Visual Recognition using DL HW4 Report

Author: 司徒立中 (111550159) | GitHub Link

1. Introduction

In this task, the dataset is designed for image restoration, consisting of 1600 paired training images for each degradation type (rain and snow), where each pair includes a degraded image and its corresponding clean image. The test set contains 100 images with unknown degradation types (either rain or snow). The objective is to perform image restoration by predicting clean images from degraded inputs, effectively handling both deraining and desnowing tasks.

The performance is evaluated solely using Peak Signal-to-Noise Ratio (PSNR). PSNR is a widely used metric that measures the quality of reconstructed images by comparing them to their ground-truth clean versions, calculated as the ratio between the maximum possible pixel value and the mean squared error between the restored and clean images. Higher PSNR values indicate better image restoration quality.

Additionally, the strategies for this task are subject to specific constraints. First, no external data is allowed, ensuring the focus remains on model architecture design rather than data augmentation or collection. Second, no pre-trained weights can be used, meaning the model must be trained from scratch. The chosen model for this task is PromptIR[1], which is capable of handling multiple degradation effects (rain and snow) simultaneously within a unified framework.

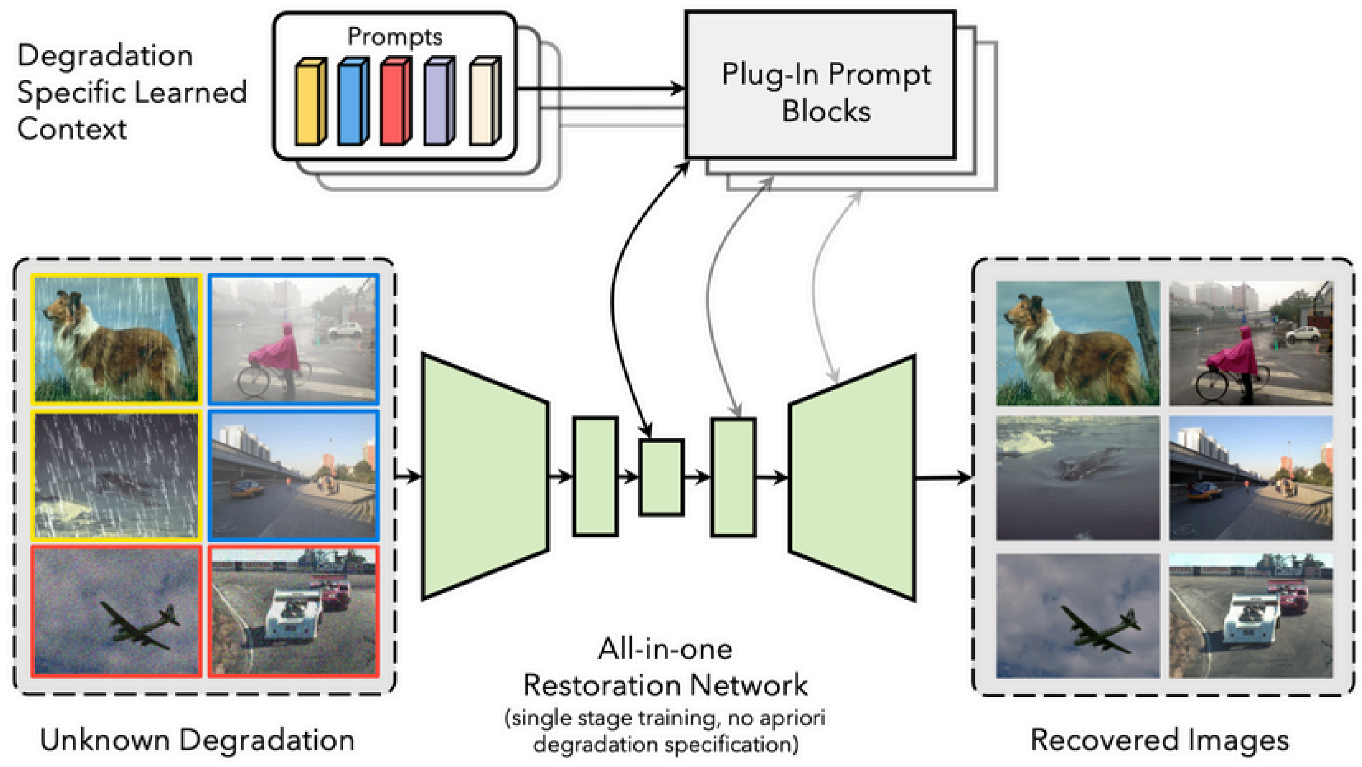

Figure1. **Pipeline Overview.** The model utilizes degradation-specific learned context prompts (top left) and plug-in prompt blocks to process images with unknown degradation types (bottom left). The all-in-one restoration network performs single-stage training without prior degradation specification, producing recovered images (right).

2. Methods

In my approach used for the image restoration task, I leverage the PromptIR[1] model with specific modifications to address both rain and snow degradation. PromptIR[1] is a versatile deep learning model designed to handle multiple image degradation types through a unified framework as Figure 1, using prompts to guide restoration for specific degradation effects. To adapt PromptIR for this task, several modifications were made to the input processing, model configuration, and training strategy.

First, the input types were expanded to include both deraining and desnowing tasks. The original deraining input pipeline was retained, while a desnowing component was newly integrated to handle snow degradation, enabling the model to process both degradation types simultaneously. To capture more contextual information and improve restoration quality, the input patch size was increased from 128×128 to 256×256 pixels. Additionally, test-time augmentation (TTA) was incorporated to enhance model robustness, applying transformations including rotations (90°, 180°, 270°), horizontal flipping, and vertical flipping, with the results aggregated to produce the final output.

To stabilize the training process, the amount of training data per epoch was increased by a factor of five, achieved through data augmentation techniques, ensuring the model encounters diverse samples during training. The hyperparameters were set as follows. The model was trained for 100 epochs with an initial learning rate of 2e-4, using a Warmup Cosine Annealing scheduler to adjust the learning rate dynamically for better convergence. These modifications collectively aim to improve the model's ability to restore images degraded by rain and snow, optimizing performance as measured by PSNR, which achieves approximately 32 on public testing dataset.

3. Results

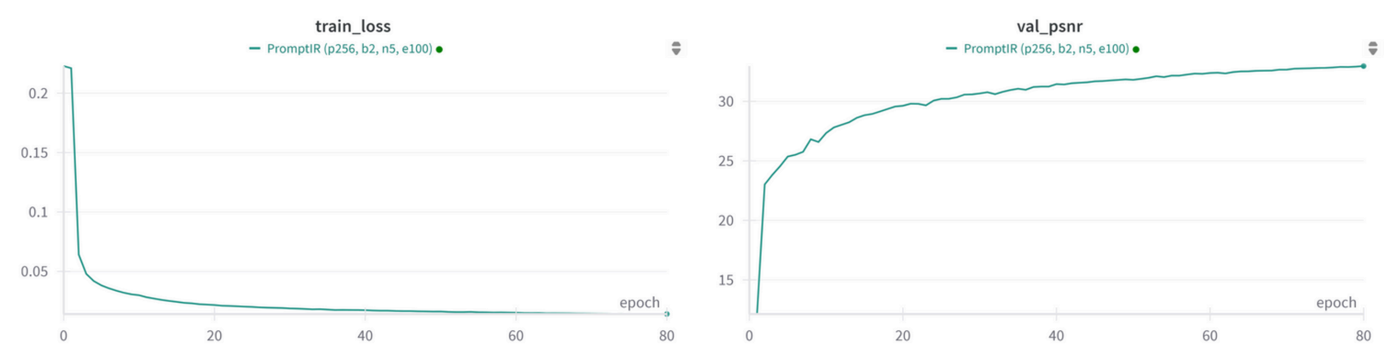

Figure2. **Training Curve.** It depicts the training loss (left) and the validation PSNR (right).

As shown in Figure 2, these two line charts illustrate the model's learning dynamics: the left chart shows the training loss steadily declining over epochs, indicating that the model is effectively optimizing its parameters; the right chart shows the validation PSNR consistently rising, demonstrating the model's improving ability to generalize and restore images accurately.

The effectiveness of the PromptIR model in restoring images degraded by rain is demonstrated through the predicted outcomes. As shown in Figure 3, the model successfully removes rain artifacts, producing visually clear images that closely resemble their clean counterparts. The top row of Figure 3 displays the degraded images, while the bottom row presents the restored images, achieving an impressive PSNR of 32. This high PSNR value indicates the

(PSNR ≈ 32)

Figure3. **Predicted Outcome.** The figure presents a comparison between degraded images (top row) and their corresponding predicted restored images (bottom row) using the PromptIR model. The restoration effectively removes rain degradation, achieving a high PSNR of 32, indicating strong performance in recovering clean images.

model's strong capability to recover fine details and maintain image quality, validating the effectiveness of the proposed modifications in handling rain degradation.

4. Additional Experiments

Figure 4. **Comparison of Image Restoration Results.** The figure shows restored images using PromptIR (top row) and AdaIR (bottom row) for the same degraded inputs, highlighting differences in restoration quality across various scenes.

In the additional experiment section, I explored the performance of AdaIR[2], a deep learning model designed for adaptive image restoration that adjusts its parameters dynamically based on the degradation type, as an alternative to PromptIR. As illustrated in Figure 4, the test results from both models appear visually similar across various scenes. However, a closer analysis of the validation PSNR reveals distinct differences: PromptIR achieved a validation PSNR of 32.89, while AdaIR scored 31.77. This indicates that, despite the comparable visual outcomes, PromptIR demonstrates a slight edge in restoration quality, reinforcing its suitability for the task.

References

[1] V. Potlapalli, S. W. Zamir, S. Khan, and F. Khan, “PromptIR: Prompting for All-in-One Image Restoration,” in Proc. 37th Conf. Neural Inf. Process. Syst. (NeurIPS), 2023.

[2] Y. Cui, S. W. Zamir, S. Khan, A. Knoll, M. Shah, and F. S. Khan, “AdaIR: Adaptive All-in-One Image Restoration via Frequency Mining and Modulation,” in Proc. 13th Int. Conf. Learn. Represent. (ICLR), 2025.

Code Reliability

In this implementation, I have made targeted modifications to the original PromptIR codebase sourced from its GitHub repository. All changes adhere to the Flake8 coding standard, ensuring clean, reliable, and maintainable code. For a detailed overview of the modifications and their impact on code reliability, please refer to the Code Reliability section of this repository.