NTIRE 2024 Restore Any Image Model (RAIM) in the Wild Challenge

Jie Liang, Radu Timofte, Qiaosi Yi, Shuaizheng Liu, Lingchen Sun, Rongyuan Wu, Xindong Zhang, Hui Zeng, Lei Zhang, Yibin Huang, Shuai Liu, Yongqiang Li, Chaoyu Feng, Xiaotao Wang, Lei Lei, Yuxiang Chen, Xiangyu Chen, Qiubo Chen, Fengyu Sun, Mengying Cui, Jiaxu Chen, Zhenyu Hu, Jingyun Liu, Wenzhuo Ma, Ce Wang, Hanyou Zheng, Wanjie Sun, Zhenzhong Chen, Ziwei Luo, Fredrik K. Gustafsson, Zheng Zhao, Jens Sjölund, Thomas B. Schön, Xiong Dun, Pengzhou Ji, Yujie Xing, Xuquan Wang, Zhanshan Wang, Xinbin Cheng, Jun Xiao, Chenhang He, Xiuyuan Wang, Zhi-Song Liu, Zimeng Miao, Zhicun Yin, Ming Liu, Wangmeng Zuo, Shuai Li

Abstract

In this paper, we review the NTIRE 2024 challenge on Restore Any Image Model (RAIM) in the Wild. The RAIM challenge constructed a benchmark for image restoration in the wild, including real-world images with/without reference ground truth in various scenarios from real applications. The participants were required to restore the real-captured images from complex and unknown degradation, where generative perceptual quality and fidelity are desired in the restoration result. The challenge consisted of two tasks. Task one employed real referenced data pairs, where quantitative evaluation is available. Task two used unpaired images, and a comprehensive user study was conducted. The challenge attracted more than 200 registrations, where 39 of them submitted results with more than 400 submissions. Top-ranked methods improved the state-of-the-art restoration performance and obtained unanimous recognition from all 18 judges. The proposed datasets are available at https://drive.google.com/file/d/1DqbxUoiUqkAIkExu3jZAqoElr_nu1IXb/view?usp=sharing *and the homepage of this challenge is at* https://codalab.lisn.upsaclay.fr/competitions/17632.

1. Introduction

Image restoration, aiming at recovering high-quality images from their low-quality counterparts, is one of the most popular low-level vision tasks in the research community. However, there has been a large gap between **A**cademic research and **I**ndustrial application for a long time. For example, the image signal processing (ISP) systems on digital cameras always face mixed and complex degradations, yet most methods in academic research are designed and evaluated based on simulated and limited degradation. How to design and train a model that can be generalized to practical applications is a challenging yet highly valuable problem.

The deep learning techniques have significantly advanced the performance of image restoration. Recently, generative adversarial networks show good performance in approximating distributions of real photos in image restoration tasks, while the large-scale pre-trained generative diffusion models have provided powerful priors to further improve the quality of image restoration outputs.

This challenge aims to provide a platform for industrial and academic participants to test and evaluate their algorithms and models on real-world imaging scenarios, bridging the gap between academic research and practical photography. The objectives of this RAIM challenge are:

- Construct a benchmark for image restoration in the wild, including real-world images with/without reference ground-truth in various scenarios and objective/subjective evaluation methods;
- Promote the research and development of RAIMs with strong generalization performance to images in the wild.

This challenge is one of the NTIRE 2024 Workshop associated challenges on: dense and non-homogeneous dehazing [1], night photography rendering [2], blind compressed image enhancement [14], shadow removal [11], efficient super resolution [10], image super resolution ($\times 4$) [4], light field image super-resolution [13], stereo image super-resolution [12], HR depth from images of specular and transparent surfaces [15], bracketing image restora-

Jie Liang, Radu Timofte, Qiaosi Yi, Shuaizheng Liu, Lingchen Sun, Rongyuan Wu, Xindong Zhang, Hui Zeng and Lei Zhang are the organizers of the NTIRE 2024 challenge, and other authors are the participants.

The Appendix lists the authors' teams and affiliations.

NTIRE 2024 website: https://cvlai.net/ntire/2024/

tion and enhancement [16], portrait quality assessment [3], quality assessment for AI-generated content [8], restore any image model (RAIM) in the wild [7], RAW image super-resolution [5], short-form UGC video quality assessment [6], low light enhancement [9], and RAW burst alignment and ISP challenge.

2. NTIRE 2024 RAIM Challenge

2.1. Training Data

In this challenge, participants can train their models using any data they can collect and any pre-trained models they can reach.

2.2. Validation and Test Data

To facilitate the design and development of RAIM by participants, we provide two types of validation and test data: paired data with reference ground truth (R-GT), and unpaired data. All data is available now at https://drive.google.com/file/d/1DqbxUoiUqkAIkExu3jZAqoElr_nu1IXb/view?usp=sharing.

2.2.1 Paired Data with R-GT

To facilitate the model validation, we first provide some paired data in the following scenarios, where both the input low-quality image and the high-quality R-GT can be collected. Examples can be found in Figure 1.

Image denoising. Compromising with the size and cost, the photosensitivity of imaging sensors, especially on mobile phone products, is limited. Meanwhile, the illumination of shooting scenes can be poor, especially in low-light imaging. Image denoising is a very fundamental requirement for image restoration.

Image super-resolution. The focal length of mobile phone cameras is limited, making it hard to meet the needs of continuous and ultra-long magnification zoom. Therefore, mobile cameras will be equipped with digital zoom algorithms, namely super-resolution algorithms. Out-of-focus restoration. The autofocus (AF) algorithm cannot guarantee 100% focus accuracy. In fleeting moments of excitement, such as blowing birthday candles, fireworks, etc, the image restoration algorithms are expected to remedy slight defocusing shots.

Motion deblur. Due to the limitations in aperture size and sensor capability, mobile phone cameras face a trade-off between shutter time and motion blur. A longer shutter may enhance the noise reduction performance, but it is prone to motion blur when encountering foreground object motion or handheld motion. An effective motion deblur algorithm is demanded.

The combinations of the above. When capturing a real-world photo in the wild, the above issues are usually triggered simultaneously by several factors, such as motion blur in high magnification super-resolution, out-of-focus in low light environments, etc. When multiple problems appear in the image, a strong model is needed to solve them jointly. In this challenge, we use these data to calculate the full-reference metrics to partially measure the effectiveness of the algorithms and screen the top performers in the early stage.

2.2.2 Data without R-GT

In many practical scenarios, the R-GT is very difficult to collect, and the image restoration performance is hard, if not possible, to be measured by full-reference metrics. In this challenge, we also provide the data with the following commonly encountered issues in practice. Examples can be found in Figure 2.

Smoothed details and textures. Limited by hardware and on-chip computing power, images captured by mobile phone cameras often face a trade-off between noise/artifacts reduction and details/texture richness, impacting the visual quality. However, due to the lack of effective quantitative measures, evaluation can only be done through subjective observation.

Text stroke adhesion in super-resolution. In the telephoto mode (e.g., equivalent focal length larger than 230mm), shooting small text from a distance is an important yet highly challenging task. Text stroke adhesion, or super-resolution errors (i.e., presenting wrong characters), will greatly deteriorate the user experience.

High light edge and color artifacts. The optical system of mobile phones is limited, and prone to purple edges, green edges, halos, and fake textures in high-light areas. This problem occurs frequently in reflective scenes, backlight scenes, night scenes, etc., which greatly affects the user's perception.

Low-frequency color noise/blocks/bandings. The low SNR of the input in mobile phone cameras demands a heavy denoising algorithm to output a clean image. However, due to factors such as computing power and storage, the bit-width of the ISP system is limited. When transitioning from the linear domain to the nonlinear domain, visual color noise, blocks and bandings often appear.

High-frequency aliasing and Moire pattern. Due to the resolution of imaging sensors, the Moire pattern can appear at specific distances and frequencies. Although users have certain expectations (or understanding) about the appearance of Moire patterns, they still hope to reduce the probability and severity of Moire patterns without affecting the clarity of the image.

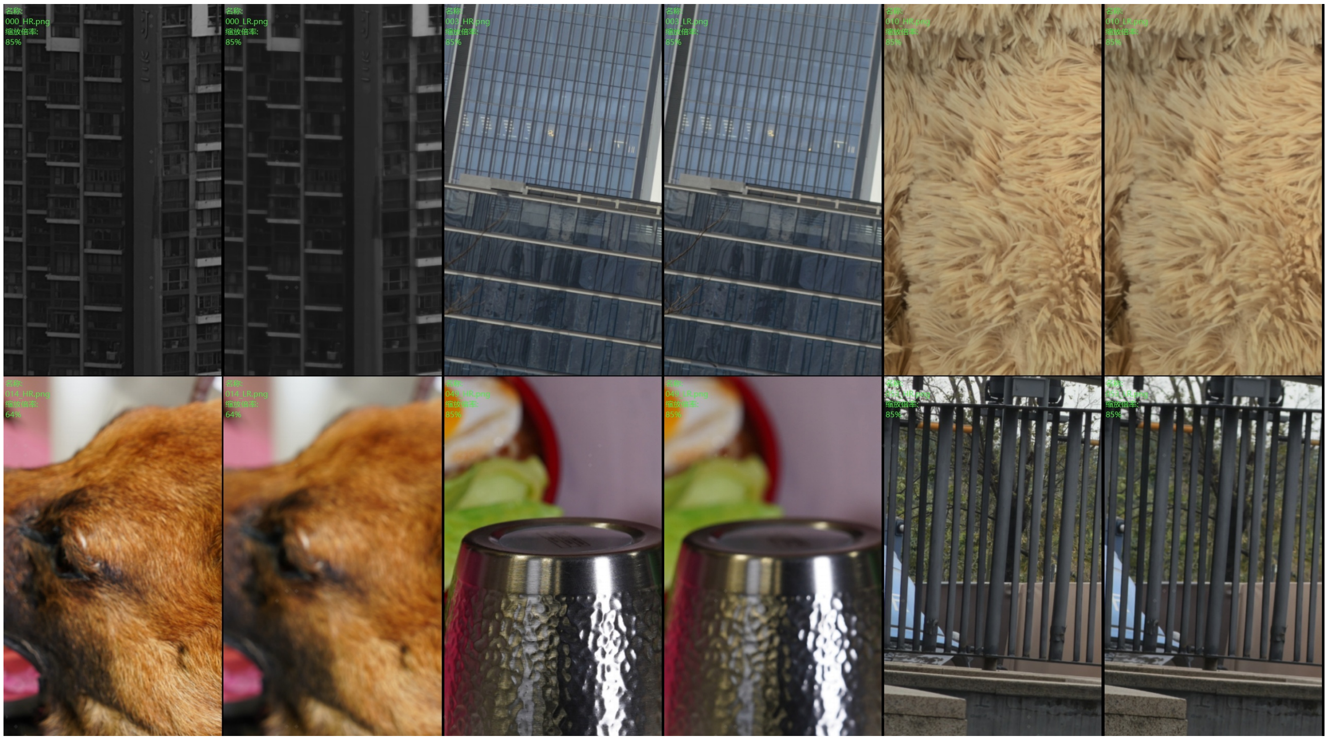

Figure 1. Example data pairs we have provided.

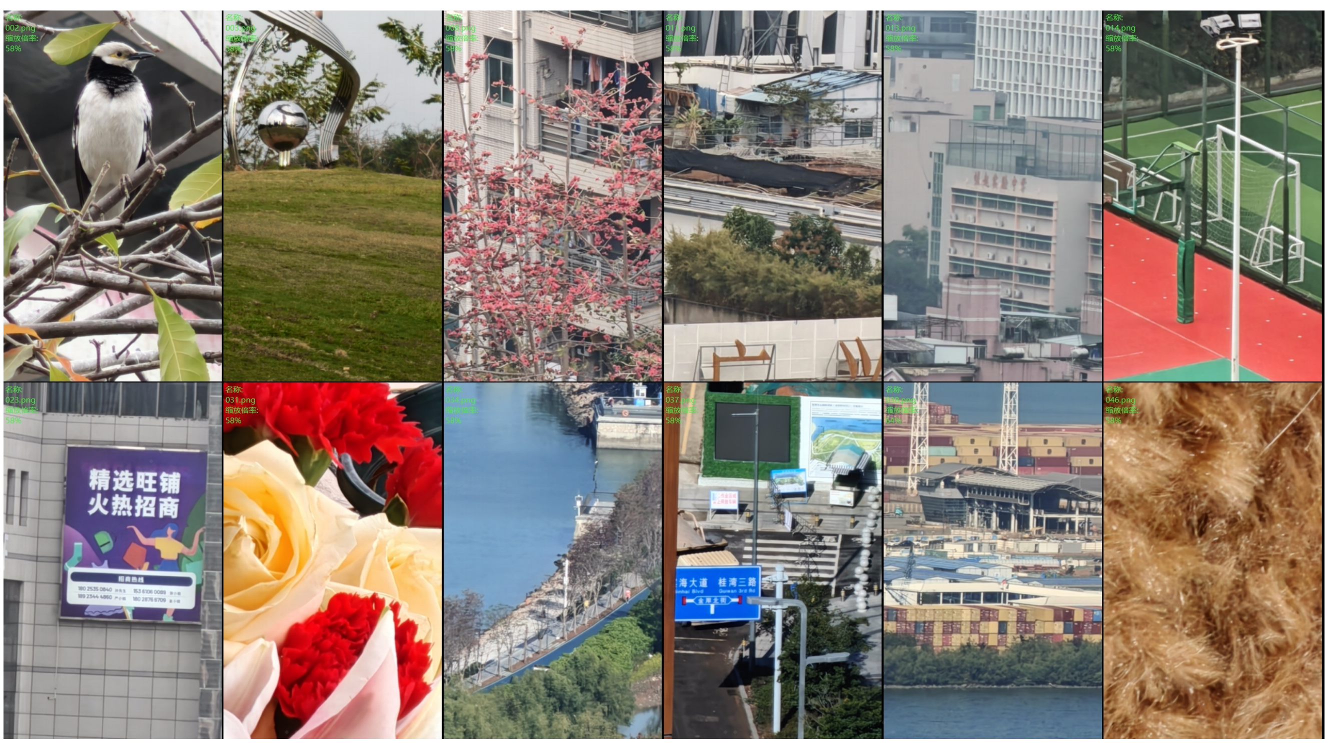

Figure 2. Example input images we have provided without R-GT.

2.3. Evaluation Measures

We evaluate the effectiveness of the models with both quantitative measures and subjective evaluation.

2.3.1 Quantitative Measure

Following prior arts, we employed the PSNR, SSIM, LPIPS, DISTS and NIQE measures to evaluate the models quantitatively by using the data with R-GT. The evaluation score is computed as follows:
$SCORE = 20\times\frac{PSNR}{50}+15\times\frac{SSIM-0.5}{0.5}+20\times\frac{1-LPIPS}{0.4}+40\times\frac{1-DISTS}{0.3}+30\times\frac{1-NIQE}{10}$.

2.3.2 Subjective Evaluation

For the test data without R-GT, we judged the perceptual quality of the restored results by visual inspection. Specifically, we invited 18 experienced practitioners and conducted a comprehensive user study. The following features were considered in the evaluation:

Textures and details. The restored image should have fine and natural textures and details.

Noise. Noise, especially color noise, should be eliminated. Some luminance noise can be kept to avoid oversmoothness in flat areas.

Artifacts. Various artifacts, such as worm-like artifacts, color blocks, bandings, over-sharpening, and so on, should be reduced as much as possible.

Fidelity. The restored image should be loyal to the given input. More details have been discussed during the competition with all participants by referencing specific images and model outputs.

2.4. Phases

2.4.1 Phase 1: Model Design and Tuning

In this phase, participants can analyze the given data and tune their models accordingly. We provided:

- 100 pairs of paired data (i.e., input with R-GT), which can be used to tune the models based on the quantitative measures.
- 100 images without R-GT, which can be used to tune the model according to visual perception.

2.4.2 Phase 2: Online Feedback

In this phase, participants can upload their results and get official feedback. We provide:

- the input low-quality images of another 100 pairs of paired data.

Only the low-quality input images are provided, and the participants can upload the restoration results to the server and get the quantitative scores online. Users can also upload their results of the images without the R-GT provided in Phase 1 to seek feedback. The organizers will provide feedback to a couple of teams that get the highest quantitative scores of the images with R-GT.

2.4.3 Phase 3: Final Evaluation

In this phase, we provide:

- another 50 images without R-GT for subjective evaluation.

In this phase, we select the top ten teams according to the quantitative score of the 100 images with R-GT in Phase 2, and then arrange a comprehensive user study on their results of the above 50 images without R-GT. The final ranks of the ten teams will be decided based on both the quantitative scores and the subjective user study, with the weight being 40% and 60%, respectively.

2.5. Awards

The following awards of this challenge are provided:

- One first-class award (i.e., the champion) with a cash prize of **US$1000**;
- Two second-class awards with cash prizes of **US$500 each**;
- Three third-class awards with cash prizes of **US$200 each**.

2.6. Important Dates

- 2024.02.07: Released data of phase 1. Phase 1 began;
- 2024.02.25: Released data of phase 2. Phase 2 began;
- 2024.03.17: Released data of phase 3. Phase 3 began;
- 2024.03.22: Phase 3 results submission deadline;
- 2024.03.27: Final rank announced.

3. Challenge Results

In total, the challenge received 200+ registrations, where 39 of them have submitted results in phase 2 with 400+ submissions. In phase 3, we invited the top 12 teams in phase 2 and received 9 valid submissions. Brief illustrations of the methods from participating teams are provided in Section 4, while the team information is provided in Section 6.

3.1. Phase 2: quantitative comparison on paired data with R-GT

In phase 2, we got submissions from 30+ teams, where the quantitative results of top-ranked teams are shown in Table 1. The evaluation measure is described in Section 2.3.1.

The script of this measure is available at https://drive.google.com/file/d/1Q1CvlbGo-WOgqya5GulS5eYIi2Rgcj5l/view?usp=sharing.

Table 1. Result of phases 2 and 3, as well as final scores and ranks. We only show teams that participated in phase 3.

Team	Score in Phase 2	Score in Phase 3	Final Score	Rank
MiAlgo	79.13	57	91.65	1
Xhs-IAG	81.96	47	82.07	2
So Elegant	79.69	46	80.09	3
IIP_IR	80.03	14	45.94	4
DACLIP-IR	78.65	9	40.03	5
TongJi-IPOE	72.99	11	39.91	6
ImagePhoneix	78.93	4	34.79	7
HIT-IIL	69.80	1	27.92	8

3.2. Phase 3: qualitative comparison on unpaired data

In stage three, we invite 18 low-level vision-related students/engineers, who are required to select the top three results of each of the 50 samples. They follow a unified principle as demonstrated in Section 2.3.2 and the feedback to individual participants. The team information is hidden and the results are randomly shuffled to make fair comparisons. By checking the results of each scorer, we found their opinions are similar so the results are valid. The final score S_{final} is calculated by

$$S_{final} = 0.4 \times S_2 + 0.6 \times S_3^n, \tag{1}$$

where S_2 indicates the score in phase 2 and S_3^n denotes the normalized score in phase 3.

For calculating S_3^n, we first calculate the score in phase 3, *i.e.*, S_3, where the team is rewarded with 3 points when selected to be top 1, 2 points for the top 2, and 1 point for the top 3. The scores are averaged by 18. Then, we calculate S_3^n by

$$S_{3\,teami}^n = \frac{S_{3teami} - min(S_3)}{max(S_3) - min(S_3)}. \tag{2}$$

We then show some example visual comparisons in Figures 3, 4, 5 and 6. All visual results in phase 3 are available at https://drive.google.com/file/d/1_vxF2s-WRm59F8Vn1nquE7q4R2zHZTmm/view?usp=sharing.

4. Teams and Methods

Due to the space limitation, we describe the participating teams and their proposed methods in the **supplementary material**.

5. Acknowledgments

This work was partially supported by the Humboldt Foundation. We thank the NTIRE 2024 sponsors: Meta Reality Labs, OPPO, KuaiShou, Huawei and University of Würzburg (Computer Vision Lab).

6. Appendix: Teams and affiliations

NTIRE 2024 Team

Challenge:
NTIRE 2024 Restore Any Image Model (RAIM) in the Wild

Organizers:
Jie Liang[1] (liang27jie@gmail.com)
Qiaosi Yi[1,2] (qiaosi.yi@connect.polyu.hk)
Shuaizheng Liu[1,2] (shuaizhengliu21@gmail.com)
Lingchen Sun[1,2] (ling-chen.sun@connect.polyu.hk)
Xindong Zhang[1] (17901410r@connect.polyu.hk)
Hui Zeng[1] (cshzeng@gmail.com)
Prof. Lei Zhang[1,2] (cslzhang@comp.polyu.edu.hk)
Prof. Radu Timofte[3] (radu.timofte@uni-wuerzburg.de)
Affiliations:
[1] OPPO Research Institute
[2] The Hong Kong Polytechnic University
[3] Computer Vision Lab, University of Würzburg, Germany

Team MiAlgo
Members:
Yibin Huang (huangyibin@xiaomi.com)
Shuai Liu, Yongqiang Li, Chaoyu Feng, Xiaotao Wang, Lei Lei
Affiliations:
Xiaomi Inc., China

Team Xhs-IAG
Members:
Yuxiang Chen[1] (chenyuxiang@xiaohongshu.com)
Xiangyu Chen[2,3], Qiubo Chen[1]
Affiliations:
[1] Xiaohongshu,
[2] University of Macau,
[3] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

Team So Elegant
Members:
Jiaxu Chen
Fengyu Sun (sunfengyu@s.upc.edu.cn), Mengying Cui
Affiliations:
China University of Petroleum (East China)

Team IIP_IR
Members:
Zhenyu Hu (zhenyuhu@whu.edu.cn),
Jingyun Liu, Wenzhuo Ma, Ce Wang, Hanyou Zheng, Wanjie Sun, Zhenzhong Chen
Affiliations:
School of Remote Sensing and Information Engineering, Wuhan University

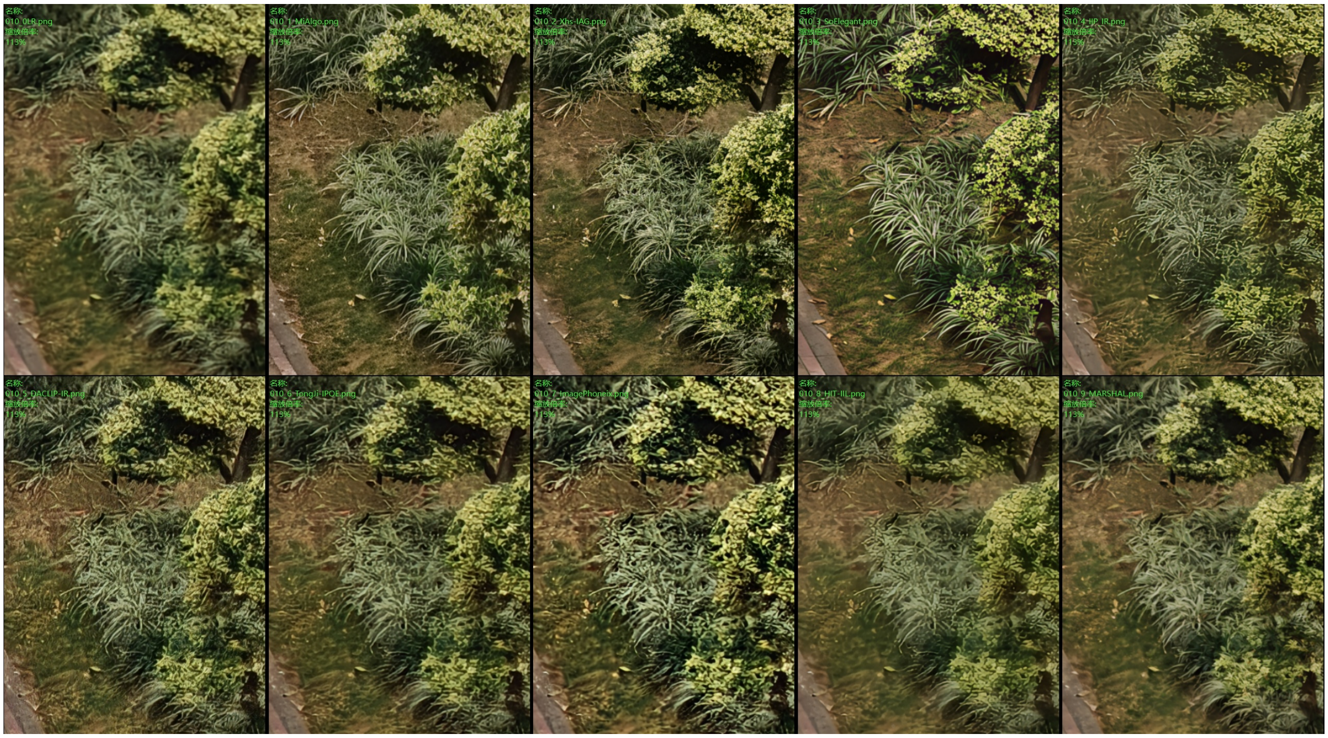

Figure 3. Visual comparisons of the input LR image (top left) and results from participated teams (others) in phase 3.

Figure 4. Visual comparisons of the input LR image (top left) and results from participated teams (others) in phase 3.

Figure 5. Visual comparisons of the input LR image (top left) and results from participated teams (others) in phase 3.

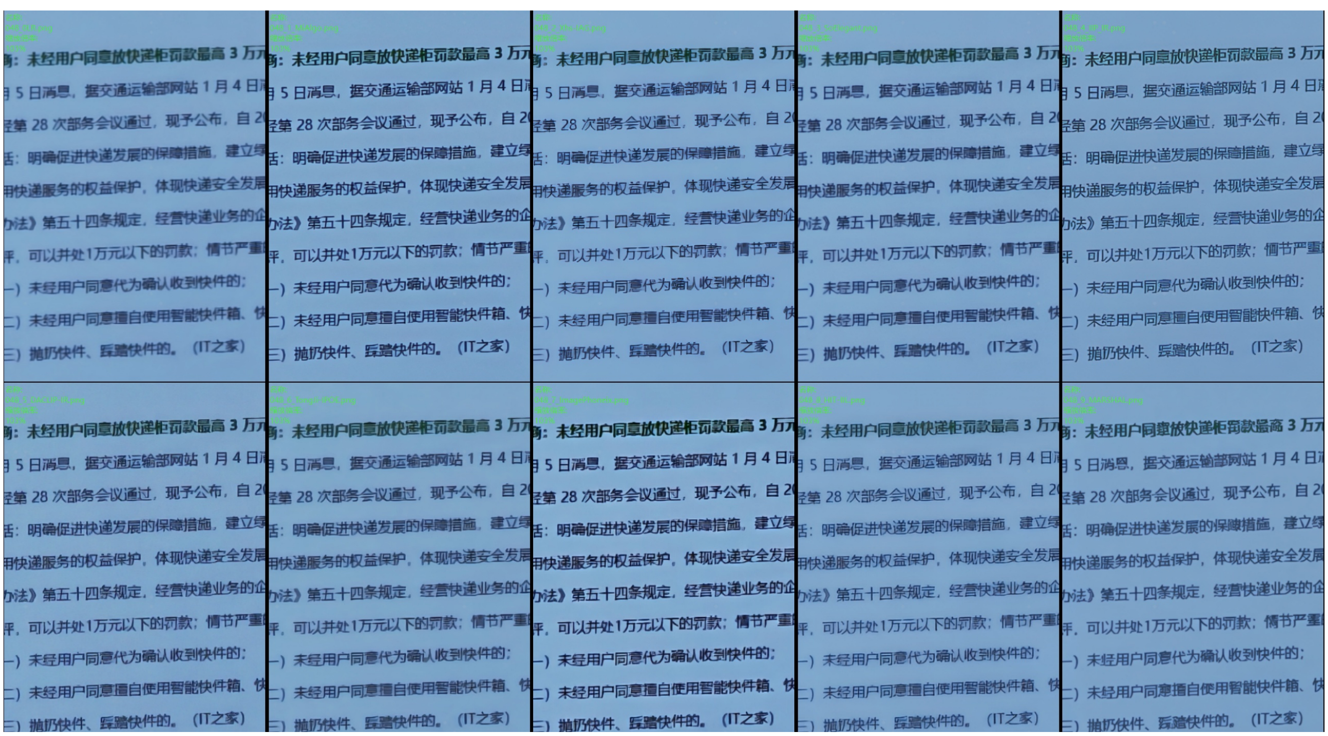

Figure 6. Visual comparisons of the input LR image (top left) and results from participated teams (others) in phase 3.

Team DACLIP-IR
Members:
Ziwei Luo (ziwei.luo@it.uu.se)
Fredrik K. Gustafsson, Zheng Zhao, Jens Sjölund, Thomas B. Schön
Affiliations:
Department of Information Technology, Uppsala University

Team TongJi-IPOE
Members:
Xiong Dun
Pengzhou Ji (jipengzhoudrew@163.com), Yujie Xing, Xuquan Wang, Zhanshan Wang, Xinbin Cheng
Affiliations:
Institute of Precision Optical Engineering, School of Physics Science and Engineering, Tongji University

Team ImagePhoneix
Members:
Jun Xiao[1] (jun.xiao@connect.polyu.hk)
Chenhang He[1], Xiuyuan Wang[1], Zhi-Song Liu[2]
Affiliations:
[1] The Hong Kong Polytechnic University
[2] Lappeenranta-Lahti University of Technology

Team HIT-IIL
Members:
Zimeng Miao (2214177602@qq.com)
Zhicun Yin, Ming Liu, Wangmeng Zuo
Affiliations:
School of Computer Science and Technology, Harbin Institute of Technology

Team MARSHAL
Members:
Rongyuan Wu (rong-yuan.wu@connect.polyu.hk)
Shuai Li
Affiliations:
The Hong Kong Polytechnic University

References

[1] Cosmin Ancuti, Codruta O Ancuti, Florin-Alexandru Vasluianu, Radu Timofte, et al. NTIRE 2024 dense and non-homogeneous dehazing challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[2] Nikola Banić, Egor Ershov, Artyom Panshin, Oleg Karasev, Sergey Korchagin, Shepelev Lev, Alexandr Startsev, Daniil Vladimirov, Ekaterina Zaychenkova, Dmitrii R Iarchuk, Maria Efimova, Radu Timofte, Arseniy Terekhin, et al. NTIRE 2024 challenge on night photography rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[3] Nicolas Chahine, Marcos V. Conde, Sira Ferradans, Radu Timofte, et al. Deep portrait quality assessment. a NTIRE 2024 challenge survey. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[4] Zheng Chen, Zongwei WU, Eduard Sebastian Zamfir, Kai Zhang, Yulun Zhang, Radu Timofte, Xiaokang Yang, et al. NTIRE 2024 challenge on image super-resolution (×4): Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[5] Marcos V. Conde, Florin-Alexandru Vasluianu, Radu Timofte, et al. Deep raw image super-resolution. a NTIRE 2024 challenge survey. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[6] Xin Li, Kun Yuan, Yajing Pei, Yiting Lu, Ming Sun, Chao Zhou, Zhibo Chen, Radu Timofte, et al. NTIRE 2024 challenge on short-form UGC video quality assessment: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[7] Jie Liang, Qiaosi Yi, Shuaizheng Liu, Lingchen Sun, Rongyuan Wu, Xindong Zhang, Hui Zeng, Radu Timofte, Lei Zhang, et al. NTIRE 2024 restore any image model (RAIM) in the wild challenge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[8] Xiaohong Liu, Xiongkuo Min, Guangtao Zhai, Chunyi Li, Tengchuan Kou, Wei Sun, Haoning Wu, Yixuan Gao, Yuqin Cao, Zicheng Zhang, Xiele Wu, Radu Timofte, et al. NTIRE 2024 quality assessment of AI-generated content challenge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[9] Xiaoning Liu, Zongwei WU, Ao Li, Florin-Alexandru Vasluianu, Yulun Zhang, Shuhang Gu, Le Zhang, Ce Zhu, Radu Timofte, et al. NTIRE 2024 challenge on low light image enhancement: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[10] Bin Ren, Yawei Li, Nancy Mehta, Radu Timofte, et al. The ninth NTIRE 2024 efficient super-resolution challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[11] Florin-Alexandru Vasluianu, Tim Seizinger, Zhuyun Zhou, Zongwei WU, Cailian Chen, Radu Timofte, et al. NTIRE 2024 image shadow removal challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[12] Longguang Wang, Yulan Guo, Juncheng Li, Hongda Liu, Yang Zhao, Yingqian Wang, Zhi Jin, Shuhang Gu, Radu Timofte, et al. NTIRE 2024 challenge on stereo image super-resolution: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[13] Yingqian Wang, Zhengyu Liang, Qianyu Chen, Longguang Wang, Jungang Yang, Radu Timofte, Yulan Guo, et al. NTIRE 2024 challenge on light field image super-resolution: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[14] Ren Yang, Radu Timofte, et al. NTIRE 2024 challenge on blind enhancement of compressed image: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[15] Pierluigi Zama Ramirez, Fabio Tosi, Luigi Di Stefano, Radu Timofte, Alex Costanzino, Matteo Poggi, et al. NTIRE 2024 challenge on HR depth from images of specular and transparent surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[16] Zhilu Zhang, Shuohao Zhang, Renlong Wu, Wangmeng Zuo, Radu Timofte, et al. NTIRE 2024 challenge on bracketing image restoration and enhancement: Datasets, methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.