

FLIGHT Mode On: A Feather-Light Network for Low-Light Image Enhancement

Mustafa Özcan, Hamza Ergezer, Mustafa Ayazaoglu
Aselsan Research
Ankara, Turkey

{mustafaoczcan, hergezer, mayazoglu}@aselsan.com.tr

Abstract

Low-light image enhancement (LLIE) is an ill-posed inverse problem due to the lack of knowledge of the desired image which is obtained under ideal illumination conditions. Low-light conditions give rise to two main issues: a suppressed image histogram and inconsistent relative color distributions with low signal-to-noise ratio. In order to address these problems, we propose a novel approach named FLIGHT-Net using a sequence of neural architecture blocks. The first block regulates illumination conditions through pixel-wise scene dependent illumination adjustment. The output image is produced in the output of the second block, which includes channel attention and denoising sub-blocks. Our highly efficient neural network architecture delivers state-of-the-art performance with only 25K parameters. The method's code, pretrained models and resulting images will be publicly available.

1. Introduction

A low-light image can be defined as an image captured in deficient illumination conditions that do not fully excite the detector. Hence, the output image is not even close to have ideal histogram distributions. In such conditions, a dedicated algorithm enhancing the image is needed to present a better image and to help increasing the performance of consequent blocks which are trained or tuned under normal lightning conditions, such as object detection, etc.

Global or local image histogram equalization techniques [4, 36] are the first-thought candidates to solve the low-light image enhancement problem. However, they do not employ spatial information and work in pixel level that does not include surrounding content information. On the other hand, using deep neural network architectures, spatial information can be utilized and combined with color information in different scales. Therefore, deep neural architectures recently provide superior performance for low-light image enhance-

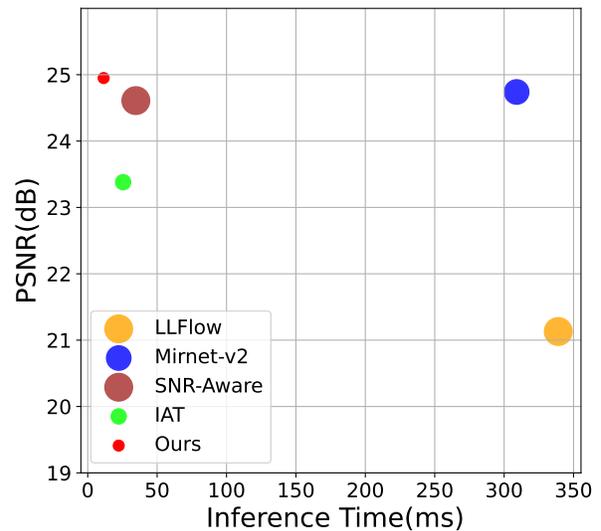


Figure 1. Performance Comparison on LOL-v1 Dataset. The diameters of the circles are proportional to the number of model parameters.

ment problem as in most of other low-level and high-level vision problems.

For most of the applications, low-light image enhancement should be implemented in image signal processing (ISP) framework of a camera. ISP frameworks process the data in real-time, and include some certain blocks. Hence, it can be argued that the most crucial requirements for LLIE should be computational memory and complexity while not sacrificing from the visual quality. In recent years, there are some studies [19, 30] that handles all the blocks of ISP blocks in a single network. However, it is not easy to deploy such a network of multi-million parameters in an edge device, i.e. surveillance camera. Therefore, it is crucial to have lightweight blocks to handle unusual conditions in ISP framework. The proposed solution is a good candidate for such cases to deploy in an edge device.

Inspired by several recent studies [3, 30], we propose

a feather-light network with carefully designed blocks to match the problem’s nature at the hand instead of throwing all of the information into a huge network and hope to get the right output by some “deep magic”. As with the previous studies, we model the LLIE as an image generated as reflectance multiplied by illumination. For this purpose, first, we strive to achieve pixelwise scene illumination. Using such an approach, we achieve an equalized the image histogram while handling uneven illumination at the same time. Since the input signal has very low SNR, there exist inherent noise in the input and this becomes visible with the illumination adjustment, the noise and color inconsistencies should be taken care of by the subsequent blocks. To solve these problems, we utilize the ideas presented in recent approaches [8, 10, 30] that addresses these problems using channel and spatial attention blocks [7, 25]. In simple terms these attention mechanisms enable selective feature extraction for image enhancement and denoising.

We propose a novel efficient neural network architecture named FLIGHT-Net is proposed for low-light image enhancement problem. It is shown that FLIGHT-Net gives outstanding results compared the state of the art. To the best of our knowledge, it is the lightest network that achieves great balance between run-time and performance among supervised learning LLIE methods.

2. Related Work

As in high-level vision tasks, the solutions based on deep neural architectures provide most successful results for image restoration tasks. Before deep learning era, traditional methods also give satisfactory results up to a point. Global and local histogram equalization methods [4, 36] are most well-known solutions for the LLIE problem. Furthermore, Retinex theory increased the understanding of the problem from a more theoretical point of view. Following Retinex theory, in LIME [6], the overall solution is based on illumination map estimation (IME) approach. Although the IME approach is inspiring, being a traditional method LIME [6] failed to generalize well on different scenarios. On the other hand, IME block kept its existence on different deep learning approaches such as [3, 29] and it is also a part of our proposed solution.

Starting from the pioneer work [17], the deep learning based LLIE methods can be categorized according to general deep learning strategies. In other words, deep LLIE methods can be divided into four main categories as supervised learning [3, 22, 23, 26, 34], semi-supervised learning, zero-shot learning [5, 12, 35] and unsupervised learning. Although the number of other types of methods is also notable, the core part of the literature is formed by supervised and zero-shot learning based methods.

In supervised learning based methods, there are two main approaches. The first approach is to extract the en-

hanced image by using a single end-to-end network [14, 22]. The second and mostly utilized approach is to design the subnetworks according to the Retinex theory [11]. In these approaches, subnetworks are designed to reconstruct the illumination and reflectance parts of the enhanced image. In [23], two main blocks called Decom-Net and Enhance-Net are used to extract illumination and reflectance maps and then adjust the illumination according to the decomposed maps. Illumination adjustment is handled by pixelwise enhancement block in [3], while color correction is solved using a transformer block.

It is not easy to build a setup with the ground truth and low-light image pair required by the problem. Therefore, more recently, zero-shot learning methods [5, 12, 15, 35] are proposed for LLIE problem. Zero-DCE and its extended version [5, 12] solve the problem by predicting a set of high-order curves for a given image. In [35], a generative strategy is applied to decompose the illumination and reflectance components. After the decomposition, enhanced image is obtained by processing the illumination component. RUAS [15] utilizes again Retinex theory and neural architecture search strategy to determine the basic blocks called illumination estimation and noise removal module. Low-light enhancement problem is also elaborated as a sub-problem in some recent image restoration studies [20, 31, 32].

3. Proposed Method

LLIE can be posed as a reconstruction of the ideal image which is under the ideal light conditions. In this section, the motivation and formulation of FLIGHT-Net are introduced. Then the details of the network are presented.

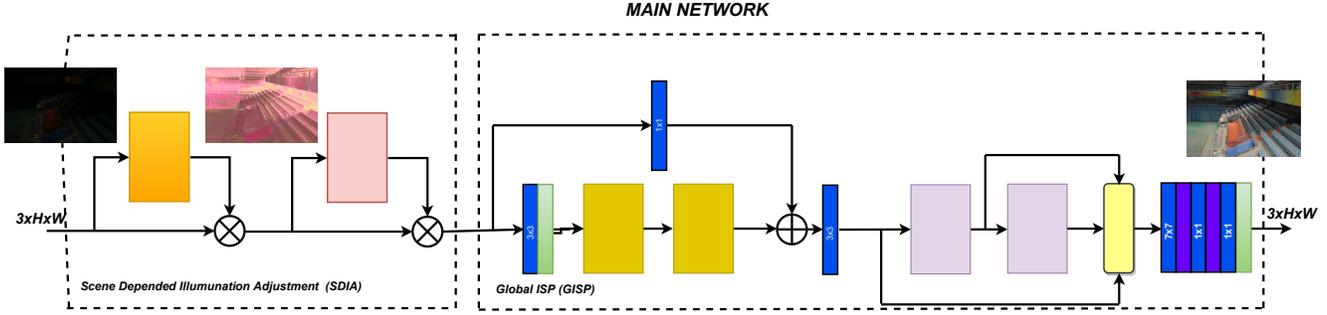
3.1. Problem Formulation

The proposed method and formulation are inspired by Retinex theory [11] and ISP framework [9]. As it is known, in the Retinex theory, an image consists of reflection and illumination via equation 1:

$$I = R.L \tag{1}$$

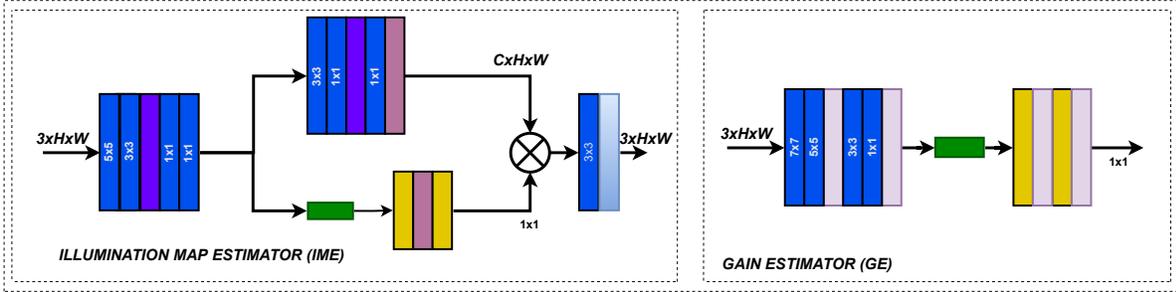
where I, R, L donate image, reflectance and illumination. Retinex-based deep learning methods estimate an illumination map using low-light image [35]. Estimated illumination maps which is required to obtain a normal light image from a low light image, transform the image on a pixel-by-pixel basis. Furthermore, considering the image acquisition from the camera, a series of transformations are applied to the linear raw RGB image, which makes obtaining normal-light image from low-light image more challenging. The transformation can be expressed as in equation 2.

$$I = f_{ISP}(S) \tag{2}$$

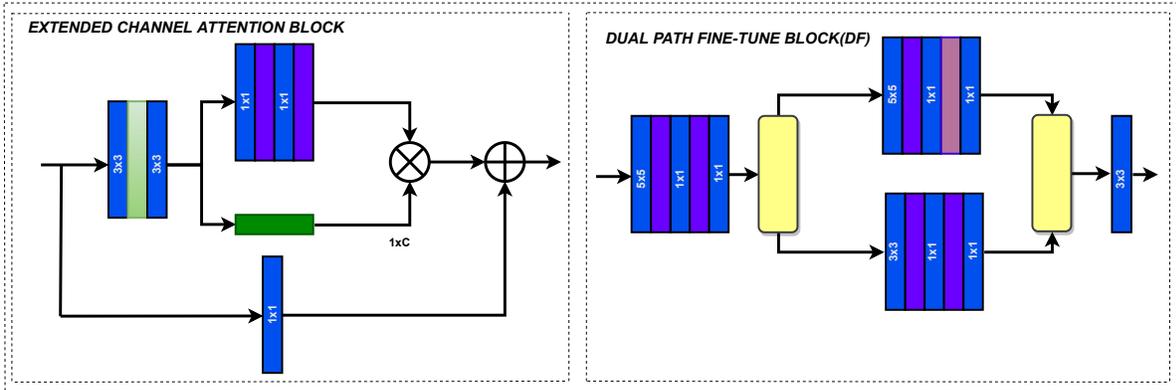


(a) Main Network

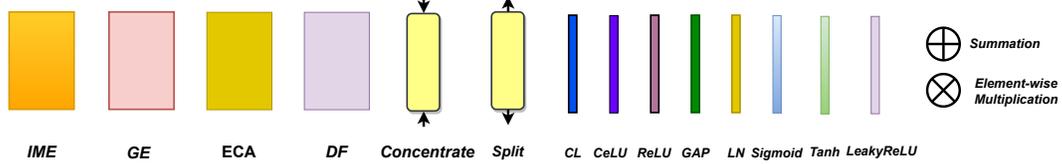
Scene Dependent Illumination Adjustment (SDIA)



(b) SDIA Block



(c) GISP Block



(d) IME: Illumination Map Estimation, GE: Gain Estimator, ECA: Extended Channel Attention, DF: Dual Path Fine-Tune, CL: Convolutional Layer, CeLU: Continuously Differentiable Exponential Linear Unit, ReLU: Rectified Linear Unit, GAP: Global Average Pooling, LN: Linear Layer

Figure 2. FLIGHT-NET

where S denotes camera input and f_{ISP} denotes the whole transformation function obtaining sRGB image from sensor input, i.e. raw data. The quality of a low-light image is affected by two main factors: low illumination and in-camera noise. The sensor input in the sensor layer produces a Raw linear RGB image that is linearly related to the am-

bient light. To obtain an image similar to a standard light image, the raw linear RGB image can be enhanced using an appropriate gain ratio, eliminating the first degradation of the low-light image. However, the second degradation, in-camera noise, increases linearly with the gain ratio and distorts the image.

Another crucial issue when working with low-light sRGB images is that raw image is processed through a series of non-linear operations such as white balance, gamma correction, noise reduction, contrast enhancement, and edge enhancement, hence it is no longer possible to obtain a normal-light sRGB image with the appropriate gain coefficient using the low-light sRGB image. Cui et.al [3] explains this by stating that the actual luminance degradation occurs in the raw-RGB space in the ISP framework and proposed the IAT network design, which is inspired and characterized the ISP process.

Inspired by Retinex theory [11], the ISP framework [9] and IAT network [3], a new deep learning network called FLIGHT-Net for low-light image enhancement is proposed. FLIGHT-Net is comprised of two primary network blocks: the Scene Dependent Illumination Adjustment (SDIA) block, and the Global ISP block (GISP). The SDIA block modifies the input in a pixel-wise manner, while the GISP block transforms the image globally. The formulation of the proposed method is shown in equation 3:

$$NLI = f_{GISP}(f_{SDIA}(LLI)) \quad (3)$$

where LLI , NLI , f_{SDIA} , f_{GISP} represent low-light image, normal-light image, SDIA and GISP blocks respectively.

SDIA block is composed of two separate blocks: illumination map estimator and gain estimator. As stated earlier, a low-light image is an image in which the sensor receives less light than a normal-light image. The gain estimator predicts the required gain ratio necessary to facilitate image enhancement. However, an accurate gain ratio for LLIE is not possible due to the ISP applied when capturing the sRGB image. Therefore, the image is transformed using a suitable map (IM), which is the output of the illumination map estimator (IME) block to increase the effect of the gain value. As a result, the input image is mapped to the latent feature space by utilizing the SDIA.

During the process of obtaining the normal-light image from the low-light image, in the SDIA, noise becomes more apparent. This is due to the inherent noise in the low-light image due to low PSNR and when low-light is illumination adjusted the noise is also becomes visible and deteriorates the visual quality. However, since the noise is visible along with the scene itself another network might be selectively diminish the noise while preserving or even enhancing the scene. Indeed our GISP block suppresses the noise while keeping structures intact through embodied attention mechanisms as shown 2c. We can consider GISP block as a transformation from a latent feature space to sRGB space. In other words, color correction, denoise and white balance operations of a typical ISP framework are mimicked via GISP block.

3.2. Network Framework

FLIGHT-Net consists of two main network block as shown in Figure 2a. The first block that converts input locally is Scene Depended Illumination Adjustment (SDIA) Network Block, and other block Global ISP Network Block (GISP) that transform its input globally. SDIA include IME and GE block whose outputs are multiplied by the input image on a pixel-wise basis. On the other hand, the output of the GISP block is used directly as the network output.

As previously stated, the low-light image is characterized by insufficient light exposure onto the sensor, resulting in a dark and low dynamic range projection. Multiplying low-light image with a single gain coefficient can be considered as a naive way for obtaining a more pleasing image. However, this is usually not enough due to the non-linear effects introduced in the ISP framework such as local histogram enhancement where each pixel are subject to different illumination and hence need to be individually corrected, so, simply adjusting the gain coefficient alone is not sufficient. Since it may cause over-illumination in some parts, while leaving other areas in darkness. To address this issue more effectively, an IME block has been devised to estimate the illumination adjustment coefficients required for each pixel. Through this estimation process, the input image can be better prepared for gain adjustment block, ultimately leading to a more efficient utilization of the gain coefficient.

The IME block, illustrated in Figure 2b, includes CNN blocks for feature extraction subblocks and local gain coefficient. After feature extraction, the illumination map can be estimated effectively by multiplying map features and the local gain coefficient estimated with the help of LN by the features. At the end of the IME block, the sigmoid activation function is preferred because the IME block is designed to transform the image where the gain coefficient works efficiently.

The GE block in Figure 2b is used to estimate the appropriate gain coefficient depending on its input which might be captured varying light conditions. It consists of CNN blocks to extract features and a linear layer to estimate the required gain coefficient using these extracted features. After gain adjustment, the input image is converted to the latent feature, which is needed for GISP to deliver a normal-light sRGB image on its output.

The GISP block is the second main block of FLIGHT-Net. As stated previously, the main purpose of this block is color correction and denoising. It consists of extended channel attention (ECA) block and dual path fine tune (DF) blocks. The ECA block is the extended version of the CA block in [33]. This block mainly strengthens the information like structures and patterns in the necessary channels while surpassing the unwanted information like noise. In the DF block, the extracted features in the channel are di-

| Method | SNR-ALLIE [26] | RetiNexNet [24] | MBLLEN [18] | DRBN [27] | KIND [34] | MAXIM [20] | IAT [3] | Ours |
|--------|----------------|-----------------|-------------|-----------|-----------|--------------|---------|--------------|
| PSNR | 24.61 | 16.77 | 17.90 | 19.55 | 20.86 | 23.43 | 23.38 | 24.96 |
| SSIM | 0.842 | 0.562 | 0.70 | 0.746 | 0.790 | 0.863 | 0.809 | 0.85 |

Table 1. Comparative performance results for LOL-v1 dataset

| Method | SNR-ALLIE [26] | Retinex [15] | IPT [2] | Sparse [29] | Band [28] | MIR-Net [31] | LPNet [13] | Ours |
|--------|----------------|--------------|---------|-------------|-----------|--------------|------------|--------------|
| PSNR | 21.48 | 18.37 | 19.80 | 20.06 | 20.29 | 20.02 | 17.80 | 21.71 |
| SSIM | 0.849 | 0.723 | 0.813 | 0.816 | 0.831 | 0.820 | 0.792 | 0.834 |

Table 2. Comparative performance results for LOL-v2-Real dataset

| Method | SNR-ALLIE [26] | Retinex [15] | IPT [2] | Sparse [29] | Band [28] | MIR-Net [31] | LPNet [13] | Ours |
|--------|----------------|--------------|---------|-------------|-----------|--------------|------------|--------------|
| PSNR | 24.14 | 16.55 | 18.30 | 22.05 | 23.22 | 21.94 | 19.51 | 24.92 |
| SSIM | 0.928 | 0.652 | 0.811 | 0.905 | 0.927 | 0.876 | 0.846 | 0.93 |

Table 3. Comparative performance results for LOL-v2-Synthetic dataset

| Method | Zero-DCE [5] | LIME [6] | Retinex-Net [24] | KinD [34] | MBLLEN [18] | GLADNet [21] | MIR-Net [31] | Ours |
|--------|--------------|----------|------------------|-----------|-------------|--------------|--------------|--------------|
| PSNR | 12.99 | 14.95 | 15.43 | 15.84 | 17.52 | 21.09 | 21.62 | 22.44 |
| SSIM | 0.44 | 0.45 | 0.34 | 0.49 | 0.60 | 0.69 | 0.77 | 0.794 |

Table 4. Comparative performance results for Rellisur dataset

vided into two in order to transform features with different receptive field, different activation functions and kernel sizes in convolution operations. In addition, foremental information are carried forward and information at every stage is protected. The details of the GISP block are shown in Figure 2c.

Last but not least we tried to reflect the current network design language to our design by using some of the suggestions for CNNs for reaching Transformer like performance as suggested in [16]. It is reported that CNN architectures can obtain the success of the transformers with a proper selection of some parameters. For instance, as it is known in classical CNN design, convolution kernel size 3x3 is mostly preferred. In the FLIGHT-Net, 5x5 and 7x7 kernel sizes are also preferred in the convolution operation, especially at the beginning of the blocks used for feature extraction. In conclusion, it seems that the results in [16] is indeed be beneficial for LLIE.

| Dataset | Training Image | Validation Image |
|------------------|----------------|------------------|
| LOL [24] | 485 | 15 |
| LOL-v2-Syn [27] | 900 | 100 |
| LOL-v2-Real [27] | 689 | 100 |
| Rellisur [1] | 3610 LL | 215 LL |
| | 722 NL | 43 NL |

Table 5. Datasets used in experiments *low-light (LL), normal-light(NL)

4. Experiments

LOL-v1 [24], LOL-v2-real [27], LOL-v2-synthetic [27] and Rellisur [1] datasets are chosen for our experiment. The

| Method | PSNR | SSIM | #P(M) | AT(ms) |
|-----------------|-------|-------|-------|--------|
| Zero-DCE++ [12] | 14.83 | 0.531 | 0.01 | 1.02 |
| LLFlow [22] | 21.13 | 0.853 | 38.86 | 339.69 |
| Mirnet-v2 [32] | 24.74 | 0.851 | 5.86 | 309.68 |
| SNR-Aware [26] | 24.61 | 0.842 | 39.12 | 34.67 |
| IAT [3] | 23.38 | 0.809 | 0.09 | 25.50 |
| Ours | 24.96 | 0.85 | 0.025 | 11.47 |

Table 6. Comparative comparison of the methods according to the number of parameters (#P) and average computation time (AT) on LOL-v1 dataset without using GT images.

LOL-v1 dataset is the first real dataset for low-light image enhancement and is used to test many state-of-art network [3, 22, 23, 26, 27, 34]. The LOL-v2 dataset is the enhanced version of the LOL-v1 dataset. It contains two different training and validation image pairs, a real captured image and a synthetically acquired image. Finally, Rellisur dataset is the first multi-purpose dataset for the problems of low-light image enhancement and super-resolution. Details of the training and validation image pairs of the datasets used in the experiment are summarized in the Table 5.

During training, all experiments are performed in Tesla V100 GPU 16GB. ADAMW optimizer is selected in the training process as in [16]. Batch size is set to 16 and initial learning rate is 8×10^{-4} . The network is trained with initial learning rate for 6000 epoch and it reduced linearly 1/10th of it between 6000-12000 epoch. For fine tuning phase, previous scheduling is applied by starting from 4×10^{-4} .

The loss function consists of two different components. Smooth L1 loss is preferred to ensure that the estimated image pixels are close to the normal-light images. The second

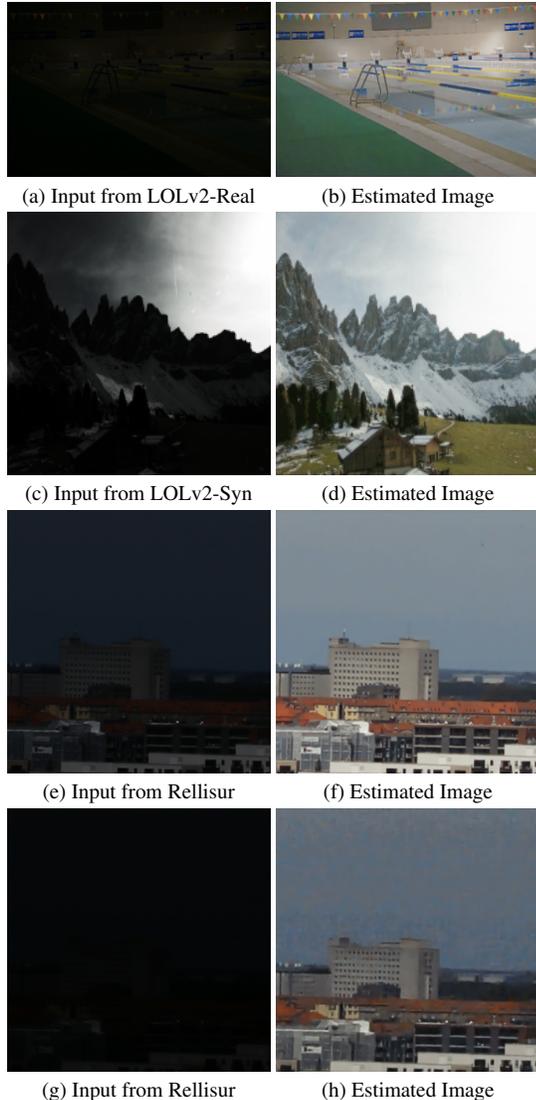


Figure 3. Visual results for LOL-v2-real, LOL-v2-synthetic and Rellisur datasets.

loss is Multi-scale Structure Similarity Measure (MS-SSIM), which enforces the network to predict more visually pleasing image. The total loss function is calculated as:

$$L_{TOTAL} = \alpha_1 * L_{L1} + \alpha_2 * L_{MS-SSIM} \quad (4)$$

α_1 and α_2 are hyperparameters for balancing the loss functions.

4.1. Comparative LLIE Results

State-of-the-art LLIE [3, 22, 23, 26, 27, 34] along with some image restoration networks [20, 32] are selected for comparison. We report PSNR and SSIM results for LOL-v1, LOL-v2-real, LOL-v2-synthetic and Rellisur datasets

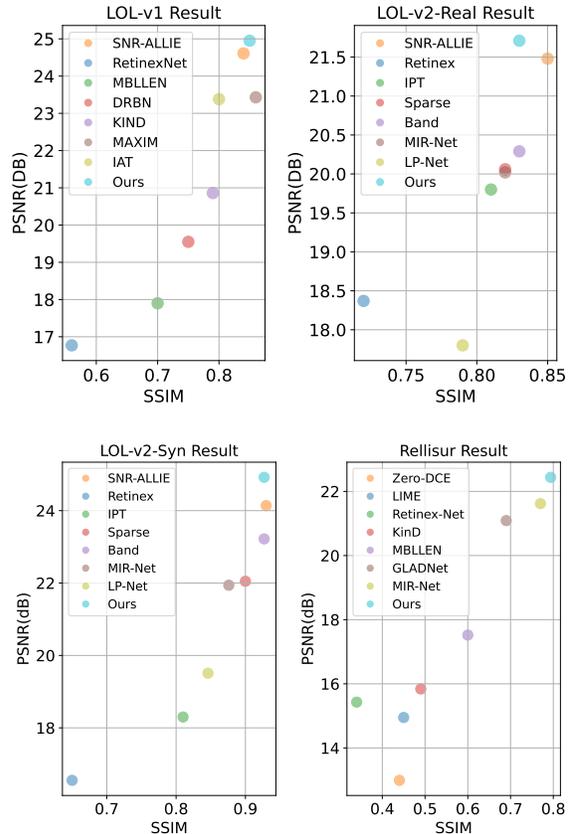


Figure 4. Performance Comparison

in Tables 1, 2, 3 and 4 respectively and their corresponding comparative plots are given in Figure 4.¹ Our method achieve best PSNR results except the LOL-v2-real dataset. For SSIM metric, again, we achieve best performance in three of four datasets with the help of MS-SSIM loss.

Comparative visual results on two images from LOL-v1 dataset for qualitative analysis are presented in Figure 5. In the first image, the effect of our SDIA block can be observed in very dark regions. The output of second image proves the success of our color correction and denoising blocks. The color distribution for this image is the best by far compared to LLIE methods and better than Mirnet-v2 [32]. Also, our method does not produce any artifact like in the result of [26] in third and fifth tassels from the left side.

In order to show the effectiveness of the proposed solution in computation, we compare our method with the state-of-the-art techniques. We utilize a mobile workstation which has a GPU of NVIDIA GeForce RTX 3070M 8GB. The total parameters and inference times of our method

¹Note that in the literature, there are some methods [22, 34] which are using some information coming from GT images while reporting performance values. However, for the sake of fairness, the results reported in this section assume that there is no GT information for any of the methods considered.



Figure 5. Visual comparison of our method on LOL-v1 dataset

and selected approaches are given in Table 6. Computation times are obtained by averaging over 100 runs for an image size of 600x400 which is the size of images in LOL-v1 dataset. As expected, our method is the fastest of all of the supervised networks. When PSNR values are compared with Zero-DCE [12], the difference is larger than 10dB on LoL-v1 dataset. Another notable result is the computation time of the work in [26]. Since it works on patches, its computation time is much lower than the other methods [22] which has similar number of parameters.

Although FLIGHT-Net has the lowest parameters with 25K compared to its main competitors [3, 20, 26], it has the best PSNR value with 24.96 and second-best SSIM value with 0.85 on LOL-v1 dataset. The best result for SSIM belongs to MAXIM [20] has 14.14M parameters which much more than the number of parameters of our method. IAT [3] can be considered as the main competitor of FLIGHT-Net when the number of parameters is taken into account since it is the only method with good performance among com-

petitors with less than 100K parameters, however, its PSNR value is 23.38 which is much lower than the PSNR value of FLIGHT-Net. Zero-DCE has 10k parameters, but its performance is very far from state-of-art.

Depending on our observation, the reason of slightly lower PSNR values in LOL-v2-real dataset is the small misalignment between ground truth and training images. This fact is also the reason of not reporting the performance values of IAT [3] since it follows a different strategy during training to eliminate this misalignment for LOL-v2-real dataset.

4.2. Ablation Study

In order to demonstrate the effectiveness of our sub-blocks and loss functions, five ablation studies are performed on LOL-v1 dataset. The quantitative results of the study are presented in Table 7 and visual results are given in Figure 6.

In the first stage of the ablation study, SDIA and GISP

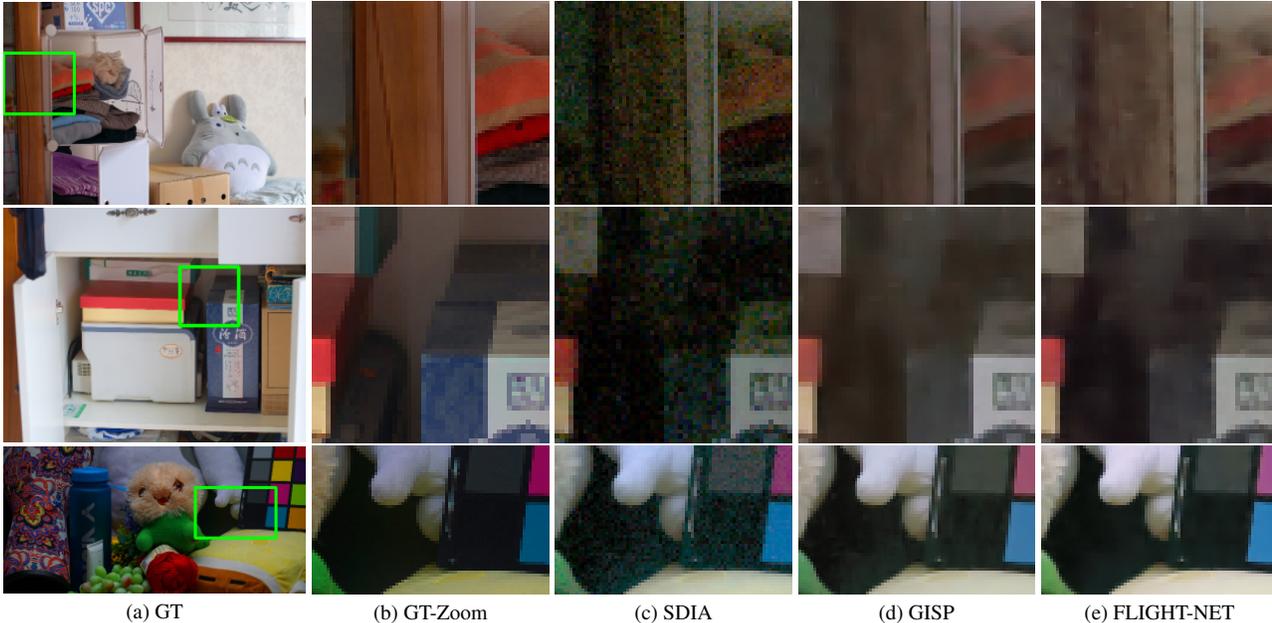


Figure 6. Visual results of ablation study.

blocks are trained separately. As expected, the SDIA branch improved the illumination but also shows the necessity of GISP block. In output image obtained with SDIA alone, we have still spatial noise and color inconsistency and therefore low PSNR and SSIM values. The GISP module handles these issues as mentioned earlier. We also experience the performance of GISP module alone, however, its performance is far behind the overall proposed network. As expected, the SDIA branch increases visibility and transforms the image into the latent feature space required for GISP to fit the target image.

We also test the effects of smooth L1 and MS-SSIM loss functions. PSNR and SSIM values for different combinations of loss functions are reported in Table 7. Training with only smooth L1 or MS-SSIM loss functions are not enough to get the optimum results. PSNR values in the case of using smooth L1 and MS-SSIM loss functions are 22.51 and 22.94 respectively while PSNR value is 24.96 in the case of using both loss functions. As a result, FLIGHT-Net achieves state-of-the-art performance when trained with the combination of smooth L1 and MS-SSIM loss.

| Blocks | Smooth L1 | MS-SSIM | PSNR | SSIM |
|-----------|-----------|---------|-------|------|
| SDIA | ✓ | ✓ | 20.78 | 0.73 |
| GISP | ✓ | ✓ | 23.70 | 0.83 |
| SDIA+GISP | ✓ | ✗ | 22.51 | 0.80 |
| SDIA+GISP | ✗ | ✓ | 22.94 | 0.84 |
| SDIA+GISP | ✓ | ✓ | 24.96 | 0.85 |

Table 7. Results of the ablation study on LOL-v1 dataset

5. Conclusion

In this paper, we have proposed a lightweight yet effective framework for low-light image enhancement problem. Our solution has achieved a state-of-the-art performance for the problem on several datasets with one of the lightest network in the literature. We believe that carefully investigating the forward problem formulation and image signal processing framework and designing the blocks accordingly helps reducing the number of parameters and boosting performance rather than hoping a pure huge bulky DNN to extract and solve for the hidden relations among the features. Lastly, as a future work, we have plans to extend our approach for the processing of low-light videos by considering temporally and spatially varying light conditions.

References

- [1] Andreas Aakerberg, Kamal Nasrollahi, and Thomas B Moeslund. Rellisur: A real low-light image super-resolution dataset. In *Thirty-fifth Conference on Neural Information Processing Systems-NeurIPS 2021*, 2021. 5
- [2] Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua Liu, Siwei Ma, Chunjing Xu, Chao Xu, and Wen Gao. Pre-trained image processing transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12299–12310, 2021. 5
- [3] Ziteng Cui, Kunchang Li, Lin Gu, Shenghan Su, Peng Gao, Zhengkai Jiang, Yu Qiao, and Tatsuya Harada. You only need 90k parameters to adapt light: A light weight transformer for image enhancement and exposure correction. In *BMVC*, 2022. 1, 2, 4, 5, 6, 7

- [4] Rafael C Gonzalez. *Digital image processing*. Pearson Education, 2009. 1, 2
- [5] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1780–1789, 2020. 2, 5, 7
- [6] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on image processing*, 26(2):982–993, 2016. 2, 5
- [7] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018. 2
- [8] Zhixiong Huang, Jinjiang Li, Zhen Hua, and Linwei Fan. Underwater image enhancement via adaptive group attention-based multiscale cascade transformer. *IEEE Transactions on Instrumentation and Measurement*, 71:1–18, 2022. 2
- [9] Andrey Ignatov, Luc Van Gool, and Radu Timofte. Replacing mobile camera isp with a single deep learning model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 536–537, 2020. 2, 4
- [10] Zhiying Jiang, Zhuoxiao Li, Shuzhou Yang, Xin Fan, and Risheng Liu. Target oriented perceptual adversarial fusion network for underwater image enhancement. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(10):6584–6598, 2022. 2
- [11] Edwin H Land. An alternative technique for the computation of the designator in the retinex theory of color vision. *Proceedings of the national academy of sciences*, 83(10):3078–3080, 1986. 2, 4
- [12] Chongyi Li, Chunle Guo Guo, and Chen Change Loy. Learning to enhance low-light image via zero-reference deep curve estimation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 2, 5, 7
- [13] Jiaqian Li, Juncheng Li, Faming Fang, Fang Li, and Guixu Zhang. Luminance-aware pyramid network for low-light image enhancement. *IEEE Transactions on Multimedia*, 23:3153–3165, 2020. 5
- [14] Seokjae Lim and Wonjun Kim. Dslr: Deep stacked laplacian restorer for low-light image enhancement. *IEEE Transactions on Multimedia*, 23:4272–4284, 2020. 2
- [15] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan 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*, pages 10561–10570, 2021. 2, 5
- [16] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11976–11986, 2022. 5
- [17] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Ll-net: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61:650–662, 2017. 2
- [18] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mblen: Low-light image/video enhancement using cnns. In *BMVC*, volume 220, page 4, 2018. 5
- [19] Ardhendu Shekhar Tripathi, Martin Danelljan, Samarth Shukla, Radu Timofte, and Luc Van Gool. Transform your smartphone into a dslr camera: Learning the isp in the wild. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VI*, pages 625–641. Springer, 2022. 1
- [20] Zhengzhong Tu, Hossein Talebi, Han Zhang, Feng Yang, Peyman Milanfar, Alan Bovik, and Yinxiao Li. Maxim: Multi-axis mlp for image processing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5769–5780, 2022. 2, 5, 6, 7
- [21] Wenjing Wang, Chen Wei, Wenhan Yang, and Jiaying Liu. Gladnet: Low-light enhancement network with global awareness. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, pages 751–755. IEEE, 2018. 5
- [22] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 2604–2612, 2022. 2, 5, 6, 7
- [23] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, 2018. 2, 5, 6
- [24] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018. 5
- [25] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018. 2
- [26] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17714–17724, 2022. 2, 5, 6, 7
- [27] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3063–3072, 2020. 5, 6
- [28] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. Band representation-based semi-supervised low-light image enhancement: Bridging the gap between signal fidelity and perceptual quality. *IEEE Transactions on Image Processing*, 30:3461–3473, 2021. 5
- [29] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021. 2, 5
- [30] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Cycleisp: Real image restoration via improved data synthesis. In *CVPR*, pages 2696–2705, 2020. 1, 2

- [31] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*, pages 492–511. Springer, 2020. [2](#), [5](#)
- [32] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for fast image restoration and enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):1934–1948, 2022. [2](#), [5](#), [6](#), [7](#)
- [33] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)*, pages 286–301, 2018. [4](#)
- [34] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM international conference on multimedia*, pages 1632–1640, 2019. [2](#), [5](#), [6](#)
- [35] Zunjin Zhao, Bangshu Xiong, Lei Wang, Qiaofeng Ou, Lei Yu, and Fa Kuang. Retinexdip: A unified deep framework for low-light image enhancement. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(3):1076–1088, 2021. [2](#)
- [36] Karel Zuiderveld. Contrast limited adaptive histogram equalization. *Graphics gems*, pages 474–485, 1994. [1](#), [2](#)