for evaluating the efficacy

of the non-reference image enhancement technique, are covered in Section 4. The system was verified with face detection from enhanced images, proved to be very effective, and provided promising outcomes.

The architecture of the DCE low-light image-enhancement system is shown in Fig 2. Based on the light estimate curves produced by DCE-Net, the low-light image was improved. An improved output image is produced by repeatedly applying the estimation curves to the RGB channels of the input image.

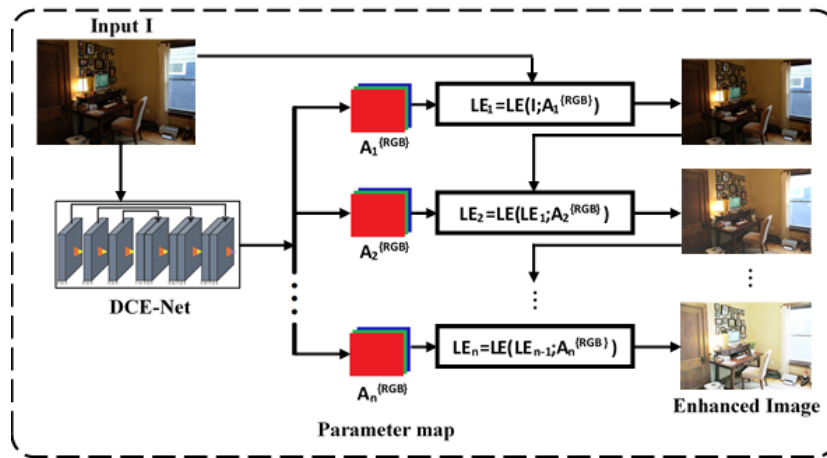


Fig 2. Proposed Low light enhancement system architecture

Light Estimation Curve

The automatic modification of the light estimation over the input image serves as the foundation for low-light image improvement. While keeping the pixel values within the normalized interval ranging from 0 to 1, the light estimating technique seeks to enhance the supplied image. Thus, data loss from overflow truncation was prevented. The light estimation curve minimally preserves the contrast between neighboring pixels. In addition, the light estimation curve should be differentiable and minimize the cost function as much as possible during the gradient backpropagation process. Eq. (1) shows a quadratic curve to achieve the aim of the light estimation curve.

$$PQ(P(x); \alpha) = P(x) + \alpha P(x)(1 - P(x)) \tag{1}$$

The pixel coordinates are indicated by x in the following equation: $PQ(P(x); \alpha)$ represents an improved form of $P(x)$. The Light Estimate Curve magnitude and exposure amount are controlled by the attainable curve factor $\alpha \in [-1, 1]$. All actions were performed pixel by pixel, and every pixel was normalized to $[0, 1]$. To prevent oversaturation and maintain the natural color of each RGB channel, a light estimation curve was applied independently to each channel.

The light estimation curve was iteratively applied to the input picture pixels in a dynamic range of values to improve the pixels. An image can be adjusted over a larger dynamic range using a higher-order curve. The dynamic range of the supplied image was adjusted using the value of the light estimation curve that best fits each pixel. Local areas cannot be over- or under-enhanced owing to pixel-by-pixel mapping, whereas global estimation affects the entire image. The bright areas of the input image were retained, while the dark areas were boosted to display the content. Eq. (1) can be rewritten in a higher order, as shown in Eq. (2), to address a wider range.

$$PQ_n(x) = PQ_{n-1}(x) + A_n(x) PQ_{n-1}(x) (1 - PQ_{n-1}(x)) \tag{2}$$

Variable n in Eq. (2) represents the number of recurrences. It was set to eight to consider all possible circumstances.

DCE-Net

Using a DCE-Net, the low-light image enhancement system discovers the mappings between the input image and the light estimate curve parameters that most closely match it. After receiving the input image, the DCE-Net determines which pixel-by-pixel estimation value best suits the corresponding higher order curves. A simple CNN was used to learn the mappings with seven symmetrically combined convolutional layers. The ReLU activation function was positioned after $32 \ 3 \times 3$ convolutional kernels with a stride of 1. Batch normalization and downsampling layers are avoided because they disrupt the relationships between nearby pixels. The Tanh activation function is used once the last convolutional layer is completed. It creates twenty-four parameter maps in eight rounds ($n = 8$), requiring 3 curve parameter maps for each of the three RGB channels. Notably, given an image input size of $256 \times 256 \times 3$, DCE-Net produces 79,416 trainable parameters using 5.21 GFlops.

DCE-Net enables us to assess the enhanced image quality by facilitating learning with a group of unique non-reference losses [2]. Any image enhancement method preserves the spatial variations between the original image and the improved image to enable the detection of objects therein. The spatial consistency measure helps preserve these variations. Exposure is another important factor that determines how dark or light an image is. It is important to control the exposure level in an image when it is subjected to enhancement. Similar to exposure, color is another important aspect that contributes to the composition and visual appeal of an image. Noise is one of the reasons for the attenuation of perceptual quality, which affects the imaging system. Illumination variations in an image provide hints on the object present through shades and shadows. These parameters are important for any image enhancement system. Thus, the images were trained using the four categories of loss discussed below.

Spatial Consistency Loss. To recognize objects or faces, the spatial variation between adjacent picture elements in the original image and the improved image must be preserved. The spatial coherence between the input and improved photos is guaranteed by the Spatial Consistency Loss (L_{SCL}), which maintains the differentiation among neighboring regions. The loss of spatial consistency between the input and improved pictures was calculated and controlled using Eq. (3).

$$L_{SCL} = \frac{1}{N} \sum_{i=1}^N \sum_{j \in RE(i)} (|X_i - X_j| - |(P_i - P_j)|)^2 \quad (3)$$

The nearby regions (up, below, left side, and right side) centered at region i are denoted by $RE(i)$, where N denotes the total number of nearby regions. The local area average intensity value is represented by symbols X and P in the improved form of the image and image input, respectively. For experimental reasons, the local area was 4×4 in size.

Exposure Control Loss. There should be no overexposure or underexposure in the corrected image. Exposure levels were controlled during the learning process using Exposure Control Loss (L_{ECL}). The difference between the average intensity values of a local area and the exposure level EXP was measured using the L_{ECL} . The constant value of EXP was 0.6. The loss, or L_{ECL} , was calculated using Eq. (4).

$$L_{ECL} = \frac{1}{N} \sum_{k=1}^N |X_k - EXP| \quad (4)$$

where X is the average intensity value of a local region in the improved image and N is the number of non-overlapping 16×16 local areas.

Color Constancy Loss. Perceiving the color of the objects in the images, which is primarily independent of the color of the light source, is crucial. Color Constancy Loss (L_{CCL}) is computed to evaluate scene lighting and adjust to eliminate illumination influence, Color Constancy Loss (L_{CCL}) is computed. Maintaining the relationships between the three channels and preventing potential color variations in enhanced photos are two benefits of L_{CCL} . To compute the color constancy loss (L_{CCL}), we used Eq. (5).

$$L_{CCL} = \sum_{\forall (x,y) \in \varepsilon} (Q^x - Q^y)^2, \varepsilon = \{(R, G), (R, B), (G, B)\}, \quad (5)$$

In the enhanced image, (x, y) indicates two channels, and Q_x and Q_y are the averages of the intensity of each channel.

Illumination Smoothness Loss. For every DCE parameter map A , an illumination smoothing loss (L_{ISL}) was applied to maintain the normal correlations between the surrounding pixels. Eq. (6) is used to compute the Illumination Smoothness Loss (L_{ISL}).

$$L_{ISL} = \frac{1}{M} \sum_{n=1}^M \sum_{c \in \xi} (|\nabla_x A_n^c| + |\nabla_y A_n^c|)^2, \xi = \{R, G, B\}, \quad (6)$$

M indicates the count of repetitions, and ∇_x and ∇_y represent the horizontal and vertical gradient operations, respectively.

Total Loss. The combined loss calculation is described in Eq. (7).

$$L_{TOTAL} = L_{SCL} + L_{ECL} + W_{CCL}L_{CCL} + W_{ISL}L_{ISL}, \quad (7)$$

where W_{CCL} and W_{ISL} are the weights of the losses.

The graphs pertaining to the loss calculations for various datasets are discussed in Section IV.

Implementation details

Due to the growing popularity of deep learning for object identification, the majority of CNN-based models use paired information for training the networks [13], [14], [15]. In this study, the experiment was conducted on four different datasets. The details of each dataset are provided in Section III.

Three paired datasets were used to analyze the low-light improvement method discussed in the previous section. An evaluation of low-light picture improvement and quantification of face detection from the improved photos was performed using the non-reference (unpaired) dataset Dark Face. Section 4 presents an experimental description and results. The

training and test sets included both overexposed and low-light photos. The details of the data split used for training and testing with each dataset are discussed, along with the dataset description in Section 3.

A 512×512 -pixel resizing was applied to the training images. We used a TensorFlow framework on an NVIDIA A100 GPU. It used an eight-block batch size. The filter weights for each layer were determined using a Gaussian function with a standard zero mean and standard deviation of 0.02. A constant value was used to start bias. An ADAM optimizer was used for the network optimization with standard settings and a constant learning rate of $1e-4$.

Datasets

The proposed low-light enhancement system was tested for its improved capabilities using various datasets, and the results are discussed in the Results section. For various loss-function calculations and visual and perceptual comparisons, we used four datasets: Dark face [16], LOL [17], RELISUR [18], and ExDark [19]. The Dark Face dataset was utilized in face identification tests as well as for quantitative comparisons of SSIM and PSNR. From the literature, it is evident that the data used in this study are benchmark datasets available for investigating the performance of image enhancement in images captured in uncontrolled environments. The LOL, RELISUR, and ExDark datasets have ground truth to verify the quality of the enhanced images. The aim of using a Dark Face was to investigate the quality of enhanced images without ground truths by detecting faces in enhanced images. Figure 3 shows the samples from the four datasets used.

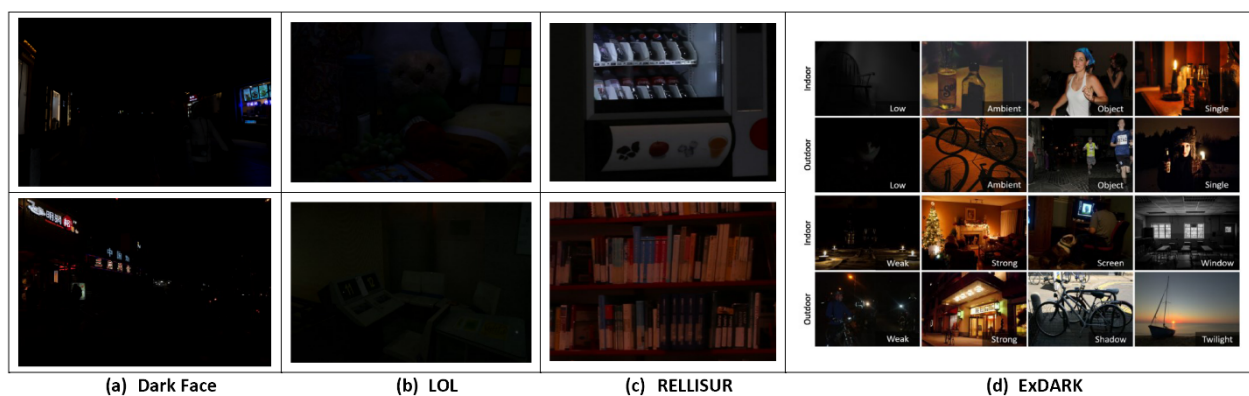


Fig 3. Example images from the datasets (a) Dark face (b) LOL (c) RELISUR (d) ExDARK

Dark Face

For the Dark Face [16] dataset, the main training and/or validation sets include 6,000 real-world low-light nighttime images of university campuses, including buildings used for instruction, roadways, overbridges, overpasses, children's parks, etc. The bounding boxes for the human faces are labelled in every image. 9,000 lowlight, raw images from the same site are also included. Furthermore, 789 distinct pairs of low-light/normal-light photos taken under controlled real-lighting circumstances (but inevitably featuring faces) were included in the training data at the participants' discretization. The dataset's human face bounding boxes were annotated using a separate testing set of 4,000 low-light photos. A total of 6000 photos from the Dark Face dataset were used in this study, of which 4800 were used as training sets, while the remaining 600 were used for testing and validation. Figure 3(a) shows sample images from the Dark Face dataset.

LOw-Light (LOL)

Among the 500 lowlight and normal light picture sets in the LOw-Light (LOL) dataset [17], 485 pairs and 15 pairs were used for training and testing, respectively. The noise generated during photo production appeared in the low-light shots. Most of the pictures show the interior settings. The low-light images and normal-light images from the LOL dataset were used for training purposes. Images of 400×600 pixels were used for testing purposes. In this experiment, we used 500 images, comprising 400 images used for training, 85 for validation, and the remaining 15 for testing purposes. Figure 3 shows a sample image and its upgraded version. Sample photographs from the LOL dataset are shown in Fig. 3 (b).

Real LowLight Image Super-Resolution Dataset (RELLISUR)

The dataset RELISUR [18] contains image pairings with extremely low light, poor resolution, normal light, and high quality. The collection includes a large number of DSLR images captured both indoors and outdoors. The collection consisted of five real lowlight photographs and true sequences of 1-, 2-, and 4-scale normal light photographs. Low-resolution photographs are represented by the 1 and 2 scale levels, and high-resolution ground-truth reference images are represented by four scale levels. The low-light photos are likewise regarded as having low resolution because they are taken at scale 1. The RELISUR dataset contained 850 unique sequences. Overall, there are 2550 normal-light low-resolution and high-resolution pairings with three different scale levels. Because each set of five underexposed photos correlates to the same normal light reference image, there are 4250 lowlight / normal light image pairs for each of the three scale levels. Consequently, the total number of lowlight/normal light image pairs in RELISUR is 12750. The 4250 images in the

RELLISUR dataset utilized in this study were divided into three categories: training (3610 images), validation (215 images), and testing (425 images). Figure 3(c) shows images from the RELLISUR dataset.

Exclusively Dark (ExDark)

An image is classified as lowlight in the ExDark [19] lowlight object image dataset if discernible or minor lighting variations exist. 7363 photos from 12 different classes—bikes, boats, bottles, buses, cars, cats, chairs, cups, dogs, motorbikes, people are included in the dataset. The majority of the low-light photos were taken via Google Search and several other websites and search engines, including Photobucket, Deviantart, Flickr, Imgur, Flickr, and Getty images. The ExDark dataset that we used for this study contains 936 images in total; training uses 597 of those images, 15 of which are used for validation, and the 324 remaining images for testing. Figure 3(d) shows the selection of example images from the ExDark collection.

IV. RESULTS AND DISCUSSION

Loss calculation impact

Learning in DCE-Net is supported by employing a collection of distinguishable non reference losses discussed in Section III, allowing us to gauge the quality of the improved images. Figures 4 to 7 show the graphs pertaining to the loss calculations of the DCE-Net training and validation on four datasets using different non reference losses. Compared to the entire output, the contrast of the improved image without spatial consistency loss (L_{SCL}) is lower. Thus, the L_{SCL} preserves the localized differences between the high- and low-light images. The L_{ECL} helps to restore the low-light area. Loss of color constancy prevents the development of color casts. Unwanted objects emerge because of the illumination smoothing loss, L_{ISL} , which disrupts the interactions between nearby pixels.

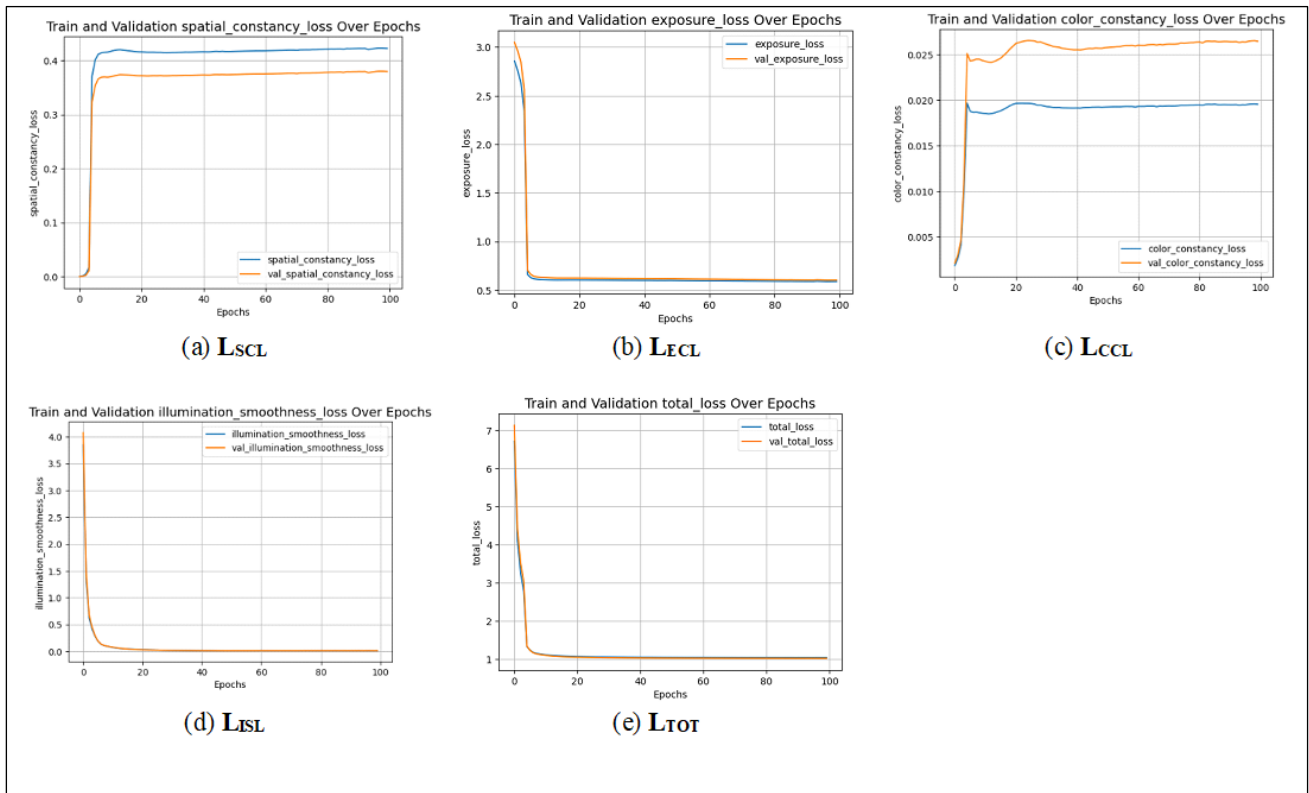


Fig 4. Loss calculations for Dark face (a) L_{SCL} (b) L_{ECL} (c) L_{CCL} (d) L_{ISL} (e) L_{TOT}

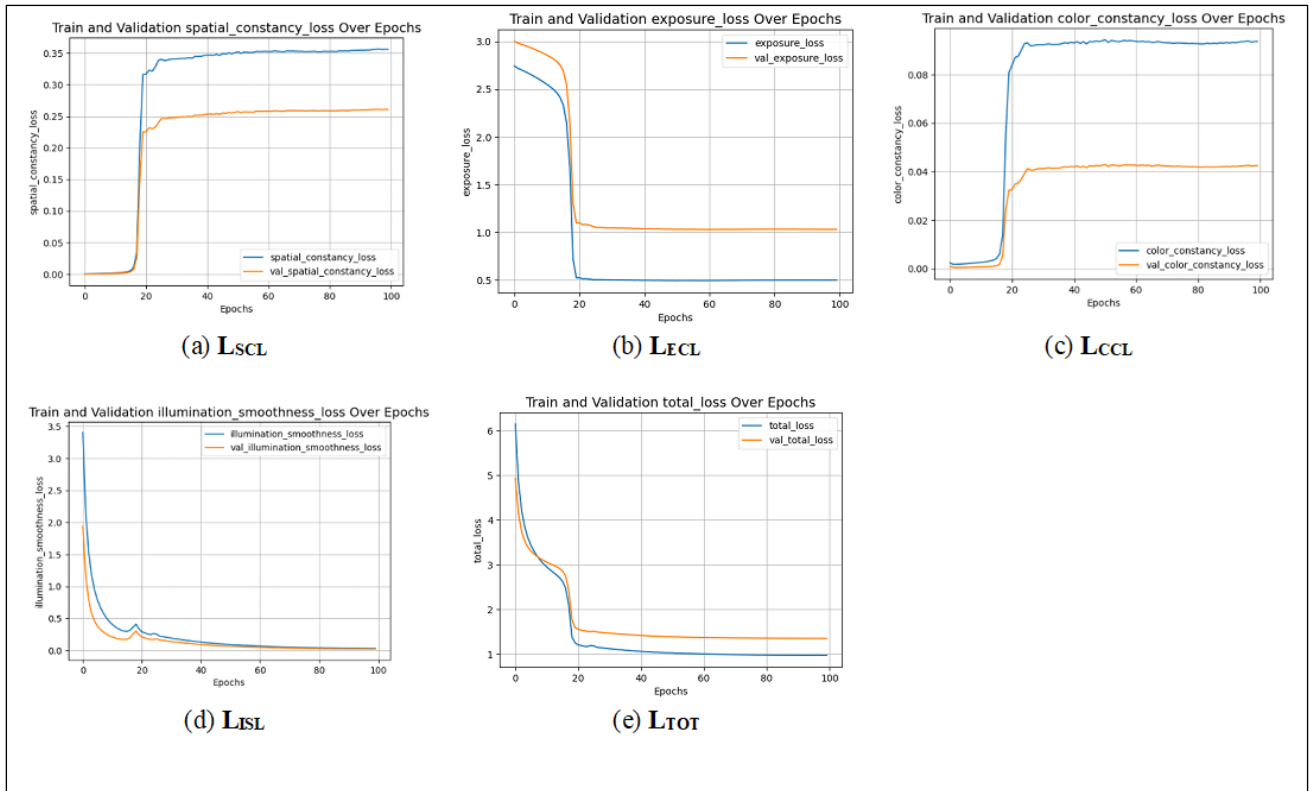


Fig 5. Loss calculations for LOL (a) LsCL (b) LECL (c) LCCL (d) Lisl (e) LTOT

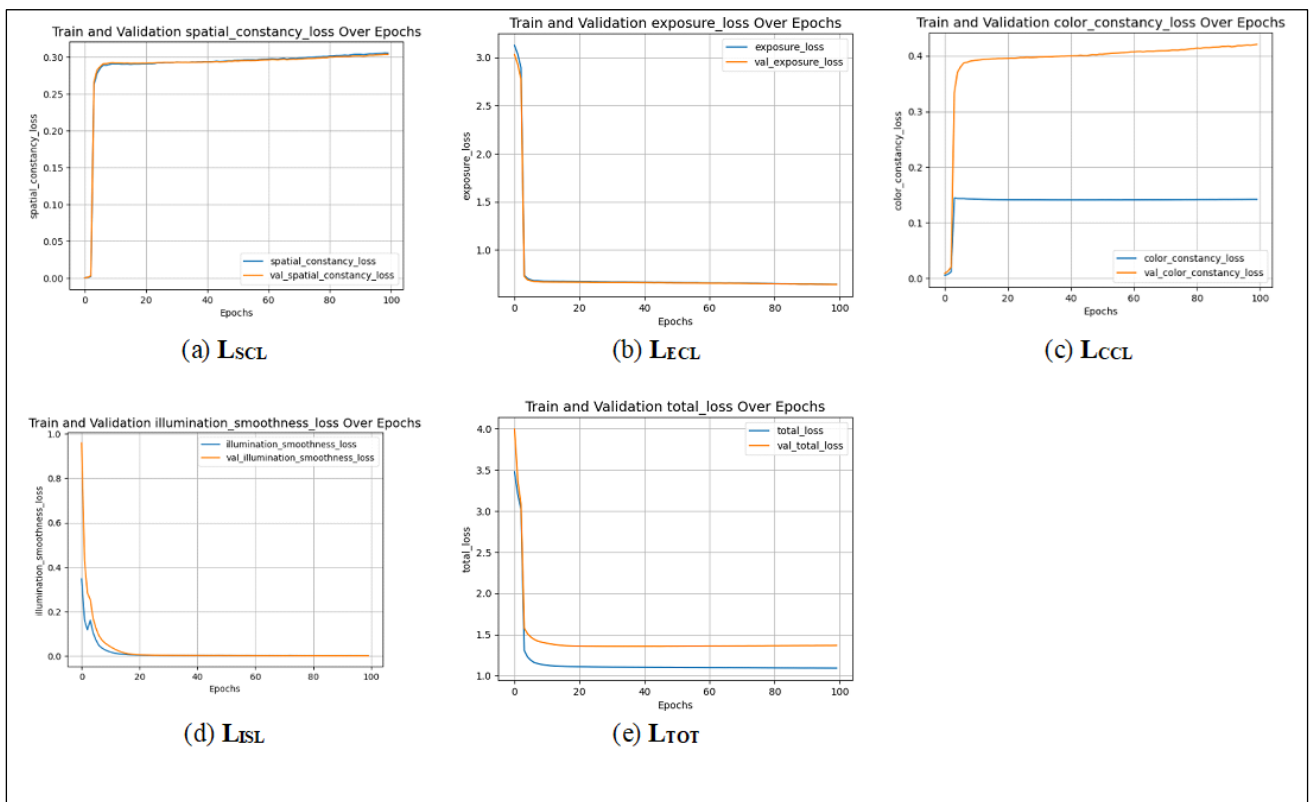


Fig 6. Loss calculations for REllISUR (a) LsCL (b) LECL (c) LCCL (d) Lisl (e) LTOT

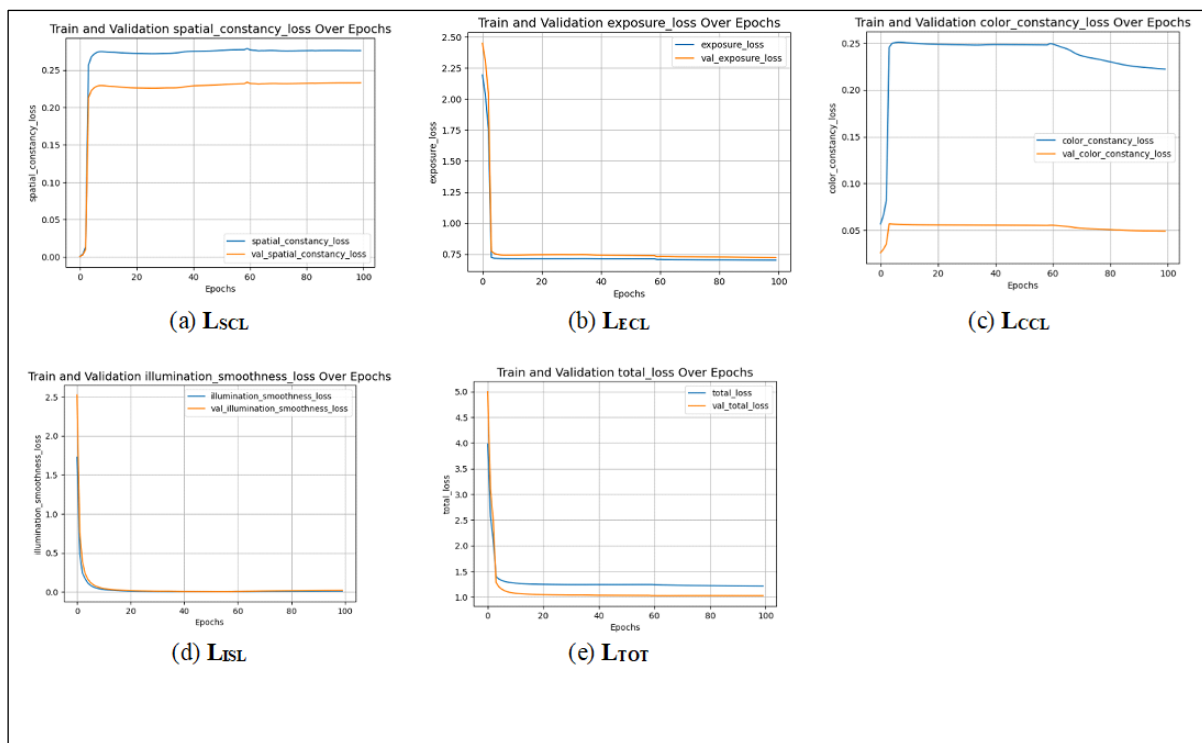


Fig 7. Loss calculations for ExDark (a) LsCL (b) LeCL (c) LcCL (d) Lisl (e) LtOT

Visual and perceptual results

The visual results of applying DCE-Net-based image enhancement to various datasets are shown in Fig 8. It is obvious that this enhancement is evident in normal eyes.



Fig 8. Visual results: (a) Dark face (b) LOL (c) RELLISUR (d) ExDark

The visual and perceptual results help understand the quality of enhanced images based on the human visual system.

Quantitative analysis

To assess the quality of the images, the PSNR and SSIM values were computed for datasets with paired images: LOL, RLLISUR, and ExDark. The quantitative metric SSIM provides a measure of the similarity between two images, the input and enhanced image, quantifying the image quality degradation caused by any processing. The PSNR is another measure of image quality. These metrics show a promising trend for the DCE-Net technique, which yields superior results, as shown in Table 1.

Table 1. Quantitative analysis of paired image datasets

| | LOL | RELLISUR | ExDark | Average |
|------|------|----------|--------|---------|
| SSIM | 0.63 | 0.63 | 0.58 | 0.61 |
| PSNR | 16.7 | 16.4 | 16.6 | 16.56 |

The DCE-Net technique performs very efficiently with easy light estimation curve mapping and a simple lightweight network, leading to reduced computational requirements and time.

Detecting faces in Dark Face dataset

To experiment with the face detection task, we used dark face dataset. Because there were no bounding box reference images for comparison, we used 600 test images to perform the face detection task. A multitask cascaded convolutional neural network (MTCNN) [20] face detector was used for face detection. Owing to different stances, lighting conditions, and occlusions, face detection in an unrestricted context is far more difficult. Deep learning methods have become popular for face detection owing to recent improvements in object detection. This three-stage cascaded architecture creates an image pyramid by resizing an image to various scales using the image as input from joint face detection. In the first stage, candidate face window and bounding box regression vectors were obtained using a convolutional proposition network (P-Net). The obtained bounding-box regression vectors were then used to calibrate the candidates and use non-maximum suppression (NMS) to continue merging candidates with a high degree of overlap.

In the second step, every candidate is routed to a different CNN known as the refined network (R-Net), which carries out NMS, calibrates via bounding box regression, and removes a significant percentage of incorrect selections.

In the third step, the network outputs the locations of the five landmarks on the face. It is well known that certain filters lose their discriminative power during the convolution process owing to a lack of variety, and using a 3×3 filter lowers the computational burden. Following the convolution and completely linked layers, a PReLU (excluding output layers) was employed as the nonlinearity activation function.

Examples of face detection experiments using MTCNN are shown in Fig 9 and Fig 10.

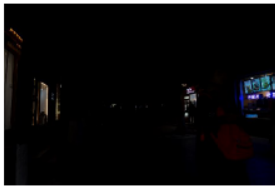
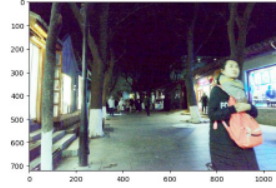
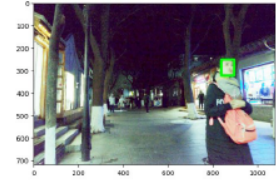



| Original image | Enhanced image | Detected face |
|---|---|--|
|  |  |  |
|  |  |  |

Fig 9. Single face detection

| Original image | Enhanced image | Detected faces |
|---|---|--|
|  |  |  |
|  |  |  |
|  |  |  |

Fig 10. Multiple face detection

The proposed image enhancement technique fails to recover depth details from a very dark image, leading to a loss of contextual information in the background. During the face detection process, this loss of contextual information may hinder the detection of tiny faces. Unconsented facial collections raise privacy concerns. Misuse of personal information raises another threat due to unprotected facial data.

V. CONCLUSION

The curve estimation values of the input image are used by the DCE-Net-based low-light image enhancement system to improve the images. The curve estimation followed by enhancement adjustments was abetted by the non-reference loss functions used. The experiments show better performance of the non-referential method compared to the existing referential methods. The face detection experiment using MTCNN on the enhanced images from the Dark Face dataset provides an excellent performance improvement by the DCE-Net-based system over the existing referential methods. Because the DCE-Net-based system requires only an input image and not a reference image, it reduces the computation and storage requirements for the system, resulting in an improved low-light image enhancement using deep learning. The proposed image enhancement technique can be investigated with more datasets for performance, and its capabilities can be measured using other existing face detection techniques.

Author Contributions

Conceptualization- RG, and KAK; methodology- RG, and KAK; software- RG; validation- RG; writing—original draft preparation- RG; writing—review and editing- RG, and KAK; All authors have read and agreed to the published version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

This research received no external funding.

Competing Interests

There are no competing interests.

References

- [1]. Y. Feng, S. Yu, H. Peng, Y. -R. Li and J. Zhang, "Detect Faces Efficiently: A Survey and Evaluations," in *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 4, no. 1, pp. 1-18, Jan. 2022, doi: 10.1109/TBIOM.2021.3120412.
- [2]. C. Li, C. Guo and C. C. Loy, "Learning to Enhance Low-Light Image via Zero-Reference Deep Curve Estimation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 4225-4238, 1 Aug. 2022, doi: 10.1109/TPAMI.2021.3063604.
- [3]. H. Liang, A. Yu, M. Shao, and Y. Tian, "Multi-Feature Guided Low-Light Image Enhancement," *Applied Sciences*, vol. 11, no. 11, p. 5055, May 2021, doi: <https://doi.org/10.3390/app11115055>.
- [4]. J. Bhattacharya, S. Modi, L. Gregorat, and G. Ramponi, "D2BGAN: A Dark to Bright Image Conversion Model for Quality Enhancement and Analysis Tasks Without Paired Supervision," *IEEE Access*, vol. 10, pp. 57942–57961, 2022, doi: <https://doi.org/10.1109/access.2022.3178698>.
- [5]. J. Hai et al., "R2RNet: Low-light Image Enhancement via Real-low to Real-normal Network," *Journal of Visual Communication and Image Representation*, vol. 90, p. 103712, Feb. 2023, doi: <https://doi.org/10.1016/j.jvcir.2022.103712>.
- [6]. J. Liang et al., "Recurrent Exposure Generation for Low-Light Face Detection," *IEEE Transactions on Multimedia*, vol. 24, pp. 1609–1621, 2022, doi: <https://doi.org/10.1109/tmm.2021.3068840>.
- [7]. W. Wang, W. Yang, and J. Liu, "HLA-Face: Joint High-Low Adaptation for Low Light Face Detection," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, doi: <https://doi.org/10.1109/cvpr46437.2021.01593>.
- [8]. W. Wang, X. Wang, W. Yang, and J. Liu, "Unsupervised Face Detection in the Dark," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 1250–1266, Jan. 2023, doi: <https://doi.org/10.1109/TPAMI.2022.3152562>.
- [9]. L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward Fast, Flexible, and Robust Low-Light Image Enhancement," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2022, doi: <https://doi.org/10.1109/cvpr52688.2022.00555>.
- [10]. J. Hai, Y. Hao, F. Zou, F. Lin, and S. Han, "Advanced RetinexNet: A fully convolutional network for low-light image enhancement," *Signal Processing: Image Communication*, vol. 112, p. 116916, Mar. 2023, doi: <https://doi.org/10.1016/j.image.2022.116916>.
- [11]. H. Gasparyan, S. Hovhannisyanyan, S. V. Babayan, and S. S. Agaian, "Iterative Retinex-Based Decomposition Framework for Low Light Visibility Restoration," *IEEE Access*, vol. 11, pp. 40298–40313, Jan. 2023, doi: <https://doi.org/10.1109/access.2023.3269719>.
- [12]. Q. Lu and P. Gan, "Low-Light Face Recognition and Identity Verification Based on Image Enhancement," *Traitement du Signal*, vol. 39, no. 2, pp. 513–519, Apr. 2022, doi: <https://doi.org/10.18280/ts.390213>.
- [13]. S. Zhang, C. Chi, Z. Lei, and S. Z. Li, "RefineFace: Refinement Neural Network for High Performance Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 11, pp. 4008–4020, Nov. 2021, doi: <https://doi.org/10.1109/tpami.2020.2997456>.
- [14]. C. Li, J. Guo, F. Porikli, and Y. Pang, "LightenNet: A Convolutional Neural Network for weakly illuminated image enhancement," *Pattern Recognition Letters*, vol. 104, pp. 15–22, Mar. 2018, doi: <https://doi.org/10.1016/j.patrec.2018.01.010>.
- [15]. W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling, and R. Yang, "Salient Object Detection in the Deep Learning Era: An In-depth Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2021, doi: <https://doi.org/10.1109/tpami.2021.3051099>.
- [16]. W. Yang et al., "Advancing Image Understanding in Poor Visibility Environments: A Collective Benchmark Study," *IEEE Transactions on Image Processing*, vol. 29, pp. 5737–5752, 2020, doi: <https://doi.org/10.1109/tip.2020.2981922>.

- [17]. W. Chen, W. Wang, W. Yang, and J. Liu, “Deep Retinex Decomposition for Low-Light Enhancement,” arXiv (Cornell University), p. 155, Aug. 2018, doi: <https://doi.org/10.48550/arXiv.1808.04560>.
- [18]. A. Aakerberg, K. Nasrollahi, and T. B. Moeslund, ‘RELLISUR: a real low-light image super-resolution dataset’, in Thirty-fifth Conference on Neural Information Processing Systems-NeurIPS 2021, 2021, doi: 10.5281/zenodo.5234969.
- [19]. Y. P. Loh and C. S. Chan, “Getting to know low-light images with the Exclusively Dark dataset,” *Computer Vision and Image Understanding*, vol. 178, pp. 30–42, Jan. 2019, doi: <https://doi.org/10.1016/j.cviu.2018.10.010>.
- [20]. K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks,” *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016, doi: <https://doi.org/10.1109/lsp.2016.2603342>.