

Article

Deep Refinement Network for Natural Low-Light Image Enhancement in Symmetric Pathways

Lincheng Jiang ^{1,2}, Yumei Jing ^{3,*}, Shengze Hu ¹, Bin Ge ¹ and Weidong Xiao ¹

¹ Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073, China; linchjiang@gmail.com (L.J.); shengzehucn@gmail.com (S.H.); bingge@nudt.edu.cn (B.G.); wdxiao@nudt.edu.cn (W.X.)

² Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

³ School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

* Correspondence: yumeijing@pku.edu.cn

Received: 31 August 2018; Accepted: 11 October 2018; Published: 12 October 2018



Abstract: Due to the cost limitation of camera sensors, images captured in low-light environments often suffer from low contrast and multiple types of noise. A number of algorithms have been proposed to improve contrast and suppress noise in the input low-light images. In this paper, a deep refinement network, LL-RefineNet, is built to learn from the synthetic dark and noisy training images, and perform image enhancement for natural low-light images in symmetric—forward and backward—pathways. The proposed network utilizes all the useful information from the down-sampling path to produce the high-resolution enhancement result, where global features captured from deeper layers are gradually refined using local features generated by earlier convolutions. We further design the training loss for mixed noise reduction. The experimental results show that the proposed LL-RefineNet outperforms the comparative methods both qualitatively and quantitatively with fast processing speed on both synthetic and natural low-light image datasets.

Keywords: low-light image; image enhancement; deep refinement network

1. Introduction

Video surveillance is of vital importance for various automated decision-making tasks including security application and medical care [1–3]. The surveillance systems consist of three basic modules: real-time monitoring and reconnaissance, collecting data from the dynamic surveillance zone and processing the information contained in the images. The images captured in the day time with enough light often have good visual quality and significant surveillance effects. However, the images captured in night or other insufficient light conditions are hard to satisfy for surveillance purposes for the dark pixels which make the visibility greatly reduced. Therefore, it is very important to design effective algorithms to process and enhance the low-light images in video surveillance systems. In addition, the low-light image enhancement algorithms are also in demand in other applications, such as unmanned vehicles, intelligent robot and object recognition [4,5].

A lot of enhancement algorithms have been proposed to enhance image contrast, reduce image noise and improve image brightness for low-light images. Histogram equalization (HE) algorithm is a very early approach which focuses on the problem [6]. HE is a simple and classic strategy that reassigns the value of each pixel to make them satisfy uniform distribution. The algorithm often leads to significant loss of image details. Under the inspiration of the Retinex theory, some new algorithms are proposed. The Retinex-based algorithms adopt Gaussian low-pass filtering in illumination components and logarithmic transformation in reflectance components. In [7], the input images are separated into the two components and then the enhanced image is generated by utilizing

the reflectance part. However, it is difficult to separate different components in this method, which may lead to unpleasant artifacts to the final results. In addition, an enhancement algorithm is proposed by [8], which implements a de-hazing process for the input images and produces the corresponding high-quality enhanced images. However, the algorithm may make the images over enhanced and exaggerated. Most enhancement methods are proposed on the basis of the three classic algorithms above and achieve some improvement.

In recent years, methods based on deep learning framework have drawn enormous interest due to their effectiveness [9–21]. These approaches are superior to previous traditional algorithms in multiple areas of computer vision, including object recognition [22], object detection [23], semantic segmentation [3], super resolution [24], and image denoising [25].

In this paper, we design a novel deep refinement network, LL-RefineNet, to deal with the enhancement of input low-light images, which learns from the synthetical low-light images and performs enhancement for both synthetical and real low-light images. We train the proposed network to effectively learn from the input low-light images and extract useful features. Then, the input images are enhanced by the trained network. In our model, structural features of the objects are learned, which are robust to noise environments. To generate the training data in our experiment, we modify source images collected from the Internet to simulate the natural low-light images. Similarly, we add noise and darkening to produce test images, on which we perform the evaluations and compare the proposed model to other relative state-of-the-art methods. Natural low-light images are further collected using cell-phone cameras, which are used to test the performance of the proposed network on real data. The flowchart of the proposed method is shown in Figure 1.

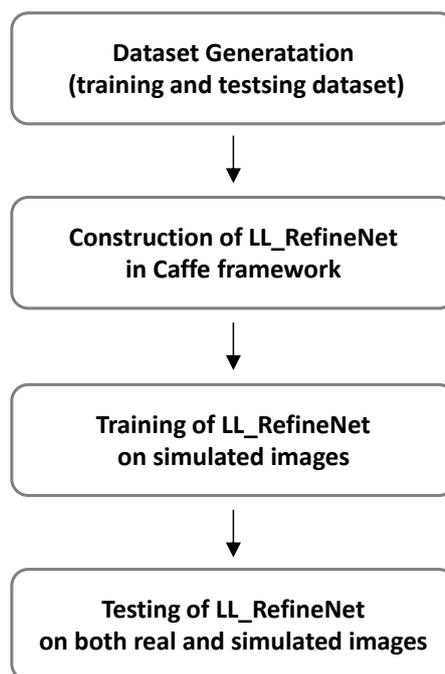


Figure 1. The flowchart of the proposed method.

We summarize the major contributions by the following four ingredients:

- To the best of our knowledge, we are the first to apply deep refinement network to low-light image enhancement. In our approach, high-level features with global information are fused with the low-level features with local information. By means of the long-range residual links and connections, the proposed network can effectively propagate the gradient backwards to the earlier layers in the network;

- An effective loss for mixed noise is also designed to improve the performance of our method under the real low-light condition.
- In addition, whole images are used for our training process instead of image patches, which avoids the problem of dealing with small patches and helps our method to achieve better results.
- Through comprehensive experiments on natural and synthetic low-light images, the proposed LL-RefineNet is demonstrated to outperform state-of-the-art models compared with benchmark algorithms both qualitatively and quantitatively with high processing speed.

In the following sections, the content of the paper is organized as: related work is reviewed in Section 2, which introduces recent algorithms for low-light image improvement and deep learning based image processing tasks. The framework of the proposed LL-RefineNet is discussed in Section 3. Results of our experiments are provided in Section 4 and the conclusions are in Section 5.

2. Related Work

2.1. Traditional Low-Light Image Enhancement Methods

In this paper, we refer to the algorithms which do not apply deep learning networks as traditional algorithms in the area of low-light image enhancement. A variety of approaches have been proposed for image enhancement. Histogram equalization (HE) is an easy but classic approach among these algorithms. Some algorithms based on HE were proposed to solve problems in the original version and improve the effect later, such as CLAHE [26], BBHE [27], DHE [28], QBHE [29], and so on. Then, Wu et al. introduced the OCTM [30] optimization method, which utilized the mathematical transfer function (MTF) to perform the mapping of contrast-tone for the input low quality image. In addition, other methods based on nonlinear functions such as the gamma function were also presented to perform image enhancement. Furthermore, a number of denoising algorithms have been proposed by researchers, including methods based on nonlinear filters [31], the well known BM3D algorithm [32] and the sparsity representation based K-SVD algorithm [33].

Low-light image enhancement is more challenging due to severe noise and big areas of dark pixels. Some algorithms were designed especially for these images. Zhang et al. [34] enhanced very low-light images by improving the de-haze algorithm for an inverted image, which made the dark regions become brighter and the contrast enhanced. A joint bilateral filter was also used in the paper to suppress the noise. An image enhancement algorithm was presented in [35] by means of sparse representations. In this approach, enhanced images were generated based on the sparse representations of input patches in the constructed dictionary. In [36], a novel method was introduced to deblur low-light images by making use of light streaks in the images. The algorithm can find useful light streaks automatically and the experimental results showed excellent effect on challenging real-world datasets. Guo et al. [37] proposed a low-light image enhancement (LIME) method, which estimated the illumination of each pixel and utilized illumination map estimation. In [38], dual-tree complex wavelet transform was used to decompose the input low-light image into two parts, including low-pass subbands and high-pass subbands. The algorithm enhanced the images by further processing the two parts.

2.2. Deep Learning Methods in Image Process Applications

In the area of image processing, great success has been achieved by deep learning-based approaches, thanks to the availability of large datasets as well as the development of computing capabilities. In [39], a general framework with a deep learning method was taken for image representation by inverting the images. The results showed that the algorithm retained photographically accurate information about the image. Chan et al. [40] proposed a simple and efficient deep learning network named PCANet for image classification on the basis of some fundamental data processing parts. In [41], a deep learning method was applied to the field of fast image retrieval in which an outstanding deep learning framework was designed to generate binary hash codes.

In the field of image enhancement, VGG filters were used in the very deep super-resolution (VDSR) model [42] for image super resolution, which applied 20 convolutional layers to generate the final enhanced result. In [25], a similar network structure was used to construct the denoising convolutional neural network (DnCNN) model, in which the batch normalization layers were placed after convolutional layers. This method reported better results compared to the conventional denoising algorithms in terms of peak signal-to-noise ratio (PSNR) and visual performance. The structure of denoising autoencoders was introduced by [43] to extract information contained in the noisy inputs, in which convolutional neural networks (CNN) were utilized to remove noise in input images. Furthermore, a multi-column framework was proposed by [44], which performed robust image denoising. The training and testing images were generated with different noise levels and diverse types of noise. Authors in [45] designed a stacked architecture of denoising autoencoders. In this model, encoding layers of the widely-used multilayer perceptron (MLP) were explored to achieve the reconstructions of high-quality images.

Recently, Lin et al. [46] proposed a novel deep learning method which designed a multi-path refinement network to improve the effect of deep convolutional neural networks. Experimental results showed its great advantage over deep convolutional neural networks.

2.3. Deep Learning Methods for Low-Light Image Enhancement

Lore et al. [47] adopted deep neural networks named LLNet to enhance low-light images, which is also the first one to introduce a deep learning method to the field of low-light image enhancement. The network was a variant of the stacked-sparse denoising and contrast-enhancement autoencoder. Results showed significant promotion of the method compared with traditional low-light image enhancement methods. In LLNet, there are no convolutional layers.

In [48], a deep network framework, low-light convolutional neural network (LLCNN), was designed to perform enhancement for input low-light images and improve the effect of LLNet. This network was constructed based on the structure of a residual convolutional neural network. In this approach, low-light images were simulated by adding noise and darkening to the clear source images using a nonlinear method. The training process was conducted on this dataset, and then this model can enhance both the simulated and real low-light images.

In this paper, we designed a new application of deep refinement network, LL-RefineNet, to deal with the enhancement of input low-light images, which will be introduced in detail in the following sections.

3. The Proposed Model

In this section, the proposed LL-Refinement network is presented as well as the training strategy and dataset generation.

3.1. Refinement in Symmetric Pathways

In the proposed LL-RefineNet, multi-level image features are learned and extracted to predict high-resolution image enhancement results. In our approach, high-level features with global information are fused with the low-level features with local information to perform enhancement for the input low-light images. By means of the long-range residual links and connections, the proposed network can effectively propagate the gradient backwards to the earlier layers in the network. This is of vital importance to our method, which enables our network to achieve efficient end-to-end training.

The proposed LL-RefineNet consists of two components in symmetric pathways: forward encoder part and backward decoder part. In our model, a 4-cascaded framework is designed, including four sub-nets in the decoder part. Each of these sub-nets takes the output of the related layers in the earlier encoding part as input. In addition, the output of the preceding sub-net is also connected to the following sub-net. In spite of the fact that the structures of all the sub-nets are similar, the parameters of the sub-nets are not tied, which enables each network to perform flexible adaptation.

Now, we follow the content in Figure 2 in a bottom-up manner and discuss how our model works. The right part of Figure 2 shows the proposed LL-RefineNet, while the left part of Figure 2 displays the detailed refinement sub-net. Beginning from RefineNet-1 that contains one input, the refinement part starts to process the features from the decoder part. In the next stage, the outputs of RefineNet-1 and early layer conv6 are passed to the following RefineNet-2 as the two inputs. High-resolution features generated from the corresponding encoder layer are used by RefineNet-2 to perform refinement for the features from RefineNet-1 which are of low-resolution.

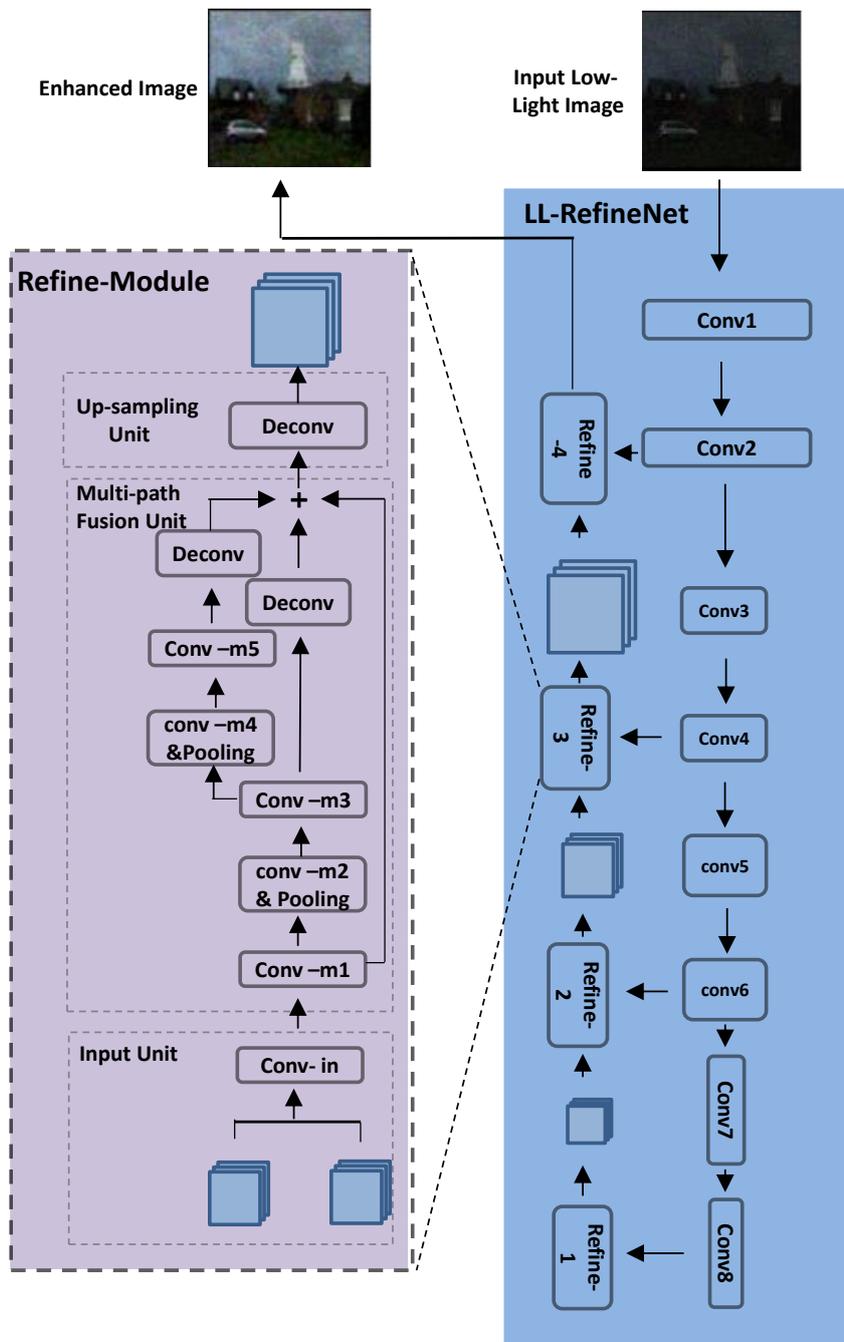


Figure 2. Architecture of the proposed LL-RefineNet.

In a similar manner, low-resolution features of the later layers and high-resolution features of the earlier layers are fused to repeat the refinement process for RefineNet-3 and RefineNet-4. In the end, the network generates the final enhanced result which is of high-resolution.

Effective end-to-end learning is achieved for our model. The residual connections contained in the whole network is of vital importance to the training process. In the forward process, low-level features containing detailed information are passed by the residual connections to refine the coarse feature maps with high-level information. On the other hand, gradients are directly propagated to earlier layers with the help of these connections. This further enables an effective end-to-end learning process.

3.2. LL-RefineNet

The architecture of our proposed network is illustrated in Figure 2. First, high-dimensional low-resolution feature maps are generated by the forward pathway, which contain appearance and structure information. Then, these features and the intermediate features from the forward pass are gradually fused in the backward refinement pass. Four cascaded sub-net blocks (Refine-1 to 4 in Figure 2) are contained in the refinement decoder part. For each of these four sub-nets, the output generated by the previous sub-net together with the feature maps from the encoder part are taken as inputs, while the output is passed to the next sub-net. Each sub-net increases the spatial resolution of feature maps by the factor of 2 and the final high resolution result is generated by the network. In detail, each sub-net in the decoder pass consists of three units: input unit, multi-path fusion unit and up-sampling unit. We will provide detailed descriptions as follows:

1. Input unit:

Two inputs are passed to each input unit, which are the feature maps generated by the previous refinement sub-net and the intermediate feature maps from the encoder part. The input unit performs concatenation for the two inputs, and then processes the inputs by a convolution layer (conv-in in Figure 2).

2. Multi-path fusion unit:

The network then passes the feature maps output by the input unit to the next module: the multi-path fusion unit. In this unit, context information is further learned from a larger receptive range. Features are extracted with the help of increasing window sizes, which are then fused to produce output feature maps for subsequent processing. We are motivated by the fact that multi-scale feature fusion performed by deep learning approaches can help the network to be more robust to the scales of images. The multi-path fusion unit is designed as a combination of multiple convolution layers (conv-m1 to conv-m5). Among these layers, two convolution layers (conv-m2 and conv-m4) are followed by pooling operations that are utilized to achieve a larger receptive field. Conv-m2 and conv-m3 further re-use the feature maps from conv-m1 to generate more complex feature representations. In a similar manner, conv-m4 and conv-m5 continue to use the feature maps generated by conv-m3. In the end, feature maps obtained from multiple paths are combined by summation, where deconvolution operations are used to achieve the same spatial resolutions.

For the proposed multi-path fusion unit, the computation of output y with given input x is expressed as:

$$y = U(x) = f_1(x) + f_3(f_2(f_1(x))) + f_5(f_4(f_3(f_2(f_1(x))))), \quad (1)$$

where f represents the corresponding convolution function. In the proposed multi-path fusion unit, the multiple paths from the input to the output contribute to learning complex features. Furthermore, the various residual links help to perform direct gradient propagation and achieve more efficient end-to-end learning processes.

3. Up-sampling unit:

Following the multi-path fusion unit, the spatial resolution of the output features maps is enlarged

by a deconvolution layer. The resulting features are then passed to the next refinement sub-net for further processing.

3.3. Loss for Mixed Noise

In contrast to most traditional image enhancement methods which only deal with the single type of Gaussian noise using L_2 loss, we further introduce L_1 loss for sparse impulse noise, which is closer to the enhancement of real low-light images. The final loss for mixed noise is computed as the weighted sum of L_1 and L_2 loss:

$$L = \frac{1}{2} \|x - x_{gt}\|_2^2 + \lambda \|x - x_{gt}\|_1, \quad (2)$$

where x is the estimated image, x_{gt} is the corresponding ground-truth image and λ is the balancing coefficient. In our experiments, we find that our network produces the best results when λ is set to be 0.5.

The backward process of the combined loss is as follows:

$$\frac{dL}{dx} = x - x_{gt} + \lambda \text{sign}(x - x_{gt}), \quad (3)$$

where $\text{sign}(x)$ is defined as:

$$\text{sign}(x) = \begin{cases} +1 & x_i > 0 \\ -1 & x_i < 0 \\ [-1, 1] & x_i = 0. \end{cases} \quad (4)$$

During the training process, the reconstructed image is compared against the clean version by computing this combined loss.

3.4. Network Training and Training Data Generation

To train the proposed LL-RefineNet, caffe framework [49] is used with modifications. The Adam [50] algorithm is chosen as the optimization algorithm due to its faster convergence in our method compared to the stochastic gradient descent algorithm. Parameters of the Adam algorithm are set as $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The starting learning rate is set as $\lambda = 1 \times 10^{-4}$ and then the learning rate is divided by 2 every 100 k iterations. Although we train the proposed LL-RefineNet on synthetic low-light images, we use both synthetic and real low-light images to test the performance of our model for image enhancement.

To generate the training data, we modify the clear source images to generate low-light image samples as done in LLCNN [48]. First, the values of image pixels are normalized to [0,1] as the pre-processing strategy. Then, multiple levels of darkening are added to the source images that simulate the low-contrast in the low-light environment. In the real low-light environment, the low-light images often contain various noise types including impulse and Gaussian noise. While the former methods only add Gaussian noise to simulate low-light images, we generate the training images with both impulse and Gaussian noise to improve the robustness of our model. Given the clean source image I_0 , the generation model for low-light image I_{LL} is as follows:

$$I_{LL} = I_0 * L + G + N_I, \quad (5)$$

where L is a contrast degradation factor, while G and N_I represent the matrix for added Gaussian noise G and impulse noise N_I , respectively. According to the data generation method, we first decrease the contrast of the source images by multiplying with contrast degradation level $L \in (0, 1)$, and then add the mixed Gaussian and impulse noises, with the standard deviation σ varying from 10 to 40 for Gaussian noise and the percentage of corrupted pixels varying from 5% to 20% for impulse noise. The impulse noise is generated in uniform distribution with the intensity from 0 to 255. We generate

20,200 low-light image samples and divide these images into two parts: 18,000 for training samples and 2200 for testing samples. The training samples are randomly shuffled in the training process.

4. Experiments and Results Analysis

In this section, descriptions of competitive enhancement methods are provided. Then, the performance of the proposed method is evaluated both qualitatively and quantitatively. We further compare different methods in terms of running time. The workstation on which we run our experiments is shown in Figure 3, which contains four K80 Graphics Processing Units (GPUs).



Figure 3. The workstation for our experiments.

4.1. Compared Methods

The compared methods for low-light image enhancement are described in this section.

1. Contrast-limiting adaptive histogram equalization (CLAHE)
Different from the conventional histogram equalization algorithm, the contrast-limiting adaptive histogram equalization [26] (CLAHE) adds contrast limiting to further improve the enhancement result. In this method, each neighborhood in the input image is processed by the contrast limiting strategy instead of a global process in the traditional histogram equalization algorithm. The CLAHE approach also outperforms the previous histogram equalization approach as it suppresses the over-amplification of noise.
2. Contrast-limiting adaptive histogram equalization with 3D block matching (CLAHE + BM3D)
As one of the state-of-the-art algorithms for image denoising, BM3D [32] is also applied in the low-light enhancement task. Based on the 3D array of grouped image patches, the BM3D algorithm first utilizes Wiener filter in a collaborative form and then jointly performs the denoising process for the grouped patches. In the HE + BM3D method, equalize the image contrast using the CLAHE algorithm first and then apply BM3D to denoise the resulting images and get the final result.
3. Gradient-based Total Variation
This work [51] studies gradient-based schemes for image denoising problems based on the discretized total variation (TV) minimization model with constraints. An acceleration of the well known dual approach is combined with the denoising problem with a novel monotone version of a fast iterative shrinkage/thresholding algorithm.

4. Multiresolution Bilateral Filtering

This method [52] presents a new image denoising framework based on multi-resolution bilateral filtering, which turns out to be very effective in eliminating noise in real noisy images.

5. Low-light convolutional neural network (LLCNN)

In the LLCNN [48] framework, deep residual convolutional neural network is designed to enhance the input low-light images. In the training process, darkening and Gaussian noise are added to clear source images to generate the training data. The resulting network is used to enhance the low-light images. Our approach is closely related to LLCNN, while the proposed model shows superior performance and robustness to mixed noise.

4.2. Quantitative Evaluation

In our experiments, we adopt the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and root-mean square error (RMSE) to compare different methods quantitatively:

1. Peak signal-to-noise ratio (PSNR). The PSNR metric [53] is used to quantify the extent of damage of source image containing noise and approximate human perception of the observed image. Generally, higher PSNR values correspond to better quality of enhanced images.

2. Structural similarity index (SSIM). The SSIM metric is used to obtain the perceived quality of enhanced images [54], which can measure how similar one image is to another one. While the PSNR metric is able to quantify the absolute values of errors between the enhanced image and the reference image, it may not effectively quantify the structural similarity. Different from PSNR values, the SSIM metric can help to compare the structural change based on the incorporation of both the pixel intensities and image structures.

3. Root-Mean-Square Error (RMSE). The RMSE metric is a frequently used measure of the differences between values predicted by a model or an estimator and the values observed. Generally, lower RMSE values correspond to better quality of enhanced images.

Based on these metrics, the proposed method together with other compared methods are evaluated in terms of average PSNR, SSIM and RMSE results on the test simulated images. Note that the average PSNR, SSIM and RMSE results are calculated for various levels of contrast degradation and noise intensity. The PSNR, SSIM and RMSE results are illustrated in Table 1. PSNR-S, SSIM-S and RMSE-S refer to the results on simulated images, while PSNR-R, SSIM-R and RMSE-R refer to the results on real images.

It is shown that the proposed method is significantly superior to compared methods on PSNR, SSIM and RMSE. To further demonstrate the effectiveness of our mixed noise loss, another experiment with L_2 loss is added with the same network structure and training settings. Compared to the proposed network using L_2 loss, our model with the mixed noise loss achieves better PSNR, SSIM and RMSE results, which is shown in Table 1. This can be explained by the fact that the L_1 term helps the training process to be more robust to impulse noise. Therefore, the proposed method performs better in real low-light environments with mixed noise.

Table 1. PSNR, SSIM and RMSE results for different algorithms.

| Comparative Items | PSNR-S (dB) | SSIM-S | RSME-S | PSNR-R (dB) | SSIM-R | RSME-R |
|-----------------------------|----------------|---------------|----------------|----------------|---------------|----------------|
| Low-Light | 12.1566 | 0.2443 | 19.5874 | 9.8208 | 0.2529 | 21.0653 |
| CLAHE | 20.7378 | 0.5573 | 14.3576 | 19.6647 | 0.5486 | 15.0286 |
| Bilateral Filtering | 20.8287 | 0.5531 | 14.2782 | 20.0634 | 0.5467 | 14.7386 |
| Total Variation | 20.1785 | 0.5602 | 14.7139 | 20.4581 | 0.5495 | 14.4862 |
| CLAHE + BM3D | 20.9753 | 0.5545 | 14.1687 | 20.3876 | 0.5487 | 14.5440 |
| LLCNN | 21.6134 | 0.5661 | 13.8470 | 20.5389 | 0.5380 | 14.4333 |
| Proposed + L_2 loss | 22.5876 | 0.6299 | 13.2358 | 21.4700 | 0.5954 | 13.9007 |
| Proposed + mixed noise loss | 22.7723 | 0.6549 | 13.0996 | 21.7579 | 0.6259 | 13.6957 |

4.3. Qualitative Evaluation

We show the qualitative results in Figures 4–6. First, we provide the visual performance on simulated low-light images. Figure 4 presents results with $L = 0.4, \sigma = 20, r = 5\%$, while Figure 5 presents those with $L = 0.2, \sigma = 40, r = 15\%$. The CLAHE method may generate low-quality results as a result of the fact that the pixels in the input noisy image are equalized, leading to degradation of the results. In addition, as the equalization strategy alters the noise structure, the BM3D based method cannot perform denoising in an effective manner. The multi-resolution bilateral filtering approach cannot deal with large amount of noise. The total variation based method can suppress some noise content, but it may cause over-smoothing effect to the results. Although LLCNN achieves better performance than these two algorithms, there still exists significant noise in the final results. Overall, our model achieves the best visual results compared to these methods on the simulated test images.

In contrast to tests on low-light synthetic images which can obtain accurate reference images, experiments on the real low-light images lack such reference data. Similar to the previous low-light image enhancement approaches, we use cell-phone cameras to collect corresponding clean images with lights on. The captured high-light image can serve as an approximate reference image for the evaluation of different algorithms on real low-light images. The results of all the methods on natural low-light images are shown in Figure 6. Although the image contrast can also be enhanced by the compared methods, the output images still suffer from a larger amount of noise. Furthermore, over-amplification is observed in the resulting images both locally and globally for these approaches. In contrast, our method can still achieve satisfactory results on the real low-light images, which demonstrates that our model trained on simulated data can generalize well to real applications.

In summary, the proposed LL-RefineNet model outperforms the benchmark methods qualitatively on both synthetic images and real low-light images.

4.4. Running Time

We further compare the proposed method with others on the running time. We find it difficult to accelerate previous non-deep learning approaches using hardware-based parallelization strategy. Therefore, they tend to take longer computing time on the test images. In contrast, our method can take advantage of the non-iterative forward pass which is performed on GPU. Although the LLCNN method can also benefit from GPU computing, the structure of residual network still requires more computational resource than the proposed method. The average running time for each algorithm on 256×256 test image is illustrated in Table 2. The proposed method is much faster than the benchmark algorithms.

Table 2. Runtime of different algorithms on our test images.

| Methods | Average Run Time (Seconds) |
|---------------------|----------------------------|
| CLAHE | 8.64 (Intel i5-4200U CPU) |
| Bilateral Filtering | 3.22 (Intel i5-4200U CPU) |
| Total Variation | 15.17 (Intel i5-4200U CPU) |
| CLAHE + BM3D | 10.13 (Intel i5-4200U CPU) |
| LLCNN | 1.26 (K80 GPU) |
| Proposed method | 0.85 (K80 GPU) |

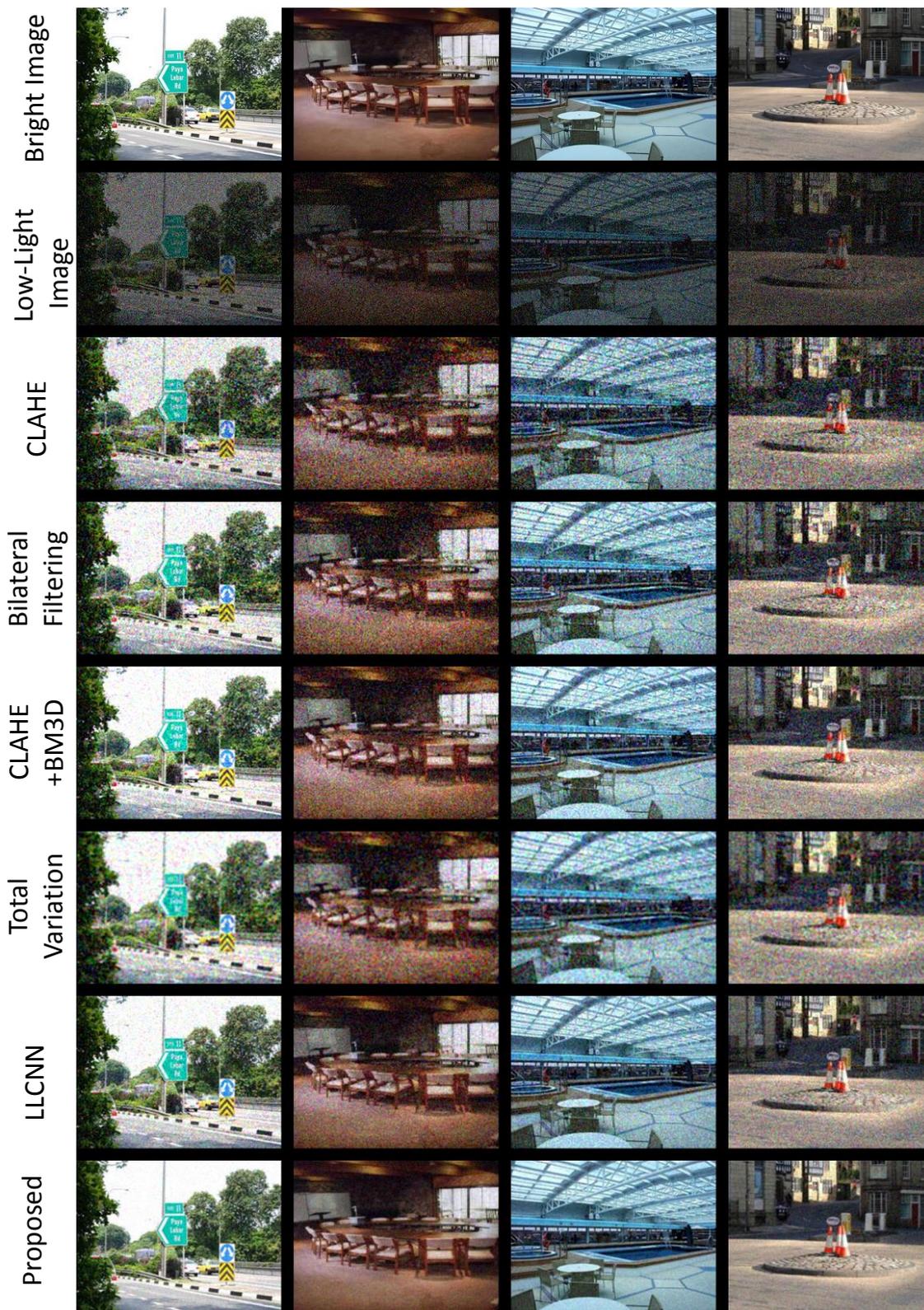


Figure 4. Performance of different algorithms on synthesized low-light images ($L = 0.4$, $\sigma = 20$, $r = 5\%$). From top to bottom: bright reference images, low-light images, results of CLAHE, Bilateral Filtering, CLAHE + BM3D, Total Variation, LLCNN and the proposed method.

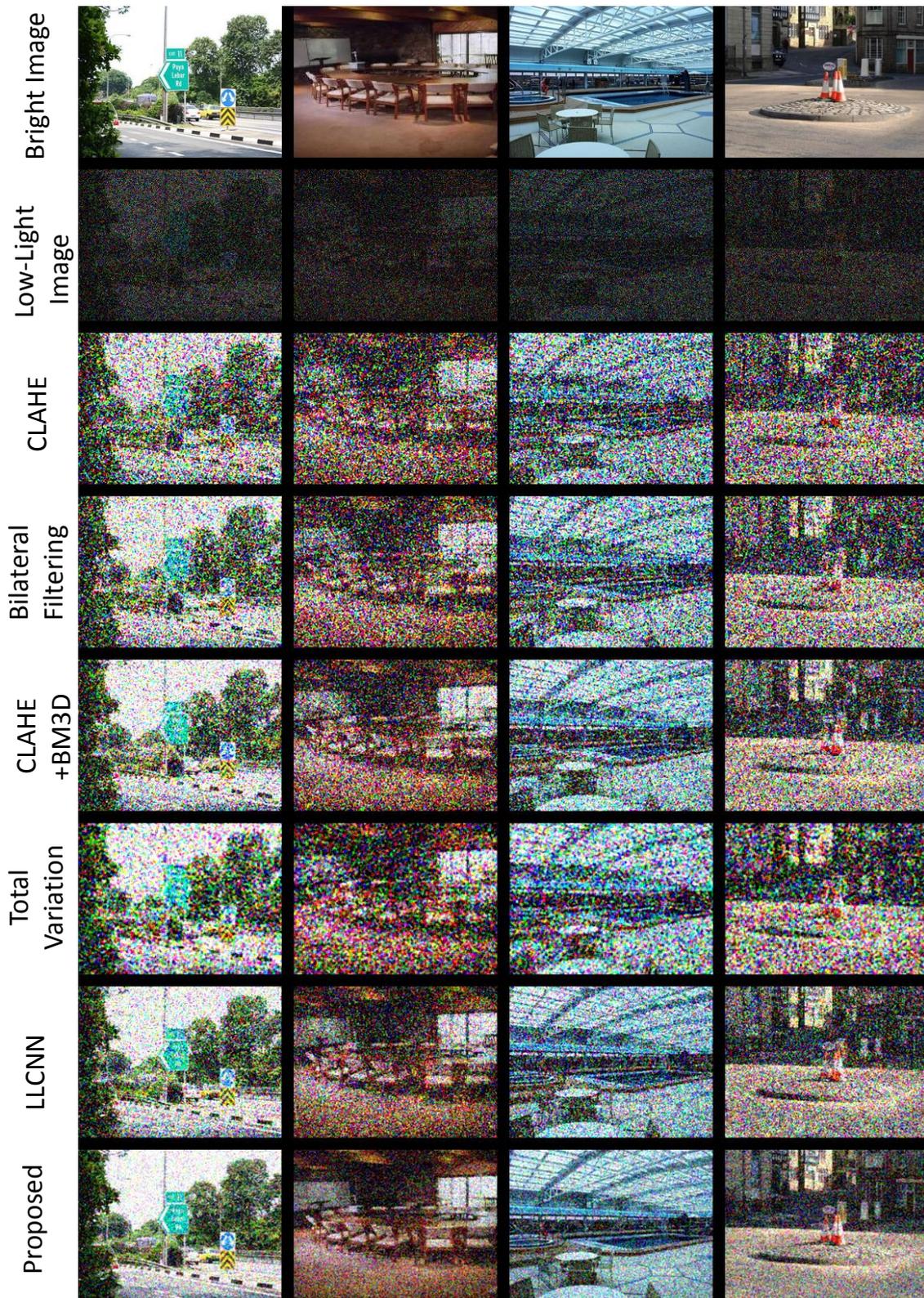


Figure 5. Performance of different algorithms on synthesized low-light images ($L = 0.2$, $\sigma = 40$, $r = 15\%$). From top to bottom: bright reference images, low-light images, results of CLAHE, Bilateral Filtering, CLAHE + BM3D, Total Variation, LLCNN and the proposed method.

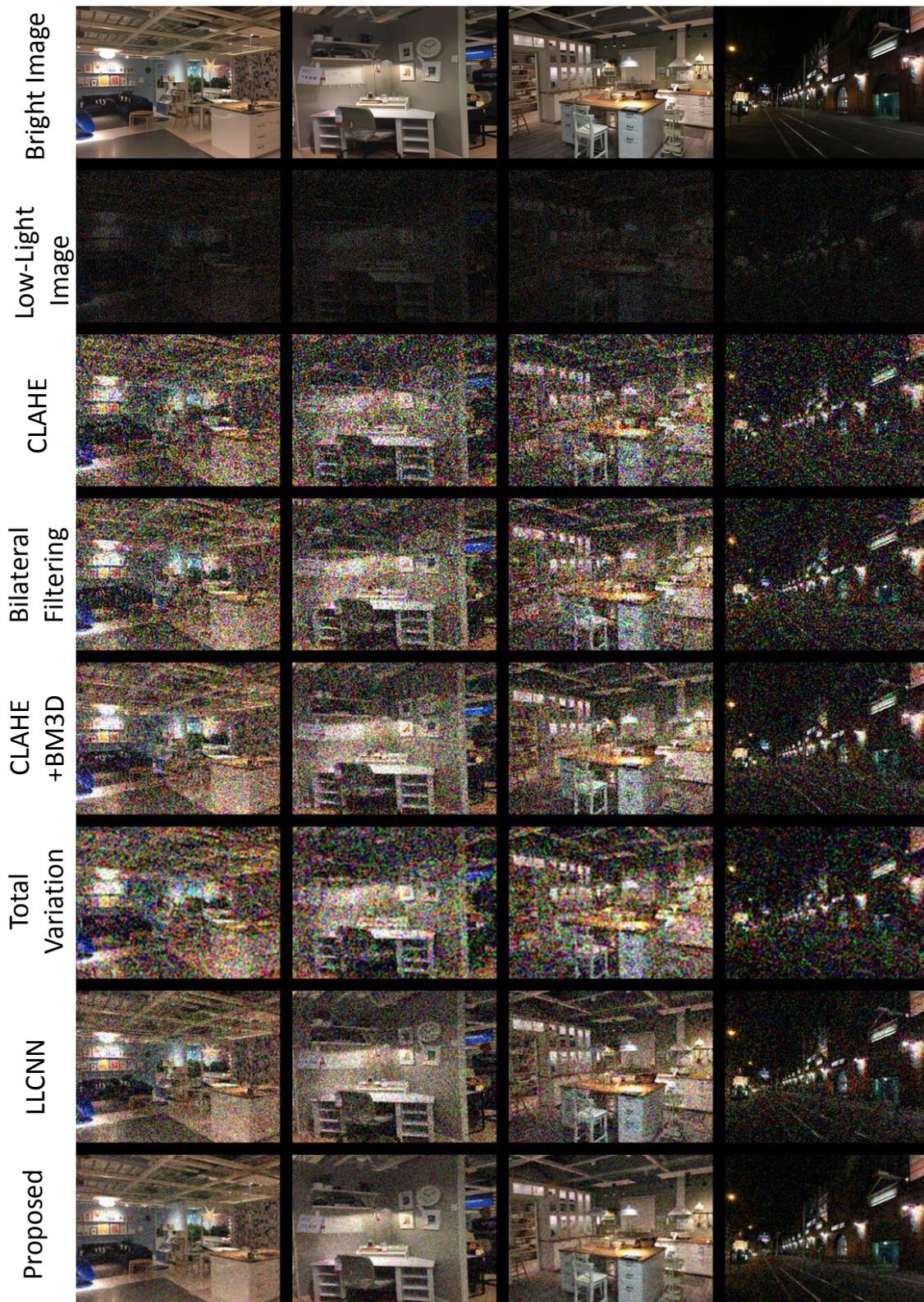


Figure 6. Performance of different algorithms on natural low-light images. From top to bottom: bright reference images, low-light images, results of CLAHE, Bilateral Filtering, CLAHE + BM3D, Total Variation, LLCNN and the proposed method.

5. Conclusions

In this paper, a deep refinement network, LL-RefineNet, is designed to learn from the synthetical low-light images and perform enhancement for both synthetical and real low-light images. The proposed network utilizes all the features and information from the encoding part and produces the high-resolution enhancement result in the refinement pass. In our model, global features generated by deeper layers are gradually fused and refined using local features from earlier convolutions. A novel training loss is further designed to handle mixed noise. It is shown in the experimental results that the proposed approach outperforms the benchmark methods both qualitatively and quantitatively with high processing speed. In the future, we will continue designing novel network architectures for the task of low-light image enhancement, and improve both the computing speed and network performance. Robust loss function for mixed noise will be further explored. Our models will be applied as pre-processing strategies for multiple tasks including semantic segmentation, object detection and object recognition.

Author Contributions: Conceptualization, L.J., Y.J. and S.H.; Methodology, L.J.; Software, L.J.; Validation, S.H.; Formal Analysis, L.J. and Y.J.; Resources, W.X.; Data Curation, L.J.; Writing—Original Draft Preparation, L.J. and Y.J.; Writing—Review and Editing, S.H. and W.X.; Visualization, L.J. and B.G.; Supervision, W.X.; Project Administration, B.G.; Funding Acquisition, S.H. and W.X.

Funding: This research was funded by the National Natural Science Foundation of China under Grant Nos. 61872446, 71690233, and 71331008.

Acknowledgments: We are pleased to thank the Editor and the Referees for their useful suggestions. This work was supported by the National Natural Science Foundation of China under Grant Nos. 61872446, 71690233 and 71331008.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
2. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *6*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
3. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 8–10 June 2015; pp. 3431–3440.
4. Rezaee, H.; Abdollahi, F. A decentralized cooperative control scheme with obstacle avoidance for a team of mobile robots. *IEEE Trans. Ind. Electron.* **2014**, *61*, 347–354. [[CrossRef](#)]
5. Luo, R.C.; Lai, C.C. Multisensor fusion-based concurrent environment mapping and moving object detection for intelligent service robotics. *IEEE Trans. Ind. Electron.* **2014**, *61*, 4043–4051. [[CrossRef](#)]
6. Pizer, S.M.; Amburn, E.P.; Austin, J.D.; Cromartie, R.; Geselowitz, A.; Greer, T.; ter Haar Romeny, B.; Zimmerman, J.B.; Zuiderveld, K. Adaptive histogram equalization and its variations. *Comput. Vis. Gr. Image Process.* **1987**, *39*, 355–368. [[CrossRef](#)]
7. Yamasaki, A.; Takauji, H.; Kaneko, S.I.; Kanade, T.; Ohki, H. Denighting: Enhancement of nighttime images for a surveillance camera. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Tampa, FL, USA, 8–11 December 2008; pp. 1–4.
8. Dong, X.; Wang, G.; Pang, Y.; Li, W.; Wen, J.; Meng, W.; Lu, Y. Fast efficient algorithm for enhancement of low lighting video. In Proceedings of the 2011 IEEE International Conference on Multimedia and Expo, Barcelona, Spain, 11–15 July 2011.
9. Połap, D.; Winnicka, A.; Serwata, K.; Kęsik, K.; Woźniak, M. An Intelligent System for Monitoring Skin Diseases. *Sensors* **2018**, *18*, 2552. [[CrossRef](#)] [[PubMed](#)]
10. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)] [[PubMed](#)]

11. Glowacz, A.; Glowacz, Z. Diagnosis of the three-phase induction motor using thermal imaging. *Infrared Phys. Technol.* **2016**, *81*, 7–16. [[CrossRef](#)]
12. Wozniak, M.; Polap, D. Adaptive neuro-heuristic hybrid model for fruit peel defects detection. *Neural Netw.* **2018**, *98*, 16–33. [[CrossRef](#)] [[PubMed](#)]
13. Wozniak, M.; Polap, D.; Capizzi, G.; Sciuto, G.L.; Kosmider, L.; Frankiewicz, K. Small lung nodules detection based on local variance analysis and probabilistic neural network. *Comput. Methods Progr. Biomed.* **2018**, *161*, 173–180. [[CrossRef](#)] [[PubMed](#)]
14. Woźniak, M.; Połap, D. Bio-inspired methods modeled for respiratory disease detection from medical images. *Swarm Evol. Comput.* **2018**, *41*, 69–96. [[CrossRef](#)]
15. Wozniak, M.; Polap, D.; Komider, L.; Clapa, T. Automated fluorescence microscopy image analysis of *Pseudomonas aeruginosa* bacteria in alive and dead stadium. *Eng. Appl. Artif. Intell.* **2018**, *67*, 100–110. [[CrossRef](#)]
16. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
17. Ganovska, B.; Molitoris, M.; Hosovsky, A.; Pitel, J.; Krolczyk, J.B.; Ruggiero, A.; Krolczyk, G.M.; Hloch, S. Design of the model for the on-line control of the AWJ technology based on neural networks. *Indian J. Eng. Mater. Sci.* **2016**, *23*, 279–287.
18. Tadeusiewicz, R. Neural networks in mining sciences-general overview and some representative examples. *Arch. Min. Sci.* **2015**, *60*, 971–984. [[CrossRef](#)]
19. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131. [[CrossRef](#)] [[PubMed](#)]
20. Jia, F.; Lei, Y.; Lu, N.; Xing, S. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. *Mech. Syst. Signal Process.* **2018**, *110*, 349–367. [[CrossRef](#)]
21. Gajewski, J.; Vališ, D. The determination of combustion engine condition and reliability using oil analysis by MLP and RBF neural networks. *Tribol. Int.* **2017**, *115*, 557–572. [[CrossRef](#)]
22. Li, C.; Min, X.; Sun, S.; Lin, W.; Tang, Z. Deepgait: A learning deep convolutional representation for view-invariant gait recognition using joint bayesian. *Appl. Sci.* **2017**, *7*, 210. [[CrossRef](#)]
23. Ren, Y.; Zhu, C.; Xiao, S. Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. *Appl. Sci.* **2018**, *8*, 813. [[CrossRef](#)]
24. Pan, Z.; Jiang, G.; Jiang, H.; Yu, M.; Chen, F.; Zhang, Q. Stereoscopic Image Super-Resolution Method with View Incorporation and Convolutional Neural Networks. *Appl. Sci.* **2017**, *7*, 526. [[CrossRef](#)]
25. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [[CrossRef](#)] [[PubMed](#)]
26. Pisano, E.D.; Zong, S.; Hemminger, B.M.; DeLuca, M.; Johnston, R.E.; Muller, K.; Braeuning, M.P.; Pizer, S.M. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *J. Digit. Imaging* **1998**, *11*, 193. [[CrossRef](#)] [[PubMed](#)]
27. Kim, Y.T. Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Trans. Consum. Electron.* **1997**, *43*, 1–8.
28. Abdullah-Al-Wadud, M.; Kabir, M.H.; Dewan, M.A.A.; Chae, O. A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **2007**, *53*, 593–600. [[CrossRef](#)]
29. Kaur, M.; Kaur, J.; Kaur, J. Survey of contrast enhancement techniques based on histogram equalization. *Int. J. Adv. Comput. Sci. Appl.* **2011**, *2*, 84–90. [[CrossRef](#)]
30. Wu, X. A linear programming approach for optimal contrast-tone mapping. *IEEE Trans. Image Process.* **2011**, *20*, 1262–1272. [[PubMed](#)]
31. Chan, R.H.; Ho, C.W.; Nikolova, M. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Trans. Image Process.* **2005**, *14*, 1479–1485. [[CrossRef](#)] [[PubMed](#)]
32. Dabov, K.; Foi, A.; Katkovnik, V. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [[CrossRef](#)] [[PubMed](#)]

33. Elad, M.; Aharon, M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **2006**, *15*, 3736–3745. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, X.; Shen, P.; Luo, L.; Zhang, L.; Song, J. Enhancement and noise reduction of very low light level images. In Proceedings of the 21st International Conference on Pattern Recognition, Tsukuba, Japan, 11–15 November 2012; pp. 2034–2037.
35. Fotiadou, K.; Tsagkatakis, G.; Tsakalides, P. Low light image enhancement via sparse representations. In Proceedings of the 2014 International Conference on Image Analysis and Recognition, Vilamoura, Algarve, Portugal, 22–24 October 2014; pp. 84–93.
36. Hu, Z.; Cho, S.; Wang, J.; Yang, M.H. Deblurring low-light images with light streaks. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 3382–3389.
37. Guo, X.; Li, Y.; Ling, H. LIME: Low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process.* **2017**, *26*, 982–993. [[CrossRef](#)] [[PubMed](#)]
38. Jung, C.; Yang, Q.; Sun, T.; Fu, Q.; Song, H. Low light image enhancement with dual-tree complex wavelet transform. *J. Vis. Commun. Image Represent.* **2017**, *42*, 28–36. [[CrossRef](#)]
39. Mahendran, A.; Vedaldi, A. Understanding deep image representations by inverting them. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5188–5196.
40. Chan, T.H.; Jia, K.; Gao, S.; Lu, J.; Zeng, Z.; Ma, Y. PCANet: A simple deep learning baseline for image classification? *IEEE Trans. Image Process.* **2015**, *24*, 5017–5032. [[CrossRef](#)] [[PubMed](#)]
41. Lin, K.; Yang, H.F.; Hsiao, J.H.; Chen, C.S. Deep Learning of Binary Hash Codes for Fast Image Retrieval. In Proceedings of the IEEE International Conference on Vision and Pattern Recognition, Deep Vision Workshop, Boston, MA, USA, 7–12 June 2015; pp. 27–35.
42. Tai, Y.; Yang, J.; Liu, X. Image super-resolution via deep recursive residual network. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
43. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.A. Extracting and composing robust features with denoising autoencoders. In Proceedings of the 25th international conference on Machine learning, Helsinki, Finland, 5–9 June 2008; pp. 1096–1103.
44. Agostinelli, F.; Anderson, M.R.; Lee, H. Adaptive multi-column deep neural networks with application to robust image denoising. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–10 December 2013; pp. 1493–1501.
45. Burger, H.C.; Schuler, C.J.; Harmeling, S. Image denoising: Can plain neural networks compete with BM3D? In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2392–2399.
46. Lin, G.; Milan, A.; Shen, C.; Reid, I.D. RefineNet: Multi-path Refinement Networks for High-Resolution Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, HI, USA, 21–26 July 2017; pp. 5168–5177.
47. Yang, W.; Tan, R.T.; Feng, J.; Liu, J.; Guo, Z.; Yan, S. Deep joint rain detection and removal from a single image. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1357–1366.
48. Tao, L.; Zhu, C.; Xiang, G.; Li, Y.; Jia, H.; Xie, X. LLCNN: A convolutional neural network for low-light image enhancement. In Proceedings of the 2017 IEEE Conference on Visual Communications and Image Processing, Saint Petersburg, FL, USA, 10–13 December 2017; pp. 1–4.
49. Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM International Conference on Multimedia, Orlando, FL, USA, 3–7 November 2014; pp. 675–678.
50. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
51. Beck, A.; Teboulle, M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.* **2009**, *18*, 2419–2434. [[CrossRef](#)] [[PubMed](#)]
52. Zhang, M.; Gunturk, B.K. Multiresolution bilateral filtering for image denoising. *IEEE Trans. Image Process.* **2008**, *17*, 2324–2333. [[CrossRef](#)] [[PubMed](#)]

53. Santoso, A.J.; Nugroho, L.E.; Suparta, G.B.; Hidayat, R. Compression ratio and peak signal to noise ratio in grayscale image compression using wavelet. *Int. J. Comput. Sci. Technol.* **2011**, *2*, 7–11.
54. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).