

NTIRE 2024 Challenge on Night Photography Rendering

Egor Ershov Artyom Panshin Oleg Karasev Sergey Korchagin Shepelev Lev
 Alexandr Startsev Daniil Vladimirov Ekaterina Zaychenkova Nikola Banić
 Dmitrii R Iarchuk Maria Efimova Radu Timofte Arseniy Terekhin

Abstract

This paper presents a review of the NTIRE 2024 challenge on night photography rendering. The goal of the challenge was to find solutions that process raw camera images taken in nighttime conditions, and thereby produce a photo-quality output images in the standard RGB (sRGB) space. Unlike the previous year’s competition, the challenge images were collected with a mobile phone and the speed of algorithms was also measured alongside the quality of their output. To evaluate the results, a sufficient number of viewers were asked to assess the visual quality of the proposed solutions, considering the subjective nature of the task. There were 2 nominations: quality and efficiency. Top 5 solutions in terms of output quality were sorted by evaluation time (see Fig. 1). The top ranking participants’ solutions effectively represent the state-of-the-art in nighttime photography rendering. More results can be found at <https://nightimaging.org>.

1. Introduction

In-camera processing is widely used to process raw images obtained directly from the sensor into photographs encoded in a standard color space, such as sRGB. The main objective of this processing is to produce images that are visually pleasing and that simultaneously realistically represent the captured scene. However, nighttime photography presents unique challenges that are not typically encountered in daytime photography. For example, while a single illuminant can often be assumed for daytime images, there are typically multiple illuminants present in nighttime scenes and these can be significantly different. This makes it difficult to determine which illuminant(s) should be primarily taken into account during scene color correction. Moreover, common photo-finishing strategies used for daytime images may not be appropriate for night images due to differences in lighting conditions.

Additionally, commonly used image metrics such as SSIM [39], LPIPS [44], or MetaQA [46]) do not appropriately assess the quality of night images. Furthermore, there

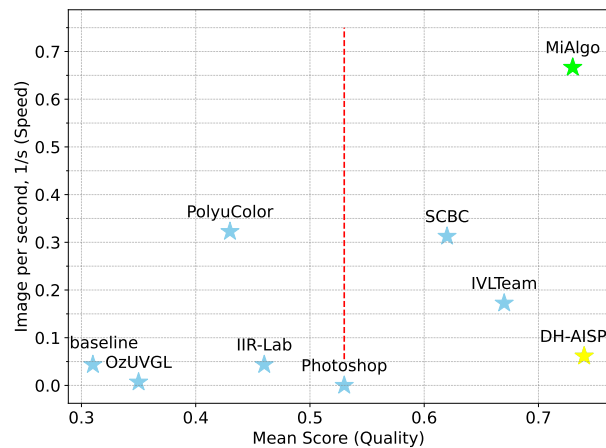


Figure 1. Illustration of final leader board scores and speed of algorithms. Red line separates top-5 quality solutions that were rearranged by inference speed. Winner of efficiency challenge is colored green, team that provided solution with best quality is colored yellow.

is a dearth of published research focused specifically on image processing for night photography, resulting in fewer established “best practices” than for daytime photography. Having all this in mind, the main objective of this challenge is, similarly to the previous ones, to further encourage research into image processing techniques for night photography. The following sections provide a detailed description of the NTIRE challenge and the solutions proposed by the participating teams: Section 2 lays out the challenge setup, Section 3 describes the obtained results, Section 4 presents a discussion on the obtained results, and Section 5 lists the teams with their members and their affiliations.

2. Challenge

In the challenge, the participating teams were required to develop automated solutions capable of producing visually appealing nighttime images. The challenge diverged from its predecessors as it incorporated raw images captured by mobile phones. This shift is prompted by the prevalent de-



Figure 2. Example of the impact of high noise level and (color) vignetting on rendering a nighttime image when using only the baseline pipeline.

mand for processing such images and the distinctive features they possess in contrast to the conventional camera images employed in prior challenges.

In addition to the already mentioned challenges, the competition participants faced the following ones as well:

1. **High noise level.** Mobile phone camera pixel is 100 times smaller than high-quality DSLR camera.
2. **Vignetting and color vignetting.** Modern mobile lens systems not only suffer from traditional vignetting problems, but also experience significant color degradation from the center to the periphery.
3. **Computational resources restriction.** Due to the limited computational resources on board of mobile phone, an extra ranking criteria has been introduced.

The impact of the first and second challenge can be better understood by taking a look at Fig. 2 where the baseline pipeline has been used for image rendering and the mentioned problems were not taken into account.

As for the third challenge, the final ranking specifically prioritized the execution time of the top-performing solutions from the regular ranking, which is traditionally based solely on average quality as determined by the mean opinion score.

The teams' objective was to submit the corresponding rendered sRGB images obtained by the using the respective teams' solutions. Given the subjective nature of this task, the submissions were evaluated using mean opinion scores assigned by observers who were presented with pairs of two different renderings of the same scene and who then had to choose the rendering that they deemed visually more appealing.

This challenge is one of the NTIRE 2024 Workshop ¹ associated challenges on: dense and non-homogeneous de-hazing [3], blind compressed image enhancement [40], shadow removal [35], efficient super resolution [32], image super resolution ($\times 4$) [13], light field image super-resolution [38], stereo image super-resolution [37], HR depth from images of specular and transparent surfaces [43], bracketing image restoration and enhancement [45], portrait quality assessment [10], quality assessment for AI-generated content [27], restore any image model (RAIM) in the wild [25], RAW image super-resolution [14], short-form UGC video quality assessment [24], low light enhancement [28], and RAW burst alignment and ISP challenge.

2.1. Challenge Data

The raw images of night outdoor and low-light indoor scenes were captured using Huawei Mate 40 Pro and encoded in 16-bit PNG files with additional meta-data provided in JSON files. The challenge started with an initial 200 images provided to participants for algorithm development and testing. Three validation datasets with 125 images each were provided during the challenge. Moreover, the calibration white low-light scenes were provided to participants as well. A baseline code was provided to emulate the basic in-camera rendering as a starting point.

The majority of the images were taken in the urban area and this category can be divided into images of open and closed spaces. There are also indoor images. These can all be designated as categories 1, 2, and 3. The first validation dataset contains more images of open spaces, with a ratio of images between these categories being 100/25/0. The second validation dataset already includes indoor images, with a ratio of 108/6/11. The third dataset has an even greater number of indoor images, with a ratio of 89/3/33.

The final dataset is focused on category 1 images, which dominate the data, but it includes other categories as well. The ratio of categories in it is 39/7/4. The dataset was made publicly available.²

2.2. Evaluation

The evaluation consisted of three validation checkpoints during the contest and a final evaluation to determine the winners. Mean opinion scores were obtained using Yandex Tasks (a service similar to Mechanical Turk) for the checkpoints and final evaluation. Yandex Tasks users ranked their preferred solutions in a forced-choice manner with a question: "Which image is more pleasant?". The answer options were: "left", "right" or "they are the same". To ensure basic quality control, all Yandex Tasks users who chose "left" of "right" for a pair of same images have been banned,

¹<https://cvlai.net/ntire/2024/>

²<https://zenodo.org/records/10931755>

while all their previous answers have been declined. It is worth noting that in our setup Yandex Tasks mainly relied on users from Eastern Europe to perform the image ranking. As a result, there may be a cultural bias in terms of the preferred image aesthetics by the users. All solutions have been anonymized to guarantee unbiased results.

As recently shown [18], the Yandex Tasks platform generates stable and reproducible results. Also, we chose only 10% best voters and conducted filtering of votes in order to exclude any fraud scenarios.

As we implement the pairwise approach for our experiments, the raw data obtained for the pairwise comparisons of the style renderings, i.e., versions of a single image can be denoted as A_{ijt} where i and j denote the i -th and j -th image versions, i.e., team renderings that were shown to the voters and t enumerates the participants who evaluated the (i, j) image pair. A_{ijt} is equal to 1 if the t -th participant evaluated the i -th image version to be better than the j -th and 0 otherwise. Then, the score for an image is calculated as $S_i = \frac{1}{NT} \sum_{j=1}^N \sum_{t=1}^T A_{ijt}$, where N is total number of evaluated solutions and T is total number of voters. The final score for team's solution is obtained as $S = \frac{1}{N} \sum_{i=1}^N S_i$.

To measure speed, all submitted solutions were executed on the same computer with the following technical specifications:

- **CPU:** Intel(R) Core(TM) i7-4790 CPU @ 3.60GHz
- **RAM:** 16.0 GB
- **GPU:** MSI GeForce RTX 2060 12Gb

Only the actual image processing time was measured, excluding image loading and saving.

During each validation checkpoint, 125 new test images were provided, and each participating team was allowed to submit up to two distinct solution image sets, each consisting of exactly 125 images. The purpose of having three validation sets was to allow participants to test different solutions' behavior and receive feedback on their solution's quality.

For the final submission, only one solution was allowed, and 50 hidden test images were used for the final validation. The user study images were generated using the code provided by the participants by means of Docker. Only open and reproducible results were accepted. The top 5 solutions according to Yandex Tasks were further sorted by performance speed.

In the context of self-assessment by people, algorithms from the MIALGO and DH-AISP teams proved to be the best. The downside of the second solution is the excessive lighting of dark images, as well as a very long processing time, namely 11 times longer than the first algorithm. An example of final solutions' images is presented in Fig. 3.

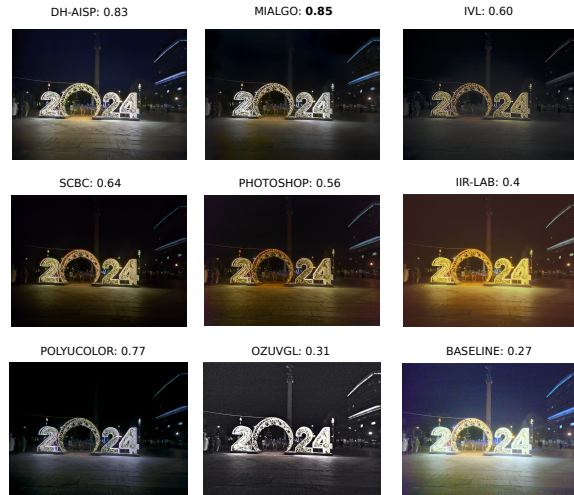


Figure 3. One scene from the final validation set and all results provided by participants. The best quality image is in bold.

3. Results

The section presents the ranking results obtained using the Yandex Tasks service and performance evaluation.

3.1. People's Choice and Discussion

Table 1 provides the ranking of the mean opinion reported by Yandex Tasks users for the different teams' final submissions (quality challenge). Table 2 provides ranking for efficiency challenge. Table 2 is made up from Table 1 with sorting top 5 solutions by speed performance.

Rank	Team	Mean Score	Time, s
1	DH-AISP	0.74	16.3
2	MiAlgo	0.73	1.5
3	IVLTeam	0.67	5.8
4	SCBC	0.62	3.2
5	Manual enhancement	0.53	∞
6	IIR-Lab	0.46	23.0
7	PolyuColor	0.43	3.1
8	OzUVGL	0.35	144.8
9	baseline	0.31	23.0

Table 1. People's choice ranking results. Quality leaderboard.

This year's competitors have presented a diverse range of solutions that produce visually appealing images.

3.2. Teams' solutions

This section contains brief description of the submitted solutions by participants. More detailed description and

Rank	Team	Mean Score	Time, s
1	MiAlgo	0.73	1.5
2	SCBC	0.62	3.2
3	IVLTeam	0.67	5.8
4	DH-AISP	0.74	16.3
5	Manual enhancement	0.53	∞

Table 2. Top 5 solutions are sorted by inference time. Efficiency leaderboard.

code will be available on the challenge repository³.

3.2.1 Baseline

In this year’s challenge, two baseline methods were given to the participants to use: a simple classic ISP and manual image enhancement. The simple classic pipeline involved debayering using linear interpolation, white balancing by using the Gray World method [9], a mean throughout train dataset matrix for CST, and a standard transform from XYZ to sRGB. This pipeline was also provided as a baseline for the participants⁴.

To enhance the images manually, we employed the Adobe Camera RAW application and invited non-professional photographers to participate. Each image was corrected individually within a short span of 3 to 5 minutes. The corrections comprised of adjusting the temperature to cool down the image, adding a violet tint, increasing the brightness via exposure adjustment, enhancing the contrast, reducing the highlights, brightening the shadows, and reducing the whites. Finally, the built-in noise reduction and color mixer were used to correct the hue and intensity of red, orange, and yellow (and sometimes blue and purple).

3.2.2 MiAlgo

We made improvements based on Deep-FlexISP [26], and the overall pipeline is shown in Figure 4.

First, we pre-process the input image, including packing raw to 4 channels, down-sampling the image (to 768×1024), correcting the black level (adjusting value based on the noise profile), normalizing (to 0-1), and correcting lens shading (adjusting shading mask based on the provided calibration image and the noise profile). It is worth noting that we downsample the image to the required size at the beginning, which ensures that our whole processing is as fast as possible.

Next, we use the pre-trained Unet [26, 33] for raw domain denoising, and use the pre-trained FC4 [20] and meta info (as_shot_neutral) for white balance parameters estimation.

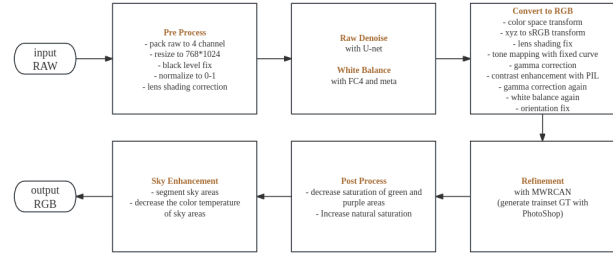


Figure 4. Overall pipeline of MiAlgo team.

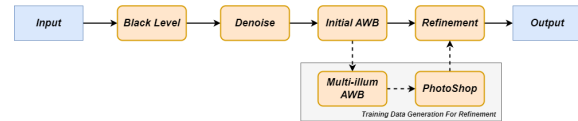


Figure 5. SCBC team pipeline scheme.

Then we convert the raw image to RGB image with color space transform (fixed matrix), XYZ to sRGB transform (fixed matrix), lens shading fix (increase shading of dark scenes and decrease saturation), tone mapping (fixed curve), gamma correction (fixed parameter), contrast enhancement (python-pillow-autocontrast [1]), gamma correction again (fixed parameter), white balance correction again (grayness index [31]), and orientation fix.

We use neural networks for the refinement of the RGB images. The model structure is based on MWRCAN [21]. The ground truth of the training data is generated using Photoshop.

Then we post-process the image, decreasing the saturation of green and purple areas (uncommon and weird colors at night), and increasing the natural saturation of the whole image.

Finally, we train a segmentation model [15, 36] to segment the sky areas and decrease the color temperature, which makes the sky more blue and cold.

3.2.3 SCBC

SCBC team proposed a multi-stage ISP pipeline as shown in Fig. 5, including pre-processing, Raw Image Denoising and Image refinement with Color Correction. For raw image denoising, pretrained models are used [2]. After having the initial images, we process the images with a multi-illumination color constancy method to solve the complex white balance introduced by artificial light source. The model is trained on RGB image with multi-illum lighting condition [23]. The result are further manually adjusted and serve as the target image for refinement network training.

³<https://github.com/createcolor/nightimaging24>

⁴available at challenge repository

3.2.4 IVLTeam

The solution proposed by IVLTeam is illustrated in Fig. 6. It relies on conventional image processing techniques and consists of five stages. It is based on previous works [47, 48], where the parameters of several processing steps are adapted to the challenge images and requirements.

The **first stage** works in the RAW domain and consists of five steps: black and white levels image normalization, raw demosaicing, image resizing to the target resolution, white balancing using the illuminant provided in the metadata, and conversion from the camera-sensor color space to the sRGB color space.

The **second stage** consists of a denoising operation using the Non-local means algorithm [8]. The intensity of denoising is proportional to the noise standard deviation estimated in the image using the method in [16]. Stronger denoising is applied to the color channels than to the luma channel to effectively remove color noise while preserving image details and edges.

The **third stage** is a sequence of algorithms that enhance image contrast by manipulating the histogram distribution. First, the Local Contrast Correction (LCC) algorithm in [30] is applied. As this process tends to decrease the overall contrast and saturation, the next step consists of a contrast and saturation enhancement using the approach proposed in [34]. Then, three steps to improve the image color appearance are applied. The first one adjusts the contrast by stretching the pixel values by a β factor around their mean. The second one is the application of the S-curve defined in [22], where the center of the curve is set to zero, transforming the function into a gamma-like operation. The third one consists of a histogram stretching operation that increases the dynamic range and improves the overall contrast. After these operations, an extra conditional contrast correction operation, consisting of an additional S-curve or gamma correction, is applied depending on the mean value of the histogram. This improves visibility for very dark images, and restores the mood of nighttime scenes when the processed image is too bright.

The **fourth** and **fifth stages** perform sharpening and white balancing, respectively. Unsharp masking is used to enhance image details, which may have been flattened by the denoising operation in the second stage. White balancing is performed using the Grayness Index (GI) algorithm [31] to further reduce color casts. As GI is sensitive to noise, the illuminant is estimated on a blurred version of the image and later applied to the sharp version.

3.2.5 DH-AISP

Our main goal is to develop a technology for creating realistic and visually pleasing photographs of night scenes. By considering the data quality and modality, we outline the

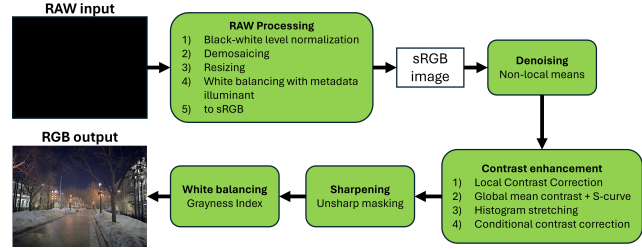


Figure 6. IVLTeam pipeline schema.



Figure 7. NISP-net

network structure in Fig. 7. It contains three parts, involving the data pre-processing module, RAW to RGB module, and HDR module, which are elaborated in the following parts.

Data pre-processing. This module contains four steps: black-and-white balance correction, automatic white balancing, adaptive gain adjustment, and binning. The adaptive gain adjustment module generates a set of underexposed, normal exposed, and overexposed images by calculating the mean of the raw data. Binning is a method that sacrifices resolution to achieve higher signal-to-noise ratio. With these modules, we can obtain a corrected RAW data, which is a necessary step to achieve normal and better results.

RAW to RGB. We have trained a U-Net [33] structure to obtain images under different exposure gains.

HDR. Night images often have multiple illuminants, and light distribution of contents varies with the locations. Consequently, learning a unified light distribution in the existing methods produce undesired results. In this way, we propose to a controllable parameter to adjust the output of raw2rgb module, which allows us to produce the underexposed and overexposed candidates with different light distribution. Following that, an exposure fusion model based on U-Net to adaptively learn the fusion weights, and generates an image with satisfactory overall brightness. Finally, the CCM algorithm is introduced to further optimize the color distribution of the final output.

3.2.6 IIR-Lab

Our main goal is to develop a technology for creating realistic and visually pleasing photographs of night scenes. By considering the data quality and requirements, we construct a new ISP pipeline, outlined in Fig. 8. It contains three parts, involving the RAW denoising module, raw to RGB module, and color enhancement module, which are elabo-



Figure 8. Illustration of our proposed ISP framework.

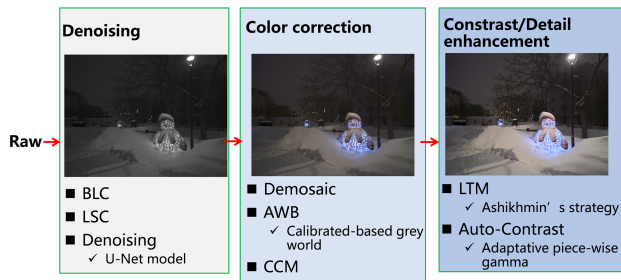


Figure 9. PolyuColor team pipeline schema.

rated in the following parts.

RAW denoising. To capture more details with better visual effects, a higher gain is commonly used for the sensor, but simple gain multiplication causes noise degradation. To solve this problem, we first processed the Bayer domain denoising module, which utilizes NAF [12].

RAW to RGB. Since the provided json file contains information such as AWB and CCM, this part of our implementation is done using traditional methods.

Color Enhancement. Night images often have complex noise, and multiple illuminants, and light distribution of contents varies with the locations. Therefore, we collect a set of nighttime images containing different light sources and render the corresponding ground truth using PS tools. We propose a UNet-based rendering network that enables real-time rendering on most devices. The final results show that our rendering network replaces the traditional modules of local and global Tonemapping, color enhancement, and sharpening.

3.2.7 PolyuColor

Nighttime imaging is challenging due to its low signal-to-noise ratio (SNR) and complicated lighting conditions. However, the three key elements that significantly impact image quality include denoising, color correction, and contrast/detail enhancement. In our method, a DNN-based approach is used for denoising, and traditional ISP algorithms are used for the other tasks, as shown in Fig. 9. We aim to maintain superior image quality while minimizing computational resource consumption.

Denoising. For RAW images, we initially apply black/white level (BLC) normalization and lens shading correction (LSC) as outlined in the baseline [17]. Following this, a U-Net-based [11] framework is employed to develop a model

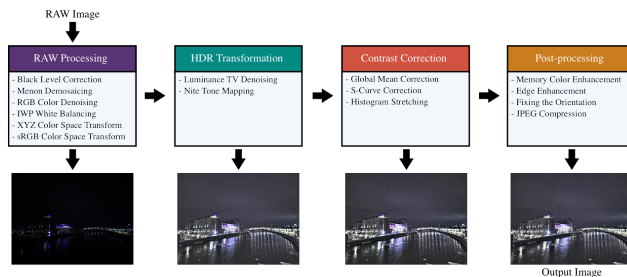


Figure 10. OzU-VGL pipeline.

trained on the SID dataset [11], with the long-exposure images serving as the ground truth, supplemented by a synthetic noise model for comprehensive training.

Color correction. We follow the methods used by the baseline for demosaicing and the color correction matrix (CCM). For auto white balance (AWB), our approach utilizes a calibrated-based grey world method [9]. By manually calibrating a set of white points and clustering them to approximate the sources on the Planckian locus, we define the correlated color temperature range within a circle, which is similar to the method in [41,42]. The image is segmented into several patches, with the grey world assumption adopted to derive the white point for each patch. The final white point is the weighted average of the cluster centers based on the distance between the white point of each patch and the corresponding cluster center(s).

Contrast/Detail enhancement. After obtaining the sRGB images, we apply local tone mapping (LTM) utilizing Ashikhmin’s method [4]. A scaling operation is performed to enhance the saturation by applying unequal gains on the three channels, which is different from the original method. An adaptive piece-wise gamma adjustment is finally applied to further enhance the global contrast.

3.2.8 OzU-VGL

In our image signal processor (ISP) pipeline for this challenge, we mainly focus on reproducing the previous year’s winner solution [47] and improving it by including the improved version of the current state-of-the-art tone mapping operators (*i.e.*, Flash and Storm [6]) where image statistics dynamically adjusts their scale parameter. Moreover, we applied simple tricks to avoid color casting in bright regions and the noise in dark regions in the final output. Fig. 10 presents the proposed ISP pipeline for this challenge.

RAW Processing. The RAW input images are given in PNG format as the data, and we first applied black-level correction to the input by using the values in the metadata. We utilized the Directional Filtering algorithm [29], also referred to as *Menon*, in our pipeline for demosaicing the RAW input, offering a more sophisticated alternative to the default CFA interpolation. Removing the noise on color

channels in images with low grayscale mean value can be critical for avoiding undesired color casting on bright regions, so we applied simple Gaussian denoising to the color channels in YCbCr color space. Then, we integrated the random subsampling-based White Patch algorithm [5] into our pipeline to mitigate the yellowish effect induced by the Gray World algorithm, particularly evident in images with low grayscale mean values. We have implemented thresholds for the first and third diagonal values of the illuminant estimation matrix to avoid the occurrence of a greenish tint in bright areas caused by randomness. It is worth noting that, due to the randomness, it may be hard to reproduce the exact outputs for consecutive runs. We have submitted illuminance estimation matrices found by the final run during the final test to ensure reproducibility. The subsequent procedure involves converting the image from the raw-RGB to the sRGB. This is achieved by initially transforming raw-RGB to the XYZ color space, then converting XYZ to the sRGB using the specified color matrix tailored to Huawei Mate 40 Pro. At this point, we point out that some images produce salt-paper-like noise in dark regions, dramatically amplified by any tone mapping operator in the final output.

HDR Transformation. To remove this noise factor, we applied total variation denoising [19] to the luminance channel in YCbCr color space before applying HDR transformation. Apart from [47], we refrained from using local or conditional contrast correction operators. Instead, we introduced a new tone mapping technique named *Nite*, specifically designed to enhance Storm’s performance in low-light conditions. The operator employs adaptive adjustment of the scaling parameter α within the Naka-Rushton equation, utilizing image statistics.

Contrast correction. To adjust the overall image color appearance, we applied three operations: (1) the global mean contrast operation enhances image contrast by scaling the values of RGB channels around their respective mean. (2) The S-curve correction applies to each RGB channel the S-curve proposed in [22]. (3) The histogram stretching operation stretches the image histogram, thereby augmenting the dynamic range and enhancing overall contrast.

Post-processing. The memory color enhancement algorithm [7] is employed to balance the colors of the sky, grass, and other specific hues in the contrast-corrected images. Subsequently, unsharp masking is applied to enhance edge sharpness, followed by aligning the image orientations based on metadata. Finally, the results are resized to the specified dimensions (*i.e.*, 1024×768) and saved in JPG format to produce the final sRGB output.

4. Discussion

A notable trend from this year’s competition is that three out of the top five solutions, including the winning one, were enhancements of the solutions from prior years. This

highlights the value of iterative improvement and underscores a continuous progression in this challenging field.

Importantly, the efficiency-winning solution of this year not only outperformed its top five peers in speed but was also the fastest overall. In terms of image quality, it ranked second, narrowly trailing behind the top contender, which represents a desirable combination of qualities for use in production.

Furthermore, it’s intriguing to observe that two solutions employing deep learning techniques surpassed a traditional method in computational efficiency. This may denote a shift towards the readiness of employing such deep learning models in image processing pipelines of mobile phones.

5. Teams and Affiliations

Team:

Organizers

Members:

Artyom Panshin¹, Egor Ershov², Maria Efimova¹, Dima Yarchuk¹, Nikola Banić³, Radu Timofte⁴, Oleg Karasev¹, Sergey Korchagin¹, Arseniy Terekhin¹, Alexandr Startsev¹, Daniil Vladimirov¹, Ekaterina Zaychenkova¹, Shepelev Lev¹

Affiliations:

¹: Institute for Information Transmission Problems (Kharkevich institute), Russia

¹: Artificial Intelligence Research Institute (AIRI), Russia

³: Gideon Brothers, Croatia

⁴: Swiss Federal Institute of Technology in Zurich, Switzerland and University of Würzburg, Germany

Team:

PolyuColor

Members:

Shuwei Yue¹, Yuyang Liu¹, Minchen Wei^{1,2}

Affiliations:

¹: Color, Imaging, and Illumination Laboratory, The Hong Kong Polytechnic University, Kowloon, Hong Kong

²: Color, Imaging, and Metaverse Research Center, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Team:

SCBC

Members:

Lu Xu, Chao Zhang, Yasi Wang

Affiliations:

SRC-Beijing

Team:

OzU-VGL

Members:

Furkan Kınlı^{1,3,4}, Doğa Yılmaz^{1,3}, Barış Özcan^{2,3}, Furkan Kırac^{1,3}

Affiliations:

¹: Vision and Graphics Laboratory, Özyeğin University, Türkiye

²: Department of Computer Engineering, Bahçeşehir University, Türkiye

³: Fishency Innovation, Norway

⁴: T-Fashion Inc., Türkiye & Canada

Team:

MiAlgo

Members:

Shuai Liu, Jingyuan Xiao, Chaoyu Feng, Hao Wang, Guangqi Shao, Yuqian Zhang, Yibin Huang, Wei Luo, Liming Wang, Xiaotao Wang, Lei Lei

Affiliations:

Xiaomi Inc., China

Team:

IVLTeam

Members:

Simone Zini¹, Claudio Rota¹, Marco Buzzelli¹, Simone Bianco¹, Raimondo Schettini¹

Affiliations:

¹: University of Milano - Bicocca, Italy

Team:

DH-AISP

Members:

Jin Guo, Tianli Liu, Mohao Wu, Ben Shao

Affiliations:

Zhejiang Dahua Technology, China

Team:

IIR-Lab

Members:

Qirui Yang^{1,2}, Xianghui Li¹, Qihua Cheng², Fangpu Zhang¹, Zhiqiang Xu², Jingyu Yang¹, Huanjing Yue¹

Affiliations:

¹: School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

²: Shenzhen MicroBT Electronics Technology Co., Ltd, Shenzhen 518000, China

6. Acknowledgements

This work was supported by Artificial Intelligence Research Institute (AIRI). 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).

References

- [1] Pillow. [GitHub - python-pillow / Pillow : Python Imaging Library \(Fork\)](https://github.com/python-pillow/Pillow). 4

- [2] Abdelrahman Abdelhamed, Mahmoud Afifi, Radu Timofte, and Michael S Brown. NTIRE 2020 Challenge on Real Image Denoising: Dataset, Methods and Results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020. 4
- [3] 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
- [4] Michael Ashikhmin. A Tone Mapping Algorithm for High Contrast Images. pages 145–156, 01 2002. 6
- [5] Nikola Banić and Sven Lončarić. Improving the white patch method by subsampling. In *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014. 7
- [6] Nikola Banić and Sven Lončarić. Flash and Storm: Fast and Highly Practical Tone Mapping based on Naka-Rushton Equation. In *VISIGRAPP (4: VISAPP)*, pages 47–53, 2018. 6
- [7] Karen M Braun. Memory color enhancement algorithm. In *International Congress of Imaging Science*, 2006. 7
- [8] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2. IEEE, 2005. 5
- [9] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26, 1980. 4, 6
- [10] 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. 2
- [11] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to See in the Dark. In *CVPR*, 2018. 6
- [12] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. Apr 2022. 6
- [13] Zheng Chen, Zongwei WU, Eduard Sebastian Zamfir, Kai Zhang, Yulun Zhang, Radu Timofte, Xiaokang Yang, et al. NTIRE 2024 challenge on image super-resolution (x4): Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024. 2
- [14] 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. 2
- [15] MMsegmentation Contributors. MMsegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark. <https://github.com/open-mmlab/msegmentation>, 2020. 4
- [16] David L Donoho and Iain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994. 5

- [17] Egor Ershov et al. NTIRE 2022 Challenge on Night Photography Rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2022. 6
- [18] Egor Ershov, Artyom Panshin, Ivan Ermakov, Nikola Banić, Alex Savchik, and Simone Bianco. Reliability and Stability of Mean Opinion Score for Image Aesthetic Quality Assessment Obtained Through Crowdsourcing. In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VISAPP*, pages 365–372. INSTICC, SciTePress, 2024. 3
- [19] Pascal Getreuer. Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman. *Image Processing On Line*, 2:74–95, 2012. 7
- [20] Yuanming Hu, Baoyuan Wang, and Stephen Lin. FC4: Fully convolutional color constancy with confidence-weighted pooling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 4
- [21] Andrey Ignatov, Radu Timofte, Zhilu Zhang, Ming Liu, Haolin Wang, Wangmeng Zuo, Jiawei Zhang, Ruimao Zhang, Zhanglin Peng, Sijie Ren, et al. AIM 2020 challenge on learned image signal processing pipeline. In *European Conference on Computer Vision*. Springer, 2020. 4
- [22] Sing Bing Kang, Ashish Kapoor, and Dani Lischinski. Personalization of image enhancement. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010. 5, 7
- [23] Dongyoung Kim, Jinwoo Kim, Seonghyeon Nam, Dongwoo Lee, Yeonkyung Lee, Nahyup Kang, Hyong-Euk Lee, ByungIn Yoo, Jae-Joon Han, and Seon Joo Kim. Large scale multi-illuminant (Ismi) dataset for developing white balance algorithm under mixed illumination. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2410–2419, 2021. 4
- [24] 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. 2
- [25] 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. 2
- [26] Shuai Liu, Chaoyu Feng, Xiaotao Wang, Hao Wang, Ran Zhu, Yongqiang Li, and Lei Lei. Deep-FlexISP: A Three-Stage Framework for Night Photography Rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022. 4
- [27] 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. 2
- [28] 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. 2
- [29] Daniele Menon, Stefano Andriani, and Giancarlo Calvagno. Demosaicing with directional filtering and a posteriori decision. *IEEE Transactions on Image Processing*, 16:132–141, 2007. 6
- [30] Nathan Moroney. Local color correction using non-linear masking. In *Color and Imaging Conference*, volume 2000.1. Society for Imaging Science and Technology, 2000. 5
- [31] Yanlin Qian, Joni-Kristian Kamarainen, Jarno Nikkanen, and Jiri Matas. On finding gray pixels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019. 4, 5
- [32] 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. 2
- [33] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, 2015. 4, 5
- [34] Raimondo Schettini, Francesca Gasparini, Silvia Corchs, Fabrizio Marini, Alessandro Capra, and Alfio Castorina. Contrast image correction method. *Journal of Electronic Imaging*, 19(2):023005, 2010. 5
- [35] 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. 2
- [36] Qiang Wan, Zilong Huang, Jiachen Lu, Gang Yu, and Li Zhang. Seaformer: Squeeze-enhanced axial transformer for mobile semantic segmentation. *arXiv preprint arXiv:2301.13156*, 2023. 4
- [37] 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. 2
- [38] 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. 2
- [39] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. 1

- [40] 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. 2
- [41] Shuwei Yue and Minchen Wei. Color constancy from a pure color view. *J. Opt. Soc. Am. A*, 40(3):602–610, Mar 2023. 6
- [42] Shuwei Yue and Minchen Wei. Effective cross-sensor color constancy using a dual-mapping strategy. *J. Opt. Soc. Am. A*, 41(2):329–337, Feb 2024. 6
- [43] 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. 2
- [44] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 1
- [45] 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. 2
- [46] Hancheng Zhu, Leida Li, Jinjian Wu, Weisheng Dong, and Guangming Shi. MetaIQA: Deep Meta-Learning for No-Reference Image Quality Assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020. 1
- [47] Simone Zini, Claudio Rota, Marco Buzzelli, Simone Bianco, and Raimondo Schettini. Back to the future: a night photography rendering ISP without deep learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2023. 5, 6, 7
- [48] Simone Zini, Claudio Rota, Marco Buzzelli, Simone Bianco, and Raimondo Schettini. Shallow Camera Pipeline for Night Photography Enhancement. In *International Conference on Image Analysis and Processing*, pages 51–61. Springer, 2023. 5