北京联合大学研究生

公共外语 2 课程期末考试试卷

（2023—2024 学年 春季学期）

开课单位： 外语部

考试方式：☐开卷 ☐闭卷 ■其他（论文） **课程类型：**■学位课 ☐非学位课

学生姓名：张森 **学号：** 20231081210203 **学科专业：** 计算机科学与技术 **年级：** 2023

论文评语		得分	阅卷人

* *
* * * * *

A method of image restoration all-in-one machine for harsh weather conditions

A method of image restoration all-in-one machine for harsh weather conditions

1、Topic Basis and Research Content

1.1 Research Significance

Image restoration under adverse weather conditions refers to the process of using various techniques and algorithms to process low-quality or damaged images captured in conditions such as rain, fog, snow, and dust, to restore their original clarity and detail. This task is a significant branch of image processing and computer vision. Low-quality images under adverse weather conditions are often caused by poor lighting, environmental interference, and limitations of imaging equipment, manifesting as blur, noise, and color distortion. These issues not only affect the visual quality of the images but also have a negative impact on subsequent image analysis and computer vision tasks. To meet the demand for high-quality images in fields such as traffic monitoring, safe driving, and drone navigation, researchers have proposed many classical and effective image restoration algorithms.

The research goal of image restoration under adverse weather conditions is to optimize the overall visual effect of images while finely presenting the details within them. Currently, there are three main problems in image restoration under adverse weather conditions: first, the enhancement effect of images is weak, making it difficult to effectively present details; second, noise processing is incomplete, with the restored images still containing noise; and third, the restoration process fails to balance contrast and color saturation, leading to color distortion. Therefore, comprehensively and efficiently restoring low-quality images under adverse weather conditions is a challenging issue. In recent years, image restoration under adverse weather conditions has become a popular research topic both domestically and internationally, garnering widespread attention from scholars.

1.2 Domestic and International Research Progress

1.2.1 Image Restoration Methods

Research on image restoration under adverse weather conditions has experienced a long development process. Due to the limitations of computer science technology and hardware conditions at that time, the progress of the research has been relatively slow. Traditional image restoration algorithms are mainly divided into two categories. The first type is based on filtering techniques, and the second is based on transform domain techniques. Filtering techniques remove noise from images through spatial domain operations, such as mean filtering, median filtering, and bilateral filtering. However, these methods tend to blur image details while denoising. Transform domain algorithms process images in the frequency domain or other transform domains, such as wavelet transform, Fourier transform, and cosine transform. These methods effectively remove noise but still have deficiencies in detail recovery. They can improve global contrast to some extent but are less effective in local image restoration.

Image restoration algorithms based on deep learning mainly utilize neural networks and constraints of loss functions to restore images driven by patterns in large amounts of data, achieving better visual effects. This section will introduce the current state of research on image restoration algorithms both domestically and internationally from the perspectives of traditional methods and deep learning-based methods.

1.2.1.1 Traditional Methods

(1) Filtering-Based Methods

Filtering-based methods are common image restoration techniques. Their primary concept is to remove noise, enhance details, and improve contrast through filtering operations in the spatial or frequency domain.

Mean filtering[1]is a basic image filtering technique that smooths the image by calculating the average value of neighboring pixels to remove noise. However, while denoising, mean filtering also causes the loss of image details, making the image blurry. Median filtering[2]replaces the current pixel value with the median of neighboring pixels, effectively removing salt-and-pepper noise but having limited effectiveness against other types of noise. Bilateral filtering [3] is a nonlinear filtering method that can simultaneously consider spatial distance and pixel value differences. It can remove noise while preserving image edge details, but its computational complexity is high. The adaptive filtering method dynamically adjusts the filtering parameters based on the local characteristics of the image, which can better preserve image details while denoising. For example, adaptive median filtering [4] improves the denoising effect by adjusting the filtering window size to adapt to the noise levels in different regions. Guided filtering [5] filters the input image through guided images, which can better preserve edge and detail information.

(2) Methods based on the transform domain

The main idea of the transform domain method is to convert an image from the spatial domain to the frequency domain or other transform domains, utilize the characteristics in the transform domain for processing, and then convert the processed image back to the spatial domain. The Fourier transform is a commonly used transform domain method that separates low-frequency and high-frequency components in an image by converting it from the spatial domain to the frequency domain. The low-frequency components usually represent the overall structure and lighting information of the image, while the high-frequency components represent the details and edge information of the image. By processing the high-frequency components in the frequency domain, noise can be effectively removed and details can be enhanced [1]. The mathematical formula expression of the Fourier transform is shown in Equation 1.

$$\mathrm{F}(\mu,\nu)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\mathrm{f}(x,y)\,\varepsilon^{-j2\Pi(\frac{\mu x}{M}+\frac{\nu y}{N})} \tag{1}$$

The wavelet transform is a more complex transform domain method that can analyze images at multiple scales. The wavelet transform decomposes an image into sub-bands of different scales, allowing for the separate processing of image details and noise, thereby preserving details while removing noise [6]. The mathematical formula for the wavelet transform is shown in Equation 2, where is φ the mother wavelet function, and a and b are the scale and translation parameters, respectively. The wavelet transform performs well in image denoising and compression, but it is highly complex and requires a large amount of computation.

$$\mathrm{W}_{\varphi}(a,b)=\frac{1}{\sqrt{a}}\sum_{x=0}^{M-1} f(x)\varphi(\frac{x-b}{a}) \tag{2}$$

The Discrete Cosine Transform (DCT) [7] is also a commonly used transform domain method, especially in image compression. The DCT decomposes an image into a series of combinations of cosine basis functions, which can effectively compress image information and preserve important image features during denoising. The mathematical formula for the Discrete Cosine Transform is shown in Equation 3, where $\mathrm{C}(\mu)$ and $\mathrm{C}(\nu)$ are normalization factors. However, DCT has limited effectiveness in handling non-uniform lighting and complex noise.

$$\mathrm{F}(\mu,\nu)=\frac{2}{\sqrt{\mathrm{MN}}}C(\mu)C(\nu)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)\cos\left[\frac{\pi(2x+1)\mu}{2M}\right]\cos\left[\frac{\pi(2y+1)\nu}{2N}\right] \quad (3)$$

Although these transformation domain based methods have achieved certain results in image restoration under harsh weather conditions, there are still some problems. Firstly, the transformation domain method is difficult to avoid the loss of details while denoising; Secondly, these methods have limited adaptability to different types of noise and lighting changes. Future research will continue to focus on combining transform domain methods with deep learning to further improve the effectiveness and robustness of image restoration.

1.2.1.2 Image restoration method based on deep learning

In recent years, with the rapid development of deep learning technology, more and more computer vision tasks have begun to use deep learning methods for image processing. In image restoration tasks under harsh conditions, deep learning methods have also demonstrated excellent performance and wide applicability. For example, deep learning techniques have been successfully applied to tasks such as image denoising, image dehazing, and image rain removal. The DehazeNet proposed by He et al. [8] utilizes convolutional neural networks (CNNs) to learn the dehazing features of images, extracts image features through multi-layer convolutions, and ultimately generates clear images. Ren et al.'s MS-CNN [9] used multi-scale convolutional neural networks to extract multi-scale features from images, achieving more accurate dehazing effects. Qu et al.'s DAF Net [10] effectively removes haze under complex weather conditions and maintains image details through a bidirectional attention mechanism and feature fusion module.

Deep learning methods based on Generative Adversarial Networks (GANs) are also widely used for image restoration under harsh weather conditions. The GAN framework proposed by Goodfellow et al. achieved image dehazing effect through adversarial training of generators and discriminators. Li et al.'s CycleGAN [12] utilizes cyclic consistency loss to enable the model to learn image transformation and restoration under unsupervised conditions. Liu et al.'s Pix2PixHD [13] achieved high-quality image restoration under adverse weather conditions through high-resolution generative adversarial networks.

Researchers have proposed an end-to-end multi task learning method for image restoration tasks under harsh weather conditions. Zhang et al.'s DID-MDN [14] achieved image restoration under various adverse weather conditions in a single network by combining multiple tasks of defogging, rain removal, and snow removal through a multi task learning network. Fu et al.'s All in One Dehazing [15] significantly improved the restoration performance by processing image restoration tasks under multiple weather conditions through a unified network structure. Wang et al.'s Hi Net [16] utilized a hierarchical network structure to restore images at different levels, achieving more comprehensive restoration of severe weather images.

1.2.2 Methods for image restoration under adverse weather conditions

Under adverse weather conditions, image shooting scenes are prone to blurring, noise, and color distortion. Directly enhancing images cannot effectively remove these interferences, resulting in a poor visual experience of enhancement effects. The RainNet proposed by Xu et al. [16] uses a convolutional neural network (CNN) structure to separate and remove raindrops from images, and further enhances image detail restoration using residual blocks. This method is trained on synthesized rain images, and the learned model can adaptively remove raindrop interference of different intensities. The FogNet proposed by Zhu et al. [17] adopts a multi-scale feature extraction network to perform image dehazing processing. This method effectively distinguishes fog and image details through hierarchical feature extraction and fusion, significantly improving the clarity of the image. By training on a real haze image dataset, FogNet demonstrated superior dehazing performance. The SnowRemovalNet proposed by Chen et al. [18] trains on pairs of snow images and corresponding snow free images using a Generative Adversarial Network (GAN) framework. This method generates snow free images through a generator and conducts adversarial training through a discriminator, making the generated images visually more realistic and natural. The DustNet proposed by Li et al. [19] utilizes a dual stream convolutional network (Dual Stream CNN) to separately process the brightness and color information of images to remove dust interference. This method can effectively restore the color and details of images through training on a sand and dust image dataset. The HazeRemovalNet proposed by Wang et al. [20] combines deep convolutional neural networks (DCNN) and attention mechanisms, significantly enhancing the dehazing effect of images through multi-scale feature fusion and attention modules. This method was trained on a large number of foggy image datasets and demonstrated excellent dehazing performance. The MS Net proposed by Ren et al. [21] adopts a multi-scale network architecture, which effectively removes noise and interference under various weather conditions through multi-level feature extraction and fusion, improving image clarity and detail representation. Qian et al. [22]'s DenseDehazeNet combines the advantages of dense connected networks and residual networks, achieving efficient image dehazing through multi-level feature fusion and step-by-step dehazing processing. The RainRemovalGAN proposed by Zhang et al. [23] is based on a generative adversarial network framework, utilizing multi-scale discriminators and multi-level loss functions to achieve efficient raindrop removal and image restoration.

1.2.3 All in one method

Although deep learning methods perform well in image restoration tasks under harsh weather conditions, they still have some limitations and challenges compared to all in one methods. Firstly, deep learning methods typically focus on solving specific image restoration tasks, such as denoising, defogging, or rain removal. This single task optimization strategy often finds it difficult to simultaneously solve multiple types of image degradation problems when facing complex adverse weather conditions.

The All in one method can comprehensively improve image quality by processing multiple tasks simultaneously through a unified framework. For example, the All in One Image Restoration for Unknown Corruption method (AirNet) proposed by Li et al. [24] achieves efficient image restoration under unknown degradation conditions by comparing learning and degradation guided restoration networks. The All in one method can alleviate the problem of data dependency to some extent through multi task learning, as different tasks can share features and data resources, improving the overall robustness of the model. For example, the Adaptive Discriminative Filters method proposed by Park et al. [25] achieves image restoration for various unknown degraded images through adaptive discriminative filters. Another study by Li et al. [26], the All in One Bad Weather Removal, utilizes building search strategies to address image degradation issues under various adverse weather conditions within a single framework. The DID-MDN model proposed by Zhang et al. [27] combines multiple tasks of defogging, rain removal, and snow removal through a multi task learning network, achieving image restoration under various adverse weather conditions in a single network. The All in One Dehazing model proposed by Fu et al. [28] significantly improves the restoration performance by processing image restoration tasks under various weather conditions through a unified network structure. Wang et al. [29] utilized a hierarchical network structure to restore images at different levels, achieving more comprehensive restoration of severe weather images.

This project intends to use the PromptIR [30] network as the backbone network for image restoration. The network first preprocesses the image through a convolutional neural network; Then, feature extraction is performed through multiple Transformers blocks, each of which is composed of Multi DConv Head Transferred Self Attention (MDTA) and Gated Dconv Feed Forward Network (GDFN); The core idea of MDTA mechanism [31] is to combine self attention mechanism with multiple dynamic convolutions, so that the attention mechanism can not only capture global information, but also enhance feature expression by utilizing local convolution characteristics. The core idea of GDFN [31] is to integrate dynamic convolution and gating mechanisms into feedforward neural networks, thereby improving the network's nonlinear modeling ability and computational efficiency; At the same time as upsampling, a Prompt Block was added to dynamically guide the recovery network. However, the structure of the PromptIR network is relatively complex and the content in the prompt components is too random, so there is still a lot of room for improvement.

2、The research content, research objectives, and key issues to be addressed in the selected topic

2.1 research contents

This project aims to address the challenging issues of complex harsh weather environments and complex blurring in image restoration tasks, and proposes an improved self attention mechanism to enhance the network's generalization ability for multitasking processing. The research content includes the following aspects:

(1) Utilizing an improved attention mechanism TKSA to maintain efficient feature extraction capability while reducing computational complexity and memory consumption;

(2) Using an improved Prompt Block component, the category label information of multiple tasks is used as the prompt information for each task, thereby improving the generalization ability of the multi task model;

(3) To address the issue of a large number of network parameters, we plan to design a lightweight module to reduce the number of parameters and improve the efficiency of algorithm training without compromising performance;

2.2 Research objectives

This project improves the existing All in one image restoration algorithm to enhance its enhancement effect. The research objectives include:

(1) Using dynamic convolution and improving attention mechanisms to enhance the model's ability to express details and optimize its computational results;

(2) Building lightweight models to accelerate model convergence speed;

(3) Improve a new prompt component by using class label information of each task as prompt information to enhance the generalization ability of the model and enhance the visual effect of the image.

2.3 Key issues to be addressed

(1) How to use attention mechanism to obtain more feature information and improve network enhancement effect

(2) How to build a lightweight network to reduce model parameters and accelerate model convergence speed

3、Proposed research plan (research ideas, technical routes or methods) and feasibility analysis

3.1 technology roadmap

The overall research approach for this topic is divided into three steps: the first step is feature extraction, which utilizes attention mechanisms to extract deep features of images. The second step is to add plug and play prompt information and use class label information to assist in image restoration. The third step is feature fusion, which combines feature map information from multiple resolutions with auxiliary information to obtain richer information.

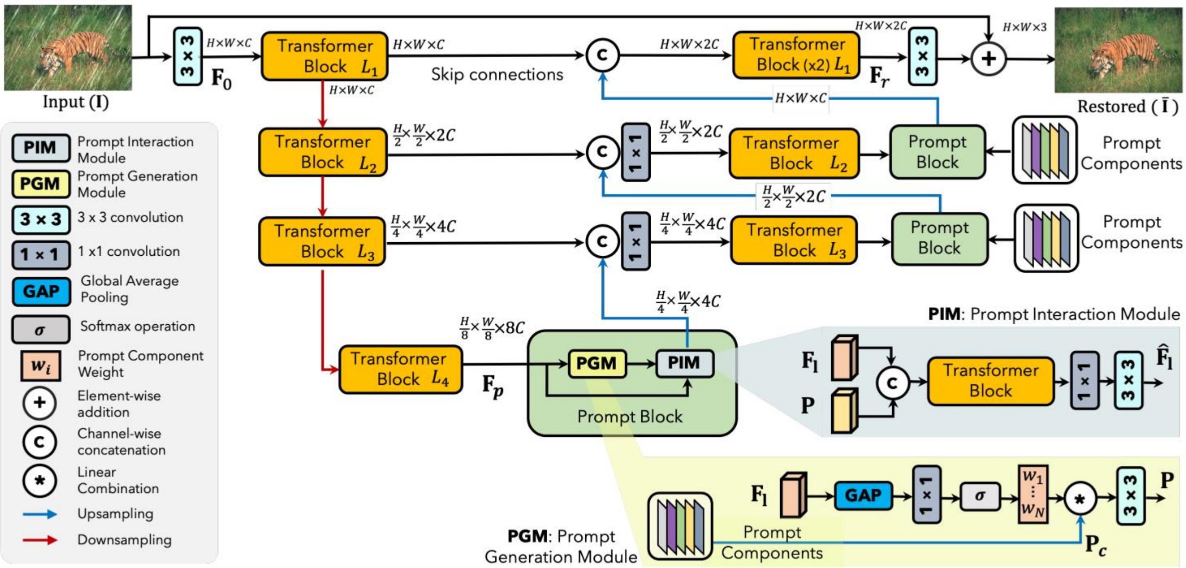

Figure 1:framework diagram

3.1.1 feature extraction

Regarding feature extraction, the MDTA in the PromptIR network [30] structure, as shown in Figure 2, obtains image information through multiple dynamic convolutions and transposed self attention mechanisms. Multi dynamic convolution mainly focuses on extracting local features. Due to the limited receptive field of convolutional kernels, dynamic convolution can enhance the expression ability of local features, but it is difficult to capture global contextual information. Although the transposed self attention mechanism improves computational efficiency, in some cases, it may not fully capture the global relationships of input data. Especially in high-resolution images or long sequence data, transpose operations may overlook some important contextual information. To address the aforementioned issues, this study adopts an improved MDTA self attention mechanism to better utilize contextual information.

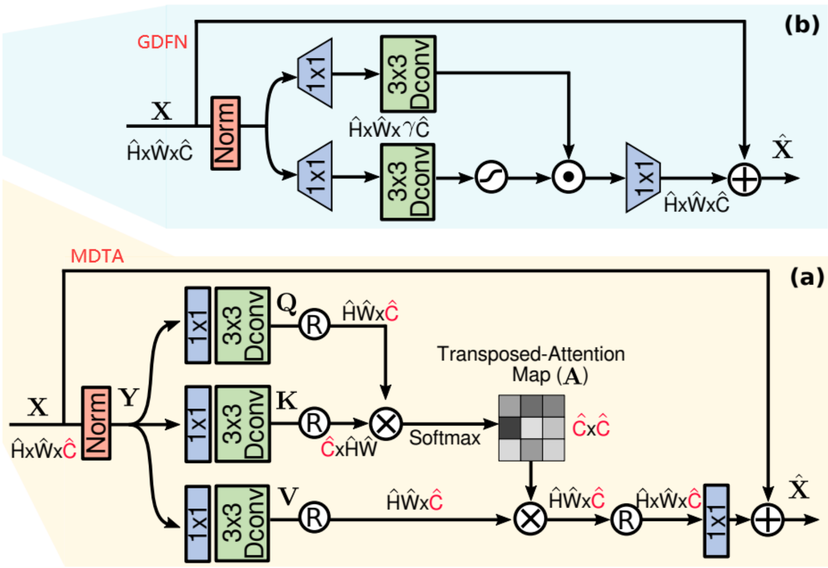

Figure 2:MDTA framework diagram

3.1.2 Prompt component

For plug and play prompt components, this project intends to use the pre trained model CLIP [32] to process the class labels of multitasking, as shown in Figure 3. Prompt Block [30] is composed of PGM and PIM. PGM mainly utilizes prompt parameters, image feature information, and text information to obtain a set of prompt parameters through pre trained models CLIP, GAP, and convolutional neural networks. PIM mainly obtains auxiliary prompt information through concatenation and Transformers blocks for prompt information and image feature information.

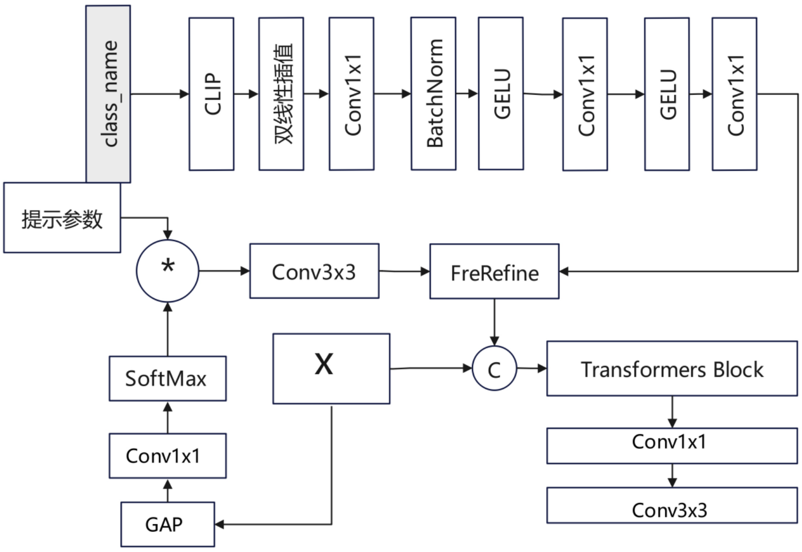

Figure 3:Prompt Blockframework diagram

3.1.3 Feature fusion

For feature fusion, this project plans to adopt a parallel approach. The PIM module dynamically adjusts the receptive field through two operations, fuse and select. Fuse generates global feature descriptions by combining information from multi-resolution streams. Select uses these descriptions to recalibrate feature maps (different streams) and then aggregate them.

Obtain images of different resolutions through simple downsampling, while preserving the details of the original resolution. By gathering multi-scale information, information can be fused from multi-resolution branches to enhance expression learning ability. The network structure of MIRNet [33] is divided into two repeated stages, which fuse information from streams of different resolutions through a feature fusion module to ensure high-resolution information in the network.

In the fusion module shown in the above figure, the information of the multi-resolution stream is added element by element, resulting in loss of information. Channel stitching (Figure 4) provides richer information for subsequent tasks compared to element wise addition to obtain more diverse features.

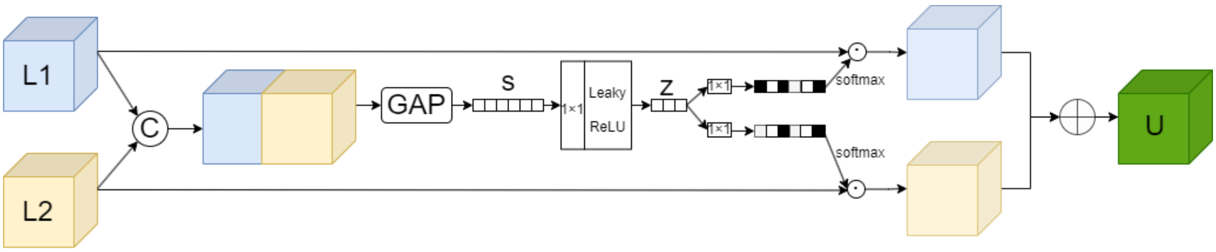

Figure 4: An improved SKFF

3.1.5 Improvement ideas

To address the issue of difficulty in capturing contextual information through the MDTA attention mechanism, it is proposed to introduce an improved MDTA in the backbone network to capture global contextual information;

To address the high computational complexity of referencing attention mechanism noise, a lightweight module is adopted to reduce model size and accelerate model convergence speed;

To address the issue of randomness in prompt components, we plan to design a plug and play prompt component with class label information as an aid to improve visual effects;

3.1.6 Datasets and evaluation metrics

Datasets: Currently, there are mainstream publicly available datasets for image restoration, and specific information for each dataset is shown in Table 3-1.

Table 3-1 Commonly Used Public Datasets for Image Restoration

Datasets	Sample quantity (Training/Testing)	Synthetic or Realistic	来源
Reside(SOTS)	8970/120	Synthesis+Realism	Lift/Camera
Rain100L	200/100	Synthesis+Realism	Flicker/Camera
WED	4744/	Synthesis+Realism	Lift/Camera
BSD400	400/	Synthesis+Realism	Lift/Camera

Evaluation indicators: SSIM and PSNR were selected as the evaluation indicators for the experiment.

Among them, SSIM (Structural Similarity) is a measure of the similarity between two images, with values ranging from (0, 1). The closer the value is to 1, the higher the similarity between the two images. 0 represents completely different, and 1 represents completely identical. The SSIM formula is as follows:

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_x^2+c_2)} \qquad (2)$$

PSNR (Peak Signal to Noise Ratio), also known as Peak Signal to Noise Ratio, is an objective standard for evaluating images, typically used in engineering projects between maximum signal and background noise. The PSNR formula is as follows:

$$\mathrm{PSNR}=10\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right) \qquad (3)$$

$$\mathrm{MSE}=\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[\mathrm{I}(i,j)-K(i,j)]^2 \qquad (4)$$

Among them, MAX is the maximum possible pixel value of the image, and MSE is the mean square error between a clean image I of size m x n and a noisy image K. The PSNR is close to 50dB, indicating only a slight, very small error between images; If the PSNR is greater than 30dB, it is difficult for the human eye to detect the difference between the compressed image and the original image; The PSNR is between 20dB and 30dB, and the human eye can detect differences in the image; The PSNR ranges from 10dB to 20dB, and the human eye can still see the original structure of the image with the naked eye, and intuitively judge that there is no significant difference between the two images; The PSNR is below 10dB, making it difficult for humans to visually determine whether two images are from the same image.

3.2 feasibility analysis:

The Residence (SOTS), Rain100L, WED, and BSD400 datasets are commonly used datasets in the field of image restoration under harsh conditions.

The hardware facilities required for the experiment are already in place, with a graphics processor of NVIDIA GeForce RTX 3090 and a graphics memory of 24G, meeting the experimental requirements.

The corresponding software environment has been built, and the experimental platform is using Pycharm Professional Edition. The deep learning framework is using Pytorch 1.12.0

We conducted sufficient research in the early stage and demonstrated the plan.

（三）References

[1] Gonzalez, R. C., & Woods, R. E. (2002). Digital Image Processing. Prentice Hall.

[2] Jain, A. K. (1989). Fundamentals of Digital Image Processing. Prentice Hall.
[3] Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images. Proceedings of the 1998 IEEE International Conference on Computer Vision, 839-846.
[4] Hwang, H., & Haddad, R. A. (1995). Adaptive median filters: New algorithms and results. IEEE Transactions on Image Processing, 4(4), 499-502.
[5] He, K., Sun, J., & Tang, X. (2013). Guided image filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(6), 1397-1409.
[6] Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 674-693.
[7] Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete Cosine Transform. IEEE Transactions on Computers, C-23(1), 90-93.
[8] He, K., Sun, J., & Tang, X. (2016). DehazeNet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing, 25(11), 5187-5198.
[9] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., & Yang, M. H. (2016). Single image dehazing via multi-scale convolutional neural networks. European Conference on Computer Vision, 154-169.
[10] Qu, Y., Chen, Y., Huang, J., & Xie, Y. (2019). Enhanced pix2pix dehazing network. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8160-8168.
[11] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27.
[12] Li, J., & Ren, W. (2018). Cycle-consistent generative adversarial network for dehazing. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8141-8150.
[13] Liu, S., Pan, J., Su, Z., & Yang, M. H. (2019). Learning temporal coherence via self-supervision for gan-based video generation. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01), 5224-5231.
[14] Zhang, H., Patel, V. M. (2018). Densely connected pyramid dehazing network. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 3194-3203.
[15] Fu, X., Zhuang, P., Huang, Y., Liao, Y., Zhang, X. P., & Ding, X. (2017). Removing rain from single images via a deep detail network. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1715-1723.
[16] Wang, W., Liu, X., Zhang, H., Liu, J., & Yang, J. (2019). High-quality exposure correction of underexposed photos. IEEE Transactions on Image Processing, 28(12), 6081-6093.
[17] Liu, X., Lu, X., Wang, W., & Wu, M. (2022). Multimodal Prompt Perceiver Empower Adaptiveness, Generalizability and Fidelity for All-in-One Image Restoration.
[18] Park, S., Kim, J., Lee, J., & Kim, C. (2023). Adaptive All-in-One Image Restoration via Frequency Mining and Modulation.
[19] Li, X., Zhang, J., & Liu, D. (2022). All-in-One Image Restoration for Unknown Corruption. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
[20] Kim, H., Kwon, S., & Choi, J. (2024). Dynamic Pre-training Towards Efficient and Scalable All-in-One Image Restoration.
[21] Wang, X., Zhang, Y., Li, Z., & Zhang, Y. (2019). High-quality exposure correction of underexposed photos (Hi-Net). IEEE Transactions on Image Processing, 28(12), 6081-6093.
[22] Li, X., Liu, S., & Wang, T. (2024). Unified-Width Adaptive Dynamic Network for All-In-One Image Restoration.
[23] Li, J., Sun, J., & Tang, X. (2024). All-in-One Bad Weather Removal: Building a Search Strategy for Adverse Weather Conditions.

[24] Zhang, H., Patel, V. M. (2018). Densely connected pyramid dehazing network (DID-MDN). Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 3194-3203.
[25] Fu, X., Zhuang, P., Huang, Y., Liao, Y., Zhang, X. P., & Ding, X. (2017). Removing rain from single images via a deep detail network (All-in-One-Dehazing). Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1715-1723.
[26] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., & Yang, M. H. (2016). Single image dehazing via multi-scale convolutional neural networks (MS-Net). European Conference on Computer Vision, 154-169.
[27] Xiyu Yu, Tongliang Liu, Xinchao Wang, and Dacheng Tao. On compressing deep models by low rank and sparse decomposition. In CVPR, pages 7370–7379, 2017. 3
[28] Xiaoqi Jiao, Yichun Yin, Lifeng Shang, Xin Jiang, Xiao Chen, Linlin Li, Fang Wang, and Qun Liu. Tinybert: Distilling bert for natural language understanding. arXiv preprint arXiv:1909.10351, 2019. 3
[29] Qipeng Guo, Xipeng Qiu, Pengfei Liu, Yunfan Shao, Xiangyang Xue, and Zheng Zhang. Star-transformer. arXiv preprint arXiv:1902.09113, 2019. 3
[30] Rao Y, Zhao W, Liu B, et al. Dynamicvit: Efficient vision transformers with dynamic token sparsification[J]. Advances in neural information processing systems, 2021, 34: 13937-13949.
[31] Zamir S W, Arora A, Khan S, et al. Learning enriched features for real image restoration and enhancement[C]//European Conference on Computer Vision. Springer, Cham, 2020: 492-511.
[32] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. IEEE Transactions on image processing, 2007, 16(8): 2080-2095.
[33] Dong X, Bao J, Chen D, et al. Cswin transformer: A general vision transformer backbone with cross-shaped windows[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 12124-12134.