

## 金字塔渐进融合低照度图像增强网络

余映,徐超越\*,李淼,何鹏浩,杨昊

(云南大学信息学院,云南昆明 650500)

**摘要:**针对现有低照度图像增强网络对不同尺度特征信息存在感知与表达能力不足的问题,提出金字塔渐进融合低照度图像增强网络模型。网络对图像进行多次下采样操作以组成特征金字塔,通过在特征金字塔的三个不同分支上加入跳跃连接,将不同尺度的特征图进行相互融合。通过精细恢复模块进一步提取精炼信息,将特征图恢复到正常的光照图像。结果表明,该网络模型不但能有效地提升低照度图像的整体亮度,而且能很好地保持图像中的细节信息和清晰的物体边缘轮廓,同时能够有效地抑制图像中的暗部噪声,使增强后的图像整体画面真实自然。

**关键词:**低照度图像增强;深度学习;特征金字塔;多尺度特征;跳跃连接

中图分类号:TP391 文献标志码:A 开放科学(资源服务)标识码(OSID):

文章编号:1001-2486(2024)02-224-14



与作者互动  
听语音  
聊科研

## Pyramid asymptotic fusion low-illumination image enhancement network

YU Ying, XU Chaoyue\*, LI Miao, HE Penghao, YANG Hao

(School of Information Science and Engineering, Yunnan University, Kunming 650500, China)

**Abstract:** Since existing low-illumination image enhancement networks have insufficient ability to perceive and express feature information of different scales, a low-illumination image enhancement network model based on pyramid asymptotic fusion was proposed. The network performed multiple down-sampling operations on the image to form a feature pyramid. It fused the feature maps at different scales by adding skip connections to three different branches of the feature pyramid. Fine recovery module further extracted the refined information, and restored the feature map to a normal light image. Results indicate that, the network model not only effectively enhances the brightness of the overall low-illumination image, but also maintains the detailed information and clear edge contours of the objects in the image. Moreover, it can effectively suppress the dark noise, and make the overall enhanced image realistic and natural.

**Keywords:** low-illumination image enhancement; deep learning; feature pyramid; multi-scale features; skip connection

随着数字成像设备的普及,人们可以方便地获取大量的数字图像,但在拍摄场景光照条件较差或图像采集设备性能受限等情况下,容易产生低照度图像。低照度图像的整体亮度和对比度较低,将影响其实用价值。为提高低照度图像的实用性和清晰度,需要采取低照度图像增强方法进行处理。

传统低照度图像增强算法主要有直方图均衡化的方法、频域的方法和 Retinex 计算理论的方法。直方图均衡化的方法可以分为全局直方图均衡化<sup>[1]</sup>、局部直方图均衡化<sup>[2]</sup>、自适应直方图均衡化<sup>[3]</sup>、限制对比度自适应直方图均衡化<sup>[4]</sup>等,

该类方法容易受到图像中伪影的影响。频域的方法主要利用小波变换<sup>[5]</sup>和傅里叶变换<sup>[6]</sup>来增强处理低照度图像,但该类方法需要人工参与变换参数的选择。Retinex 计算理论的方法<sup>[7-9]</sup>利用颜色恒常知觉计算理论,通过获得反映物体固有属性的反射分量来实现图像增强处理。该类方法可以获得较好的图像增强效果,然而其模型计算复杂度高,且容易产生局部区域过曝现象。此外,余春艳等<sup>[10]</sup>和 Dong 等<sup>[11]</sup>采用去雾方法来实现低照度图像增强,但增强后的图像中物体边缘处容易出现伪影。这些传统算法的图像增强能力有限,模型泛化能力较差。

收稿日期:2022-01-06

基金项目:国家自然科学基金资助项目(62166048,61263048);云南省应用基础研究计划资助项目(2018FB102);云南大学“中青年骨干教师培养计划”资助项目(XT412003)

第一作者:余映(1977—),男,云南大理人,副教授,博士,硕士生导师,E-mail: yuying\_mail@163.com

\*通信作者:徐超越(1997—),女,吉林长春人,硕士研究生,E-mail: xuchao Yue@mail.ynu.edu.cn

近年来,基于深度学习的低照度图像增强方法已成为研究热点。Lore 等<sup>[12]</sup>提出堆叠稀疏去噪编码器对低照度图像进行增强去噪,但其增强的图像细节模糊,模型鲁棒性弱。Li 等<sup>[13]</sup>和 Wei 等<sup>[14]</sup>将 Retinex 理论与卷积神经网络(convolutional neural networks, CNN)相结合,采用监督学习来训练网络参数以实现低照度图像增强,但容易产生边缘轮廓模糊的问题。Lü 等<sup>[15]</sup>提出 MBLLEN 网络,通过多个子网进行增强,以多分支融合策略产生输出图像,但其增强图像整体模糊。Zhang 等<sup>[16]</sup>提出 KIND 方法,具有很好的泛化力,但其图像锐化过度,局部细节不清晰。Jiang 等<sup>[17]</sup>提出生成对抗网络(generative adversarial networks, GAN)的夜间图像增强模型,无须使用图像对进行训练,但其模型鲁棒性仍然较弱。文献[18-20]采用零样本学习方案来增强图像,将低照度图像增强看成曲线估计任务,但其增强的图像整体亮度偏暗。

虽然深度学习的方法是当前图像增强处理研究的主要趋势,但都是对同一尺度的特征图进行特征提取,导致对不同尺度图像特征的表达力不足,图像细节容易丢失,噪声易被放大,边缘轮廓出现模糊。针对这些问题,有研究者将金字塔网络<sup>[21]</sup>和残差学习方式<sup>[22]</sup>引入低照度图像增强网络中,这些方法通过端到端的网络结构将特征信息进行有效的表达和传递。受此启发,本文提出了金字塔渐进融合低照度图像增强网络(pyramid asymptotic fusion low-illumination image enhancement network, PAFNet)。

对于低照度增强任务,局部信息和全局信息都是必不可少的,因为全局信息能够评估整个图像的光照状况,局部特征可以细化图像细节。PAFNet 将特征金字塔的思想应用于低照度图像增强网络中,采用特征金字塔的结构可以合并来自不同层次的多尺度特征,使得网络模型可以包含多个尺度的特征信息,并使所有尺度上的特征图都具有较强的语义信息。PAFNet 还改变现有

深度学习的低照度图像增强网络对不同分辨率特征图的连接方式,采用并行多分支网络与跨层串行相结合的方式设计出金字塔渐进融合模块(pyramid asymptotic fusion module, PAFM)。该模块不仅将不同分辨率的特征图进行并行连接,而且在其基础上加入跳跃连接,把不同分辨率的特征图进行反复融合,在不同网络分支之间保持信息交互,使整个网络能够保持高分辨率的细节和语义信息。其次,还设计了精细恢复模块(fine recovery module, FRM),采用渐进融合的方式进一步提取精炼信息,从而恢复细节更加清晰、边缘更加明朗的正常光照图。此外,该网络模型还加入残差块来降低网络训练的复杂度,以减轻训练过程中梯度消散的问题,从而提高网络模型的稳定性。

## 1 算法模型

### 1.1 网络结构

为有效利用低照度图像局部和全局特征,解决图像在增强过程中容易丢失细节的问题,提出的金字塔渐进融合低照度图像增强网络 PAFNet 的基本结构如图 1 所示,其中 Conv3-32×1 表示卷积核尺寸为 3×3 卷积层,通道数为 32,执行 1 次卷积操作。PAFNet 首先将原始低照度图像通过 3×3 卷积核以及 LReLU(leaky rectified linear units)激活函数得到特征图  $X_i$ ,其大小为  $H \times W$ ,与原始图像相同。之后对特征图  $X_i$  进行卷积核大小为 2×2、步长为 2 的最大池化操作,得到特征图  $Y_i$ (大小为  $H/2 \times W/2$ ),以提取最显著的特征信息。为了进一步提取最显著的特征信息,对特征图进行卷积核大小为 2×2、步长为 2 的最大池化操作,得到特征图  $Z_i$ (大小为  $H/4 \times W/4$ )。需要说明的是, $X_i$ 、 $Y_i$ 、 $Z_i$  是经过卷积、池化等操作得到的浅层特征图,用来组成特征金字塔。通过这种方式可以把低照度图像转换为特征图来提取浅层特征,以使特征图更加稀疏化,有利于更好地挖掘不同层次的特征信息。

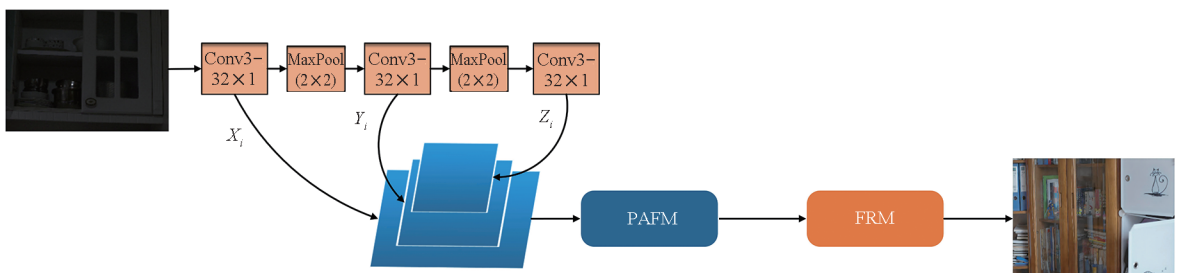


图1 模型网络结构

Fig. 1 Network structure of the model

PAFNet 把取得的浅层特征图送入金字塔渐进融合模块 PAFM 中,进行深层次的特征提取和融合,从而得到含有丰富语义信息和细节信息的特征图。最后把 PAFM 模块中获得的深层特征图输入精细恢复模块 FRM 中进行逐层融合,恢复出正常的光照图像。整个网络的计算流程可以表示为:

$$X_i = \text{Conv}_{3 \times 3}(\text{Input}) \quad (1)$$

$$Y_i = \text{Conv}_{3 \times 3}(\text{down\_sampling}(X_i)) \quad (2)$$

$$Z_i = \text{Conv}_{3 \times 3}(\text{down\_sampling}(Y_i)) \quad (3)$$

$$(X_{i+1}, Y_{i+1}, Z_{i+1}) = \text{PAFM}(X_i, Y_i, Z_i) \quad (4)$$

$$O_{\text{output}} = \text{FRM}(X_{i+1}, Y_{i+1}, Z_{i+1}) \quad (5)$$

其中:  $\text{Input}$  为原始低照度输入图像;  $\text{Conv}_{3 \times 3}$  表示大小为  $3 \times 3$ 、步长为 1 的卷积核;  $\text{down\_sampling}$  表示采用最大池化操作对输入特征图进行尺度为  $1/2$  的降采样;  $X_{i+1}, Y_{i+1}, Z_{i+1}$  为经过金字塔渐进融合模块 PAFM 输出的特征图,其大小分别为  $H \times W, H/2 \times W/2$  和  $H/4 \times W/4$ ;  $O_{\text{output}}$  为经过 FRM 模块逐层融合得到最终的正常光照图。

为了保持深层网络的学习能力,训练更高效的网络模型,PAFNet 还使用残差块来确保特征信息的有效传递。该模块采用跳跃连接,将输入特

征图直接与输出特征图进行相加,然后经过大小为  $3 \times 3$  的卷积层。与普通卷积相比,残差块可以捕捉到更多的图像特征信息,能有效改善网络模型的稳定性。

研究表明,批正则化(batch normalization, BN)可以加速网络模型的训练速度。将 BN 操作应用于残差块 Res\_Block,可以更好地解决网络模型中的梯度消散问题。该处理过程可以表示为

$$R_{\text{out}} = \text{Conv}_{3 \times 3}(\text{Add}[R_{\text{in}}, \text{BN}(\text{Conv}_{3 \times 3}(R_{\text{in}}))]) \quad (6)$$

式中,  $R_{\text{in}}$  是输入残差块的特征图,  $R_{\text{out}}$  是经过残差块输出的特征图, Add 表示逐像素相加。

## 1.2 PAFM 模块

为了提高低照度图像增强网络对不同尺度特征的表达和感知能力,PAFNet 设计了 PAFM 模块,该模块由两个阶段组成,其网络结构如图 2 所示。由于在低照度图像增强任务中挖掘和利用图像的深层特征非常重要,同时考虑到模型复杂度的问题,因此在 PAFM 模块中每个分支都包括 6 个残差块和 5 个卷积层来提取图像的深层次特征。

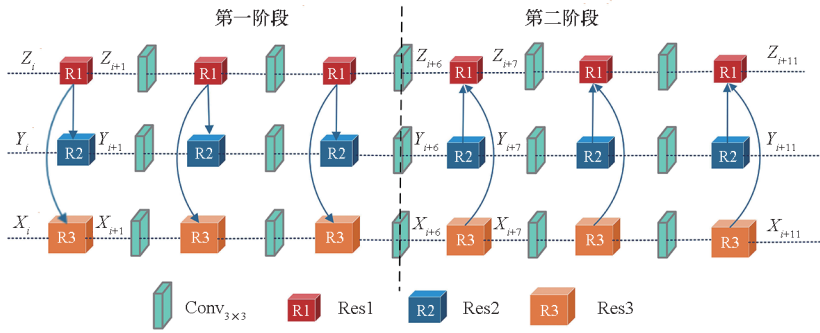


图 2 PAFM 模块结构

Fig.2 PAFM module structure

PAFM 模块有三个不同尺度的分支,其中特征图  $X_i$  所在的分支拥有高分辨率的特征图,它们具有较小的感受野,能获得小区域图像信息;特征图  $Y_i$  所在的分支拥有中分辨率的特征图,可以同时获得图像细节和整体结构信息;特征图  $Z_i$  所在的分支拥有低分辨率的特征图,具有较大的感受野,可以更好地获取图像的整体结构。

PAFM 模块在第一阶段将特征信息从低分辨率分支输入到高分辨率分支中,通过加入从低分辨率到高分辨率的重建过程,使网络获得具有结构信息的高分辨率特征图,从而提高了对图像整体结构的恢复能力。通过第一阶段,网络可以获得亮度适当增强但细节较少的粗糙图像。该阶段的

计算过程可以表示为:

$$X_{i+1} = \text{Cat}[\text{Res3}(X_i), (\text{Conv}_{3 \times 3}(\text{down\_sampling}(Z_i))) \times 2] \quad (7)$$

$$Y_{i+1} = \text{Cat}[\text{Res2}(Y_i), \text{Conv}_{3 \times 3}(\text{down\_sampling}(Z_i))] \quad (8)$$

$$Z_{i+1} = \text{Res1}(Z_i) \quad (9)$$

其中:  $X_i, Y_i, Z_i$  分别为并行的三个具有不同分辨率的特征图,其大小分别为  $H \times W, H/2 \times W/2$  和  $H/4 \times W/4$ ; Res1、Res2、Res3 分别表示输入大小为  $H \times W, H/2 \times W/2$  和  $H/4 \times W/4$  的残差块,它们具有相同的网络结构,但其输入特征图的尺寸不同;  $X_{i+1}, Y_{i+1}, Z_{i+1}$  分别是经过 Res1、Res2、Res3

得到的特征图,其大小分别为  $H \times W$ 、 $H/2 \times W/2$  和  $H/4 \times W/4$ ;  $\text{down\_sampling}$  为最大池化操作,  $\text{Cat}$  表示连接操作,“ $\times 2$ ”表示最大池化和卷积操作执行 2 次。

PAFM 模块在第二阶段将特征信息从高分辨率分支输入低分辨率分支中。经过第一阶段的多层卷积和残差块操作,图像细节信息损失严重。为解决该问题,PAFM 模块将高分辨率特征图加到低分辨率特征图中,使网络模型在恢复图像整体结构的同时,还能不断地找回丢失的细节信息。通过第二阶段,网络可以获得亮度进一步增强、细节信息更加丰富的精细图像。该阶段的计算过程可以表示为:

$$X_{i+7} = \text{Res3}(X_{i+6}) \quad (10)$$

$$Y_{i+7} = \text{Res2}(Y_{i+6}) \quad (11)$$

$$Z_{i+7} = \text{Cat}[\text{Res1}(Z_{i+6}), \text{Conv}_{3 \times 3}(\text{up\_sampling}(Y_{i+6})), \\ (\text{Conv}_{3 \times 3}(\text{up\_sampling}(X_{i+6}))) \times 2] \quad (12)$$

其中: $X_{i+6}$ 、 $Y_{i+6}$ 、 $Z_{i+6}$ 是第二阶段输入的特征图,其大小分别为  $H \times W$ 、 $H/2 \times W/2$  和  $H/4 \times W/4$ ;

$X_{i+7}$ 、 $Y_{i+7}$ 、 $Z_{i+7}$ 分别是经过 Res1、Res2、Res3 模块输出的特征图; $\text{up\_sampling}$  为大小为  $2 \times 2$ 、步长为 2 的转置卷积操作。

### 1.3 FRM 模块

为了提高网络模型的稳定性,让网络能够输出细节和边缘更加清晰的正常光照图像,PAFNet 设计了 FRM 模块,该模块的网络结构如图 3 所示。FRM 模块由 1 个 Res1、2 个 Res2、1 个 Res3 和 6 层卷积(卷积核为  $3 \times 3$ 、步长为 1)组成。首先,它将 PAFM 模块的输出特征图进行逐层融合,将通过 Res1 低分辨率特征图  $Z_{i+11}$  与通过 Res2 的中分辨率特征图  $Y_{i+11}$  进行融合,得到中分辨率特征图  $Y_{i+12}$ (大小为  $H/2 \times W/2$ )。然后,为了提取精炼信息,将通过 Res2 的中分辨率特征图  $Y_{i+12}$  和通过 Res3 的高分辨率特征图  $X_{i+11}$  进行融合,得到与原始图像一样大小的特征图  $N$ 。最后,将高分辨率特征图  $N$  输入 6 层卷积中得到正常的光照图。该模块的计算过程可以用式(13)~(15)表示。

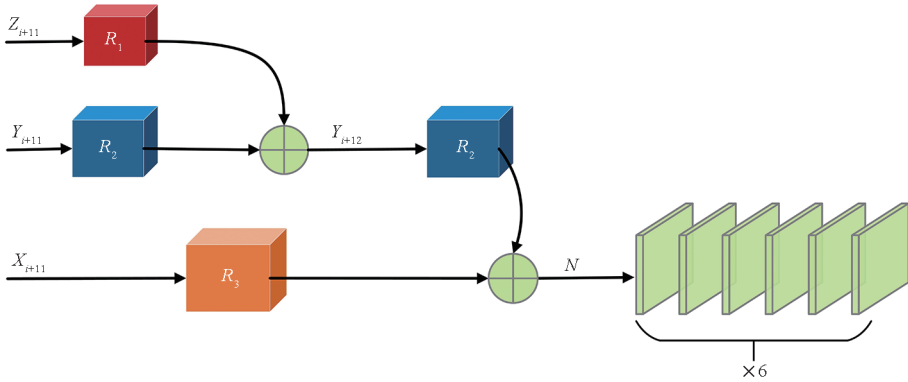


图3 FRM 模块结构

Fig. 3 FRM module structure

$$Y_{i+12} = \text{Cat}[\text{Res1}(Z_{i+11}), \text{Res2}(Y_{i+11})] \quad (13)$$

$$N = \text{Cat}[\text{Res2}(Y_{i+12}), \text{Res3}(X_{i+11})] \quad (14)$$

$$\text{Output} = (\text{Conv}_{3 \times 3}(N)) \times 6 \quad (15)$$

其中,  $\text{Output}$  为最终输出的正常光图像,“ $\times 6$ ”表示  $\text{Conv}_{3 \times 3}$  操作重复执行 6 次。

### 1.4 损失函数

为了让图像增强结果的纹理细节更清晰,亮度分布更合理,将均方误差(mean square error, MSE)、结构性相似度(structural similarity index, SSIM)和梯度损失(gradient loss, GL)结合起来作为图像增强模型的联合训练损失,其计算公式如下:

$$L_{\text{Loss}} = \alpha \times L_{\text{MSE}} + \beta \times L_{\text{SSIM}} + \gamma \times L_{\text{GL}} \quad (16)$$

MSE 损失函数是网络训练时常用的回归损失函数,在训练过程中具有较好的收敛性,其计算

公式为:

$$L_{\text{MSE}} = \frac{1}{ST} \sum_{i=0}^{T-1} \sum_{j=0}^{S-1} [I(i,j) - K(i,j)]^2 \quad (17)$$

式中: $I$  表示原始正常图像,其大小为  $S \times T$ ;  $K$  表示经过网络模型增强后的图像。

为了衡量增强后的低照度图像和正常光照图像之间的结构相似性,在联合训练损失中加入了 SSIM 损失函数,其计算公式为:

$$L_{\text{SSIM}(x,y)} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (18)$$

式中, $\mu_x$  和  $\mu_y$  分别为图像  $x$  和  $y$  的均值, $\sigma_x^2$  和  $\sigma_y^2$  分别是图像  $x$  和  $y$  的方差, $\sigma_{xy}$  表示图像  $x$  和  $y$  的协方差, $c_1$  和  $c_2$  为默认调节参数。

为了避免网络在训练过程中陷入局部最优

解,在联合训练损失中加入了梯度损失函数,其计算公式为:

$$L_{GL} = \frac{1}{2} [(\nabla_x R_{low} - \nabla_x R_{high})^2 + \nabla_y R_{low} - \nabla_y R_{high})^2]$$

(19)

式中: $\nabla$ 是梯度算子,包含图像  $x$  和  $y$  两个维度方向; $\nabla_x$ 和 $\nabla_y$ 分别表示水平和垂直方向的梯度; $R_{high}$ 和 $R_{low}$ 分别代表正常光照图像和低照度图像的像素值。

## 2 实验部分

本节实验将提出的 PAFNet 与现有经典方法进行性能比较和评估,并进行消融实验。基于 LOL 数据集对比了经典的低照度图像增强算法。为了说明实验的有效性,又基于 PASCAL VOC 数据集对比了 2019 至 2021 年最新深度学习方法的低照度图像增强算法。实验平台配置 Intel Xeon W - 2102 2.90 GHz CPU、8 GB RAM 和 Nvidia 2080 GPU,实验程序在 Tensorflow 1.15 框架上运行,联合损失函数的权重参数设置为: $\alpha = \beta = \gamma = 1$ 。

## 2.1 定性分析

### 2.1.1 在 LOL 公开数据集上的测试

本小节实验将提出的 PAFNet 模型与 9 种经典低照度图像增强算法进行比较:MSRCR<sup>[7]</sup>, DONG<sup>[11]</sup>, NPE<sup>[23]</sup>, SRIE<sup>[8]</sup>, MF<sup>[9]</sup>, LIME<sup>[24]</sup>, BIMEF<sup>[25]</sup>, MBLLEN<sup>[15]</sup>, KIND<sup>[16]</sup>。其中,MBLLEN 和 KIND 为近阶段提出的深度学习的低照度图像增强方法,其他 7 种为传统的低照度图像增强方法。在本实验中,网络模型的训练和测试采用 LOL 公开数据集<sup>[14]</sup>,其包含 485 对低光/正常光训练图像和 15 幅低照度测试图像。

图 4 所示为 LOL 数据集中的 3 幅测试图像以及 10 种对比方法得到的低照度图像增强实验结果,第一幅为“游泳馆”图像,第二幅为“衣柜”图像,第三幅为“玩偶”图像。其中 LOW 和 HIGH 分别表示低照度图像和正常光照图像。可以看出,本文模型可以有效地增强低照度图像的亮度,还原图像细节,同时还能有效去除暗部区域中的噪声。

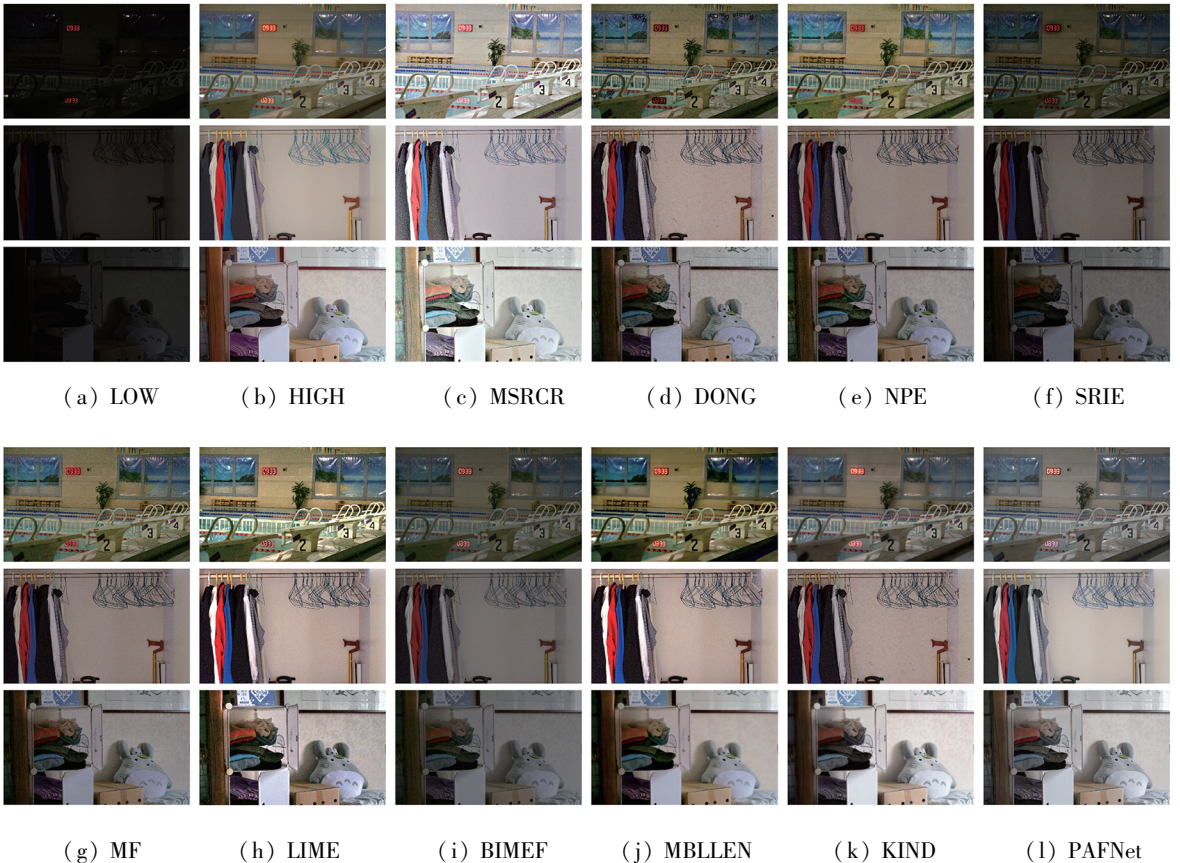


图 4 LOL 数据集实验对比

Fig. 4 Experimental comparison on the LOL dataset

图 5、图 6、图 7 分别给出了“游泳馆”“衣柜”和“玩偶”3 幅测试图像的局部细节对比。

可以看出,DONG、SRIE 和 BIMEF 方法计算得到的增强图像整体亮度仍然偏暗;MSRCR 方法在

其增强图像的局部区域出现了过曝现象; BIMEF、MBLLEN 和 KIND 方法在其增强图像的物体边缘会出现模糊现象; MSRCR、DONG、SRIE、BIMEF、MF、NPE 和 LIME 方法在其增强后的图像中产生了大量噪声。相比之下, PAFNet 方法不但能更好地提升增强图像的整体亮度,而且能防止出现局部区域过曝,其增强后的图像保持了丰富的物体细节信息和清晰的边缘轮廓,同时还能有效抑制暗部噪声,使整体画面比较干净、真实自然。

### 2.1.2 在 PASCAL VOC 公开数据集上的测试

本小节实验将提出的 PAFNet 模型与 5 种深

度学习的低照度图像增强算法进行比较: DRBN<sup>[26]</sup>, DSLR<sup>[27]</sup>, KIND++<sup>[28]</sup>, RRDNet<sup>[19]</sup>, EnlightenGAN<sup>[17]</sup>。其中 DSLR 和 KIND++ 为监督学习方法, DRBN 是半监督学习方法, EnlightenGAN 是无监督学习方法。RRDNet 是采用零样本学习方案。

在本实验中,网络模型的训练和测试采用 PASCAL VOC 公开数据集<sup>[29]</sup>,其包含人工合成的 16 925 幅低照度图像、144 张测试图像。该数据集在合成时对相应普通图像的每个通道进行随机伽马调整,并加入了峰值为 200 的泊松噪声。

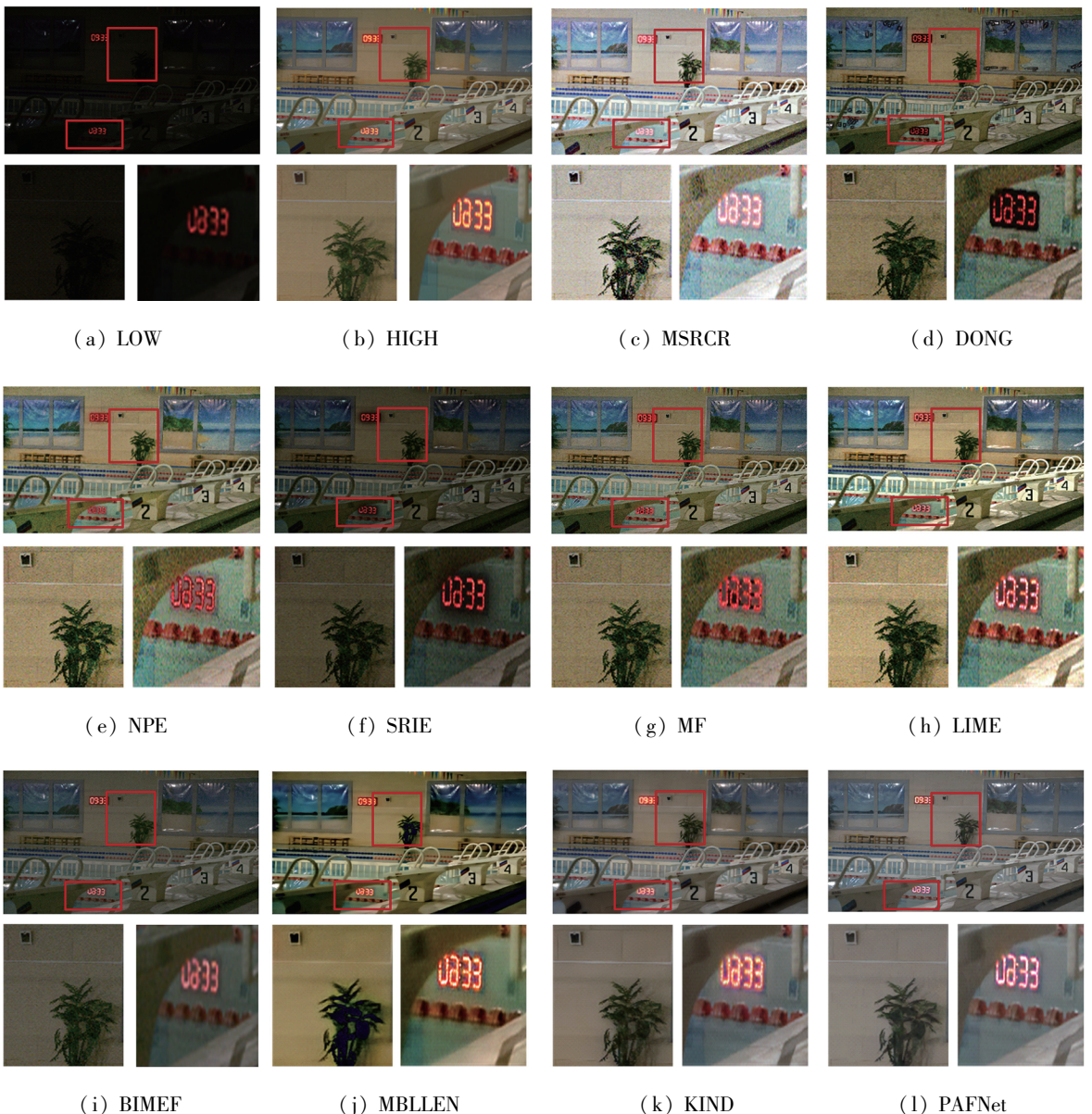


图 5 “游泳馆”图像细节对比

Fig. 5 Detailed comparison of the “natatorium” image



图 6 “衣柜”图像细节对比

Fig. 6 Detailed comparison of the “wardrobe” image

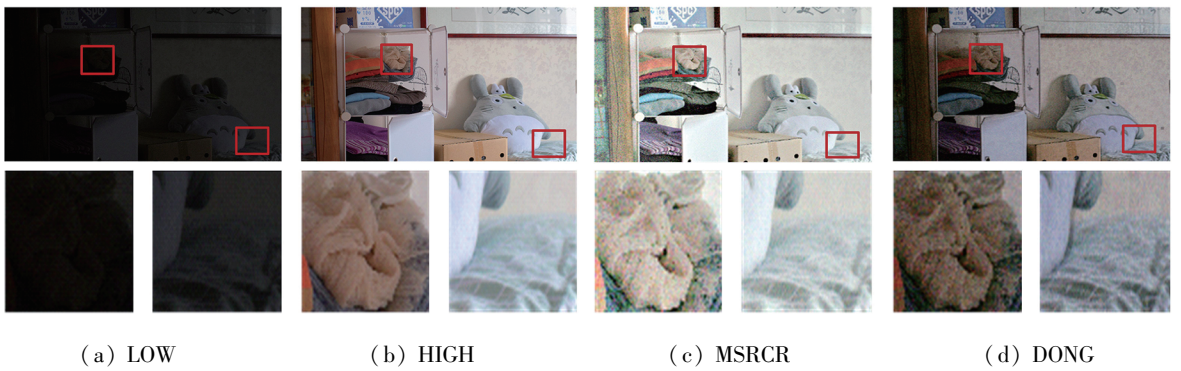




图7 “玩偶”图像细节对比

Fig.7 Detailed comparison of the “doll” image

图8所示为PASCAL VOC数据集中的4幅测试图像以及6种对比方法得到的低照度图像增强实验结果。4幅测试图像分别取名为“人”“瓶子”“墙”“飞机”。可以看出,DRBN算法图像细节不够丰富;DSLRL和RRDNet算法在图像整体亮度提升上不够理想,有些黑暗的地方仍模糊不清;EnlightenGAN算法出现了少量噪声,图像细节轮廓模糊;KIND++算法产生了过曝的现象。相比之下,PAFNet方法不但能更好地提升低照度图像的整体亮度,而且能有效地去除原始低照度图像中的噪声,同时增强后的图像保持了丰富的物体细节和清晰的边缘轮廓,整体画面表现自然真实。

## 2.2 定量分析

本节实验采用图像质量客观评价指标来评估所提出的PAFNet模型。选用的评价指标包括峰值信噪比(peak signal-to-noise ratio, PSNR)、SSIM、信噪比(signal-to-noise ratio, SNR)、视觉信息保真度(visual information fidelity, VIF)、亮度顺序误差(lightness order error, LOE)及信息保真度准则(information fidelity criterion, IFC)。其中MSE值越小则PSNR值和SNR值越大,表明图像的质量越好,失真越少;SSIM值越大,表明图像的质量越好,更符合人眼的评判标准;VIF和IFC值越大则说明图像质量越好,细节保留得越好;LOE值越小则待测图像越接近真实图像。

表1、表2、表3分别给出了对应LOL数据集中3幅测试图像的各种对比方法所得到的客观评价指标(表中加粗字体为每列最优值,后文同)。可以看出,在表1中除了PSNR、SNR指标略低于KIND外,PAFNet模型在其他评价指标上都是最优的,而且在IFC、LOE、运行时间三个评价指标上明显优于其他对比方法。在表2和表3中,除了LOE指标外,本文PAFNet模型在PSNR、SSIM、SNR、VIF、IFC、运行时间指标上明显优于其他对比方法。另外,在表4中还给出了LOL数据集中全部15幅测试图像所对应各种方法的平均指标。可以看出,PAFNet模型在整个数据集上的平均指标也明显优于其他所有对比方法。由此可见,本文该研究方法对低照度图像增强的效果在客观评价指标上明显优于其他9种对比方法。

表5给出了对应PASCAL VOC数据集中的4幅测试图像的各种基于深度学习的低照度图像增强方法所得到的客观评价指标。可以看出,对于“人”“瓶子”“飞机”三幅测试图像,PAFNet在所有6种客观评价指标上,都明显优于其他5种深度学习的方法。对于“墙”测试图像,除了LOE指标外,PAFNet在PSNR、SSIM、SNR、VIF、IFC指标上明显优于其他对比方法。由此可见,本文研究方法对各种场景类型的图像都能进行有效的低照度增强处理,在客观评价指标上优于现有的深度学习方法。



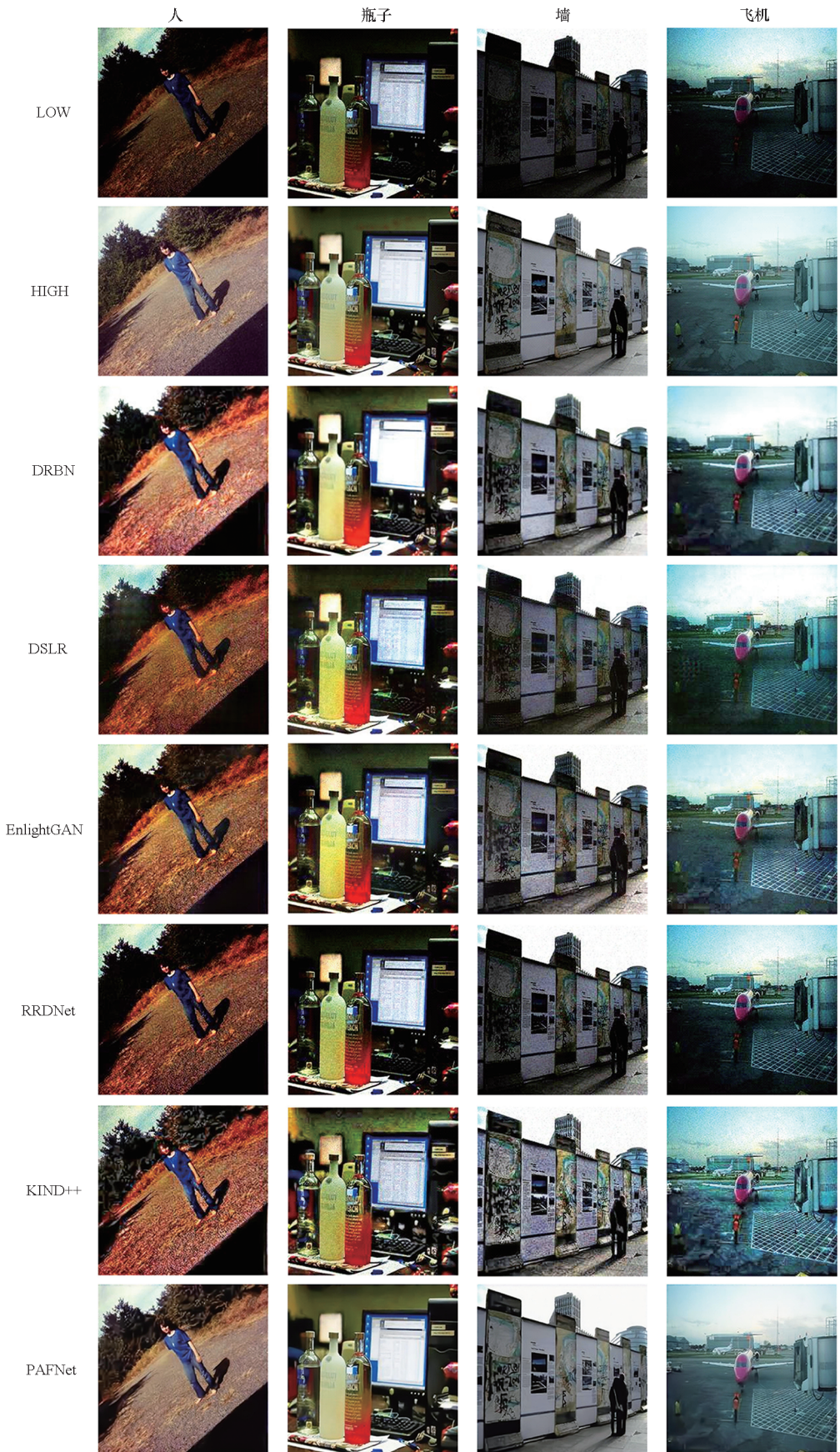


图 8 PASCAL VOC 数据集实验对比

Fig. 8 Experimental comparison on the PASCAL VOC dataset

表1 “游泳馆”图像评价指标对比

Tab.1 Comparison of the evaluation metrics of the “natatorium” image

方法	PSNR/dB	SSIM	SNR/dB	VIF	IFC	LOE	运行时间/s
MSRCR	10.51	0.40	2.62	0.55	3.09	884.43	0.653
DONG	20.60	0.49	13.39	0.45	2.40	297.31	0.199
NPE	19.20	0.46	11.69	0.56	2.98	931.63	7.815
SRIE	14.85	0.57	6.51	0.57	3.43	263.38	7.747
MF	23.16	0.52	17.10	0.52	2.78	260.71	0.439
LIME	15.99	0.42	7.82	0.63	3.28	495.80	0.454
BIMEF	18.86	0.70	10.70	0.59	3.71	265.63	0.225
MBLLEN	20.76	0.75	13.41	0.57	3.23	238.79	0.095
KIND	<b>26.81</b>	<b>0.90</b>	<b>21.16</b>	0.57	3.59	357.07	0.054
PAFNet	24.78	<b>0.90</b>	19.98	<b>0.64</b>	<b>4.12</b>	<b>231.83</b>	<b>0.016</b>

表2 “衣柜”图像评价指标对比

Tab.2 Comparison of the evaluation metrics of the “wardrobe” image

方法	PSNR/dB	SSIM	SNR/dB	VIF	IFC	LOE	运行时间/s
MSRCR	16.99	0.69	13.16	0.56	2.41	1 633.50	1.299
DONG	20.93	0.60	16.47	0.51	2.60	529.27	0.267
NPE	16.40	0.65	11.60	0.56	2.39	732.92	8.159
SRIE	11.14	0.62	6.27	0.52	2.48	478.75	9.799
MF	16.10	0.65	11.25	0.56	2.38	697.25	1.844
LIME	20.79	0.58	16.90	0.60	2.44	799.80	0.851
BIMEF	13.01	0.72	8.14	0.53	2.64	570.04	1.205
MBLLEN	23.09	0.82	19.25	0.52	2.38	<b>385.17</b>	0.074
KIND	23.30	0.88	20.24	0.52	2.46	647.45	0.098
PAFNet	<b>24.24</b>	<b>0.90</b>	<b>21.66</b>	<b>0.57</b>	<b>2.69</b>	486.11	<b>0.019</b>

表3 “玩偶”图像评价指标对比

Tab.3 Comparison of the evaluation metrics of the “doll” image

方法	PSNR/dB	SSIM	SNR/dB	VIF	IFC	LOE	运行时间/s
MSRCR	12.34	0.57	7.65	0.51	2.75	228.00	3.553
DONG	17.09	0.54	12.91	0.45	2.21	304.04	0.376
NPE	15.76	0.57	11.46	0.54	2.77	384.03	8.028
SRIE	9.93	0.53	5.31	0.49	2.88	282.46	8.056
MF	14.78	0.57	10.36	0.48	2.47	424.75	2.485
LIME	19.38	0.51	5.78	0.23	1.16	1 251.30	0.384
BIMEF	11.49	0.63	6.90	0.47	2.96	343.70	1.397
MBLLEN	20.81	0.78	17.08	0.43	2.47	<b>213.12</b>	0.089
KIND	20.66	0.88	16.42	0.52	3.12	279.50	0.089
PAFNet	<b>27.69</b>	<b>0.90</b>	<b>28.22</b>	<b>0.57</b>	<b>3.45</b>	221.86	<b>0.025</b>

表 4 整个 LOL 数据集的图像评价指标对比

Tab. 4 Comparison of the image evaluation metrics of the whole LOL dataset

方法	AVG_PSNR/dB	AVG_SSIM	AVG_SNR/dB	AVG_VIF	AVG_IFC	AVG_LOE
MSRCR	12.67	0.45	7.14	0.51	2.41	1 359.85
DONG	16.72	0.47	10.84	0.43	2.04	407.67
NPE	16.97	0.47	11.57	0.49	2.35	686.21
SRIE	11.92	0.49	5.67	0.47	2.53	367.67
MF	16.97	0.50	11.07	0.46	2.08	427.38
LIME	15.64	0.44	10.03	0.47	2.17	750.37
BIMEF	13.89	0.59	8.36	0.47	2.60	382.43
MBLLEN	17.90	0.79	13.75	0.45	2.45	385.16
KIND	17.65	0.80	13.74	0.48	2.71	436.02
PAFNet	<b>20.55</b>	<b>0.84</b>	<b>14.75</b>	<b>0.53</b>	<b>2.80</b>	<b>332.94</b>

表 5 PASCAL VOC 数据集评价指标对比

Tab. 5 Comparison of the evaluation metrics of the PASCAL VOC dataset

图像名称	方法	PSNR/dB	SSIM	SNR/dB	VIF	IFC	LOE
人	DRBN	20.64	0.66	14.16	0.40	3.04	341.67
	DSLR	14.80	0.59	8.32	0.34	2.66	175.48
	EnlightGAN	16.23	0.61	9.75	0.40	3.01	280.25
	RRDNet	13.53	0.45	7.01	0.38	2.95	179.27
	KIND + +	17.15	0.57	10.66	0.38	2.80	451.16
	PAFNet	<b>24.91</b>	<b>0.84</b>	<b>18.43</b>	<b>0.51</b>	<b>4.17</b>	<b>104.70</b>
瓶子	DRBN	19.88	0.80	13.15	0.37	2.62	910.80
	DSLR	22.54	0.74	15.81	0.37	2.59	290.45
	EnlightGAN	24.10	0.79	17.37	0.42	3.02	341.71
	RRDNet	19.47	0.60	12.74	0.41	2.88	160.59
	KIND + +	19.66	0.70	12.94	0.36	2.54	325.73
	PAFNet	<b>24.26</b>	<b>0.90</b>	<b>17.54</b>	<b>0.56</b>	<b>4.25</b>	<b>74.31</b>
墙	DRBN	20.61	0.80	15.29	0.38	2.86	365.82
	DSLR	17.97	0.71	12.65	0.33	2.53	175.48
	EnlightGAN	20.59	0.73	15.26	0.40	2.96	365.11
	RRDNet	16.43	0.60	11.10	0.38	2.83	<b>113.37</b>
	KIND + +	20.30	0.64	14.98	0.36	2.61	448.61
	PAFNet	<b>29.50</b>	<b>0.91</b>	<b>24.18</b>	<b>0.55</b>	<b>4.36</b>	139.02
飞机	DRBN	22.00	0.84	17.15	0.43	2.28	344.91
	DSLR	20.30	0.79	15.44	0.40	2.10	729.30
	EnlightGAN	7.51	0.11	2.65	0.02	0.08	305.30
	RRDNet	16.79	0.56	11.93	0.48	2.47	212.59
	KIND + +	20.48	0.61	15.63	0.44	2.19	314.78
	PAFNet	<b>33.51</b>	<b>0.94</b>	<b>28.66</b>	<b>0.58</b>	<b>3.34</b>	<b>98.54</b>

### 2.3 消融实验

本节对提出的 PAFNet 网络模型中各个网络模块进行消融实验,本实验采用 LOL 数据集。为公平比较,实验均在相同设置下进行训练,采用 PSNR、SSIM 和 MSE 三个指标来定量评估模型中各个模块的有效性。

表 6 给出了采用 LOL 数据集中的“玩偶”图像来进行消融实验所得到的客观评价指标。另外,AVG\_PSNR 是采用 LOL 数据集中所有 15 幅测试图像所得到的平均峰值信噪比。为了验证联合损失函数中所采用的 3 个损失函数的有效性,每次训练分别移除其中一个损失函数来进行消融实验。从表 6 中可以看出,移除梯度损失、SSIM 损失、MSE 损失当中的任何一个损失函数后,网络模型所得到的低照度图像增强效果的客观评价指标都会有明显劣化。特别是移除梯度损失函数会对结果产生较大的影响。若

采用全部三种损失作为联合损失函数对网络模型进行训练,会得到最好的实验结果。

为了验证模型中精细恢复模块 FRM 和残差块 Res\_Block 的有效性,每次训练分别移除其中一个网络模块来进行消融实验,如表 6 所示。从表中可以看出,加入 Res\_Block 模块和 FRM 模块能够有效提升低照度图像的增强效果。加入 FRM 模块后,AVG\_PSNR 指标提升了 3.16%;而加入 Res\_Block 模块后,AVG\_PSNR 指标提升了 6.80%。为了直观表示消融实验中评价指标的变化情况,将消融实验的各个数据绘制成散点图和折线图,如图 9 和图 10 所示。其中,图 9 所示为各消融实验在迭代 2 000 次后所得到的平均 PSNR 值。可以看出,PAFNet 的基础模型能取得最好的结果。图 10 所示为各消融实验每隔 200 次迭代所得到的平均 PSNR 值及其变化情况。由此可见 PAFNet 的基础模型在充分训练以后能逐步达到最好的实验结果。

表 6 关于损失函数和网络结构的消融实验

Tab. 6 Ablation experiments on the loss function and the network structure

序号	消融实验模型	玩偶			AVG_PSNR/dB	AVG_SSIM
		PSNR/dB	SSIM	MSE		
1	基础模型	<b>27.69</b>	<b>0.90</b>	<b>110.73</b>	<b>20.55</b>	<b>0.84</b>
2	无梯度损失	20.62	0.87	564.30	19.24	0.82
3	无 SSIM 损失	24.48	0.89	242.93	20.05	0.83
4	无 MSE 损失	25.15	0.88	198.72	20.35	0.83
5	无 Res_Block 模块	23.21	0.88	310.84	19.24	0.82
6	无 FRM 模块	24.65	0.89	222.63	19.92	0.83

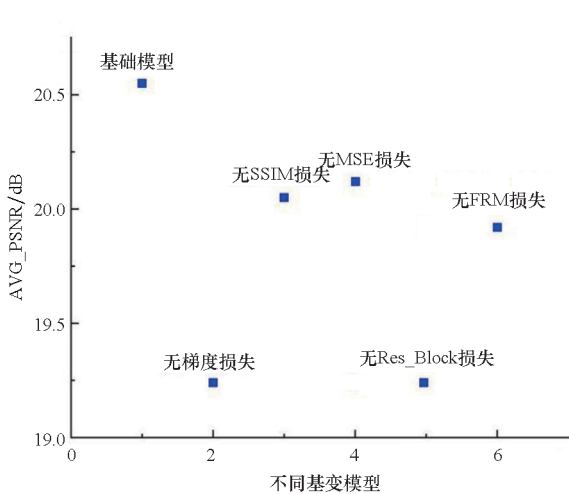


图 9 消融实验散点图

Fig. 9 Scatter diagram of the ablation experiment

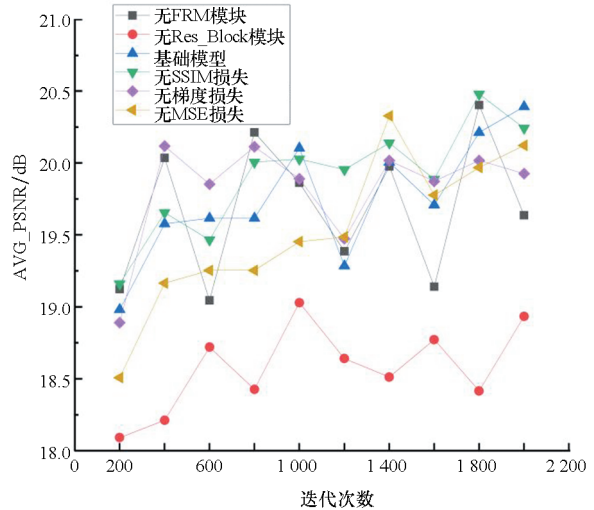


图 10 消融实验折线图

Fig. 10 Broken line diagram of the ablation experiment

PAFM 模块是 PAFNet 的核心,为了分析 PAFM 模块采用不同上下采样数量(即模型深度)对网络性能的影响,表 7 给出了在上下采样数量分别为 2、3、4、5 层时,网络模型所得到的 AVG\_PSNR、AVG\_SSIM、AVG\_MSE 指标。可以看出,当上下采样数量为 3 层时,网络模型能取得较为理想的结果。虽然 5 层的效果也很好,但其网络参数量及训练所消耗的时间会急剧增加,计算和存储成本也随之增加。因此,算法模型最终将金字塔渐进融合模块 PAFM 的上下采样数量设定为 3 层网络。

表 7 关于上下采样数量的消融实验

Tab.7 Ablation experiments on the number of upper and lower sampling

上下采样数量	AVG_PSNR/ dB	AVG_SSIM	AVG_MSE
上下采样个数 × 2 (14)	19.665	0.828	1183.251
上下采样个数 × 3 (16)	20.550	<b>0.839</b>	<b>956.133</b>
上下采样个数 × 4 (18)	20.286	0.834	1 066.069
上下采样个数 × 5 (20)	<b>21.701</b>	<b>0.839</b>	1 206.741

### 3 结论

本文提出了基于金字塔渐进融合的低照度图像增强网络模型。在网络训练过程中采用渐进融合的策略,对网络进行端到端的有监督学习。由于使用跳跃连接将高、低分辨率的特征图进行反复融合,所以金字塔结构的各网络层可以学习相邻层的图像特征信息,从而提高了网络对不同尺度特征的表达能力,而且使整个网络保持高分辨率的细节信息和语义信息。为了验证所提出网络模型的有效性,从主观视觉和客观性能指标角度评价了 PAFNet 模型对低照度图像的增强处理效果。实验结果表明,与现有方法相比,PAFNet 模型具有更优的性能。在增强图像亮度的同时,能保持更多的图像结构和纹理信息,而且可以有效地去除图像黑暗区域中的噪声。

基于研究提出的低光照图像增强网络模型,未来的研究工作将会涉及以下两个方面的应用:一方面,将该网络模型应用于云南壁画文化遗产

的数字化保护研究,针对在野外低光照情况下采集到的古代壁画图像,需要尽可能地提升壁画图像的亮度,同时还需要更真实地还原壁画图像的细节信息,以便后续的破损壁画数字化修复研究;另一方面,将进一步研究能应用于视频监控系统的低照度视频增强方法,以提高夜间或低光照情况下的实时视频监控能力。

### 参考文献 (References)

- [1] CHENG H D, SHI X J. A simple and effective histogram equalization approach to image enhancement [J]. Digital Signal Processing, 2004, 14(2): 158 - 170.
- [2] KIM J Y, KIM L S, HWANG S H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(4): 475 - 484.
- [3] PIZER S M, AMBURN E P, AUSTIN J D, et al. Adaptive histogram equalization and its variations [J]. Computer Vision, Graphics, and Image Processing, 1987, 39(3): 355 - 368.
- [4] REZA A M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement [J]. Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, 2004, 38(1): 35 - 44.
- [5] WANG X H, ISTEPANIAN R S H, SONG Y H. Microarray image enhancement by denoising using stationary wavelet transform [J]. IEEE Transactions on NanoBioscience, 2003, 2(4): 184 - 189.
- [6] AGAIAN S S, PANETTA K, GRIGORYAN A M. Transform-based image enhancement algorithms with performance measure [J]. IEEE Transactions on Image Processing, 2001, 10(3): 367 - 382.
- [7] RAHMAN Z, JOBSON D J, WOODDELL G A. Multi-scale Retinex for color image enhancement [C] // Proceedings of the 3rd IEEE International Conference on Image Processing, 1996: 1003 - 1006.
- [8] FU X Y, ZENG D L, HUANG Y, et al. A weighted variational model for simultaneous reflectance and illumination estimation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 2782 - 2790.
- [9] FU X Y, ZENG D L, HUANG Y, et al. A fusion-based enhancing method for weakly illuminated images [J]. Signal Processing, 2016, 129: 82 - 96.
- [10] 余春艳, 徐小丹, 林晖翔, 等. 应用雾天退化模型的低照度图像增强 [J]. 中国图象图形学报, 2017, 22(9): 1194 - 1205.
- [11] YU C Y, XU X D, LIN H X, et al. Low-illumination image enhancement method based on a fog-degraded model [J]. Journal of Image and Graphics, 2017, 22(9): 1194 - 1205. (in Chinese)
- [12] DONG X, WANG G, PANG Y, et al. Fast efficient algorithm for enhancement of low lighting video [C] // Proceedings of the IEEE International Conference on Multimedia and Expo, 2011: 1 - 6.
- [12] LORE K G, AKINTAYO A, SARKAR S. LLNet: a deep

- autoencoder approach to natural low-light image enhancement[J]. *Pattern Recognition*, 2017, 61: 650 – 662.
- [13] LI C Y, GUO C L, LOY C C. Learning to enhance low-light image via zero-reference deep curve estimation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(8): 4225 – 4238.
- [14] WEI C, WANG W J, YANG W H, et al. Deep Retinex decomposition for low-light enhancement[EB/OL]. [2021 – 08 – 20]. <https://arxiv.org/pdf/1808.04560.pdf>.
- [15] LYU F F, LU F, WU J H, et al. MBLLN: low-light image/video enhancement using CNNs [C]//*Proceedings of BMVC*, 2018.
- [16] ZHANG Y H, ZHANG J W, GUO X J. Kindling the darkness: a practical low-light image enhancer [C]//*Proceedings of the 27th ACM International Conference on Multimedia*, 2019: 1632 – 1640.
- [17] JIANG Y F, GONG X Y, LIU D, et al. EnlightenGAN: deep light enhancement without paired supervision [J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 2021, 30: 2340 – 2349.
- [18] ZHANG L, ZHANG L J, LIU X, et al. Zero-shot restoration of back-lit images using deep internal learning [C]//*Proceedings of the 27th ACM International Conference on Multimedia*, 2019: 1623 – 1631.
- [19] ZHU A Q, ZHANG L, SHEN Y, et al. Zero-shot restoration of underexposed images via robust Retinex decomposition[C]//*Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, 2020: 1 – 6.
- [20] LI C Y, GUO C L, HAN L H, et al. Lighting the darkness in the deep learning era[J/OL]. *Computer Vision and Pattern Recognition*, 2021[2022 – 01 – 01]. <https://arxiv.org/pdf/2104.10729v1.pdf>.
- [21] LI J Q, LI J C, FANG F M, et al. Luminance-aware pyramid network for low-light image enhancement [J]. *IEEE Transactions on Multimedia*, 1809, 23: 3153 – 3165.
- [22] WANG L W, LIU Z S, SIU W C, et al. Lightening network for low-light image enhancement [J]. *IEEE Transactions on Image Processing*, 2020, 29: 7984 – 7996.
- [23] WANG S H, ZHENG J, HU H M, et al. Naturalness preserved enhancement algorithm for non-uniform illumination images [J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3538 – 3548.
- [24] GUO X J, LI Y, LING H B. LIME: low-light image enhancement via illumination map estimation [J]. *IEEE Transactions on Image Processing*, 2017, 26(2): 982 – 993.
- [25] YING Z Q, LI G, GAO W. A bio-inspired multi-exposure fusion framework for low-light image enhancement [J]. *Journal of Latex Class Files*, 2015, 14(8): 1 – 10.
- [26] YANG W H, WANG S Q, FANG Y M, et al. From fidelity to perceptual quality: a semi-supervised approach for low-light image enhancement [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020: 3060 – 3069.
- [27] LIM S, KIM W. DSLR: deep stacked Laplacian restorer for low-light image enhancement [J]. *IEEE Transactions on Multimedia*, 2021, 23: 4272 – 4284.
- [28] ZHANG Y H, GUO X J, MA J Y, et al. Beyond brightening low-light images [J]. *International Journal of Computer Vision*, 2021, 129(4): 1013 – 1037.
- [29] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge [J]. *International Journal of Computer Vision*, 2010, 88(2): 303 – 338.

(编辑:熊立桃,杨琴)