

## Low Contrast Image Enhancement Using Convolutional Neural Network with Simple Reflection Model

Young Shik Moon\*, Bok Gyu Han, Hyeon Seok Yang, Ho Gyeong Lee

*Hanyang University, Computer Science & Engineering, 15588, Republic of Korea*

### ARTICLE INFO

*Article history:*

*Received: 20 December, 2018*

*Accepted: 23 January, 2019*

*Online : 5 February, 2019*

*Keywords:*

*Image Enhancement*

*Convolutional Neural Network*

*Reflection Model*

*Machine Learning*

### ABSTRACT

*Low contrast images degrade the performance of image processing system. To solve the issue, plenty of image enhancement methods have been proposed. But the methods work properly on the fixed environment or specific images. The methods dependent on fixed image conditions cannot perform image enhancement properly and perspective of smart device users, algorithms including iterative calculations are inconvenient for users. To avoid these issues, we propose a locally adaptive contrast enhancement method using CNN and simple reflection model. The experimental results show that the proposed method reduces over-enhancement, while recovering the details of the low contrast regions.*

### 1. Introduction

Image enhancement is an important and typical topic of computer vision. Although the performance of digital systems has improved greatly, the low quality of some images due to external factors such as environment and backlight, which may degrade the performance of image processing systems such as intelligent traffic systems, visual surveillance, and consumer electronics [1]. Especially, low contrast images reduce visibility. Therefore, many smart devices help users take pictures and improve their results with internal image processing methods, but require some input, such as a low contrast area input by the device user, or an average brightness and hue. In specific condition, it is a good solution for the device user who does not consider about the visibility of the image such as counter light images on purpose. However, only chooses the color tone and increases overall brightness is not fit for image enhancement in view of the simple device users.

While maintaining the image quality, to improve the low contrast region of input image is not easy. To solve the problem in this conditional situation, a plenty of methods have been proposed based on mathematical knowledge and traditional image processing method but the methods only perform well in fixed circumstance or make side effect such as over-enhancement or halo effect. To reduce the undesired effect, some enhancement methods use optimization techniques with image decomposition mechanism. However, the computation time is increased and the methods may produce unintended adverse effects depending on the

objective function with decomposed factors. Because of these disadvantage, the methods are not proper to simple device users.

In the past, artificial neural network methods have been one of the most difficult methods to obtain successful results due to small amounts of computation resource and data. However, artificial neural networks are widely used in many fields due to the development of devices and explosively increased amount of data. In addition, as many learning methods are studied, many effective approaches based on input data conditions have been proposed. Recently, the deep learning based methods show the best performance in computer vision sections. The methods are widely being employed on many computer vision tasks that are not easily solved with traditional methods such as image classification, image segmentation, etc. The deep learning based methods with well-labeled training data automatically abstract features of the image differently from the traditional methods and reconstruct the final result with the extracted features. However, the performance of the methods is built on well-labelled data. To overcome the conditional limits, several training techniques are proposed such as semi-supervised learning and weakly-supervised learning. From the perspective of performance, using these skills are able to lead to better results and it can be used widely in general purpose. The performance of the deep learning models trained using non-labelled or rough labelled data shows the power of representation ability.

Focus on the advantage of deep learning, we propose a locally adaptive contrast enhancement method. The method follows previous image enhancement mechanism basically, detecting low contrast region and apply enhancement algorithm. Each person

\*Young Shik Moon, 55, Hanyangdaehak-ro, Sangnok-gu, Ansan-si, Gyeonggi-do, Republic of Korea, ysmoon@hanyang.ac.kr

evaluations for the low contrast region are different. Therefore, we use deep learning mechanism to employ the advantage which extracting feature factor from dataset automatically. Even though our method is based on training with confusing data, the power of representation ability of deep learning shows well results on contrast enhancement purpose. Also, we use a simple reflection model to reduce computational burden.

The remainder of this paper is structured as follows: Section 2 provides background information on image enhancement and briefly introduces previous methods. Section 3 gives a detail of proposed method and shows the detail of our network. While Section 4 provides the experiment results of image contrast enhancement with the proposed method and the previous methods. Finally, Section 5 concludes this paper.

**2. Related works**

For image enhancement, a lot of image processing methods were proposed previously. Histogram based methods are the most popular due to its simplicity and effectiveness. Basically, histogram equalization (HE) is a well-known traditional contrast enhancement method. The main idea of HE contains automatic calculation of the uniform histogram distribution in dynamic range. But HE method does not preserve to mean intensity of input images. The mean intensity of an input image and result of HE are different and it cause undesired effect such an over-enhancement. To solve the issue, brightness preserving bi-histogram equalization (BBHE) [2] was proposed. It divides histogram of an input image into sub histograms and enhances contrast separately. Adaptive histogram equalization (AHE) is used for enhancing contrast in images. It differs from HE by adaptive method that computes several histograms and each histogram corresponding to a distinct section of an image. AHE divide an input image into small sub images and apply HE to each sub image. But because of the size of the sub image, it is too sensitive to noise. Contrast-limited adaptive histogram equalization (CLAHE) [3] is improved version of AHE. The processes of CLAHE is basically similar to AHE. CLAHE controls the sensitive noise with contrast limited distribution parameters. The neighboring tiles, sub image, are combined by bilinear interpolation. Especially in homogeneous areas, the contrast can be limited to avoid noises by distribution parameter.

Retinex-based methods assume that recognizes the relative brightness of the scene, rather than recognizing the brightness of the scene at a certain position when the human visual system recognizes the scene. Retinex based methods have been proposed widely and single-scale retinex (SSR) [4], multi-scale retinex (MSR) [5] and multi-scale retinex with color restoration (MSRCR) [6] are most well-known algorithms. SSR is a method of using the difference between the center pixel value and the convolution result around the center pixel on the log scale. Several convolution masks were proposed [7, 8] but to select the variance of convolution mask effects largely and inadequate variance value can be lead to deterioration of image quality. The MSR is proposed as a method to mitigate this problem of SSR and the retinex result is calculated as the weighted average value of SSR with multiple variance values. The MSRCR proposed that color restoration function (CRF) added to MSR result to solve a problem that come out the gray scale result if specific color is dominant in the input image.

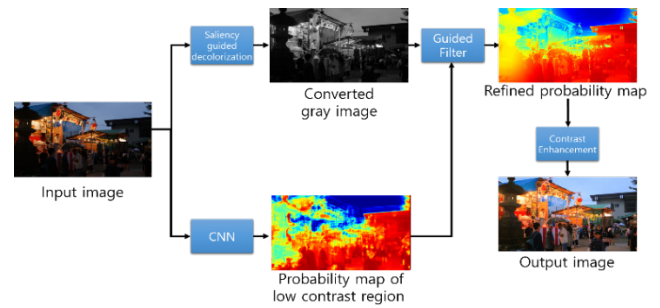


Figure 1. Overall flow of the proposed method.

In recent years, some methods are proposed to enhance low contrast images with adjustment to decomposed illumination image which obtained from various feature models. Different feature models lead to different illumination images and affect the enhanced images. R. Chouhan [9] proposed image enhancement method based on image decomposition technique which follow the discrete wavelet transform (DWT). X. Fu [10] proposed a probabilistic method based on image decomposition. The method decomposed input image into illumination and reflection model then applied the maximum a posteriori (MAP) for estimating enhanced illumination and reflection effectively. Y. ZhenQuiang [11] proposed low-light image enhancement using the camera response model. The camera response model consist of two sub models, camera response function (CRF) and brightness transform function (BTF). Based on these models and estimated exposure ratio map, they enhance image each pixel of the low light image.

The convolutional neural network (CNN) [12] proposed by Y. Lecun has been used widely in computer vision. CNN is being employed on many hard tasks and it showed the best results in many competition in the area. Basically, neural network was proposed for estimating test samples based on training data with hidden layers but calculation with just hidden layer is not suitable for some tasks which related to 2D and 3D data such as images and videos. To preserve locality and flexibility, the Y. Lecun proposed

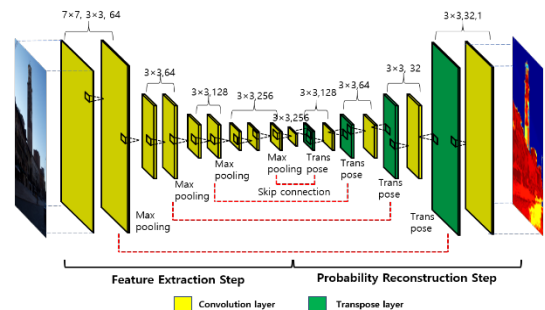


Figure 2. Structure of the convolutional neural network to detect low contrast regions.



Figure 3. Example of training data.

(a) Original image. (b) Labelled image.

basic mechanism of CNN with simple convolutional operation.

One of the most improved tasks based on deep learning in computer vision task is semantic segmentation. Traditional segmentation algorithm was based on mathematical knowledge such as specific feature extraction or the relationship of each pixel. Because of these factors it required in deep research each object which to be segmented. The features that can segment each of object in image are used in restrictive environment. But the methods based on deep learning, It works well without the need to handcraft features of the image based on network characteristics and training data. The most well-known model in semantic segmentation based on deep learning is fully convolutional networks (FCN) proposed by E. Shelhamer [13]. This method shows a basic segmentation network architecture with convolution, pooling and deconvolution. In an input image, important feature information is extracted and compressed by the trained mask through convolution layer and the pooling layer. The compressed information is upsampled in final step with deconvolution layer for matching the size of the input image and the result image. This network not only contains a basic idea for segmentation but also skip-connection mechanism through experimental results. H. Noh [14] proposed a segmentation method with deconvolution technique. The basic model is encoder-decoder network and each lost spatial data by pooling is compensated by deconvolution and unpooling layer.

### 3. Proposed method

The proposed method consists of four steps in total. The first step is low contrast estimation step that generate low contrast probability map. The second step is a create contrast gray scale step. The third step is a refining probability map step with obtained probability map and converted gray scale image. The final step is an enhancement step that enhance low contrast image with refined probability map and converted gray scale. Figure 1 shows the flow chart of the proposed method.

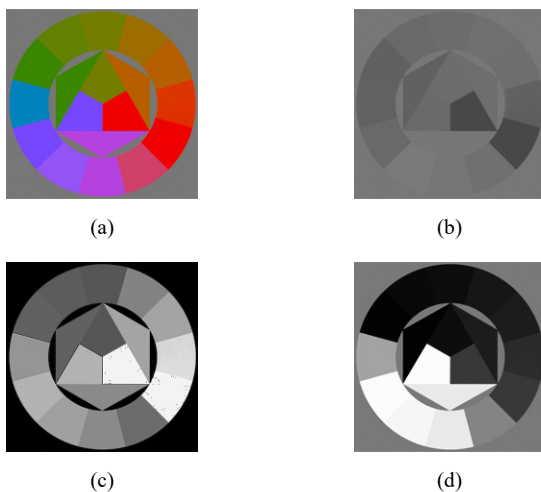


Figure 4. Comparison of gray scale images.

- (a) Input image. (b) Standard gray image.
- (c) Contrast chromatic weight. (d) Contrast gray image.

#### 3.1. Low contrast region probability map

The key idea of the proposed method is to generate probability map of low contrast region. For generation of the probability map quickly and working on small devices such as mobile and tablet PC without computation burden, we construct the network as small

as possible. The Figure 2 shows structure of the proposed method. As shown Figure 2 the network consist of 8 convolution blocks. Inspired by semantic segmentation networks, we design the structure similar to these networks, encoder-decoder networks.

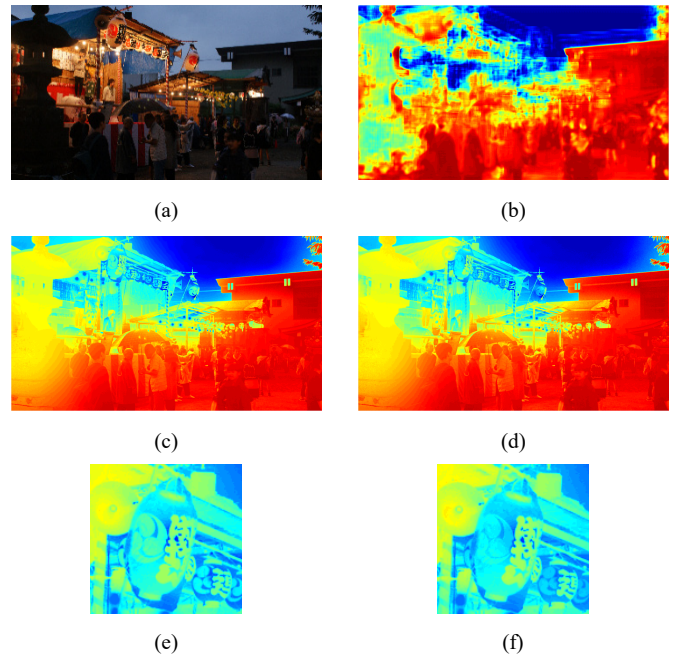


Figure 5. Comparison of refined probability maps.

- (a) Input image. (b) Probability map.
- (c) Refined the (b) with standard gray image.
- (d) Refined the (b) with contrast gray image.
- (e) Part of (c). (f) Part of (d).

From first convolution block to transpose layer, it is called feature extraction step which abstracts low contrast region's feature. All the remaining layers belong to image reconstruction step which rebuild probability map of low contrast regions. To compensate for the lost spatial information by using pooling layer, we use skip connection mechanism and transpose layers. The skip connection was proposed to improve performance of semantic segmentation network that it added the extracted spatial information from the extraction step to the reconstruction step.

We have train the network to generate probability map of low contrast region with low contrast image dataset which collected from on the internet and we used published dataset [15]. We have made ground truth manually and each label consisted of 3 label colors (red, yellow, none). The colors is converted to 1, 0.5 and 0



Figure 6. Result of enhancement step.  
(a) Input image. (b) Final result of enhancement step.

for training step. To get enough training data, we have applied data augmentation including rotation (-10, 0, and +10 degree) and mirroring (vertical, horizontal, and diagonal). As result, the size of training data is 6924. Figure 3 shows example of training data.

### 3.2. Contrast gray image

In the previous works, [1, 16] used Chromatic contrast weights [17] and standard gray scale image for refining the probability map. But only use standard gray scale was not proper because of the image contrast. The standard gray scale image is built with

constant ratio of each channel therefore the image was not be able to contain the contrast of each object's shape. To solve the issue, using the chromatic contrast weight was proposed with handled constant values such as angle of HSV offset and color opponent. The chromatic contrast weight was calculated based on these values, they showed similar values even difference color. To convert an input color image to the gray image appropriately, we use saliency-guided decolorization method [18]. The method based on sparse model and it is able to express contrast with gray value properly. Figure 4 shows gray convert result of standard, chromatic weight and saliency-guided decolorization method.

### 3.3. Refining outline of probability map

In this step, we refine the probability map with guided filter [19] and converted contrast gray image. The guided filter performs as edge-preserving smooth filter which is well known filter in the computer vision area. The guided filter needs the guided image and filtering image, we chose the probability map as a filtering image and the contrast gray image as guided image. It has performed that the boundary region of the probability map refined similar to contrast gray image. Refined the boundary region and shape prevent halo effect and undesired over-enhancement. Figure 5 shows the result of this step. In (e) and (f) of Figure 5, the refined probability map based on standard gray image shows lost detail shape but (f) shows the detail of tomoe pattern.

### 3.4. Enhancement

The final step is enhancement step. In this step we follow the reflection model :

$$I(x) = L(x)J(x)+w \tag{1}$$

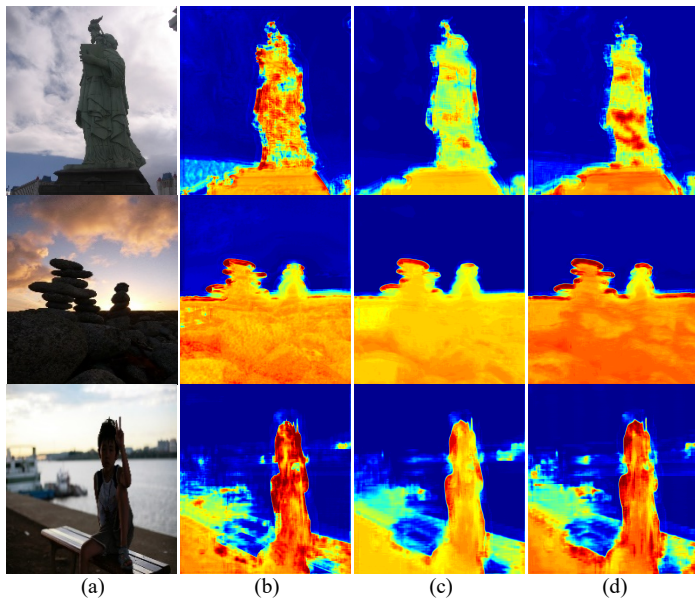


Figure 7. Comparison of the probability maps.  
 (a) Input image. (b) Probability map of proposed method.  
 (c) Without skip connection. (d) No pooling.

where  $I(x)$  is an observed image,  $J(x)$  is a desired (enhanced) image,  $L(x)$  is illuminance map which is a degraded operator, and  $w$  is additive noise. We assume that the image has no noise. So, we can get the enhanced image following (2)

$$J(x) = \frac{I(x)}{F(L(x))+\epsilon} \tag{2}$$

where  $F$  is the guided filter and  $\epsilon$  is small value which is used to avoid dividing by zero. For implementation of (2), we used (2) as

$$\text{Output}_{enhanced}^{ch} = \frac{I^{ch}(x,y)}{(1-\text{Refined map}(x,y))+\epsilon} \tag{3}$$

where  $I^{ch}$  is each color channel of the input image, and the refined map is the result computed in the previous step. Figure 6 shows result of the enhancement step.

## 4. Experiment and results

In the following section we evaluate performance of the proposed method. In Section 4.1, we compare the network performance of the proposed method by comparing two different networks. The performance of the proposed method is then evaluated by comparing the results of the proposed method and the previous methods. The experiments were performed on a PC with a 3.20 GHz Intel Pentium Quad Core Processor and GTX 1080ti graphics card.

### 4.1. Network structure experimentation

For the comparison of effect of the networks, we made two kind of networks. First network was built without skip-connection and the second network was constructed without pooling layers. Each networks were trained same setting as proposed method such as hyper parameters, size of kernel, and number of channels. The Figure 7 shows the generated probability map of each network. The probability maps in Figure 7, the proposed method can get better shape than the others because of reconstruction power in decoder section. Also, the results of no-skip network and no-pooling networks can be seen quite similar. However, from a detailed point of view, the result of no pooling network shows better shape than no-skip network because of the lost the spatial information by pooling layers. In the previous work [13], they proved the power of skip connection and pooling layer but they used precise labeled data for solving each task. Basically, the neural network is used for estimation of some test data based on training data. Even though the training data we used were not precise, we reconfirmed that the network architecture and the inner method could drive quite successful result.

### 4.2. Comparison with other methods

In order to evaluate the performance of the proposed method, we compared the performance of the proposed method and the previous methods, CLAHE [3], MSRCR [6], Fu's method [10], and Ying's method [11]. Figure 8 shows the results of each

methods. The first row of Figure 8, the test image was taken from outside which contains the light source and low contrast region such as the counter light region. In the result of CLAHE, the contrast was all most same as input image. Because of the input image contains light source and contrast region together, the size of sub image is sensitive and the distribution of the sub image is controlled by the distribution parameter. The result by MSRCR

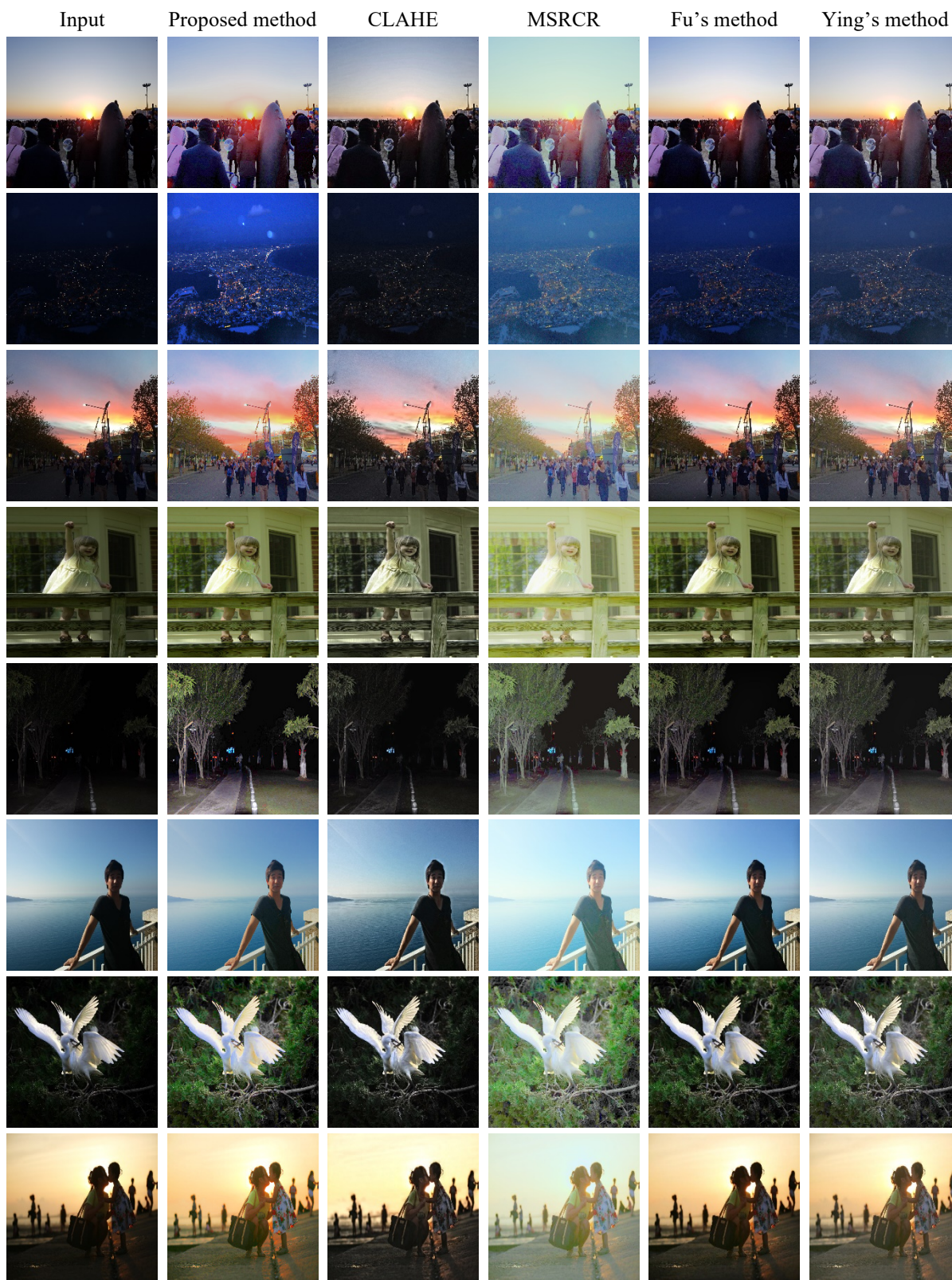


Figure 8. Results of proposed method and previous methods.



(a) (b) (c)  
 Figure 9. Zoom in the result of Fu's method, Ying's method and proposed method.  
 (a) Proposed method. (b) Fu's method.  
 (c) Ying's method.

that not only the contrast is better than CLAHE but also the shape of the statue is seen clearly. But overall brightness is increased too much and the color of the sky looks the same color. The results of Fu's method and Ying's method are improved properly in perspective of contrast and shape. But the overall brightness results are inadequate. In the second row of Figure 8, the results of overall brightness are quite dark. Therefore, the results cannot show the detail of the sky view. For more accurate qualitative comparison, Figure 9 shows zoomed in the results of Fu's method, Ying's method, and proposed method. As you can see images in Figure 9, the shade of the man's face is enhanced properly and we can distinguish the boundary between the nose and the cheek.

## 5. Conclusion

In this paper, we proposed the image enhancement algorithm with simple reflection model and CNN. In order to enhance low contrast image, plenty of methods have been proposed and result of these methods show good quality from perspective of researcher. But previous methods showed undesired effects such as over- enhancement or unnatural results in unfit circumstance such as illumination setting and image sizes. To solve these problems, CNN was employed to detect low contrast regions in order to preserve the shape of low contrast region and prevent over-enhancement. For the purpose of expression the shape of input images, we employed the saliency-guided decolorization method and the guided filter. We showed our network was suitable for image enhancement through experimental network changes and results, and proved that the results of the proposed method were superior to those of the previous methods through comparison of results with the methods. As a result, the proposed method showed successful enhancement without side effects.

## References

[1] B. K. Han, H. S. Yang, and Y. S. Moon, "Locally Adaptive Contrast Enhancement Using Convolutional Neural Network," IEEE International Conference on Consumer Electronics, Las Vegas, USA, 2018. <https://doi.org/10.1109/ICCE.2018.8326096>

[2] Y. T. Kim, "Contrast Enhancement Using Brightness Preserving Bi-Histogram Equalization," IEEE Trans., 1997. <http://doi.org/10.1109/30.580378>

[3] K. Zuiderveld, Graphics Gems IV, Academic Press, 1994.

[4] D. Marini and A. Rizzi, "Computational approach to color adaptation effects," Image and Vision Computing, 18(13), pp. 1005-1014. Oct. 2000. [https://doi.org/10.1016/S0262-8856\(00\)00037-8](https://doi.org/10.1016/S0262-8856(00)00037-8)

[5] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A Multi-Scale Retinex For Bridging the Gap Between Color Images and the Human Observation of Scenes," IEEE Trans. on Image Processing: Special Issue on Color Processing, 6, pp. 965-976, 1997. <http://doi.org/10.1109/83.597272>

[6] Z. Rahman, G.A. Woodell, and D.J. Jobson, "Multiscale Retinex for Color Image Enhancement," Proceedings of 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 1996.

[7] E. Provenzi, M. Fierro, A. Rizzi, L. De Carli, D. Gadia, and D. Marini, "Random spray retinex: A new retinex implementation to investigate the local properties of the model," IEEE Transactions on Image Processing, 16(1), pp. 162-171, Jan. 2007. <https://doi.org/10.1109/TIP.2006.884946>

[8] D.J. Jobson, Z. Rahman, and G.A. Woodell, "Properties and performance of a center/surround retinex," IEEE Transactions on Image Processing, 6(3), Mar., 1997. <https://doi.org/10.1109/83.557356>

[9] R. Chouhan, C. Pradeep Kumar, R. Kumar, and R. K. Jha, "Contrast Enhancement of Dark Images using Stochastic Resonance in Wavelet Domain," International Journal of Machine Learning and Computing, 2(5), 2012. <http://doi.org/10.7763/IJMLC.2012.V2.220>

[10] X. Fu, Y. Liao, D. Zeng, Y. Huang, X. Zhang, and X. Ding, "A Probabilistic Method for Image Enhancement With Simultaneous Illumination and Reflectance Estimation," in IEEE Transactions on Image Processing, 24(12), pp. 4965-4977, Dec., 2015. <https://doi.org/10.1109/TIP.2015.2474701>

[11] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang, "A New Low-Light Image Enhancement Algorithm Using Camera Response Model," 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, pp. 3015-3022, 2017. <http://doi.org/10.1109/ICCVW.2017.356>

[12] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, 86(11), pp. 2278-2324, Nov. 1998. <http://doi.org/10.1109/5.726791>

[13] E. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(4), pp. 640-651, 1 April 2017. <http://10.1109/TPAMI.2016.2572683>

[14] H. Noh, S. Hong and, B. Han, "Learning Deconvolution Network for Semantic Segmentation," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, pp. 1520-1528, 2015. <http://doi.org/10.1109/ICCV.2015.178>

[15] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the Exclusively Dark dataset," Computer Vision and Image Understanding, 178, pp 30-42, 2019. <https://doi.org/10.1016/j.cviu.2018.10.010>

[16] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and John Paisley, "A fusion-based enhancing method for weakly illuminated images," Signal Processing, 129, 2016. <https://doi.org/10.1016/j.sigpro.2016.05.031>

[17] C. O. Ancuti, C. Ancuti, and P. Bekaert, "Enhancing by saliency-guided decolorization," CVPR 2011, Colorado Springs, CO, USA, pp. 257-264, 2011. <http://doi.org/10.1109/CVPR.2011.5995414>

[18] C. Liu and T. Liu, "A sparse linear model for saliency-guided decolorization," 2013 IEEE International Conference on Image Processing, Melbourne, VIC, pp. 1105-1109, 2013. <http://doi.org/10.1109/ICIP.2013.6738228>

[19] K. He, J. Sun, and X. Tang, "Guided Image Filtering," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(6), pp. 1397-1409, June 2013. <http://doi.org/10.1109/TPAMI.2012.213>