Attention-model Guided Image Enhancement for Robotic Vision Applications

Ming Yi, Wanxiang Li, Armagan Elibol, and Nak-Young Chong

Abstract— **Optical data is one of the crucial information resources for robotic platforms to sense and interact with the environment being employed. Obtained image quality is the main factor of having a successful application of sophisticated methods (e.g., object detection and recognition). In this paper, a method is proposed to improve the image quality by enhancing the lighting and denoising. The proposed method is based on a generative adversarial network (GAN) structure. It makes use of the attention model both to guide the enhancement process and to apply denoising simultaneously thanks to the step of adding noise on the input of discriminator networks. Detailed experimental and comparative results using real datasets were presented in order to underline the performance of the proposed method.**

I. INTRODUCTION

Lately, robots have been taking place in daily life more and more. Cameras are one of the most widely used sensors in almost all robotic platforms. Optical data obtained are crucial for further complex tasks such as mapping, localization, object detection and recognition, and several others. While capturing an image, lighting is one of the main factors that play an essential role in obtained image quality. The image quality would be adversely affected by noise and low contrast when it is acquired in low-light environments. Such issues would degrade the image quality drastically, thus both the performance and the outcome of the aforementioned complex tasks. Hence, it is needed to develop a method to recover details for images acquired under low-light environments. Although some methods have been proposed, low-light image enhancement is still a challenging task as it needs to manipulate color, contrast, brightness, and noise simultaneously by only using a given low-quality input image.

Recently, methods using deep-learning algorithms have been proposed for low-light image enhancement. Wei et al. [1] proposed to obtain low-/normal-light image pairs by using different camera settings to train the deep neural network named RetinexNet. Retinex-based methods are aimed to recover the contrast through the estimated illumination map. However, two essential problems remain; Firstly, these methods have not considered non-uniform lighting like fusion-based methods. Secondly, these methods have been focused on restoring brightness and contrast; however, the influences of noise were neglected. The noise in the original input dark images is non-negligible, and the noise in an output contrast-enhanced image is inevitable. Some methods were proposed to overcome the low-light image noise problem and obtain a sharper output image. They include a denoising process directly as a separate component in their enhancement pipeline. In these methods, there are usually two main ideas to deal with the problem; the first one is applying denoising before contrast enhancement while the second one is applying the enhancement before denoising. However, it is a dilemma to make a simple cascade of the denoising and enhancement procedures. For example, denoising before enhancement is likely to cause blurring problems, and enhancement before denoising would cause a noise amplification problem. In order to overcome the aforementioned problems, in this paper, we propose an attention-guided Generative Adversarial Network (GAN) model to solve the noise problem and non-uniform lighting problem simultaneously. An attention-map network based on the U-Net architecture is proposed to guide the generator of the proposed GAN model to recover details hindered by non-uniform lighting in under-exposed and/or over-exposed areas of the input images. Moreover, we derive global and local discriminators from guiding the denoising according to adding random noise in the training process. Our proposal is built upon the recent method EnlightenGAN [2]. Differently, we propose to use a more sophisticated and accurate model for generating attention maps, and we also propose to add noise to the discriminators during training in order to improve the overall performance and perform denoising simultaneously.

All authors are with the School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Nomi, 923-1292, Ishikawa, Japan {ming.yi,wanxiang.li,aelibol,nakyoung} at jaist.ac.jp

II. RELATED WORKS

We present a brief overview of related works in two main categories; Low-light Enhancement and Image Denoising.

A. Low-light Enhancement

Several algorithms based on deep learning have been proposed and found themselves with overwhelming advantages. Fu et al. [3] proposed a weighted minimization algorithm in order to estimate both reflectance and illumination from an input image. Guo et al. [4] developed a structure-aware smoothing model to improve the illumination consistency of images. Lore et al. [5] proposed a deep AutoEncoder approach to extract and learn features from low-light images and then enhance those images. Li et al. [6] proposed the LightenNet, which aims at generating a mapping function between the weakly illuminated image and the corresponding illumination map to obtain the enhanced image. Meanwhile, Wei et al. [1] proposed a deep Retinex decomposition method, which can learn to decompose the observed image into reflectance and illumination in a data-driven way without

decomposing further the image of ground truth. The authors aimed to improve the algorithm to make the process of image-brightening more effective [7].

B. Image Denoising

Image denoising is one of the core research areas in image processing. Many solutions have been presented using methods such as total-variation [8], wavelet-domain processing [9], sparse coding [10], nuclear norm minimization [11], and 3D transform-domain filtering (BM3D) [12]. Especially, BM3D is a state-of-the-art classic technique for denoising. It uses Wiener filter over patches and makes the denoised patches connected back to the first images via a voting instrument, which removes the noise out from the considered area. Some denoising applications based on deep learning have also been developed [13], [14], [5]. An evaluation study with real data showed that BM3D outperformed recent techniques on real images[15]. Nevertheless, most of the existing methods, such as joint denoising and demosaicing [16], [17] have been evaluated on artificial datasets, which are image-sets with added Gaussian noise or Poisson noise rather than real images collected in extreme low-light conditions.

III. Proposed Method

We proposed a method based on the one in [2] to enhance the noisy image obtained under low- and/or non-uniform lighting conditions. Main goal is to improve the capability of robotic platforms on more complicated tasks (e.g., object detection). The proposed method uses GAN architecture, and it adopts a U-Net model as an attention-map generator to guide the generator. The global-local discriminator structure is used to direct the generator to focus on both global and local information. By adding random noise to the input of discriminators, it was also possible to train the generator to enhance not only lighting but also denoising. The overview of the proposed GAN architecture is denoted in Fig.1.

A. Attention Model

The attention model can direct the generator to focus on the important information of an image. We used the naive attention model based on the gray image of the input image using the illumination map algorithm [18] as in [2]. However, the performance of the naive attention model is inefficient in some situations (i.e., the case in Fig. 2), especially when the input image is exceptionally dark. In order to improve the performance, we adopt U-Net [19] as it performed well on semantic segmentation, image restoration, and enhancement. U-Net is employed in our method to generate the attention map. To train the U-Net, we used the output of the illumination map algorithm [18]. This provides to enhance the underexposed areas and avoid over-enhancing the normally exposed areas. The output of U-Net is an attention map indicating the regional underexposure level. An example is shown in Fig. 3.

To obtain the correct attention map, we use the L_2 error metric to measure the prediction error as:

$$L_A = ||F_A(I) - G)||^2 \tag{1}$$

where I is the input image, and $F_A(I)$ is the predicted attention map.

$$G = 1 - L \tag{2}$$

where L is the normalized illumination map of the input RGB image. The illumination map is determined by:

$$L = (0.299 \times R + 0.587 \times G + 0.114 \times B) \tag{3}$$

B. Attention-Guided Generator

Under-exposure images may still have some bright regions, and this means that equally processing every region in an image may not enhance the image. Hence, the ideal situation should be enhancing the dark regions more than bright regions. Motivated by this, we propose to adopt an attention-guided generator constructed by U-Net. The generated attention map is used to guide the generator.

C. Global-Local Discriminators

In order to achieve our goal, the generator has to output the normal light images that have a minimal distance to the real distributions under the guide of the discriminator. However, in practice, using one discriminator using the naive method judging by the image-level often fails on spatially-varying light images [2], [20]. We consider using a global-local discriminator structure [2], [20] in order both to improve the illumination of the whole image and to enhance local regions adaptively. We have an image-level global discriminator to guide the overall brightening. We use a local discriminator to enhance local regions by taking randomly cropped local patches from the output of the generator and real normal light images trying to identify whether real (i.e., the patch is from the real normal light images) or fake (i.e., the patch is from images generated by the generator). By using both the global and the local discriminator, the network can ensure that all local patches of enhanced images look like realistic normal-light ones and this allows the system to work appropriately for overexposed or underexposed images. In order to endow denoising capability to the generator, we propose to add random noise to the input of both Global and Local discriminators. This noise addition would allow the generator not only to generate images with better illumination but also to remove the noise and artifacts. One alternative would be adding another discriminator. However, having more discriminators may cause instabilities on the generator and make the training process more difficult. The following loss functions for the global discriminator D and the generator G are used [2]:

$$\begin{aligned} L_D^{Global} = {} & \mathbb{E}_{x_r \sim \mathbb{P}_{real}}[(D_{Ra}(x_r, x_f) - 1)^2] \\ & + \mathbb{E}_{x_f \sim \mathbb{P}_{fake}}[D_{Ra}(x_f, x_r)^2] \end{aligned} \tag{4}$$

$$\begin{aligned} L_G^{Global} = {} & \mathbb{E}_{x_f \sim \mathbb{P}_{fake}}[(D_{Ra}(x_f, x_r) - 1)^2] \\ & + \mathbb{E}_{x_r \sim \mathbb{P}_{real}}[D_{Ra}(x_r, x_f)^2] \end{aligned} \tag{5}$$

where D_{Ra} is the function of the discriminator defined by sigmoid function. For the local discriminator, randomly cropped 5 patches from the output and real images were used

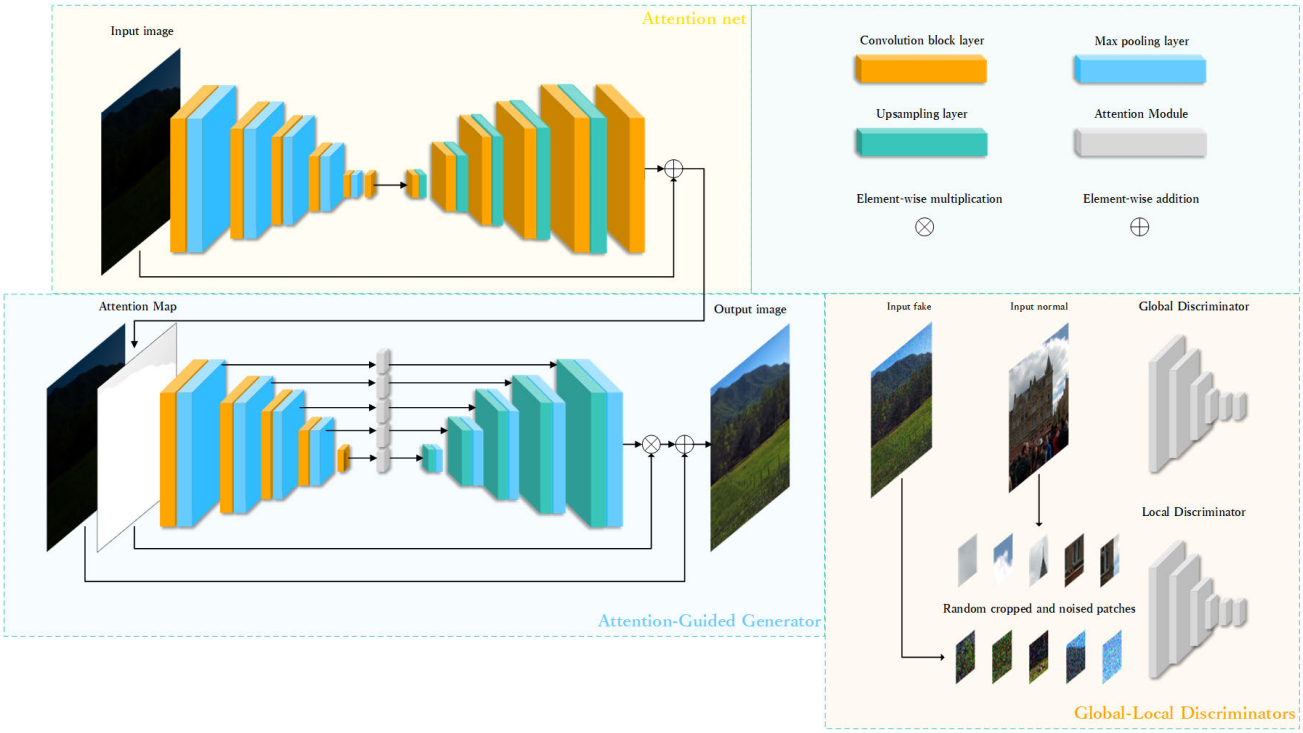
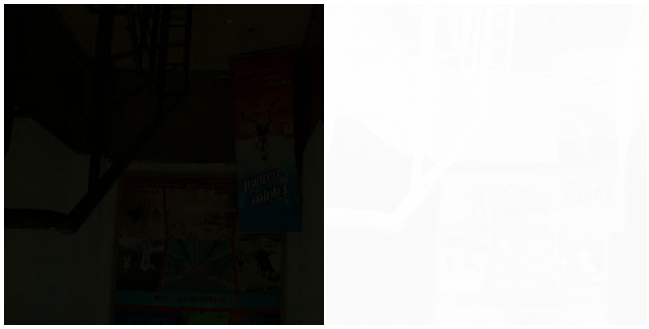
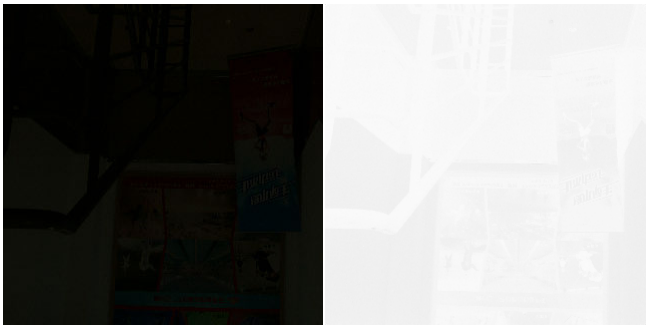


Fig. 1: Structure of the proposed GAN architecture



(a) Input Image (b) Output Illumination Map

Fig. 2: Output of naive attention model



(a) Input Image (b) Output Attention Map

Fig. 3: Output of U-Net based attention map extraction

each time and the original LSGAN [21] was adopted as the adversarial loss, defined as follows:

$$L_D^{Local} = \mathbb{E}_{x_r \sim \mathbb{P}_{real-patches}}[(D(x_r)-1)^2] + \mathbb{E}_{x_f \sim \mathbb{P}_{fake-patches}}[(D(x_f)-0)^2] \quad (6)$$

$$L_G^{Local} = \mathbb{E}_{x_r \sim \mathbb{P}_{fake-patches}}[(D(x_f)-1)^2] \quad (7)$$

D. Self Feature Preserving Loss

The perceptual similarity is used in order to preserve the textures and structures between input images and output images. The perceptual similarity proposed by Johnson et al. [22] calculates the perceptual loss by inputting two images into the pre-trained VGG. The VGG provides features so that the distance between images in the feature space can be obtained. Jiang et al. [2] proposed to use the Preserving Loss on input low-light images and its enhanced normal-light output images to make the training process unpaired and keep the textures and structures feature. The self feature preserving loss L_{SFP} is defined as the mean difference of feature maps provided by VGG network:

$$L_{SFP}(I^L) = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^L) - \phi_{i,j}(G(I^L)))^2, \quad (8)$$

where I^L and $G(I^L)$ are input and output image of the generator while $\phi_{i,j}$ is the feature map of i-th max pooling and j-th convolutional layer and it is provided by VGG network. $W_{i,j}$ and $H_{i,j}$ denote the width and height of the feature map. The overall loss function for training Pre-processing Module is thus written as:

$$Loss = L_G^{Global} + L_G^{Local} + L_{SFP}^{Global} + L_{SFP}^{Local} \quad (9)$$

IV. Experimental Results

We conducted two sets of experiments to evaluate the performance of our proposed method. The first set is to evaluate the enhancement and denoising quality of the output images using the proposed pre-processing module while the second set is to evaluate the object detection accuracy using YOLO-V3 [23] object detection network. For the training process, the proposed module has the capability of using unpaired datasets. In this paper, we opted to use the final dataset used in the EnlightenGAN [2] containing 914 low light and 1016 normal light images from several datasets released from [1], [24] and HDR sources [25], [26]. For test, besides using the test set from the final dataset in [2] (includes 148 paired low/normal light images), we also test our model using the real-world low-light images from the public NPE dataset that includes 8 low light images [27], LIME dataset that includes 10 low light images [4], MEF dataset that includes 17 low light images [28], DICM dataset that includes 69 low light images [29] and VV dataset that includes 24 low light images [30]. Since both the paired and unpaired the datasets were used, both referenced and no-referenced image evaluation methods were used. For paired images, we used referenced well-known image quality metrics, PSNR and SSIM. For unpaired images, we used no-referenced image evaluation methods, namely Perception-based Image Quality Evaluator(PIQE) [31] and Blind/Referenceless Image Spatial Quality Evaluator(BRISQUE) [32]. The proposed pre-processing module was trained for 200 epochs, first 100 epochs using the learning rate of 1×10^{-4}, and the second 100 epochs with the rate linearly decayed to 0. Adam optimizer is used as the optimizer of neural networks, and every training epoch batch size is set to 12 RGB images with a resolution of 320×320. A dropout layer was used with the dropout rate at 0.7. Moreover, batch normalization is used in most layers in order to perform a stable training process.

A. Comparison using Full-Reference Image Quality Assessment

We used SSIM and PNSR for the full-reference image quality assessment on the Final Dataset in [2]. Obtained results are presented in Table I. The numbers provided in Table I are mean values obtained through using the whole dataset. The row "Without Adding Noise" presents the obtained result of the proposed framework trained without adding random noise on the inputs of discriminators. From the result, it can be seen that the proposed method provided better mean values than other tested methods. Fig. 4 shows an example input image and resulting images with tested methods as well as its ground truth image.

B. Comparison using No-Reference Image Quality Assessment Methods

DICM[29], LIME [4], MEF [28], NPE [27] and VV[30] Datasets were used for comparison using no-reference image quality assessment methods, namely PIQE and BRISQUE. Default pre-trained PIQE and BRISQUE models were used in MATLAB© environment for the evaluations. Mean values of obtained results for each dataset and method are presented in Tables II and III. From the result of PIQE and BRISQUE evaluations, the proposed method performs better than the majority of the tested existing methods.

TABLE I: Comparison using Full-Reference Quality Methods

Method \ Metric	PSNR	SSIM
Input Image	10.370	0.300
BIMEF [33]	18.040	0.757
Dong [34]	16.952	0.670
LIME [4]	14.695	0.610
MF [35]	17.677	0.720
MultiscaleRet [36]	12.204	0.511
NPE [27]	17.824	0.667
SRIE [3]	16.650	0.705
RetinexNet [1]	11.100	0.535
EnlightenGAN [2]	17.314	0.757
Without Adding Noise	17.499	0.752
Proposed Method	**18.180**	**0.766**

TABLE II: Comparative Results using PIQE Score

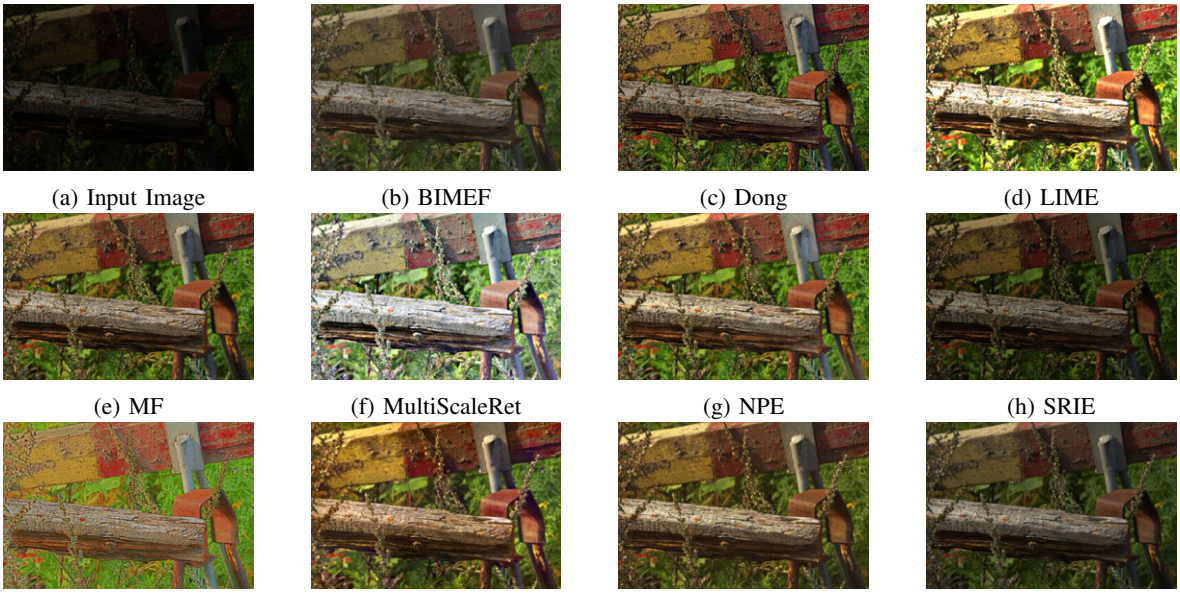
Method \ Dataset	LIME	DICM	MEF	NPE	VV
Original	32.971	35.367	42.827	35.564	25.154
BIMEF	36.228	36.676	34.386	33.351	21.327
Dong	38.449	36.087	37.044	**31.927**	**18.462**
LIME	41.032	41.339	40.024	36.807	21.797
MF	36.381	35.833	34.132	34.335	21.367
MultiScaleRet	40.342	40.199	38.441	35.716	19.909
NPE	36.481	37.179	35.825	31.968	19.504
SRIE	34.799	39.355	37.787	35.382	23.182
RetinexNet	43.118	37.927	41.216	34.499	32.509
EnlightenGAN	34.226	33.438	32.257	33.960	25.490
Proposed Method	**31.802**	**31.231**	**31.037**	33.114	24.335

TABLE III: Comparative Results using BRISQUE Score

Method \ Dataset	LIME	DICM	MEF	NPE	VV
Original	25.142	28.115	29.066	25.137	29.380
BIMEF	23.245	26.811	**19.287**	24.517	22.444
Dong	26.223	26.733	26.495	23.168	29.665
LIME	22.309	26.884	24.058	26.134	26.203
MF	22.323	25.672	22.619	26.292	22.780
MultiScaleRet	22.699	26.010	23.014	24.205	23.270
NPE	22.157	25.316	23.769	24.462	22.818
SRIE	24.181	27.698	22.088	25.635	24.435
RetinexNet	25.648	26.657	26.037	26.873	20.391
EnlightenGAN	**20.134**	26.282	23.641	27.340	19.305
Proposed Method	20.558	**24.395**	21.582	**23.552**	**17.591**

C. Object Detection Experiments

In order to verify that the proposed module helps to improve the object detection performance, the object detection module (same as YOLO V3 [23]) is trained with COCO [37] dataset for 100 epochs with a batch size of 8 RGB images of 416×416 using Adam optimizer and learning rate of 1×10^{-4}. We input low-light images (referred to as original), ground truth images, enhanced low-light images (referred to the as enhanced original), and enhanced ground truth images into the object detection module, and the obtained results are given in Table IV. The last row of the table shows the result of the original low-light images and ground truth



(a) Input Image (b) BIMEF (c) Dong (d) LIME

(e) MF (f) MultiScaleRet (g) NPE (h) SRIE

(i) RetinexNet (j) EnlightenGAN (k) Proposed (l) Ground Truth

Fig. 4: Sample input image, its ground truth and resulting images using different tested methods and the proposed one.

images. In the table, the "Total" column shows the output counts of the object detection network; the column Correct shows the total number of correctly detected objects while the column "Incorrect" represents the total number of the wrong detection. From the table, it can be seen that the proposed

TABLE IV: Object Detection Performance Comparison

Method \ Data	Enhanced Original			Enhanced Ground Truth		
	Total	Correct	Incorrect	Total	Correct	Incorrect
Dong	171	162	9	158	148	10
BIMEF	176	165	11	198	187	11
EnlightenGAN	175	**168**	7	183	170	13
Proposed Method	170	164	**6**	195	**188**	**7**
Original Data	96	91	5	197	188	9

pre-processing module helps to improve the performance of object detection module. We also present the results of object detection on images obtained by applying a pre-processing module to the ground truth images. From the results, it can be seen that the proposed method maintains the image quality similar level without causing any artifacts due to over enhancement. Some examples of object detection results are given in Fig. 5.

V. CONCLUSIONS

In this paper, a novel method based on GAN to improve image quality is proposed. The proposed method makes use of attention models to guide the enhancement process in order to detect and perform enhancement on local-regions where it is most needed. It also applies denoising simultaneously improving the lighting. Different experiments were carried out on several datasets, and comparative results with existing methods were reported. The proposed method showed promising performance.

REFERENCES

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

[2] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep light enhancement without paired supervision," *arXiv preprint arXiv:1906.06972*, 2019.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] Y. Shi, X. Wu, and M. Zhu, "Low-light image enhancement algorithm based on retinex and generative adversarial network," *arXiv preprint arXiv:1906.06027*, 2019.

[8] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[9] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans Image Processing*, vol. 12, no. 11, 2003.

[10] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[11] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 2862–2869.

[12] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[13] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *Advances in neural information processing systems*, 2012, pp. 341–349.

[14] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE transactions*

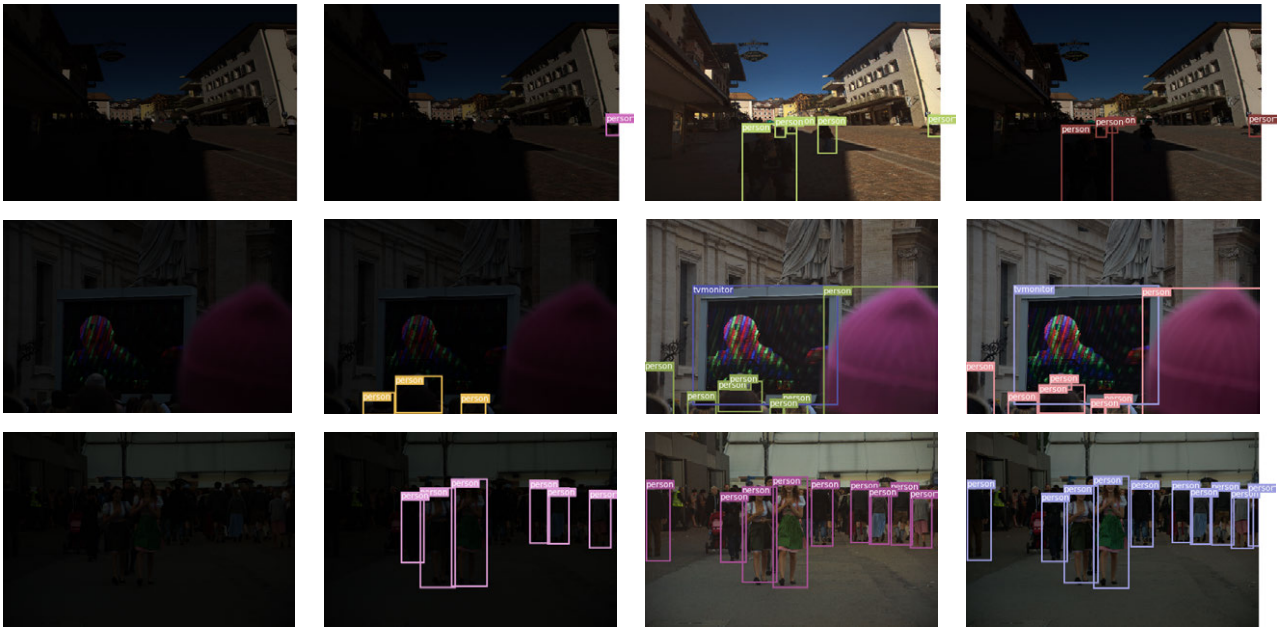


Fig. 5: Samples of Object Detection results using original, enhanced and ground truth images. First Column denotes the input image while the second column shows the result of object detection using original image. The third column is for the object detection results using enhanced images by the proposed method and the final column shows the object detection using ground truth images

on pattern analysis and machine intelligence, vol. 39, no. 6, pp. 1256–1272, 2016.

[15] T. Plotz and S. Roth, "Benchmarking denoising algorithms with real photographs," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1586–1595.

[16] K. Hirakawa and T. W. Parks, "Joint demosaicing and denoising," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2146–2157, 2006.

[17] M. Gharbi, G. Chaurasia, S. Paris, and F. Durand, "Deep joint demosaicking and denoising," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, p. 191, 2016.

[18] J. Song, L. Zhang, P. Shen, X. Peng, and G. Zhu, "Single low-light image enhancement using luminance map," in *Chinese Conference on Pattern Recognition*. Springer, 2016, pp. 101–110.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241.

[20] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "DeblurGAN-v2: Deblurring (orders-of-magnitude) faster and better," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 8878–8887.

[21] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2794–2802.

[22] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European conference on computer vision*. Springer, 2016, pp. 694–711.

[23] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[24] D.-T. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: a raw images dataset for digital image forensics," in *Proceedings of the 6th ACM Multimedia Systems Conference*, 2015, pp. 219–224.

[25] 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.

[26] N. K. Kalantari and R. Ramamoorthi, "Deep high dynamic range imaging of dynamic scenes." *ACM Trans. Graph.*, vol. 36, no. 4, pp. 144–1, 2017.

[27] 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.

[28] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3345–3356, 2015.

[29] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation," in *19th IEEE International Conference on Image Processing*. IEEE, 2012, pp. 965–968.

[30] V. Vonikakis, D. Chrysostomou, R. Kouskouridas, and A. Gasteratos, "A biologically inspired scale-space for illumination invariant feature detection," *Measurement Science and Technology*, vol. 24, no. 7, p. 074024, 2013.

[31] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *2015 Twenty First National Conference on Communications (NCC)*. IEEE, 2015, pp. 1–6.

[32] A. Mittal, A. K. Moorthy, and A. C. Bovik, "Blind/referenceless image spatial quality evaluator," in *2011 conference record of the forty fifth asilomar conference on signals, systems and computers (ASILOMAR)*. IEEE, 2011, pp. 723–727.

[33] Z. Ying, G. Li, and W. Gao, "A bio-inspired multi-exposure fusion framework for low-light image enhancement," *arXiv preprint arXiv:1711.00591*, 2017.

[34] X. Dong and J. Wen, "Low lighting image enhancement using local maximum color value prior," *Frontiers of Computer Science*, vol. 10, no. 1, pp. 147–156, 2016.

[35] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," *Signal Processing*, vol. 129, pp. 82–96, 2016.

[36] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image processing*, vol. 6, no. 7, pp. 965–976, 1997.

[37] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.