Cam-Bench: A Benchmark for Image-based Camera Parameter Estimation

Quanhong Peng*
School of Computer, Qinghai Normal University
Key Laboratory of Tibetan Information Processing
Xining, China
202333331056@stu.qhnu.edu.cn

Dan Zhang*†
School of Computer, Qinghai Normal University
Key Laboratory of Tibetan Information Processing
Xining, China
danz@mail.bnu.edu.cn

Dong Zhao
School of Computer, Qinghai Normal University
Key Laboratory of Tibetan Information Processing
Xining, China
zhaodong@mail.sdu.edu.cn

Jianpeng Zhang
School of Computer, Qinghai Normal University
Xining, China
202333331015@stu.qhnu.edu.cn

Meihua Song
School of Computer, Qinghai Normal University
Xining, China
202333331038@stu.qhnu.edu.cn

Chenlei Lv†
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, China
chenleilv@mail.bnu.edu.cn

Abstract

For camera-based image capturing, the impact of exposure or camera parameters (ISO sensitivity, shutter speed, and aperture F-number) on imaging quality is decisive. Such parameters interact in a coupled manner during the imaging process to determine the exposure quality and the degree of blur in a photograph. Naturally, decoupling such parameters from images holds significant value for applications like image quality assessment and illumination optimization. However, there has been no systematic research dedicated to this topic. In this paper, we propose a new benchmark, Cam-Bench, for estimating camera parameters on images directly. It collects an image dataset Cam-10K with various indoor scenes and accurate labels of camera parameters. Based on Cam-10K, we propose a camera parameter estimation network to decouple and regress recorded exposure information. To the best of our knowledge, Cam-Bench is the first benchmark for camera parameter estimation. Experiments demonstrate that it can enhance the performance of various downstream applications.The source code has been made publicly available at: https://github.com/pengquanhong/CamBench.

CCS Concepts

• **Computing methodologies** → **Computer vision**; **Image processing**.

Keywords

Camera Parameter Estimation, Lighting Enhancement, Exposure

*Co-first authors;†Corresponding authors.

ACM ISBN 979-8-4007-2035-2/2025/10
https://doi.org/10.1145/3746027.3754875

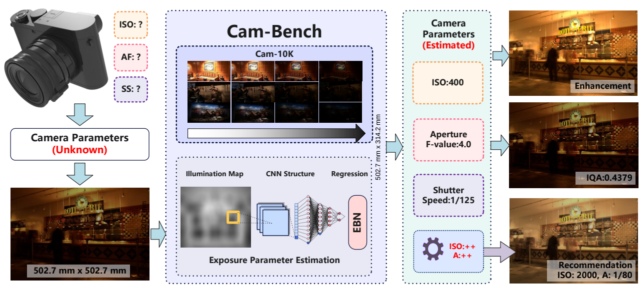

Figure 1: Illustration of Cam-Bench. With a collected dataset Cam-10k, Cam-Bench implements a camera parameter estimation that supports related downstream tasks (lighting enhancement, IQA, and auto-parameter recommendation).

ACM Reference Format:
Quanhong Peng*, Dan Zhang*†, Dong Zhao, Jianpeng Zhang, Meihua Song, and Chenlei Lv†. 2025. Cam-Bench: A Benchmark for Image-based Camera Parameter Estimation. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25), October 27–31, 2025, Dublin, Ireland.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3746027.3754875

1 Introduction

For computational photography, optimizing camera parameters (ISO sensitivity, shutter speed, and aperture F-number) is important to achieve high-quality visual outputs. The ISO sensitivity directly affects brightness and noise levels, whereas the aperture F-number controls light intake and depth of field, and the shutter speed determines exposure duration and motion blur effects. These parameters collectively define the exposure quality of captured image. Some suboptimal combinations can introduce irreversible artifacts such as sensor noise (high ISO), compressed dynamic range (small aperture F-numbers), or motion-induced blur (slow shutter speeds). Although traditional photography techniques are based on empirical knowledge and most camera devices offer automated parameter recommendations, their effectiveness still has certain limitations, particularly in dynamic scenes where manual exposure and contrast adjustments remain indispensable even for professionals. Therefore, conducting research on the impact of camera parameters on image quality is essential, both in practical applications and theoretical frameworks.

The accurate prediction of intrinsic camera parameters remains a fundamental challenge in computational photography, as it essentially requires decoupling camera parameters from the captured image. Currently, most research works focus on illumination modeling and optimization, including low-light recovery [22] and HDR reconstruction [15]. Such solutions attempt to improve exposure and image perceptual quality, typically using parametric curves or convolutional neural operators at the pixel level. However, they are essentially data-driven post-processing operations that lack explicit modeling of the physical imaging process, making it difficult to accurately quantify the independent effects of individual decoupled parameters. Naturally, the lack of research on the estimation of decoupled parameters reduces the controllability of the evaluation and optimization of image exposure quality.

To address this challenge, we propose Cam-Bench, a novel benchmark for image-based camera parameter estimation, as shown in Figure 1. The benchmark comprises: 1) **Cam-10K**, the first large-scale image dataset with accurate camera parameter priors covering complex illumination scenarios (e.g., museums and laboratories); 2) a dedicated **parameter estimation framework** built upon a Retinex-inspired Convolutional Neural Network (CNN) architecture. The framework utilizes a dual-stream architecture to extract scene-specific features (illumination gradients, exposure characteristics) and learn camera-specific parameter mappings. A cross-exposure fusion module enables joint pixel-level optimization of parameter-content adaptation. Crucially, we introduce a physics-derived latent metric, the exposure brightness number (EBN) to enable dynamic parameter adaptation across diverse illumination scenarios, which estimates camera parameters with semantic consistency. To our knowledge, Cam-Bench represents the pioneering effort for independent camera parameter estimation. As demonstrated in Figure 1, it establishes new baselines for illumination-related downstream tasks, including lighting-aware IQA, adaptive exposure optimization, etc. The contribution can be concluded as:

- We construct an image dataset **Cam-10K** that is a comprehensive indoor(e.g., museums and laboratories) dataset specifically designed for precise camera parameter estimation It contains 10,000 real-world scenes captured under controlled illumination conditions (overexposure, underexposure, and normal exposure). Each image is annotated with accurate photographic metadata including ISO sensitivity, aperture (F-number), and shutter speed, providing enough samples for data-driven camera parameter analysis.
- We present a camera parameter estimation network that is a pioneering approach to the best of our knowledge. The proposed dual-branch architecture employs a CNN backbone to implement a novel cross-exposure fusion module for joint parameter-content optimization. A key innovation is our physics-derived Exposure Brightness Number (EBN), which enables dynamic parameter adaptation across diverse illumination conditions.
- We illustrate the significant improvement of Cam-Bench in a set of downstream tasks, including illumination transfer, parameter-adaptive lighting optimization, exposure quality assessment, and auto-parameter recommendation for photo capturing. Comprehensive experiments validate that Cam-Bench establishes new state-of-the-art performance in both parameter estimation accuracy and related applications. For lighting optimization task, PSNR-based IQA can be increased by 24.41% with Cam-Bench.

2 Related Work

Estimating camera parameters from an image can be understood as a decoupling process of the illumination. Although such direct estimation is not a common research topic, studies on illumination modeling and exposure optimization are highly relevant. In this section, we review studies on illumination modeling and exposure optimization.

Statistical Scheme. Traditional image illumination enhancement techniques use statistical approaches such as histogram equalization (HE) to optimize illumination conditions through dynamic expansion of the image's intensity range. These techniques operate through histogram redistribution through either global [3, 9] or local [14, 19] levels. Global histogram equalization methods uniformly adjust the pixel intensities across the entire image to achieve statistically balanced results. However, such methods are limited in flexibility for illumination optimization due to the requirement of maintaining global consistency [17, 18]. In contrast, local strategies offer greater flexibility as local variants apply adaptive processing within image sub-regions to preserve local contrast [4]. The trade-off is that they cannot guarantee semantic consistency in images, potentially leading to erroneous color bleeding. These approaches perform coupled analysis of image color and illumination conditions, but cannot support inverse parameter parsing.

Retinex-based Solution. An alternative paradigm builds upon the Retinex theory [13], which formulates image enhancement through intrinsic decomposition into reflectance (scene albedo) and illumination components. Assuming reflectance remains lighting-invariant, these methods focus on accurate illumination estimation. Recent advances include naturalness-preserving framework for non-uniform illumination [21], joint reflectance-illumination estimation via weighted variational modeling [5], structure-guided illumination refinement [7], and noise-aware optimization formulation [16]. Compared to HE-based approaches, these methods provide physically-grounded enhancement by explicitly modeling illumination effects, thereby better maintaining natural appearance while suppressing artifacts.

Deep Feature-based Scheme. Recent deep learning methods for illumination optimization have evolved along two main directions. The first direction focuses on enhancing image spatial representation through advanced network architectures, exemplified by Retinex-based CNN[23, 28, 29] that decompose images into illumination and reflectance components, as well as GAN-based approaches like EnlightenGAN [11] that enable unpaired training through adversarial learning. Although these approaches have shown promise, they often struggle with noise amplification and color distortion in challenging lighting conditions. Alternative solutions like Zero-DCE [6] reformulate the problem as image-specific curve estimation, eliminating the need for reference data but still operating within the limitations of processed RGB images.

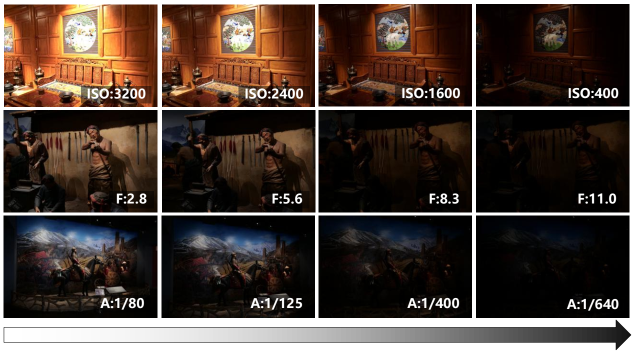

Figure 2: Some collected instances according to the "fix-two-predict-one" strategy. First row: keeping shutter speed & aperture and changing ISO sensitivity; second row: keeping shutter speed & ISO sensitivity and changing aperture; third row: keeping ISO sensitivity & aperture and changing shutter speed. In each row, smooth changing in image brightness is clearly observable with parameter value increasing.

The second direction addresses these constraints by directly processing RAW sensor data. Classical work [2] demonstrated that end-to-end neural networks could reconstruct high-quality images while avoiding traditional image signal processing(ISP) pipeline constraints. Subsequent advances like ParamISP [12] achieved more accurate modeling by explicitly incorporating camera parameters such as exposure time and ISO sensitivity. Parallel developments in reversible ISP architectures [25, 27] have improved bidirectional mapping between RAW and sRGB spaces, though most methods still neglect the critical role of camera-specific parameters in the imaging process.

In this paper, we fully draw upon the research achievements of previous works and propose Cam-Bench that directly extracts camera parameters from images. It is equivalent to obtain a decoupled representation of illumination model through feature analysis from the image. Driven by collected dataset with prior parameters, the estimation becomes compatible with semantic information. In following parts, we introduce implementation details of Cam-Bench.

3 Dataset Collection

To implement the camera parameter estimation, we collect a new dataset Cam-10K as the data part of Cam-Bench. It takes more than 10,000 images with comprehensive illuminations to advance computational photography research, particularly in camera parameter estimation, enhancement, and quality assessment. Each sample of Cam-10K should take clear prior parameters to be the ground truth labels. It captures eight diverse scenes (e.g., museums, libraries) under varying natural/artificial lighting, enabling systematic study of how camera parameters (ISO, shutter speed(A), aperture(F)) impact image quality. Each parameter was adjusted across 10 levels (ISO:400-3200 in 10 EV steps, F:2.8–11 with 1-stop increments, A:1/80s–1/640s in 10 logarithmic steps), generating total 10^3 unique parameter combinations per scene.

A critical challenge arises from the inherent many-to-one mapping in photographic parameter spaces: different parameter combinations (e.g., high ISO vs. long exposure) can produce radiometrically equivalent images under fixed scene illumination. It raises an issue: if we train a supervised network using GT annotations to predict camera parameters, the stability of the predicted results are inherently ambiguous. To address the issue, we employ a "fix-two-predict-one" strategy to record annotations, where two parameters are held constant to predict the third one. It allows to explicit brightness analysis, as shown in Figure 2. Cam-10K is split into two parts: 8,000 training images for parameter learning across lighting conditions, and 2,000 test images for evaluation. The test image set combines real captures, synthetic images (from lighting software), and curated MIT-Adobe FiveK [1], ensuring the generalization in practice. Based on the collected Cam-10K, we can implement camera parameter estimation based on a supervised or semi-supervised framework. The dataset stores parameter variation and scene diversity offer a robust benchmark for computational photography. The large-scale training data supports model adaptability, while the hybrid test set validates real-world applicability. By bridging controlled experiments and practical scenarios, this dataset significantly accelerates research in camera optimization and image enhancement. Next, we explain the implementation of camera parameter estimation.

4 Parameter Estimation

Based on the Cam-10k, we propose the parameter estimation framework that can be regarded as the implementation part of Cam-Bench. It firstly extracts brightness and exposure features from images, then integrates them with three core camera parameters (ISO, F and A) through a feature fusion module. This process yields an intermediate metric, the Exposure Brightness Number (EBN), which enables accurate parameter prediction. The whole architecture comprises: an improved Retinex decomposition module that separates images into reflectance (content-preserving) and illumination (brightness-encoding) maps using multi-scale decoupling and cross/self-attention mechanisms; a camera optical parameter fusion module that normalizes and nonlinearly processes ISO, F and A into a robust feature vector z, employing dropout for generalization; camera parameters are optimally predicted by minimizing the Mean Squared Error (MSE) between the Exposure Brightness Number (EBN) and the normalized feature vector z.

4.1 Network Architecture

We firstly introduce the architecture of the camera parameter estimation network. It consists of two core modules include Retinex decomposition module and feature fusion module. The Retinex decomposition module can decompose the input image into a reflectance map and an illumination map. Through multi-scale feature decoupling techniques, such maps retain semantic features and illumination information. The reflectance component preserves the material and texture features of the scene, while the illumination component encodes the brightness distribution features. The feature fusion module integrates prior camera parameters (ISO sensitivity, aperture F-number, and shutter speed) for fused feature vector generation, which supervise network training.

The two key steps are the training phase and the testing phase. The main difference between the two is that the training phase introduces the feature fusion module to constrain the training process. During the training phase, a convolutional neural network (CNN) is first used to extract features from the illumination map processed by the Retinex decomposition module, thereby accurately obtaining

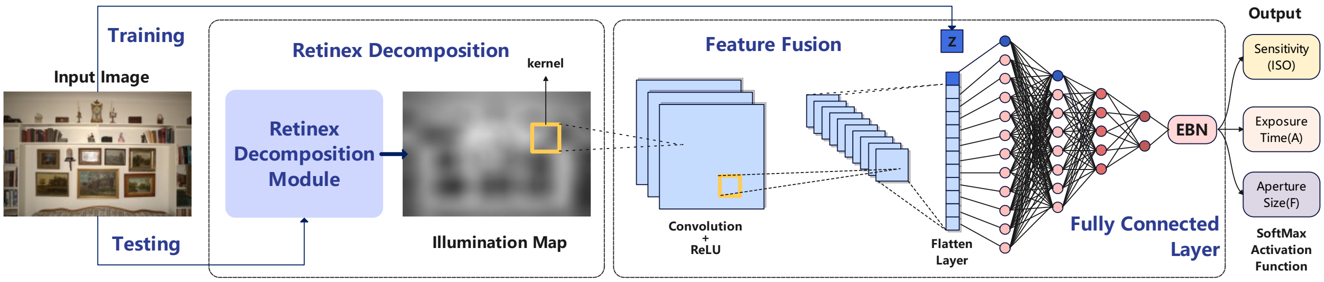

Figure 3: The pipeline of Cam-Bench for camera parameter estimation.

brightness and exposure features. Subsequently, through a fusion operation, the brightness and exposure features are combined to generate the Exposure Brightness Number (EBN).

To optimize model performance, feature separation regularization is introduced to ensure the independence and robustness of brightness and exposure features. During training, the model optimizes feature extraction by aligning the EBN with the fused camera optical features using mean squared error loss. It ensures input images with the same camera parameters produce a consistent EBN. The relationship between EBN and camera parameters can be formulated as:

$$EBN(f(x,y)) = E(ISO(f(x,y)), A(f(x,y)), F(f(x,y))), \quad (1)$$

where $EBN(\cdot)$ denotes the EBN of the image $f(x,y)$, x and y are coordinates in the image, $ISO(f(x,y))$ represents the sensor sensitivity, $A(f(x,y))$ indicates the shutter speed, and $F(f(x,y))$ stands for the aperture. The function E describes the nonlinear relationship among these four parameters. Based on the formulation, camera parameters can be accurately predicted. It not only achieves a precise correlation between the physical camera parameters and the image brightness characteristics but also lays a solid foundation for downstream tasks. The specific workflow of the estimation network is shown in Figure 3. In following parts, we introduce details of mentioned modules.

4.2 Retinex Decomposition Module

Even two images are captured with same camera parameters, pixel-level differences can still cause discrepancies between the exposure brightness index and the actual image visualization, impacting the accuracy of parameter estimation. To address the issue, we employ a Retinex decomposition module based on the Retinex theory [24]. It can be formulated as:

$$I(x,y) = R(x,y) \cdot L(x,y), \quad (2)$$

where $R(x,y)$ is the reflectance map (content image), $L(x,y)$ is the illumination map (brightness image), and operation $\cdot$ denotes element-wise multiplication. The decomposition is to decouple reflectance and illumination maps from the input image.

For decoupling, we firstly encode latent features from the input image. Then, such features are decomposed using cross-attention and self-attention mechanisms to generate a reflectance map rich in content details and an illumination map containing only brightness information [10], as shown in Figure 4. Specifically, we estimate the initial reflectance and illumination maps according to [5], that can be formulated as:

$$\tilde{L}(x) = \max_{c\in[0,\ C)} F_c(x),\ \tilde{R}(x) = F(x)/(\tilde{L}(x) + \tau), \quad (3)$$

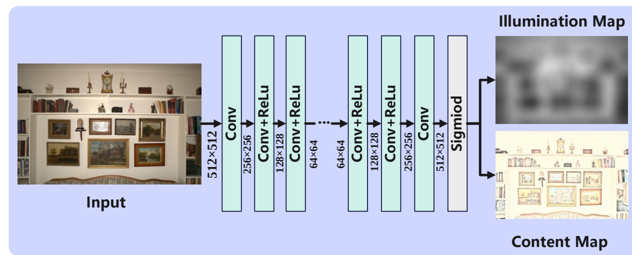

Figure 4: Schematic diagram of the retinex decomposition module for illumination and content decoupling.

where τ is a small constant to avoid division by zero, x denotes the pixel position in the image, C represents the channel set and c is channel index, $F_c(x)$ is the feature value at x for the c-th channel in the feature map output by the encoder, $\tilde{L}(x)$ is the initial estimated illumination map, which takes the maximum value of $F_c(x)$ across all channels, $F(x)$ is the spatial feature vector, containing information from C; $\tilde{R}(x)$ is the initial estimated reflectance map.

To refine the estimated reflectance and illumination maps, two separate branches are employed. Convolutional blocks are used to extract embedded features, denoted as $L' = \text{Convs}(\tilde{L})$ and $R' = \text{Convs}(\tilde{R})$. Then, a cross-attention (CA) module [8] is utilized to enhance the content information in the reflectance map under the guidance of the illumination map, resulting in $R'' = \text{CA}(R', L')$. Additionally, a self-attention (SA) module [20] is applied to further extract content information from the illumination map, denoted as $L'' = \text{SA}(L')$, which is then supplemented into the reflectance map. The final outputs of the content map R and the illumination map L can be achieved:

$$R = \text{Convs}(R'' + L''),\ L = \text{Convs}(L' - L''). \quad (4)$$

The decomposition is performed in the latent space, which can more effectively separate the image content under complex illumination conditions. The generated reflectance map is rich in content details, while the illumination map contains only brightness information that is unaffected by detailed content. Subsequent training utilizes decoupled illumination maps in conjunction with camera parameters to facilitate further learning.

4.3 Feature Fusion Module

Based on the decoupled illumination map, we propose a feature fusion module to process camera parameters of input images. It generates feature vectors based on the network and indirectly output camera parameters. The purpose is to regulate EBN values, ensuring the network produces consistent camera parameters for different images with various exposure conditions. The module employs nonlinear processing and normalization to address the differences in magnitude and nonlinear distributions of camera parameters, while also utilizing a random dropout mechanism to prevent overfitting and ensure the robustness and generalization capability of the model.

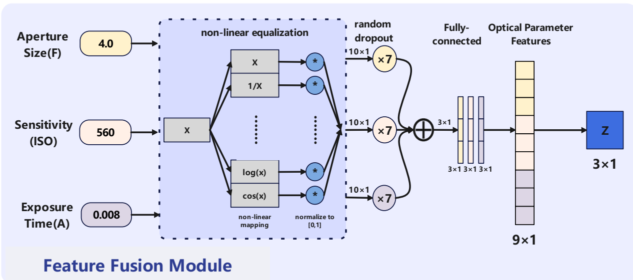

Figure 5: Schematic diagram of the feature fusion module for camera parameter encoding.

Firstly, recorded camera parameters of the image are taken as inputs. Due to the significant differences in magnitude and the nonlinear distributions of these parameters, we initially process them through a nonlinear equalization layer [12]. Let x represents the value of the camera parameter. Within the nonlinear equalization layer, multiple nonlinear mapping functions are applied to each camera parameter, including $\frac{1}{x}$, $\sqrt{x}$, $x^{-1/2}$, $x^{1/4}$, $x^{-1/4}$, $\log(x)$, $\sin(\log(x))$, $\cos(\log(x))$, $\sin(c \cdot x)$, $\cos(c \cdot x)$, and c controls the frequency of the sinusoidal functions. These functions process the input camera parameters through various nonlinear transformations to accommodate different data distributions, introduce nonlinear features, enhance feature diversity, and reduce the differences among parameters through normalization, thereby providing richer and more stable information for subsequent model training. We empirically chose three different values, and the results are normalized to the range $[0, 1]$ to compensate for the differences in magnitude among the parameters. To mitigate overfitting caused by the differences in parameter magnitudes and insufficient training data, we implement random dropout on the equalized feature vectors of each optical parameter during training. It is randomly discarding certain parameter feature vectors with a certain probability (e.g., 70%). Finally, the equalized feature vectors output by the nonlinear equalization layer are fed into a fully connected layer to generate the optical parameter feature vector z. The specific workflow is shown in Figure 5. Combined with the mentioned EBN, camera parameters can be estimated.

4.4 Loss Function

For parameter estimation network training, we employ the Mean Squared Error (MSE) as the primary loss function, which aims to bring the model's predictions to align ground-truth. It penalizes the deviations between the predicted and ground-truth values, thereby guiding the model to gradually optimize its predictions. The MSE loss can be formulated as:

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2, \tag{5}$$

where y_i is the true exposure brightness indicator, $\hat{y}_i$ is the model predicted value, N is the number of samples in the current batch. The loss is used to guide the model to align ground-true values.

Using MSE alone may lead to overfitting on certain samples and inconsistent performance on samples with similar or identical features. In the training data, there may exist cases where different images have identical camera parameters. Although these images have different content, their predicted target brightness values should be the same from the perspective of camera parameters. The consistency loss function aims to constrain the consistency of model outputs among samples with identical camera parameters. It can be formulated as:

$$L_{\text{Consistency}} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} b(y_i,\ y_j) \cdot \|\hat{y}_i - \hat{y}_j\|^2, \tag{6}$$

where $b(\cdot)$ is a binary function, $b(y_i, y_j) = 1$ if $y_i \approx y_j$ ($|y_i - y_j| < min(y_i, y_j) \cdot 5\%$), otherwise $b(y_i, y_j) = 0$, $\hat{y}_i$ and $\hat{y}_j$ are predicted outputs of the model, and N is the number of sample pairs that meet the condition. The consistency loss keeps the deviation of the model's output within 10% of the ground-truth values when the input parameters are exactly the same.

There may be subtle differences in parameters of different images (for instance, ISO 400 and 500 share similar visualization). Although such differences are not significant, they should be represented and learned in a physical perspective. We utilize the contrastive loss function to enhance the model's perceptual capability for such differences. It is represented as:

$$L_{\text{Contrastive}} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} b'(y_i, y_j) \cdot \max(1 - |\hat{y}_i - \hat{y}_j|, 0), \tag{7}$$

where $b'(\cdot)$ is a binary function, $b'(y_i, y_j) = 1$ if $\mathbf{1}(y_i \neq y_j)$, otherwise $b'(y_i, y_j) = 0$, $\left|\hat{y}_i - \hat{y}_j\right|$ represents the absolute difference of the model's predicted values. The loss function is to enable the model to distinguish samples corresponding to different parameters while avoiding the predicted values from being too similar. Finally, we achieve the loss function based on the mentioned three items:

$$L_{\text{Total}} = L_{\text{MSE}} + \alpha L_{\text{Consistency}} + \beta L_{\text{Contrastive}}. \tag{8}$$

where α and β are hyperparameters controlling the weights of the consistency loss and contrastive loss, respectively. Combined the three loss functions, the parameter estimation network can better capture the complex characteristics for exposure learning and output more accurate and stable camera parameters. It can fit the true exposure brightness values while improve the rationality and discriminative ability of predictions under specific constraints. Experiments will demonstrate advantages of Cam-Bench.

5 Experiments

To evaluate the performance of Cam-Bench, we build a series of camera parameter prediction tests and report their performance in downstream tasks. The experimental environment equipped with an Intel i9 3.0GHz CPU and NVIDIA 3090 GPU, using PyCharm and PyTorch for network training. We firstly evaluate the accuracy of Cam-Bench for camera parameter estimation. Then, we report its performance in some related downstream tasks, including illumination transfer, lighting enhancement, exposure-based IQA (image quality assessment), and auto-parameter recommendation. Experimental results show that our method can significantly enhance the modeling and processing capabilities for illumination.

Table 1: Camera parameter estimation accuracy of Cam-Bench for different test cases.

Known Parameters			Predicted Parameters			Accuracy
ISO	A	F	ISO	A	F	(%)
	✓	✓	✓			84.6
✓		✓		✓		81.3
✓	✓				✓	82.8
✓				✓	✓	87.9
			✓	✓	✓	75.8

Table 2: Qualitative results on mixed test dataset by different lighting enhancement methods.

Methods	PSNR ↑	SSIM ↑	BRISQUE ↓	Cam-IQA ↑
ZeroDCE [6]	17.41	0.6792	0.4929	0.6487
EnlightenGan [11]	18.81	0.6777	0.5474	0.6011
Uretinex-net [24]	20.46	0.7226	0.6825	0.5921
HVI [26]	18.93	0.7351	0.4377	0.6513
ZeroDCE + CamBench	**21.66**	**0.7905**	**0.3859**	**0.7371**

5.1 Evaluation for Parameter Estimation

The most straightforward validation of the proposed method is to verify its accuracy in camera parameter estimation. We conduct a comprehensive and rigorous validation of Cam-Bench, encompassing multiple scenarios of optical parameter prediction. It should be noticed that while our network accounts for the impact of parameter variations, simultaneously estimating all three camera parameters remains challenging. According to camera imaging principles, theoretically similar pixel-level results can be obtained through fine-tuning such parameters. Fortunately, if we fix certain camera parameters and estimate the remaining ones based on the input, such ambiguity can be effectively constrained. It mirrors the principle of using priority modes (e.g., aperture-priority or shutter-priority) in conventional photography.

Based on these considerations, we examine the following test cases: ISO sensitivity prediction with known shutter speed and aperture values; shutter speed prediction with known ISO sensitivity and aperture values; aperture value prediction with known ISO sensitivity and shutter speed; simultaneous prediction of both shutter speed and aperture value given only ISO sensitivity. These test scenarios cover common parameter combinations in photographic practice, designed to validate the model's predictive capability and accuracy under varying conditions, given the complexity and diversity of camera parameters, we consider a prediction to be successful if the predicted value is within 10% fluctuation of the ground-truth value. Through these evaluations, we gain deeper insights into the model's performance in practical applications, facilitating further optimization. The detailed results are presented in Table 1.

5.2 Applications

Illumination Transfer. Retinex decomposition implemented by Cam-Bench decouples the illumination map that can represent the illumination information of input image. It solely captures the lighting conditions of the scene and is useful for illumination-based downstream tasks. With reference to color transfer [18], we propose an illumination map-based lighting transfer application. It

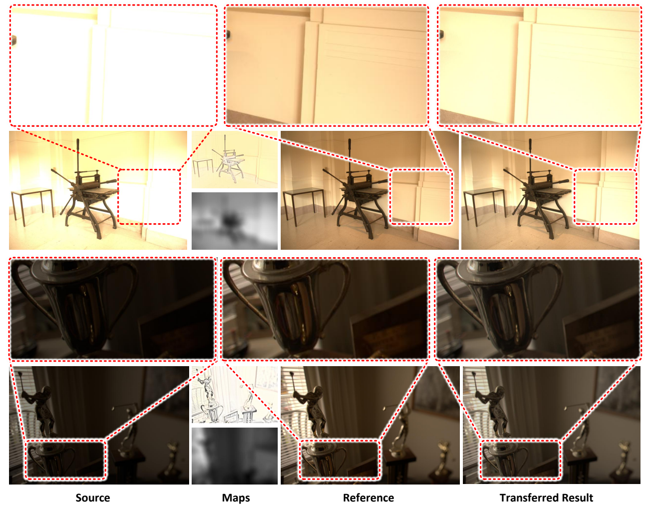

Figure 6: Instances of illumination transfer. Illumination information can be mapped from reference images to source ones with similar content.

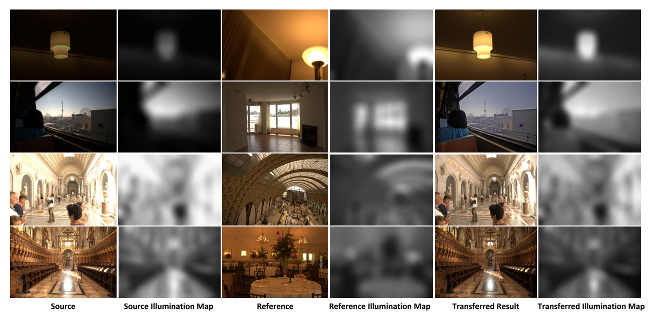

Figure 7: Instances of illumination transfer. Illumination information can be mapped from reference illumination maps to source ones without similar content.

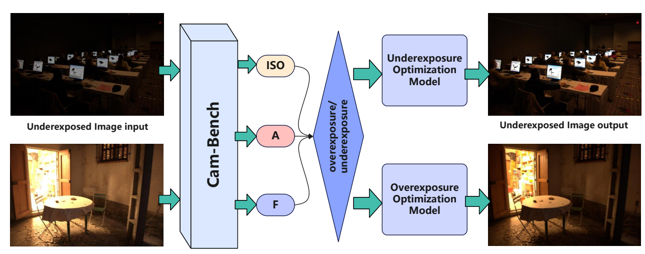

Figure 8: Cam-Bench for lighting enhancement. Estimated camera parameters can be used to drive an auto-model selection for images with various exposure information.

can transfer illumination information from the reference image to the source image, regardless of whether their content is similar. In Figure 6, we show some instances of illumination transfer between images with similar content. Illumination maps of reference images are synthesized with source contents to generate transferred results. In Figure 7, we show some instances of illumination transfer between images without similar content. Illumination maps of source images are edited according to the reference. Then, the illumination

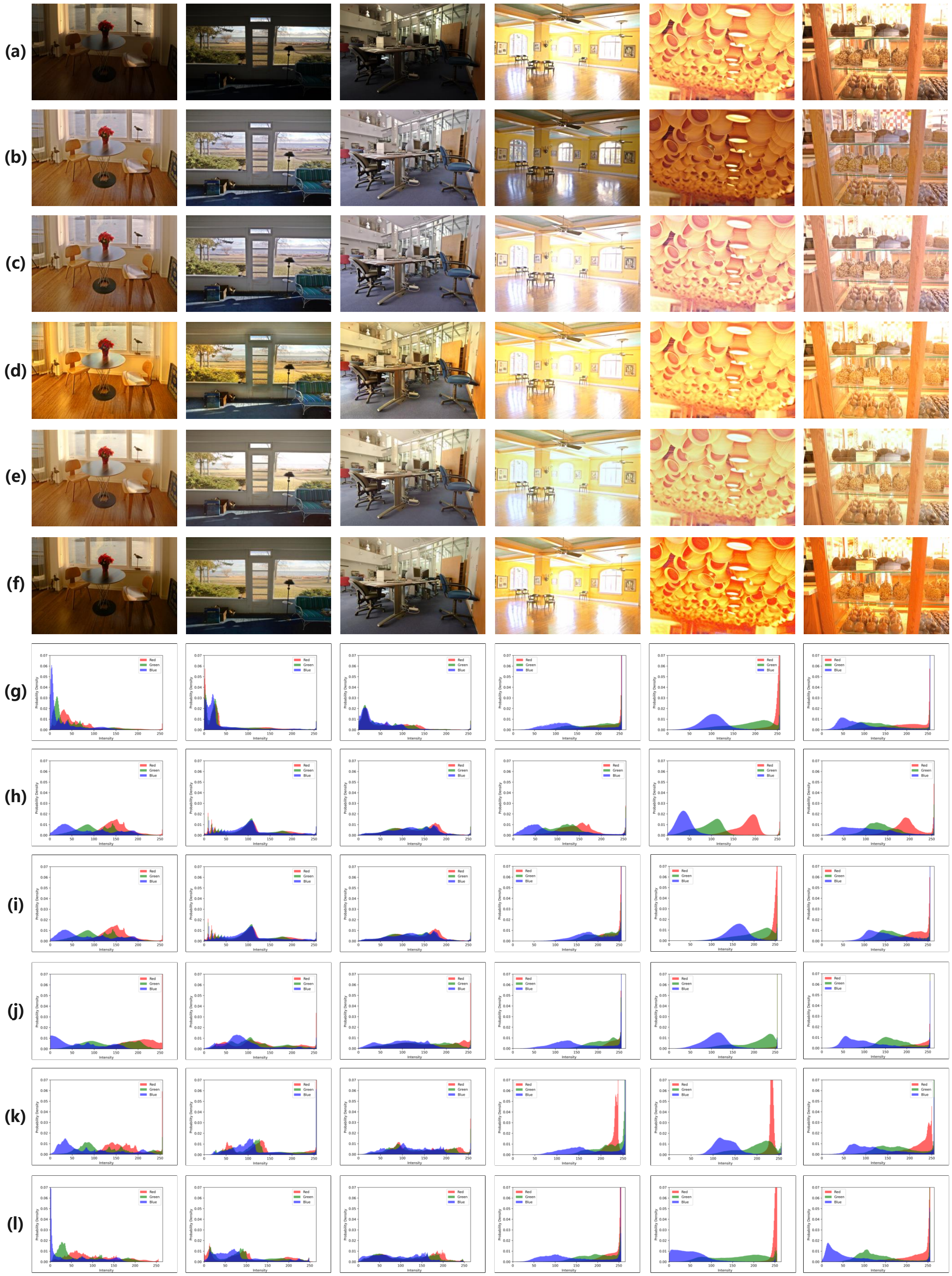

Figure 9: Comparisons of different lighting enhancement methods. (a) original input; (b) ZeroDCE [6] with Cam-Bench; (c) ZeroDCE [6]; (d) Enlightengan [11]; (e) Uretinex-net [24]; (f) HVI [26]; (g-l) RGB histograms corresponding to (a-e).

Table 3: Qualitative results of exposure-based IQA by different methods on indoor images. ISO$<$ 400 means underexposure; ISO$\in [800, 1200]$ means normal exposure; ISO$>$ 3200 means overexposure.

ISO Level	PSNR ↑	SSIM ↑	BRISQUE ↓	Cam-IQA ↑
$<$ 400	9.51	0.4648	0.6530	0.4691
400 ∼ 800	10.94	0.6113	0.7452	0.6971
800 ∼ 1200	22.46	0.7962	0.2164	0.8255
1200 ∼ 1600	21.44	0.8346	0.2469	0.8043
1600 ∼ 3200	13.77	0.6402	0.6544	0.6716
$>$ 3200	11.87	0.5615	0.7046	0.5394

transfer can be implemented. It optimizes the brightness and contrast of the image while preserving the naturalness and authenticity of the content, thus achieving high-quality image enhancement for complex lighting environments.

Lighting Enhancement. Implementing a lighting enhancement model based on estimated camera parameter naturally serves as a downstream application, when the exposure information of an image can be explicitly represented. The advantage lies in the fact that camera parameters can directly represent exposure information and automatically link to a corresponding pre-trained lighting enhancement model, as shown in Figure 8. It is crucial for most illumination optimization methods [6, 10, 11, 24], as their performance is inherently constrained by the training data distribution. Under a fixed parameter configuration, such methods can typically handle only a single type of exposure adjustment task. In Figure 9, we compare different lighting enhancement methods and show enhanced results with histograms. Even integrating Cam-Bench with earlier solution [6], superior results can still be achieved on the test dataset with overexposed and underexposed images. In Table 2, quantitative results provide reliable empirical evidence.

Exposure-based IQA. From a photographic perspective, the characterization of image exposure quality can be directly mapped to a set of camera parameters. For instance, by fixing the shutter speed, and aperture F-number, ISO sensitivity estimation can serve as a direct quantitative descriptor of the image's exposure quality. Therefore, we propose another downstream application of Cam-Bench, which is the exposure-based IQA (Cam-IQA). As mentioned before, generated EBN is used to estimate camera parameters. We utilize the EBN to output scores, which directly corresponds the ISO sensitivity (other parameters are fixed). In Table 3, quantitative results for images with different exposure levels are reported. Using ISO: 800 ∼ 1200 as the baseline for normal exposure (indoor scenes typically exhibit darker exposures that require a little higher ISO values), scores of Cam-IQA can accurately characterize the progressive impact of overexposure and underexposure on image quality. Under normal exposure conditions as the baseline, Cam-IQA exhibits an overall linear variation when increasing or decreasing the ISO, statistically demonstrating a precise reflection of exposure quality changes based on ISO adjustment. In contrast, both SSIM and BRISQUE exhibit score fluctuations when ISO values are smoothly changed.

Auto-parameter Recommendation. The exposure assessment capability of Cam-Bench inherently supports the inverse optimization of camera parameters for a given image, thereby enabling the development of an auto-parameter recommendation framework. Inspired by the parameter priority modes (e.g., aperture-priority, shutter-priority) implemented in conventional camera systems, Cam-Bench provides equivalent programmable control capabilities through its exposure optimization framework. It can provide parameter modification recommendation when an inexperienced user manually adjusts camera parameters (non-auto mode) for photography. In Figure 10, two instances are shown. Auto-parameters are provided for exposure correction.

Figure 10: Instances of auto-parameter recommendation for photo capturing. Exposure quality can be optimized.

Limitations. Collected Cam-10K focuses on indoor scenes with various illumination conditions. In terms of generalization capability, the parameter prediction accuracy exhibits instability when applied to outdoor scenarios. From a generalizability perspective, Cam-Bench exhibits performance degradation in parameter prediction for outdoor scenarios. In addition, more accurate results can in fact be obtained by leveraging pixel-level random noise analysis and foreground-background blur metrics. Cam-Bench lacks dedicated processing modules for these features, which necessitates resolution in the future work.

6 Conclusion

In this paper, we propose a new benchmark for camera parameter estimation from image directly. It can be regarded as a pioneering work in this topic. The main contributions consist of two components: a specifically collected image dataset Cam-10K with various exposure conditions and ground truth (GT) annotations of camera parameters; a camera parameter estimation network for image-based parameter prediction directly. Cam-10K provides a comprehensive collection of indoor images under complex lighting conditions, enabling high-precision camera parameter parsing. The proposed estimation network utilizes a Retinex decomposition module to split illumination and content maps. Then, it uses a feature fusion module with a EBN-based mechanism to generate camera parameters. Benefited from the design, Cam-Bench can accurately estimate camera parameters while addressing both the fluctuations induced by minor parameter variations and the coupling effects between parameters. Furthermore, it provide a camera parameter-based solution for illumination analysis. A series of downstream tasks can achieve performance improvements with the assistance of Cam-Bench. Experiments demonstrate that it takes significant research value and warrants further investigation.

References

[1] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. 2011. Learning Photographic Global Tonal Adjustment with a Database of Input / Output Image Pairs. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition.

[2] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. 2018. Learning to see in the dark. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 3291–3300.

[3] Dinu Coltuc, Philippe Bolon, and J-M Chassery. 2006. Exact histogram specification. IEEE Trans. on Image Processing 15, 5 (2006), 1143–1152.

[4] Yuming Fang, Pengwei Yuan, Chenlei Lv, Chen Peng, Jiebin Yan, and Weisi Lin. 2024. Saliency Guided Deep Neural Network for Color Transfer With Light Optimization. IEEE Trans. on Image Processing 33 (2024), 2880–2894.

[5] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. 2016. A weighted variational model for simultaneous reflectance and illumination estimation. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 2782–2790.

[6] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. 2020. Zero-reference deep curve estimation for low-light image enhancement. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 1780–1789.

[7] Xiaojie Guo, Yu Li, and Haibin Ling. 2016. LIME: Low-light image enhancement via illumination map estimation. IEEE Trans. on Image Processing 26, 2 (2016), 982–993.

[8] Ruibing Hou, Hong Chang, Bingpeng Ma, Shiguang Shan, and Xilin Chen. 2019. Cross attention network for few-shot classification. Proc. Conf. on Neural Information Processing Systems 32 (2019).

[9] Haidi Ibrahim and Nicholas Sia Pik Kong. 2007. Brightness preserving dynamic histogram equalization for image contrast enhancement. IEEE Trans. on Consumer Electronics 53, 4 (2007), 1752–1758.

[10] Hai Jiang, Ao Luo, Xiaohong Liu, Songchen Han, and Shuaicheng Liu. 2025. Lightendiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. In Proc. Euro. Conf. on Computer Vision. Springer, 161–179.

[11] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. 2021. Enlightengan: Deep light enhancement without paired supervision. IEEE Trans. on Image Processing 30 (2021), 2340–2349.

[12] Woohyeok Kim, Geonu Kim, Junyong Lee, Seungyong Lee, Seung-Hwan Baek, and Sunghyun Cho. 2024. ParamISP: Learned Forward and Inverse ISPs using Camera Parameters. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 26067–26076.

[13] Edwin H Land. 1977. The retinex theory of color vision. Scientific american 237, 6 (1977), 108–129.

[14] Chulwoo Lee, Chul Lee, and Chang-Su Kim. 2013. Contrast enhancement based on layered difference representation of 2D histograms. IEEE Trans. on Image Processing 22, 12 (2013), 5372–5384.

[15] Huafeng Li, Zhenmei Yang, Yafei Zhang, Dapeng Tao, and Zhengtao Yu. 2024. Single-image HDR reconstruction assisted ghost suppression and detail preservation network for multi-exposure HDR imaging. IEEE Trans. on Computational Imaging (2024).

[16] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. 2018. Structure-revealing low-light image enhancement via robust retinex model. IEEE Trans. on Image Processing 27, 6 (2018), 2828–2841.

[17] Chenlei Lv and Dan Zhang. 2024. Palette-based Color Transfer between Images. arXiv preprint arXiv:2405.08263 (2024).

[18] Chenlei Lv, Dan Zhang, Shengling Geng, Zhongke Wu, and Hui Huang. 2024. Color transfer for images: A survey. ACM Trans. on Multimedia Computing, Communications and Applications 20, 8 (2024), 258:1–258:29.

[19] J Alex Stark. 2000. Adaptive image contrast enhancement using generalizations of histogram equalization. IEEE Trans. on Image Processing 9, 5 (2000), 889–896.

[20] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. Proc. Conf. on Neural Information Processing Systems 30 (2017).

[21] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. 2013. Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Trans. on Image Processing 22, 9 (2013), 3538–3548.

[22] Wenjing Wang, Huan Yang, Jianlong Fu, and Jiaying Liu. 2024. Zero-reference low-light enhancement via physical quadruple priors. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 26057–26066.

[23] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. 2018. Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560 (2018).

[24] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. 2022. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 5901–5910.

[25] Yazhou Xing, Zian Qian, and Qifeng Chen. 2021. Invertible image signal processing. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 6287–6296.

[26] Qingsen Yan, Yixu Feng, Cheng Zhang, Guansong Pang, Kangbiao Shi, Peng Wu, Wei Dong, Jinqiu Sun, and Yanning Zhang. 2025. HVI: A New color space for Low-light Image Enhancement. arXiv preprint arXiv:2502.20272 (2025).

[27] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. 2020. Cycleisp: Real image restoration via improved data synthesis. In Proc. IEEE/CVF Conf. on Computer Vision & Pattern Recognition. 2696–2705.

[28] Yonghua Zhang, Xiaojie Guo, Jiayi Ma, Wei Liu, and Jiawan Zhang. 2021. Beyond brightening low-light images. Int. J. Computer Vision 129 (2021), 1013–1037.

[29] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. 2019. Kindling the darkness: A practical low-light image enhancer. In Proc. ACM Int. Conf. on multimedia. 1632–1640.