Image Contrast Enhancement using Block based CNN Learner

Pooja Patel and Arpana Bhandari

June 19, 2020

Image Contrast Enhancement using Block based CNN Learner

Pooja Patel
M.Tech Scholar
Department of CSE
SIRT, Bhopal
patel.pooja83@rediffmail.com

Arpana Bhandari
Assistant Professor
Department of CSE
SIRT, Bhopal
arpanabhandari08@gmail.com

Abstract: Image enhancement is one of the difficult problems of image processing. The purpose of image enhancement is to process an image so that the result is more suitable for a particular application than the original image. Digital image enhancement techniques offer various ways to improve the visual quality of images. The appropriate selection of these techniques is very important. Producing the natural scene with good contrast, vivid color and rich details is an essential goal of digital photography. The acquired images, however, are often low contrast because of poor lighting conditions and the limited dynamic range of imaging device. Contrast enhancement is thus an important step to improve the quality of recorded images and make the image details more visible. In this paper, a block-based methodology is proposed for contrast enhancement of low contrast indoor and outdoor images with reference images. The low contrast images are adjusted with CNN Adaptive Bilateral Enhancer with reference image blocks.

Keywords: Image Enhancement, Digital Image Processing, Convolution Neural Network, Bilateral Filter, Quality Measure.

I. INTRODUCTION

Image enhancement is image processing to improve certain features of an image. Image enhancement essentially improves the perception of information or the interpretation of images for the human viewer and provides better input for other automated image processing techniques. The image captured from a natural environment with a high dynamic range includes light and dark areas [1]. Due to the increasing dynamic detection range of the human eye, these images are very difficult for human eyes to perceive [2]. Image enhancement is a technique for improving the quality of these images and changing the properties of an image to make it more suitable for a particular task and viewer. Accentuates edges, edges and contrasts to make image display more useful for analysis and visualization [3]. The extension does not increase the content of the data, while helping to increase the dynamic range of the functions selected for improvement so that they can be easily recognized. There are numerous techniques that represent an image enhancement process that require improving the visual appearance of an image or converting an image into a better image that is easy for humans or machines to analyze [4]. Image enhancement is mainly used when you want to remove noise, enhance dark images and highlight the edges of the object in an image. For some specific applications, the result obtained is more suitable than the original image [5]. Image enhancement techniques can be divided into two main categories:

A. *Spatial Domain Enhancement Method*

The name domain refers to the collection of pixels composing an image. Spatial domain techniques are procedures that work directly on these composed pixels.

In the spatial domain image enhancement technique, transformations are directly applied on the pixels. The pixel values are manipulated to achieve desired enhancement. Spatial domain processes can be expressed as:

$$g(x, y) = T [f (x, y)],$$

where $f(x, y)$ is the input image

$g(x, y)$is the processed image

and T is an operator on f,

defined over the neighborhood of (x, y).

Spatial domain techniques like the logarithmic transforms, power law transforms, histogram equalization are based in the direct manipulation of the pixels in image.

Spatial domain methods can again be divided into two categories:

- point processing

- spatial filtering operations

B. Frequency Domain method

Frequency domain image enhancement is straightforward. The frequency filters developed an image in the frequency domain. This type filtering technique is very simple:

1. Transform the image into the Fourier domain.

2. Multiply the image by the filter.

3. Take the inverse transform of the image.

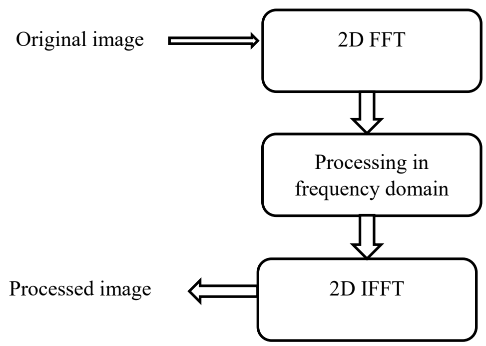


Fig.1.The frequency filters process

The overall visibility or quality of the image can be improved without introducing an unrealistic visual facade and irrelevant artifacts. The normal method of improving overall contrast substantially increases luminance for bright pixels and decreases luminance for dark pixels [6]. Therefore, it is desirable to improve the contrast according to the neighborhood in order to obtain sufficient contrast to improve the image without losing the compression of the dynamic range [7]. It can be divided into several categories as follows:

A. Linear Contrast Enhancement

Increasing linear contrast is also called contrast stretching. The original image can be extended linearly to a new distribution. The full sensitivity range of the digital camera can be used by increasing the original value of an image. This enhancement method can be used primarily in remote sensing images [8][9].

B. Non Linear Contrast Enhancement

Using an algorithm, nonlinear contrast enhancement implies the histogram equalization method. The limitation of nonlinear contrast enhancement is that each value in the input image has multiple values in the output image, causing the original object to lose its exact brightness [10].

II. LITERATURE REVIEW

Steffi et al. [1] proposed a principal component analysis framework to enhance low-light-level images with decomposed luminance chrominance components and obtained SSIM is about 76% and PSNR value is about 18. This algorithm fails to recover the details for night vision images as well similarity value between enhanced and original image is low.

X. Guo et al. [2] proposed low-light image enhancement (LIME) method. The illumination of each pixel is first estimated individually by finding the maximum value in R, G, and B channels. Furthermore, they refine the initial illumination map. Lightness Order Error (LOE) obtained is 2.394 with high error rate.

Kede Ma et al. [3] proposed a patch decomposition approach for multi-exposure image fusion (MEF). They decomposed an image into three independent components: Signal strength, Signal structure, and Mean intensity. Upon fusing these three components separately, they reconstruct a desired patch and place it back into the fused image. The quality obtained is 0.982 with high computational complexity.

X Fu et al. [4] proposed a weighted variational model to estimate both the reflectance and the illumination from an observed image. Average computational times in seconds is 13.5 and PSNR is approx. 17 that exhibits some color distortions.

Gomez-Ojeda et al. [5] this article proposes a small size convolutional neural network (CNN) that can function faster. Finally, author validated the extended representations by evaluating the sequences generated by the two architectures in different advanced VO algorithms such as ORB-SLAM and DSO.

Xiao B. et al. [6] this article suggests histogram-based image enhancement techniques to improve contrast in images. However, most histogram-based image enhancement methods cannot freely adjust the brightness and contrast of the enhanced image. In this paper, two new image enhancement algorithms based on a histogram are proposed. The proposed algorithms offer the possibility to control the brightness and contrast of an improved image by adapting two parameters. The principles for selecting parameters are also discussed in this document. The experimental results show a better performance of the proposed methods in terms of quality of perception and evaluation of image quality compared to existing methods based on the histogram.

III. Methodology

In this technology era, there are several improvements as well as developments in the field of image sensors and cameras. In near future, this technology will be as much developed that how human visual system perceives natural scenes, cameras can capture similar images. However, the range of light intensity that any camera can capture is in between $2^8 - 2^{14}$. This intensity is measured in term stop which is base 2 logarithm of the dynamic range. It is known that the DSLR camera can capture with an about 8 to 12 stops dynamic range whereas the human visual system can perceive the dynamic range more than 24 stops. With the high dynamic range of the camera, there may causes the underexposed as well as over-exposed conditions which needs to be adjusted [3].

The dynamic range of the cameras are quite low as compared to the natural scenes.

To achieve the objectives mentioned above, following steps will be performed:

Data collection: The training data is collected by different cameras and from different scenes.

Reference Image Generation: Having the candidate sequences, high-quality reference images are generated.

CNN-BASED learning: With the constructed dataset, we can design a CNN based bilateral enhancer to learn with its corresponding reference image to achieve this goal.

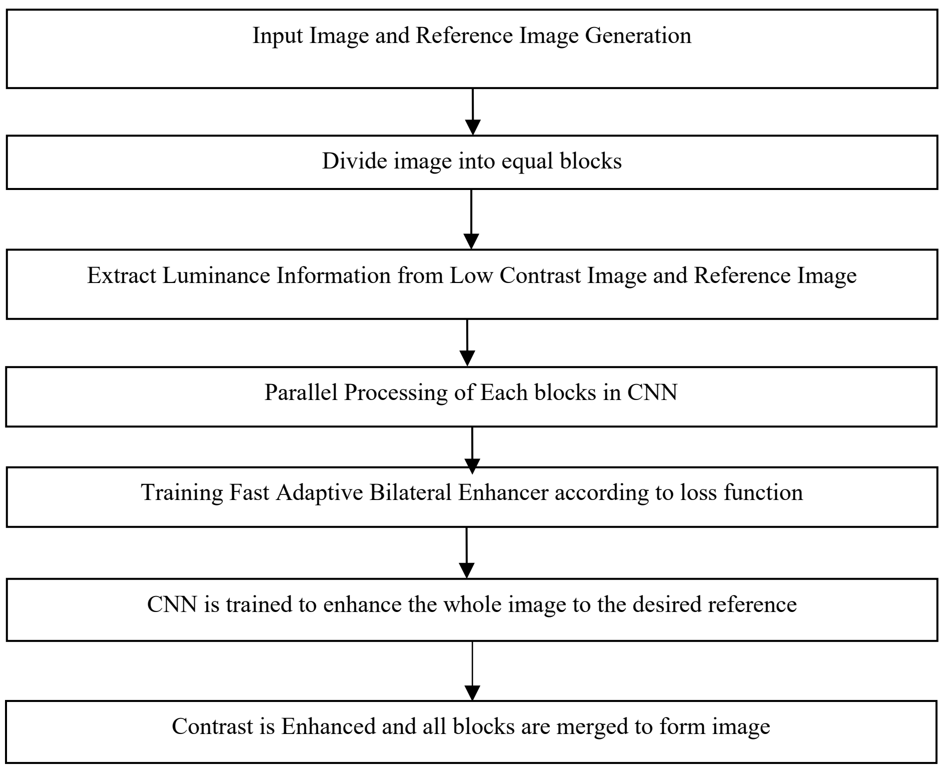


Fig. 2. Flowchart of Proposed Methodology

A. Input Image

In order to design robust and efficient technique, it is required to collect input images from real-world natural scene. In this research, different image sequences are collected from different camera and collected as a common dataset. For creating low contrast image dataset, the exposure value of the camera are set and collected different sequences of indoor and outdoor scenes.

B. Reference Image Generation

For reference image generation, High Dynamic Range (HDR) algorithm is used to generate under exposed and over exposed high-quality reference images.

C. Divide image into equal blocks

In this step, input image is resized into 1024*1024 pixels which is further divided into eight equal blocks of 128*128 pixels.

D. Extract Luminance Information

Luminance information is extracted and processed in Convolution Neural Network (CNN). Luminance measures perceived “gray-level” of pixel and is calculated as:

$$I(x, y) = L(x, y) + H(x, y)$$

Where, $L(x , y)$ = Low frequency luminance Information

$H(x, y)$ = High frequency luminance Information

CNN learn the mapping function between the $L(x , y)$ and $H(x, y)$ of low contrast input image $I(x, y)$ and its corresponding reference image $I_{ref}(x, y)$.

E. *Parallel Processing of Each blocks in CNN*

Each image blocks are individually processed in CNN blocks parallelly. To combat the issues of fully connected neural network, convolutional neural networks (CNNs) were created and will be used as the foundation of this dissertation. These are a type of neural network which is designed specifically to be used with images, and differ slightly from traditional neural network structure. Another key property of CNNs is that they are not fully connected, where every node in a layer is connected to every node in the previous layer. There are four main elements to a CNN:

- Convolutional layer
- Rectified Linear Unit
- Pooling layer
- Fully connected layer

Stride Convolution: The feature map generated according to the network is reduced by the convolutional operations. Padding is done to the output images before performing convolution, as output image has to be of same size as that of input image [5]. But this padding may cause artifact in the input image. So, this network is designed with deconvolution layer to make the output size be similar to input size. This deconvolution layer not only decreases the artifacts as well as reduces the computational overhead by applying filters.

Rectified Linear Unit: Rectified Linear Unit (ReLU) are used in many CNN architectures as an activation function for the network. In this activation function, the negative co-efficient are replaced with zero value which is represented by the local features of the input image. The function is represented as:

$$f(x) = \begin{Bmatrix} 0 \; for \; x < 0 \\ x \; for \; x \; \geq 0 \end{Bmatrix}$$

Some of the neurons dropped because they do not contribute to forward passage and do not participate in backpropagation. Every time an input is presented, the neural network analyzes another architecture, but all these architectures share a common weight. This technique reduces the complex adaptations of neurons because a neuron cannot rely on the presence of some other neurons.

Pooling Layer: The pooling layer is used only to reduce the dimensionality of the previous level so that it is more suitable for the next level of the network. This is generally done with the maximum grouping, in which the maximum value of the window that the filter is observing is folded on the image.

Fully-Connected Layer: The fully connected layer of CNN is a normal neural network and is generally used as a last step in a convolutional network.

Table I shows the CNN model configuration.

Table I: Proposed CNN Model Configuration

Layer	Filters	Kernel Size	Stride	Output size
Conv	96	11*11	4	11*11*96
Pooling	N/A	3*3	2	N/A
Conv	256	5*5	1	5*5*256
Pooling	N/A	3*3	2	N/A
Conv	384	3*3	1	3*3*384
Pooling	N/A	3*3	2	N/A
Conv	384	3*3	1	3*3*384
Pooling	N/A	3*3	2	N/A
Conv	256	3*3	1	3*3*256
Pooling	N/A	3*3	2	N/A
Fully connected	N/A	N/A	N/A	3*3*256

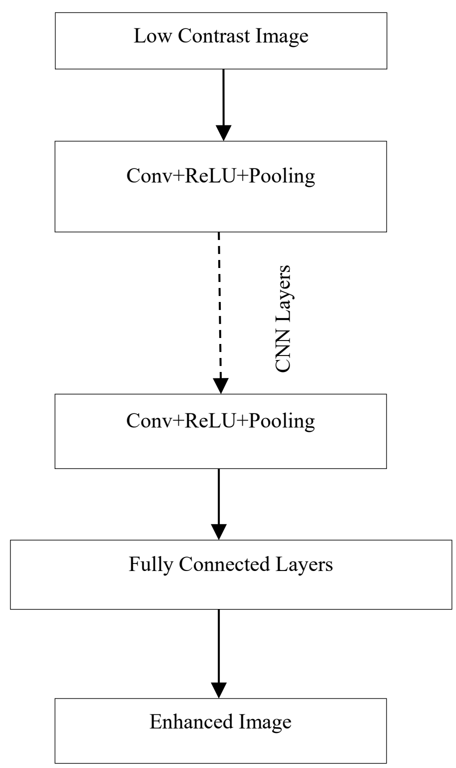


Fig.3. Proposed CNN Architecture

F. Training Fast Adaptive Bilateral Enhancer

With the constructed dataset, the proposed work will designed a CNN Adaptive Bilateral Enhancer to know and map a equalization function among low contrast image I(x, y) and its respective reference images $Iref_{under\text{-}exposed}$(x, y) and $Iref_{over\text{-}exposed}$(x, y). Further, the work is proceeded to train a CNN network accordingly loss function i.e. Structural dissimilarity (DSSIM).

Bilateral enhancer is a technique to smooth images while preserving edges. Each pixel is replaced by a weighted average of its reference Images.

After feature extraction of the input low contrast image. The intensity or luminosity values are updated. So, CNN Adaptive Bilateral Enhancer is designed in which weight is updated as:

$$I_{enhanced}(x,y) = \sum_{n=1}^{N} W_n(x,y) * I_n(x,y)$$

Where,

N = the number of images in a set of multi-exposure images (under exposed and over-exposed)

I_n(x, y) = pixel intensity of the image

W_n(x, y) = weight extracted out of luminance of the input image I_n(x, y).

$$W_n(x,y) = \exp\left(\left(-I_n(x,y) - (1 - m_n^2)\right)/2\,\sigma_n^2\right)$$

Where, m_n = mean luminance value of I_n(x,y)

σ_n = Standard deviation of luminance value of I_n(x,y)

G. Whole Image Enhancement Network

By using the weight Updation using CNN features the proposed methodology is used for enhancing the luminance values of the pixels of the low contrast image blocks. As the input images is low-contrast that contains both dark as well as bright pixel values. By using the CNN network, the low contrast image is enhanced to high contrast image by shifting the color intensity value. Therefore, the proposed methodology merges the under-exposed and over-exposed components and enhance the image and introduce CNN adaptive filter to refine it to the respective reference image blocks. Finally, contrast is enhanced and all blocks are merged to form image.

IV. DATABASE DESCRIPTION

The Images were collected from different resources such as :

i. VV: This dataset is collected by Vassilios Vonikakis in his daily life to provide the most challenging cases for enhancement. Each image in the dataset has a part that is correctly exposed and a part that is severely under/over-exposed. A good enhancement algorithm should enhance the under/overexposed regions while not affect the correctly exposed one [11].

ii. LIME-data: This dataset contains 10 low-light images used in [12].

iii. NPE3: This dataset contains 85 low-light images downloaded from Internet. NPE-data contains 8 outdoor nature scene images which are used in [13]. NPE-ex1, NPE-ex2 and NPE-ex3 are three supplementary datasets including cloudy daytime, daybreak, nightfall and nighttime scenes.

iv. DICM4: It contains 69 captured images from commercial digital cameras collected by [14].

v. MEF5: This dataset was provided by [15]. It contains 17 high-quality image sequences including natural sceneries, indoor and outdoor views and man-made architectures. Each image sequence has several multi-exposure images, we select one of poor-exposed images as input to perform evaluation.

V. RESULT ANALYSIS

A. Performance Parameters

In this research work two performance parameters are used for image quality assessment. These parameters are:

1) Peak Signal to Noise Ratio (PSNR)

PSNR represents the degradation of the enhanced image with reference images i.e. under exposed and over exposed. It is expressed as a decibel scale. Higher the value of PSNR higher the quality of image. PSNR is represented as:

$$PSNR = 10log10(\frac{(X * Y)}{MSE}$$

Where,

X and Y are height and width respectively of the image.

MSE= Mean Square Error between enhanced image and reference images

2) Structural Similarity Index (SSIM)

The structural similarity (SSIM) index is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. SSIM is used for measuring the similarity between two images. The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality is based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve on traditional methods such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE). The SSIM index is calculated on various windows of an image. The measure between two windows x and y of common size N×N is:

$$SSIM\ (x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Where, μ_x= mean of x

μ_y = mean of y

σ_x^2 = variance of x

σ_y^2= variance of y

σ_{xy}= co-variance of x and y

c_1 and c_2 are variables to stabilize the division with weak denominator.

3) Feature Similarity Index (FSIM)

Feature-similarity (FSIM) index is based on the fact that human visual system (HVS) understands an image mainly according to its low-level features. Specifically, the phase congruency (PC), which is a dimensionless measure of the significance of a local structure, is used as the primary feature in FSIM. The image gradient magnitude (GM) is employed as the secondary feature in FSIM. PC and GM play complementary roles in characterizing the image local quality [16].

The feature similarity is calculated as the measurement between $f_1(x)$ and $f_2(x)$ into two components, each for PC or GM. First, the similarity measure for $PC_1(x)$ and $PC_2(x)$ is defined as:

$$S_{pc}(x) = \frac{2PC_1(x) * PC_2(x) + T}{PC_1^2(x) + PC_2^2(x) + T}$$

Where, S_{pc} = similarity measure for phase congruency

PC_1= phase congruency of low contrast image

PC_2= phase congruency of reference image

T = A +ve constant to increase the stability of S_{PC}

Similarly, the similarity measure for $GM_1(x)$ and $GM_2(x)$ is defined as:

$$S_{GM}(x) = \frac{2GM_1(x) * GM_2(x) + T}{GM_1^2(x) + GM_2^2(x) + T}$$

Where, S_{GM} = similarity measure for gradient magnitude

GM_1= gradient magnitude of low contrast image

GM_2= gradient magnitude of reference image

T = A positive constant to increase the stability of S_{GM}

B. *Result Analysis*

The experimental result is performed and tested on different exposure images. All these images, are collected from different resources such as some images are of indoor and some are outdoor, with or without a lighting fixture. The under exposed as well as overexposed images are created for references. After result analysis, the proposed method is compared to the existing methods on the basis of image quality measure. Table II shows some examples of low contrast images and enhanced contrast images that verifies the effectiveness of the proposed algorithm.

Table II: Result of Low Contrast and Enhanced Images

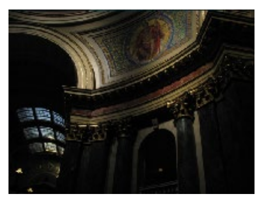
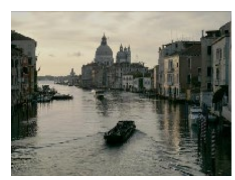
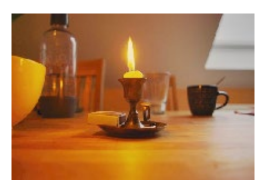
Low Contrast Image	**Enhanced Image**
	
	
	
	
	

Table III, Figure 4 and 5 represents the comparative performance of proposed work with existing work.

Table III: Performance Evaluation of Proposed Methodology

Input Image	FSIM	SSIM	PSNR	MSE	Computation Time (in Sec)
1	0.997	0.93	23.0597	0.0016	6.8
2	0.99	0.920	19.520	0.0037	8.67
3	0.994	0.90	19.3701	0.0039	6.18
4	0.996	0.915	21.45	0.0024	5.86
5	0.995	0.9534	19.302	0.0039	6.27
6	0.993	0.92	20.87	0.0027	6.31
7	0.993	0.91	21.42	0.0024	6.34
8	0.996	0.959	20.84	0.0027	6.18
9	0.995	0.940	23.05	0.0017	6
10	0.997	0.94	22.85	0.0017	7.08
Average	0.995	0.928	21.17	0.0026	6.569

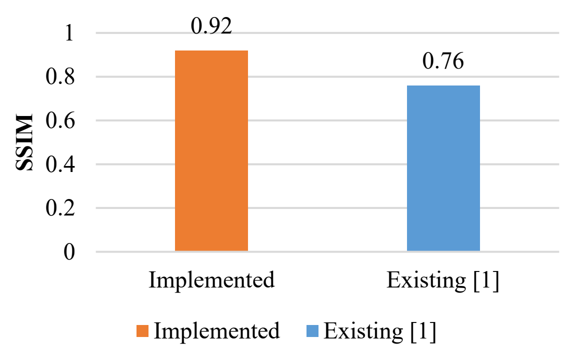


Fig. 4. SSIM Comparative Performance Evaluation

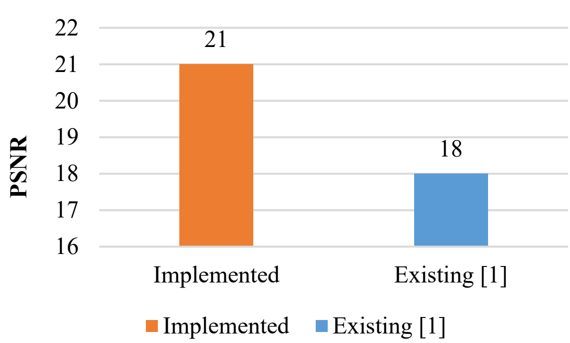


Fig. 5. PSNR Comparative Performance Evaluation

VI. CONCLUSION

Image enhancement techniques have become an important preprocessing tool for digital image processing applications. The proposed CNN-based SICE enhancer, which is capable of adaptively generating high quality enhancement result for a single over-exposed or underexposed input image and may significantly outperforms better to existing work. After result analysis, the proposed method is compared to the existing methods on the basis of image quality measure such as FSIM, SSIM and PSNR values. It is observed that average FSIM obtained is approx. 0.99 whereas average SSIM obtained is approx. 0.92. Similarly, average PSNR obtained is approx. 21. The result analysis shows that the proposed methodology significantly outperforms better as compared to existing work about 3%. In future work, this work would be extended on other application areas such as remote sensing applications, geo-satellite images, MRI images as well biometrics, etc. Apart from contrast enhancement, the proposed methodology will be extended on deblurring and denoising of image.

REFERENCES

[1] Steffi Agino Priyanka, Yuan-kai Wang, Shih-yu Huang, "Low-Light Image Enhancement by Principal Component Analysis", IEEE Access, Vol. 7, 2019, pp. 3082-3092.

[2] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," IEEE Trans. Image Process., vol. 26, no. 2, pp. 982–993, Feb. 2017.

[3] K. Ma, H. Li, H. Yong, Z.Wang, D. Meng, and L. Zhang, "Robust multiexposure image fusion: A structural patch decomposition approach," IEEE Trans. Image Process., vol. 26, no. 5, pp. 2519–2532, May 2017.

[4] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 2782–2790.

[5] Gomez-Ojeda, R., Zhang, Z., Gonzalez-Jimenez, J., & Scaramuzza, D. "Learning-based image enhancement for visual odometry in challenging HDR environments". *In 2018 IEEE International Conference on Robotics and Automation (ICRA)* pp. 805-811, 2018.

[6] Xiao, B., Tang, H., Jiang, Y., Li, W., & Wang, G. "Brightness and contrast controllable image enhancement based on histogram specification". *Neuro computing*, vol. 275, pp. 2798-2809, 2018.

[7] Wan, M., Gu, G., Qian, W., Ren, K., Chen, Q., & Maldague, X. "Infrared image enhancement using adaptive histogram partition and brightness correction". *Remote Sensing*, vol. 10, issue 5, pp. 682-691, 2018.

[8] Ma, J., Fan, X., Yang, S. X., Zhang, X., & Zhu, X. "Contrast limited adaptive histogram equalization-based fusion in YIQ and HSI color spaces for underwater image enhancement". *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, issue 7, 2018.

[9] Ridho Dwisyah Putra, Tito Waluyo Purboyo "A Review of Image Enhancement Methods" *International Journal of Applied Engineering Research* Vol. 12, issue 23, pp. 13596-13603, 2017.

[10] Premkumar, S., Parthasarathi, K.A.: An efficient approach for colour image enhancement using Discrete Shearlet Transform, published in 2nd International Conference on Current Trends in Engineering and Technology (ICCTET), IEEE, pp. 363–366, 2014.

[11] https://sites.google.com/site/vonikakis/datasets
[12] http://cs.tju.edu.cn/orgs/vision/~xguo/LIME.htm
[13] http://blog.sina.com.cn/s/blog a0a06f190101cvon.html
[14] http://mcl.korea.ac.kr/projects/LDR/LDR TEST IMAGES DICM.zip
[15] https://ece.uwaterloo.ca/~k29ma
[16] Jianrui Cai, Shuhang Gu, and Lei Zhang, "Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images", IEEE Transactions on Image Processing, Vol. 27, No. 4, April 2018.