

Lighting Every Darkness in Two Pairs:

A Calibration-Free Pipeline for RAW Denoising

靳鑫^{1*} 肖嘉文^{1*} 韩凌昊¹ 郭春乐^{1†}

Ruixun Zhang² 刘夏雷¹ 李重仪^{1,3}

¹VCIP, CS, Nankai University ²Peking University ³S-Lab, Nanyang Technological University

{xjin, xiaojw, lhhan}@mail.nankai.edu.cn, zhangruixun@pku.edu.cn,

{guochunle, xialei, lichongyi}@nankai.edu.cn

<https://srmeo.github.io/projects/led-iccv23>

Abstract

基于标定的方法在极低光照环境下的 RAW 图像去噪中占主导地位。然而，这些方法存在几个主要缺陷：1) 标定过程费力且耗时，2) 不同相机的降噪网络难以迁移，3) 合成噪声和真实噪声之间的差异被高倍数字增益放大。为了克服上述缺点，我们提出了一种无需标定的流水线——**Lighting Every Darkness (LED)**，无论数字增益或相机传感器如何变化。我们的方法无需标定噪声参数并进行重复训练，仅需少量配对数据和微调，即可适配目标相机。此外，在两个阶段中精心设计的结构变化有效缩小了合成噪声与真实噪声之间的域间差距，而无需任何额外的计算成本。仅使用每个附加数字增益 2 对数据（总共 6 对）以及 0.5% 的迭代次数，我们的方法便能超越其他基于标定的方法，取得卓越的性能。

1. 引言

噪声是图像捕捉中一个不可避免的话题，近年来已被系统地研究 [5, 62, 49, 39, 2, 8, 53]。与标准 RGB 图像相比，RAW 图像在去噪方面有两个巨大优势：可操作的原始噪声分布 [53] 以及更高的位深度，有助于区分信号和噪声。基于学习的方法在使用配对的真实数据集进行 RAW 图像去噪方面取得了显著进展 [63, 21, 60, 32]。然而，对于每个单独的相机型号，收集大规模的真实 RAW 图像数据集是不可行的。因此，越来越多的研究开始关注在合成数据集上部署基于学习的方法 [1, 57, 31, 53, 64, 42, 38]。

基于标定的噪声合成与基于物理的模型已被证明在拟合真实噪声方面非常有效 [51, 53, 64, 43, 65, 17]。通常，这些方法包括以下几个步骤：首先，根据电子成像流程建立一个精心设计的噪声模型。然后，选择一个特定的目标相机，并仔细标定预定义噪声模型的参数。最后，生成用于训练去噪网络的合成配对数据。此外，一些方法还借助基于深度神经网络 (DNN) 的

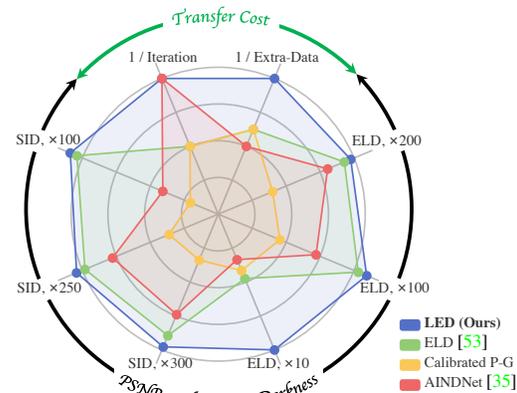


Figure 1. LED 在各种光照条件下（不同数字增益和相机传感器）相比于基于标定或迁移学习的方法，取得了最先进的性能。此外，所提流程应用于新相机模型时，所需的成本最小。详细信息请参见 Sec. 4.

生成模型进行噪声参数的标定 [43, 65]。

尽管这些方法已取得了出色的性能，但它们仍然受到以下三个主要缺点的限制，如 Fig. 2 (a) 所示：1) 标定专用的数据收集需要稳定的光照环境和复杂的后处理，导致过程繁琐且耗时；2) 针对特定相机训练的去噪网络难以迁移到其他相机上，这导致网络与相机之间的强依赖关系，需要为不同的目标相机进行重复的标定和训练；3) 某些噪声分布可能未包含在噪声模型中，称为超出模型噪声 [53, 64, 17]。换句话说，合成噪声 (SN) 和真实噪声 (RN) 之间的域间差距仍然存在。尽管最近的研究 [65] 主要集中于通过基于 DNN 的标定来减轻标定成本，但耦合问题和超出模型的噪声仍然增加了训练开销，并限制了其性能。

为了解决基于标定方法的上述三个问题，我们提出了一种无标定的流程——**Lighting Every Darkness (LED)**。如 Fig. 2 (b) 所示，我们的框架无需任何标定数据或操作。此外，为了消除去噪网络与特定目标相机之间的强依赖关系，我们提出了一个预训练和微调框架。对于虚拟¹与目标相机之间的差距，以及超出

*同等贡献。

†郭春乐是通讯作者。

¹“虚拟”相机不对任何真实的相机型号，但具有预定义噪声模型的合理噪声参数。

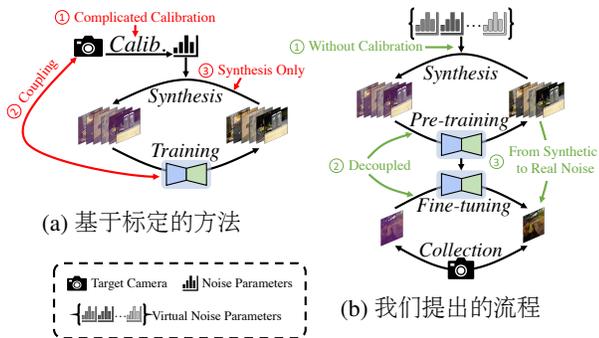


Figure 2. 基于标定的方法和我们提出的 LED 流程的缩略图。“ \rightarrow ”表示基于标定方法的问题，“ \dashrightarrow ”突出显示我们为了解决上述问题提供的方案。标定代表标定操作，包括预定义噪声模型、收集专用标定数据、后处理和计算噪声参数。在 LED 中，收集过程仅捕获少量配对数据，从而减轻部署成本。

模型噪声的影响，我们提出了一个重参数化噪声去除 (RepNR) 模块。在预训练过程中，RepNR 模块配备了多个相机特定对齐 (CSA) 模块。每个 CSA 负责学习虚拟相机的相机特定信息，并将特征对齐到共享空间。然后，去噪卷积学习所有模型内（假定为噪声模型一部分的组件）噪声的共同知识。在微调过程中，我们将所有虚拟相机的 CSA 平均作为目标相机的初始化。此外，添加了一个并行卷积分支用于去除超出模型噪声 (OMNR)。每个数字增益比率仅使用目标相机拍摄的 2 对图像（共 6 对原始图像），用于学习去除其真实噪声（关于为何每个比率使用 2 对图像的讨论请见 Sec. 5）。在部署过程中，所有 RepNR 模块可以结构化地重参数化 [15, 16, 10] 为一个简单的 3×3 卷积，无需额外的计算成本，从而得到一个简洁的 UNet [47]。

我们的主要贡献总结如下：

- 我们提出了一种无标定的流程——Lighting Every Darkness，避免了标定噪声参数的所有额外成本。
- 设计的 CSA 解除去噪网络与相机型号之间的耦合，而 OMNR 通过学习不同传感器的超出模型噪声实现少量样本迁移。
- 与最先进的方法相比，仅需每个比率 2 对原始图像和 0.5% 的迭代次数。

2. 相关工作

使用配对真实数据进行训练。自从 SIDD [2] 的开创性工作以来，RAW 数据在图像去噪中的潜力已经被广泛探索。最近的工作将重点从普通光照下的图像去噪转向了极低光照环境，例如 SID [8] 和 ELD [53]。尽管基于真实噪声的方法取得了有前景的结果 [9, 11, 58, 59]，但收集大规模的配对（低质量和高质量配对）真实数据集仍然是其应用的瓶颈。像 Noise2Noise [39] 和 Noise2NoiseFlow [42] 等方法通过使用配对的低质量

RAW 图像，可以避免费时费力地收集带噪声与清晰图像的配对数据。然而，这些方法在极度噪声的场景中，如极其黑暗的场景中，往往失败 [8, 53]。我们的 LED 旨在通过在极低光照环境下使用少量配对图像来补充真实噪声去除的知识，从而缓解数据收集中的困难。

基于标定的去噪。基于合成噪声的方法可以避免收集配对数据集的繁琐工作，但仍然存在实际约束。广泛使用的噪声模型，泊松噪声和高斯噪声，特别是在极低光照环境中 [8, 53]，与真实噪声分布差距较大²。因此，基于标定的方法，通过模拟电子成像流程中的每个噪声成分 [4, 23, 20, 29, 37]，因其可靠性而蓬勃发展。ELD [53] 提出了一个能较好地拟合真实噪声的噪声模型，在黑暗场景下表现出色。Zhang 等人 [64] 认识到信号无关噪声的来源过于复杂，无法建模，进而提出了一种从暗帧中随机采样信号无关噪声的方法。然而，它仍然需要对信号相关噪声的参数进行标定，例如整个系统增益。Kristian 等人 [43] 将基于物理的噪声模型与生成对抗框架 [19] 相结合，构建了噪声生成器。Zou 等人 [65] 通过使用对比学习 [12, 22] 进行参数估计，旨在实现更准确简洁的标定。尽管基于标定的方法取得了出色的性能，但仍然需要稳定的光照环境（例如亮度和温度）、专用的标定数据收集（例如每个相机设置几十张图像）和复杂的后处理（例如对齐、定位和统计）来估计噪声参数。此外，由于参数的多样性和预定义噪声模型的不均匀性，每个相机都需要重复进行标定和训练 [50, 20, 37, 41]。同时，合成噪声和真实噪声之间的领域差距也没有得到有效解决。我们的 LED 通过无标定流程、预训练和微调框架以及提出的 RepNR 模块解决了上述问题。

从合成噪声到真实噪声。当在合成数据上训练而在真实数据上测试时，真实噪声与合成噪声之间的领域差距是一个不可避免的挑战。随着 AdaIN [27, 34] 和少样本学习 [24, 56, 26] 的进展，最近的研究主要集中在利用迁移学习 [35] 或领域自适应 [45] 技术来缓解领域差距。然而，在极暗的场景中，由于极端噪声和额外数字增益引起的数值不稳定性，这些方法会在信号重建上失败。我们提出的相机特定对齐 (CSA) 能够避免数值不稳定性，同时仍能解耦相机特定信息和噪声模型的共同知识。此外，与实例或层归一化 [48, 3] 相比，这些对齐操作可以重新参数化为卷积操作，如定制批量归一化 [28]，因此不会增加额外的计算成本。

3. 方法

在本节中，我们首先介绍了我们提出的无标定 RAW 图像去噪的整体流程。然后，我们将介绍我们的重参数化噪声去除 (RepNR) 模块。整体去噪流程如图 3 所示。

²在极低光照场景下进行去噪需要对输入图像应用额外的数字增益（最多 300 \times ），这加剧了真实噪声与合成噪声之间的领域差距。

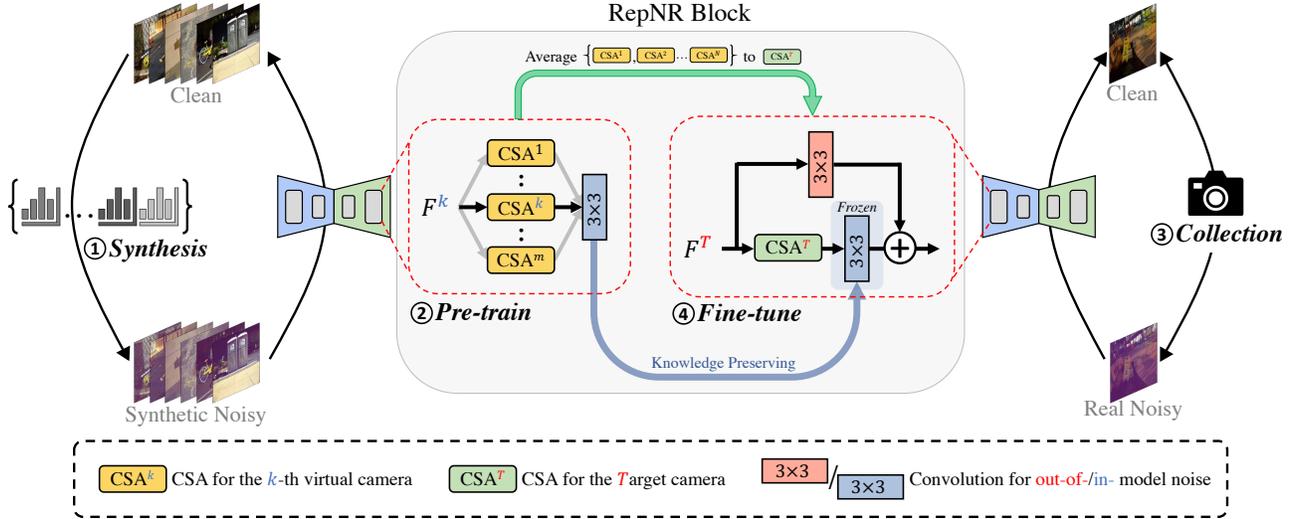


Figure 3. 我们提出的 LED 及 RepNR 模块的示意图。整体流程分为四个部分：1) 采样一组 N 个虚拟相机，负责后续合成噪声；2) 使用 N 个 CSA 和合成配对图像对去噪网络进行预训练，每个 CSA 对应一个虚拟相机；3) 使用目标相机收集少量真实噪声图像对；4) 使用真实噪声数据对预训练的去噪网络进行微调，专门化网络以适应目标相机。中间部分展示了我们为 RepNR 模块的不同训练阶段提供的不同优化策略。

3.1. 预备知识与动机

在 RAW 图像空间中，捕获的信号 D 总是被视为清晰图像 I 和噪声成分 N 的和，公式表示为式 (1)：

$$D = I + N, \quad (1)$$

其中 N 被假设为一个噪声模型，

$$N = N_{shot} + N_{read} + N_{row} + N_{quant} + \epsilon, \quad (2)$$

其中 N_{shot} 、 N_{read} 、 N_{row} 和 N_{quant} 分别表示快门噪声、读取噪声、行噪声和量化噪声。 ϵ 表示模型外的噪声。除了模型外的噪声，其他噪声是从特定的分布中采样的：

$$\begin{aligned} N_{shot} + I &\sim \mathcal{P}\left(\frac{I}{K}\right)K, \\ N_{read} &\sim TL(\lambda; \mu_c, \sigma_{TL}), \\ N_{row} &\sim \mathcal{N}(0, \sigma_r), \\ N_{quant} &\sim U\left(-\frac{1}{2}, \frac{1}{2}\right), \end{aligned} \quad (3)$$

其中 K 表示整体系统增益。 \mathcal{P} 、 \mathcal{N} 和 U 分别表示泊松分布、高斯分布和均匀分布。 $TL(\lambda; \mu, \sigma)$ 表示 Tukey-lambda 分布 [33]，其形状参数为 λ ，均值为 μ ，标准差为 σ 。此外， (K, σ_{TL}) 和 (K, σ_r) 的联合分布存在一个线性关系，可以表示为：

$$\begin{aligned} \log(K) &\sim U(\log(\hat{K}_{min}), \log(\hat{K}_{max})), \\ \log(\sigma_{TL}) | \log(K) &\sim \mathcal{N}(a_{TL} \log(K) + b_{TL}, \hat{\sigma}_{TL}), \\ \log(\sigma_r) | \log(K) &\sim \mathcal{N}(a_r \log(K) + b_r, \hat{\sigma}_r), \end{aligned} \quad (4)$$

在这种情况下，一台相机可以近似表示为一个十维坐标 \mathcal{C} ：

$$\mathcal{C} = (\hat{K}_{min}, \hat{K}_{max}, \lambda, \mu_c, a_{TL}, b_{TL}, \hat{\sigma}_{TL}, a_r, b_r, \hat{\sigma}_r). \quad (5)$$

以前的方法专注于通过标定来调整坐标 \mathcal{C} ，然而这过程劳动密集且存在巨大的领域差距（即模拟噪声与真实噪声之间的差距）。此外，由于神经网络与相机之间的强耦合，仍然需要重复训练过程。我们的目标是摒弃复杂的标定过程，并减弱网络与相机之间的强耦合。此外，我们充分考虑了模型外的噪声，且通过我们 RepNR 模块的结构化修改可以缓解这一问题。总的来说，我们的动机是使网络成为一个快速适配器 [46, 18]。

3.2. 通过相机特定对齐进行预训练

预处理。为了促进网络成为一个快速适配器，我们首先利用虚拟相机进行网络的预训练。给定虚拟相机的数量 m 和参数空间（表示为 \mathcal{S} ），对于第 k 个相机，我们选择每个参数范围内的第 k 个 m 二分点，并将它们结合起来得到一个虚拟相机。通过合成噪声增强的数据，我们可以基于多个虚拟相机对网络进行预训练，使网络学习通用知识。

相机特定对齐。如 Fig. 3 所示，在预训练过程中，我们引入了相机特定对齐 (CSA) 模块，该模块专注于调整输入特征分布。在基准模型中，一个 3×3 的卷积层后接 leaky-ReLU [55] 是主要的组成部分。为了将来自不同虚拟相机的特征映射到共享空间中，在每个卷积层之前插入了一个多路径对齐层。每个路径对应于第 k 个相机的 CSA，将第 k 个相机特定的特征分布对齐到共享空间中。设第 k 个虚拟相机的特征为 $F = (f_1, \dots, f_c) \in \mathcal{R}^{B \times C \times H \times W}$ 。形式上，第 k 个分

支包含一个权重 $W^k = (w_1^k, \dots, w_c^k) \in \mathcal{R}^C$ 和一个偏置 $b^k = (b_1^k, \dots, b_c^k) \in \mathcal{R}^C$, 对 F 进行通道级别的线性投影, 表示为 $Y = W^k F + b^k$. $W^k (k = 1, \dots, m)$ 初始化为 $\mathbf{1}$, $b^k (k = 1, \dots, m)$ 初始化为 $\mathbf{0}$, 在开始时对 3×3 的卷积没有影响。在训练过程中, 通过第 k 个虚拟相机的噪声增强数据将被输入到第 k 个路径进行对齐, 并输入共享的 3×3 卷积进行进一步处理。详细的预训练流程如 Algorithm 1 所示。

3.3. 通过少量 RAW 图像对进行微调

在预训练之后, 模型已准备好用于实际的去噪任务。我们提出采用少量样本策略, 特别是仅使用 6 对原始图像 (每个比率 2 对, 共 3 种比率) 来对预训练模型进行微调。假设 3×3 卷积已经学会了足够的特征来处理由 CSA 对齐后的特征。为了更好地利用预训练获得的模型参数, 卷积层在进一步微调时保持冻结。为了解决真实噪声问题, 我们用一个新的 CSA 层替换了多分支 CSA, 这个新层被称为 CSA^T (目标相机的 CSA)。与预训练中的多分支 CSA 不同, CSA^T 层是通过预训练的 CSA 进行平均初始化的, 以实现更好的泛化。上述的 $CSA^T-3 \times 3$ 分支称为模型内噪声去除分支 (IMNR)。

然而, 真实噪声不仅包含建模部分, 还包含一些模型外的噪声。由于我们的 CSA 层仅设计用于对合成噪声增强的特征进行对齐, 因此 IMNR 能够处理的噪声和真实噪声之间仍然存在差距 (即公式 Eqn. (2) 中的 ϵ)。因此, 我们提出添加一个新的分支, 称为模型外噪声去除分支 (OMNR), 以学习真实噪声和建模噪声之间的差距。以前的工作已经展示了并行卷积分支在迁移学习和持续学习中的潜力 [61]。OMNR 仅包含一个 3×3 卷积, 旨在从少量的原始图像对中抓取真实噪声的结构先验。考虑到我们对噪声剩余部分 ϵ 没有先验知识, 我们将 OMNR 的权重和偏置初始化为全零的张量。将 IMNR 和 OMNR 结合得到我们提出的 RepNR 块。需要注意的是, 首先学习模型内噪声, 然后学习模型外噪声是更合理的。因此, 我们将优化过程分为两个步骤: 首先训练 IMNR, 然后训练 OMNR。遵循这一过程, 二步微调的迭代只占预训练的 0.5%, 在实践中很容易实现。详细的微调流程如 Algorithm 2 所示。

关于 CSA^T 初始化的分析。如 Sec. 3.3 中所述, 我们

Algorithm 1 LED的预训练流程

Require: 模型 Φ, m, \mathcal{S} , 干净的数据集 D

$\Phi_{\text{pre}} \leftarrow \text{insert-multi-CSA}(\Phi)$

$\{c_k\}_{k=1}^m \leftarrow \text{generate-virtual-camera}(\mathcal{S})$

while 未收敛 **do**

 从数据集 D 中随机采样小批次 x_i

$k \leftarrow \text{random}(1, m)$

$\tilde{x}_i \leftarrow \text{augment}(c_k, x_i)$

$\text{train}(\Phi_{\text{pre}}, \{\tilde{x}_i, x_i\})$

end while

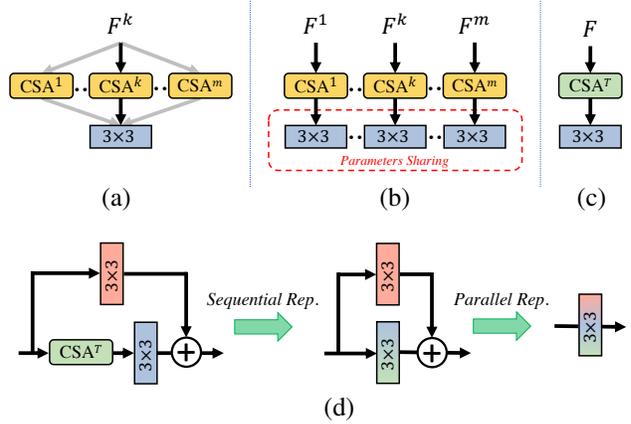


Figure 4. CSA^T 初始化策略和重参数化过程的说明。(a) 预训练过程中的 RepNR 块。(b) 我们的 RepNR 块可以看作是 m 个共享的参数块, 每个用于特定的虚拟相机。(c) 我们通过平均预训练的 CSA 来初始化 CSA^T , 这可以视为模型集成。(d) 部署过程中的重参数化。Rep. 表示重参数化。

通过对多分支 CSA 层中的预训练 CSA 进行平均来初始化 CSA^T 。由于每个卷积是由多分支 CSA 中的每条路径共享的, 因此该初始化可以视为 m 个模型的集成, 其中 m 是路径的数量。正如 [7, 30, 54] 中所述, 不同模型的权重平均可以显著提高模型的泛化能力。这符合我们将模型泛化到目标噪声域的动机。

另一个原因是, CSA 几乎是由坐标 \mathcal{C} 决定的。从这个角度来看, 多个 CSA 的平均值可以看作是这些坐标的重心。同时, 测试相机的坐标, 无论是在 SID [8] 还是 ELD [53] 中, 都包含在参数空间 \mathcal{S} 中。在这种情况下, 平均预训练的 CSA 似乎是一个很好的起点。

3.4. 部署

当微调完成后, 模型的部署对未来的应用具有重要意义。直接将 3×3 卷积替换为我们的 RepNR 块, 必然会导致参数数量和计算量的增加。然而, 值得注意的是, 我们的 RepNR 块仅由串行与并行线性映射组成。此外, RepNR 块中每个分支的感受野为 3。因此, 利用结构重参数化技术 [14, 15, 16], 我们的 RepNR 块在部署时可以转化为一个普通的 3×3 卷积, 如 Fig. 4 (d) 所示。这意味着我们的模型在应用过程中不会产生额外的成本, 同时也能够与其他方法进行公平比较。

4. 实验与分析

在本节中, 我们详细介绍我们的实现方法, 说明数据集和评估指标, 进行对比实验, 并展示消融实验。

4.1. 实现细节

与大多数去噪方法 [57, 13] 类似, 我们使用 $L1$ 损失函数作为训练目标。为了公平比较, 我们使用与先前方法相同的 UNet [47] 架构, 唯一的不同是我们将 UNet 中的卷积块替换为我们提出的 RepNR 块。如 Sec. 3.4 所

Table 1. 在SID [8] Sony子集上的定量结果。最好的结果用**粗体**表示，第二好的用下划线表示。当迁移到新的目标相机时，额外的数据需求和迭代次数 (K) 已计算。基于DNN模型的方法需要为目标相机训练噪声生成器，因此导致更大的迭代要求。AINDNet*表示AINDNet是使用我们提出的噪声模型而不是AWGN进行预训练的。值得注意的是，除了AINDNet之外，所有方法都使用相同的UNet架构进行训练，而我们保持AINDNet与他们的论文相同，参数几乎是UNet的两倍。

类别	方法	额外数据需求	迭代次数 (K)	×100		×250		×300	
				PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
基于DNN模型	Kristina <i>et al.</i> [43]	~1800 噪声-清晰图像对	327.6	38.7799	0.9120	34.4924	0.7900	31.2971	0.6990
	NoiseFlow [1]	~1800 噪声-清晰图像对	777.6	37.0200	0.8820	32.9457	0.7699	29.8068	0.6700
基于标定	Calibrated P-G	~300 标定数据	257.6	39.1576	0.8963	33.8929	0.7630	31.0035	0.6522
	ELD [53]	~300 标定数据	257.6	<u>41.8271</u>	<u>0.9538</u>	38.8492	0.9278	35.9402	0.8982
	Zhang <i>et al.</i> [64]	~150/~150 for calib./database	257.6	40.9232	0.9488	38.4397	0.9255	35.5439	0.8975
基于真实数据	SID [8]	~1800 噪声-清晰图像对	257.6	41.7273	0.9531	<u>39.1353</u>	<u>0.9304</u>	37.3627	0.9341
	Noise2Noise [39]	~12000 noisy pairs	257.6	39.2769	0.8993	34.1660	0.7824	31.0991	0.7080
	AINDNet [35]	~300 噪声-清晰图像对	1.5	40.5636	0.9194	36.2538	0.8509	32.2291	0.7397
	AINDNet*	~300 噪声-清晰图像对	1.5	39.8052	0.9350	37.2210	0.9101	34.5615	0.8856
	LED (我们的)	6 噪声-清晰图像对	1.5	41.9842	0.9539	39.3419	0.9317	<u>36.6728</u>	<u>0.9147</u>

Table 2. 在ELD [53]数据集的两个相机模型，SonyA7S2和NikonD850上的定量结果。最好的结果用**粗体**表示。

相机	比率	Calibrated P-G PSNR/SSIM	ELD [53] PSNR/SSIM	LED (我们的) PSNR/SSIM	相机	比率	Calibrated P-G PSNR/SSIM	ELD PSNR/SSIM	LED (我们的) PSNR/SSIM
SonyA7S2	×1	54.3710/0.9977	52.8120/0.9957	51.9547/0.9968	NikonD850	×1	50.6207/ 0.9949	50.5628/0.9925	50.6222/0.9939
	×10	49.9973/0.9891	50.0152/0.9913	50.1762/0.9945		×10	48.3461/0.9884	48.3667/0.9890	48.0684/ 0.9894
	×100	41.5246/0.8668	44.9865/0.9707	45.3574/0.9779		×100	42.2231/0.9046	43.6907/0.9634	43.5620/ 0.9667
	×200	37.6866/0.7818	42.5440/0.9430	42.9747/0.9577		×200	39.0084/0.8391	41.3311/0.9364	41.3984/0.9482

Algorithm 2 LED的微调与部署流程

Require: 预训练模型 Φ_{pre} , 真实数据集 D_{real}

$\Phi_{ft} \leftarrow \text{freeze} 3 \times 3$ 卷积(Φ_{pre})

$\Phi_{ft} \leftarrow \text{average-CSA}(\Phi_{ft})$

while 未收敛 **do**

从 D_{real} 中采样小批次图像对 $\{x_i, y_i\}$

训练($\Phi_{ft}, \{x_i, y_i\}$)

end while

$\Phi_{ft} \leftarrow \text{freeze}(\Phi_{ft})$

$\Phi_{ft} \leftarrow \text{add-OMNR}(\Phi_{ft})$

while 未收敛 **do**

从 D_{real} 中采样小批量图像对 $\{x_i, y_i\}$

训练($\Phi_{ft}, \{x_i, y_i\}$)

end while

$\Phi_{final} \leftarrow \text{部署}(\Phi_{ft})$

述，RepNR块可以通过结构重参数化转化为简单的卷积块，而不产生额外的计算开销。在预训练过程中，使用与ELD [53]相同的数据预处理和优化策略。我们使用SID [8]训练子集中的长曝光原始图像进行噪声合成。关于数据预处理，我们将Bayer图像打包为4个通道，然后裁剪长曝光数据，补丁大小为512 × 512，无重叠，将每个epoch的迭代次数从161增加到1288。我们的实现基于PyTorch [44]和MindSpore。

我们使用200个epoch (257.6K次迭代) 和Adam优化器 [36] ($\beta_1 = 0.9, \beta_2 = 0.999$) 进行训练，优化过程中不应用权重衰减。初始学习率设置为 10^{-4} ，然后在第100个epoch (128.8K次迭代) 时减半，最终在第180个epoch (231.84K次迭代) 时降至 10^{-5} 。

在微调过程中，我们首先冻结 3×3 卷积并将多分支CSA平均作为 CSA^T 的初始化。在使用 10^{-4} 的学习率训练 CSA^T 1K次迭代后，我们添加了OMNR分支（一个并行的 3×3 卷积），并冻结网络中剩余的所有参数。最后，我们用 10^{-5} 的学习率训练OMNR分支500次迭代。整个训练过程完成后，我们通过将RepNR块重参数化为卷积来部署我们的模型。

4.2. 数据集与评估指标

我们在两个RAW去噪数据集上对我们提出的LED进行了基准测试，分别是SID [8]和ELD [53]。这两个数据集中包含了四种不同的相机型号：Sony A7S2、Nikon D850、Canon EOS70D、Canon EOS700D，以及7种不同的额外数字增益，范围从×1到×300。对于SID数据集，我们从每种额外数字增益 (×100、×250、和×300) 中随机选择两对数据作为少量样本训练数据集。对于ELD数据集，前两个场景的配对RAW图像用于微调预训练网络。整个训练过程完成后，SID [8] Sony子集的测试集和ELD [53]数据集的剩余场景被用来验证我们提出的LED的有

效性。我们还在Canon相机（Canon EOS70D和Canon EOS700D）上对LED进行了评估，并取得了最先进的性能。相关结果将在更新版本中发布。

我们将PSNR和SSIM [52]作为定量评估指标，用于像素级和结构级的评价。需要注意的是，低光RAW图像的像素值通常位于比sRGB图像更小的范围内，即归一化后为[0, 0.5]，这会导致较低的均方误差和较高的PSNR。

4.3. 与最先进方法的比较

我们在两个数据集上评估了我们的LED方法，分别是SID [8]的Sony子集和ELD数据集 [53]，以评估LED在室外和室内场景中的泛化能力。我们将最先进的RAW去噪方法与LED进行比较，这些方法在极低光环境下表现优越，包括：

- 基于DNN模型的方法：Kristina *et al.* [43] 和 Noise-Flow [1]。这些方法首先在配对的真实RAW图像上进行训练，以学习如何为特定相机生成噪声，因此在此部署到新相机模型时需要更多的迭代次数。
- 基于校准的方法：ELD [53]、Zhang *et al.* [64] 和 Calibrated P-G。这些方法需要一个耗时且繁琐的标定过程。
- 基于真实数据的方法：使用噪声-清晰图像对（SID [8]）、噪声-噪声图像对（Noise2Noise [39]）和迁移学习（AINDNet [35]）进行训练。

为了公平比较，上述所有方法的去噪网络都使用与ELD [53]相同的设置进行训练，如Sec. 4.1中所述。

定量评估。如Tab. 1和Tab. 2所示，我们的方法在极低光环境下优于之前的基于标定的方法。合成噪声和真实噪声之间的领域差距会随着增益比例的增大（ $\times 250$ 和 $\times 300$ ）而加剧，导致在使用合成噪声进行训练时性能下降，正如ELD [53]与SID [8]之间的比较所示。此外，基于DNN模型的方法通常会比基于校准的方法更多的差异。特别是，Kristina *et al.* [43]没有考虑不同的系统增益。然而，我们的方法通过使用少量真实数据进行微调，缓解了这一差异，因此在 $\times 100$ 和 $\times 250$ 数字增益下实现了更好的性能，如Tab. 1所示。AINDNet [35]在极暗的场景下也能通过噪声模型的较小偏差获得更好的性能。噪声模型的偏差不会影响在小额外数字增益下的去噪能力，如Tab. 2所示。然而，我们的方法在极低光场景中表现出优越性，并且能够在不同的相机模型中展现强大的鲁棒性。值得注意的是，LED与其他方法相比，减少了训练成本，包括数据需求和训练迭代次数。

定性评估。Fig. 5和Fig. 6分别展示了我们的方法与其他最先进方法在SID [8]和ELD [53]数据集上的比较。在极低光条件下成像时，强烈的噪声会严重干扰色调。如Fig. 5所示，输入图像呈现绿色或紫色色偏，大多数对比方法无法恢复正确的色调。得益于隐式噪声

Table 3. 关于RepNR块的消融研究。提供的指标是在微调策略下计算的，如Fig. 3中的③所示。

配置			$\times 100$	$\times 250$	$\times 300$
U-net	CSA	OMNR	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
✓			41.518/0.951	39.140/0.923	36.273/0.898
✓	✓		41.866/0.954	39.201/0.931	36.499/0.912
✓	✓	✓	41.984/0.954	39.342/0.932	36.673/0.915

Table 4. 关于预训练策略的消融研究。带有*的模型表示采用与PMN [17]相同的训练策略进行去噪器训练，而LED*则利用该策略进行预训练。

方法	$\times 100$	$\times 250$	$\times 300$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
LED	41.984/0.954	39.342/0.932	36.673/0.915
ELD* [53]	42.081/0.955	39.461/0.934	36.870/0.920
LLD* [6]	42.100/0.955	39.760/0.933	36.760/0.912
LED*	42.396/0.955	39.843/0.939	36.997/0.923

建模和多样的采样空间，LED能够高效地恢复受严重噪声干扰的信号，呈现准确的色彩还原和丰富的纹理细节。此外，对比方法难以识别模型外的噪声，这些噪声会以固定模式或特定位置破坏结果图像。而在微调阶段，LED额外学习去除这些相机特有的噪声，从而实现了更好的视觉质量和更强的鲁棒性。

4.4. 消融研究

重参数化噪声去除模块（RepNR）。我们对重参数化噪声去除（RepNR）模块中的不同组件进行了消融实验。如Tab. 3所示，我们的RepNR在三个不同的增益比下表现更好，且RepNR块中的每个组件对整体流程都有积极贡献。

使用先进策略的预训练。如Tab. 4所示，使用SGDR [40]优化器和更大批次大小（与PMN [17]相同）进行预训练，能够在相同的微调成本下进一步提高性能（每个增益比使用2对图像和1.5K迭代），验证了所提LED方法的可扩展性。此外，与LLD [6]（CVPR23中的同期工作）相比，LED在数据成本和时间成本上表现出更好的性能。关于时间成本，ELD* [53]在我们的实现中训练时间大约为一天，而LED微调只需要不到4分钟（快了367倍）。

目标相机CSA的初始化。由于我们按照Sec. 3.3中的方法初始化了CST^T，因此我们展示了(1, 0)初始化和模型平均之间的PSNR/SSIM差异。从结果可以观察到，模型平均在大多数场景下获得了更好的性能。此外，SID [8]中Sony A7S2的表现最能代表泛化能力，因为该数据集的规模较大。

使用更多图像进行微调。我们展示了关于微调图像数量的消融研究，以展示我们提出的LED的前景。如Fig. 7所示，随着配对数据量的增加，性能会逐渐提高。此外，LED在使用2对噪声-清晰配对图像进行微调时，优于ELD [53]。我们将在Sec. 5中提供更多讨论

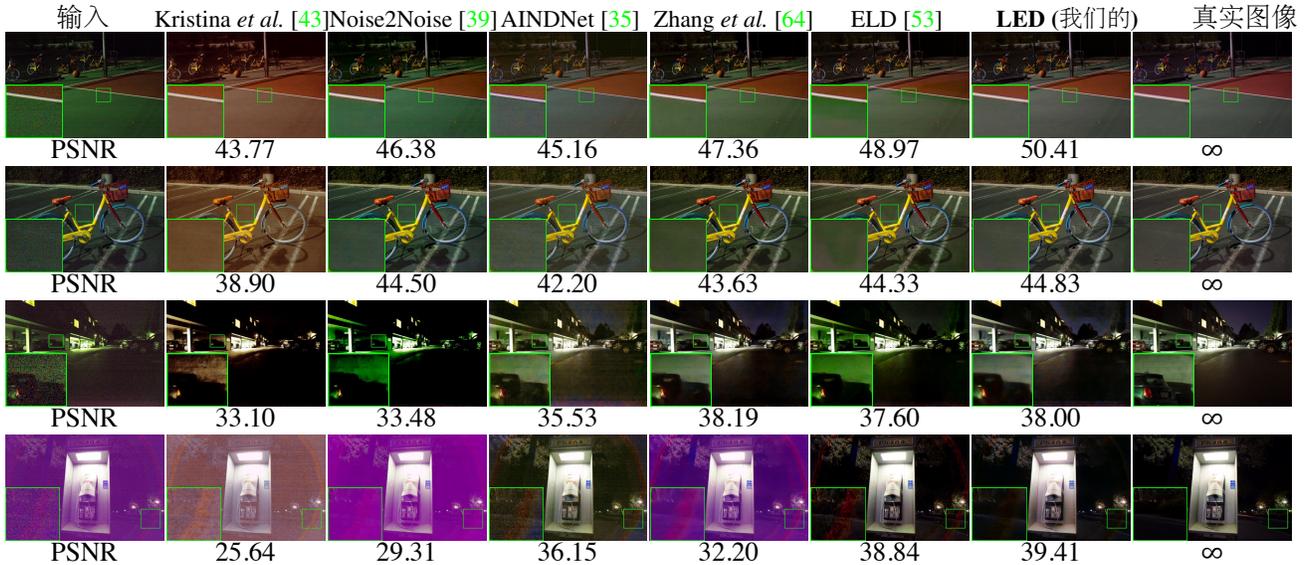


Figure 5. 在SID [8]数据集上, 我们的LED与其他最先进方法的视觉比较(请放大以获得最佳查看效果)。我们对输入图像进行了放大并使用与ELD [53]相同的ISP进行后处理。



Figure 6. 在ELD [53]数据集上的视觉比较。

Table 5. 关于目标相机CSA初始化策略的消融研究。Sony A7S2#表示在SID [8]数据集上进行微调 and 测试, 其他则基于ELD [53]数据集。

初始化	指标	Sony		Nikon		Canon	
		A7S2#	A7S2	D850	EOS700D	EOS70D	
(1, 0)	PSNR	39.015	47.310	45.790	41.409	42.344	
	SSIM	0.9307	0.9809	0.9737	0.9408	0.9520	
平均值	PSNR	39.161	47.616	45.903	41.516	42.495	
	SSIM	0.9322	0.9817	0.9743	0.9412	0.9524	

Table 6. 关于在合成数据集上进行微调和测试的配对数的消融研究。 N 表示使用 N 对具有相似整体系统增益的图像进行微调。 N^* 表示使用具有轻微不同整体系统增益的图像对。

比率	1	2	4	2*
$\times 100$	41.295/0.9480	41.704/0.9523	41.432/0.9466	43.795/0.9648
$\times 250$	39.239/0.9350	39.410/0.9351	39.327/0.9367	41.311/0.9457
$\times 300$	38.314/0.9229	38.486/0.9216	38.499/0.9240	39.190/0.9278

论。

5. 讨论

为什么每个比率使用2对数据? 如 Eqn. (4)所示, 噪声

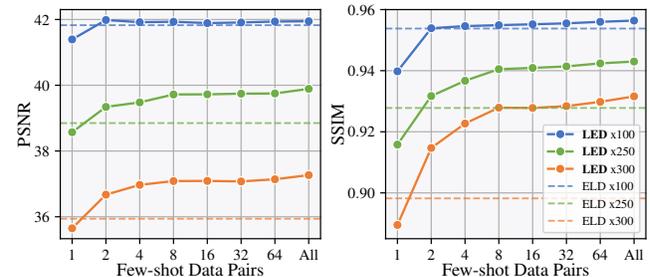


Figure 7. 关于微调数据量的消融实验。LED 在每个增益比仅使用2对图像时就能实现更好的性能。

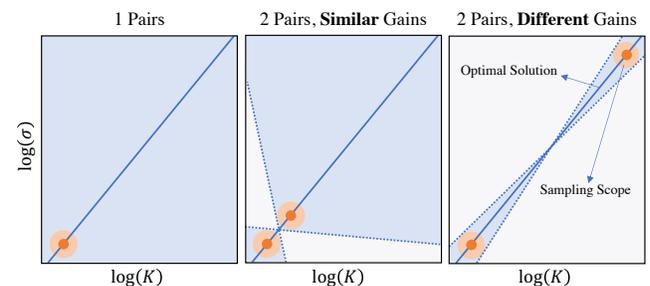


Figure 8. 线性关系下整体系统增益 $\log(K)$ 和噪声方差 $\log(\sigma)$ 的可行解空间的示意图(蓝色区域), 展示了不同采样策略的效果。

方差 $\log(\sigma)$ 与整体系统增益 $\log(K)$ 之间存在线性关系。只有一对数据时, 无法正确找到线性关系, 导致最差的性能, 如 Tab. 6所示。此外, 使用两对或更多具有相似系统增益的配对数据, 也无法准确建模线性关系, 因为由于采样范围的误差 (δ 在 Eqn. (4)中), 无法精确拟合线性关系, 如图 8所示。利用两点确定直线的原则, 我们采用两对具有轻微不同系统增益的配对数据来建模线性关系, 从而显著提高去噪能力。此外, 如图 7所示, 随着配对数据数量的增加, 线性关

系能够更加准确地拟合，进一步消除了回归误差。

6. 结论

为了缓解基于标定方法固有的缺陷，我们提出了一种无标定的管道，用于照亮每一片黑暗。得益于相机特定的对齐方法，我们将显式的标定过程替换为隐式的学习过程。CSA通过解耦相机特定的信息和噪声模型的通用知识，使得快速适应目标相机成为可能。此外，我们设计了一个并行卷积机制，用于学习去除模型外的噪声。通过对每个比率使用2对数据（共6对）和1.5K次迭代，我们实现了比现有方法更优的性能。

7. 致谢

本研究得到了国家自然科学基金（编号：62225604）、中央高校基础研究基金（南开大学，070-63233089）、中国博士后科学基金（编号：2021M701780）的资助。南开大学超级计算中心提供了计算支持。我们还得到了CAAI-华为MindSpore开放基金（CAAIXSJLJJ-2022-024A）的赞助。

References

- [1] Abdelrahman Abdelhamed, Marcus A Brubaker, and Michael S Brown. Noise flow: Noise modeling with conditional normalizing flows. In *ICCV*, 2019. 1, 5, 6
- [2] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *CVPR*, 2018. 1, 2
- [3] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv:1607.06450*, 2016. 2
- [4] Robert A. Boie and Ingemar J. Cox. An analysis of camera noise. *TPAMI*, 1992. 2
- [5] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *CVPR*, 2005. 1
- [6] Yue Cao, Ming Liu, Shuai Liu, Xiaotao Wang, Lei Lei, and Wangmeng Zuo. Physics-guided iso-dependent sensor noise modeling for extreme low-light photography. In *CVPR*, 2023. 6
- [7] Junbum Cha, Sanghyuk Chun, Kyungjae Lee, Hanchchol Cho, Seunghyun Park, Yunsung Lee, and Sungrae Park. Swad: Domain generalization by seeking flat minima. *Advances in Neural Information Processing Systems*, 34:22405–22418, 2021. 4
- [8] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *CVPR*, 2018. 1, 2, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15
- [9] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *ECCV*, 2022. 2
- [10] Linwei Chen, Ying Fu, Kaixuan Wei, Dezhi Zheng, and Felix Heide. Instance segmentation in the dark. *IJCV*, 2023. 2
- [11] Liangyu Chen, Xin Lu, Jie Zhang, Xiaojie Chu, and Chengpeng Chen. Hinet: Half instance normalization network for image restoration. In *CVPR*, 2021. 2
- [12] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *ICML*, 2020. 2
- [13] Shen Cheng, Yuzhi Wang, Haibin Huang, Donghao Liu, Haoqiang Fan, and Shuaicheng Liu. Nbnnet: Noise basis learning for image denoising with subspace projection. In *CVPR*, 2021. 4
- [14] Xiaohan Ding, Yuchen Guo, Guiguang Ding, and Jungong Han. Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks. In *ICCV*, 2019. 4
- [15] Xiaohan Ding, Xiangyu Zhang, Jungong Han, and Guiguang Ding. Diverse branch block: Building a convolution as an inception-like unit. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10886–10895, 2021. 2, 4
- [16] Xiaohan Ding, Xiangyu Zhang, Ningning Ma, Jungong Han, Guiguang Ding, and Jian Sun. Repvgg: Making vgg-style convnets great again. In *CVPR*, 2021. 2, 4
- [17] Hansen Feng, Lizhi Wang, Yuzhi Wang, and Hua Huang. Learnability enhancement for low-light raw denoising: Where paired real data meets noise modeling. In *ACM MM*, 2022. 1, 6
- [18] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *ICML*, 2017. 3
- [19] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *NeurIPS*, 2014. 2
- [20] Ryan D Gow, David Renshaw, Keith Findlater, Lindsay Grant, Stuart J McLeod, John Hart, and Robert L Nicol. A comprehensive tool for modeling cmos image-sensor-noise performance. *IEEE TED*, 2007. 2
- [21] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *CVPR*, 2019. 1
- [22] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *CVPR*, 2020. 2
- [23] Glenn E Healey and Raghava Kondepudy. Radiometric ccd camera calibration and noise estimation. *TPAMI*, 1994. 2
- [24] Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *TPAMI*, 2021. 2
- [25] Mu Hu, Junyi Feng, Jiashen Hua, Baisheng Lai, Jianqiang Huang, Xiaojin Gong, and Xian-Sheng Hua. Online convolutional re-parameterization. In *CVPR*, 2022. 10
- [26] Gabriel Huang, Issam Laradji, David Vazquez, Simon Lacoste-Julien, and Pau Rodriguez. A survey of self-supervised and few-shot object detection. *TPAMI*, 2022. 2
- [27] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *ICCV*, 2017. 2
- [28] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *ICML*, 2015. 2

- [29] Kenji Irie, Alan E McKinnon, Keith Unsworth, and Ian M Woodhead. A technique for evaluation of ccd video-camera noise. *IEEE TCSVT*, 2008. [2](#)
- [30] Pavel Izmailov, Dmitrii Podoprikin, Timur Garipov, Dmitry Vetrov, and Andrew Gordon Wilson. Averaging weights leads to wider optima and better generalization. *arXiv:1803.05407*, 2018. [4](#)
- [31] Geonwoon Jang, Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. C2n: Practical generative noise modeling for real-world denoising. In *CVPR*, 2021. [1](#)
- [32] Xin Jin, Ling-Hao Han, Zhen Li, Chun-Le Guo, Zhi Chai, and Chongyi Li. Dnf: Decouple and feedback network for seeing in the dark. In *CVPR*, 2023. [1](#)
- [33] Brian L Joiner and Joan R Rosenblatt. Some properties of the range in samples from tukey’s symmetric lambda distributions. *Journal of the American Statistical Association*, 1971. [3](#)
- [34] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *CVPR*, 2019. [2](#)
- [35] Yoonsik Kim, Jae Woong Soh, Gu Yong Park, and Nam Ik Cho. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization. In *CVPR*, 2020. [1](#), [2](#), [5](#), [6](#), [7](#), [10](#), [12](#), [13](#), [14](#), [15](#)
- [36] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980*, 2014. [5](#)
- [37] Mikhail Konnik and James Welsh. High-level numerical simulations of noise in ccd and cmos photosensors: review and tutorial. *arXiv:1412.4031*, 2014. [2](#)
- [38] Shayan Kousha, Ali Maleky, Michael S Brown, and Marcus A Brubaker. Modeling srgb camera noise with normalizing flows. In *CVPR*, 2022. [1](#)
- [39] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. *CVPR*, 2018. [1](#), [2](#), [5](#), [6](#), [7](#), [10](#), [12](#), [13](#), [14](#), [15](#)
- [40] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *ICLR*, 2017. [6](#)
- [41] Matteo Maggioni, Enrique Sánchez-Monge, and Alessandro Foi. Joint removal of random and fixed-pattern noise through spatiotemporal video filtering. *IEEE TIP*, 2014. [2](#)
- [42] Ali Maleky, Shayan Kousha, Michael S Brown, and Marcus A Brubaker. Noise2noiseflow: Realistic camera noise modeling without clean images. In *CVPR*, 2022. [1](#), [2](#)
- [43] Kristina Monakhova, Stephan R Richter, Laura Waller, and Vladlen Koltun. Dancing under the stars: video denoising in starlight. In *CVPR*, 2022. [1](#), [2](#), [5](#), [6](#), [7](#), [10](#), [12](#), [13](#), [14](#), [15](#)
- [44] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *NIPS Workshops*, 2017. [5](#)
- [45] K Ram Prabhakar, Vishal Vinod, Nihar Ranjan Sahoo, and R Venkatesh Babu. Few-shot domain adaptation for low light raw image enhancement. In *BMVC*, 2021. [2](#)
- [46] Sachin Ravi and Hugo Larochelle. Optimization as a model for few-shot learning. In *ICLR*, 2016. [3](#)
- [47] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015. [2](#), [4](#), [10](#)
- [48] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv:1607.08022*, 2016. [2](#)
- [49] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *CVPR*, 2018. [1](#)
- [50] Hans Wach and Edward R Dowski Jr. Noise modeling for design and simulation of computational imaging systems. In *Visual Information Processing XIII*, 2004. [2](#)
- [51] Yuzhi Wang, Haibin Huang, Qin Xu, Jiaming Liu, Yiqun Liu, and Jue Wang. Practical deep raw image denoising on mobile devices. In *ECCV*, 2020. [1](#)
- [52] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 2004. [6](#)
- [53] Kaixuan Wei, Ying Fu, Yinqiang Zheng, and Jiaolong Yang. Physics-based noise modeling for extreme low-light photography. 2021. [1](#), [2](#), [4](#), [5](#), [6](#), [7](#), [10](#), [11](#), [12](#), [13](#), [14](#), [15](#)
- [54] Jia-Wen Xiao, Chang-Bin Zhang, Jiekang Feng, Xialei Liu, Joost van de Weijer, and Ming-Ming Cheng. Endpoints weight fusion for class incremental semantic segmentation. In *CVPR*, 2023. [4](#)
- [55] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. Empirical evaluation of rectified activations in convolutional network. *arXiv:1505.00853*, 2015. [3](#)
- [56] Han-Jia Ye, Lu Ming, De-Chuan Zhan, and Wei-Lun Chao. Few-shot learning with a strong teacher. *TPAMI*, 2022. [2](#)
- [57] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Cycleisp: Real image restoration via improved data synthesis. In *CVPR*, 2020. [1](#), [4](#)
- [58] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. In *ECCV*, 2020. [2](#)
- [59] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *CVPR*, 2021. [2](#)
- [60] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for fast image restoration and enhancement. *IEEE TPAMI*, 2022. [1](#)
- [61] Chang-Bin Zhang, Jia-Wen Xiao, Xialei Liu, Ying-Cong Chen, and Ming-Ming Cheng. Representation compensation networks for continual semantic segmentation. In *CVPR*, 2022. [4](#)
- [62] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE TIP*, 2017. [1](#)
- [63] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Pfdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE TIP*, 2018. [1](#)
- [64] Yi Zhang, Hongwei Qin, Xiaogang Wang, and Hongsheng Li. Rethinking noise synthesis and modeling in raw denoising. In *ICCV*, 2021. [1](#), [2](#), [5](#), [6](#), [7](#), [10](#), [11](#), [12](#), [13](#), [14](#), [15](#)

Appendix

A. 网络架构

我们在Fig. 9中展示了所提LED的详细网络架构，该架构采用UNet风格 [47]，包含五个阶段。在编码器和解码器中，每个阶段都由两个顺序的RepNR块组成。值得注意的是，除了AINDNet [35]外，所有其他方法都与SID [8]共享相同的UNet架构。此外，经过重参数化后，LED最终会生成与SID [8]及其他方法相同的架构，以便进行公平比较（见Sec. B）。

B. 结构重参数化过程

在本节中，我们将详细介绍结构重参数化的过程。如Sec. 3.4中所述，RepNR块由串行和并行线性映射组成，这些映射可以融合为一个单一的操作。具体来说，RepNR块可以转换为一个简单的 3×3 卷积。形式上，该架构包含两个 3×3 的卷积，权重为 $\{W_0, W_1\}$ ，偏置为 $\{b_0, b_1\}$ ，其中一个卷积后跟随一个CSA层。设 $CSA(x) = kx + b$ ，其中 k, b 表示CSA层的权重和偏置。那么，输入 x 的结果可以表示为：

$$\begin{aligned}\tilde{x} &= W_0(CSA(x)) + b_0 + W_1x + b_1 \\ &= W_0(kx + b) + b_0 + W_1x + b_1 \\ &= (W_0k + W_1)x + (W_0b + b_0 + b_1) \\ &= \tilde{W}x + \tilde{b},\end{aligned}\tag{6}$$

这是整个部署过程的公式化表示。它表明我们的RepNR块可以转换为一个简单的 3×3 卷积，并且在推理过程中不会增加额外的成本。值得注意的是，我们采用了与 [25]相同的在线重参数化策略，因此在训练和测试之间没有任何性能差距。

C. 更多视觉结果

LED在细节恢复方面优于ELD [53]（基于标定的方法），如图10所示。如图11所示，LED在去除模型外噪声方面优于其他基于标定的方法 [53, 64]。在图??和图12-15中，我们提供了两个基准（ELD [53]和SID [8]）的更多结果。我们展示了Kristina *et al.* [43]、Noise2Noise [39]、AINDNet [35]、Zhang *et al.* [64]和ELD [53]的修复结果以供比较。

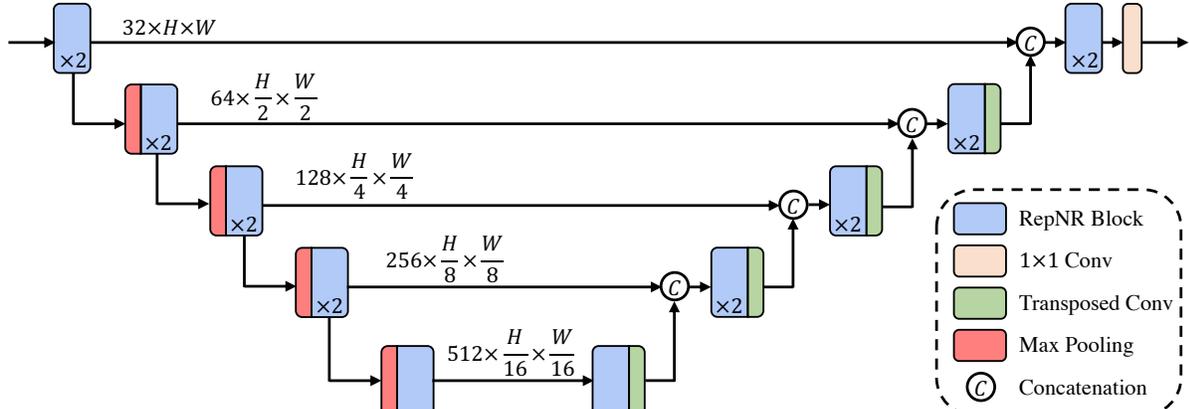


Figure 9. 我们提出的LED的详细网络架构。箭头上标注的 $\hat{C} \times \hat{H} \times \hat{W}$ 格式表示对应阶段的特征尺寸。 $H \times W$ 为输入分辨率。RepNR块中的 $\times 2$ 表示两个RepNR块按顺序堆叠。经过结构重参数化 (Sec. B)，我们的方法输出与SID [8]和其他方法相同的结构，以便进行公平比较。

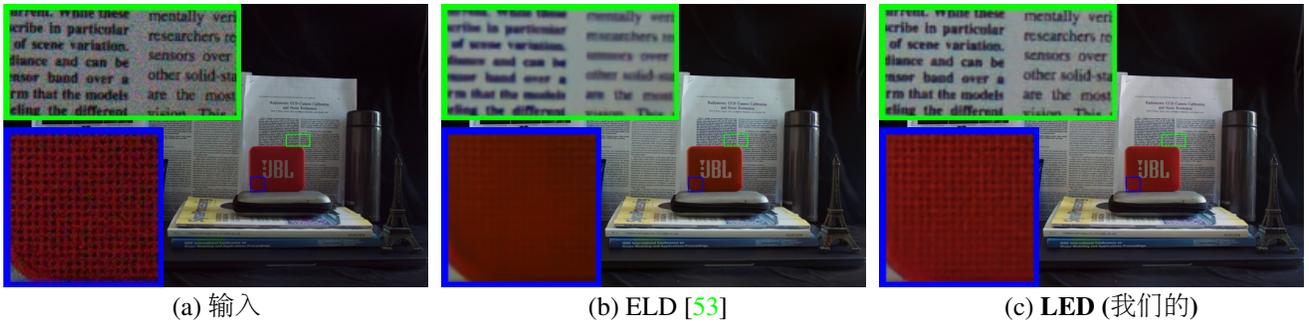


Figure 10. 提出的LED在细节恢复方面显著优于现有的最先进方法。

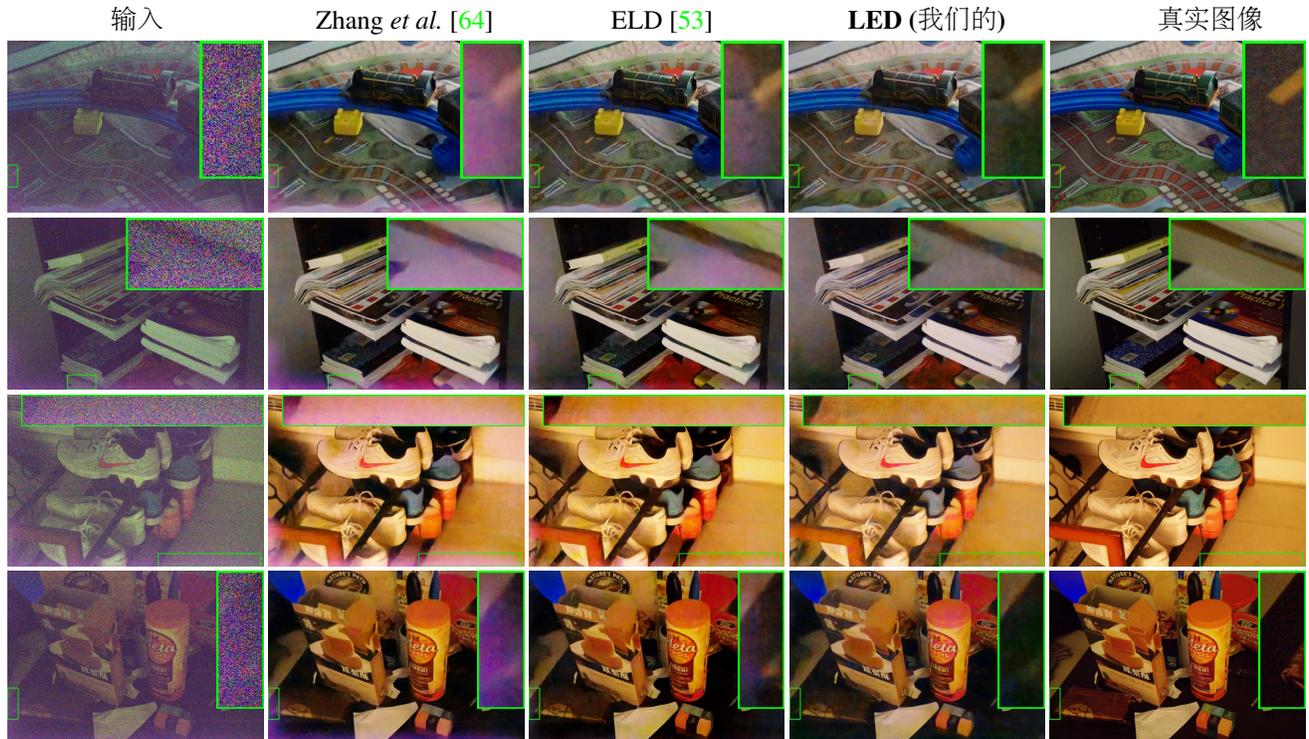


Figure 11. 与最先进的基于校准的方法 (ELD [53] 和 Zhang et al. [64]) 相比，提出的LED能够有效去除模型外的噪声 (放大查看效果最佳)。

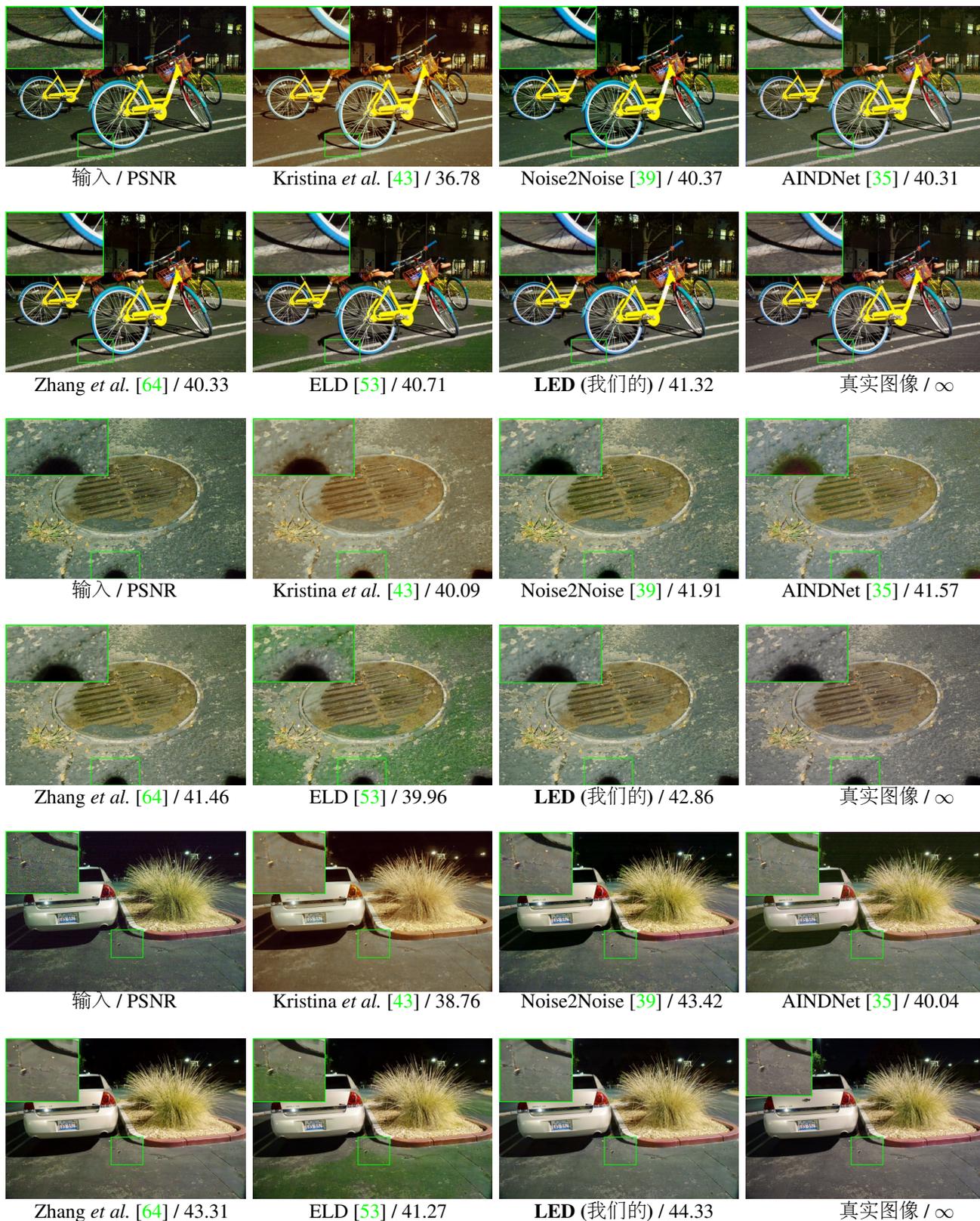


Figure 12. 我们的LED与其他最先进方法在SID [8]数据集上的视觉对比（最佳查看请放大）。我们使用与ELD [53]相同的ISP对输入图像进行放大和后处理。

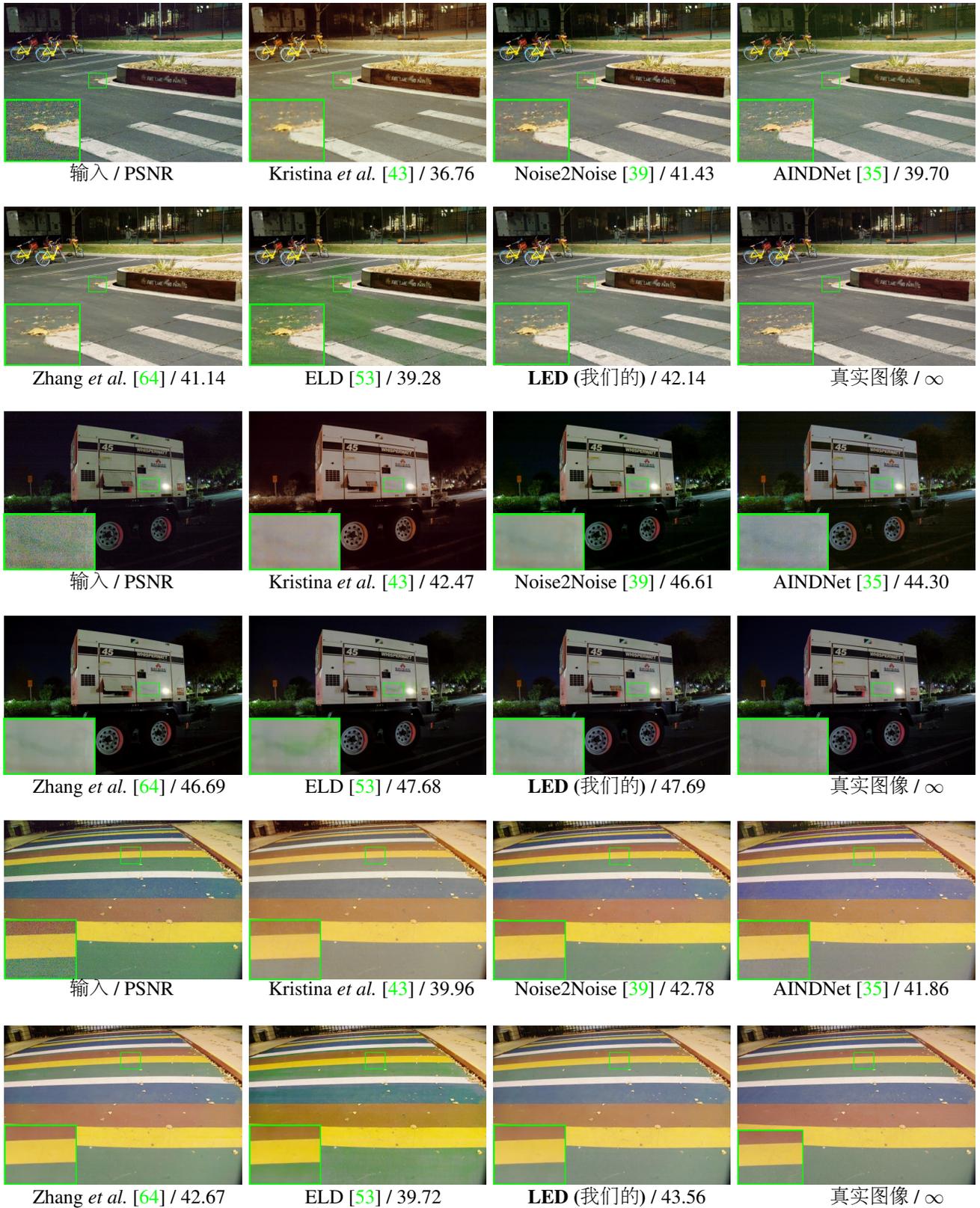


Figure 13. 我们的LED与其他最先进方法在SID [8]数据集上的视觉对比（最佳查看请放大）。我们使用与ELD [53]相同的ISP对输入图像进行放大和后处理。

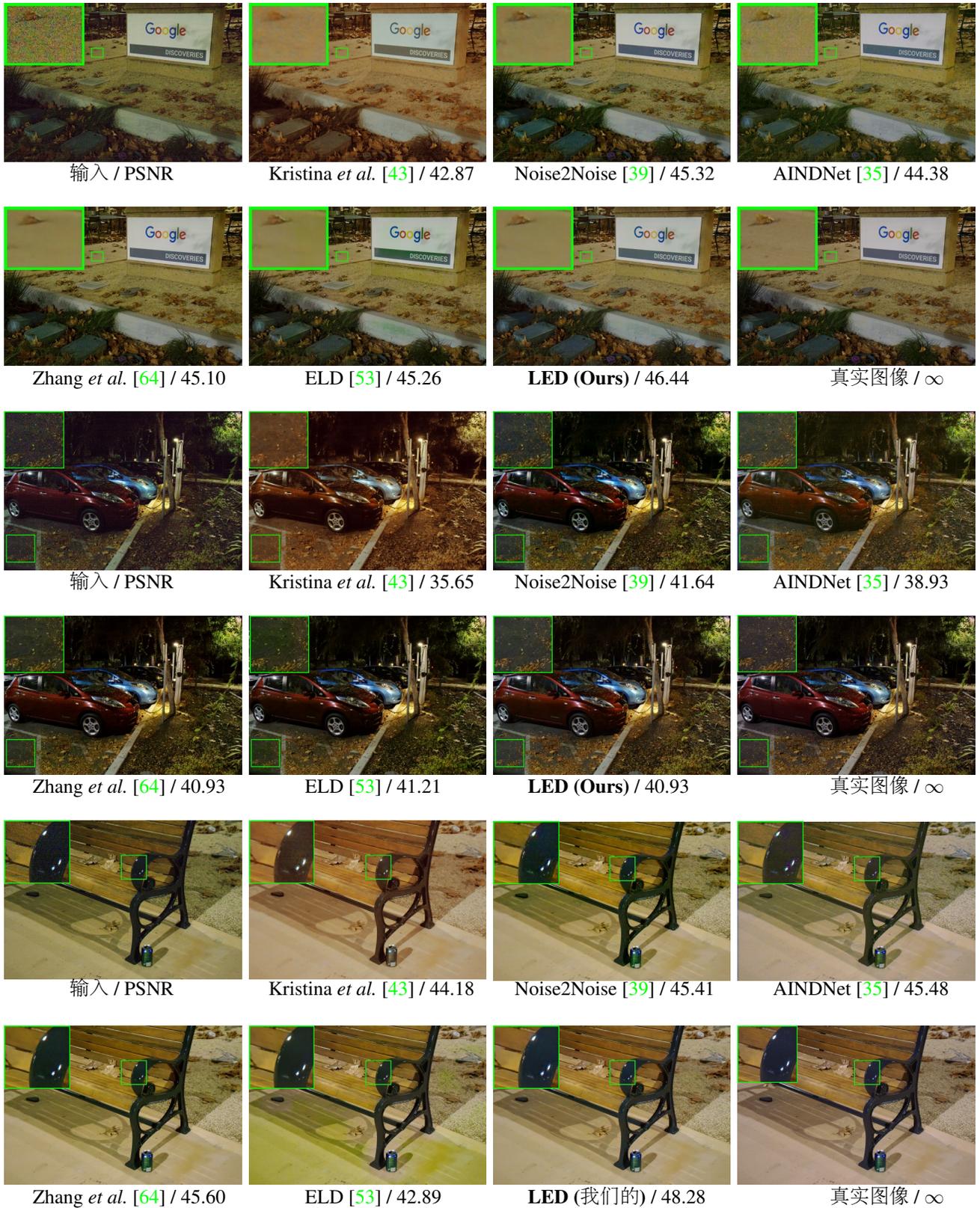


Figure 14. 我们的LED与其他最先进方法在SID [8]数据集上的视觉对比（最佳查看请放大）。我们使用与ELD [53]相同的ISP对输入图像进行放大和后处理。

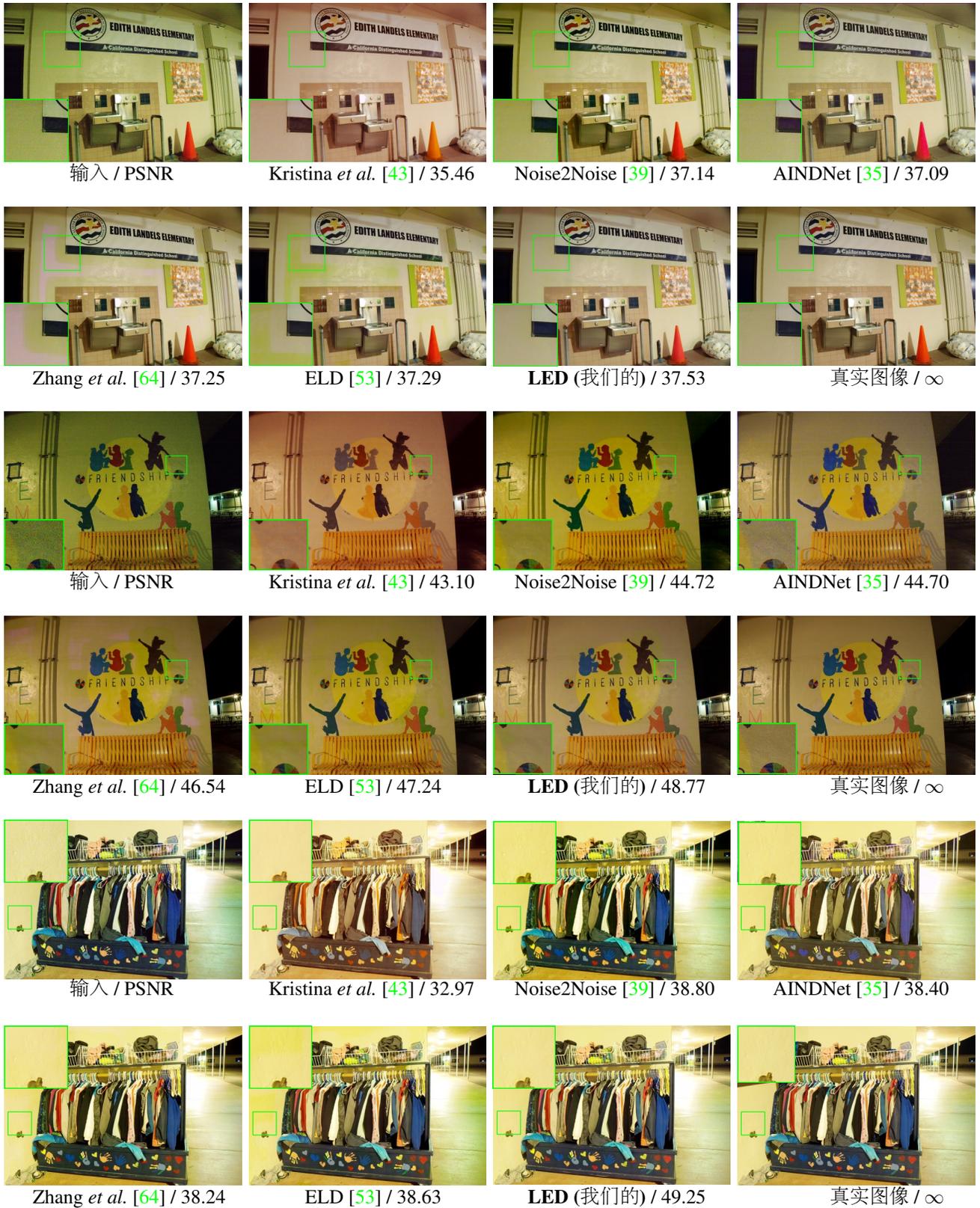


Figure 15. 我们的LED与其他最先进方法在SID [8]数据集上的视觉对比（最佳查看请放大）。我们使用与ELD [53]相同的ISP对输入图像进行放大和后处理。