Raw Image Based Over-Exposure Correction

Shizhan Zhao, Yanzhao Zhang, Libo Yan, Xiaoqiang Lu, Licheng Jiao, Yuwei Guo
Intelligent Perception and Image Understanding Lab, Xidian University
Xi'an, Shannxi Province, 710071, China
23171214628, 23171214596, 23171214439@stu.xidian.edu.cn

Abstract

Our team, with the username CVCV on codalab, scored 20.56 on the Test-Ratio=3 leaderboard, ranking second. In this report, we will cover the technical details of solving this task. The task of the competition is based on overexposure correction of RAW images. Overexposure can lead to washed-out regions, where important details are lost, colors appear faded or completely bleached, and the overall contrast of the image is compromised. Therefore, the development of overexposure correction technology has become an important effort in the field of image processing. We used the Channel-Guidance Network (CGNet) model for overexposure correction. Since the input of CGNet is RGGB four-channel, the training data is converted from CR2 format to RGGB four-channel PNG format, and the test data is converted from mat format to RGGB four-channel PNG format. We input the overexposed images of four ratios (x3, x5, x8, x10) into CGNet for training. After training the model, we use the training weights for inference to get the predicted results. Multiplying the pixels of the predicted result by the appropriate magnification will further improve the score. We also tried the multi-scale test, left and right flip test, after the weighted average model fusion, the effect is not ideal. The experimental results show that our strategy is effective and we have achieved an excellent score of 20.56 on the Real-world Paired Over-exposure (RPO) dataset.

1 Introduction

This competition is supported by the 4th Workshop on Physics Based Vision meets Deep Learning (PBDL2024), which provides two directions: the Low-Light Enhancement and Detection Challenge and the High Dynamic Range Imaging Challenge.

Overexposed images present a unique set of challenges that can severely undermine image integrity. Overexposure can lead to washed-out regions, where important details are lost, colors appear faded or completely bleached, and the overall contrast of the image is compromised. This degradation can be particularly detrimental for applications relying on fine-grained visual cues, such as pattern recognition and texture analysis. In light of these issues, the advancement of overexposure correction techniques has become a crucial endeavor in the field of image processing, aiming to retrieve lost information in overexposed areas, rebalance color distributions, and reinstate the dynamic range that is crucial for maintaining visual fidelity. This work is integral to improving the robustness of computer vision systems when faced with images that have been excessively exposed to light.

To propel research in this field forward, the organizers will utilize the RAW image-based Real-world Paired Over-exposure (RPO) dataset, introduced by Prof. Fu's team in [1]. The RPO dataset is the first specialized collection to systematically study the generality and practicality of over-exposure correction models. The dataset encompasses both indoor and outdoor scenes, captured in daylight or under direct illumination to avoid flickering. Each short-exposure (normal-exposure) image is paired with long-exposure (over-exposure) images with 4 ratios (x3, x5, x8, x10). The use of mirrorless cameras like the Canon EOS 5D Mark IV, equipped with a full-frame CMOS sensor, ensured high-resolution captures and minimized vibrations. The training dataset consists of 300 ground truths and corresponding overexposed images (ratio=3,5,8,10). The resolution of both RAW image and corresponding sRGB image is 6744×4502. The validation and test set processes RAW images into four-channel (RGGB)images, crops and saves them as .mat files. Unlike training set, validation and test set only includes input files and does not have ground truth.

2 Methods

Our training process for the entire task is shown in **Figure 1**. For the original training dataset, it is first converted from CR2 to RGGB four-channel PNG. Then the data is input into the CGNet model and supervised learning is carried out by the GroundTruth of the RGB three-channel. Finally, the overexposed image can be corrected to make the overexposed image return to normal.

Our test process for the entire task is shown in **Figure 2**.For the original validation and test data set, it is first converted from mat format to RGGB four-channel PNG. Then the data is input into the CGNet model, the trained weights are loaded for model inference, and the inference results of RGB three-channel are obtained.

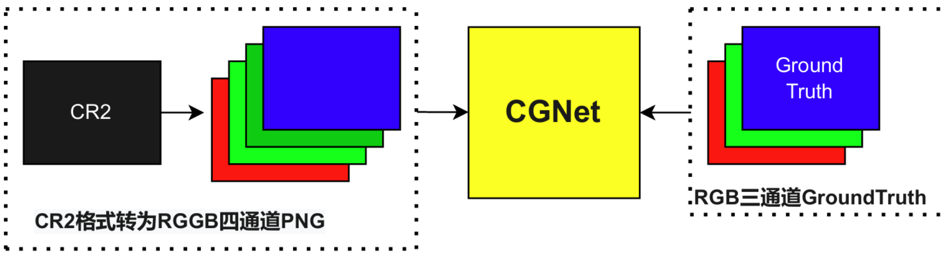

Figure 1: Training process diagram

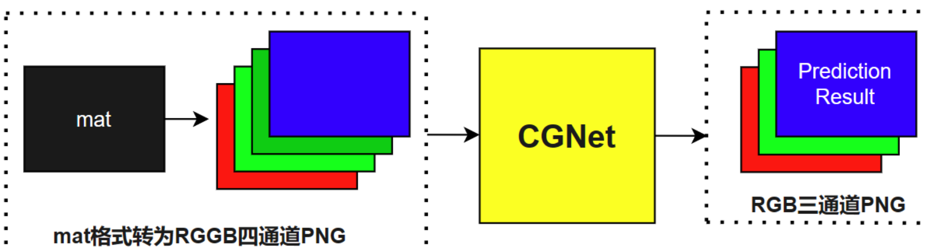

Figure 2: Test process diagram

Most existing methods of overexposure in image correction have been developed based on sRGB images, which can lead to complex and non-linear degradation due to the image signal processing pipeline. Compared to sRGB-based technologies, RAW images are characterized by a near-linear correlation with scene brightness and exhibit superior performance due to the rich information content due to higher bit depth. Traditional digital camera sensors are designed to have a higher response ratio and relative spectral sensitivity to green channels. Therefore, in RAW images captured by most digital camera systems, the green channel is usually more likely to be overexposed in bright scenes than the red or blue channel. The red and blue channels of RAW images show more appropriate brightness and richer texture details than the green channels. This indicates that the green channel in the RGGB RAW image is more saturated than the red or blue channel and requires stronger correction.

Channel-Guidance Network(CGNet), which takes advantage of RAW images for overexposure correction. CGNet estimates correctly exposed sRGB images directly from overexposed RAW images in an end-to-end manner. Specifically, they introduce a RAW based channel guide branch into the U-Net-based backbone, which utilizes color channel intensity priors of RAW images to achieve superior overexposure correction performance.

2.1 Data preprocessing

Our team chose CGNet model for overexposure image correction, and the model default input format is RGGB four-channel. In order to maintain the performance of the model, we decided to convert the original images in the dataset into RGGB format for training. The original training image is stored in CR2 format, and the rawpy library is directly called to batch convert CR2 files to RGB three-channel format. Then copy the green channel and convert it to RGGB four-channel format. Then, the overexposed images of four ratios (3,5,8,10) were input into CGNet together and divided into the training set and validation set according to the ratio of 8:2.

The storage format of the original test image is mat. After reading the mat file, it is found that the pixel value is between 0 and 1, and the four-channel is RGBG. Therefore, the pixels are multiplied by 255, and the four channels are converted to RGGB, and the processed PNG image is obtained.

2.2 CGNet

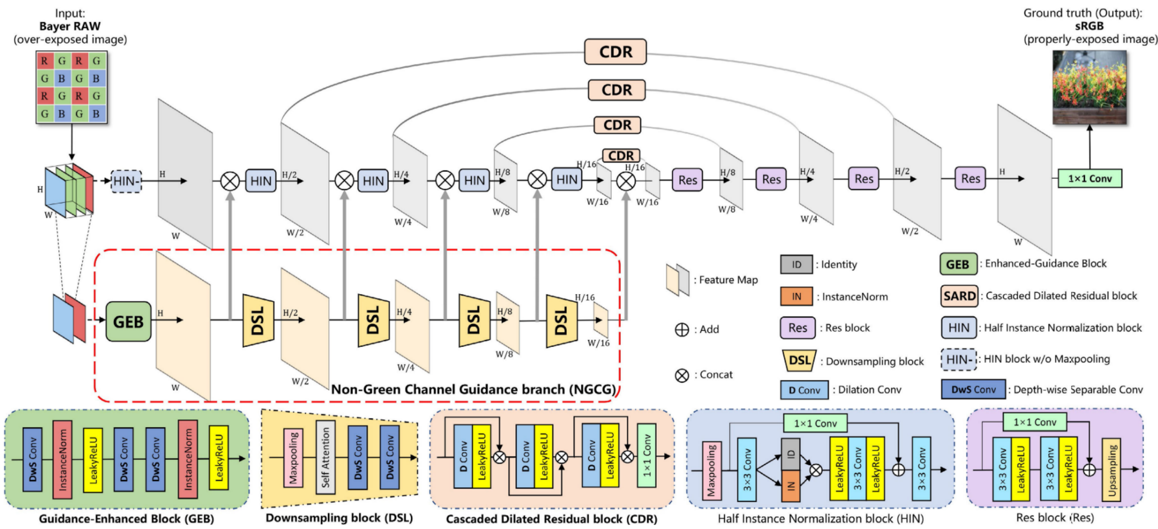

Figure 3: **Architecture of our Channel-Guidance Network (CGNet) for image over-exposure correction.** Given an over-exposed RAW image, they pad it into a four-channel RGGB image, and then feed it into their CGNet. Their CGNet is based on a U-Net backbone. The encoder consists of "Half-Instance Normalization", while the decoders are Residual blocks. They replace the original skip connection with Cascaded Dilated Residual (CDR) blocks. The red and blue channels are input to a Non-Green Channel Guidance (NGCG) branch for texture detail reconstruction. Their CGNet is pre-trained on their synthetic RAW image-based dataset, and fine-tuned on their Real-world Paired Over-exposure dataset.

As shown in **Figure 3**, the main branch is based on a basic U-net with four encoder (downsample) and decoder (upsample) stages. Specifically, they first extract the initial features from the four-channel RAW images using a standard 3×3 convolution. In the encoder section, the HIN blocks are utilized to broaden the receptive field and enhance the robustness of features at various scales. During the downsampling operation, they double the number of channels in the feature maps. Moving on to the decoder part, residual blocks are employed to capture high-level features more effectively. For the skip connection, they introduce a novel Cascaded Dilated Residual (CDR) block to extract multi-scale features, which are then merged with the encoder's features to mitigate the loss of detail information resulting from downsampling. The proposed NGCG branch integrates the prior knowledge of blue and red channels into each scale of the main branch encoder, aiding the main branch in recovering over-exposed areas more effectively.

Non-Green Channel Guidance: Firstly, the pixels pertaining to the red (or blue) channel are extracted from their respective positions within each 2×2 block of a Bayer image. Subsequently, these red and blue channels are input into the NGCG branch, which then generates an initial prediction of the corresponding components in the output sRGB image. The NGCG branch is structured with a Guidance-Enhanced Block (GEB) and four downsampling blocks, serving to guide the five corresponding encoder blocks. Within the GEB, there are two 3×3 convolutional layers, with an Instance Normalization and LeakyReLU following each layer, as well as a 3×3 convolutional operation in between. It is worth noting that depth-wise separable convolution is utilized instead of the traditional 3×3 convolution to efficiently capture local information. The downsampling blocks comprise a maxpooling layer, a modified self-attention structure, and two depth-wise separable convolution layers. This arrangement allows the NGCG branch to aid the primary backbone network in restoring over-exposed RAW images in a multi-scale manner.

Cascaded Dilated Residual Block: In detail, each CDR block incorporates three residual connections that feature dilated convolution and the LeakyReLU activation function. Subsequently, a 1×1 convolutional layer follows, which enables the CDR block to effectively utilize the features extracted from each stage of the encoder and adequately explore local texture information. Additionally, it is worth noting that the dilated convolution employed in this configuration effectively enhances the receptive field of the CDR block, thereby facilitating multi-scale contextual feature extraction.

3 Experiments

3.1 Implement Detail

We chose the CGNet model for training. Since the training data set contains overexposed images with four ratios (3,5,8,10), we input the overexposed images with four ratios into CGNet and divide them into the training set and the verification set according to the ratio of 8:2.

The training parameters include a batch size of 24, a learning rate of 0.0001, 4 undersampling layers, 9 residual layers, the optimizer selects Adam, trains a total of 500 epoches, and executes the learning rate reduction strategy from 100 epoches. Train using a single NVIDIA V100 GPU without loading pre-trained weights.

3.2 Results

The experimental results are presented in **Table 1**. After loading the weights obtained from the training, the processed test set is predicted, and the pixels of the predicted result are multiplied by a suitable multiplier to improve the score. Through experiments, it is found that the highest scores of overexposed images with ratio=3,5,8,10 are 20.56,20.49,20.54,20.03, respectively.

Table 1: Experiment Results

Ratio	argument	score
ratio=3	2.20	20.56
ratio=5	1.65	20.49
ratio=8	1.27	20.54
ratio=10	1.10	20.03

4 Conclusion

The main task of this competition is to correct overexposed images. The report details our team's approach to data processing and details of model training and model predictions. The experiment verifies that the data converted from CR2 or mat format to four-channel RGGB will not be affected by performance degradation when CGNet model training is used. In addition, by multiplying the pixels of the predicted results, we improved the quality of the picture and got a higher score. The experiment verifies that our strategy to solve this problem is reasonable and effective. In the end, we scored 20.56 on the RPO dataset, placing us in second place.

References

[1] Ying Fu, Yang Hong, Yunhao Zou, et al, ”Raw Image Based Over-Exposure Correction Using Channel-Guidance Strategy,” IEEE Trans. on Circuits and Systems for Video Technology, 2023.

[2] L. Chen, X. Lu, J. Zhang, X. Chu, and C. Chen, “HINet: Half instance normalization network for image restoration,” IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun.2021.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016.

[4] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017.