# Shallow Camera Pipeline for Night Photography Enhancement

Simone Zini(✉), Claudio Rota, Marco Buzzelli, Simone Bianco, and Raimondo Schettini

Department of Informatics, Systems and Communication, University of Milano–Bicocca, Milan, Italy
{simone.zini,marco.buzzelli,simone.bianco,raimondo.schettini}@unimib.it, c.rota30@campus.unimib.it

**Abstract.** Enhancing night photography images is a challenging task that requires advanced processing techniques. While CNN-based methods have shown promising results, their high computational requirements and limited interpretability can pose challenges. To address these limitations, we propose a camera pipeline for rendering visually pleasing photographs in low-light conditions. Our approach is characterized by a shallow structure, explainable steps, and a low parameter count, resulting in computationally efficient processing. We compared the proposed pipeline with recent CNN-based state-of-the-art approaches for low-light image enhancement, showing that our approach produces more aesthetically pleasing results. The psycho-visual comparisons conducted in this work show how our proposed solution is preferred with respect to the other methods (in about 44% of the cases our solution has been chosen, compared to only about 15% of the cases for the state-of-the-art best method).

**Keywords:** Night photography enhancement · Low-light image enhancement · Psycho-visual image quality assessment

## 1 Introduction

A digital camera processing pipeline is a series of steps that a digital camera performs to process the raw data captured by its image sensor into a final image. The pipeline is responsible for applying various adjustments to the image, such as corrections for lens distortion, white balancing, noise reduction, sharpening, and color enhancement. Although camera manufacturers may use varying processing algorithms and stages in their pipelines, these basic steps are commonly involved in most digital camera processing pipelines [9]. The parameters of the single processing modules are usually optimized by manufacturer for daylight or flashlight illuminated scenes. The problem addressed in this paper is the enhancement of night scenes when the rendering intent is not only the visibility of spatial details, but also to keep the naturalness of the scene depicted and possibly to improve


G. L. Foresti et al. (Eds.): ICIAP 2023, LNCS 14233, pp. 51–61, 2023.
https://doi.org/10.1007/978-3-031-43148-7_5

the aesthetic of the photos. From a technical point of view, when processing nighttime images multiple challenges occur in comparison to the processing of daytime scenes. Nighttime scenes are typically much darker: this can result in images with a low signal-to-noise ratio, making it difficult to extract useful information from the image. Color casts are generally present, due to the sometimes different artificial lighting sources in the scenes. Also, due to the extended exposure time necessary for shooting in dark scenarios, motion blur and sensor noise are likely to occur in the final results. To address these problems, recent methods that leverage deep neural networks to enhance low-light images have been proposed, achieving remarkable results [7,8,11,19,22]. Zhang *et al.* [22] designed a Convolutional Neural Network (CNN) that decomposes images into two components responsible for light adjustment and degradation removal, respectively, and is trained with paired images shot under different exposure conditions. Jiang *et al.* [8] proposed to enhance low-light images using Generative Adversarial Networks (GANs), exploiting unpaired data for the training of the model. Yang *et al.* [19] proposed a semi-supervised model that integrates CNNs and GANs to enhance low-light images in two stages: the first one learns a coarse-to-fine band representation and infers different band signals jointly, while the second one recomposes the band representation using adversarial learning. Recently, Guo *et al.* [7] proposed zero-shot learning methods to eliminate the requirement of paired and unpaired data, which was later improved by Li *et al.* [11]. However, these approaches suffer from several problems, such as over-enhancement, color distortion, and loss of details. Moreover, their heavy computational complexity, memory consumption, and energy requirements may not always meet the constraints for onboard deployment in digital cameras.

In this paper, we introduce a camera pipeline for the rendering of visually pleasing photographs in low-light conditions, containing several algorithms that address the challenges presented by low-light images and characterized by a shallow structure and by a low parameter count. At a time when the resolution of most imaging problems is delegated to the direct or indirect use of neural networks, we propose a "traditional" processing pipeline, whose main modules are designed on the basis of our knowledge of the mechanisms of human vision, and on the basis of our knowledge of the main limitations of traditional imaging devices. The few parameters of the different modules are heuristically set by the authors according to their personal preferences [3], without any reference to existing datasets of low-light images corrected by human experts or automatic approaches. This low parametric dependency means that our solution is flexible, as it can be potentially tuned to match individual users' preferences and to different sensors.

To prove the effectiveness of our method we adopted the dataset used in the NTIRE2022 Night Photography Rendering challenge [6]. Psycho-visual experimental results involving real user evaluations show that our solution produces more pleasing results with respect to several CNN-based state-of-the-art methods for low-light image enhancement.

# 2 Proposed Method

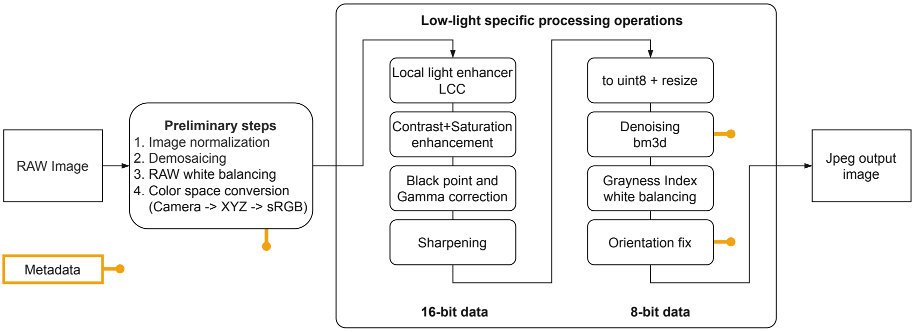


**Fig. 1.** Overview of the complete proposed pipeline. The entire pipeline can be divided in two parts: preliminary data preparation steps and low-light processing steps. Metadata extra information is exploited in the steps marked with the orange dot. (Color figure online)

The scheme of the proposed solution is depicted in Fig. 1. Our pipeline can be divided into two parts: the preliminary steps, which are the basic stages of a typical camera processing pipeline, and the low-light specific part, which instead contains steps to specifically handle night images. We refer multiple times within our pipeline to image metadata, which are indicated in the following using *italic text*.

## 2.1 Preliminary Steps

The first part of our pipeline is made of four steps working in the RAW domain. The first step is image normalization: the *black_ level* as provided in the image metadata is subtracted, and the image values are rescaled so that the *white_ level* is set to one. The demosaicing operation converts the single-channel RAW image into the three-channel RGB image using the appropriate color filter array pattern (*cfa_ array_ pattern*). Then, a preliminary automatic white balance step is performed using the Gray World algorithm [4], in order to provide a first approximate correction of the image cast. Finally, a color transformation step converts the image from the camera-specific color space to XYZ (obtained as the inverse of *color_ matrix_ 1*) and finally to the sRGB color space.

## 2.2 Low-Light Specific Processing Operations

The second part of our pipeline has been specifically designed to handle images taken by night in low-light conditions.

The first step of this second part is the use of the Local Contrast Correction (LCC) algorithm by Moroney [15]. Here the local correction is performed on the Y channel of the YCbCr color space using a pixel-wise gamma correction, whose values are determined using a mask $M$ obtained by blurring the luminance channel Y with a Gaussian filter in order to brighten dark areas and to not clip pixels that are already bright. The corrected $\hat{Y}$ channel image is obtained as

$$\hat{Y} = Y^{\gamma^{\frac{0.5-(1-M)}{0.5}}}, \tag{1}$$

where $M$ is computed as previously described, and $\gamma$ is the value of the exponent for gamma correction. According to Schettini *et al.* [18], we computed $\gamma$ as:

$$\gamma = \begin{cases} \frac{\ln(0.5)}{\ln(\bar{Y})} \text{ if } \bar{Y} \geq 0.5 \\ \frac{\ln(\bar{Y})}{\ln(0.5)} \text{ otherwise} \end{cases}, \tag{2}$$

where $\bar{Y}$ is the average value of the Y channel. Since $1-M$ inverts the computed mask, bright areas are darkened by a gamma value lower than 1, and dark areas are brightened by a gamma value greater than 1.

The application of LCC tends to reduce the overall contrast and saturation, as noted by Schettini *et al.* [18]. Therefore, as subsequent steps, we perform contrast and saturation enhancement.

The contrast enhancement step adaptively stretches and clips the image histogram based on how the distribution of dark pixels changes before and after the contrast correction of LCC. Every histogram computed has 256 bins. The histogram range used for stretching and clipping is defined as follows: let any given pixel be "dark" if, in the YCbCr color space, its Y value is lower than 0.14 and its chroma radius, as defined in [18], is lower than 0.07. The lower range for histogram stretching is defined by the number of dark pixels after the application of LCC. If there is at least one pixel, the lower range is given by the difference of the bins corresponding to 30% of dark pixels in the cumulative histogram of $\hat{Y}$ and $Y$, which represent the output and input of LCC in Eq. 1. If there are no dark pixels, the lower range corresponds to the 2nd percentile value of the $\hat{Y}$ histogram. Concerning the "bright" pixels, the upper range for histogram stretching always corresponds to the 98th percentile value of the $\hat{Y}$ histogram. For both ranges, the maximum number of bins to clip is 50. Using the determined range, the image histogram is stretched and the histogram bins that fall outside are clipped.

For the saturation enhancement step, we correct each RGB channel as suggested by Sakaue *et al.* [17]:

$$\hat{C} = 0.5 \times \frac{\hat{Y}}{Y} \times (C + Y) + C - Y, \tag{3}$$

where $C$ stands for each RGB channel, $\hat{C}$ is the corresponding output channel, $\hat{Y}$ and $Y$ are the output and input Y channels used in Eq. 1.

After contrast and saturation enhancement, a black point correction step is performed in order to restore the natural aesthetics of night images, since LCC adjusts local statistics but produces an overall washed-out result. This operation

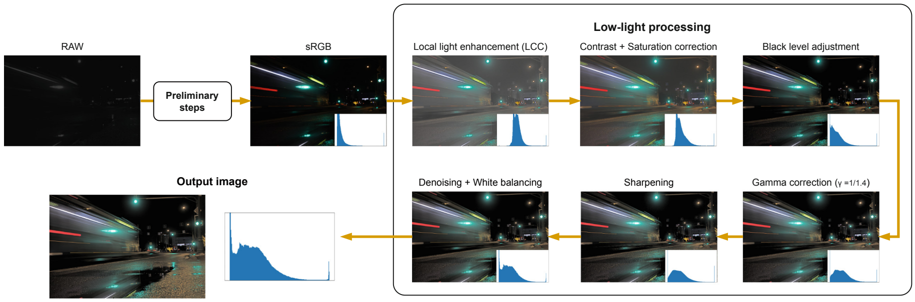


**Fig. 2.** Step-by-step results of the proposed pipeline. Along with images, we also reported histograms to show how the global pixel distribution changes.

is performed by clipping to zero all pixels below the 20th percentile value of the value channel V in the HSV color space. After this operation, a global gamma correction is performed with a gamma value set to $\frac{1}{1.4}$, followed by a sharpening operation using unsharp masking.

The image is then converted to 8-bit encoding, resized to match the predefined output size (imposed by the challenge to be $1300 \times 866$ for landscape orientation and $866 \times 1300$ for portrait one), and processed with the Block-Matching and 3-D Filtering (BM3D) denoising algorithm [5] to remove noise introduced by the poor light conditions typical of night scenes. Here the *noise_profile* value from the image metadata is used to determine the strength of the denoising operation, which is controlled by BM3D through a parameter $\sigma$ that encodes an estimate of the noise standard deviation, used internally to control the parameters of the method. According to the distribution of the *noise_profile* values in the training data, we defined three classes representing different noise intensities, and we empirically assigned a $\sigma$ value (0.2, 0.6 and 0.8) to each class. Since noise is more visible in dark regions rather than in bright regions and BM3D removes part of the high-frequency information, we performed a blending operation in RGB using a mask generated by blurring the luminance channel Y of the original noisy image in the YCbCr color space with a Gaussian filter. The final denoised image $\hat{D}$ is computed as

$$\hat{D} = I_{BM3D} \times (1 - mask \times u) + I \times (mask \times u), \tag{4}$$

where $I_{BM3D}$ is the image denoised with BM3D, $I$ is the original noisy image and the $u$ parameter, empirically set to 0.6, controls the denoising effect in bright areas.

A second automatic white balance step is performed, this time on non-linear processed RGB data, in order to reduce color casts in those scenarios where the initial Gray World approach may have failed. Here the Grayness Index (GI) algorithm [16] is used. GI is very sensitive to noise, hence we estimated the image illuminant on the image $I_{BM3D}$, then we normalized it by its maximum value and applied it to the image $\hat{D}$ obtained after the blending operation in Eq. 4. The image is then rotated in relation to the *orientation* information stored in the metadata and finally saved as JPEG image at quality 100.

As illustrative example, Fig. 2 shows intermediate images and histograms of the proposed pipeline after each step. The application of LCC [15] improves the local contrast but centers the image histogram and reduces the overall saturation, hence the contrast and saturation enhancement step is necessary to correct this behavior. Yet, the obtained histogram is still biased towards the center of the dynamic range, and a black level adjustment is fundamental to restore the natural anesthetic of the image. Here a gamma correction can increase the overall brightness. Since BM3D [5] effectively removes noise but also part of the details, a preliminary sharpening operation that strengthens high frequencies helps preventing this problem.

## 3 Experiments

### 3.1 Dataset

We adopted the dataset used in the NTIRE2022 Night Photography Rendering challenge [6], which provides 250 RAW-RGB images of night scenes captured using a Canon EOS 600D device and encoded in 16-bit PNG files. Each RAW image has a resolution of $3646 \times 5202$ pixels. Image metadata are also available in JSON format. Due to the nature of the challenge, ground truth images are not available. According to the challenge organization, 50 images are provided as train set, 50 as the first validation set, 50 as the second validation set, and the remaining 100 as the final validation set (among these 100, only 50 were selected for the final evaluation). Since our solution does not need a training procedure, we used the train set to empirically select the few parameters required by our pipeline and used all validation sets to validate the results.

### 3.2 Results and Discussion

We evaluated our pipeline by comparing it with a subset of state-of-the-art approaches for low-light image enhancement and with other solutions that participated in the NTIRE2022 Night Photography Rendering challenge [6] using psycho-visual comparisons and Mean Opinion Score (MOS).

We selected eight recent state-of-the-art approaches for low-light image enhancement and performed a psycho-visual evaluation test using the same 50 images from the validation set, processed by the selected methods. More precisely, we selected DRBN [19], Kind [22] and Kind++ [21], TBEFN [14], EnlightenGAN [8], ExCNet [20], Zero-DCE [7] and Zero-DCE++ [11]. Since these methods expect images to be in sRGB color space, we used the preliminary steps described in Sect. 2 to convert the RAW images into sRGB images and applied these enhancement methods to them. The resulting images have been obtained using the LLIE platform [10]. The comparison has been performed by

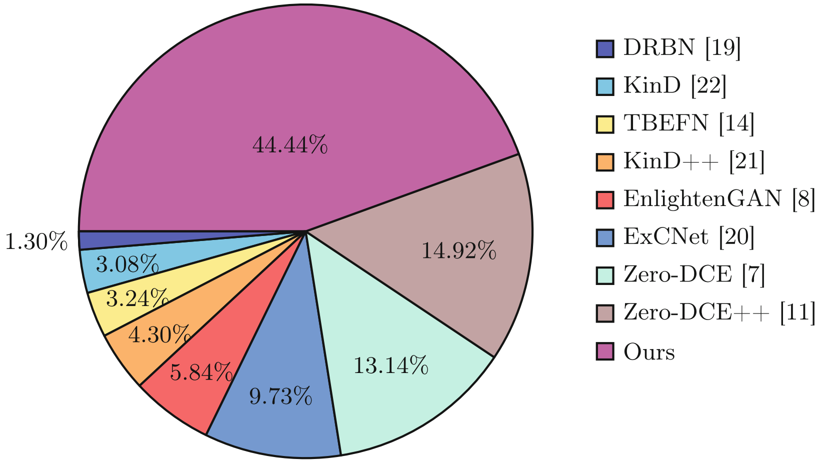


**Fig. 3.** Pie chart reporting the distribution of the results of the psycho-visual test, in terms of number of preferences, obtained comparing our proposed method with other state-of-the-art low-light image enhancement approaches.

a total of 31 users. Each user involved in the evaluation was shown a $3 \times 3$ grid containing the same image enhanced by the nine methods and was asked to click on the preferred one. This process was repeated for each of the 50 images. Grid composition and image order were randomly generated. The evaluation was done on monitors between 24 and 27 in. under controlled lighting conditions, and the images were shown on black background.

Figure 3 shows the pie chart with the results. As can be seen, in almost 45% of the cases the proposed approach is preferred with respect to the other ones. Zero-DCE [7] and Zero-DCE++ [11] obtained almost the same number of votes, followed by ExCNet [20]. From this first analysis, it is easy to notice how the proposed approach leads to more appreciated images with respect to the other methods. In order to provide a visual comparison for the readers, some results of the proposed pipeline are shown in Fig. 4. We also reported the same images corrected using different state-of-the-art deep learning-based low-light image enhancement approaches. In this figure are reported four different cases: the first two are cases in which our approach received the highest consensus in terms of user votes, while the last two are a mid-case scenario and a worst-case scenario, respectively. We can observe how our solution is better at removing noise, increasing sharpness, reducing color cast and preserving the mood that is typical of night scenes. This produces more pleasing results, as also confirmed by the vote distribution in Fig. 3. It is worth noting that the number of votes received by our solution is considerably higher than the votes received by other methods when it obtained the highest score. Instead, when other methods were preferred to ours, their vote counts are comparable to the number of votes received by our solution.

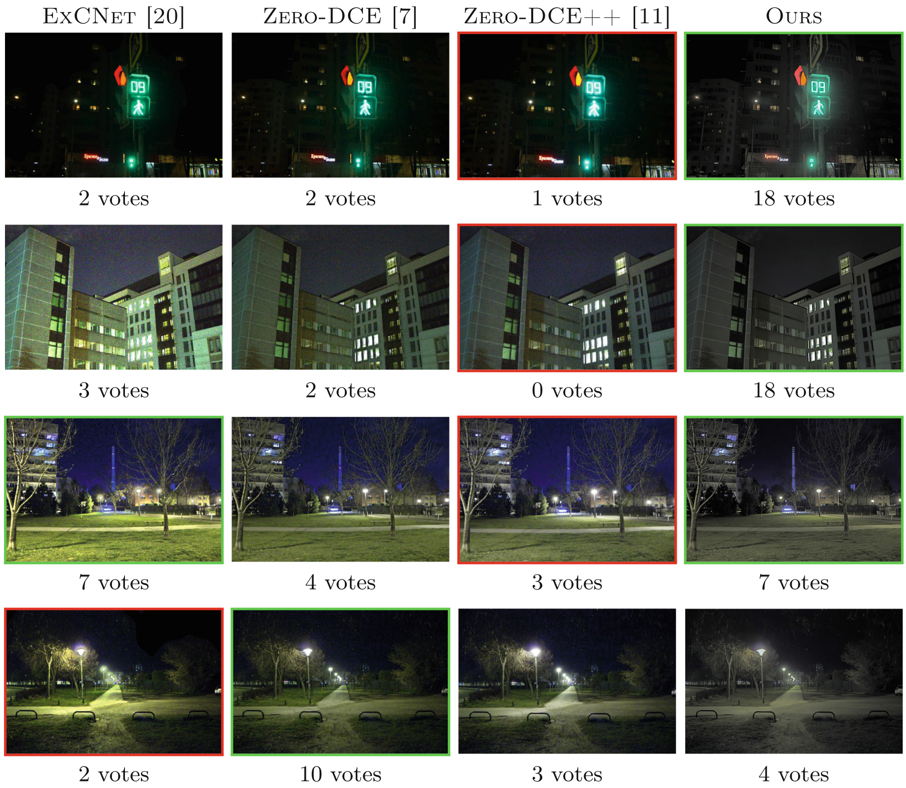


**Fig. 4.** Visual comparison between the proposed method and three highest score state-of-the-art approaches. Our solution produces sharper results, better reduces noise and color cast, and better maintains the mood of night photographs while the others tend to over-saturate colors and light casts. For each row, images framed in green are the ones with the highest score while the ones framed in red are the ones with the worst score. (Color figure online)

For what concerns the NTIRE2022 challenge, MOS results are obtained through visual comparison on the Yandex Toloka platform. Here every submission, consisting of 50 images of the final validation set, was included in 3250 comparisons. The results of the final leaderboard are reported in Table 1. As shown, our pipeline won the fifth place in the challenge obtaining 1935 votes. Note that our solution received only 112 fewer votes than the second winning solution (i.e. about 5% fewer votes) that uses different neural models for most of the operations in its pipeline [12].

In Table 1 we also add a further column, named significance score. First of all, for each method we compute the 95% confidence interval of the Score using the Binomial test. The significance score for each solution corresponds to the number

**Table 1.** Final leaderboard of the NTIRE2022 Night Photography Rendering challenge [6]. Every submission (50 images) was included in 3250 comparisons using the Yandex Toloka platform. Our team is highlighted in bold.

| Rank | Team | Score | Votes | Sign. Score |
|---|---|---|---|---|
| 1 | MIALGO | 0.8009 | 2603 | 12 |
| 2 | Sorashiro | 0.6298 | 2047 | 11 |
| 3 | Feedback | 0.6089 | 1979 | 11 |
| 4 | OzU-VVGL | 0.6045 | 1964 | 11 |
| **5** | **IVLTeam** | **0.5955** | **1935** | **10** |
| 6 | NoahTCV | 0.5742 | 1866 | 9 |
| 7 | NTU607QCO | 0.4798 | 1559 | 6 |
| 8 | Winter | 0.4631 | 1505 | 6 |
| 9 | Sigma_WHU | 0.4411 | 1433 | 5 |
| 10 | Namecantbenull | 0.3965 | 1288 | 3 |
| 11 | BISPL | 0.3683 | 1197 | 3 |
| 12 | Baseline | 0.2734 | 888 | 1 |
| 13 | Low Light Hypnotize | 0.0182 | 59 | 0 |

of solutions with respect to which it is statistically better or equivalent, i.e. the number of confidence intervals that are lower or overlap with the current one. The significance score highlights how the result achieved by the first solution [13] is statistically better than all the others, while the solutions ranked from the second position to the fourth one are actually statistically equivalent and therefore rank in the second place. They are followed by our solution, which ranks in the third place and is statistically better than all the remaining solutions. The results have been additionally evaluated by a professional photographer, who awarded our solution with the sixth place in the final leaderboard [6].

## 4 Conclusions

We have proposed a low-complexity handcrafted camera pipeline for the rendering of visually pleasing night photographs. Our solution includes several processing steps that address the challenges presented by low-light images and depends on a small number of free parameters, which we empirically set according to our personal preferences. However, the optimal parameters could be easily found with optimization methods if one had a suitable training set (whose cardinality however should not be so high as in the case of neural networks).

The effectiveness of our pipeline was validated through experiments involving real users, which demonstrated that our method produces more visually appealing results than other state-of-the-art methods for low-light image enhancement. These results have been further evaluated in the context of the NTIRE2022

Night Image Rendering challenge [6], also demonstrating that traditional imaging pipelines can compete with modern deep learning-based methods.

An interesting next step could be the parametrization and optimization of the proposed pipeline, adopting unsupervised training solutions to model user preferences [23]. Another promising research direction is that of exploiting saliency [1] to perform spatially varying enhancement or to exploit no reference image aesthetic metrics [2] to drive model parameters selection.

## References


1. Bianco, S., Buzzelli, M., Ciocca, G., Schettini, R.: Neural architecture search for image saliency fusion. Inform. Fusion **57**, 89–101 (2020)
2. Bianco, S., Celona, L., Napoletano, P., Schettini, R.: Predicting image aesthetics with deep learning. In: Blanc-Talon, J., Distante, C., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2016. LNCS, vol. 10016, pp. 117–125. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-48680-2_11
3. Bianco, S., Cusano, C., Piccoli, F., Schettini, R.: Personalized image enhancement using neural spline color transforms. IEEE Trans. Image Process. **29**, 6223–6236 (2020)
4. Buchsbaum, G.: A spatial processor model for object colour perception. J. Franklin Inst. **310**(1), 1–26 (1980)
5. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007)
6. Ershov, E., et al.: NTIRE 2022 challenge on night photography rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (2022)
7. Guo, C., et al.: Zero-reference deep curve estimation for low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1780–1789 (2020)
8. Jiang, Y., et al.: Enlightengan: deep light enhancement without paired supervision. IEEE Trans. Image Process. **30**, 2340–2349 (2021)
9. Karaimer, H.C., Brown, M.S.: A software platform for manipulating the camera imaging pipeline. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 429–444. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_26
10. Li, C., Guo, C., Han, L., Jiang, J., Cheng, M.M., Gu, J., Loy, C.C.: Low-light image and video enhancement using deep learning: A survey. IEEE Trans. Pattern Anal. Mach. Intell. **44**(12), 9396–9416 (2021)
11. Li, C., Guo, C., Loy, C.C.: Learning to enhance low-light image via zero-reference deep curve estimation. IEEE Trans. Pattern Anal. Mach. Intell. **44**(8), 4225–4238 (2021)
12. Li, Z., Yi, S., Ma, Z.: Rendering nighttime image via cascaded color and brightness compensation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 897–905 (2022)
13. Liu, S., et al.: Deep-flexisp: a three-stage framework for night photography rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1211–1220 (2022)

14. Lu, K., Zhang, L.: Tbefn: A two-branch exposure-fusion network for low-light image enhancement. IEEE Trans. Multimedia **23**, 4093–4105 (2020)
15. Moroney, N.: Local color correction using non-linear masking. In: Color and Imaging Conference, vol. 2000, pp. 108–111. Society for Imaging Science and Technology (2000)
16. Qian, Y., Kamarainen, J.K., Nikkanen, J., Matas, J.: On finding gray pixels. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8062–8070 (2019)
17. Sakaue, S., Nakayama, M., Tamura, A., Maruno, S.: Adaptive gamma processing of the video cameras for the expansion of the dynamic range. IEEE Trans. Consum. Electron. **41**(3), 555–562 (1995)
18. Schettini, R., Gasparini, F., Corchs, S., Marini, F., Capra, A., Castorina, A.: Contrast image correction method. J. Electron. Imaging **19**(2), 023005 (2010)
19. Yang, W., Wang, S., Fang, Y., Wang, Y., Liu, J.: 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, pp. 3063–3072 (2020)
20. Zhang, L., Zhang, L., Liu, X., Shen, Y., Zhang, S., Zhao, S.: Zero-shot restoration of back-lit images using deep internal learning. In: Proceedings of the 27th ACM International Conference on Multimedia, pp. 1623–1631 (2019)
21. Zhang, Y., Guo, X., Ma, J., Liu, W., Zhang, J.: Beyond brightening low-light images. Int. J. Comput. Vision **129**, 1013–1037 (2021)
22. Zhang, Y., Zhang, J., Guo, X.: Kindling the darkness: A practical low-light image enhancer. In: Proceedings of the 27th ACM International Conference on Multimedia, pp. 1632–1640 (2019)
23. Zini, S., Buzzelli, M., Bianco, S., Schettini, R.: A framework for contrast enhancement algorithms optimization. In: 2022 IEEE International Conference on Image Processing (ICIP), pp. 1431–1435. IEEE (2022)