

Lighting up NeRF via Unsupervised Decomposition and Enhancement

Haoyuan Wang, Xiaogang Xu, Ke Xu and Rynson W.H. Lau

2023

摘要

给定场景的一组图像和相应的相机姿势，神经辐射场 (NeRF) 能有效地合成该场景的新视图。然而，从低光场景拍摄的图像很难用于训练 NeRF 模型以产生高质量的结果，因为它们的像素强度低，噪声大并且颜色失真。虽然可以考虑将现有的低光图像增强方法与 NeRF 方法相结合，但由于单个二维增强过程导致的视图不一致，效果也不理想。在本文中，作者提出了一种新的方法，称为 Low-Light NeRF(或 LLNeRF)，以增强场景表现，并以无监督的方式直接从 sRGB 低光图像中合成正常光的新视图。其方法的核心是辐射场学习过程的分解，它能够与 NeRF 优化过程联合起来从而增强光照，降低噪声和纠正扭曲的颜色。而我的复现工作是，修改其增强网络提高学习能力，并且引入了一种即插即用的随机结构性损失作为辅助优化策略，能够使增强后的场景表现效果更好，有更好的结果表现。同时，我还在另一具有多视角的低光/正常光图像数据集上进行实验，也取得较好的效果。

关键词：神经辐射场；低光增强

1 引言

神经辐射场 (Neural Radiance Field, NeRF) [15] 是一种通过学习场景表示作为隐式函数来渲染新视图的强大方法。这些隐式函数由多层感知器 (MLP) 参数化，并通过测量输入视图的比色误差来优化。因此，高质量的输入图像是 NeRF 高质量结果的前提。换句话说，训练 NeRF 模型通常要求输入图像具有高可见度，并且几乎所有像素都真实地表示场景照明和物体颜色。但是，在低光条件下拍摄照片时，图像质量得不到保证。低光图像通常具有较低的能见度。由于光子较低，相机发出的噪声也相对放大，这进一步掩盖了场景细节并扭曲了物体颜色。低光图像的这种特性使得现有的 NeRF 模型无法生成高质量的新视图图像。

可以注意到，最近提出了一些从退化的输入图像训练 NeRF 模型的方法 [8, 12, 14]。Ma 等人 [12] 提出了一种从正常光照场景的模糊输入合成新视图的方法。Mildenhall 等人 [14] 表明当使用高动态范围 RAW 数据进行训练时，NeRF 对低光照输入图像的零均值噪声具有鲁棒性。Huang 等人 [8] 提出 HDR- Nerf，它从一组在不同已知曝光水平下拍摄的低动态范围 (LDR) 输入图像中产生高动态范围 (HDR) 的新视图。后两种方法利用 RAW 图像中记录的 HDR 信息和元数据 (即曝光水平) 来增强场景表现。然而，这些方法不适用于由相机在低光

场景中拍摄的 sRGB 图像 (8 位/通道)。与 RAW 数据不同, sRGB 图像是由相机 ISP 过程生成的。它们具有低动态范围和低信噪比。

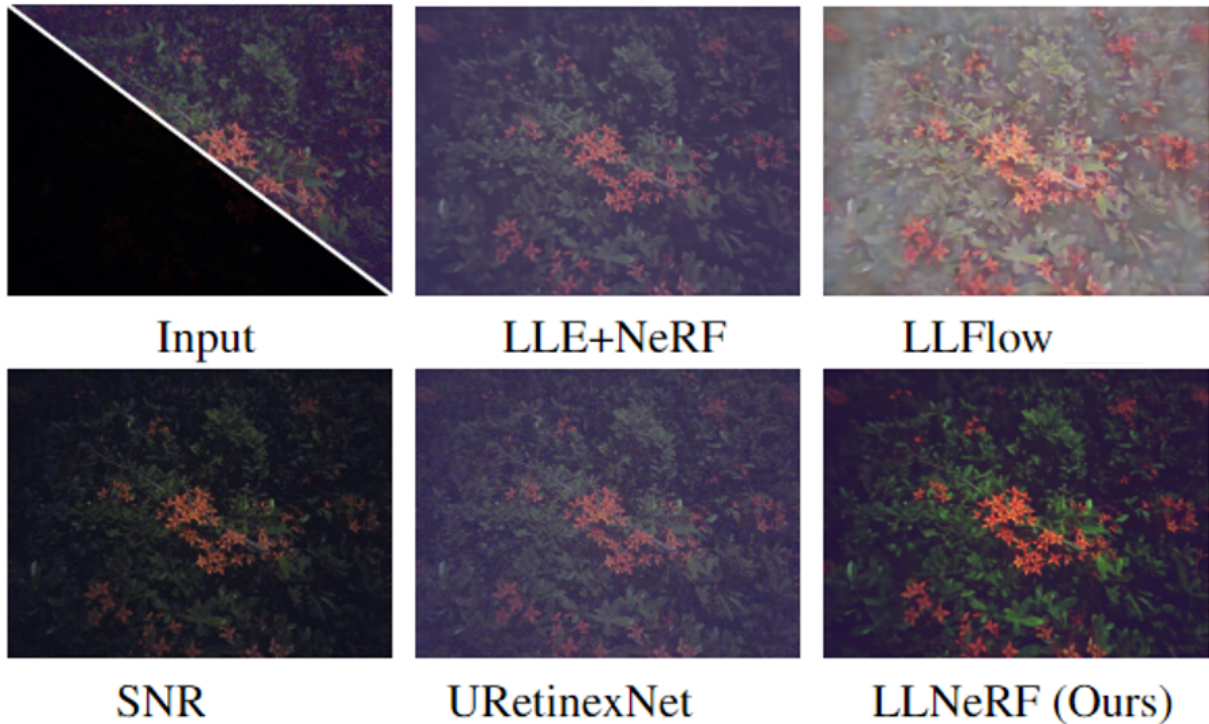


图 1. baseline 模型 (LLE+NeRF)、SOTA 低光增强模型和本文模型的比较

这个问题的一个简单解决方案是首先增强低光输入图像, 然后使用增强的结果来训练 NeRF 模型。然而, 虽然这可能能够提高亮度, 但现有的低光增强模型没有考虑如何保持多视图图像的一致性。此外, 这些基于学习的增强方法倾向于从它们自己的训练数据中学习亮度的特定映射, 这可能不能很好地推广到室外场景中。由于 NeRF 的视图相关优化, 这两个原因导致 NeRF 在不同视图学习中学习有偏差的信息, 从而产生不切实际的新视角图像。如图1所示, 其展示了 baseline 模型即 LLE+NeRF, SOTA 的二维低光增强模型以及本文作者的模型对输入的低光图像进行增强的效果。可以看到, 本文作者的模型效果较好。

本文的作者提出了一种新的方法, 在没有真实图像的情况下, 从一组 8 位的低光 sRGB 图像中渲染出新的正常光图像。其关键的解决方案是: 在 NeRF 优化中, 3D 点的颜色可以解耦为与视角相关和与视角无关的分量, 并且与视图相关的分量主要收到光照的影响。因此, 对与光照相关的视角相关分量的操作能够在保持场景的纹理和结构的同时提高亮度、校正颜色和减少噪声。实验表明, 作者提出的方法优于最先进的 NeRF 模型和 baseline (即将 NeRF 与最先进的增强方法相结合)。

总之, 本文作者提出了第一种从 LDR 低光图像集合中重建合适照明的 NeRF 模型的方法。主要贡献包括:

1. 建议将 NeRF 分解为与视角相关和与视角无关的颜色分量以进行增强。其分解过程不需要真实图像。
2. 提出了一种无监督的方法来增强照明并校正颜色, 同时渲染无噪声的新视角图像。

3. 制作了该任务下的一个真实世界的数据集，并进行了广泛的实验，在真实世界的场景中证明其有效性。

2 相关工作

2.1 神经辐射场

神经辐射场通过参数化的隐式函数表示 3D 场景，并允许渲染高质量的新视角图像。然而，NeRF 对输入图像敏感，因为它依赖于输入图像的色度优化。一些方法侧重于通过使用飞行时间数据 [2]、潜在外观建模 [13]、相机自校准 [10]、深度估计 [6, 18] 和语义标签 [24] 来提高 NeRF 对野外动态场景的鲁棒性。

其他的一些方法 [8, 12, 14] 提出了从退化的输入图像训练 NeRF 模型。Ma 等人 [12] 提出了一种可变形的稀疏核模块，用于去模糊，同时从模糊输入合成新视角图像。Mildenhall 等人 [14] 提出直接在相机捕获的原始图像上训练 NeRF 以处理低光场景的低能见度和噪声。Huang 等人 [8] 提出 HDR-NeRF 从不同曝光水平的 LDR 图像集合中合成新视角的 HDR 图像，使用色调映射器隐式处理曝光融合。与上述方法不同，本文的作者旨在解决使用一组低光 sRGB 图像来训练 NeRF 的问题，而由于这些低光图像能见度低、动态范围低、噪声大、颜色失真高，因此更具挑战性。

2.2 低光增强

低光增强旨在提高从低光场景拍摄的图像的内容可见性。一系列基于深度学习的增强方法学习从低光图像到经过专家修饰的图像或使用高端相机捕获的图像的特定映射。这些方法提出了不同的先验和技术，旨在提高神经网络学习这种映射的能力，例如使用 HDR 信息 [7]，生成对抗学习 [9]，深度参数过滤器 [16]，以及强化学习 [17]。一些方法提出将图像分解为光照层和细节层 [4]、不同频率分量层 [21] 和不同曝光区域 [1] 以增强低光图像。最近，Xu 等人 [22] 提出将 transformer 和 CNN 结合起来，从而对低光增强的长期相关性进行建模。

本文作者的工作在精神上更接近于基于 Retinex 理论的增强方法 [11, 19, 23]。这些方法首先将输入图像分解为照明层和反射层，然后增强图像的照明层。虽然这些方法从通常缺乏几何信息的 2D 图像中学习这种分解，但本文方法是在辐射场中工作，从而产生更真实的分解和增强。

3 预备知识和分析

这一部分首先总结了神经辐射场（NeRF）在正常光场景下的工作原理，然后分析和解释了 NeRF 在处理低光场景时面临的挑战。

3.1 NeRF 预备知识

给定一组多视角的训练图像，NeRF [15] 通过学习从而渲染光线 r 上每一个像素的颜色 \mathbf{c}_r ，而光线 r 可以通过相机索引和 2D 像素坐标唯一识别。NeRF 通过辐射场来表示场景，该辐射场将任意单个光线投射到 $r(t) = \mathbf{o} + t\mathbf{d}$ 作为输入，其中 \mathbf{o} ， \mathbf{d} ， t 分别是光线原点、光线

方向和沿光线的距离。渲染过程有三个步骤: (1) NeRF 在近端和远端图像平面之间采用分层采样策略沿着光线 $r(t)$ 采集 n 个采样点, 其中 $t_i \in \mathbf{t}$, \mathbf{t} 是一个 n 维向量; (2) NeRF 将一个可选的变换函数 $\psi(\cdot)$ 应用到沿光线采样的坐标向量 \mathbf{t} 上; (3) NeRF 使用多层感知机 MLPs 的 $F_{\text{density}}, F_{\text{color}}$ 从 \mathbf{t} 和视角方向 \mathbf{d} 学习沿光线上采样点的体密度和颜色, 分别用 σ 和 \mathbf{c} 表示为:

$$\begin{cases} (\tau, \sigma) = F_{\text{density}}(\psi(r(t_i)), \mathbf{d}; \Theta_{F_{\text{density}}}), t_i \in \mathbf{t} \\ \mathbf{c} = F_{\text{color}}(\tau, \mathbf{d}; \Theta_{F_{\text{color}}}), \end{cases} \quad (1)$$

其中 τ 是神经网络学习到的中间特征。不同的 NeRF 实现可能有不同版本的转换函数 $\psi(\cdot)$ 。原始 NeRF 的实现 [15] 使用了频率位置编码函数作为 $\psi(\cdot)$, 而在 Mip-NeRF [3] 中, $\psi(\cdot)$ 被实现为区间分割和集成的位置编码。在本文中, 作者使用 Mip-NeRF [3] 的实现方法, 像素颜色通过下面公式进行渲染:

$$\mathbf{c}_r = \sum_i w_i \mathbf{c}_i = \sum_i (1 - e^{-\sigma \delta_i}) e^{-\sum_{j < i} \sigma_j \delta_j} \mathbf{c}_i, \quad (2)$$

其中 $\delta_i = t_{i+1} - t_i$ 。 \mathbf{c}_r 是对应光线 $r(t)$ 的最终渲染的 3 通道像素颜色。然后在训练图像的真实像素颜色 $\tilde{\mathbf{c}}_r$ 的监督下对 NeRF 进行优化。

3.2 挑战

由于 NeRF 模型直接根据 2D 投影图像优化其隐式辐射场, 因此使用低光 sRGB 图像训练 NeRF 模型存在两个挑战。首先, NeRF 无法处理低光图像的低像素强度, 只能产生暗图像作为新视角图像。其次, 虽然 RawNeRF [14] 表明 NeRF 由于其本质的积分过程, 在 Raw 域对零均值噪声具有鲁棒性, 但相机处理后的 sRGB 图像的信噪比远低于 Raw 图像。此外, 相机的 ISP 过程改变了 Raw 图像的线性特性, 并且得到的 sRGB 图像混合了噪声。因此, 当在低光 sRGB 图像上进行训练时, NeRF 不能处理噪声和颜色失真。

为了获得正常光 NeRF, 将低光增强方法与 NeRF 相结合 (LLE+NeRF) 可能是一种解决方案。然而, 由于现有的低光增强方法主要基于特定的训练数据来学习从低光到正常光的映射。这种映射可能不能很好地推广到训练数据分布之外的新场景。因此, 使用通过这些现有方法增强的图像来训练 NeRF 模型可能会产生低质量的新视角图像。另一方面, 同时拍摄低光和正常光的多视角图像作为训练数据是不现实的。

作者的目标是开发一种方法, 以无监督的方式从低光场景中生成高质量的新视角图像。

4 本文方法

4.1 本文方法概述

本文作者的主要思想是分解 NeRF 的隐式辐射场, 然后利用先验来增强光照、减少噪声和校正新视角图像的颜色。图2(c) 显示了本文方法的流程。

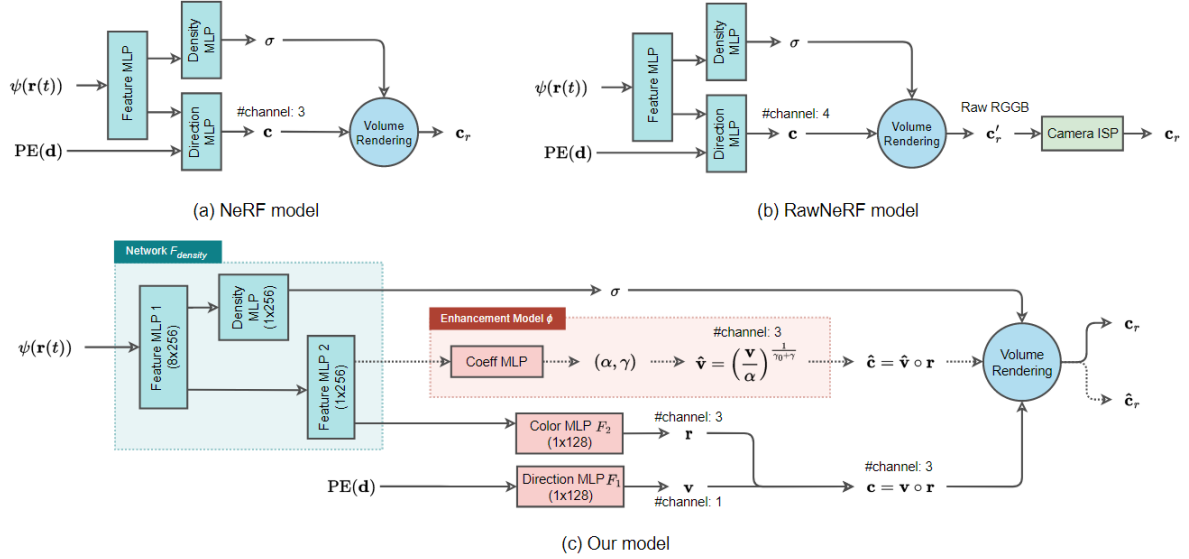


图 2. NeRF 模型 [15](a)、RawNeRF 模型 [14](b) 和本文模型 (c) 的图示。本文无监督增强方法的数据流显示在虚线内。本文模型共同学习新视角图像及增强射线上所有采样点的输出。每个最终被增强的像素都是使用体渲染公式渲染的，如公式2所示

4.2 神经辐射场的分解

如图3所示，当静态场景中的一个 3D 点 \mathbf{x} 投影到两个视角上的两个像素 ($\tilde{\mathbf{c}}_{ri}$ 和 $\tilde{\mathbf{c}}_{rj}$) 时，两个像素的颜色可能会出现不同，因为物体表面可能不是各向同性的，照明也不均匀。然而，这两个像素的颜色仍然在色谱的相同范围内。这表明一个 3D 点 \mathbf{x} 的颜色可以分解为与视角无关的基本分量和与视角相关的分量。与视角无关的基础分量表示本征颜色，它决定了 \mathbf{x} 的颜色的光谱范围。与视角相关的分量考虑了可能导致不同视角之间颜色差异的因素 (在大多数情况下，光照是主要因素，这取决于光源的位置和颜色以及 3D 点 \mathbf{x} 所在表面的朝向)。

受此启发，作者提出将颜色 \mathbf{c} 分解为与光照相关的视角相关分量 \mathbf{v} 及其表示基色的互补分量 \mathbf{r} 的乘积。利用 NeRF 来约束 \mathbf{v} 使其与视角相关，并进一步将其表述为专注于操纵光照强度的单通道表示。

考虑公式2中对图像 \mathbf{I} 的像素 \mathbf{c}_r 的渲染。由于任意图像的像素 \mathbf{c}_r 是沿光线上的所有 $\{\mathbf{c}_i\}_{i=1}^n$ 与视角相关颜色的加权累加，因此将沿光线的每个 \mathbf{c} 分解为 \mathbf{v} 和 \mathbf{r} ，并通过学习来增强颜色：

$$\begin{cases} \mathbf{v} = F_1(\tau, \mathbf{d}; \Theta_{F_1}) & \text{and} & \mathbf{r} = F_2(\tau; \Theta_{F_2}), \\ \mathbf{c} = \mathbf{v} \circ \mathbf{r} & \text{and} & \hat{\mathbf{c}} = \phi(\mathbf{v}) \circ \mathbf{r}, \end{cases} \quad (3)$$

其中， $\hat{\mathbf{c}}$ 为增强的颜色， ϕ 为神经网络参数化的增强函数， \circ 为像素乘法。 F_1 和 F_2 是两个 MLP。从而增强后的图像 \mathbf{I}_e 为：

$$\mathbf{I}_e = \{\hat{\mathbf{c}}_r\}, \hat{\mathbf{c}}_r = \sum_i w_i \phi(\mathbf{v}_i) \circ \mathbf{r}_i. \quad (4)$$

这种方法使模型能够学习合理的分解，该分解具有简单的形式，但在跨视角监督未增强的颜色 \mathbf{c} 时具有很强的约束。

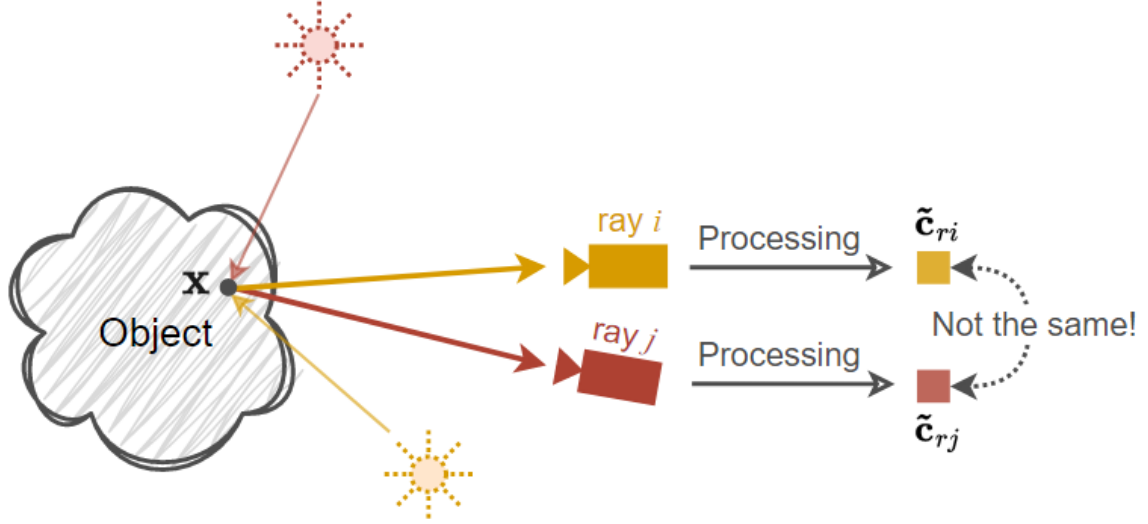


图 3. 同一空间点 \mathbf{x} 的二维投影 $\tilde{\mathbf{c}}_{ri}, \tilde{\mathbf{c}}_{rj}$ 不完全相同，但在同一光谱中。颜色的变化取决于视角，即观察到的颜色中与视角相关的成分主要受到光照的影响。

与基于图像的分解的区别。

4.3 无监督增强

在无监督分解的基础上，作者提出了一种无监督增强方法来增强光照 NeRF 模型。

对于每个空间坐标和视角方向，使用公式4来增强光线上每个采样点的 \mathbf{v} ，即 $\hat{\mathbf{v}} = \phi(\mathbf{v})$ 。作者提出在渲染的 RGB 值 $\hat{\mathbf{c}}_r$ 的约束下使用动态伽马校正来增强 \mathbf{v} ，如：

$$\hat{\mathbf{v}} = \phi(\mathbf{v}) = \left(\frac{\mathbf{v}}{\alpha}\right)^{\frac{1}{\gamma_0 + \gamma}} \quad (5)$$

其中 α 是一个标量， γ 是一个三维向量。这两个系数都是增强网络 ϕ 的输出。 γ_0 是初始化非线性变换的固定值。 α 定义为全局调节光照增益的标量；并且通过在先验损失函数的约束下，通过对三个颜色通道的 \mathbf{v} 进行小置换，将 γ 定义为一个三维向量，用于颜色失真校正。

通过应用公式5， \mathbf{v} 沿光线增强，而 \mathbf{r} 不变。因此，本文模型可以在保持场景几何信息的同时调整场景的照明和颜色。虽然模型允许应用更复杂的转换函数，但通过实验发现，公式5在性能和计算成本之间取得了很好的平衡。

4.4 优化策略

本文方法是以端到端的方式训练模型，如图4所示。通过训练数据集上的光线对模型进行迭代优化的同时，提供了三种监督信号：基于灰色世界先验的比色监督和基于平滑先验的监督用于优化增强网络，基于数据的监督用于优化辐射场。

灰色世界先验的色度监督。作者制定了一个简单但有效的基于灰色世界先验的损失 L_c 来约束增强网络 ϕ 的学习，以产生逼真的图像：

$$L_c = \mathbb{E}[(\hat{\mathbf{c}}_r - e)^2] + \lambda_1 \mathbb{E}\left[\frac{\text{var}_c(\hat{\mathbf{c}}_r)}{\beta_1 + \text{var}_c(\mathbf{r}_r)}\right] + \lambda_2 \|\gamma\|_2, \quad (6)$$

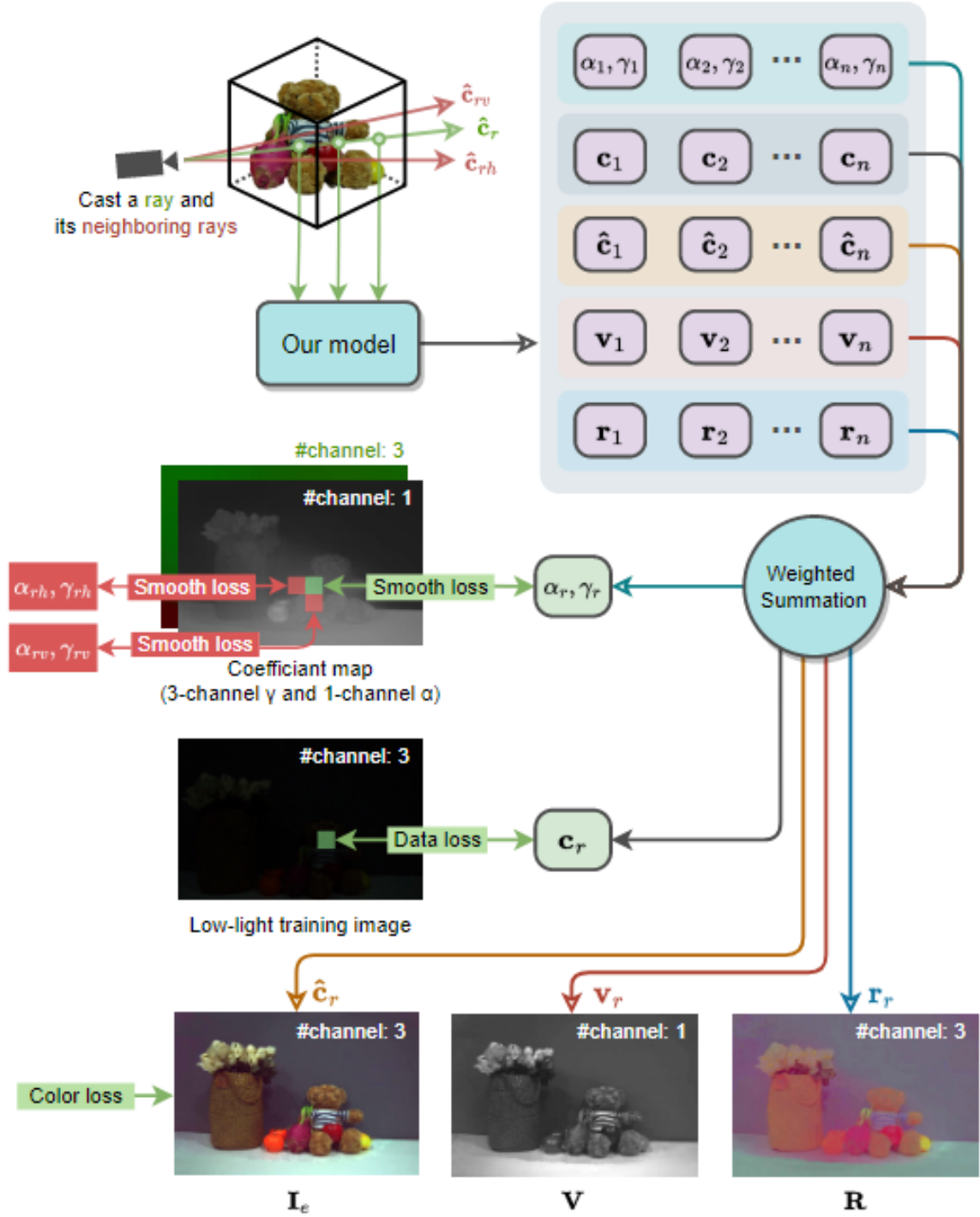


图 4. 训练过程和损失函数的图示。

其中 $e, \beta_1, \lambda_1, \lambda_2$ 是超参数, 并且 \mathbf{var}_c 表示通道的方差。公式6的第一项是提高像素的亮度 (其中 $e = 0.55$)。第二项是基于灰度世界先验对颜色进行校正, 通过减小三个通道间的方差, 将失真的颜色推向自然分布。为了防止渲染像素收敛为灰色, 我们进一步根据沿光线的加权基色 \mathbf{r} 的颜色添加动态权重, 以放松对高度饱和颜色的约束。第三项是防止过拟合的正则化项。

基于平滑先验的监督。为了在增强的辐射场中保留场景的颜色和结构, 并约束系数 (α 和 γ) 的学习, 我们期望集成的系数产生局部平滑的映射。因此, 我们将这两个系数的加权和相对于图像空间中集成的 \mathbf{v}_r 的梯度约束为:

$$L_s = \underbrace{\mathbb{E}[(\frac{\partial \alpha_r}{\partial \mathbf{v}_r})^2]}_{L_{sa}} + \underbrace{\mathbb{E}[(\frac{\partial \gamma_r}{\partial \mathbf{v}_r})^2]}_{L_{sg}}. \quad (7)$$

由于训练中光线是随机采样的, 直接从公式7中很难获得所需的梯度信息。因此写出公式7中的 L_{sa} 成的离散近似为:

$$L_{sa} = \frac{1}{2} \left[\frac{(\alpha_r - \alpha_{rh})^2}{(\mathbf{v}_r - \mathbf{v}_{rh})^2 + \epsilon_1} + \frac{(\alpha_r - \alpha_{rv})^2}{(\mathbf{v}_r - \mathbf{v}_{rv})^2 + \epsilon_1} \right], \quad (8)$$

其中 $\alpha_{rh}, \alpha_{rv}, \mathbf{v}_{rh}, \mathbf{v}_{rv}$ 为图像空间中水平方向和垂直方向上相邻射线集成的 α 和 \mathbf{v} 。为了利用平滑性约束, 如图4所示, 在每个优化步骤中对光线及其相邻光线进行采样。 L_{sg} 的获取方法与 L_{sa} 相似。

数据监督。为了学习场景几何, 作者应用了 [14] 中的数据损失, 即 $\mathbb{E}[\eta(\tilde{\mathbf{c}}_r) - \eta(\mathbf{c}_r)]$ 的线性化, 其中 $\eta(y) = \log(y + \epsilon_2)$ 。由于训练图像中的大多数像素具有低强度, 因此使用色调映射函数 η 来放大暗区域的误差, 以促进学习过程。

5 复现细节

5.1 与已有开源代码对比

我的工作是在本文的代码中进行修改的, 主要是加入了即插即用的随机结构性损失用于辅助优化, 以提升细节。而随机结构性损失是根据 [20] 中的代码进行修改。因为该代码是用 Pytorch 框架进行编写, 而本文的代码是在 Jax 库环境下进行编写的, 所以我主要将其代码框架进行转换, 以适配本文代码的 Jax 库环境从而进行模型训练。

5.2 实验环境搭建

本文的实验均在 RTX 3090 显卡上进行, 代码框架为 Jax。Cuda 版本 11.3, 并且 jax 的版本为 0.3.15。

5.3 数据集

5.3.1 本文数据集

作者收集了一个真实世界的数据集作为模型学习和评估的基准。为了获得具有真实噪声分布的真实低照度图像, 其在夜间室外场景或包含多种物体的低照度室内场景中进行拍摄。由

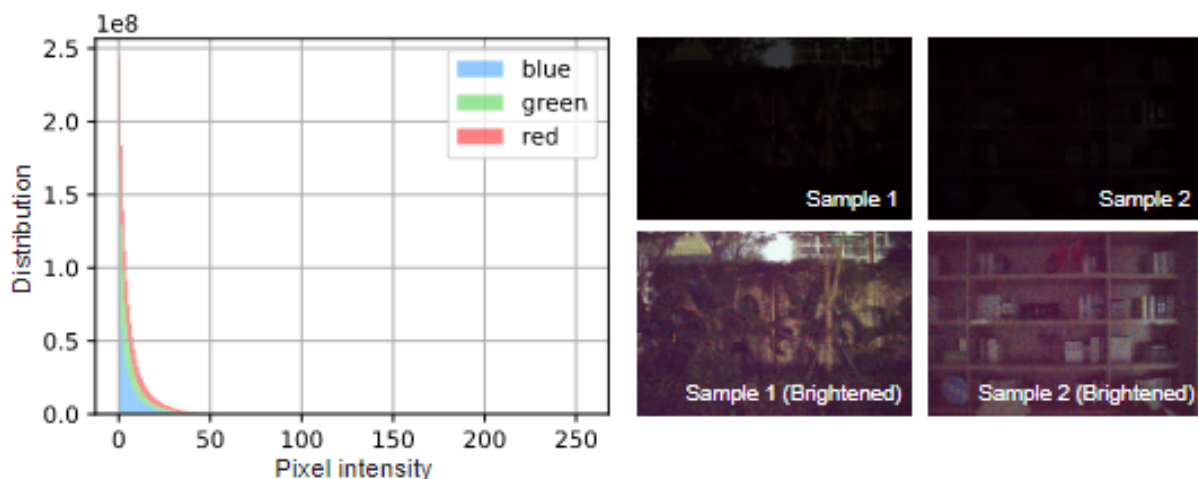


图 5. 本文数据集的强度分布和样本图像。收集了室内和室外场景的低光图像。这些图像通常具有低像素强度，明显的颜色失真和重噪点。

于 ISP 操作依赖于设备，并且设备之间的噪声分布也不同，因此作者使用手机相机和数码单反相机收集数据，以丰富数据集的多样性。图5中展示了数据集的一些样本和统计数据。如图所示，本文数据集的平均亮度非常低 (大多数像素的亮度低于 255 中的 50)。此外，这些图像中的噪声和颜色失真程度非常高，这使得该任务非常具有挑战性。

5.3.2 Aleth-NeRF 数据集

而因为本文的数据集没有低光图像对应的正常光图像，因此我无法进行对模型进行完整的定量和定性分析。因此我在这里引用来自 Aleth-Nerf [5] 这篇文章中的数据集。如图6所示，其在现实世界中收集了 5 个场景（“布偶”、“椅子”、“沙发”、“自行车”、“灌木”）。每个场景包括 25-65 张图像，每个场景都包含低光和正常光 (通过调整曝光时间和 ISO 来消除曝光过度)，而相机的其他配置是固定的。图像原始的分辨率为 3000×4000 。为了方便，对原始分辨率进行了下采样 8 倍为 375×500 ，并在正常光图像上通过采用 COLMAP 生成真实标签图像的视角和角度信息。对于数据集分割，在每个场景中，选择 3-5 张图像作为测试集，1 张图像作为验证集，和其他图像作为训练集。通过这个具有多视角的低光/正常光图像对的数据集，我就能够对本文的模型更好地进行定量分析和定性分析。

5.4 创新点

而我主要的工作就是，针对作者提出的模型在各种场景上的表现进行改进。包括：

1. 引入一种即插即用的随机结构性损失作为辅助优化策略
2. 在另一具有多视角的低光/正常光数据集中验证方法的有效性

从而能够提高模型的学习能力，并且考虑场景的全局信息，能在增强结果上有更好的结果。

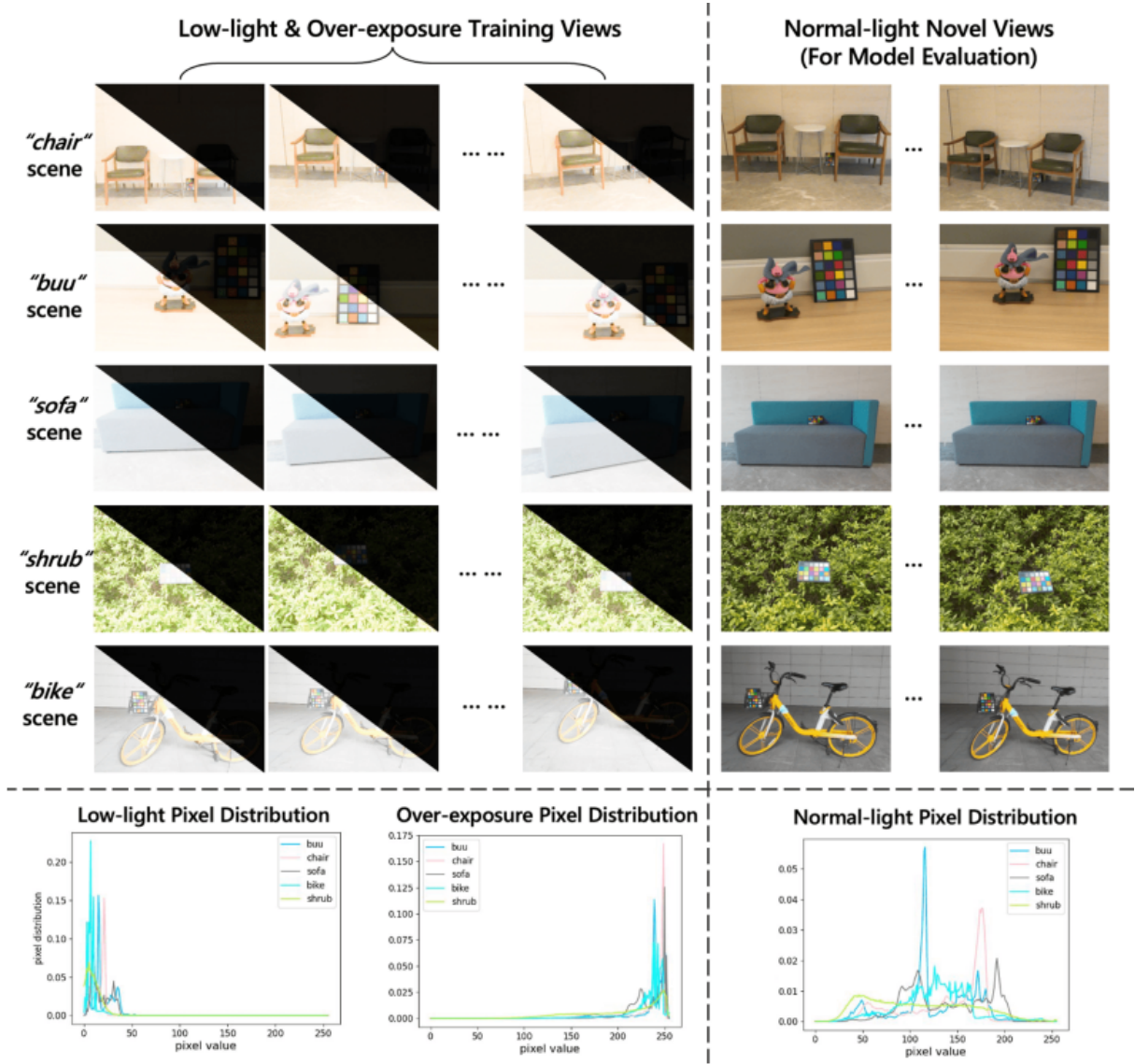


图 6. Aleth-NeRF [5] 收集了包含 5 个场景的数据集。每个场景包含低光和过度曝光训练视图用于模型训练，以及用于性能评估的真实正常光新视角图像。



(a) *LLNeRF*

(b) *LLNeRF_S3IM*

图 7. 优化前与优化后的对比图

6 实验结果分析

由图7可知，在本文数据集中的一个场景进行了实验，发现优化后能够一定程度地提升细节，但相比之下会损失一定的亮度，因此这也是后面的一个改进方向。除此之外，因为上面提到本文数据集没有提供参考图像，因此我还在另一数据集中进行了定量分析，如表1所示。

表 1. 优化前与优化后的定量分析

Method	bike	buu	chair	shrub	sofa	Avg
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Aleth-NeRF [5]	17.56/0.719	19.14/0.839	16.97/0.847	17.55/0.484	16.97/0.847	17.64/ 0.747
LLNeRF(origin)	21.38/ 0.784	19.24/0.873	15.88/0.819	14.25/0.343	16.52/ 0.869	17.45/0.738
LLNeRF(S3IM)	23.16/0.777	21.11/0.878	16.08/0.809	14.75/0.394	16.35/0.855	18.29/0.743

由该表可以知道，与优化前相比，加入了随机结构性损失作为辅助优化的指标结果更好，PSNR 大约提升 0.8dB，并且 SSIM 指标的结果也更好。其次还与 Aleth-NeRF 进行了比较，优化后的模型能在 PSNR 指标上超过该方法，而优化前的则要低于 Aleth-NeRF。总而言之，随机结构性损失能提升模型的细节表现和性能。

7 总结与展望

在本文中，作者提出了一种从低光 sRGB 图像中训练 NeRF 模型的新方法，以产生具有高可见度，鲜艳色彩和细节的新视图图像。基于对成像过程的观察，本文模型以无监督的方式将神经辐射场分解为与光照相关的视角相关分量和与视角无关的颜色基分量。模型在基于

先验的损失函数的监督和没有参考图像的情况下增强了照明。同时我进行了大量的实验来分析方法的特性，证明它与现有方法相比是有效的。其次，我对其进行了优化，改善了图像细节，得到较好的结果。最后我还在另一多视角的低光/正常光数据集上验证了方法的有效性。

而对于未来的计划，可以在训练速度更快的框架上进行集成，因为本文的模型还是基于普通 Nerf 的框架，因此其训练速度还是非常慢。其次，还考虑自制数据集，在更多的场景上进行实验更好地分析模型性能。

参考文献

- [1] Mahmoud Afifi, Konstantinos G. Derpanis, Bjorn Ommer, and Michael S. Brown. Learning multi-scale photo exposure correction. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2021.
- [2] Benjamin Attal, Eliot Laidlaw, Aaron Gokaslan, Chang-Il Kim, Christian Richardt, James Tompkin, and Matthew O’ Toole. Törf: Time-of-flight radiance fields for dynamic scene view synthesis. *Neural Information Processing Systems, Neural Information Processing Systems*, Dec 2021.
- [3] Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P. Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct 2021.
- [4] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, page 2049–2062, Apr 2018.
- [5] Ziteng Cui, Lin Gu, Xiao Sun, Xianzheng Ma, Yu Qiao, and Tatsuya Harada. Aleth-nerf: Illumination adaptive nerf with concealing field assumption, 2023.
- [6] Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised nerf: Fewer views and faster training for free. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2022.
- [7] Michaël Gharbi, Jiawen Chen, Jonathan T. Barron, Samuel W. Hasinoff, and Frédo Durand. Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics*, page 1–12, Aug 2017.
- [8] Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Xuan Wang, and Qing Wang. Hdr-nerf: High dynamic range neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2022.
- [9] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, and Kenneth Vanhoey. Dslr-quality photos on mobile devices with deep convolutional networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

- [10] Yoonwoo Jeong, Seokjun Ahn, Christopher Choy, Animashree Anandkumar, Minsu Cho, and Jaesik Park. Self-calibrating neural radiance fields. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct 2021.
- [11] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. *Cornell University - arXiv, Cornell University - arXiv*, Dec 2020.
- [12] Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and PedroV. Sander. Deblur-nerf: Neural radiance fields from blurry images. *Cornell University - arXiv, Cornell University - arXiv*, Nov 2021.
- [13] Ricardo Martin-Brualla, Noha Radwan, Mehdi S. M. Sajjadi, Jonathan T. Barron, Alexey Dosovitskiy, and Daniel Duckworth. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2021.
- [14] Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P. Srinivasan, and Jonathan T. Barron. Nerf in the dark: High dynamic range view synthesis from noisy raw images. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2022.
- [15] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.
- [16] Sean Moran, Pierre Marza, Steven McDonagh, Sarah Parisot, and Gregory Slabaugh. Deeplpf: Deep local parametric filters for image enhancement. *arXiv: Computer Vision and Pattern Recognition, arXiv: Computer Vision and Pattern Recognition*, Mar 2020.
- [17] Jongchan Park, Joon-Young Lee, Donggeun Yoo, and InSo Kweon. Distort-and-recover: Color enhancement using deep reinforcement learning. *arXiv: Computer Vision and Pattern Recognition, arXiv: Computer Vision and Pattern Recognition*, Apr 2018.
- [18] Yi Wei, Shaohui Liu, Yongming Rao, Wang Zhao, Jiwen Lu, and Jie Zhou. Nerfing-mvs: Guided optimization of neural radiance fields for indoor multi-view stereo. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct 2021.
- [19] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, Jianmin Jiang, and Hong Kong. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement.
- [20] Zeke Xie, Xindi Yang, Yujie Yang, Qi Sun, Yixiang Jiang, Haoran Wang, Yunfeng Cai, and Mingming Sun. S3im: Stochastic structural similarity and its unreasonable effectiveness for neural fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 18024–18034, 2023.

- [21] Ke Xu, Xin Yang, Baocai Yin, and Rynson W.H. Lau. Learning to restore low-light images via decomposition-and-enhancement. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2020.
- [22] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement.
- [23] Qing Zhang, Ganzhao Yuan, Chunxia Xiao, Lei Zhu, and Wei-Shi Zheng. High-quality exposure correction of underexposed photos. In *Proceedings of the 26th ACM international conference on Multimedia*, Oct 2018.
- [24] Shuaifeng Zhi, Tristan Laidlow, Stefan Leutenegger, and Andrew J. Davison. In-place scene labelling and understanding with implicit scene representation. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct 2021.