LiCENt: Low-light image enhancement using the light channel of HSL

Atik Garg\textsuperscript{1}, Xin-Wen Pan\textsuperscript{2}, and Lan-Rong Dung\textsuperscript{2}, Member, IEEE
\textsuperscript{1}EECS International Graduate Program, National Yang-Ming Chiao Tung University, Hsinchu, Taiwan
\textsuperscript{2}Department of Electrical and Computer Engineering, National Yang-Ming Chiao Tung University, Hsinchu, Taiwan

Corresponding author: Lan-Rong Dung (e-mail: lennon.dung@gmail.com).

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grant MOST 109-2221-E-009-051.

ABSTRACT Images captured in low-brightness environments often lead to poor visibility and exhibit artifacts such as low brightness, low contrast, and color distortion. These artifacts not only affect the visual perception of the human eye but also decrease the performance of computer vision algorithms. Existing deep learning-based image enhancements methods are quite slow and usually require extensive hardware specifications. Conversely, lightweight enhancement approaches do not provide satisfactory performance as compared to state-of-the-art methods. Therefore, we proposed a fast and lightweight deep learning-based algorithm for performing low-light image enhancement using the light channel of Hue Saturation Lightness (HSL). LiCENt stands for Light Channel Enhancement Network that uses a combination of an autoencoder and convolutional neural network (CNN) to train a low-light enhancer to first improve the illumination and later improve the details of the low-light image in a unified framework. This method used a single channel lightness ‘L’ of HSL color space instead of traditional RGB color channels which helps in reducing the number of learnable parameters by a factor of 8.92, at the most. LiCENt also has significant advantages for the Brilliance Perception Adjustment, which enables the model to avoid issues including over-enhancement and color distortion. The experimental results demonstrate that our approach generalizes well in synthetic and natural low-light images and outperforms other methods in terms of qualitative and quantitative metrics.

INDEX TERMS Deep learning, image processing, low-light enhancement, encoder-decoder architecture, image enhancement, computer vision, convolutional neural network

I. INTRODUCTION

In the last decade, the growth of social platforms such as Facebook, Twitter, and YouTube have created a trend to capture images or videos of everyday activities and share them with others via the internet\cite{1}. However, several uncontrollable factors can degrade the performance of visual devices due to factors such as poor lighting conditions, indoors, at night, or on cloudy days. An inadequate amount of light can cause a loss of detail in the captured images, which not only severely affects the human eyes subjective cognitive abilities but visual device capabilities. Insufficient light significantly reduces the performance of visual device applications, such as object classification, object detection, object tracking, and other technologies that rely on clear details or outlines in the images. Therefore, it is necessary to build a fast, lightweight, and effective method which can be used in conjunction with other computer vision applications and work well on diverse illumination scenarios.

In the current literature, enhancements methods traditionally used histogram equalization\cite{2}, inverse domain operation\cite{3}, and retinex decomposition\cite{4}, to improve low-light images. These methods show acceptable performance in various scenarios; however, enhanced images often remain unsatisfactory in terms of visual perception. Recently, the rapid rise of Convolutional Neural Network (CNN) technology has helped neural/deep learning methods to make great progress in low-light enhancement; however, its limitations are also apparent. These learning-based methods are generally slow and employ complex deep-network structures for learning operations. Large hardware components, such as several GPUs and big RAM, are required for this process to run for a long period. In addition, it is very difficult to procure the training set itself because image data must be obtained for both low-light and regular illumination. Thus, to facilitate learning-based networks, Chen \textit{et al.}\cite{5} proposed raw low-light short-
exposure images, including accompanying reference images acquired from a long exposure. Whereas others synthesized the dataset using gamma adjustment to normal image patches, which may cause undesirable enhancement. To overcome these problems, Lore et al. [6] proposed a stacked-sparse denoising autoencoder that can benefit from instances of synthetic degraded training samples to adaptively enhance low-light images. Similarly, Wang et al. [7] suggested a Deep Lightning Network (DLN) comprising Lightning back-projection (LBP) blocks to iteratively execute lightning and blackening processes for natural light projections to determine the residuals. Wang also proposed a Feature Aggregation (FA) block that adaptively blends the outcomes of various LBPs to effectively use global and regional features. EnlightenGAN [8] uses a Generative Adversarial Network (GAN) to perform unpaired low-light image enhancement by using gray-scale images as an attention map. More recently, a lightweight deep network, Zero-DCE [9] introduced non-reference loss functions, which approximated the performance of enhancement and improved learning for different lighting through incisive and simple nonlinear curve mapping.

Although existing models employ diverse methods for processing low-light images, because of their weakly illuminated structures, they still exhibit under- or over-enhancement artifacts. In the quest to preserve the underlying structures rather than exaggerate color effects, recent methods use a complex mesh of neural structures and require many hardware resources, although it often smoothens out the texture and surface details of objects. In this study, a fast and lightweight method for low-light image enhancement is introduced with the help of an illumination channel. Most of the earlier works used RGB color space to enhance low-light images and retrieve spatial information with the help of a neural model. However, it is difficult to maintain a proper balance between the three color channels of RGB at the output because the three channels are interdependent for color information [10]. This often causes color distortion and information loss. Therefore, the proposed model Light Channel Enhancement Network (LiCEN) used an independent lightness ‘L’ channel to retrieve the lost details from the neural model, while using the unchanged color information from the hue ‘H’ and saturation ‘S’ channels. This enables the model to avoid the problem of over-enhancement and color distortion by observing the issue as Brilliance Perception Adjustment (BPA) while reducing the number of parameters required by the model to 8.92x. The experimental findings show that the proposed approach improves the contrast and the details of low-light, which is significantly better than most existing state approaches. The main contributions of this study are summarized as follows.

- We proposed a lightweight and fast method with the hybrid combination of an autoencoder and CNN layers to improve the illumination and details of the low-light image in a unified framework.
- We proposed that enhancing only the lightness component in the HSL color space has significant advantages for the BPA over other color spaces.
- Our method achieved significantly higher accuracy and reduced the learnable parameters by up to 8.92x as compared to state-of-the-art methods.

The rest of this paper is organized as follows. Section II summarizes related works. Section III introduces the methodology and architecture of the proposed model. Section IV shows the experimental details, optimization, and evaluation results. Section V concludes the paper.

II. RELATED WORK

A. CONVENTIONAL LOW-LIGHT IMAGE ENHANCEMENT

Early low-light enhancement-based techniques applied several handcrafted methods. Some traditional approaches used a variation in HE. Specifically, Yeong-Taeg Kim [11] proposed an intelligent method to preserve the mean brightness of an image to reduce unnecessary visual deterioration, LDR [2] layered the dissimilar representations of...
2D histograms to improve the image contrast by increasing the gray-level differences between neighboring pixels. Turgay and Tardi[12] proposed a method based on inter-pixel contextual information CLAHE[13] and its variants[14] used adaptive histogram equalization to improve image enhancement. HE methods tend to improve the image illumination by compelling the output image to fall into a specific range and often experience information loss and color distortions. Similarly, gamma correction[15, 16] methods map the luminance intensities to compensate for the nonlinear luminance effect but frequently run into the problem of over-enhancement.

Retinex-based methods believe an image is made of reflection and illumination, and use this principle to improve image illumination. In the past, many methods, such as SSR[17], MMR[18], and the McCannps Retinex algorithm[19] used this principle to retrieve and use the illumination map for low-light image enhancement. LIME[20] determined the brightness of each pixel individually by determining the highest value in RGB channels, and JED[21] introduced a retinex model-based decomposition to sequentially assess piece-wise smoothed illumination and noise-suppressed reflectance. Xuesong, Hongxun, et al.[22] proposed a mask-weighted least-squares method to improve the quality of dark images and reduce noise and artifacts without needless reinforcement. Recently, Liu et al.[23] introduced an intensity projection strategy to calculate the transmission based on a rank-one transmission prior for the real-time scene recovery. However, prior-based strategies are do not adjust very well to diverse scenarios, despite priors are usually obtained through a lot of statistics[24].

### B. DEEP LEARNING-BASED

Over the past decade, deep learning-based techniques have shown great promise for low-level vision problems, such as denoising[25], dehazing[23, 26, 27], and tone mapping[28]. Generative adversarial network (GAN)-based methods have also gained considerable popularity for image-enhancement applications. Deep photo enhancer[29] used a GAN-based unpaired learning method for image enhancement, whereas EnlightenGAN[8] proposed an unsupervised GAN to train without low/normal-light image pairs. However, training is unstable in GAN-based methods and require a large amount of data and learnable parameters to obtain the desired output. Lately, several methods, such as Retinex-Net[30], LVENet and LightenNet[31] have used the Retinex principle to approximate the illumination map to design their CNN models. Retinex-Net[30] used an end-to-end image decomposition model with a successive low-light enhancement network to improve illumination, whereas LightenNet[31] used an incredibly small network to enhance the images. Similarly, LVENet[32] proposed to use retinex
theory to estimate the illumination component using a lightweight depthwise separable convolutional. However, their performance is often unsatisfactory, and LightNet[31] unnaturally brightens the center of the image, which varies from the ground truth. Furthermore, the definition of the ground-truth illumination and reflectance elements is not clear, which makes it difficult to guide the training process[33]. Recently, Zero-DCE[9] proposed incisive and simple nonlinear curve mapping with the help of non-reference loss functions; however, as the method only depends on the non-reference loss function, oversaturation of the colors can often be observed in the enhanced results. In contrast, the proposed method uses a very small number of learnable parameters to enhance the low-light image, which is much closer to natural illumination.

III. METHODOLOGY

Conventional low-light enhancement methods have focused on improving global image illumination, which leads to poor local illumination, causing a loss of details in the enhanced image. CNN methods have attempted to solve this issue; however, CNN operations are complex, and deep networks with skip connections are needed to preserve the details, which makes them computationally expensive. Therefore, the proposed method uses a simple but effective architecture that uses the HSL color space to enhance the low-light images.

A. INCLUSION OF COLOR SPACE

The RGB color space is the most widely used and composed of three color components: red, green, and blue. Most of the earlier works[34, 35] used the RGB color space to enhance low-light images. However, as RGB is based on three color channels, it is not very effective for low-light enhancement algorithms to maintain a proper balance between the color channels at the output[36] which results in color distortion. Furthermore, when a dark image contains several small details, it becomes more difficult for the low-light enhancement method to recover the detailed information. In this study, we use the HSL color space to eliminate this problem.

![RGB and HSL color space](image.png)

FIGURE 3. RGB and HSL color space. (Images generated by tuning the code created by Massimo Ciacci [37]).

As shown in Fig. 3, RGB color space only used the ratio between three inter-dependent channels to determine the color/information, but HSL color space separates the color information between hue(‘H’), saturation(‘S’), and lightness(‘L’). The lightness component ‘L’ differs from and unrelated to the chrominance component ‘H’ in the HSL color space., i.e., the color details of a picture and saturation component ‘S’, i.e., the perception of color in an image. When chrominance and saturation remain unchanged, lightness determines the image information. Thus, the proposed method only utilizes the lightness channel to enhance the low-light image while maintaining the same chromaticity of ‘H’ and color saturation ‘S’. Fig. 2(a) shows a flowchart of this enhancement approach. This causes the hue and saturation components to be completely independent of the enhancement process, and the manipulation of the lightness component does not affect the color information of the low-light image. Other color spaces, such as YCbCr and HSV, also provide similar options, but the former causes unnatural enhancement, which often leads to oversaturation in images[38], and the latter has a lower illuminance range [39]. Thus, we preferred to use HLS in this study.

![Effects of Brilliance Perception Adjustment on real-world and synthetically generated images](image.png)

FIGURE 4. Effects of Brilliance Perception Adjustment on real-world and synthetically generated images.

In addition, we take advantage of phenomena that are slightly similar to the retinex theory. We termed this phenomenon the ‘Brilliance Perception Adjustment’ (BPA). In this phenomenon, the perception of brilliance changes to the human eye with the change in the amount of incident light, even though its saturation does not change. For example, the first row in Fig. 4 (a) shows real-world images, which are captured in a time difference of a couple of minutes from the same device, but the human eye can perceive a difference in the color of grass and building due to a change in the amount of sunlight. We perceive this occurrence as a change in brilliance rather than a color change. Using the same concept of nature, we avoided manipulating the chrominance component or even the saturation component directly but adjusted the lightness component of HSL (as shown in Fig. 3(b)), the dark green color denoted by A1 becomes grassy green denoted by A2 with the increase in light), termed as BPA, in low-light
images to obtain images that are much closer to actual natural light images. In the second row of Fig.4(b), we synthetically altered the lightness component of the HSL to obtain a change in the effect of sunlight.

B. NETWORK ARCHITECTURE

Traditional CNN architectures use cascaded structures with zero paddings to maintain the image size, but they cannot efficiently extract useful features from the input image and thus cause distortion in color and structure information in the output image. Therefore, the proposed architecture comprises two parts: brightness enhancement and detail enhancement, as shown in Fig. 2(b), to solve this issue. The details of these two parts are as follows.

1) BRIGHTNESS ENHANCEMENT

Brightness enhancement is an autoencoder network that contains three stages: downsampling, bottleneck, and upsampling.

The autoencoder is an encoder-decoder network that extracts the main feature of the target from the input image and transforms it into a low-dimensional feature representation to reconstruct it into a high-dimensional representation with the help of a decoder[40]. U-Net[41] uses an autoencoder to train unclear medical images to segment the structures within it. This is attained by a shrinking path to capture and a corresponding expanding path to enable precise localization. In this study, we used a similar concept to enhance low-light images. The autoencoder network used in U-Net[41] is implemented in different domains of pixel values, and if we use the original network, it will not provide the maximum performance or might be suitable for our task. Therefore, optimizing the layers and hyperparameters is required before it can be used for low-light image enhancement.

As shown in Fig. 2(a), a low light RGB image is first converted into HSL color space, and then the ‘L’ component is reduced to a fixed training size before being passed into the encoder, while ‘H’ and ‘S’ component is used to change back the output to RGB color space after low-light enhancement process. The encoder network contains seven convolutional layers, and the output size of each layer can be estimated using Eq.(1).

\[
O = \frac{W - K + 2P}{S} + 1
\]  

(1)

where \(O\) is the output size of the layer, \(W\) is the input layer size, \(K\) is the filter size, \(P\) is padding, and \(S\) is stride. In this study, we set \(K, P, S\) as \(3 \times 3, 0\), and 2, respectively. The filter size was fixed to a \(3 \times 3\) configuration because a small filter size better detects smaller details in the image, and a small filter also decreases the number of required parameters for the low-light enhancement network. Each convolutional layer was followed by a ReLU layer. After each layer, the size of the feature map is decreased to perform the encoder downsampling process with the help of convolution operations. We used a stride of 2 to avoid padding because, as mentioned by Cai et al.[42], padding can lead to synthetic artifacts around the boundary of the target image and unnecessary computation burden. It should be noted that the layers are manually optimized to obtain the maximum performance, as shown in Table I. To achieve a better loss during training, the performance of the encoder is transferred through the bottleneck layer, which allows the network to compact feature representations to fit best in the available space. In addition, the bottleneck layer reduces the number of feature maps in the network with the help of \(1 \times 1\) convolutions because it contains fewer output channels than input channels.

The decoder network follows the bottleneck layer and provides a coarse result of the low-light image. The filter size of the decoder network was also fixed to \(3 \times 3\), to better reconstruct smaller details. Some previous studies [40, 42] used deconvolution to reconstruct the feature map, but it caused checkerboard artifacts in the output results. Thus, to avoid such artifacts, we performed bilinear interpolation to increase the size, followed by a convolution layer, as shown in Fig. 5.

As shown in Table II, long skip connections were introduced in the decoder network from its respective mirror encoder network to enable the network to learn residual
features. Residual feature learning helps the network recover high-quality feature maps by providing an alternate backpropagation path to the gradient. Skip connections also solve the issue of gradient explosion/vanishing and decrease the training time of the network by improving its convergence speed.

(a) Ground Truth images

(b) Deconvolution results

(c) Bilinear interpolation results

FIGURE 5. Comparison between Deconvolution & Bilinear Interpolation.

2) DETAIL ENHANCEMENT

Brightness enhancement acts as a coarse enhancement block and improves global illumination; however, smaller details are lost in the output image. The detailed enhancement network solves this issue using a series of convolution layers. The autoencoder output is concatenated with a low-light input image, such that it is possible to retain and relay both unprocessed low-light information and the illumination calculation of brightness enhancement. These concatenated layers were followed by four convolution layers with ReLU to reconstruct a better-quality image with enhanced details.

C. LOSS FUNCTION

The loss function is an important aspect of improving the network to enhance the image under specific illumination conditions. Many previous studies have used the mean absolute error (MAE)[42] and mean squared error (MSE)[31] loss functions. The MAE and MSE loss functions are known as L1 and L2 loss functions, and MAE is used to minimizing the distance, which is the summation of all the absolute errors between the ground truth and the predicted value. Similarly, MSE is used to minimize the distance, which is the summation of all squared differences between the ground truth and the predicted value. These two loss functions work well in many cases, but they often experience artifacts because they only consider the difference between two-pixel values and not the details in the image. This causes the loss function to be stuck in the local optimal solution, which is not suitable for our task. As low-light enhancement does not require specific light conditions, we used a loss function that is not only oriented towards preserving the texture and details but also endures the irregular illuminance around the ground truth. Recently, the SSIM loss function has been used in low-light enhancement, as it considers the human visual system instead of the distance between pixel values. It considers the difference between two luminance signals and normalizes the error/similarity scale in the range 0 to 1. The loss function for the SSIM for pixel p is defined as follows, where ‘N’ is the number of sampled images:

\[
L_{SSIM}(p) = \frac{1}{N} \sum_{x \in X, y \in Y} 1 - SSIM(x_o, y_L) \quad (2)
\]

where 

\[
SSIM(x_o, y_L) = \frac{(2\mu_{x_o}\mu_{y_L} + c_1)\cdot(2\sigma_{x_o y_L} + c_2)}{(\mu_{x_o}^2 + \mu_{y_L}^2 + c_1)\cdot(\sigma_{x_o}^2 + \sigma_{y_L}^2 + c_2)}
\]

SSIM(x_o, y_L) depicts the structural similarity between two images ‘x_o’ & ‘y_L’. The \(\mu_x\) and \(\mu_y\) are the means of the predicted ‘x_o’ and ground truth ‘y_L’ images, respectively. And \(\sigma_{x_o}^2\), \(\sigma_{x_o y_L}\) and \(\sigma_{y_L}\) are the variance-and-covariance-predicted and ground-truth images, respectively. Constants \(c_1\) and \(c_2\) are used to prevent the denominator from being zero (\(c_1 = 0.0001, c_2 = 0.0009\) is the same as that used in [33]). The SSIM loss function requires grayscale images. It does not always depict an accurate representation of the error between the predicted and ground truth, but this issue is solved using our method, resulting in better learning. Detailed training and inference procedures are presented in Algorithm 1.

Algorithm 1: Low light enhancer Network

Pre-process: Convert RGB images into HSL color space and separate them into input \((x_h, x_s, x_l)\) and ground truth \((y_h, y_s, y_L)\) channels.

Input: the lightness channel of low light image x_L

Pre-processing: resize the \(x_L\) to 192x192

Subject: Brightness enhancement network BE, Detail Enhancement network DE

1: repeat until convergence

2. Employ the BE on the \(x_L\) to get coarse output \(x_CO\)
3. Resize the image into input image size
4. Concatenate \(x_CO\) and \(x_L\) to get \(x_CO \_L\) for re-inserting lost information
5. Pass the \(x_CO \_L\) into DE layer to get predicted output \(x_O\)
6. Calculate error with Eq. (2) and update weights

7. After obtaining the final output, \((x_{hp}, x_{hs}, x_{ho})\) is converted back into RGB color space to get enhanced output

IV. EXPERIMENTS AND DISCUSSION

In this section, we describe the implementation details of the proposed low-light enhancement method and evaluate its performance using state-of-the-art methods. First, we present
the dataset and training parameter settings and share the optimization details to improve the performance of the network by adjusting the hyperparameters and layers. We then discuss the details of the experiment to determine the best-suited color space for low-light enhancement. Finally, in terms of improvement, we numerically and visually evaluated the approaches using several state-of-the-art alternatives. Test images of various scenes from the datasets provided by RetinexNet[43], GLADNet[34], LIME[44], and Part2 of the SICE dataset [42].

A. EXPERIMENT DETAILS

In the field of deep learning, it is quite difficult to establish a complete and perfect dataset that can satisfy the diversity in the real world: typically, more extensive training data can enhance the performance of the low-light image enhancement model. The low-light enhancement dataset requires sample images with low illuminance and reference images with a normal illuminance range. However, it is difficult to define low and normal brightness values. Furthermore, the resulting image data will be limited to stationary objects, as capturing dynamic images or street scenes for the collection of training datasets and test datasets are particularly difficult. To this end, normal illuminance images are artificially synthesized to simulate low-brightness situations, while ensuring that scene diversity covers a wide variety of scenes, characters, and brightness situations. In this study, we customized our dataset by combining two publicly available datasets, RetinexNet[43] and GLADNet[34]. RetinexNet[43] and GLADNet[34] have 1 K and 5 K low-light images, respectively, with their ground truth in the normal brightness range. Both datasets emphasize different types of scenes: RetinexNet[43] has mostly outdoor images covering various landscapes, for example, mountains, houses, campuses, clubs, and streets, whereas GLADNet[34] has more indoor images with people performing various activities. Therefore, we combined RetinexNet + GLADNet to obtain a total of 5400 images for training and verification and 600 images for testing. Furthermore, to address the lack of training data, data augmentation was used to solve the problem of insufficient data. We randomly rotated and flipped the images and increased the diversity of our datasets through different permutations and combinations.

All experiments were performed in the Tensorflow-1.9 environment, and other application programming interfaces (APIs) were imported to facilitate our programs. The computer had an Intel(R)Core(TM)i5-8400 processor with a frequency of up to 2.80GHz. It had 16GB of RAM along with an NVIDIA GeForce GTX 1060 GPU with 3GB of memory. For the training process, we used the Adam optimizer to update the network weights and used a hyperparameter configuration largely similar to some of the previous works. We used a batch size of five, and our initial learning was 0.001, which gradually decreased with the help of the exponential decay method. Typically, a fixed learning rate that is either too large or too small may prevent the method from achieving the best solution. For every 100 batches, we reduced the learning rate by 0.96 times, the total number of training epochs was set to 100, which took approximately 3.5hrs.

B. NETWORK OPTIMIZATION

Convolutional networks can search for effective features from low-light images through the convolutional layers to reconstruct enhanced images with better contrast, but an excessive number of layers and kernels can easily increase the computation and skew the corresponding results, affecting the enhanced image quality. Therefore, the enhancement model was optimized to reduce the computational cost. We alienated the optimization process into two parts to optimize our model: the number of encoding and decoding layers in brightness enhancement and the number of convolutional layers in detail enhancement. With the help of optimized neural network parameters, LiCENt achieved better model performance than its alternatives, without being more complicated. The details of the optimization process are as follows.

1) BRIGHTNESS ENHANCEMENT OPTIMIZATION

The autoencoder design in this study was inspired by U-net[41]. According to [44], the deeper the network architecture, the better is the accuracy. Therefore, this study attempts to experiment with six different combinations of layers, from (3,3) to (8,8). These six different combinations ensure that the subsequent feature maps gradually improve the receptive fields of the neural network model such that the bottleneck layer contains information and features from the entire low-brightness image. Therefore, among the options, the size of the 1 × 1 convolutional kernel does not improve the receptive field, and the even-sized convolution kernel does not guarantee the size of the input and output of the neural network. A convolution kernel size of 5 × 5 or larger has large receptive fields, and a single 5 × 5 convolution kernel requires more parameters than two 3 × 3 convolution kernel layers. In addition, two 3 × 3 convolution kernel layers have nonlinear function mappings to improve the prediction capability of the autoencoder.

In Fig. 6, we trained the algorithm using different sets of autoencoder layers (3,3), (8,8). Each set of layers used settings similar to those listed in Tables I and II. In our analysis, the SSIM loss obtained by the model first sharply decreased, but after some epochs, it eventually became stable. However, various sets of autoencoder layers perform differently. A smaller value of the loss function indicates that the trained model can effectively enhance a low-brightness image closer to the reference image. The combination of (6,6) performs worse, whereas (4,4) and (7,7) perform the best among the combinations. Therefore, four - and seven-layer autoencoder networks with a 3 × 3 convolution kernel are used with different versions to effectively improve the
training of the model for better low-light enhancement.

![SSIM Loss vs Epoch for the autoencoder layers](image1.png)

**FIG 6.** SSIM Loss vs Epoch for the autoencoder layers.

### 2) DETAILS ENHANCEMENT OPTIMIZATION

The enhanced image obtained will significantly improve the image brightness and its overall contrast; however, the contours and image details are still blurry if we convert it back to the RGB color space. Therefore, a detail enhancement (DE) block is added and optimized to improve the details with a series of convolutional layers. As shown in Fig. 7, the model performance was analyzed for different numbers of layers. ‘Conv Layer 1’ represents only one layer is used, while ‘Conv Layer 6’ represents six layers are used in the DE block. In our experiment, we observed that when we used a series of four layers (Conv Layer 4) in the DE block, the model provided the minimum loss for 100 epochs. Table III shows the average IQA values and runtimes for different filter number settings and four layers of filter sizes, in which 16-16-16-16 shows the best timing performance and 64-64-64-64 shows the best IQA performance. Therefore, we constructed two architectures in which performance-oriented LiCENt used the (7,7) autoencoder with 64-64-64-64 detail enhancement layers, while an optimized version LiCENt mini used the (4,4) autoencoder with 16-16-16-16 detail enhancement layers.

![SSIM Loss vs Epoch for the Details enhancement layers](image2.png)

**FIGURE 7.** SSIM Loss vs Epoch for the Details enhancement layers

### TABLE III

| Filter number | PSNR | SSIM | FSIM | Runtime (sec) |
|---------------|------|------|------|---------------|
| 16-32-64-128  | 18.5688 | 0.8053 | 0.9409 | 24.42         |
| 128-64-32-16  | 18.8086 | 0.8085 | 0.9416 | 26.88         |
| 16-16-16-16   | 128-19.8633 | 0.8079 | 0.9399 | 13.08         |
| 32-32-32-32   | 18.8157 | 0.8097 | 0.9436 | 15.96         |
| 64-64-64-64   | 19.0033 | 0.8127 | 0.9471 | 24.06         |
| 128-128-128-128 | 18.402 | 0.8120 | 0.9456 | 36.02         |

### C. COLOR SPACE DETERMINATION

The color space is an important but often neglected aspect of image enhancement. As detailed in Section III- (A), RGB color space is the most popular selection for image enhancement. The three-color channels of RGB contain both color and spatial information, which is first compressed and later retrieved with the help of a neural network model. However, some of the information is lost and cannot be retrieved, resulting in the loss of spatial information and color distortion. Some alternative color spaces, such as HSV, HSL, and YCbCr, represent color and illumination information with independent channels. YCbCr, HSV, and HSL define one channel for illuminance intensity reorientation, and the other two channels define the color characteristics. Therefore, to prevent color distortions in RGB, we only enhanced the brightness characteristics in this study while keeping the information of the two color channels separate. To demonstrate the relationship between the channels of the enhanced image, we performed a separate experiment to determine the color space best suited for low-light image enhancement.

![Flowchart for color space determination](image3.png)

**FIGURE 8.** Flowchart for color space determination

For this experiment, we randomly select 2000 bright images from the RetinexNet and GLADNet datasets. Initially, we considered each normal-brightness RGB image as a reference and artificially synthesized the low-brightness image, as shown in Fig. 8. Then, both RGB images are converted into one of the three color spaces, and the converted image has \( I, J, \) and \( K \) channels to represent the image. \( K_I \) and \( K_R \) represent the illuminance channel for low and referenced brightness images, respectively, and the respective representations are used for the two-color channels. In our architecture, the CNN model is designed to enhance \( K_I \) with poor contrast to be quite close to \( K_R \) with good contrast in the most ideal situation. Therefore, we replaced \( I_I, J_I, K_I \) with \( I_L, J_L, K_R \), which represents a good contrast, and the enhanced image was converted back to the RGB color space, which represents the low-brightness image that has been enhanced. Then, the PSNR and SSIM methods...
are used to evaluate the result, as shown in Table IV. In the quantitative and qualitative observations, HSL performed much better than the other alternatives, as shown in Fig. 10 and Table IV. Therefore, we selected the HSL color space for this study.

| Color space | PSNR  | SSIM  |
|-------------|-------|-------|
| YCbCr       | 25.8696 | 0.7783 |
| HSV         | 26.8882 | 0.8525 |
| HSL         | 27.4810 | 0.8767 |

**D. EVALUATION WITH BASELINE METHODS**

We performed performance comparisons with several representative state-of-the-art low-light enhancement methods. MBLLEN[35], KinD[46], DPE[29] and GLADNet[34] are CNN-based methods and also use the autoencoder, LightenNet[31], RUAS[45] and RetinexNet[43] are likewise CNN-based methods but use Retinex theory to obtain the enhanced output, whereas Zero-DCE[9] employs a plain CNN of seven convolutional layers with symmetrical concatenation. Dong[47] and LIME[44] are conventional methods based on the dehazing and retinex models, respectively. The codes of [35], [34], [31], [43], [44], [9], [46], [29], [45] and [47] are from the original authors, and they are compared with our proposed method to show quantitative superiority.

To evaluate the results quantitatively, we used several image quality assessment (IQA) metrics, such as structural
A. Garg et al.: LiCENt: Low-light image enhancement using the light channel of HSL

similarity index (SSIM), peak signal-to-noise ratio (PSNR), and feature similarity index (FSIM) and non-reference metric including Unified No-reference Image Quality and Uncertainty Evaluator (UNIQUE)[48] and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)[49] were used to evaluate the different methods. NPE[50] and LIME[44] datasets were used for non-reference-based comparisons. For an unbiased reference-based comparison, we only used artificially synthesized 600 low-light images from the normal brightness ground truth image of the RetinexNet + GLADNet dataset, as many baseline methods might not perform well on over-exposed images. We also trained and tested our model for the LOL[30] and MIT Adobe 5K[51] datasets, which include 500 and 5000 respective pairs of low/normal brightness images. We reserved 450 and 4500 training images from the respective datasets and the remaining used for the test set. To further prove the effectiveness of our method in a wide range of scenarios, we reconstructed the test environment using the zero-DCE method[9]. As specified in [9], we used the already given low-light images of the Part2[42] subset for testing, which contains 229 multi-exposure sequences and their corresponding ground truth images for each multi-exposure sequence.

1) QUANTITATIVE COMPARISON

As shown in Tables V, VI, and VII, higher SSIM, PSNR, FSIM, and UNIQUE indicate higher accuracy, while in BRISQUE lower is better. In Table V, our methods outperform the baseline methods for the SSIM, PSNR, and FSIM metrics by enhancing only a single color space channel of the HSL. This not only makes the method computationally inexpensive but also helps to preserve the natural colors of the images. Table V also shows our method display competitive performance against representative methods. On most occasions, it outperforms other methods, while at some places it shows the second best performance. In Table VII, similar to our test environment, LiCENt also shows better performance than the other alternative methods. In these evaluations, LiCENt mini(mini) was slightly lagging behind the LiCENt model, but it shows comparative performance against alternative approaches. Table VIII lists the number of learning parameters used in other state-of-the-art neural networks. The method is designed to run in a limited
hardware environment without compromising the performance. The LiCENt mini used far less (up to 8.92/0.18 = 49x appx.) learnable parameters compared to the other methods. Similarly, it also dwarfs other alternative approaches by a vast margin. It should be noted that the parameters in [43], [34], [35], [31], and LiCENt are calculated manually, whereas, in the zero-DCE and RUAS method, they are given in [9] and [45] respectively. Similarly, LiCENt also used fewer parameters than most alternatives. Approaches such as LightenNet[31], RUAS[45], and Zero-DCE[9] used fewer parameters than our performance-oriented LiCENt model, which also used fewer parameters than most alternatives. However, none of these methods are fully deep-learning-based networks. To reduce the number of parameters in their methods, they used some portion of the low-light enhancement that is part manual, whereas other methods such as MBLLEN[35] and LiCENt solely depend on deep learning-based networks.

**FIGURE 12.** (First row) Visual Comparison of different state-of-the-art methods on the no-reference LIME dataset. (Second row) Real-world low-light images. (Third row) Enhancement with the proposed LiCENt method.

**TABLE V**

| Method   | PSNR | SSIM | FSIM |
|----------|------|------|------|
| (a) Input | 11.580 | 0.439 | 0.796 |
| (b) Dong[47] | 16.566 | 0.668 | 0.892 |
| (c) LIME[44] | 16.764 | 0.748 | 0.878 |
| (d) LightenNet[31] | 15.005 | 0.720 | 0.891 |
| (e) RetinexNet[43] | 16.725 | 0.777 | 0.862 |
| (f) MBLLEN[35] | 17.218 | 0.747 | 0.934 |
| (g) GLADNet[34] | 17.452 | 0.753 | 0.939 |
| (h) RUAS[45] | 15.272 | 0.734 | 0.861 |
| (i) KinD[46] | 18.039 | 0.807 | 0.930 |
| (j) Zero-DCE[9] | 17.045 | 0.725 | 0.926 |
| (k) Ours (LiCENt) | 19.003 | 0.813 | 0.947 |
| (l) Ours (LiCENt mini) | 18.151 | 0.810 | 0.936 |

2) VISUAL COMPARISON OF IMAGES FROM OUR DATASET

Although IQA indices can quantify the degree of image quality and evaluate the performance of the method based on the objective analysis of the data, the scores on the IQA indices do not truly reflect the visual experience of our eyes. This degree is also slightly different because human vision is subjective. Therefore, qualitative comparison and evaluation of the relationship between the methods with the naked eye is also very important aspect.

To exemplify the qualitative effectiveness of our method, we performed a visual comparison between the reference image,
low-brightness image, and several state-of-the-art methods

TABLE VI

| Method       | PSNR/SSIM | UNIQUE/BRISQUE | VOLUME XX, 2017 |
|--------------|-----------|----------------|-----------------|
|               |           |                |                 |
| (a) LIME[44] | 14.783    | 0.798          | 0.773           |
| (b) RetinexNet[43] | 13.547    | 0.767          | 0.816           |
| (c) LightenNet[31] | 14.227    | 0.703          | 0.764           |
| (d) GLADNet[34]  | 16.371    | 0.758          | 0.785           |
| (e) DPE[29]     | 16.125    | 0.730          | 0.725           |
| (f) RUAS[45]    | 18.227    | 0.971          | 0.711           |
| (g) KinD[46]    | 18.407    | 0.759          | 0.798           |
| (h) Zero-DCE[9] | 17.742    | 0.828          | 0.819           |
| (i) Ours       | 18.444    | 0.835          | 0.842           |
| (j) Ours (mini)| 17.273    | 0.801          | 0.816           |

TABLE VII

| Method       | PSNR/SSIM | SRIE[52] | LIME[44] | Li et al.[33] | RetinexNet[43] | LightenNet[31] | EnlightenGAN[8] | LIME[44] | LIME[44] | LIME[44] | LIME[44] | LIME[44] |
|--------------|-----------|----------|----------|---------------|----------------|----------------|-----------------|----------|----------|----------|----------|----------|
|               |           | 14.14    | 16.17    | 15.19         | 15.99          | 13.52          | 16.21           | 17.273   | 16.809   | 13.547   | 13.52    | 13.52    |

TABLE VIII

| Method       | Params | Ratio | Type | VOLUME XX, 2017 |
|--------------|--------|-------|------|-----------------|
| LightenNet[31]| 28.9K  | 0.28x | Quasi|                 |
| RetinexNet[43]| 612.5K | 5.91x | Fully|                 |
| MBLE[N[35]    | 275K   | 2.64x | Fully|                 |
| GLADNet[34]   | 928K   | 8.92x | Fully|                 |
| RUAS[45]      | 33K    | 0.31x | Quasi|                 |
| Ours (LiCEN)  | 104K   | 1.0x  | Fully|                 |
| Ours (LiCEN mini) | 19K | 0.18x | Fully|                 |

along with the proposed method. Simultaneously, we also attach the IQA indicators (PSNR/SSIM) of each image as a basis for comparison to assist in the analysis. Fig. 9 and Fig. 10 show over-and-under-enhanced issues with alternative approaches, which are often encountered in low-brightness image enhancement. In several cases, enhanced pixel intensity values are mainly concentrated in the low-intensity value areas. Thus, the excessive enhancement will cause the enhanced image to be unsatisfactory in terms of visual perception. Conventional methods, such as Dong[47] and LIME[44] require manual adjustment of the parameters and attempt to improve the overall visibility of the scene. Therefore, they sometimes show acceptable performance in some regions, but particularly in darker and brighter regions, they often exhibit under-and over-exposure. LightenNet[31] is a retinex-based method that extracts information locally; therefore, in situations where lighting conditions are poor, it is unable to enhance the darker regions of the low-light image. Similarly, RUAS[45] also enhances darker regions poorly, but its prior architecture search mechanism often fails to exploit the latent structures low light image. Although MBLLEN[35] can present richer information and textures, there remains a significant color cast influence, resulting in a change that is not adequately natural, and plateaus appear. CNN-based methods MBLE[N[35], KinD[46] GLADNet[34], and Zero-DCE[9] use color and spatial information to enhance the low-light image, but improving both sets of information makes the situation complex, thus leading to loss of detail, texture, and color distortion. In summary, the suggested LiCEN approach achieves high performance and shows harmony between the lighting, contrast, and texture information to effectively increase the illumination of low-light images.

3) COMPARISON WITH REAL-WORLD IMAGES

We tested our methods on real low-light images and synthetically produced datasets. The proposed approach worked well on natural low-light images, was comparable to synthetic datasets, and surpassed the efficiency of existing approaches. In Fig. 12, we illustrate the visual distinction between existing methodologies and reference methods.

E. NUMERICAL EXPERIMENTS

In this section, we show the potential application for the proposed method. All the numerical experiments were conducted on Intel(R)Core (TM)i9-10900 processor. It has 23GB of RAM along with an NVIDIA GeForce RTX 2060S GPU with 8GB of memory.

OBJECT DETECTION

Object detection has observed tremendous growth with the rise of various state-of-the-art algorithms. However, the visual quality of images captured by cameras keeps changing throughout the day. This also affects the performance of object detection methods. Therefore, we compared our method with the various representative image enhancement techniques. Fig. 13 shows that different daylight conditions can affect the performance of object detection algorithms[55]. KinD[46] under-enhanced the darker regions, which leads to poor object detection, while RUAS[45] did a better job in enhancing the brighter regions but also over smoothed the images regions. This not only causes the object detection algorithm to fail to detect several objects but also decreases the detection confidence. MBLE[N[35] shows good performance in terms of visual
perception and object detection but it is a very slow method as shown in Table IX. LiCENt and LiCENtミニ show similar performance for object detection, but LiCENtミニ is quite faster than compared to other methods. The running time of all methods is calculated on the test set of our combined (RetinexNet + GLADNet) dataset. LiCENtミニ outperforms other approaches by a vast margin. It only takes 3.5ms to process the image, which makes it quite suitable to use in conjunction with object detection algorithms.

outperforms various state-of-the-art algorithms. Likewise, we show that utilizing the BPA phenomenon yields more visually pleasing enhanced images that are much closer to a natural illumination environment. Our future work will explore using a GAN-based approach for detail enhancement instead of conventional CNN layers to further improve the information in the enhanced image.

V. CONCLUSION

In this paper, we propose a solution for the low-light enhancement problem using a novel and flexible framework. The proposed LiCENt method operates on a single lightness channel of the HSL color space and generalizes well in a diverse number of imaging scenarios. Experimental observations on multiple low-light datasets indicate that, under both quantitative and qualitative measures, our system

![Image](https://example.com/image.png)

**FIG 13. Visual comparison for YOLOv4-tiny[55] results for various image enhancement methods**

| Method    | KinD[46] | MBLLEN[35] | RUAS[45] | Ours | Ours (mini) |
|-----------|----------|------------|----------|------|-------------|
| CPU(S)    | 0.5650   | 0.3283     | 0.1028   | 0.1674 | 0.0371      |
| GPU(S)    | -        | 0.0974     | 0.0089   | 0.0091 | 0.0035      |

**TABLE IX**

**COMPARISON OF AVERAGE RUN TIME FOR COMBINED (RETNEXNET + GLADNET) DATASET**

REFERENCES

[1] R. Gonzalez, R. Woods, and S. Eddins, “Segmentation using the watershed algorithm,” ed: Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2004, pp. 417-425.

[2] C. Lee, C. Lee, and C.-S. Kim, “Contrast enhancement based on layerd difference representation of 2D histograms,” *IEEE transactions on image processing*, vol. 22, no. 12, pp. 5372-5384, 2013.

[3] L. Zhang et al., “Simultaneous enhancement and noise reduction of a single low-light image,” *IET Image Processing*, vol. 10, no. 11, pp. 840-847, 2016.

[4] X. Ren, W. Yang, W.-H. Cheng, and J. Liu, “L3M: Robust low-light enhancement via low-rank regularized retinex model,” *IEEE Transactions on Image Processing*, vol. 29, pp. 5862-5876, 2020.

[5] C. Chen, Q. Chen, J. Xu, and V. Koltun, “Learning to see in the dark,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3291-3300.

[6] K. G. Lore, A. Akintayo, and S. Sarkar, “LLNet: A deep autoencoder approach to natural low-light image enhancement,” *Pattern Recognition*, vol. 61, pp. 650-662, 2017.

[7] L.-W. Wang, Z.-S. Liu, W.-C. Siu, and D. P. Lun, “Lightening network for low-light image enhancement,” *IEEE Transactions on Image Processing*, vol. 29, pp. 7984-7996, 2020.

[8] Y. Jiang et al., “Enlightengan: Deep light enhancement without paired supervision,” *IEEE Transactions on Image Processing*, vol. 30, pp. 2340-2349, 2021.

[9] C. 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.

[10] W. Wang, Z. Chen, X. Yuan, and X. Wu, “Adaptive image enhancement method for correcting low-illumination images,” *Information Sciences*, vol. 496, pp. 25-41, 2019.

[11] Y.-T. Kim, “Contrast enhancement using brightness preserving histogram equalization,” *IEEE transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1-8, 1997.

[12] T. Celik and T. Tjahjadi, “Contextual and variational contrast enhancement,” *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3431-3441, 2011.

[13] A. M. Reza, “Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement,” *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 38, no. 1, pp. 35-44, 2004.

[14] M. S. Hitam, E. A. Awaluddin, W. N. I. H. W. Yussof, and Z. Bachok, “Mixture contrast limited adaptive histogram equalization for underwater image enhancement,” in *2013 International conference on computer applications technology (ICCAT)*, 2013: IEEE, pp. 1-5.

[15] H. Farid, “Blind inverse gamma correction,” *IEEE transactions on image processing*, vol. 10, no. 10, pp. 1428-1433, 2001.

[16] Y. Shi, J. Yang, and R. Wu, “Reducing illumination based on nonlinear gamma correction,” in *2007 IEEE International Conference on Image Processing*, 2007: IEEE, pp. I-529-I-532.

[17] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, “Properties and performance of a center/surround retinex,” *IEEE transactions on image processing*, vol. 6, no. 3, pp. 451-462, 1997.

[18] Z.-u. Rahman, D. J. Jobson, and G. A. Woodell, “Multiscale retinex for color rendition and dynamic range compression,” in *Applications of Digital Image Processing XIX*, 1996: vol. 2847: International Society for Optics and Photonics, pp. 183-191.

[19] J. McCann, “Lessons learned from monodr:voids applied to real images and color gamuts,” in *Color and imaging conference*, 1999, vol. 1999, no. 1: Society for Imaging Science and Technology, pp. 1-8.
[20] X. Guo, Y. Li, and H. Ling, “LIME: Low-light image enhancement via illumination map estimation,” IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982-993, 2016.

[21] X. Ren, M. Li, W.-H. Cheng, and J. Liu, “Joint enhancement and denoising method via sequential decomposition,” in 2018 IEEE International Symposium on Circuits and Systems (ISCAS), 2018, IEEE, pp. 1-5.

[22] X. Jiang, H. Yao, and D. Liu, “Nighttime image enhancement based on image decomposition,” Signal, Image and Video Processing, vol. 13, no. 1, pp. 189-197, 2019.

[23] J. Liu, W. Liu, J. Sun, and T. Zeng, “Rank-one prior: Toward real-time scene recovery,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 14802-14810.

[24] Y. Liu, H. Yin, J. Wan, Z. Liu, and A. Chong, “EdgeAware Network for Image Dehazing,” IEEE Signal Processing Letters, vol. 29, pp. 174-178, 2021.

[25] J. Chen, J. Chen, H. Chao, and M. Yang, “Image blind denoising with generative adversarial network based noise modeling,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3155-3164.

[26] B. Cai, X. Xu, K. Jia, C. Qiang, and D. Tao, “Dehazenet: An end-to-end system for single image haze removal,” IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187-5198, 2016.

[27] Q. Zhu, J. Mai, and L. Shao, “A fast single image haze removal algorithm using color attenuation prior,” IEEE transactions on image processing, vol. 24, no. 11, pp. 3522-3533, 2015.

[28] A. Rana, P. Singh, G. Valenzise, F. Dufaux, N. Komodakis, and A. Smolic, “Deep tone mapping operator for high dynamic range images,” IEEE Transactions on Image Processing, vol. 29, pp. 1285-1298, 2019.

[29] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, “Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6306-6314.

[30] C. Wei, W. Wang, W. Yang, and J. Liu, “Deep retinex decomposition for low-light enhancement,” arXiv preprint arXiv:1808.04560, 2018.

[31] 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, 2018.

[32] Y. Guo, Y. Lu, and R. W. Liu, “Lightweight deep network-enabled real-time low visibility enhancement for promoting vessel detection in maritime video surveillance,” The Journal of Navigation, pp. 1-21.

[33] X. Zhang and X. Wang, “MARN: Multi-scale attention retina network for low-light image enhancement,” IEEE Access, 2021.

[34] W. Wang, C. Wei, W. Yang, and J. Liu, “GLADNet: Low-light enhancement network with global awareness,” in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018, IEEE, pp. 751-755.

[35] F. Lv, F. Lu, J. Wu, and C. Lim, “MBLLEN: Low-Light Image/Video Enhancement Using CNNs,” in BMVC, 2018, p. 220.

[36] W. Wang, X. Wu, X. Yuan, and Z. Gao, “An Experiment-Based Review of Low-Light Image Enhancement Methods,” IEEE Access, vol. 8, pp. 87884-87817, 2020.

[37] M. Ciaucci. “3D color spaces.” MathWorks. https://www.mathworks.com/matlabcentral/fileexchange/60939-3d-color-spaces (accessed December 10, 2021).

[38] J.-S. Chiang, C.-H. Hsia, H.-W. Peng, and C.-H. Lien, “Color image enhancement with saturation adjustment method,” Journal of Applied Science and Engineering, vol. 17, no. 4, pp. 341-352, 2014.

[39] M. Sarfuddin and R. Missaoui, “A new perceptually uniform color space with associated color similarity measure for content-based image and video retrieval,” in Proc. of ACM SIGIR 2005 workshop on multimedia information retrieval (MMIR 2005), 2005, pp. 1-8.

[40] G. Eifertsen, J. Kronander, G. Danes, R. K. Mantituk, and J. Unger, “HDR image reconstruction from a single exposure using deep CNNs,” ACM transactions on graphics (TOG), vol. 36, no. 6, pp. 1-15, 2017.

[41] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in International Conference on Medical image computing and computer-assisted intervention, 2015, Springer, pp. 234-241.

[42] J. Cai, S. Gu, and L. Zhang, “Learning a deep single image contrast enhancer from multi-exposure images,” IEEE Transactions on Image Processing, vol. 27, no. 4, pp. 2049-2062, 2018.

[43] C. Wei, W. Wang, W. Yang, and J. Liu, “Deep retinex decomposition for low-light enhancement,” arXiv preprint arXiv:1804.05560, 2018.

[44] X. Guo, “Lime: A method for low-light image enhancement,” in Proceedings of the 24th ACM international conference on Multimedia, 2016, pp. 87-91.

[45] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, “Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 10561-10570.

[46] Y. Zhang, J. Zhang, and X. Guo, “Kindling the darkness: A practical low-light image enhancement,” in Proceedings of the 27th ACM international conference on multimedia, 2019, pp. 1632-1640.

[47] X. Dong et al., “Fast efficient algorithm for enhancement of low lighting video,” in 2011 IEEE International Conference on Multimedia and Expo, 2011, pp. 1-6.

[48] W. Zhang, K. Ma, G. Zhai, and X. Yang, “Uncertainty-aware blind image quality assessment in the laboratory and wild,” IEEE Transactions on Image Processing, vol. 30, pp. 3474-3486, 2021.

[49] A. Mittal, A. K. Morthoy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” IEEE Transactions on image processing, vol. 21, no. 12, pp. 4695-4708, 2012.

[50] S. Wang, J. Zheng, H.-M. Hu, and B. Li, “Naturalness preserved enhancement algorithm for non-uniform illumination images,” IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3538-3548, 2013.

[51] V. Bychkovskiy, S. Paris, E. Chan, and F. Durand, “Learning photographic global tonal adjustment with a database of input/output image pairs,” in CVPR 2011, 2011: IEEE, pp. 97-104.

[52] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, “A weighted variational model for simultaneous reflectance and illumination estimation,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2782-2790.

[53] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, “Structure-revealing low-light image enhancement via robust retinex model,” IEEE Transactions on Image Processing, vol. 27, no. 6, pp. 2828-2841, 2018.

[54] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, “Underexposed photo enhancement using deep illumination estimation,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 6849-6857.

[55] A. Bochkovskiy. “DarkNet: Open Source Neural Networks in Python.” https://github.com/AlexeyAB/darknet (accessed January 15, 2022).

ATIK GARG is currently pursuing his Ph.D. degree in Computer Science from the National Yang-Ming Chiao-Tung University, Hsinchu, Taiwan. He received his B.E. degree in Instrumentation and Control Engineering from Sant Longowal Institute of Engineering and Technology, Longowal, Sangrur, Punjab, India, in 2016 and the M.S. degree from the National Chiao Tung University, Hsinchu, Taiwan, in 2019. From 2016 to 2017, he was a Research Assistant in Hong Kong polytechnic University, Hong Kong. His research interests include computer vision with deep learning and image/video processing.

Xin-Wen Pan received the B.S. degree from Department of Electrical Engineering, National Chi Nan University, Nantou, Taiwan, in 2018. In 2021, he received the M.S. degree in Electrical and Control Engineering, College of Electrical and Computer Engineering from National Chiao Tung University, Hsinchu, Taiwan. His research interests include computer vision and image processing with deep learning.
Lan-Rong Dung received a BSEE and the Best Student Award from Feng-Chia University, Taiwan, in 1988, an MS in electronics engineering from National Chiao Tung University, Taiwan, in 1990, and Ph.D. in electrical and computer engineering from Georgia Institute of Technology, in 1997. He joined the faculty of National Chiao Tung University, Taiwan in 1999 where he is currently a full professor in the Department of Electrical and Control Engineering. He is a member of Circuits and Systems and Signal Processing societies of the IEEE.

From 1997 to 1999 he was with Rockwell Science Center, Thousand Oaks, CA, as a Member of the Technical Staff. His current research interests include multimedia architecture, digital image processing, and low-power VLSI design.

Prof. Dung received the VHDL International Outstanding Dissertation Award celebrating in Washington DC in October 1997.