Transition-constant Normalization for Image Enhancement

Jie Huang[1,*], **Man Zhou**[1,2,*], **Jinghao Zhang**[1], **Gang Yang**[1], **Mingde Yao**[1], **Chongyi Li**[3], **Zhiwei Xiong**[1], **Feng Zhao**[1†]

[1]University of Science and Technology of China, China
[2]Nanyang Technological University, Singapore
[3]Nankai University, China

`{hj0117, manman, jhaozhang, mdyao, yg1997}@mail.ustc.edu.cn,`
`lichongyi25@gmail.com, {zwxiong, fzhao956}@ustc.edu.cn,`

Abstract

Normalization techniques that capture image style by statistical representation have become a popular component in deep neural networks. Although image enhancement can be considered as a form of style transformation, there has been little exploration of how normalization affect the enhancement performance. To fully leverage the potential of normalization, we present a novel Transition-Constant Normalization (TCN) for various image enhancement tasks. Specifically, it consists of two streams of normalization operations arranged under an invertible constraint, along with a feature sub-sampling operation that satisfies the normalization constraint. TCN enjoys several merits, including being parameter-free, plug-and-play, and incurring no additional computational costs. We provide various formats to utilize TCN for image enhancement, including seamless integration with enhancement networks, incorporation into encoder-decoder architectures for downsampling, and implementation of efficient architectures. Through extensive experiments on multiple image enhancement tasks, like low-light enhancement, exposure correction, SDR2HDR translation, and image dehazing, our TCN consistently demonstrates performance improvements. Besides, it showcases extensive ability in other tasks including pan-sharpening and medical segmentation. The code is available at *https://github.com/huangkevinj/TCNorm*.

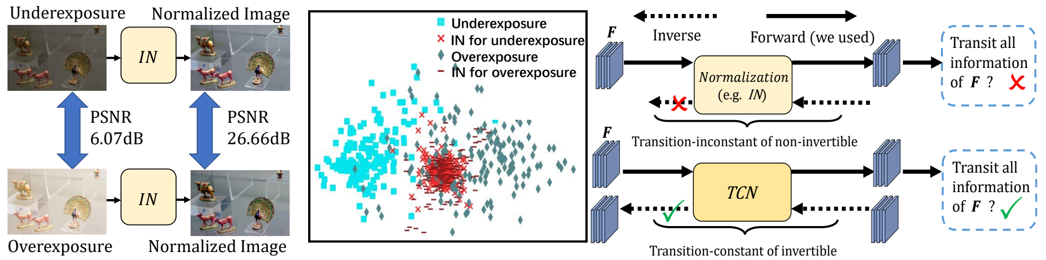
1 Introduction

Image enhancement is an important task in machine vision, which aims to improve the quality of low-visibility images captured under unfavorable light conditions (i.e., low light) by adjusting contrast and lightness. The last decades have witnessed quantities of approaches designed for image enhancement based on various hand-crafted priors [1, 2, 3, 4, 5]. However, the complex and variant adjustment procedures make it a challenging group of tasks. In addition to the common low-light image enhancement, efforts have also been directed toward solving image enhancement-like tasks, including exposure correction, image dehazing, and SDR2HDR translation.

Very recently, the deep-learning paradigm exhibits remarkable success in the image enhancement field than traditional methods [6, 7, 8]. Despite the progress, most of them focus on roughly constructing complex deep neural architectures and have not fully explored the intrinsic characterizes of lightness in networks. In fact, the lightness variants could bring difficulties to their learning procedures. This

*Both authors contributed equally to this research.
†Corresponding author.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).



(a) Apply normalization for different exposure images. (b) Transition-constant property of normalization.

Figure 1: In (a), the instance normalization (IN) captures lightness consistency representations across exposures, thus bridging their distribution gaps as shown in t-SNE. In (b), normalization techniques meet transition-inconstant problem, while proposed TCN exhibits differently with invertible ability.

motivates us to delve into the working mechanism of current neural networks that learn lightness adjustment and prescribe the right medicine customized for image enhancement.

On the other hand, normalization family such as batch and instance normalization, is specially designed for promoting the learning procedure for deep networks. It involves computing statistical representation and normalizing the corresponding distribution, which has been shown to capture image style through statistical representation [9]. Image enhancement, which aims to restore lightness-corrupted images to their normal versions, can be viewed as a style transformation inherently linked to normalization techniques (see Fig. 1 (a)). However, existing methods have rarely explored the potential of the normalization technique. Inspired by the inborn connection, we thus focus on developing the normalization technique tailored for image enhancement.

In this work, we propose a novel operation called transition-constant normalization (TCN) for image enhancement tasks. The TCN operation aims to normalize partial representations to ensure consistent learning while preserving constant information for image reconstruction. As illustrated in Fig. 1 (b), we construct the TCN within an invertible format to enable seamless information transmission to subsequent layers for reconstructing enhanced results. The TCN is designed with two key rules (see Fig. 2): (1) We organize the operations in the normalization layer into two streams, following the principle of invertible information transmission, thereby maintaining constant information transition. (2) We incorporate a subsampling operation that divides the features into two streams with consistent statistical properties. One stream provides the statistics for normalizing the other stream, satisfying the normalization requirement. Notably, the TCN requires no parameters, making it a convenient and orthogonal addition to existing enhancement architectures for improving their performance.

To facilitate its application, we present multiple usage formats for the TCN in image enhancement: (1) Integration into existing enhancement networks, allowing for seamless incorporation and performance improvement; (2) Plug-in capability in encoder-decoder architectures for downsampling and recomposition of information; (3) Construction of a lightweight architecture based on TCN, striking a balance between performance and computational cost. Through extensive experiments across various image enhancement tasks, we consistently observe performance gains by integrating our TCN.

The contributions of this work are summarized as follows: 1) We present a novel perspective on image enhancement using a dedicated normalization technique. This technique enhances the learning of lightness adjustments by modeling consistent normalized features, while ensuring their complementarity for reconstructing the results. 2) We construct the Transition-constant Normalization (TCN) by organizing normalization operations to satisfy the invertible mechanism, ensuring constant feature normalization and information transition. 3) Our proposed TCN is compatible with existing enhancement architectures, allowing for convenient integration and performance improvement. Furthermore, we can derive multiple implementation formats for TCN and explore its applicability in various tasks, highlighting its potential for wide-ranging applications.

2 Related Work

Image enhancement tasks. Image enhancement tasks aim to improve the quality of low-visibility images by adjusting lightness and contrast components (e.g., illumination, color, and dynamic range).

Recent years have witnessed rapid development in the related areas [10, 11, 12, 13]. For low-light image enhancement, algorithms are designed to enhance the visibility of images captured under low-light conditions [14, 15, 16, 17, 18, 19, 7, 20, 21, 22, 23]. In the exposure correction task, methods are focused on correcting both underexposure and overexposure to normal exposure [15, 24, 25, 26, 27]. For SDR2HDR translation, this task aims to design methods to convert images from a low-dynamic range to a high-dynamic range [28, 29, 30, 31, 32]. While in image dehazing, this task requires methods to enhance the contrast and recover the color shift problems [33, 34, 35, 36, 37]. To this end, image enhancement tasks cover variant scenes and remain challenges to be solved.

Normalization techniques. Normalization techniques have been studied for a long time [38, 39, 40, 41]. Batch Normalization (BN) [38] normalizes the features along the batch dimension that stabilizes the optimization procedure. Instance Normalization (IN) [9] focuses on normalizing the instance-level statistics of features, which has been widely employed in style transfer tasks [42, 43]. Some other variants of normalization, including Layer Normalization (LN) [39], Group Normalization (GN) [44], and Position Normalization (PN) [45] have been proposed for facilitating the application of networks.

3 Method

In this section, we first briefly revisit the normalization techniques and then detail the design and mechanism of the proposed TCN. Finally, we present the variants of the TCN as implementation.

3.1 Preliminaries

Given a batch of features $x \in \mathbb{R}^{\mathrm{N}\times\mathrm{C}\times\mathrm{H}\times\mathrm{W}}$, where N, C, H and W represent batch size, channel numbers, the spatial height and width, respectively. Let x_{ncij} and $\hat{x}_{ncij}$ denote a pixel before and after normalization, where $\mathrm{n} \in [1, \mathrm{N}], \mathrm{c} \in [1, \mathrm{C}], \mathrm{i} \in [1, \mathrm{H}], \mathrm{i} \in [1, \mathrm{W}]$. Without taking considering into affined parameters, we can express normalization operation as:

$$\hat{x}_{ncij} = \mathrm{Norm}(x_{ncij}) = \frac{x_{ncij} - \mu_k}{\sqrt{\sigma_k^2 + \epsilon}}, \tag{1}$$

where μ_k and σ_k denote the feature mean and standard deviation, ϵ is a small constant to preserve numerical stability. $k \in \{$IN, BN, LN, GN$\}$ is to distinguish different normalization formats. Within the above normalization family, the calculation of μ_k and σ_k is different and are expressed as:

$$\mu_k = \frac{1}{|I_k|} \sum_{n,c,i,j \in I_k} x_{ncij}, \quad \sigma_k = \sqrt{\frac{1}{|I_k|} \sum_{n,c,i,j \in I_k} (x_{ncij} - \mu_k)^2}, \tag{2}$$

1) **IN**: $I_k : I_{IN} = \{(i, j) | i \in [1, \mathrm{H}], \mathrm{j} \in [1, \mathrm{W}]\}$;
2) **BN**: $I_k : I_{BN} = \{(n, i, j) | n \in [1, \mathrm{N}], \mathrm{i} \in [1, \mathrm{H}], \mathrm{j} \in [1, \mathrm{W}]\}$;
3) **LN**: $I_k : I_{LN} = \{(c, i, j) | c \in [1, \mathrm{C}], \mathrm{i} \in [1, \mathrm{H}], \mathrm{j} \in [1, \mathrm{W}]\}$;
4) **GN**: $I_k : I_{GN} = \{(c, i, j) | c \in [\mathrm{g}, \mathrm{g} + \mathrm{C}/\mathrm{G}], \mathrm{i} \in [1, \mathrm{H}], \mathrm{j} \in [1, \mathrm{W}] | \mathrm{g} \in [1, \mathrm{G}]\}$.

where I_k is a set of pixels, $|I_k|$ denotes the number of pixels and G is the group division.

Within deep neural networks, Eq. 1 is often affined with scaling and shifting parameters α and β:

$$\hat{x}_{ncij} = \alpha \cdot \mathrm{Norm}(x_{ncij}) + \beta = \alpha \frac{x_{ncij} - \mu_k}{\sqrt{\sigma_k^2 + \epsilon}} + \beta. \tag{3}$$

It is well-known that lightness can be considered as a kind of style and Instance normalization can facilitate consistent style information [9] and image enhancement network optimization due to bridging the gap of different lightness representations.

Verifying normalization effect for lightness. Given an image x_a and its lightness-adjusted version x_b, the relationship between x_a and x_b can be expressed in correction procedure [46, 47, 48] as:

$$x_b = \Lambda x_a^\gamma, \tag{4}$$

where Λ is a linear transformation and γ is for global non-linear adaption and is close to 1 when content is not severely changed. Therefore, the p-norm distance between their normalized versions is:

$$||\mathrm{f}(x_a) - \mathrm{f}(x_b)\|_p = |\frac{x_a - \mu_a}{\sigma_a} - \frac{\Lambda x_a^\gamma - \mu_b}{\sigma_b}\|_p \approx |\frac{x_a - \mu_a}{\sigma_a} - \frac{\Lambda x_a - \Lambda \mu_a}{\Lambda \sigma_a}\|_p \ll \|x_a - \Lambda x_a^\gamma\|_p. \tag{5}$$

Therefore, normalization reduces the distance of different lightness, which is also validated in Fig. 1(a). However, normalizing the statistics itself often blocks the information flow, which hinders the network from reconstructing the final results [49] and we describe it as follows.

Transition-inconstant of Normalization. Referring to $\mathrm{Norm}(\cdot)$ in the Eq. 1 as f, its Jacobian matrix is expressed as:

$$J_f(x) = \begin{bmatrix} \frac{\partial \hat{x_0}}{x_0} & \cdots & \frac{\partial \hat{x_0}}{x_0} \\ & \ddots & \vdots \\ \frac{\partial \hat{x_0}}{x_0} & & \frac{\partial \hat{x_0}}{x_0} \end{bmatrix}_{m\times m} = \begin{bmatrix} \frac{1}{\sigma_k} & \cdots & \frac{1}{\sigma_k} \\ & \ddots & \vdots \\ \frac{1}{\sigma_k} & & \frac{1}{\sigma_k} \end{bmatrix}_{m\times m} \tag{6}$$

Therefore, the calculation of the above Jacobian matrix is **det** $(J_f(x))$ = 0, denoting the normalization operation is not invertible and resulting in transition-inconstant. Meanwhile, in practice, previous works have demonstrated that IN would lead to severe information loss [42] and huge representation ability changes [50, 51], while LN and BN can keep almost all of the original information representation ability. However, IN is more suitable than BN and LN for image enhancement tasks due to its strong capability of capturing and affecting style information, which is crucial for image enhancement. To this end, the main goal of this paper is to introduce a new mechanism that enables the IN can keep the information representation ability for image enhancement.

3.2 Transition-constant Normalization (TCN)

Based on the above analysis, we aim to refresh the normalization technique to enable it to transmit information constantly while normalizing the features. To this end, we introduce the TCN as shown in Fig. 2, which is free of parameters and is convenient to implement. Since IN can normalize different lightness effectively and thus is useful for image enhancement, we design the TCN based on the IN as its default implementation format in this paper.

Operation description. We construct the TCN by applying the normalization operations with a two streams flow design with subsampled features, where one stream provides the statistical information for normalizing another stream. Specifically, the feature $F \in \mathbb{R}^{\mathrm{B}\times\mathrm{C}\times\mathrm{H}\times\mathrm{W}}$ is firstly subsampled to $F_s \in \mathbb{R}^{\mathrm{B}\times 4\mathrm{C}\times\frac{\mathrm{H}}{2}\times\frac{\mathrm{W}}{2}}$ according to the unshuffle operation [52] as shown in Fig. 2 (a), which is:

$$F_s^{ab} = F[:, :, a ::, b ::], \quad a, b \in \{0, 1\}, \tag{7}$$

where a and b denote the subsampling index. We divide F_s into two features F_1 and F_2 with two groups according to the sampling index (i, j) as:

$$F_1 = \mathrm{Concat}(F_s^{01}, F_s^{10}), F_2 = \mathrm{Concat}(F_s^{00}, F_s^{11}) \tag{8}$$

where $\mathrm{Concat}(\cdot, \cdot)$ denotes the concatenate operation along the channel dimension.

Then, we calculate the mean μ_2 and standard deviation σ_2 of one stream feature F_2 in IN format, which are derived by setting I_k in Eq. 2 as I_{IN}:

$$\mu_2 = \frac{1}{\mathrm{HW}} \sum_{i\in[1,\mathrm{H}], \mathrm{j}\in[1,\mathrm{W}]} F_{ij}, \quad \sigma_2 = \sqrt{\frac{1}{\mathrm{HW}} \sum_{i\in[1,\mathrm{H}], \mathrm{j}\in[1,\mathrm{W}]} (F_{ij} - \mu_2)^2}. \tag{9}$$

These statistics are utilized to normalize the feature F_1 as the output in this stream:

$$\hat{F}_1 = \frac{F_1 - \mu_2}{\sqrt{\sigma_2^2 + \epsilon}}. \tag{10}$$

Next, we subtract the feature F_2 and $\hat{F}_1$ and obtain the output of another stream:

$$\hat{F}_2 = F_2 - \hat{F}_1 = F_2 - \frac{F_1 - \mu_2}{\sqrt{\sigma_2^2 + \epsilon}}. \tag{11}$$

Finally, the two stream features are sampled with pixel shuffle operation to the original resolution with the shape of $[\mathrm{N}, \frac{\mathrm{C}}{2}, \mathrm{H}, \mathrm{W}]$. They are further concatenated in the channel dimension as $\hat{F} \in \mathbb{R}^{\mathrm{B}\times\mathrm{C}\times\mathrm{H}\times\mathrm{W}}$, which is the output of the TCN. This procedure is expressed as:

$$\hat{F} = \mathrm{Pixshuffle}(\mathrm{Concat}(\hat{F}_1, \hat{F}_2)), \tag{12}$$

where $\mathrm{Pixshuffle}(\cdot)$ denotes the pixel shuffle operation [53]. We verify the above procedures satisfy the transition-constant and normalization ability as below, respectively.

1) Verify the transition-constant ability. We validate the above two-stream design satisfies the invertible procedure and thus is transition-constant. To this end, Eq. 10 and Eq. 11 are re-written as:

$$\begin{aligned} \hat{F}_1 &= (F_1 - M(F_2)) \oslash S(F_2), \\ \hat{F}_2 &= F_2 - \hat{F}_1, \end{aligned} \tag{13}$$

where $S(\cdot)$ and $M(\cdot)$ denote standard deviation and mean functions, $\oslash$ denotes element division.

Inspired by the proof in RealNVP's [54] transformation, we need to calculate the Jacobian matrix of Eq. 13 (denote it as g), which is more intuitively written as:

$$\begin{aligned} \hat{F}_1 &= (F_1 - M(F_2)) \oslash S(F_2), \\ \hat{F}_2 &= F_2 - F_1 \oslash S(F_2) + M(F_2) \oslash S(F_2), \end{aligned} \tag{14}$$

We derive its Jacobian matrix (detailed in the supplementary) as:

$$J_g = \begin{bmatrix} \frac{\partial \hat{F}_1}{\partial F_1} & \frac{\partial \hat{F}_1}{\partial F_2} \\ \\ \frac{\partial \hat{F}_2}{\partial F_1} & \frac{\partial \hat{F}_2}{\partial F_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\partial S(F_2)} & 0 \\ \\ -1 & 1 \end{bmatrix} \tag{15}$$

Here, the above Jacobian matrix is further calculated as:

$$\mathbf{det}(J_g) = \frac{1}{\partial S(F_2)} \neq 0 \tag{16}$$

Upon the $\mathbf{det}(J_g) \neq 0$, it indicates that J_g is full rank, verifying the invertible property of TCN and further the transition-constant ability. To highlight, the TCN is an invertible function and would not block information flow, leading to the information transition constant for image reconstruction. We further present the relation of the TCN and the invertible operation more directly in the supplementary.

Formats	Redefine I_k to Eq. 9 as
TCN (IN)	I_{IN} (default I_k of Eq. 9)
TCN (BN)	$I_{BN} = \{(n,i,j) \mid n \in [1,N],$ $i \in [1,H], j \in [1,W]\}$
TCN (LN)	$I_{LN} = \{(c,i,j) \mid c \in [1,\mathrm{C}],$ $i \in [1,\mathrm{H}], \mathrm{j} \in [1,\mathrm{W}]\}$
TCN (GN)	$I_{GN} = \{(c,i,j) \mid$ $c \in [\mathrm{g}, \mathrm{g}+\mathrm{C/G}], \mathrm{i} \in [1,\mathrm{H}],$ $j \in [1,\mathrm{W}] \mid \mathrm{g} \in [1,\mathrm{G}]\}$

Table 1: The TCN family with different μ_2,σ_2 calculation in Eq. 9.

2) Verify the normalization ability. The normalization ability of the TCN is guaranteed by the pixel unshuffle operation in Eq. 7, leading to the same statistics of F_s^{ab} [55, 56]. Therefore, we have $\mu_2 \approx \mu_1, \sigma_2 \approx \sigma_1$, and the Eq. 10 is thus converted to:

$$\hat{F}_1 = \frac{F_1 - \mu_2}{\sqrt{\sigma_2^2 + \epsilon}} \approx \frac{F_1 - \mu_1}{\sqrt{\sigma_1^2 + \epsilon}}. \tag{17}$$

Therefore, it has the same format as Eq. 1, demonstrating that the operation in Eq. 10 has the normalization ability as the IN. Further, we verify the above rules by the toy experiment (see Sec. 4.1) in Fig. 4 and Fig. 5.

Discussion. The core of the TCN is the statistic calculation manner of μ_2 and σ_2 in Eq. 9 which can be generalized in a unified calculation manner in Eq. 2 and derive from other normalization formats of TCN, shown in Table 1. Note that, although GN and LN would less affect information representation ability [50], we experimentally find introducing the transition-constant design would improve their performance in the supplementary. We provide more discussions in the supplementary.

3.3 Variants of TCN for Image Enhancement

Upon the above principles of TCN, we provide the following implementation variants within image enhancement task.

The original TCN. We construct the original TCN (see Fig. 2 (a)) for image enhancement based on calculating statistics in Eq. 9 of IN format, which is plug-and-play for networks.

The affined TCN. We extend the original TCN by introducing affined parameters α and β to the normalization procedure, resulting in the affined TCN (Fig. 2 (b)). We incorporate learnable shifting

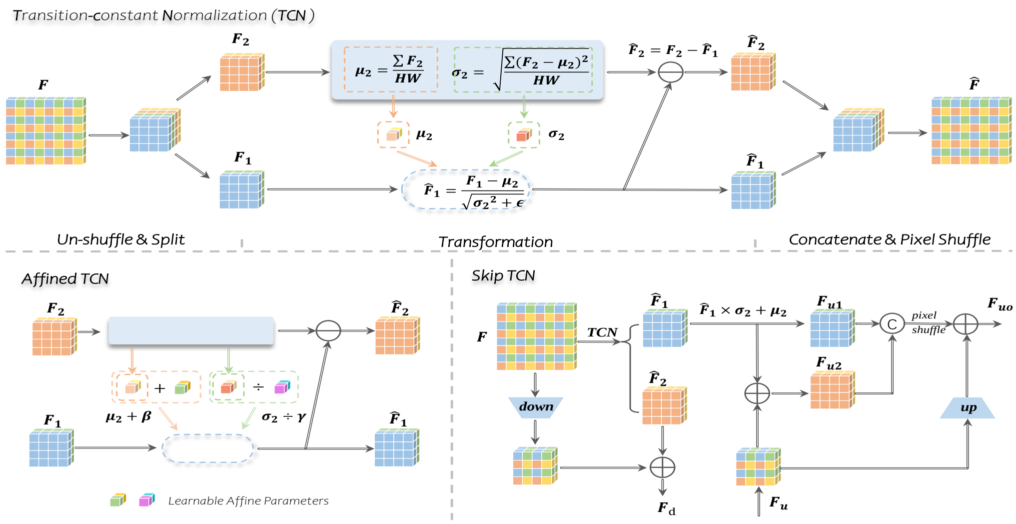


Figure 2: The illustration of the TCN operation and other TCN variants for image enhancement.

parameter β and scaling parameter γ into μ_2 and σ_2 in Eq. 9:

$$\mu_2' = \mu_2 + \beta, \quad \sigma_2' = \frac{\sigma_2}{\gamma} \tag{18}$$

where μ_2' and σ_2' represent the affined statistics. Then, we substitute Eq. 18 to Eq. 1:

$$\hat{F}_1' = \frac{F_1 - {\mu'}_2}{\sqrt{{\sigma'}_2^2 + \epsilon}} = \frac{F_1 - \mu_2 - \beta}{\sqrt{\frac{\sigma_2^2}{\gamma^2} + \epsilon}} \tag{19}$$

Since ϵ is a small constant near to 0, the Eq. 19 can be approximated as:

$$\hat{F}_1' \approx \gamma \frac{F_1 - {\mu'}_2 - \beta}{\sqrt{{\sigma'}_2^2 + \epsilon}} = \gamma \frac{F_1 - {\mu'}_2}{\sqrt{{\sigma'}_2^2 + \epsilon}} + \beta', \quad \beta' = \frac{-\gamma\beta}{\sqrt{{\sigma'}_2^2 + \epsilon}}. \tag{20}$$

Eq.20 shares a format similar to the affined normalization, with γ and β' as the learnable scaling and shifting parameters in Eq. 3. The affined TCN seamlessly integrates into image enhancement networks, serving as a plug-and-play solution. Notably, it maintains the information transition-constant property, as discussed in detail in the supplementary material.

The skip TCN. From Eq.10 and 11, the TCN generates two types of features: a domain-invariant lightness consistent feature $\hat{F}_1$ and a domain-variant lightness inconsistent feature $\hat{F}_2$. Previous studies [57, 58] have demonstrated the effectiveness of incorporating the domain-variant component into deep encoder-decoder networks while skipping the domain-invariant component to the decoder layer. In this work, we propose the skip TCN architecture, illustrated in Fig. 2 (c).

Given a feature F in an encoder layer, we convey its lightness inconsistent feature $\hat{F}_2$, obtained from Eq. 10, to the downsampled deeper layer that derives F_d using the following expression:

$$F_d = \mathrm{Down}_2(\mathrm{F}) + \hat{\mathrm{F}}_2, \tag{21}$$

where Down_2 means downsampling with a factor of 2. While for the lightness consistent feature $\hat{F}_1$ derived in Eq. 11, we skip it to the corresponding decoder layer feature F_u with the statistic μ_2 and σ_2 derived in Eq. 9. We integrate them by inverting the operation of Eq. 10 and Eq. 11:

$$\begin{aligned} F_{u1} &= \hat{F}_1 \cdot \sigma_2 + \mu_2, \quad F_{u2} = \hat{F}_1 + F_u, \\ F_{uo} &= \mathrm{Up}_2(\mathrm{F_u}) + \mathrm{Pixshuff}(\mathrm{Concat}(\mathrm{F_{u1}}, \mathrm{F_{u2}})), \end{aligned} \tag{22}$$

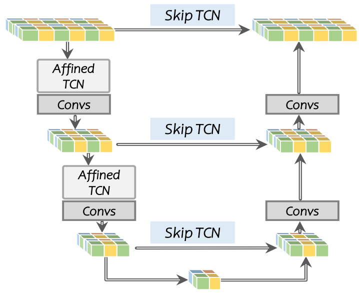


Figure 3: The overview of the TCN-Net.

Here, the upsampling operation with a factor of 2 is denoted as Up_2, and F_{uo} represents the skip TCN result. The skip TCN prioritizes the processing of the lightness component while preserving lightness invariant features, mitigating learning difficulties. Further discussion is provided in the supplementary material.

Construct a very efficient TCN-based Network. We introduce TCN-Net, an efficient network architecture depicted in Fig. 3, which combines affined TCN and skip TCN. This framework adopts an encoder-decoder-based architecture with vanilla convolution blocks to depict the effectiveness of the TCN. Further details and discussions are available in the supplementary material.

4 Experiments

In this section, we validate the effectiveness and scalability of our proposed TCN on various image enhancement tasks. We provide more experimental results in the supplementary material.

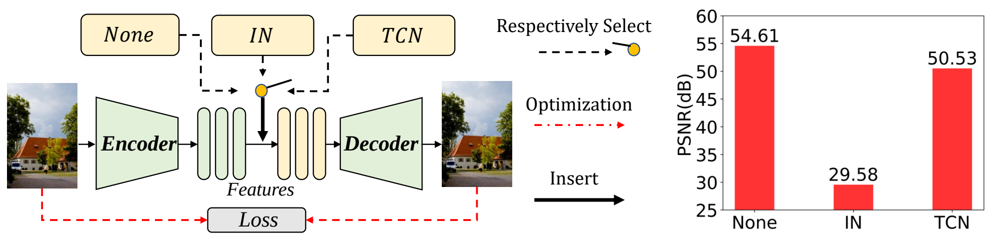


Figure 4: Toy experiment of self-reconstruction. The left is the setting of toy experiments with inserting different operations, and the right presents the self-reconstruction PSNR of testing images.

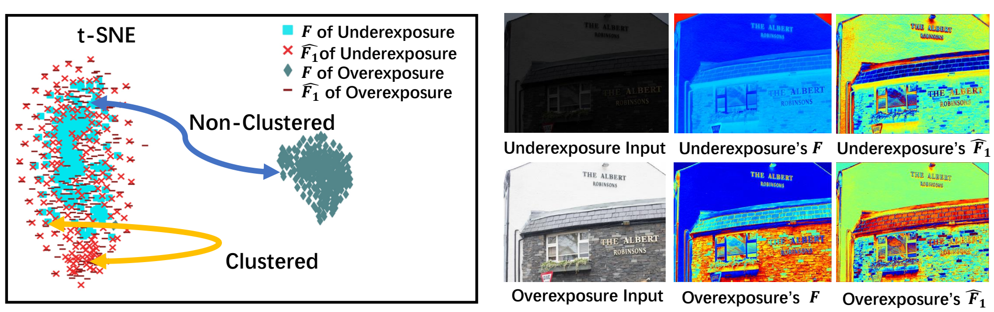


Figure 5: Feature visualization of toy experiment. In left and right parts, we show the feature in the TCN of underexposure and overexposure samples when testing them with inserting TCN in Fig. 4.

4.1 Toy Experiment

To illustrate the proposed TCN has the ability to normalize the features while transition-constant, we introduce a toy experiment as shown in Fig. 4 : we construct an encoder-decoder-based architecture for reconstructing the input image, where the TCN and other normalization formats are inserted between the encoder and decoder as different versions. Then we train this self-reconstruction architecture on 1000 samples from MIT-FiveK dataset [59] until its convergence and test the self-reconstruction effect on another 100 samples from the same dataset. The quantitative results in the right of Fig. 4 indicate that our TCN reconstructs the input image better than directly inserting IN, demonstrating the information transition-constant property. Furthermore, we test 100 underexposure samples and 100 overexposure samples from the SICE dataset, and we provide the feature distribution of F and $\hat{F}_1$ (input and normalized output of the TCN) in the left part of Fig. 5 , as well as feature maps in the right part of Fig. 5. The different exposure features processed by the TCN get to be

Settings	#Param	Flops (G)	LOL	Huawei	FiveK
DRBN (Baseline) [15]	0.53M	39.71	19.95/0.7712	20.64/0.6136	22.11/0.8684
+IN	0.53M	39.71	20.73/0.7986	21.01/0.6200	22.93/0.8727
+Original TCN	0.53M (+0)	39.71 (+0)	21.15/**0.8190**	21.12/**0.6242**	**23.98**/0.8851
+Affined TCN	0.53M (+0)	39.71 (+0)	21.29/0.8167	21.04/0.6231	23.92/**0.8858**
+Skip TCN	0.53M (+0)	39.77 (+0.07)	**21.52**/0.8271	**21.15**/0.6195	23.82/0.8832
SID (Baseline) [14]	7.40M	51.06	20.85/0.7845	19.68/0.6050	21.49/0.8425
+IN	7.40M	51.06	20.51/0.7858	20.09/0.6034	21.75/0.8453
+Original TCN	7.40M (+0)	51.06 (+0)	21.43/0.7913	20.53/0.6067	23.11/0.8581
+Affined TCN	7.40M (+0)	51.06 (+0)	21.35/0.7867	20.62/0.6077	23.20/0.8624
+Skip TCN	7.41M (+0.01)	51.42 (+0.36)	**21.92/0.8056**	**20.76/0.6083**	**23.61/0.8704**
TCN-Net	0.012M	0.97	22.08/0.7895	20.99/0.6121	23.47/0.8663

Table 2: Comparison over low-light image enhancement in terms of PSNR/MS-SSIM.

Settings	MSEC	SICE
DRBN (Baseline) [15]	19.52/0.8309	17.65/0.6798
+IN	21.98/0.8463	20.15/0.6947
+Original TCN	22.37/0.8533	20.74/0.7133
+Affined TCN	22.41/0.8504	**20.85/0.7192**
+Skip TCN	**22.48/0.8572**	20.65/0.7159
SID (Baseline) [14]	19.04/0.8074	18.15/0.6540
+IN	21.36/0.8373	19.81/0.6667
+Original TCN	22.31/0.8522	20.51/0.6745
+Affined TCN	**22.43**/0.8542	**20.68**/0.6757
+Skip TCN	22.36/**0.8603**	20.64/**0.6852**
TCN-Net	22.19/0.8480	20.72/0.7024

Table 3: Comparison over exposure correction.

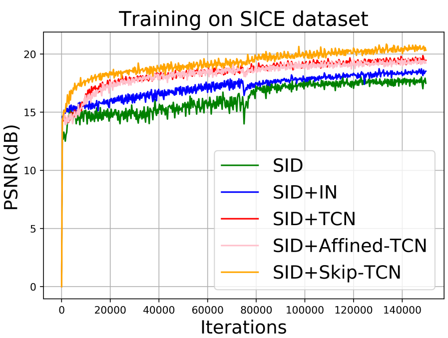


Figure 6: Training PSNR on exposure correction.

clustered, demonstrating the normalization ability of the TCN for extracting the lightness-consistence representation of different samples. We provide more discussions in the supplementary material.

4.2 Experimental Settings

Low-light Image Enhancement. Following previous works [60, 61], we employ three widely used datasets for evaluation, including LOL dataset [7], Huawei dataset [60] and MIT-FiveK dataset [59]. We employ two different image enhancement networks, DRBN [15] and SID [14] as baselines.

Exposure Correction. Following [62], we adopt MSEC dataset [24] and SICE dataset [63] for evaluations. The above two architectures, i.e., DRBN [15] and SID [14] are regarded as baselines.

SDR2HDR Translation. Following [30], we choose the SRITM dataset [31] and HDRTV dataset [30] for evaluation. We employ the structures of NAFNet [64] with its three basic units as the baseline in the experiments.

Image Dehazing. Following [33], we employ the RESIDE dataset [65] consisting of Indoor and Outdoor parts for evaluations. We adopt the network of PFFNet [66] as the baseline for validation.

4.3 Implementation Details

Since there exist three TCN formats in Sec. 3.3, we respectively integrate them into the baseline to conduct experiments. For comparison, we perform the experiments of baseline networks and the integration of the IN operation. Additionally, the TCN-Net in Sec. 3.3 is also performed in experiments. We train all baselines and their integrated formats following the original settings, and our TCN-Net until it converges. More implementation details are provided in the supplementary.

Settings	SRITM	HDRTV
NAFNet (Baseline) [64]	33.44/**0.9537**	36.49/0.9706
+IN	33.62/0.9491	36.62/0.9683
+Original TCN	**33.69**/0.9505	**36.94**/0.9712
+Affined TCN	<u>33.65</u>/0.9495	36.55/<u>0.9716</u>
+Skip TCN	33.51/<u>0.9513</u>	<u>36.64</u>/**0.9720**
TCN-Net	32.48/0.9439	36.78/0.9744

Table 4: Comparison over SDR2HDR translation.

Settings	Indoor	Outdoor
PFFNet (Baseline) [66]	21.74/0.8452	24.47/0.9274
+IN	23.13/0.8583	25.61/0.9309
+TCN	<u>23.57</u>/0.8635	25.63/0.9311
+Affined TCN	**23.71**/<u>0.8652</u>	**25.84**/<u>0.9312</u>
+Skip TCN	23.21/**0.8708**	<u>25.63</u>/**0.9315**
TCN-Net	24.06/0.8645	23.72/0.8572

Table 5: Comparison over image dehazing.

Figure 7: The visual comparison of low-light image enhancement on the MIT-FiveK dataset.

4.4 Comparison and Analysis

Quantitative Comparison. The model comparisons are conducted over different configurations, as illustrated in the implementation details. We present the quantitative results from Table 2 to Table 5, where the best and second-best results are highlighted in bold and underlined. As can be seen, almost all formats of the TCN that we incorporate have improved the performance across the datasets in all tasks, validating the effectiveness of our method. Specifically, integrating variants of TCN helps improve the training performance of baseline as shown in Fig. 6. In contrast, naively integrating the IN could not always bring performance improvement (i.e., the results of SID in Table 2). All the above results suggest the effectiveness of our proposed method without introducing any parameters. Moreover, the proposed TCN-Net achieves effective performance with efficiency. All the above evaluations prove the convenience of applying the TCN in image enhancement tasks.

Qualitative Comparison. We report the visual results of low-light image enhancement on the MIT-FiveK dataset [59] due to the limited space. As shown in Fig. 7, the integration of the TCN leads to a more visually pleasing effect with less lightness and color shift problems compared with the original baseline. We provide more visual results in the supplementary material.

4.5 Extensive Applications

The TCN can also be applied to other machine vision tasks that demonstrate its extensibility. Since TCN is proposed to extract lightness (a kind of style) invariant feature while keeping information transition-constant, we introduce another two tasks that are also related to style information, including pan-sharpening and medical segmentation. For pan-sharpening, it aims to fuse two style images, and we hope TCN can extract their invariant information with information preserving; For medical segmentation, there often exists a style domain gap between training and testing sets.

Extension on medical segmentation. We apply the TCN on the UNet [67] and AttUNet [68] in the medical segmentation task. We train the baseline and its integrated version on the heart segmentation task of Medical Segmentation Decathlon challenge dataset [69]. As shown in Table 6, our TCN improves and keeps the performance of U-Net and Att-Unet, respectively, while IN brings a significant performance drop. The results suggest the scalability of the TCN compared with the IN.

Extension on pan-sharpening. We apply the original TCN to the GPPNN [70] and PANNet [71] baselines in the pan-sharpening task, which is a common task in guided image super-resolution. We integrate it when extracting pan and multi-spectral features, and experimental results on WorldView II dataset [72, 73] in Fig. 8 suggest the effectiveness of the TCN.

Settings	Dice ↑	HD95↓
UNet(baseline)	0.9162	**3.9188**
UNet(+IN)	0.9171	7.7305
UNet(+TCN)	**0.9204**	<u>4.0171</u>
AttUNet(baseline)	<u>0.9182</u>	**3.5453**
AttUNet(+IN)	**0.9193**	6.8549
AttUNet(+TCN)	0.9180	<u>3.6241</u>

Table 6: Comparison over medical segmentation in terms of dice and HD95.

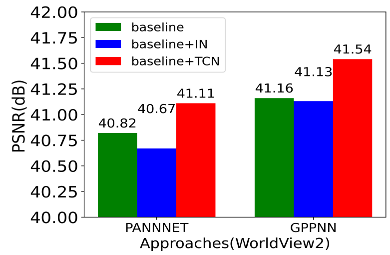


Figure 8: Comparison over pan-sharpening for GPPNN and PANNet.

5 Limitation and Discussion

Firstly, we validate the effectiveness of TCN in image enhancement tasks, while the investigation of applying TCN to other image restoration tasks will be explored in the future, such as the all-in-one image restoration task that meets similar challenges like image enhancement tasks, which has been pointed in some related works [74, 75]. Second, dedicated to image enhancement tasks, we mainly discuss the IN format of the TCN. However, other normalization formats can be future explored for other tasks. Moreover, the design formats of the TCN could inspire some areas that also require transition-constant, such as image fusion tasks [76]. Finally, the TCN could introduce very few computation burdens although it is free of parameters, which is negligible compared with its bring performance improvement. Note that the focus of this work is beyond introducing a plug-and-play operation to existing networks for performance gain. The introduced TCN can be a new choice of normalization and feature disentanglement, which excavate consistent representations while preserving information when developing a new model that requires this property.

6 Conclusion

In this paper, we introduce a new perspective that develops the normalization technique tailored for image enhancement approaches. We propose the TCN that transits the information constantly with the invertible constraint, meanwhile, it keeps the normalization ability for capturing lightness consistence representations. The proposed TCN is a general operation that can be integrated into existing networks without introducing parameters. Extensive experiments demonstrate the effectiveness and scalability of applying the TCN and its variants in various image enhancement tasks.

Broader Impact

Image enhancement is an important task that improves the quality of these images, exhibiting a high value of research and application. Our method introduces a normalization operation with information transition-constant property, which shows promising results that improve the learning ability of networks for image enhancement tasks conveniently. However, there could be negative effects brought by the proposed methodology. For example, some people may prefer the image with a dim light effect, which would be eliminated by the image enhancement algorithm. In these cases, it is suggested to combine the users' preferences to achieve customized image enhancement effects.

Acknowledgements.

This work was supported by the JKW Research Funds under Grant 20-163-14-LZ-001-004-01, and the Anhui Provincial Natural Science Foundation under Grant 2108085UD12. We acknowledge the support of GPU cluster built by MCC Lab of Information Science and Technology Institution, USTC.

References

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

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

[3] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2782–2790, 2016.

[4] Bolun Cai, Xianming Xu, Kailing Guo, Kui Jia, Bin Hu, and Dacheng Tao. A joint intrinsic-extrinsic prior model for Retinex. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4000–4009, 2017.

[5] Xueyang Fu, Delu Zeng, Yue Huang, Yinghao Liao, Xinghao Ding, and John Paisley. A fusion-based enhancing method for weakly illuminated images. *Signal Processing*, 129:82–96, 2016.

[6] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61:650–662, 2017.

[7] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep Retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018.

[8] Jiaying Liu, Dejia Xu, Wenhan Yang, Minhao Fan, and Haofeng Huang. Benchmarking low-light image enhancement and beyond. *International Journal of Computer Vision*, 129:1153–1184, 04 2021.

[9] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.

[10] Jie Huang, Zhiwei Xiong, Xueyang Fu, Dong Liu, and Zheng-Jun Zha. Hybrid image enhancement with progressive laplacian enhancing unit. In *ACM MM*, pages 1614–1622, 2019.

[11] Chongyi Li, Chunle Guo, Ruicheng Feng, Shangchen Zhou, and Chen Change Loy. Cudi: Curve distillation for efficient and controllable exposure adjustment. *arXiv preprint arXiv:2207.14273*, 2022.

[12] Naishan Zheng, Jie Huang, Man Zhou, Zizheng Yang, Qi Zhu, and Feng Zhao. Learning semantic degradation-aware guidance for recognition-driven unsupervised low-light image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3678–3686, 2023.

[13] Man Zhou, Jie Huang, Chun-Le Guo, and Chongyi Li. Fourmer: An efficient global modeling paradigm for image restoration. In *International Conference on Machine Learning*, pages 42589–42601. PMLR, 2023.

[14] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[15] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In *Proceedings of the IEEE/CVF international conference on computer vision (CVPR)*, pages 3063–3072, 2020.

[16] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5901–5910, June 2022.

[17] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the ACM International Conference on Multimedia*, pages 1632–1640, 2019.

[18] Yonghua Zhang, Xiaojie Guo, Jiayi Ma, Wei Liu, and Jiawan Zhang. Beyond brightening low-light images. *International Journal of Computer Vision*, 129(4):1013–1037, 2021.

[19] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020.

[20] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, and Yicong Zhou. Zero-shot restoration of underexposed images via robust Retinex decomposition. In *IEEE International Conference on Multimedia and Expo*, pages 1–6, 2020.

[21] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6849–6857, 2019.

[22] Naishan Zheng, Jie Huang, Qi Zhu, Man Zhou, Feng Zhao, and Zheng-Jun Zha. Enhancement by your aesthetic: An intelligible unsupervised personalized enhancer for low-light images. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 6521–6529, 2022.

[23] Naishan Zheng, Man Zhou, Yanmeng Dong, Xiangyu Rui, Jie Huang, Chongyi Li, and Feng Zhao. Empowering low-light image enhancer through customized learnable priors. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12559–12569, October 2023.

[24] Mahmoud Afifi, Konstantinos G Derpanis, Bjorn Ommer, and Michael S Brown. Learning multi-scale photo exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9157–9167, 2021.

[25] Haoyuan Wang, Ke Xu, and Rynson W.H. Lau. Local color distributions prior for image enhancement. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[26] Jie Huang, Yajing Liu, Feng Zhao, Keyu Yan, Jinghao Zhang, Yukun Huang, Man Zhou, and Zhiwei Xiong. Deep fourier-based exposure correction network with spatial frequency interaction. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[27] Jie Huang, Feng Zhao, Man Zhou, Jie Xiao, Naishan Zheng, Kaiwen Zheng, and Zhiwei Xiong. Learning sample relationship for exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9904–9913, June 2023.

[28] Xiangyu Chen, Yihao Liu, Zhengwen Zhang, Yu Qiao, and Chao Dong. HDRUnet: Single image hdr reconstruction with denoising and dequantization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 354–363, 2021.

[29] Xuan Dong, Xiaoyan Hu, Weixin Li, Xiaojie Wang, and Yunhong Wang. Miehdr cnn: Main image enhancement based ghost-free high dynamic range imaging using dual-lens systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 1264–1272, 2021.

[30] Xiangyu Chen, Zhengwen Zhang, Jimmy S Ren, Lynhoo Tian, Yu Qiao, and Chao Dong. A new journey from sdrtv to hdrtv. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4500–4509, 2021.

[31] Soo Ye Kim, Jihyong Oh, and Munchurl Kim. Deep sr-itm: Joint learning of super-resolution and inverse tone-mapping for 4k uhd hdr applications. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3116–3125, 2019.

[32] Jingwen He, Yihao Liu, Yu Qiao, and Chao Dong. Conditional sequential modulation for efficient global image retouching. In *European Conference on Computer Vision*, pages 679–695. Springer, 2020.

[33] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2154–2164, 2020.

[34] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *ICCV*, pages 4770–4778, 2017.

[35] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. Contrastive learning for compact single image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10551–10560, 2021.

[36] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7314–7323, 2019.

[37] Xiang Chen, Zhentao Fan, Pengpeng Li, Longgang Dai, Caihua Kong, Zhuoran Zheng, Yufeng Huang, and Yufeng Li. Unpaired deep image dehazing using contrastive disentanglement learning. In *European Conference on Computer Vision*, pages 632–648. Springer, 2022.

[38] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015.

[39] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

[40] Jingjing Xu, Xu Sun, Zhiyuan Zhang, Guangxiang Zhao, and Junyang Lin. Understanding and improving layer normalization. *Advances in Neural Information Processing Systems*, 32, 2019.

[41] Zachary Nado, Shreyas Padhy, D Sculley, Alexander D'Amour, Balaji Lakshminarayanan, and Jasper Snoek. Evaluating prediction-time batch normalization for robustness under covariate shift. *arXiv preprint arXiv:2006.10963*, 2020.

[42] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision*, pages 1501–1510, 2017.

[43] Yongcheng Jing, Xiao Liu, Yukang Ding, Xinchao Wang, Errui Ding, Mingli Song, and Shilei Wen. Dynamic instance normalization for arbitrary style transfer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020.

[44] Yuxin Wu and Kaiming He. Group normalization. In *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

[45] Boyi Li, Felix Wu, Kilian Q Weinberger, and Serge Belongie. Positional normalization. In *Advances in Neural Information Processing Systems*, 2019.

[46] Francesco Banterle, Alessandro Artusi, Kurt Debattista, and Alan Chalmers. *Advanced high dynamic range imaging*. CRC press, 2017.

[47] Steve Mann and Rosalind W. Picard. Being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures. 1994.

[48] Gabriel Eilertsen, Rafal Konrad Mantiuk, and Jonas Unger. A comparative review of tone-mapping algorithms for high dynamic range video. In *Computer graphics forum*, volume 36, pages 565–592. Wiley Online Library, 2017.

[49] Xin Jin, Cuiling Lan, Wenjun Zeng, Zhibo Chen, and Li Zhang. Style normalization and restitution for generalizable person re-identification. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3140–3149, June 2020.

[50] Antoine Labatie, Dominic Masters, Zach Eaton-Rosen, and Carlo Luschi. Proxy-normalizing activations to match batch normalization while removing batch dependence. *Advances in Neural Information Processing Systems*, 34:16990–17006, 2021.

[51] Ekdeep S Lubana, Robert Dick, and Hidenori Tanaka. Beyond batchnorm: towards a unified understanding of normalization in deep learning. *Advances in Neural Information Processing Systems*, 34:4778–4791, 2021.

[52] Bin Sun, Yulun Zhang, Songyao Jiang, and Yun Fu. Hybrid pixel-unshuffled network for lightweight image super-resolution. *arXiv preprint arXiv:2203.08921*, 2022.

[53] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016.

[54] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp. *International Conference on Learning Representations*, 2017.

[55] Yuqian Zhou, Jianbo Jiao, Haibin Huang, Yang Wang, Jue Wang, Honghui Shi, and Thomas Huang. When awgn-based denoiser meets real noises. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019.

[56] Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. Ap-bsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[57] Mengping Yang, Zhe Wang, Ziqiu Chi, and Wenyi Feng. Wavegan: Frequency-aware gan for high-fidelity few-shot image generation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 1–17, 2022.

[58] Jie Liang, Hui Zeng, and Lei Zhang. High-resolution photorealistic image translation in real-time: A laplacian pyramid translation network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9392–9400, 2021.

[59] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 97–104, 2011.

[60] Jiang Hai, Zhu Xuan, Ren Yang, Yutong Hao, Fengzhu Zou, Fang Lin, and Songchen Han. R2rnet: Low-light image enhancement via real-low to real-normal network. *arXiv preprint arXiv:2106.14501*, 2021.

[61] Lin Zhao, Shao-Ping Lu, Tao Chen, Zhenglu Yang, and Ariel Shamir. Deep symmetric network for underexposed image enhancement with recurrent attentional learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12075–12084, 2021.

[62] Jie Huang, Yajing Liu, Xueyang Fu, Man Zhou, Yang Wang, Feng Zhao, and Zhiwei Xiong. Exposure normalization and compensation for multiple-exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6043–6052, 2022.

[63] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, 27(4):2049–2062, 2018.

[64] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 17–33. Springer, 2022.

[65] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2019.

[66] Kangfu Mei, Aiwen Jiang, Juncheng Li, and Mingwen Wang. Progressive feature fusion network for realistic image dehazing. In *Asian Conference on Computer Vision (ACCV)*, pages 203–215. Springer, 2019.

[67] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[68] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

[69] Amber L Simpson, Michela Antonelli, Spyridon Bakas, Michel Bilello, Keyvan Farahani, Bram Van Ginneken, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv preprint arXiv:1902.09063*, 2019.

[70] Shuang Xu, Jiangshe Zhang, Zixiang Zhao, Kai Sun, Junmin Liu, and Chunxia Zhang. Deep gradient projection networks for pan-sharpening. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1366–1375, June 2021.

[71] Junfeng Yang, Xueyang Fu, Yuwen Hu, Yue Huang, Xinghao Ding, and John Paisley. Pannet: A deep network architecture for pan-sharpening. In *IEEE International Conference on Computer Vision*, pages 5449–5457, 2017.

[72] Man Zhou, Keyu Yan, Jie Huang, Zihe Yang, Xueyang Fu, and Feng Zhao. Mutual information-driven pan-sharpening. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1798–1808, June 2022.

[73] Man Zhou, Jie Huang, Keyu Yan, Hu Yu, Xueyang Fu, Aiping Liu, Xian Wei, and Feng Zhao. Spatial-frequency domain information integration for pan-sharpening. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XVIII*, pages 274–291. Springer, 2022.

[74] Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17452–17462, 2022.

[75] Wenxin Wang, Boyun Li, Yuanbiao Gou, Peng Hu, and Xi Peng. Relationship quantification of image degradations. *arXiv preprint arXiv:2212.04148*, 2022.

[76] Jinyuan Liu, Xin Fan, Zhanbo Huang, Guanyao Wu, Risheng Liu, Wei Zhong, and Zhongxuan Luo. Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5802–5811, 2022.