ISSN 2616-5775 Vol. 7, Issue 11: 27-33, DOI: 10.25236/AJCIS.2024.071104

Research on portability and effectiveness based on MPR-Net

Jingyao Zhang[1,*]

[1]*School of Optical Engineering, China Academy of Space Technology, Beijing, 100094, China*
[*]*Corresponding author: 673675215@qq.com*

Abstract: *As one of the well-known algorithms in the field of image restoration, Multi-Stage Progressive Image Restoration (MPR-Net) achieved quite great results in the fields of computer vision and image denoising. MPR-Net introduces the attention mechanism into the traditional encoder-decoder network and integrates the idea of multi-stage learning, so that the network can better learn the feature information of the image and improve the effect of image restoration. This paper takes the MPR-Net model as the main research object to introduce the principle of MPR-Net. In addition, this paper adopts a self-made binocular camera to collect the data set, and the degraded images in the data set are actually captured, which is different from the public GOPRO data set, which uses an average of 15 frames of clear images to obtain degraded images. The MPR-Net model is successfully applied to the self-made data set to realize image restoration, which verifies the effectiveness and portability of the MPR-Net model. Meanwhile, compared with the GOPRO data set, the image features of the self-made data set in this paper are seriously degraded, and the restoration effect of the MPR-Net model on this data set is greatly reduced. In this paper, the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used as measurement indicators. The training result of the MPR-Net model on the self-made dataset is 0.7304 and the PSNR is 25.79dB, which is 0.2 and 5.96dB less than the training result on the public GOPRO dataset, respectively.*

Keywords: *MPR-Net, Image Restoration, binocular camera, PSNR, SSIM, GOPRO*

1. Introduction

Today, with the rapid development of technology, the application of convolutional neural networks (CNNS) [3] extends from target location and image classification to more complex natural language processing (NLP) and image restoration. With deep learning as the background, the field of image restoration is booming. Image restoration technology is initially a simple single-scale encoder-decoder model. Under the background of supervised learning, it learns the features of the input image through convolutional neural network, then compares the output result with the sharp image, minimizes the loss function to reach a specific threshold, then backpropagates and repeats, and finally obtains the corresponding weight set. The degraded image is restored to a sharp image under the mapping of this weight set.

Image features include color features, texture features and shape features. There are many cases in which the image features are not obvious, such as the phenomenon that the color difference between the target and the background in the image is small. It is not enough to learn image features only through single-scale single-stage [8] network, and the features of large and small objects cannot be learned well at the same time, so new types of image restoration models are constantly emerging, and MPR-Net [1] model is one of them. The MPR-Net model uses the idea of multi-scale and multi-stage learning, and integrates attention mechanism and feature fusion mechanism into the model, so that the image restoration ability of the model is improved well.

To put it simply, MPR-Net divides the model into three periods. The first two periods are the encoder-decoder network of traditional image restoration technology, which can learn the image features of large receptive fields and extract the image features from a macro perspective. The final third stage belongs to the high-resolution branch, which can learn the potential features of images in a more detailed way. Finally complete the image restoration task [4].

In this paper, a self-made binocular camera will be used to construct a dataset of degraded image-sharp image to test the restoration effect of the MPR-Net model on this data set and prove its portability

ISSN 2616-5775 Vol. 7, Issue 11: 27-33, DOI: 10.25236/AJCIS.2024.071104

and usability in different scenes. At the same time, the restoration results are compared with the restoration effect of the MPR-Net model on the GOPRO data set, so as to comprehensively evaluate the performance and characteristics of the MPR-Net model. Finally, the MPR-Net model is summarized, hoping to make the model more perfect and play a beneficial role in its development.

2. Related work

2.1 Binocular camera Introduction

The shooting equipment of the data set in this paper is a binocular camera, which is composed of two commercial detectors A and B. The physical objects of the binocular camera is as shown in the Figure 1, and the performance parameters of the detector is as shown in the Table 1.

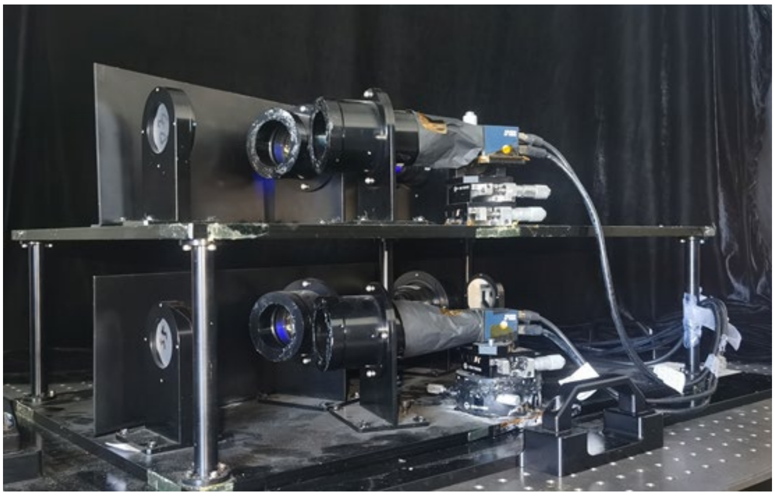


Figure 1: Binocular camera

Table 1: Performance specification of detector

Performance specification	Parameter
Model	MER-530-20GM-P NIR
Interface	GigE
Resolution	2592(H)×2048(V)
Frame frequency	20fps
Size of pixel	4.8μm×4.8μm
Depth of pixel	8bit,10bit
Time of exposure	5μs~1s

In the B camera, a custom anamorphic mirror is added to the first optical element to simulate the error of a low-precision mirror, resulting in a degraded image. Parameters of the deformable mirror are shown in the Table 2.

Table 2: Performance specification of anamorphic mirror

Performance specification	Parameter
Model	50
Range of voltage/V	-250~+300
Reflectance	99.9%
Diameter of effective area/mm	51
Thickness of mirror/mm	1.6
Weight/kg	0.85

The robustness of the B camera can be verified by simulation.

2.2 Data set creation

When camera A and camera B shoot, they ensure the same illumination, overlooking Angle, elevation Angle and atmospheric path. [8] The images taken by camera A are used as sharp images, and the images taken by camera B are used as degraded images. According to ORB [2] calibration and other pretreatment

ISSN 2616-5775 Vol. 7, Issue 11: 27-33, DOI: 10.25236/AJCIS.2024.071104

means, the two forms 1200 image pairs, including buildings, cars, signal towers, advertisements and so on. A screenshot of a partial homemade dataset is as shown in the Figure 2.

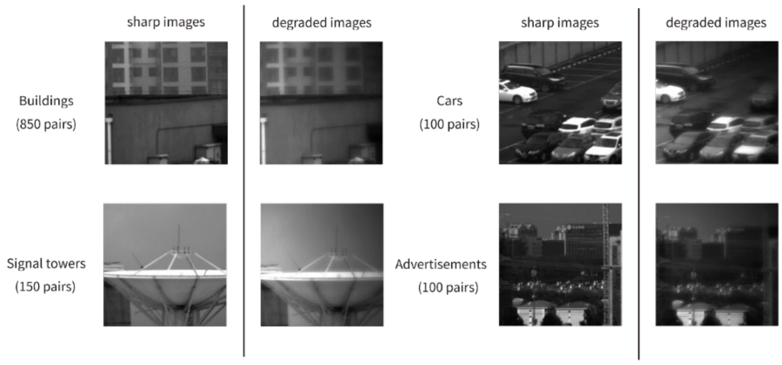


Figure 2: Self-made data set

The data set is divided into train set, test set and validation set, and the specific number of partitions is shown in the Table 3:

Table 3: Segmentation of data set

Data set	Amount
The number of train set	960
The number of validation set	120
The number of test set	120

2.3 MPR-Net

MPR-Net uses a new collaborative design approach to achieve a scientific balance between spatial details and high-level context information. The image features are learned step by step in a multi-stage way, so that the restored images retain more detailed features at the same time of high resolution. [1] The structure of the MPR-Net model is as shown in the Figure 3.

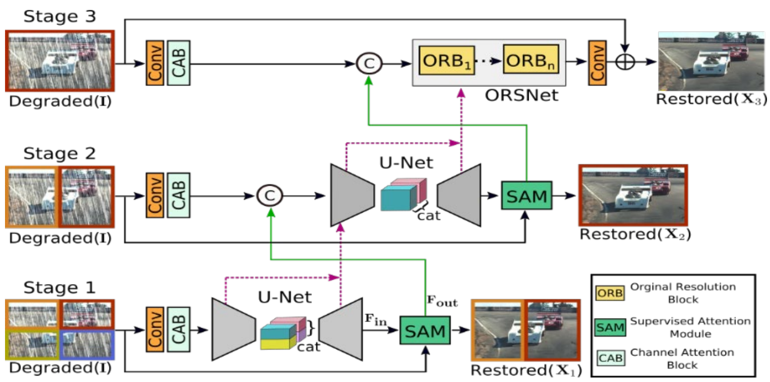


Figure 3: The structure of MPR-Net

Where, the Original Resolution Block (ORB) [2] structure diagram is as shown in the Figure 4.

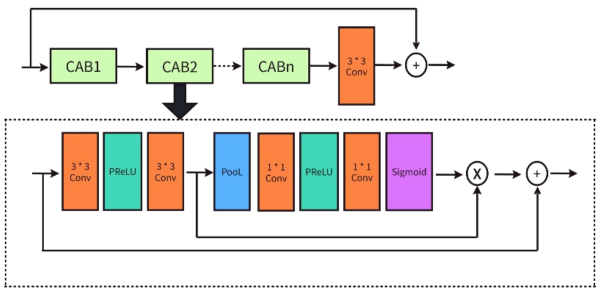


Figure 4: ORB in the ORSNet

ISSN 2616-5775 Vol. 7, Issue 11: 27-33, DOI: 10.25236/AJCIS.2024.071104

The Supervised Attention Module (SAM) [6] structure diagram is as shown in the Figure 5.

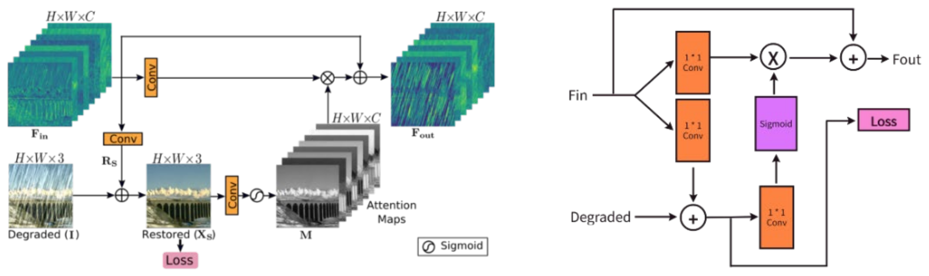


Figure 5: SAM

Encoder-decoder subnetwork [1] [5] [7] structure diagram is as shown in the Figure 6.

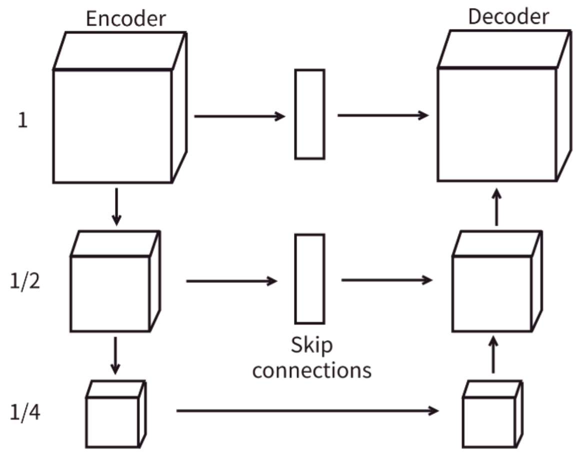


Figure 6: Decoder-Encoder Subnetwork

The overall process is as follows:

Stage 1:

Step 1: The input image is divided into four Patches.

Step 2: Each Patch expands its dimension by 3×3 convolution, so as to extract more abundant image features later.

Step 3: The temporary output passes through the CAB module and features on each dimension are extracted through the attention mechanism.

Step 4: Encoder process: Image features of three scales can be encoded, multi-scale context features can be extracted, and deeper semantic features can also be extracted.

Step 5: For deep-level features, the same scale features of the four patches are merged into the left and right scales.

Step 6: Decoder process: Extract the merged features for each scale.

Step 7: Eventually, The output1 of Stage 1 is obtained through SAM.

Stage 2:

Step 1: The input image is divided into two Patches.

Step 2: Each Patch expands its dimension by 3×3 convolution, so as to extract more abundant image features later.

Step 3: The temporary output passes through the CAB module and features on each dimension are extracted through the attention mechanism.

Step 4: Feature concat with output1.

ISSN 2616-5775 Vol. 7, Issue 11: 27-33, DOI: 10.25236/AJCIS.2024.071104

Step 5: Encoder-Decoder process.

Step 6: Eventually, The output2 of Stage 2 is obtained through SAM.

Stage 3:

Step 1: The input image is no longer divided.

Step 2: Each Patch expands its dimension by 3×3 convolution, so as to extract more abundant image features later.

Step 3: The temporary output passes through the CAB module and features on each dimension are extracted through the attention mechanism.

Step 4: Feature concat with output2.

Step 5: Enter into ORSNet.

Step 6: Eventually, The output3 of Stage 3 is obtained through convolution.

In short, this model is divided into three stages, Stage 1 and Stage 2 global learning of a wide range of image features, in Stage 3 no more subsampling operation, keep the input resolution of the network and the input image the same, so that the output image of the network to retain fine texture.

Add supervised attention modules to Stage 2 and Stage 3. Under the supervision module, feature fusion is carried out between the feature-map of the previous stage and the scaled feature-map of this stage, so as to consolidate the contextual image features. Each stage carries the semantic information of the previous stage, so as to realize multi-stage and multi-scale image restoration.

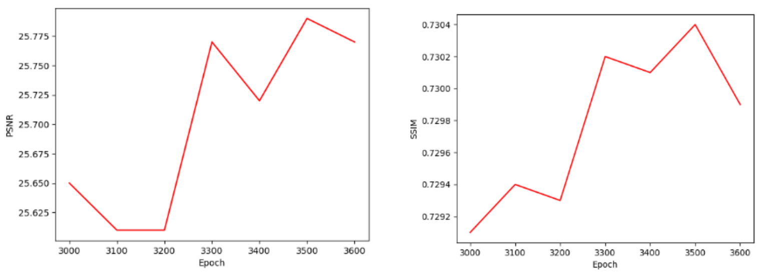
The loss function adopted by the MPR-Net model is as follows:

$$Loss = \sum_{S=1}^{3}[\sqrt{||X_S - Y||^2 + \varepsilon^2} + \lambda\sqrt{||\Delta(X_S) - \Delta(Y)||^2 + \varepsilon^2} \quad (1)$$

Where, $X_S = I + R_S$, at any given stage S, instead of instead of directly predicting a restored image X_S, MPR-Net predicts a residual image R_S. Y represents the ground-truth image. ε is set to 10^{-3}. Δ denotes the Laplacian operator. λ is set to 0.05.

3. Results

The PSNR and SSIM evaluation indexes of the MPR-Net model after 3000 rounds are as shown in the Figure 7.



(Left side is test set PSNR, right side is test set SSIM)

Figure 7: Test set PSNR and SSIM on self-made data set

After 3500 rounds, both PSNR and SSIM decreased, which may have caused overfitting. Therefore, the test index value of 3500 rounds was adopted: PSNR was 25.79dB, SSIM was 0.7304. The effect diagram of image restoration is as shown in the Figure 8(three pairs were randomly selected).

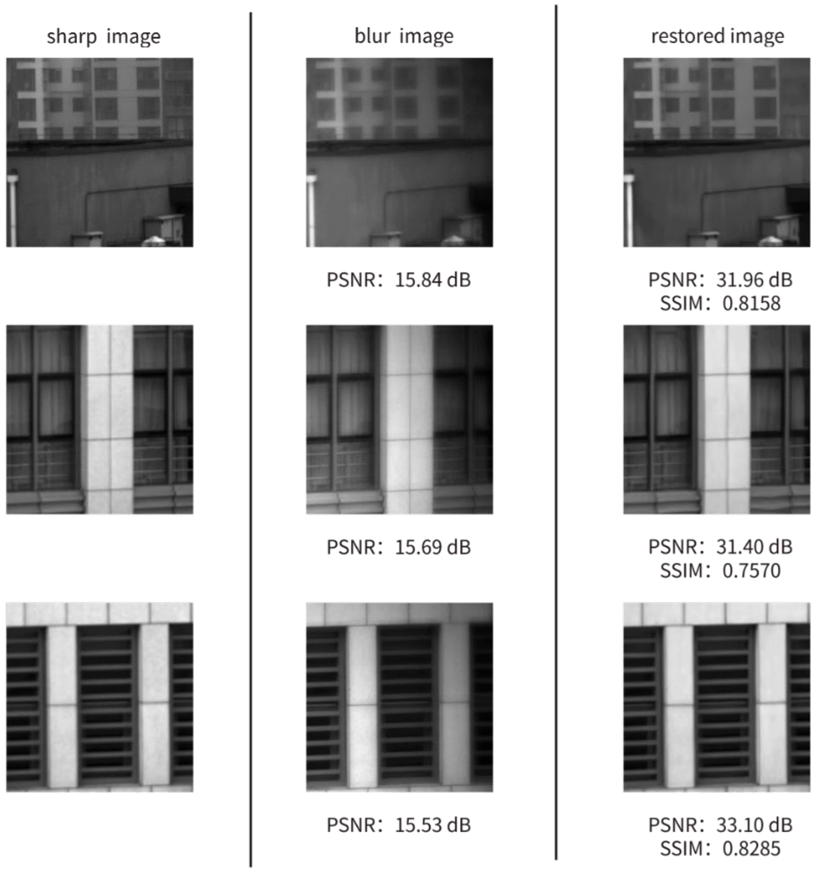


Figure 8: Restored Images

According to the restoration effect diagram, it can be seen that the MPR-Net model can restore the self-made data set in this paper, so that its look and feel are close to the sharp images captured, and the deblurring effect is good, which verifies the portability and effectiveness of the MPR-Net model, and also verifies that the degraded images in the self-made data set can be successfully restored to the sharp images. This data set can then be used for research related to image restoration.

However, the indicators of the MPR-Net model based on the GOPRO data set are respectively, PSNR is 31.75dB and SSIM is 0.9304. It can be seen that the PSNR and SSIM indexes based on the self-made data set in this paper are reduced by 0.2 and 5.96dB respectively, which positively reflects that the image quality of degraded images and sharp images in the data set will directly affect the restoration effect of the model.

4. Conclusions

The MPR-Net model can complete the image restoration task in both the GOPRO data set and the self-made data set, which verifies its portability and effectiveness. However, in the face of highly fuzzy data sets, the image restoration effect of the MPR-Net model will be reduced, so there is still room for improvement in this model. Other modules are added to the network for trial testing, so as to enhance its ability to learn image features.

References

[1] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-Stage Progressive Image Restoration [C]. arXiv:2102.02808, 2021.
[2] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, ORB: an efficient alternative to SIFT or SURF [C]. In ICCV 2011.
[3] Chen Chao, Qi Feng. Review on the development of Convolutional neural networks and their

ISSN 2616-5775 Vol. 7, Issue 11: 27-33, DOI: 10.25236/AJCIS.2024.071104

applications in computer vision [J]. Computer science, 2019, 46 (3): 11. DOI: 10.11896 / j.i SSN. 1002-137 - x. 2019.03.008.

[4] Saeed Anwar and Nick Barnes. Real image denoising with feature attention[C]. ICCV, 2019.

[5] Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: convolutional networks for biomedical image segmentation[C]. In MICCAI, 2015.

[6] Yulun Zhang, Kunpeng Li, Kai Li, Bineng Zhong, and Yun Fu. Residual non-local attention networks for image restoration[C]. In ICLR, 2019.

[7] Kong J Y, Zhang H S, Improved U-Net network and its application of road extraction in remote sensing image [J]. Chinese Science and Technology, 2022, 42(3): 105-113. (in Chinese)

[8] Chen Ruilin, Zhang Bo, Duan Xikai, et al. Remote Sensing Image Resolution Enhancement Technology Based on Single-Pixel Imaging [J]. Spacecraft Recovery & Remote Sensing, 2023, 44(6): 130-139. (in Chinese)