Semantic Pre-supplement for Exposure Correction

Zhen Zou
University of Science and Technology of China
He Fei, China
zouzhen@mail.ustc.edu.cn

Wei Yu
University of Science and Technology of China
He Fei, China
patrick914y@mail.ustc.edu.cn

Jie Huang
University of Science and Technology of China
He Fei, China
hj0117@mail.ustc.edu.cn

Feng Zhao
University of Science and Technology of China
He Fei, China
fzhao956@ustc.edu.cn

Abstract

Exposure correction tasks are dedicated to recovering the brightness and structural information of overexposed or underexposed images. The recovery difficulty of areas with different exposure levels is different, as severely exposed areas are more difficult to recover due to severe structural information loss than commonly exposed areas. However, existing methods focus on the simultaneous recovery of global brightness and structure, ignoring that the recovery difficulty varies between areas. To address this issue, we propose a novel exposure correction strategy named "Inpainting Assisted Exposure Correction"(IAEC), which pre-performs image structure repair on severely exposed areas to guide the exposure correction process. This method is based on the observation that the contextual semantic information contained in the image structure can effectively help the overall image recovery, and the lack of contextual semantic information in severely incorrectly exposed areas is very severe. The pre-performed structural repair by the inpainting model can well supplement the insufficient contextual semantic information caused by severe exposure. Therefore, we use an inpainting model to perform pre-structure repair on severely exposed areas to obtain supplementary contextual semantic information and then align the structure-repaired image with the improperly exposed input at the feature level. Extensive experiments demonstrate that our method gets superior results than the state-of-the-art methods and has the potential to be applied to other tasks with similar context loss problems.

1. Introduction

Exposure in photography refers to the camera's management of light during the shooting process[1]. Since various scenes have distinct light conditions, the appropriate exposure varies accordingly. In areas with excessively improper exposure, seriously improper brightness distribution destroys the structural features of the image and leads to the loss of context semantics[25]. The purpose of exposure correction is to correct images taken under non-ideal lighting conditions, bringing them to a standard exposure level to achieve pleasing visual effects or to facilitate subsequent advanced visual tasks.

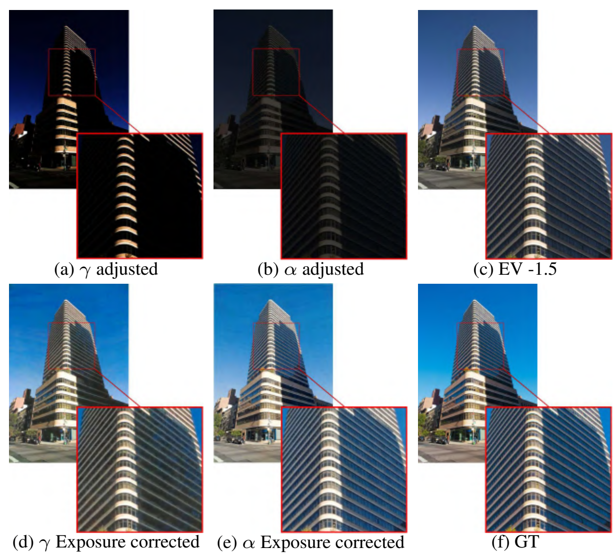


Figure 1. Images rich in structural information are easier to correct. (a) the γ adjusted image has fewer structural details than (b) the α adjusted image, which has the same average brightness as the former. (c) EV -1.5 original underexposure image. Compared to (d) the γ adjusted image after exposure correction, (e) the α adjusted image after exposure correction has better structure reconstruction (such as windows of the building).

Leveraging the powerful learning capabilities of neural

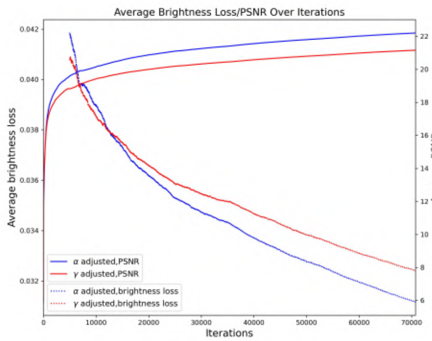


Figure 2. As the training progresses, the PSNR of α adjusted images grows faster and has a higher upper limit than that of γ adjusted images. At the same time, the average absolute brightness loss between α adjusted images and the original images also decreases faster and has a lower limit than that of γ adjusted images.

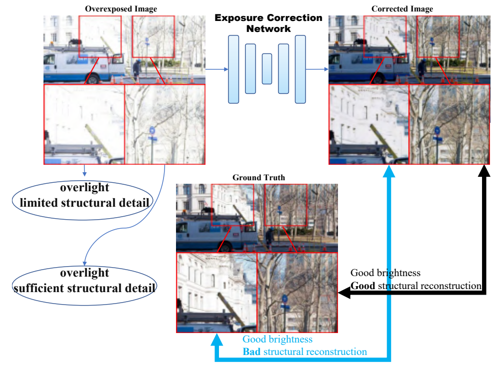


Figure 3. Illustration of the standard exposure correction process. Severe overexposure in some areas greatly damages the structural information, making structural reconstruction of these areas difficult. In some areas, even the brightness situation is very bad, but sufficient structural information makes recovery not difficult.

networks and well-designed network architectures, recent methods have demonstrated commendable performance in most scenarios[1, 7, 8, 19]. However, their effectiveness often wanes when faced with severely exposed areas where structural details are limited. As illustrated in Fig. 3, different regions retain unequal structural details, which brings uneven recovery difficulty. However, existing methods treat the entire image equally, ignoring this unevenness. Inappropriate exposure in most areas only changes the statistical distribution of image brightness, while in severely improperly exposed areas the structural features of the image are also erased, resulting in a lack of semantics.

In high-level computer vision tasks, the semantics within the image structure allow the model to understand the image content and achieve good results. Sufficient structural information can also help image restoration in low-level tasks, and exposure correction is no exception. We conducted an experiment and confirmed this. We extracted 1050 images from the MSEC test set with a relative EV of -1.5 and gamma-adjusted them by varying the scaling factor (α) and gamma parameter (γ) in the formula $I_{\text{gamma}} = \alpha \cdot I^{\gamma}$. Both sets of images had their average brightness reduced to one-fourth of the original underexposed brightness, corresponding to $\alpha = \frac{1}{4}$ with an adjustment in γ. Images in Fig. 1 reveal that the image adjusted with α better preserves dark details compared to the image adjusted with γ.

Subsequently, exposure correction was performed on the two sets. During the training process, we tracked the PSNR and absolute brightness loss, which calculates the overall brightness difference between the image and its GT, rather than on an element-wise basis. Fig. 2 illustrates that the set adjusted with α experiences a more rapid decrease in average absolute brightness loss compared to the set adjusted with γ. This trend is reversed in the PSNR values, suggesting that images with more structural details are easier to restore because they retain more contextual semantic information. The correction results are shown in Fig.1.

In exposure correction tasks, it is inevitable to lose structural information due to severely improper exposure in specific areas. The severe lack of structural information requires the network to possess not only exposure correction capabilities but also reasoning capabilities for unknown regions. The image inpainting task is specifically designed to reconstruct missing portions of an image, aligning with our requirements. We thus cast the exposure correction of severely exposed areas as an image repair problem, treating these areas as repair targets. We leverage the reconstructed structural information from these regions to guide the subsequent exposure correction process.

Building upon the above observation and analysis, we aim to enhance the performance of exposure correction networks by rich semantic information, which comes from preperformed structural repair on severely improperly exposed areas by an inpainting module. In light of this, we introduce our Inpainting Assisted Exposure Correction (IAEC) method, comprising an inpainting module, a fundamental exposure correction network, and an auxiliary training regularization term. The inpainting module repairs severely exposed areas of the image, and brings sufficient semantic information. The regularization term uses the distillation method to implicitly embed the function of the inpainting module into the exposure correction network, which allows the method to completely abandon the inpainting part during the inference process. Extensive experiments conducted various datasets demonstrate that our IAEC method consistently outperforms state-of-the-art methods.

The main contributions of this work are summarized as:

1. This study explores the impact of semantic deficiency on exposure correction resulting from structural loss due to

severe exposure. The extreme lack of semantic information in severely exposed areas makes the recovery of both structure and brightness very difficult.
2. We introduced the Inpainting Assisted Exposure Correction (IAEC) method, leveraging semantic pre-repair. It employs an inpainting module to perform structural repair of severely improper exposure. The module output acts as a guide to generate an additional regular term, assisting the model optimization. It does not introduce any additional overhead in the inference phase.
3. Experiments on ME, SICE, and LCDP datasets show that our method surpasses state-of-the-art methods in performance while maintaining lightweight.

2. Related Work

2.1. Learning-based Image Inpainting

Pathak et al. [20] first try to fix holes with the context semantics of the image, using a simple Encoder-Decoder architecture network and training it in an adversarial way. Compared with traditional methods, this is the first time that high-level semantic understanding is introduced into the inpainting task, allowing it to generate more reasonable content. [33, 36, 39] focus on contextual attention to enable the network to fully grasp high-level semantics. Different convolution [15, 37] strategies have also been designed to better extract image information. Xie et al. [29] introduce learnable bidirectional attention maps with partial convolution [15] , making it easier to fill irregular holes.

In addition to the above, foreground contours [30], object edges [18], image structures [11, 16, 22], etc. are also widely used as intermediate clues to improve image restoration effects. Liu et al. [16] fuse the texture features (shallow layers of the encoder) and structure features (deep layers of the encoder) via feature equalization in order to restore both structure and texture. Liao et al. [14]introduce the use of semantic segmentation maps to guide the inpainting operation. However, it's notable that an additional semantic segmentation step is required during the training process. Xiong et al. [30] propose using foreground contours as image structures rather than object edges, a new perspective.

Image inpainting is also used as a tool to assist in the execution of other tasks [13, 21, 38]. Zavrtanik et al. [38] cast anomaly detection as a reconstruction-by-inpainting problem. They generate to-be-inpainted maps by randomly masking images with disjoint but complete holes and perform inpainting on them, recombining the inpainted parts from all images into a single reconstructed image to evaluate an anomaly map. Li et al. [13] find that pre-trained shadow removal networks on the image inpainting dataset can achieve high restoration quality with only a simple encoder-decoder network, which inspires us a lot.

2.2. Deep Image Exposure Correction

In recent times, deep learning-based methods have demonstrated remarkable performance, leveraging the potent representation capabilities inherent in deep neural networks. Existing methods fall into two main groups: those based on physical priors and those focusing on learning image-to-image mapping.

Certain methods endeavor to utilize physical priors based on the Retinex theory for data-driven image decomposition [6, 27, 41]. RetinexNet[27, 35] introduces multistage subnetworks that decompose images into illumination and reflection components, subsequently performing adjustments for illumination and reflection construction. Wang et al. [24] propose estimating illuminance maps at a lower resolution and then enhancing them through an adaptable bilateral grid interpolation process. Wu et al. [28] employ deep neural networks to unfold the optimization process of Retinex decomposition.

Other methods learn image-to-image mappings without explicit consideration of physics prior. Various techniques have been developed to improve model performance, such as Laplacian pyramid[1, 12], Fourier transformation[7]. Afifi et al. [1] design a coarse-to-fine multi-scale network based on the Laplacian pyramid decomposition. FECNet[7] introduces a deep Fourier-based network for interactions in the spatial and frequency domains. Huang et al. [9] introduce the ERL framework that establishes connections between the optimization processes of samples by learning the sampling relationships within the batch dimension.

Some methods consider modeling image structures, such as edges, as priors to guide restoration[31, 32, 42]. In contrast, our approach, for the first time, places emphasis on severely improper exposure regions where structural information is lost.

3. Method

Underexposed/Overexposed images suffer from severe structural detail loss in extreme exposure regions. Our experiment reveals that structural detail loss not only complicates the reconstruction of the region but also adversely impacts overall brightness recovery. Most existing methods perform the restoration of the entire image simultaneously. while the semantic restoration of the structural missing region in advance is often overlooked. In reality, even though these semantic repairs may not precisely align with the original ground truth, they have a positive effect on the overall reconstruction of the image.

3.1. Overview

In this section, we introduce the proposed Inpainting Assisted Exposure Correction framework. An overview of our framework can be viewed in Fig. 4. With the original in-

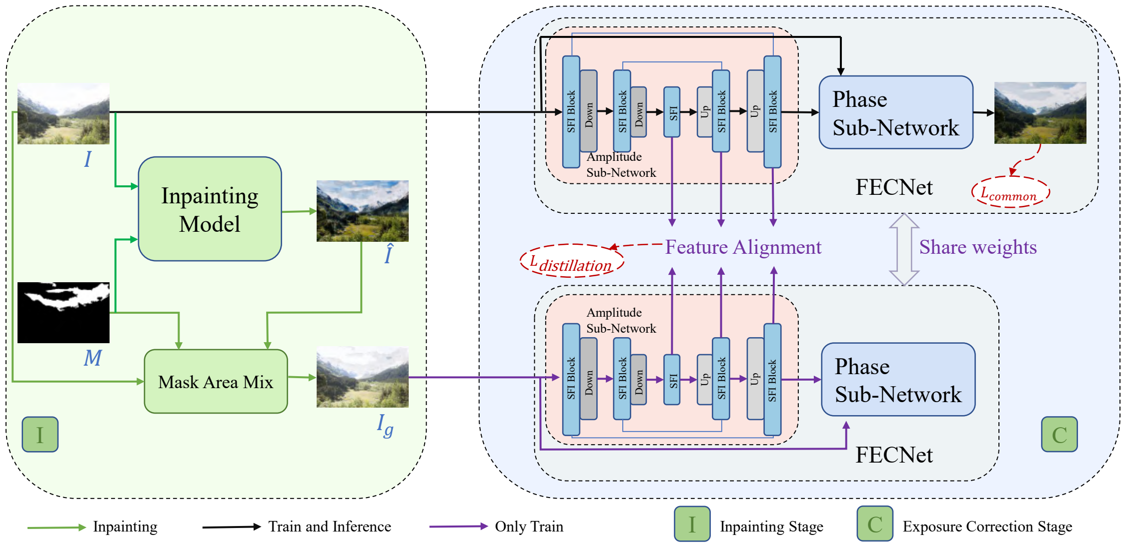


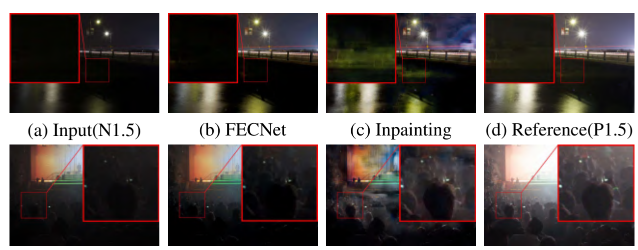
Figure 4. The illustration of IAEC, which uses pre-performed context inpainting to assist the base model in correcting improperly exposed images. Specifically, the output of the inpainting model is mixed with the original input in the mask area to generate a guidance map. The original input and guidance are aligned at the feature level so that the model receives the contextual structural information contained in the guidance map. In the inference phase, only the base model participates, so there will be no additional operations introduced.

tention of using pre-structurally repaired improper exposed images to guide the exposure correction process, our implementation method is very clear. First, we train an inpainting network using the exposure dataset with masks (indicating areas requiring repair). The output of the inpainting network is then mixed with the original input in the masked area to generate a guidance image. Finally, the guidance image and the original input are respectively fed into the exposure correction network with shared parameters and the intermediate features are aligned to obtain an additional regularization term added to the loss function. During the inference phase, we use the exposure correction network without introducing any additional operations. Detailed aspects of the modules will be expounded upon in the ensuing sections.

3.2. Image Inpainting Module

Unlike pure image inpainting tasks, inpainting exposed images faces two main challenges, determining the areas requiring inpainted and the scale of exposure dataset is not large enough for image inpainting.

To address the first challenge, considering an underexposed image, areas with missing structures typically have lower initial brightness in the original ground truth. Underexposure during the shooting process further decreases the brightness, resulting in a structure-missing block. Therefore, we identify areas in the underexposed image where brightness is less than 20% of the whole image brightness as potential structure-missing blocks. Areas with excessively small sizes (less than two-thousandth of the original image area) are discarded to prevent an abundance of small patches. Masks generated by the above process and the underexposed images form the image pairs used in the inpainting task.



(a) Input(N1.5) (b) FECNet (c) Inpainting (d) Reference(P1.5)
(e) Input(N1.5) (f) FECNet (g) Inpainting (h) Reference(P1.5)

Figure 5. Display of the effects of the inpainting module. The second and third columns are the results of using the correction network and the inpainting module separately.

For the second point, inspired by [13], we employ a two-step training strategy. Initially, the network is trained on an inpainting dataset, followed by fine-tuning on our exposure-inpainting dataset. This approach equips the model with the ability to reason about content missing areas in improperly exposed images.

To achieve effective structural restoration, we leverage a straightforward encoder-decoder structure with a fusion block proposed in [13] for image inpainting. The specific structure of the module will be shown in the supplementary material. As depicted in Fig. 5, we illustrate the structural reconstruction capability of the inpainting module in regions with severe darkness, using underexposed images as examples. For clarity, we utilize the corresponding overexposed image as a reference, which contains richer structural details in dark areas. In comparison to the exposure correction network, the inpainting module excels in reconstructing dark areas where structural details are significantly lost. However, the reasoning ability of the inpainting module also causes it to generate more artifacts, it's inevitable. We address this issue by leveraging mask area mixing and the refining capability of subsequent networks

3.3. Training with Inpainting Assistance

We initially train the inpainting module on the Places2 dataset[40] and subsequently fine-tune it on our exposure-inpainting dataset. Then the inpainted images given by the inpainting module are mixed with the original inputs to generate guidance images. The image reconstructed by the inpainting model has richer structural information than the original input, which can assist the subsequent exposure correction model. However, considering that pixels in unmasked areas of the image may not align with our desired outcome, introducing pixel values in these areas may result in unexpected artifacts. To mitigate this, we mix the reconstructed image and the original input only in the mask area(the Mask Area Mix part in Fig. 4):

$$I_g = I \odot (1 - M) + (\lambda I + (1 - \lambda)\hat{I}) \odot M, \quad (1)$$

where λ is a mix coefficient and $\odot$ donates element-wise multiplication operation and $\hat{I}$ donates the output of the inpainting model. Given that GT also has a better structure than input, we do not introduce GT because it would compromise the network's original ability to correct. After all, when the input and target are both GT, the network tends to be equivalent to self-mapping.

We adopt the base model for subsequent exposure correction from FECNet [8] (right part of Fig. 4). FECNet's operation of restoring the brightness component (Amplitude Sub-Network) first coincides with our perspective that the semantic information brought by structural repair can greatly guide image brightness restoration. The Spatial-Frequency Interaction (SFI) block in FECNet interactively processes the local spatial features and the global frequency information to encourage complementary learning, which comes in both phase and amplitude forms. The phase sub-network and amplitude sub-network use corresponding forms of SFI blocks respectively. We'll show the details of FECNet in the supplement material.

Consider how to use the guidance image to assist our exposure correction process. We propose the introduction of an additional regularization term. Specifically, the sample and guidance image is fed into two parameter-shared exposure correction models, and brightness-related features are aligned to obtain a distillation loss, expressed as :

$$L_d = \|F_g - F\|_1, \quad (2)$$

where F_g and F donate the feature extracted from the exposure process of guidance images and original inputs, respectively.

The total loss function for the training process is the combination of the conventional loss and our distillation loss :

$$L = L_c + \alpha L_d, \quad (3)$$

where α is a weight factor and we set it to 5 empirically. The impact of this factor will be discussed in the following ablation study. L_c represents the common loss used in the exposure correction methods, which refers to the L1 loss of the output and the ground truth in the spatial and frequency domain here.

4. Experiments and Analysis

4.1. Experimental Settings

Datasets. We evaluate our proposed method on two representative multiple exposure datasets, including the ME dataset proposed in [1] and the SICE dataset proposed in [3]. The ME dataset contains five exposure levels for each scene, including 17,675 images for training, 750 images for validation, and 5,905 images for testing, a total of 24,330 images. We use Expert C [1] as ground truth. For the SICE dataset, following [7], we use the middle exposure subset as the ground truth, and the corresponding second and last-second exposure subsets are set as underexposed and overexposed images, respectively. We adopt 1000 images for training, 24 images for validation, and 60 images for testing.

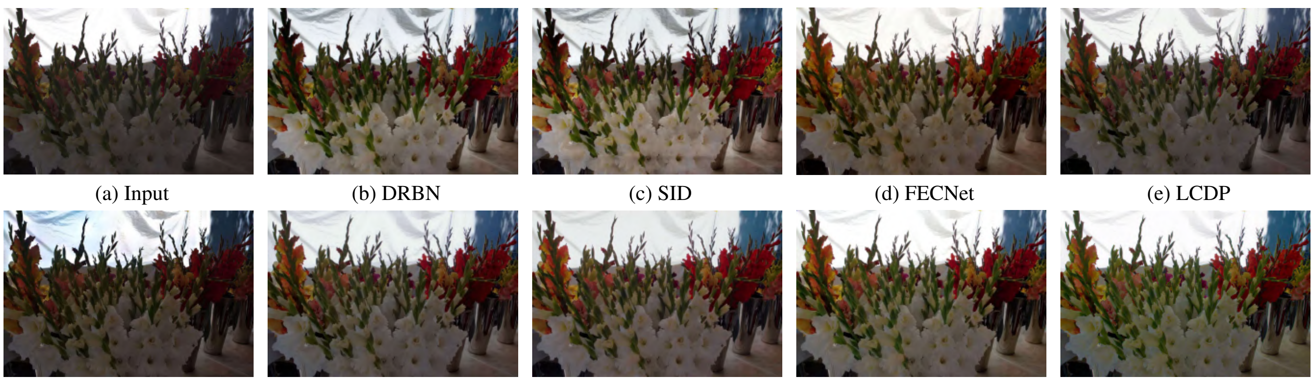
Implementation Details. The implementation of our proposed method is based on PyTorch framework with one NVIDIA 3090 GPU.

For the inpainting part, following [13] we first pretrain the model for 450,000 iterations with batch size 8 on the Places2 dataset, followed by fine-tuning 250,000 iterations under the same batch size on our exposure-inpainting dataset. The input images are resized to 256 × 256. We use Adam as the optimizer to optimize the network with a learning rate of 0.00005.

For the exposure correction part, it's important to note that since the location of the inpainting area (mask area) is not fixed for each image, we cannot obtain the training set by cropping patches from the training images (randomly cropped areas may not contain any inpainting information). We resize input images to 384 × 384 with batch size of 4 and guidance images provided by the inpainting model are also resized to the same size. For the ME and SICE datasets, the total number of iterations is set as 353,600(80 epochs) and

Table 1. Quantitative results on ME [1] and SICE [3] testing set in terms of PSNR/SSIM. **The best results are highlighted in bold.** The second-best results are highlighted with underline. #Param denotes the number of parameters.

Method	ME						SICE						#Param
	Under		Over		Average		Under		Over		Average		
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
CLAHE [23]	16.77	0.6211	14.45	0.5842	15.38	0.5990	12.69	0.5037	10.21	0.4847	11.45	0.4942	-
RetinexNet [26]	12.13	0.6209	10.47	0.5953	11.14	0.6048	12.94	0.5171	12.87	0.5252	12.90	0.5212	0.84M
Zero-DCE [5]	14.55	0.5887	10.40	0.5142	12.06	0.5441	16.92	0.6330	7.11	0.4292	12.02	0.5311	0.079M
DPED [10]	13.14	0.5812	20.06	0.6826	15.91	0.6219	16.83	0.6133	7.99	0.4300	12.41	0.5217	0.39M
DRBN [34]	19.74	0.8290	19.37	0.8321	19.52	0.8309	17.96	0.6767	17.33	0.6828	17.65	0.6798	0.53M
SID [4]	19.37	0.8103	18.83	0.8055	19.04	0.8074	19.51	0.6635	16.79	0.6444	18.15	0.6540	7.40M
RUAS [17]	13.43	0.6807	6.39	0.4655	9.20	0.5515	16.63	0.5589	4.54	0.3196	10.59	0.4393	**0.003M**
MSEC [1]	20.52	0.8129	19.79	0.8156	20.35	0.8210	19.62	0.6512	17.59	0.6560	18.58	0.6536	7.04M
CMEC [19]	22.23	0.8140	22.75	0.8336	22.54	0.8257	17.68	0.6592	18.17	0.6811	17.93	0.6702	5.40M
ENC-DRBN [7]	22.72	0.8544	22.11	0.8521	22.35	0.8530	21.77	0.7052	19.57	0.7267	20.67	0.7160	0.58M
ENC-SID [7]	22.59	0.8423	22.36	0.8519	22.45	0.8481	21.30	0.6645	19.63	0.6941	20.47	0.6793	7.45M
FECNet [8]	22.96	0.8598	23.22	0.8748	23.12	0.8688	22.01	0.6737	19.91	0.6961	20.96	0.6849	0.15M
IAEC(Ours)	**23.50**	**0.8644**	**23.44**	**0.8761**	**23.46**	**0.8714**	**22.78**	**0.7289**	**20.42**	**0.7737**	**21.60**	**0.7513**	0.15M



(a) Input (b) DRBN (c) SID (d) FECNet (e) LCDP
(f) MSEC (g) ENC-DRBN (h) ENC-SID (i) Ours (j) GT

Figure 6. Visual comparison with other methods on underexposed images from the ME dataset. There are brightness and color shift issues that exist in MSEC, DRBN, and LCDP, while SID tends to generate artifacts. Compared with other methods, our method achieves good brightness recovery and also handles structure details in the background appropriately, achieving the best visual effect.

40,000(160 epochs), respectively. To address the potential gap between the training and the test set caused by resizing, we fine-tuned our model on the original size of the training set for 176,800 and 20,000 iterations, with the batch size set to 1. The initial learning rate is $1e^{-4}$, and decays by a factor value of 0.5 at $\frac{1}{2}$ and $\frac{3}{4}$ of the total iterations. We adopt PSNR SSIM for evaluation.

4.2. Comparison with State-of-the-art

To verify the performance of proposed method, we compare it with state-of-the-art exposure correction methods, including CLAHE [23], RetinexNet [26], Zero-DCE [5], DPED [10], DRBN [34], SID [4], RUAS [17], MSEC [1], CMEC [19], ENC [7] and FECNet [8]. Note that we report the number of parameters for the different methods and we don't compare the results with methods with a large number of parameters such as LACT[2], which is more than 10 times the size of our backbone.

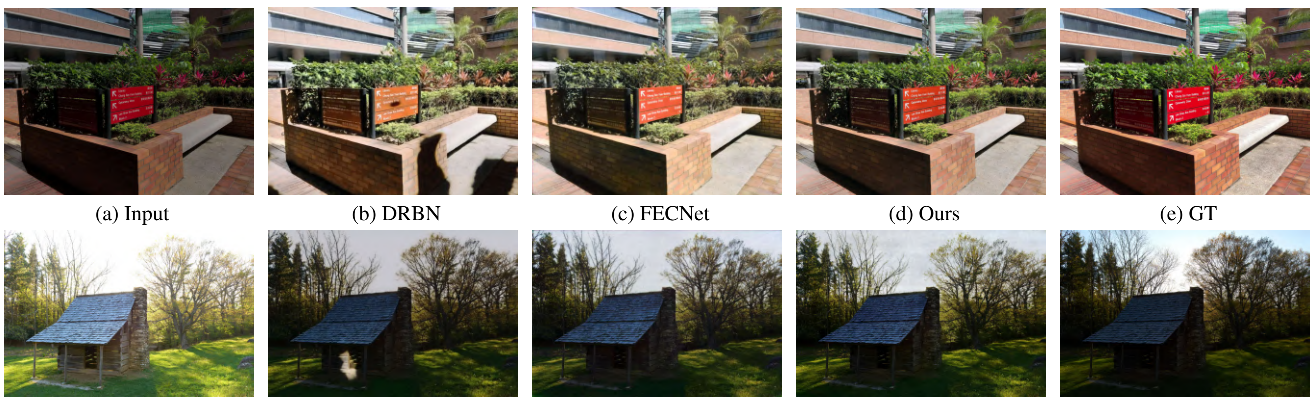
Quantitative results. Table. 1 presents the quantitative comparisons of our method with state-of-the-art methods on the ME and SICE datasets. For the ME dataset, we averaged the first two exposure levels as the underexposed subset and the remaining exposure levels as the overexposed subset. Some results of existing methods are obtained from [8]. Our method achieves the highest PSNR and SSIM on all sets, achieving a PSNR value of 23.46dB and an SSIM value of 0.8714. For the SICE dataset, our method yields the best average performance, including a significantly improved PSNR score of 21.60dB and a comparable SSIM score of 0.7513. While achieving these, our method inherits the advantages of FECNet[8] and uses a small number of parameters.

Qualitative results. We provide visual comparisons of our method with these state-of-the-art methods on several ME and SICE datasets as qualitative results. Figs. 6 and 7 show the visual results of some methods tested on underex-

(a) Input (b) DRBN (c) SID (d) LCDP (e) FECNet

(f) MSEC (g) ENC-DRBN (h) ENC-SID (i) Ours (j) GT

Figure 7. Visual comparison with other methods on overexposed images from the ME dataset.



(a) Input (b) DRBN (c) FECNet (d) Ours (e) GT

(a) Input (b) DRBN (c) FECNet (d) Ours (e) GT

Figure 8. Visualization results on the SICE dataset of (top) underexposure correction and (bottom) overexposure correction.

posed images from the ME dataset. Visualization results on the SICE dataset can be seen in Fig. 8. Due to the guidance of the reconstruction structure and the contextual semantics brought by the repaired image, IAEC not only has reasoning capabilities in areas with missing structures caused by severely improper exposure but also performs well in overall brightness restoration.

4.3. Ablation Studies

To validate the effectiveness of our method, we performed the following ablations on the ME dataset.

The effect of the Inpainting Module. The inpainting module is a simple fused block-based encoder-decoder structure network that performs pre-structure information repair. Without the inpainting module, our method would lose the pre-made semantic repair and degenerate into standard FECNet. In Fig. 9 we visualize the effect of the module. It can be seen that the module repairs the leaves that are mixed due to underexposure in the mask area, which allows our IAEC to understand nearby blocks with more semantic information to achieve better results.

Nevertheless, the inpainting module still has huge advantages over our exposure correction method in terms of structural reconstruction. Although our method partially inherits its advantages, it is not enough. We believe there are better strategies to realize the potential of the inpainting module, and we will explore such strategies in the future.

The effect of weight factor α. To explore the effect of different weight factor α, we perform experiment experiments with setting different α. As shown in Table. 2, Appropriate α has a greater performance improvement, while

tuning / α	0	5	20	50
w/o tuning	23.12	23.25	23.14	23.15
w tuning	-	23.46	23.30	23.23

Table 2. Ablation study for investigating the effect of α. The two rows are the results of without and with fine-tuning on the original size training set samples.

tuning / λ	0	0.5	0.7	1
w/o tuning	23.12	23.25	23.15	23.17
w tuning	-	23.46	23.24	23.17

Table 3. Ablation study for investigating the effect of λ

(a) Input (b) Inpainting (c) FECNet (d) IAEC(Ours) (e) GT (e) Mask

Figure 9. The reasoning capabilities of the inpainting module enable it to separate intermingled color blocks well in the mask area due to improper exposure, and our IAEC benefits from this. (e) is the mask of the image used in the inpainting module.

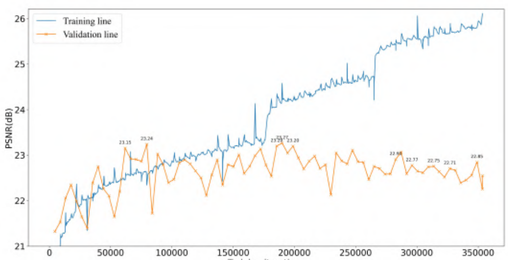


Figure 10. In the later stages of training, especially following the reduction in learning rate, the model's performance on the validation set declines instead of continuing to improve.

a larger α results in less improvement.

The effect of mix coefficient λ. We set λ to 0, 0.5, 0.7, and 1 to observe the impact of the mixing degree of the pre-repaired image and the original image on the performance.

4.4. Analysis

The randomness of the positions of severely improper exposed areas in the image means that we cannot guarantee that the randomly cropped patch contains the valid part of the inpainting module output. Therefore, we choose to resize input images so that the input of the exposure correction model must contain the valid output of the inpainting module. However, this leads to a domain gap caused by scaling between the training and the test set. This domain gap combined with the overfitting of the training set by the model in the later stages of training will have a serious negative impact. As shown in Fig. 10, in the late training period, the PSNR of the model on the training set has reached about 25.7dB, while the average PSNR on the verification set is only about 22.7dB, which is not even as good as the effect in the early and middle stages of training. To eliminate this effect we recommend fine-tuning the model on the

Method	ZERO-DCE	LCDPNet	FECNet	Ours
PSNR	12.59	23.24	22.58	**23.44**
SSIM	0.653	0.842	0.839	**0.863**

Table 4. Quantitative comparison on the LACP dataset.

original size of the image, which achieved good results as shown in Tables. 2 and 4.

Compared with FECNet, the effectiveness of the inpainting strategy is demonstrated by the great Quantitative and Qualitative improvement, which comes from the good reasoning ability displayed in the mask area. But it's noteworthy that real-world images aren't always underexposed in dark areas. Therefore, the current mask region selection method is not optimal. We will explore better mask strategies in the real world in future works.

5. Conclusion

In this study, we introduce a novel approach by leveraging structural repair to enhance the capabilities of existing exposure correction networks. This is the first attempt at the exposure correction task. The proposed framework integrates image inpainting and exposure correction, marking the first instance where structural repair assists in exposure correction. Our methodology involves training an inpainting network on the exposure-inpainting dataset, enabling it to reason effectively on mask areas. Subsequently, the output of the inpainting network is employed to guide the exposure correction network. Remarkably, a simple inpainting network exhibits robust reasoning capabilities, achieving commendable structural repair and introducing richer semantic information. This proves beneficial for existing exposure correction networks in both structural repair and brightness adjustment. Experimental results on the various datasets demonstrate the superior performance of our method compared to existing exposure correction methods.

References

[1] Mahmoud Afifi, Konstantinos G Derpanis, Bjorn Ommer, and Michael S Brown. Learning multi-scale photo exposure correction. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 9157–9167, 2021. 1, 2, 3, 5, 6

[2] Jong-Hyeon Baek, DaeHyun Kim, Su-Min Choi, Hyo-jun Lee, Hanul Kim, and Yeong Jun Koh. Luminance-aware color transform for multiple exposure correction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6156–6165, 2023. 6

[3] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Trans. Image Process.*, 27(4):2049–2062, 2018. 5, 6

[4] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3291–3300, 2018. 6

[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 *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1780–1789, 2020. 6

[6] Ke Xu Haoyuan Wang and Rynson W.H. Lau. Local color distributions prior for image enhancement. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022. 3

[7] Jie Huang, Yajing Liu, Xueyang Fu, Man Zhou, Yang Wang, Feng Zhao, and Zhiwei Xiong. Exposure normalization and compensation for multiple-exposure correction. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 6043–6052, 2022. 2, 3, 5, 6

[8] Jie Huang, Yajing Liu, Feng Zhao, Keyu Yan, Jinghao Zhang, Yukun Huang, Man Zhou, and Zhiwei Xiong. Deep fourier-based exposure correction network with spatial-frequency interaction. In *Eur. Conf. Comput. Vis.*, pages 163–180. Springer, 2022. 2, 5, 6

[9] Jie Huang, Feng Zhao, Man Zhou, Jie Xiao, Naishan Zheng, Kaiwen Zheng, and Zhiwei Xiong. Learning sample relationship for exposure correction. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9904–9913, 2023. 3

[10] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 3277–3285, 2017. 6

[11] Jingyuan Li, Fengxiang He, Lefei Zhang, Bo Du, and Dacheng Tao. Progressive reconstruction of visual structure for image inpainting. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5962–5971, 2019. 3

[12] 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. 3

[13] Xiaoguang Li, Qing Guo, Rabab Abdelfattah, Di Lin, Wei Feng, Ivor Tsang, and Song Wang. Leveraging inpainting for single-image shadow removal. *arXiv preprint arXiv:2302.05361*, 2023. 3, 4, 5

[14] Liang Liao, Jing Xiao, Zheng Wang, Chia-Wen Lin, and Shin'ichi Satoh. Guidance and evaluation: Semantic-aware image inpainting for mixed scenes. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVII 16*, pages 683–700. Springer, 2020. 3

[15] Guilin Liu, Fitsum A Reda, Kevin J Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. Image inpainting for irregular holes using partial convolutions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 85–100, 2018. 3

[16] Hongyu Liu, Bin Jiang, Yibing Song, Wei Huang, and Chao Yang. Rethinking image inpainting via a mutual encoder-decoder with feature equalizations. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16*, pages 725–741. Springer, 2020. 3

[17] 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 *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 10561–10570, 2021. 6

[18] Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Qureshi, and Mehran Ebrahimi. Edgeconnect: Generative image inpainting with adversarial edge learning. In *ICCV*, 2019. 3

[19] Ntumba Elie Nsampi, Zhongyun Hu, and Qing Wang. Learning exposure correction via consistency modeling. In *Proc. Brit. Mach. Vision Conf.*, 2021. 2, 6

[20] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A. Efros. Context encoders: Feature learning by inpainting. In *CVPR*, 2016. 3

[21] Jonathan Pirnay and Keng Chai. Inpainting transformer for anomaly detection. In *International Conference on Image Analysis and Processing*, pages 394–406. Springer, 2022. 3

[22] Yurui Ren, Xiaoming Yu, Ruonan Zhang, Thomas H Li, Shan Liu, and Ge Li. Structureflow: Image inpainting via structure-aware appearance flow. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 181–190, 2019. 3

[23] Ali M Reza. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, 38:35–44, 2004. 6

[24] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 6849–6857, 2019. 3

[25] Yang Wang, Long Peng, Liang Li, Yang Cao, and Zheng-Jun Zha. Decoupling-and-aggregating for image exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18115–18124, 2023. 1

[26] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018. 6

[27] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018. 3

[28] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5901–5910, 2022. 3

[29] Chaohao Xie, Shaohui Liu, Chao Li, Ming-Ming Cheng, Wangmeng Zuo, Xiao Liu, Shilei Wen, and Errui Ding. Image inpainting with learnable bidirectional attention maps. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 8858–8867, 2019. 3

[30] Wei Xiong, Jiahui Yu, Zhe Lin, Jimei Yang, Xin Lu, Connelly Barnes, and Jiebo Luo. Foreground-aware image inpainting. In *CVPR*, 2019. 3

[31] Kai Xu, Huaian Chen, Chunmei Xu, Yi Jin, and Changan Zhu. Structure-texture aware network for low-light image enhancement. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(8):4983–4996, 2022. 3

[32] Xiaogang Xu, Ruixing Wang, and Jiangbo Lu. Low-light image enhancement via structure modeling and guidance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9893–9903, 2023. 3

[33] Zhaoyi Yan, Xiaoming Li, Mu Li, Wangmeng Zuo, and Shiguang Shan. Shift-net: Image inpainting via deep feature rearrangement. In *Proceedings of the European conference on computer vision (ECCV)*, pages 1–17, 2018. 3

[34] 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 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. 6

[35] 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. 3

[36] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. Generative image inpainting with contextual attention. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5505–5514, 2018. 3

[37] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. Free-form image inpainting with gated convolution. In *ICCV*, 2019. 3

[38] Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. Reconstruction by inpainting for visual anomaly detection. *Pattern Recognition*, 112:107706, 2021. 3

[39] Yanhong Zeng, Jianlong Fu, Hongyang Chao, and Baining Guo. Learning pyramid-context encoder network for high-quality image inpainting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1486–1494, 2019. 3

[40] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(6):1452–1464, 2017. 5

[41] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, and Yicong Zhou. Zero-shot restoration of underexposed images via robust retinex decomposition. In *2020 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2020. 3

[42] Minfeng Zhu, Pingbo Pan, Wei Chen, and Yi Yang. Eemefn: Low-light image enhancement via edge-enhanced multi-exposure fusion network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 13106–13113, 2020. 3