

# Unsupervised Exposure Correction

Ruodai Cui<sup>1</sup>, Li Niu<sup>2\*</sup>, and Guosheng Hu<sup>3</sup>

<sup>1</sup> Qualcomm Technologies, Inc.  
ruodcui@qti.qualcomm.com

<sup>2</sup> Department of Computer Science and Engineering, MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University  
ustcnewly@sjtu.edu.cn

<sup>3</sup> University of Bristol

**Abstract.** Current exposure correction methods have three challenges, labor-intensive paired data annotation, limited generalizability, and performance degradation in low-level computer vision tasks. In this work, we introduce an innovative Unsupervised Exposure Correction (UEC) method that eliminates the need for manual annotations, offers improved generalizability, and enhances performance in low-level downstream tasks. Our model is trained using *freely* available paired data from an emulated Image Signal Processing (ISP) pipeline. This approach does not need expensive manual annotations, thereby minimizing individual style biases from the annotation and consequently improving its generalizability. Furthermore, we present a large-scale Radiometry Correction Dataset, specifically designed to emphasize exposure variations, to facilitate unsupervised learning. In addition, we develop a transformation function that preserves image details and outperforms state-of-the-art supervised methods [12], while utilizing only 0.01% of their parameters. Our work further investigates the broader impact of exposure correction on downstream tasks, including edge detection, demonstrating its effectiveness in mitigating the adverse effects of poor exposure on low-level features. The source code and dataset are publicly available at [https://github.com/BeyondHeaven/uec\\_code](https://github.com/BeyondHeaven/uec_code).

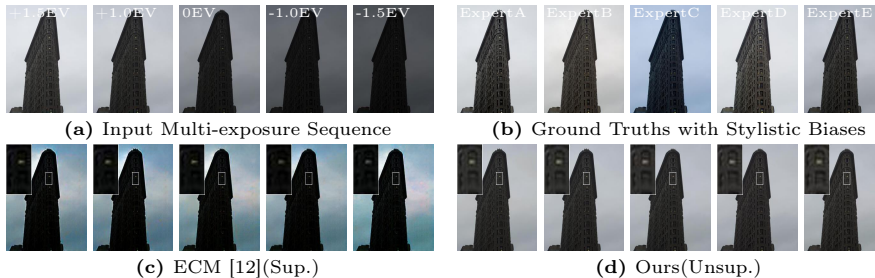
**Keywords:** Exposure correction · Unsupervised learning

## 1 Introduction

Exposure, a pivotal factor in photography, significantly impacts image quality by influencing visual clarity. In radiometry, the exposure is defined as scene irradiance, the amount of light that reaches the image sensor. This can be controlled by exposure value (EV), which combines essential factors such as aperture, shutter speed, and ISO settings. Despite advancements in Image Signal Processor (ISP), which facilitate automatic EV adjustments during image capture, challenges persist under non-ideal lighting conditions. As a result, post-processing

---

\* Corresponding author.



**Fig. 1:** Visual Comparison: ECM [12] vs. Our UEC method.

of sRGB images remains crucial. The advent of deep learning has spurred extensive research in this field, with numerous studies proposing models for exposure correction [1, 4, 14, 16, 19, 24, 25, 27, 32, 34, 42, 44], demonstrating significant achievements. Nevertheless, these methods still face three challenges.

(1) A primary challenge arises from the dependence on the paired data, where the ground truth is from proficient photographers. This process is inherently complex and labor-intensive, as it involves detailed editing and refinement of each image, demanding more manual work than just labeling as in classification tasks. (2) Previous methods often suffer from limited generalizability. To be specific, (a) as noted, the efficiency of manual adjustments is low, resulting in small datasets available in academia. (b) the manual adjustments inevitably introduce different stylistic biases. Personal preferences vary significantly between individuals, suggesting that such ground truths are intrinsically noisy. (3) Previous methodologies have predominantly focused on generating aesthetically pleasing images. However, they often yield images with notable degradations in low-level features. Such degradations render these images less suitable for various downstream computer vision tasks, like edge detection and segmentation, where preserving these features is important.

Inspired by Afifi et al.’s work of exposure dataset [2], we realize that acquiring paired data does not necessarily require extensive manual intervention. They apply an emulated Image Signal Processing (ISP) pipeline on RAW data, thereby creating multi-exposure sequences, as shown in Fig. 1a. The work [2] learns the mapping from the inputs (Fig. 1a) to the ground truth (ExpertC in Fig. 1b), which is a typical supervised learning. However, this approach introduces ambiguities due to inconsistencies in the ground truths, e.g. five experts in Fig. 1b. Unlike existing supervised learning methods [2, 12, 14], in this work, we introduce an innovative Unsupervised Exposure Correction (UEC) method. (1) To achieve this, we creatively ask the images in the same multi-exposure sequence, which can be generated freely, to mutually serve as a ground truth for learning exposure adjustments across a diverse range. Specifically, we simply generate multi-exposure sequences and EV labels via an emulated ISP pipeline. In this way, we do not need the expensive manual annotations to generate ground truth. (2) The existing methods have inconsistent ground truth which might have different

standards on colors, leading to degraded performance. In contrast, in our UEC method, the images in multi-exposure sequence do not have such “standards” because they mutually work as the ground truth. (3) We employ a pixel-wise exposure transformation in sRGB color space for preserving visual details as shown in Fig. 1d which greatly outperforms the SOTA method in Fig. 1c. This detail-preserving property is crucial for downstream computer vision tasks that rely on low-level features, such as edge detection and segmentation. To further facilitate the analysis without individual stylistic biases in exposure correction, we contribute a new dataset, Radiometry Correction Dataset, that is large and encompasses broad exposures. This dataset maintains a consistent style with only radiometry variations.

Our contributions can be summarized as follows: (1) By modeling radiometry, we introduce an innovative unsupervised learning methodology that fundamentally addresses the problem of expensive annotations. To our knowledge, this is the first unsupervised learning solution for exposure correction, which achieves competitive performance as SOTA supervised models [12] with only 0.01% of their parameters, and minimizes individual stylistic biases. (2) We propose a pixel-wise exposure transformation that can well preserve the details of images and flexibly accept different resolutions of input images. By this way, the enhanced images can improve the performance of many downstream tasks, e.g. edge detection. (3) We contribute a new dataset focusing on radiometry for better generalizability. Input images of this dataset are synthesized from ground truth data, ensuring a uniform style across the dataset.

## 2 Related Work

Exposure correction has evolved from traditional techniques to modern deep learning methods. Traditional methods, like histogram equalization [17], and Retinex-based algorithms [22, 33], laid the groundwork. Recent deep learning studies, including Ren et al. [34], Zhang et al. [44], Guo et al. [14], and Loh et al. [29], focus on low-light enhancement. However, these methods exposed a gap in addressing both underexposed and overexposed images. Afifi et al. [2] introduced a dataset featuring varying exposure attributes, enabling comprehensive methodologies.

Exposure correction methods fall into two categories: image-to-image translation and color transformation. Image-to-image translation aims to create dense translations between input and output pairs, leveraging deep learning as seen in works by Eyiokur et al. [12], Afifi et al. [2], and others [4, 8, 19]. Color transformation methods, on the other hand, predict mapping curve parameters to enhance images, employing techniques like quadratic transforms [7, 26, 37, 41], local affine transforms [13], curve-based transforms [5, 14, 20, 23, 31], filters [11, 30], lookup tables [38, 43], and MLP models [15, 28, 39]. These techniques define specific functions or models for the enhancement process, offering diverse tools for image exposure challenges.

### 3 Method

Traditionally, exposure correction learns a mapping from one input image  $I_1$  with one (usually weaker) exposure to another image  $I_2$  with a different (usually stronger and better) exposure. To learn this mapping, conventionally, people *manually* generate the well-exposed image  $I_2$  as the ground-truth of the existing  $I_1$ , which is quite expensive and is hard to scale up.

To address this, we creatively propose a way to exploit the information from the *freely* generated multi-exposure sequence from RAW data. In the same sequence, the images with different exposures can mutually work as the ground-truth to learn the mapping for exposure adjustment. In Sec. 3.1, we formulate our Unsupervised Exposure Correction (UEC). Then, the neural architectures in Sec. 3.2 and loss functions in Sec. 3.3 are introduced to achieve UEC. Following that, Sec. 3.4 outlines the testing methodology. Moreover, we propose our Radiometry Correction Dataset where the paired images are freely generated from RAW data in Sec. 4.

#### 3.1 Unsupervised Exposure Correction Modeling

The emulated ISP enables us to generate extensive multi-exposure sequences. To make full use of the large data, we introduce a task with self-supervision. A reference image guides exposure adjustments, serving as the calibration target, while a transformation curve shifts the exposure level. This method involves varying the exposure over a wide range, rather than directly aligning it to an optimal value, thereby categorizing it as a specific form of style transfer.

We initiate the process by employing a style encoder to extract the exposure feature from the image, denoted as  $E = e(I)$ . Diverging from conventional style transfer methods, which typically rely on a one-hot vector as a prior [9] or leverage features refined through triplet loss [21], we have developed an alternative optimization approach that better accommodates exposure dynamics. We discovered a sequential relationship within varying exposure images, allowing us to order them by their exposures. Consequently, to quantify the difference between any two images, we can employ a single scalar.

$$\Delta E = d(E_1, E_2) = d(e(I_1), e(I_2)). \quad (1)$$

Then, we can perform the style transformation from  $\Delta E$ :

$$I'_1 = f(\Delta E, I_1). \quad (2)$$

Having established the transformation function, , we now focus on developing optimization strategies. We identify two principles to guide the optimization process.

**Restoration Supervision for Pretext Task.** Consider a scenario where we sample two images,  $I_1$  and  $I_2$ , from the same sequence  $I$ , leading to the equation  $I'_1 = I_2$ . We call this “Restoration Supervision”, which can train  $f(\cdot)$  under the restoration guidance of  $I_2$ , as illustrated in Fig. 2a.

However, it is important to note that both  $I_1$  and  $I_2$  originate from the same sequence  $I$ . For the real task, we need to handle image pairs from different scenes. Since our method learn from data which varies solely in exposure, the trained  $f(\cdot)$  can only modify exposures. Moreover, we calculate the exposure difference in a latent space, not in pixel space, which omits low-level features. This approach enables the difference function  $d(\cdot, \cdot)$  to adapt images from different scenes to a certain degree. Therefore, when we select reference images from a different sequence, the alterations remain confined to exposure adjustment, even though the  $\Delta E$  computed may not precisely indicating the optimal exposure variation.

**Monopoly Principle for Real Task.** In order to compute this value accurately, we proceed to train the network with images from different scenes. However, comparing exposures in images from different scenes poses a challenge. Instead of directly contrasting the input and output images, we compare two outputs from the same input, ensuring scene consistency and enabling pixel-to-pixel optimization. Notably, if an over-exposed image is selected as a reference, the resulting image will be brighter; conversely, choosing an under-exposed image as a reference will lead to a darker result. Therefore, by selecting reference images with varying EVs, we can create a bright-dark image pair. We call this “Monopoly Principle” to indicate that the output’s brightness should be monopoly with the exposure of reference image.

We select two reference images,  $J_1$  and  $J_2$ , from a sequence  $J$ , distinct from  $I$ , where we sample input images  $I_1$ . We assume the EV of  $J_1$  exceeds that of  $J_2$ , which means every pixel value of  $J_1$  should be equal or over that of  $J_2$ . This condition can be expressed as:

$$\forall(x, y), \quad J_1(x, y) \geq J_2(x, y) \quad \text{with } \text{EV}(J_1) > \text{EV}(J_2). \quad (3)$$

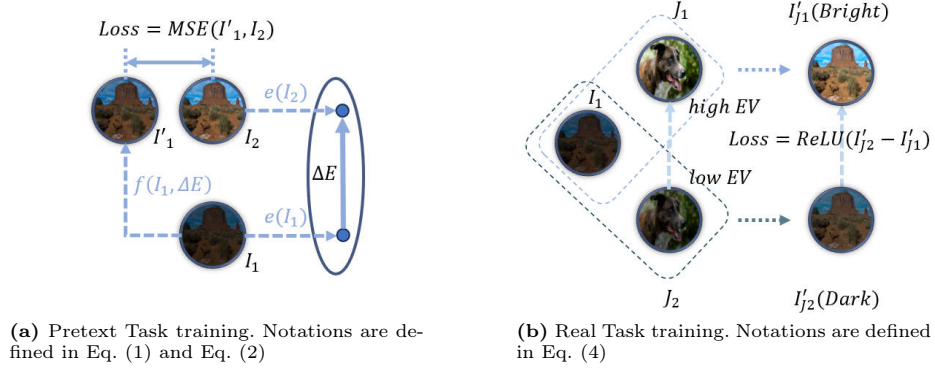
Employing  $J_1$  and  $J_2$  as reference images, we apply  $d(\cdot, \cdot)$  and  $f(\cdot)$  to the same image  $I_1$ , resulting in two transformed images,  $I'_{J_1}$  and  $I'_{J_2}$ , respectively. In this case,  $I'_{J_1}$  is expected to exhibit a brighter appearance compared to  $I'_{J_2}$ . This can be described as

$$I'_{J_1} = f(d(e(I_1), e(J_1)), I_1), \quad I'_{J_2} = f(d(e(I_1), e(J_2)), I_1). \quad (4)$$

Combine Eq. (3) and Eq. (4), we can get

$$\forall(x, y), \quad I'_{J_1}(x, y) \geq I'_{J_2}(x, y) \quad \text{with } \text{EV}(J_1) > \text{EV}(J_2). \quad (5)$$

As demonstrated in Fig. 2b, this learning criterion enables the network to align exposures across diverse scenes, thereby enhancing the applicability of our method in real-world situations.



**Fig. 2:** Schematic diagrams of training UEC.

**Summarization.** We design a transformation to adjust exposures, which can be optimized by two principles. Firstly, we select images from a single multi-exposure sequence, establishing a baseline for exposure manipulation. Secondly, we aim to adapt to varied scenes. This involves selecting two reference images from different exposures, but differing from the sequence where we sample the input image. By preserving the relative brightness in the outputs, we facilitate exposure style transformations across different scenes. We concurrently train the network towards using two specific loss functions for each goal.

### 3.2 Unsupervised Exposure Correction Architecture

We have introduced three crucial functions: (1)  $e(\cdot)$  for the extraction of exposure-related features, (2)  $d(\cdot, \cdot)$  for the computation of exposure difference between two images in a latent space, and (3)  $f(\cdot)$  for the exposure correction of images given the exposure difference. Our UEC network is structured around these functions, as illustrated in the overview shown in Fig. 3a. Specifically, our UEC framework incorporates three distinct neural networks: (1) Exposure Feature Encoder, detailed in Fig. 3b; (2) Parameter Predictor based on the difference, detailed in Fig. 3c; and (3) Exposure Corrector, detailed in Fig. 3d.

Firstly, a decoder comprises two convolutional layers followed by a global pooling layer, producing a 96D feature representation through the combination of maximum, average, and standard deviation across channels. These statistical metrics links to global attributes of images, e.g., contrast, histogram distribution. Secondly, Parameter Predictor evaluates the exposure difference between image pairs, and then computes the parameter  $\lambda$  for exposure correction. Thirdly, the exposure correction is modeled as an interpolation of direct scaling and a non-linear adjustment through a network:

$$I_{out}(x, y) = \lambda \times I_{in}(x, y) + (1 - \lambda) \times h(I_{in}(x, y)), \quad (6)$$

where  $\lambda$  modulates the transformed between the original input image  $I_{in}$  and its corresponding output  $I_{out}$ .  $h(\cdot)$  is the non-linear transformation function im-

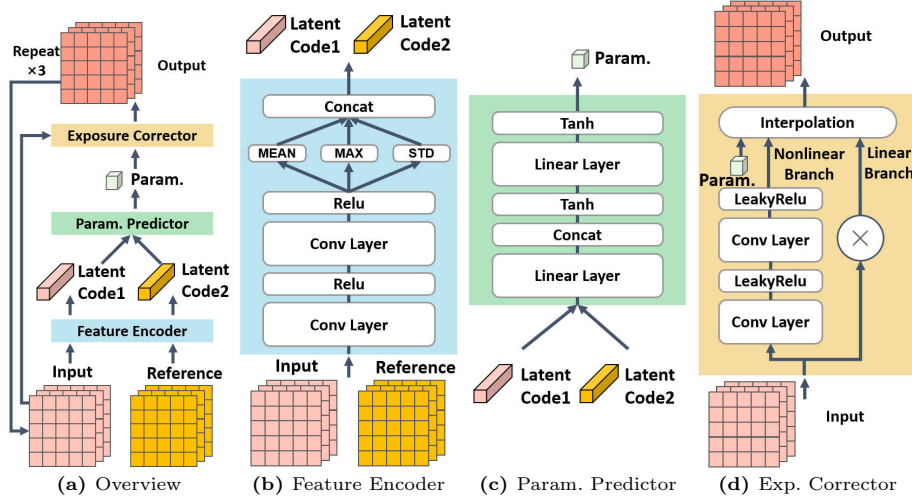


Fig. 3: The architecture of our UEC method.

plemented by  $1 \times 1$  convolution layers. This iterative process is repeated three times to enhance the correction effectiveness. For more details, please refer to supplementary material.

### 3.3 Loss Function

We jointly train the Exposure Feature Encoder, Parameter Predictor, and Exposure Corrector in an end-to-end fashion. We employ three types of losses and compute their weighted sum as the final loss.

**Restoration Loss.** Based on the Restoration Supervision discussed in Sec. 3.1, when we sample the input and reference from the same multi-exposure sequence, the synthesized exposure-adjusted image should be identical to the reference image. This is illustrated with  $I_1$  and  $I_2$  in Fig. 2a. We utilize the L2 loss for the restoration:

$$L_{\text{restoration}} = \frac{1}{CHW} \|I^{\text{out}} - I^{\text{ref}}\|_2. \quad (7)$$

Here,  $C$ ,  $H$ , and  $W$  denote the image’s channels, height, and width, respectively.  $I^{\text{out}}$  is the neural network output, and  $I^{\text{ref}}$  the reference image.

**Monopoly Loss.** Based on the Monopoly Principle discussed in Sec. 3.1, when we sample the two references from the same multi-exposure sequence, which is different from the input’s sampling source, the relative brightness relation should also be transferred to the output image pair. This is shown with  $I'_{j1}$  and  $I'_{j2}$  in Fig. 2b. We utilize the ReLU function for the monopoly loss:

$$L_{\text{monopoly}} = \frac{1}{CHW} \text{ReLU}(I^{\text{out}2} - I^{\text{out}1}) \quad \text{with } \text{EV}(I^{\text{ref}1}) > \text{EV}(I^{\text{ref}2}). \quad (8)$$

In this equation,  $I^{\text{out}1}$  and  $I^{\text{out}2}$  represent the neural network outputs corresponding to the references  $I^{\text{ref}1}$  and  $I^{\text{ref}2}$ , respectively. It is important to note that the EV of  $I^{\text{ref}1}$  is greater than that of  $I^{\text{ref}2}$ . The ReLU function ensures that the loss calculation emphasizes the positive differences, aligning with the expected brightness relation between the two outputs.

**Semantic-preserving Loss.** For the preservation of semantics, we use total variation loss [3], a proven regularization technique to maintain spatial coherence and semantic integrity. This approach smooths transitions between adjacent pixels:

$$L_{\text{semantic}} = \frac{1}{CHW} \|\nabla I^{\text{out}}\|_2. \quad (9)$$

Here,  $\nabla(\cdot)$  denotes the gradient operator, computing the image’s spatial derivative.

**Summarization.** The final loss function is expressed as:

$$L = \alpha_1 \cdot L_{\text{restoration}} + \alpha_2 \cdot L_{\text{monopoly}} + \alpha_3 \cdot L_{\text{semantic}}, \quad (10)$$

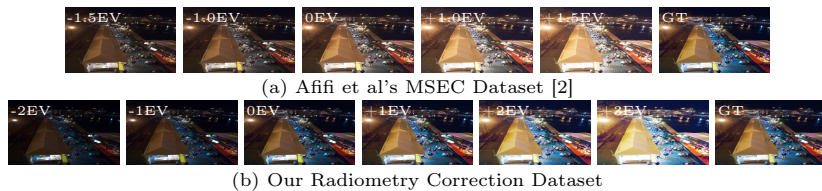
where  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  are the balancing weights. Since our framework (Exposure Feature Encoder, Parameter Predictor, Exposure Corrector) works in an end-to-end way, the total loss function, Eq. (10), is added on the end of exposure corrector to update the weights of all the modules during training.

### 3.4 Testing Details

We have outlined the training procedure for our UEC model, which employs reference images of diverse quality to address the full spectrum of exposure adjustments. For the testing phase, it is crucial to identify the optimal value for the final calibration. The testing inference largely replicates the training process as depicted in Fig. 3a, with the primary distinction being the use of a single reference image for all testing inputs. Specifically, we hard code the exposure features derived from this image across all test cases.

The efficiency of this method is underscored by its minimal data requirements. Mastering exposure adjustments is complex, but identifying an image’s optimal exposure level is considerably more straightforward. Our approach emphasizes radiometry to ensure consistency, which differs from colorimetry, thereby minimizing the data needed to such an extent that a single ground truth image is sufficient. After training, the UEC model is adept at generating a sequence of images with varying exposures for any given input. The goal during testing is





**Fig. 4:** Visualization of Datasets: (a) Afifi et al’s [2] vs. (b) Ours.

to select the most suitable image from this sequence. To enhance efficiency, our method bypasses the need to generate and then choose from multiple exposures, instead directly yielding the image with the best exposure.

## 4 Our Radiometry Correction Dataset

In Sec. 1, we have explored the benefits of disentanglement in radiometry adjustment for enhancing the generalizability of exposure correction. For this purpose, a specialized dataset is essential. Hence, we developed Radiometry Correction Dataset using MIT-Adobe FiveK Dataset [5], which includes 5,000 RAW photographs and their expertly modified sRGB images. These adjustments were performed by five specialists who worked directly with the RAW images. Drawing from their expertise, we adopted a reverse engineering approach to generate ill-exposed images. This process entails adjusting the exposure of the reference images to generate synthetic versions that are either underexposed or overexposed, while freezing other post-processing ISP procedures to keep the individual style constant. Following the established convention, we selected the versions edited by ExpertC as our reference images. Based on this, we adjusted the exposures relative to the original images, spanning  $-2\text{EV}$ ,  $-1\text{EV}$ ,  $0\text{EV}$ ,  $+1\text{EV}$ ,  $+2\text{EV}$ , and  $+3\text{EV}$ . Considering that input images are often underexposed, this range ensures a balanced exposure spectrum in relation to the reference images.

Fig. 4 provides a comparison between Afifi’s MSEC Dataset [2] and our Radiometry Correction Dataset. Evidently, in this example, the ground exhibits a blue hue as ExpertC’s individual style. For additional comparisons, please consult the supplementary material.

## 5 Experiments

### 5.1 Experiment Settings

**Datasets.** The prevailing benchmark primarily utilizes the MSEC Dataset from the study by Afifi et al. [2]. To mitigate bias stemming from individual styles on colorimetry, we introduce our Radiometry Correction Dataset. Both datasets are employed for training and testing various methods. To assess the models’ generalizability, we conduct training on the Exposure Dataset and perform testing on LOL dataset [6], which is specifically designed for low-light image enhancement.



**Fig. 5:** Visual Comparison: ECM [12] vs. Our UEC method.

**Implementation Details.** The weights in Eq. (10) are empirically set  $\alpha_1 = \alpha_2 = 1$  and  $\alpha_3 = 0.1$ . We use only one well-exposed reference image during testing. We select the second image from left to right in Fig. 9 as our reference in our evaluations. More details are described in our supplementary material.

**Evaluation Metrics.** In Sec. 1, we highlighted the importance of evaluating exposure correction in terms of both overall aesthetics and low-level features. To assess overall aesthetics, we compare the output to the ground truth images using metrics such as PSNR and SSIM [35]. For low-level feature evaluation, we analyze the output through edge detection and make comparisons using PSNR and F1 scores, applying a threshold based on median pixel values.

## 5.2 Exposure Correction Results

**Results on MSEC Dataset.** We compare our results with the ground truth, as presented in Tab. 1a. Even though our Unsupervised Exposure Correction (UEC) model has not been trained on manually calibrated data and our adjustments are confined to radiometry, the proficiency of our UEC model remains competitive when compared to supervised SOTA models like ECM [12] and Afifi et al. [2].

In Fig. 6, we present a comparison of test images using various methods, including our approach, ground truth images, and results from supervised models. Our method demonstrates competitive performance in radiometry correction, matching the SOTA supervised techniques. Furthermore, in Fig. 1 and Fig. 5, we compare our UEC method with ECM [12]. In Fig. 1, it is important to note that the ground truth colorimetry, as determined by five experts, exhibits variability, indicating that learning from human-adjusted images may introduce biases reflective of their individual styles. In Fig. 5, the ECM’s results [12] (Fig. 5c) manifest a color cast in the background, where the background turns green. Our method addresses this issue by applying radiometry correction. Furthermore, to highlight the superior detail resolution achieved by our UEC method, we have magnified a segment of the image located in the upper left corner. Further comparisons are elaborated in the supplementary material.

**Results on Generalizability.** To assess the models’ generalizability, we utilized pretrained models and evaluated their performance on LOL dataset [40].

Method	PSNR	SSIM
HDRCNN w/PS [10]	17.032	0.687
DPED (iPhone) [18]	16.274	0.629
DPED (BlackBerry) [18]	17.890	0.671
DPE(HDR) [8]	16.206	0.623
DPE(S-FiveK) [8]	17.510	0.677
Zero-DCE [14]	12.597	0.549
Affi et al. [2]	19.483	0.739
ECM [12]	<b>20.874</b>	<b>0.877</b>
Ours	18.756	0.812

(a) Testing Result

Method	PSNR	SSIM
Affi et al.(E) [2]	14.268	0.638
ECM(E) [12]	15.439	0.650
ECM(R) [12]	17.537	0.725
Ours(E)	<b>18.571</b>	<b>0.728</b>

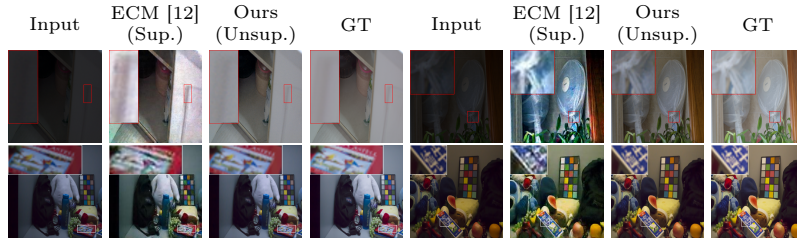
(b) generalizability Results. “E” denotes pre-training on MSEC Dataset [2]; “R” signifies pretraining on our Radiometry Correction Dataset. We use LOL dataset [40] for evaluation.

**Table 1:** Quantitative comparison of performance and generalizability.**Fig. 6:** Results on MSEC Dataset [2]. We take images from [12] and compare with our model.

LOL dataset comprises 485 pairs of training images, with each pair featuring one low-light image and one corresponding normal-light image. For our assessment, we exclusively employed the training image pairs from LOL dataset to serve as our evaluation set.

For evaluation, we selected one ground truth image from MSEC Dataset to serve as our reference image. As illustrated in Tab. 1b, while the performance of supervised models significantly decreases, our results demonstrate stability and outperform ECM [12]. It is noteworthy that ECM [12], when trained on our Radiometry Correction Dataset, displayed superior generalizability compared to its performance on MSEC Dataset. The visualization in Fig. 7 reveals that our approach ensures better color consistency without stylistic biases. These results underscores the robustness of radiometry correction.

**Results on Radiometry Correction Dataset.** To solely assess radiometric performance, we trained and evaluated exposure correction methods using our Radiometry Correction Dataset. We benchmarked our UEC method against the SOTA supervised model, ECM [12]. Both methods demonstrated similar performance in terms of PSNR; however, our UEC method significantly outperformed in terms of SSIM. The results are detailed in Tab. 2a. To ensure an fair



**Fig. 7:** Generalizability performance by ECM [12] and our UEC. We pretrain models on MSEC Dataset [2] and assess their performance on LOL Dataset [40].

		ECM [12]		Ours				ECM [12]	
EV		PSNR	SSIM	PSNR	SSIM	EV		PSNR	F1-Score
-2		<b>20.122</b>	0.718	19.475	<b>0.812</b>	-2		15.510	0.912
-1		<b>20.301</b>	0.735	20.144	<b>0.847</b>	-1		15.983	0.918
0		20.546	0.751	<b>20.968</b>	<b>0.884</b>	0		16.449	0.923
+1		20.667	0.758	<b>21.614</b>	<b>0.907</b>	+1		16.730	0.925
+2		20.588	0.758	<b>21.187</b>	<b>0.897</b>	+2		16.758	0.926
+3		<b>20.445</b>	0.743	19.901	<b>0.862</b>	+3		16.443	0.925
Avg.		20.445	0.744	<b>20.548</b>	<b>0.868</b>	Avg.		16.312	0.922

(a) Results on Radiometry Correction Dataset

		ECM [12]		Ours	
EV		PSNR	F1-Score	PSNR	F1-Score
-2		15.510	0.912	<b>20.059</b>	<b>0.959</b>
-1		15.983	0.918	<b>21.616</b>	<b>0.966</b>
0		16.449	0.923	<b>23.746</b>	<b>0.974</b>
+1		16.730	0.925	<b>25.550</b>	<b>0.978</b>
+2		16.758	0.926	<b>24.090</b>	<b>0.975</b>
+3		16.443	0.925	<b>20.929</b>	<b>0.964</b>
Avg.		16.312	0.922	<b>22.665</b>	<b>0.969</b>

(b) Edge detection results on Radiometry Correction Dataset.

**Table 2:** Quantitative comparison of performance and generalizability.

comparison, we utilized the official implementations and adhered to the default hyperparameter configurations.

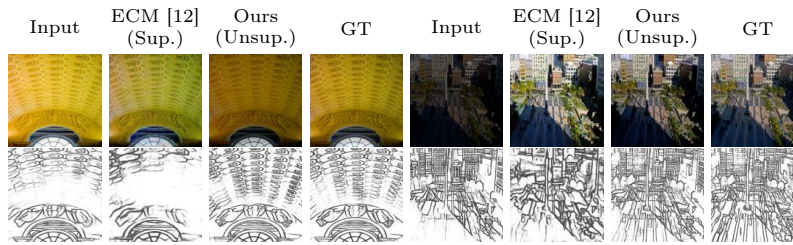
**Results of Varying Exposure Manipulation.** Compared to traditional supervised methods that learn the mapping from a poorly exposed image to a well-exposed one, our method can adjust the exposure within a certain range. To demonstrate the effectiveness of our method at varying exposure levels, we generate images with different exposures by controlling the exposure of the reference image. The results are shown in Tab. 3. Our method performs better on the MSEC dataset compared to our dataset. This discrepancy might be attributed to the wider range of exposure values in our dataset, which varies from -2EV to +3EV, presenting a greater challenge. In contrast, the MSEC dataset has a narrower exposure range of -1.5EV to +1.5EV. In our dataset, underexposure and overexposure lead to a loss of texture, making it difficult to achieve satisfactory results solely by adjusting brightness. Additionally, in our dataset, darker images yield better results, possibly because they are closer to completely black, making conversion easier. However, overexposed images are not entirely white. The large exposure range from underexposure to overexposure further increases the challenge. We provide the visualized results in our supplementary material.

EV	-1.5	-1.0	0	+1.0	+1.5
PSNR	27.764	27.733	27.823	27.959	28.097

(a) Exposure control on MSEC dataset.

EV	-2	-1	0	+1	+2	+3
PSNR	22.577	20.528	18.336	17.820	15.752	15.138

(b) Exposure control on Our dataset.

**Table 3:** Performance of exposure control on both datasets.**Fig. 8:** Edge detection performance by ECM [12] and our UEC.

**Results on Edge Detection.** Exposure correction not only enhances image aesthetics but also supports downstream tasks in computer vision. Since many algorithms rely on images with standard exposure, variations in real-world exposure can challenge their effectiveness. By improving visibility and detail clarity, exposure correction enhances image features crucial for accurate object recognition. However, it is important to maintain feature integrity to avoid compromising details with artifacts. To assess its effect on crucial features, we focused on edge detection—a task highly sensitive to exposure changes—using the LDC network [36] as a benchmark to compare performance across different exposures.

Using Radiometry Correction Dataset, we compared ECM’s [12] method against our UEC approach. The findings, detailed in Tab. 2b and illustrated in Fig. 8, highlight ECM’s underperformance, even below that of unmodified inputs. This shortfall is attributed to ECM’s tendency to lose fine details, negating the benefits of exposure correction. Conversely, our UEC method outperforms, demonstrated by higher PSNR values indicating superior image quality. Notably, our unsupervised approach excels in adapting to varied exposures for diverse computer vision tasks, requiring only the preservation of RAW files during data collection.

**Impact of Reference Images on UEC Results.** Fig. 9 demonstrates the impact of utilizing different reference images on the outcomes of our UEC model. Notably, the results show minor variations within a certain range, while main-

**Fig. 9:** Impact of Reference Images on UEC Results.

Method	Parameters	Model Size	Time(GPU)	Time(CPU)
ECM. [12]	182,443,267	695 MB	7.08ms	149.39ms
ours	19,388	0.079 MB	1.46ms	6.38ms

**Table 4:** Comparison at a resolution of  $256 \times 256$  between our UEC model and ECM [12].

taining their overall performance. This observation emphasizes the UEC model’s robust generalizability, its proficiency in extracting exposure features from diverse images, and its competence in achieving exposure alignment across a range of scenes.

### 5.3 Time and Memory Consumption

Our approach achieves comparable performance to ECM [12] with only 0.01% of their parameters. It supports real-time operation at 4K resolution, processing high-definition images directly and avoiding low-level feature loss from down-sampling. As shown in Tab. 4, our method is 4.85 times faster on GPU and 23.41 times faster on CPU compared to ECM [12]. Additional results can be found in the supplementary material.

## 6 Conclusion

In this study, we introduce an unsupervised methodology for correcting exposure, overcoming three primary limitations observed in prior methods: the dependence on labor-intensive paired datasets, constrained generalizability, and the deterioration of low-level features. To address these issues, our approach unfolds in three stages. Initially, we establish an unsupervised learning framework grounded in the inherent principles of exposure, thereby obviating the requirement for ground truth data. Subsequently, we eschew the traditional reliance on ground truth, which often incorporates subjective stylistic biases, and instead, draw insights from a diverse array of images differing solely in their radiometric properties. Finally, we employ color transformation techniques to maintain the intrinsic pixel relationships, effectively preserving low-level features and enhancing the robustness of our method.

## References

1. Affi, M., Derpanis, K.G., Ommer, B., Brown, M.S.: Learning to correct overexposed and underexposed photos. arXiv preprint arXiv:2003.11596 **13** (2020)
2. Affi, M., Derpanis, K.G., Ommer, B., Brown, M.S.: Learning Multi-Scale Photo Exposure Correction. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition pp. 9153–9163 (2021)
3. Aly, H.A., Dubois, E.: Image up-sampling using total-variation regularization with a new observation model. IEEE Transactions on Image Processing **14**(10), 1647–1659 (2005)
4. Bhattacharya, J., Modi, S., Gregorat, L., Ramponi, G.: D2bgan: A dark to bright image conversion model for quality enhancement and analysis tasks without paired supervision. IEEE Access **10**, 57942–57961 (2022)
5. Bychkovsky, V., Paris, S., Chan, E., Durand, F.: Learning photographic global tonal adjustment with a database of input/output image pairs. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 97–104. IEEE (2011)
6. Cai, J., Gu, S., Zhang, L.: Learning a deep single image contrast enhancer from multi-exposure images. IEEE Transactions on Image Processing **27**(4), 2049–2062 (2018)
7. Chai, Y., Giryes, R., Wolf, L.: Supervised and unsupervised learning of parameterized color enhancement. In: The IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 992–1000 (2020)
8. Chen, Y.S., Wang, Y.C., Kao, M.H., Chuang, Y.Y.: Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6306–6314 (2018)
9. Choi, Y., Choi, M., Kim, M., Ha, J.W., Kim, S., Choo, J.: Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8789–8797 (2018)
10. Dayley, L.D., Dayley, B.: Photoshop CS5 Bible. John Wiley & Sons (2010)
11. Deng, Y., Loy, C., Tang, X.: Aesthetic-driven image enhancement by adversarial learning. In: 2018 ACM Multimedia Conference on Multimedia Conference (MM). pp. 870–878. ACM (2018)
12. Eyiokur, F.I., Yaman, D., Ekenel, H.K., Waibel, A.: Exposure Correction Model to Enhance Image Quality. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops **2022-June**, 675–685 (2022)
13. Gharbi, M., Chen, J., Barron, J.T., Hasinoff, S.W., Durand, F.: Deep bilateral learning for real-time image enhancement. ACM Transactions on Graphics (TOG) **36**(4), 118 (2017)
14. Guo, C., Li, C., Guo, J., Loy, C., Hou, J., Kwong, S., Cong, R.: Zero-reference deep curve estimation for low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1780–1789 (2020)
15. He, J., Liu, Y., Qiao, Y., Dong, C.: Conditional sequential modulation for efficient global image retouching. In: European Conference on Computer Vision (ECCV). pp. 679–695. Springer (2020)
16. Hu, S., Yan, J., Deng, D.: Contextual information aided generative adversarial network for low-light image enhancement. Electronics **11**(1), 32 (2021)



17. Ibrahim, H., Kong, N.S.P.: Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics* **53**(4), 1752–1758 (2007)
18. Ignatov, A., Kobyshev, N., Timofte, R., Vanhoey, K., Van Gool, L.: Dslr-quality photos on mobile devices with deep convolutional networks. In: *Proceedings of the IEEE international conference on computer vision*. pp. 3277–3285 (2017)
19. Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., Wang, Z.: Enlightengan: Deep light enhancement without paired supervision. *IEEE transactions on image processing* **30**, 2340–2349 (2021)
20. Kim, H., Koh, Y., Kim, C.: Global and local enhancement networks for paired and unpaired image enhancement. In: *European Conference on Computer Vision (ECCV)*. pp. 339–354. Springer (2020)
21. Kotovenko, D., Sanakoyeu, A., Lang, S., Ommer, B.: Content and style disentanglement for artistic style transfer. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 4422–4431 (2019)
22. Land, E.H., McCann, J.J.: Lightness and retinex theory. *Josa* **61**(1), 1–11 (1971)
23. Li, C., Guo, C., Ai, Q., Zhou, S., Loy, C.C.: Flexible piecewise curves estimation for photo enhancement. *arXiv preprint arXiv:2010.13412* (2020)
24. Li, C., Guo, C., Feng, R., Zhou, S., Loy, C.C.: Cudi: Curve distillation for efficient and controllable exposure adjustment. *arXiv preprint arXiv:2207.14273* (2022)
25. Li, C., Guo, C., Loy, C.C.: Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44**(8), 4225–4238 (2021)
26. Liu, E., Li, S., Liu, S.: Color enhancement using global parameters and local features learning. In: *Proceedings of the Asian Conference on Computer Vision* (2020)
27. Liu, R., Ma, L., Zhang, J., Fan, X., Luo, Z.: Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 10561–10570 (2021)
28. Liu, Y., He, J., Chen, X., Zhang, Z., Zhao, H., Dong, C., Qiao, Y.: Very lightweight photo retouching network with conditional sequential modulation. *CoRR abs/2104.06279* (2021)
29. Loh, Y.P., Chan, C.S.: Getting to know low-light images with the exclusively dark dataset. *Computer Vision and Image Understanding* **178**, 30–42 (2019)
30. Moran, S., Marza, P., McDonagh, S., Parisot, S., Slabaugh, G.: Deeplpf: Deep local parametric filters for image enhancement. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 12826–12835 (2020)
31. Moran, S., McDonagh, S., Slabaugh, G.: Curl: Neural curve layers for global image enhancement. In: *2020 25th International Conference on Pattern Recognition (ICPR)*. pp. 9796–9803. IEEE (2021)
32. Nsampi, N.E., Hu, Z., Wang, Q.: Learning exposure correction via consistency modeling. In: *Proc. Brit. Mach. Vision Conf.* (2021)
33. Rahman, Z.u., Jobson, D.J., Woodell, G.A.: Retinex processing for automatic image enhancement. *Journal of Electronic imaging* **13**(1), 100–110 (2004)
34. Ren, W., Liu, S., Ma, L., Xu, Q., Xu, X., Cao, X., Du, J., Yang, M.H.: Low-light image enhancement via a deep hybrid network. *IEEE Transactions on Image Processing* **28**(9), 4364–4375 (2019)
35. Sara, U., Akter, M., Uddin, M.S.: Image quality assessment through fsim, ssim, mse and psnr—a comparative study. *Journal of Computer and Communications* **7**(3), 8–18 (2019)



36. Soria, X., Pomboza-Junez, G., Sappa, A.D.: Ldc: lightweight dense cnn for edge detection. *IEEE Access* **10**, 68281–68290 (2022)
37. Wang, B., Yu, Y., Xu, Y.: Example-based image color and tone style enhancement. *ACM Transactions on Graphics (TOG)* **30**(4), 1–12 (2011)
38. Wang, T., Li, Y., Peng, J., Ma, Y., Wang, X., Song, F., Yan, Y.: Real-time image enhancer via learnable spatial-aware 3d lookup tables. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 2471–2480 (October 2021)
39. Wang, Y., Li, X., Xu, K., He, D., Zhang, Q., Li, F., Ding, E.: Neural color operators for sequential image retouching. In: *European Conference on Computer Vision*. pp. 38–55. Springer (2022)
40. Wei, C., Wang, W., Yang, W., Liu, J.: Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560* (2018)
41. Yan, Z., Zhang, H., Wang, B., Paris, S., Yu, Y.: Automatic photo adjustment using deep neural networks. *ACM Transactions on Graphics (TOG)* **35**(2), 11 (2016)
42. Yang, K.F., Cheng, C., Zhao, S.X., Yan, H.M., Zhang, X.S., Li, Y.J.: Learning to adapt to light. *International Journal of Computer Vision* **131**(4), 1022–1041 (2023)
43. Zeng, H., Cai, J., Li, L., Cao, Z., Zhang, L.: Learning image-adaptive 3d lookup tables for high-performance photo enhancement in real-time. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* (2020)
44. Zhang, Y., Zhang, J., Guo, X.: Kindling the darkness: A practical low-light image enhancer. In: *Proceedings of the 27th ACM international conference on multimedia*. pp. 1632–1640 (2019)