计算机系统应用 ISSN 1003-3254, CODEN CSAOBN Computer Systems & Applications, 2024, 33(4):194-201 [doi: 10.15888/j.cnki.csa.009481] ©中国科学院软件研究所版权所有. E-mail: csa@iscas.ac.cn http://www.c-s-a.org.cn Tel: +86-10-62661041

# 复合主干融合的图像增强算法①

万嘉龙<sup>1</sup>, 况立群<sup>1,2,3</sup>, 曹亚明<sup>1,2,3</sup>, 郭 磊<sup>1,2,3</sup>, 熊风光<sup>1,2,3</sup>

<sup>1</sup>(中北大学 计算机科学与技术学院,太原 030051) <sup>2</sup>(机器视觉与虚拟现实山西省重点实验室,太原 030051) <sup>3</sup>(山西省视觉信息处理及智能机器人工程研究中心,太原 030051) 通信作者:况立群, E-mail: kuang@nuc.edu.cn



**摘** 要: 基于深度学习的微光图像增强算法所生成的图像普遍存在噪声凸显和细节丢失等问题, 而端对端深度学习 算法的性能又在很大程度上依赖于骨干网络的提取能力, 因此, 通过探索更有效的骨干网络结构可以提升微光增强 任务的性能收益. 本文提出了一种复合主干网络融合策略的图像增强算法, 将不同图像增强算法中的主干网络进行 融合, 以提高整体网络的特征提取能力. 该算法通过逐层融合来自不同主干网络的特征信息, 将复合特征引导到解 码器中, 再充分利用不同的上采样方法, 将主干网络融合的特征进行堆叠, 最终生成正常光照条件下的图像. 通过与 现有的主流算法进行定量与定性的对比实验, 结果显示, 本文方法显著提升了微光图像的亮度, 同时保留图像的细 节特征, 在峰值信噪比和结构相似性客观指标上, 在 LOL-V2 数据集上达到了 24.35 dB 和 0.871, 有效解决了图像 增强后的噪声凸显和细节丢失问题.

关键词:网络融合;图像增强;微光;复合主干

引用格式: 万嘉龙,况立群,曹亚明,郭磊,熊风光.复合主干融合的图像增强算法.计算机系统应用,2024,33(4):194-201. http://www.c-sa.org.cn/1003-3254/9481.html

## Image Enhancement Algorithm for Composite Backbone Fusion

WAN Jia-Long<sup>1</sup>, KUANG Li-Qun<sup>1,2,3</sup>, CAO Ya-Ming<sup>1,2,3</sup>, GUO Lei<sup>1,2,3</sup>, XIONG Feng-Guang<sup>1,2,3</sup>

<sup>1</sup>(School of Computer Science and Technology, North University of China, Taiyuan 030051, China) <sup>2</sup>(Shanxi Key Laboratory of Machine Vision and Virtual Reality, Taiyuan 030051, China) <sup>3</sup>(Shanxi Province's Vision Information Processing and Intelligent Robot Engineering Research Center, Taiyuan 030051, China)

Abstract: The images generated by low-light image enhancement algorithms based on deep learning generally have problems such as noise highlighting and detail loss. However, the performance of end-to-end deep learning algorithms largely depends on the extraction ability of the backbone network. Therefore, exploring more effective backbone network structures can improve the performance benefits of low-light enhancement tasks. This study proposes an image enhancement algorithm based on a composite backbone network fusion strategy, which integrates backbone networks from different image enhancement algorithms to improve the overall network's feature extraction ability. The algorithm integrates feature information from different backbone networks layer by layer and guides composite features into the decoder. It then fully utilizes different upsampling methods to stack the fused features of the backbone network, ultimately generating images under normal lighting conditions. Through quantitative and qualitative comparative experiments with existing mainstream algorithms, the results show that our method significantly improves the brightness of low-light images while preserving the detailed features of the images. In terms of objective indicators such as peak

① 基金项目:国家自然科学基金 (62272426, 62106238); 山西省科技重大专项计划"揭榜挂帅"项目 (202201150401021); 山西省科技成果转化引导专项 (202104021301055); 山西省基础研究计划 (202203021222027)

收稿时间: 2023-10-27; 修改时间: 2023-12-04; 采用时间: 2023-12-20; csa 在线出版时间: 2024-03-04 CNKI 网络首发时间: 2024-03-08

<sup>194</sup> 软件技术•算法 Software Technique•Algorithm

signal-to-noise ratio and structural similarity, it achieves 24.35 dB and 0.871 in the LOL-V2 dataset, effectively solving the problems of noise highlighting and detail loss after image enhancement.

Key words: network fusion; image enhancement; low-light; composite backbone

人类视觉系统在光线不足时通常难以有效地识别 周围环境,这对夜间行动、安全监控以及军事作战等 多个领域都带来了重大挑战. 微光增强技术能够有效 克服该问题,它通过捕获微弱的光信号,经过一系列数 字图像处理步骤,包括信号增强等,将原本几乎不可见 的场景变得清晰可辨,为各行业提供了高效的视觉支 持. 图像增强算法主要分为传统图像处理方法和基于 深度学习的方法.

传统方法可以细分为基于直方图均衡化的方法、 基于伽马校正的方法和基于 Retinex<sup>[1]</sup>的方法.基于直 方图均衡化的方法对图像的对比度提升效果较好,但 容易降低图像的灰度级,导致图像过度增强;伽马矫正 的方法可以有效增强图像较暗的部分,但会导致较亮 区域的细节丢失;而基于 Retinex 的方法增强虽然可以 有效增强微光图像,但是生成后的图像色彩失真且伴 随较大的图像噪声.

随着深度学习方法在图像领域的发展,研究人员 尝试将深度学习技术应用到图像增强上,一定程度上 解决了计算复杂度高以及泛化性能不佳等问题.基于 Retinex 思想的 RetinexNet<sup>[2]</sup>、KIND<sup>[3]</sup> 和 KIND++<sup>[4]</sup>等 方法通过对图像的反射分量和光照分量改善图像效果, 其中 RetinexNet 仅将图像分解为光照分量与反射分量, 由此得到的图像缺少图像图层之间的联系,恢复效果 不够理想,所以在 KIND 中增加了图层分解的相关信 息,KIND++在此基础上进一步增加了多尺度注意力模 块,有效解决了一些视觉缺陷的问题.运用参数估计的 ZERO-DCE<sup>[5]</sup>、ZERO-DCE++<sup>[6]</sup>等方法通过学习图像 亮度的变化曲线来完成图像增强任务,然而图片中较 亮的地方容易形成亮斑.结合 LUT 的 3D-LUT<sup>[7]</sup>等方 法通过学习查找表中的 LUT 映射关系来进行图像增 强,虽然提升了图像中的亮度但同时也放大了图像中 的噪声. 使用 GAN 网络的 EnglishGAN<sup>[8]</sup>、DBEN<sup>[9]</sup>等 方法通过不断学习微光与正常光之间的映射关系来提 高图像清晰度,但是可能产生不合理的图像光照强度. 融合 Transformer 的 Uformer<sup>[10]</sup>、IPT<sup>[11]</sup>、RCT<sup>[12]</sup>、 MAXIM<sup>[13]</sup>及 UHDFour<sup>[14]</sup>等方法通过加深图像中色素

之间空间关系保留图像的结构特征. 然而, 深度学习方 法中的单一的图像增强结构都有各自侧重的图像恢复 能力, 无法兼顾生成图像的细节与结构信息, 生成的图 像往往伴随大量的噪声并且生成的图像色彩偏暗. 为 了解决该问题, 本文提出一种复合主干融合的图像增 强方法, 通过逐层融合不同主干网络模块, 并以并行结 构的方式共同作用. 相较于简单的网络深化和拓宽策 略, 通过设计的复合主干网络融合结构, 将不同网络骨 干交叉融合, 充分发挥各自网络结构的优势, 从而获得 更强大的特征提取能力. 该网络结构可以有效改善图 像中的噪音和细节丢失问题, 进一步提高了微光图像 生成的质量. 本文主要贡献是提出了 IEM (inception enhancement module) 模块并使用 IEM 模块与简单 CNN 构建新的主干网络并验证了其有效性.

## 1 网络整体结构

本文网络整体结构主要由编码器和解码器两个关 键部分组成,如图 1 所示. 微光图像在进入编码器之前, 为增加特征的丰富性,采用卷积神经网络 (CNN) 对输 入特征进行维度增加. 然后,将特征送入编码器中提取 融合特征. 编码器由 3 层双分支结构组成,初始层采用 双 CNN 网络,其余两层采用 CNN 结构与 IEM 模块共 同作用. 为了保持融合特征的原始特性,通过 Concat 操作将输出特征堆叠在一起,并输入到解码器中. 解码 器利用不同的解码结构对不同融合层的特征进行解码, 然后通过 Concat 操作堆叠在一起,最后使用卷积操作 将特征转换为增强后的图像. 本文网络结构的创新体 现在以下两个方面.

(1)提出了一种复合主干网络融合的方法,将 CNN 网络与本文引入的 IEM 模块相结合,并汲取了其他分 支融合策略的经验,更有效地实现了对目标图像的增强.

(2) 在分支网络结构中,构建了一种 IEM 的新模 块,对 IAT (illumination adaptive Transformer)的 串行结构进行了改进,提出了更加有效的基于并行结 构的特征提取策略,对多个尺度卷积形成的图像进行 融合.



图1 复合主干融合网络的图像增强网络结构图

#### 1.1 编码器结构

图像增强的目标是将在不良光照下获得的图像 L 映射到正常光照下的图像 N<sub>i</sub>.本文通过逆映射函数 F,将 RGB 图像 L 映射到相应的 Raw-RGB 空间 F(L) 上. 然后,通过学习图像像素的映射关系 S,来动态调整 图像的亮度,生成最终的预测图像 S(F(L)),使其尽量逼 近正常的光照图像 N<sub>i</sub>.具体公式如下:

$$N_i = S(F(L)) \tag{1}$$

为了提高网络特征的提取效率,本文采用了复合 主干融合网络来学习图像像素映射关系 S. 在网络的每 一层中,提取图像特征的独立表达公式如下:

$$MID_n = C_1^n X_1 + C_1 C_2^{n-1} X_2 \tag{2}$$

其中, *C*<sub>1</sub> 为 CNN 操作, *C*<sub>2</sub> 为 IEM 操作, *n* 为特征提取的层数, *MID*<sub>n</sub>是每层提取的融合特征.

复合主干融合网络通过在编码器的不同层次和顶 层之间采用不同的特征提取方式,增加层次之间的特 征多样性.同时,为了最大程度地减少网络模型的参数 量,使用了大量简单的 CNN 结构作为基本特征提取单 元.复合主干融合的网络模型分为3 层特征提取,第1 层特征提取由两个 CNN 结构分支组成,初步学习输入 和输出之间的映射关系.为了防止信息损失,在 CNN 中采用 LeakyReLU 函数提取早期特征;第2 层和第 3 层特征提取采用 CNN 与信息融合模块 (IEM) 结合 的方式,增强解码器在特征提取方面的表现能力.复合 主干融合的分层提取方式能更好地帮助网络捕获和表

196 软件技术•算法 Software Technique•Algorithm

达特征,提高其在图像处理任务中的性能.

在分支网络的设计方面,为了更全面地满足色彩 处理的需求, 秉承了 IAT<sup>[15]</sup>中 PEM (pixel-wise enhancement module) 模块设计的灵感. 但 PEM 中的串行结构 不能很好地表达图像映射的线性关系,因此改为并行 结构的特征提取方式 IEM,以更好提取图像的结构信 息. IEM 网络结构如图 2,首先采用深度可分离卷积 DWConv 编码位置信息得到特征,然后运用 LightNorm 模块对特征进行提取,再并行使用不同大小的卷积核 处理特征,最后将特征进行融合. IEM 网络流程公式如下:

$$N = X + (K_{1 \times 1} + K_{3 \times 3} + K_{5 \times 5})(AX + B)^{r}$$
(3)

其中, *A* 为可学习的缩放因子, *B* 为偏置项, *r* 为调整参数, *K*<sub>1×1</sub>, *K*<sub>3×3</sub>, *K*<sub>5×5</sub> 为不同卷积核的特征提取模块. *N* 是 IEM 中的特征提取的结果.

在整个过程中,由于受到并行的 LightNorm 的约束,参数学习过程能够很好地保留图像的结构信息,从 而更充分地满足色彩处理的需求. IEM 的设计方法有 助于提高图像处理的效率和性能,并使网络在色彩校 正任务中表现更出色.

#### 1.2 解码器结构

解码器的目的是生成多个层次的特征,使其与真 实数据的特征更接近,为后续的损失函数提供更好的 约束关系.在复合主干网络中,为了丰富解码器之间的 特征关联,将解码器分为 CUpSample 和 BUpSample 两 种结构,为不同的主干应用不同结构的解码器.



### 1.2.1 CUpSample

CUpSample 网络结构如图 3 所示, 浅层特征图分 辨率相对较高,可以保留诸如边缘、纹理和形状等低 级别的视觉特征,能够更好地捕捉图像中的结构信息.

与此相反,深层特征图在卷积神经网络中的深层层级 中生成,可以捕捉更高级别的抽象特征,如对象的类 别、形状等更多地语义信息. 深层特征与来自于同层 级的浅层特征 MID 相结合, 可以充分利用浅层特征与 深层特征的优势. CUpSample 网络在对深层特征进行 提取的时候,经过 UpSample 输入到 DDConv 模块<sup>[16]</sup>, 用以提高特征的感知域,增强特征的表达能力,使 CNN 能够更好地捕捉到特征图中的空间信息,提高特 征的质量和多样性. 为了确保通道数的值与下一层的 通道数相匹配以及避免在特征融合过程中产生较大 的偏差,最后通过连续的卷积操作来减少特征图的通 道数,保持特征的相对均衡.通道数的均衡有助于确 保深层特征和浅层特征之间的融合不会引入过多的 噪音或偏置,更有效地结合不同层级的信息,提高模 型的性能.



CUpSample 网络结构图 图 3

#### 1.2.2 DDConv

由于感受野的局限性, 普通卷积 Conv 很难提取边 缘模糊小目标或目标的特征.因此,卷积不可避免地对 图像生成的适应性较差,泛化能力较弱.尽管可变形卷 积和动态卷积在一定程度上优于普通卷积,在处理图 像生成的时候,它们仍然不具备平衡网络性能和大小的 能力.为了解决当前卷积操作的缺点,本文采用了 CIT<sup>[16]</sup> 中的 DDConv 的卷积策略,其能够根据具体的任务和 数据分布自适应地学习核变形偏移和权重系数,从而 实现卷积核的形状和值的变化. 其特征图的输出是:

$$y = \sigma((a1 \cdot W1 + \dots + an \cdot Wn) \cdot x) \tag{4}$$

其中, n 是权重系数的数量, an 是可学习参数的权重系 数, y 是 DDConv 生成的输出特征图.

DDConv 在执行卷积操作之前, 根据相应的权重 系数线性组合不同的权重矩阵,实现了卷积核权重的

## 1.2.3 BUpSample

动态调整.

BUpSample 网络结构如图 4 所示,旨在将浅层和 深层的特征信息有机地结合在一起,以提高神经网络 的性能. 与 CUpSample 不同的是, 在进入到 BUpSample 之前,网络中已经引入相应的全局的信息,为了避免 DDConv 结构对全局特征产生影响, 在设计中取消了 对应的 DDConv 的结构. 在深度卷积神经网络中, 特征 图中通常会出现一些相似性很高的特征图, 它们之间 的差异较小. BUpSample 中的 Ghost module 模块是处 理具有高度相似性的特征图的模块,通过简单的线性 运算,能够生成更多像素的特征图,而无需引入大量的 额外参数,这意味着网络可以用更少的参数表示更多 的特征信息,从而降低计算和存储成本,提高计算效率. 与 CUpSample 一致的是, 深层特征需要与浅层特征图

的 MID 模块进行连接,完成深浅层的融合,将来自深 层结构的高级语义特征与来自浅层结构的纹理特征巧 妙地结合在一起,提高网络的表现力.值得注意的是, 在融合特征之后,为了充分地提取特征,还需要进一步 地进行特征提取.为了实现这一目标,网络两次使用 Ghost Module 模块将融合后的特征进一步提取,有效 增强了网络的表现能力.



图 4 BUpSample 网络结构图

## 2 损失函数

本文构建了 SmoothL1 与感知损失相结合的损失 函数,既缓解了梯度不平滑与爆炸的问题,同时保证了 图像质量.

在图像生成学习过程中,使用 L1 损失有助于模型 学习生成图像的每一个细节与目标之间的对应关系. 然而,L1 损失可能会引起一些梯度不平滑的问题,而 L2 损失则会导致梯度爆炸.为了解决梯度问题,使用 L1 和 L2 结合的 SmoothL1 损失,其公式如下:

$$loss_{x,y} = \begin{cases} 0.5(x-y)^2, \text{ if } |x-y| < 1\\ |x-y| - 0.5, \text{ otherwise} \end{cases}$$
(5)

其中, x 和 y 分别表示模型的输出和标签, |x-y|表示它 们之间的差异. 当|x-y|小于 1 时, 即差异较小时, 采用平 方误差, 差异较大时, 采用线性误差.

当 x 较小时,式(5)等效于 L2 损失,有助于在学习 中保持平滑;而当 x 较大时,式(5)等效于 L1 损失,在 一定程度上解决了梯度爆炸的问题.

为了使生成的图像更符合人眼对图像质量的感知, 在 SmoothL1 基础上增加感知损失. 与传统的损失函数 不同, 感知损失源自图像风格迁移, 它通过比较不同层 的特征表示来衡量图像之间的感知差异, 使生成的图 像更符合人眼对图像质量的感知. 其公式如下:

$$PerceptualLoss = \frac{1}{N} \sum_{i=1}^{N} N(F_i(x) - F_i(y))^2 \qquad (6)$$

其中, x 表示输入图像, y 表示目标图像, F<sub>i</sub>(x) 和 F<sub>i</sub>(y) 分别表示它们在预训练的神经网络中的第 i 层的特征 表示, N 是特征层数. 所以使用 SmoothL1 与感知损失 函数可以有效地保护图像的细节. 由于损失函数无论 是 SmoothL1 还是感知损失, 都更加侧重于图像的感知 质量, 以期更加满足于人眼的视觉需求, 所以在生成的 部分图像会比原图像更加明亮.

## 3 实验结果与分析

## 3.1 数据集

本文采用公开数据集 low-light dataset<sup>[2]</sup>, 它是在真 实场景中拍摄的低光与正常光照下的配对数据集, 被 广泛地应用于基于深度学习的低照度图像增强领域. 该数据集包括弱光和正常光, 通过改变曝光时间与 ISO 来收集, 包括 LOL-V1 和 LOL-V2 两个版本. LOL-V1 数据集包括 500 个图像对, 分辨率为 400×600, 其中 485 对作为训练集, 15 对作为测试训练集. LOL-V2 数据集 包括 789 对图像, 分辨率为 400×600, 其中 689 对作为 训练集, 100 对作为测试集.

## 3.2 实验与参数设置

本文实验的硬件环境为 GPU GeForce RTX 3090, CUDA 版本 11.3, 所使用的深度学习框架为 PyTorch, Python 版本为 3.8. 实验参数 batch\_size 为 3, 迭代次数 为 600, 学习率初始值为 0.001, 在 300 轮之后数值基本 稳定. 损失函数采用 SmoothL1 与感知损失联合计算, 其中感知损失的参数为 0.4. 学习率的调整使用余弦退 火算法, 参数设置为 300.

实验选取 PSNR 指标和 SSIM 指标作为图像质量 的衡量指标,作为在不同的应用场景下的判断标准. PSNR 使用峰值信噪比评价图像质量,它可以评价两幅 图像之间的相似程度.由于 PSNR 不能很好地按照人 眼的感受来衡量两张图像的相似度,因此增加 SSIM 指标,它不仅考虑了图像的亮度信息,而且还考虑了图 像的结构和纹理的相似性,能更接近人眼的感受.两个 指标可以对图像增强效果进行良好的综合评价.

### 3.3 实验结果对比及分析

本实验中使用的数据集为 LOL-V1 和 LOL-V2. 为 了验证本文所提模型的有效性,在 LOL-V1 数据集上 将本文的算法模型与 ZERO-DCE、RetinexNet、DRBN、 3D-LUT、KIND、Uformer、IPT、RCT、UHDFour、 MAXIM及IAT等主流方法进行性能对比,实验结果 见表 1. 如表 1 所示,本文方法在 PSNR 和 SSIM 上都 取得了更好的结果,比之前产生较好结果的IAT与UHD-Four在 PSNR 指标上分别提升了 0.44 dB 和 0.16 dB, 在 SSIM 指标上分别提升了 0.016 与 0.027.

表1 主干融合	网络在 LOL-V1 数据集	長下的效果
Methods	PSNR (dB)	SSIM
ZERO-DCE	14.83	0.531
RetinexNet	16.77	0.562
DRBN	19.55	0.746
3D-LUT	16.35	0.585
KIND	20.86	0.790
Uformer	16.36	0.771
IPT	16.27	0.504
RCT	22.67	0.788
UHDFour	23.10	0.820
MAXIM	23.43	0.863
IAT	23.38	0.809
Ours	23.54	0.836

在 LOL-V2 数据集上,将本文的算法模型与 ZERO-DCE、RetinexNet、DRBN、3D-LUT、KIND、Uformer、 IPT、RCT、MAXIM 及 IAT 方法进行对比,实验结果 见表 2. 如表 2 所示,本文方法在 PSNR 和 SSIM 上都 取得了更好的结果,比之前产生较好结果的 IAT 与 MAXIM 在 PSNR 指标上分别提升了 0.85 dB 和 1.49 dB, 在 SSIM 指标上分别提升了 0.053 与 0.046.

表2 主干融合图	网络在 LOL-V2 数据集	下的效果
Method	PSNR (dB)	SSIM
ZERO-DCE	14.32	0.511
RetinexNet	18.37	0.723
DRBN	20.13	0.820
3D-LUT	17.59	0.721
KIND	19.74	0.761
Uformer	18.82	0.771
IPT	19.80	0.813
RCT		0.788
MAXIM	22.86	0.818
IAT	23.50	0.824
Ours	24.35	0.871

为了直观地验证本文模型的有效性,将本文增强 后的微光图像视觉效果与LIME、ZERO-DCE、RCT、 KIND++、MAXIM和IAT算法的可视化效果进行比 较,如图 5 和图 6 所示,LIME和ZERO-DCE方法的可 视化效果最差,图像亮度提升效果不佳,大部分区域呈 现亮度不足的情况.RCT、KIND++、MAXIM和 IAT方法都可以提高图像的亮度,但是细节上均存在 缺失,无法产生令人满意的效果.本文模型增强的微光 图像,虽然整体亮度略高于真实图像,但在图像细节和 噪声去除方面都优于对比的全部方法.



图 5 非均匀光照的视觉效果对比图

#### 3.4 消融实验

为了验证双分支上采样结构的有效性,在 LOL-V1 数据集上,使用 4 种不同的双分支构成方式进行消融 实验对比.如表 3 所示, Exp1 中两个分支均是 BUp-Sample 的双分支结构; Exp2 中两个分支均是 CUpSample 的双分支结构; Exp3 中分支 1 是 CUpSample,分支 2 是 BUpSample 的双分支结构; Exp4 分支 1 是 BUp-Sample,分支 2 是 CUpSample 的双分支结构.实验结 果见表 3, 结果显示, Exp3 中分支 1 为 CUpSample 且 分支 2 为 BUpSample 的双分支结构更能提高图像增 强效果, 其在 PSNR 和 SSIM 评价指标上结果都更优, 分别为 23.54 dB 和 0.836.

为了验证损失函数对网络结构的影响,本实验在 LOL-V1数据集进行损失函数的消融实验,使用6组常 用的损失函数组合进对比,实验结果见表4. Exp1、 Exp2和 Exp3 分别使用 L1 损失函数、SmoothL1 损失

函数以及感知损失函数,由表4可见,感知损失表现比 其他的损失函数表现得更好.为了探索模型更好的拟 合程度, Exp4、Exp5 和 Exp6 分别将 L1 和 SmoothL1 损失与感知损失结合在一起,由表4中可以观测出, SmoothL1 与感知损失函数相结合是实验中效果最好 的一种组合, PSNR 和 SSIM 评价指标分别为 23.54 dB 和 0.836.

最后为了验证复合主干融合的网络优于单分支的 网络结构.本实验在 LOL-V1 数据集,以3种结构进行 消融实验对比,分别为仅使用 CNN 结构 (Exp1) 仅使 用 IEM 结构 (Exp2) 以及 CNN 与 IEM 结合的结构 (Exp3). 实验结果见表 5, 由表中的数据可以观察到, Exp3 的 PSNR 和 SSIM 评价指标比 Exp2 分别提升了 1.03 dB 和 0.145, 表明了复合主干融合网络的有效性.



图 6 均匀光照下的视觉效果对比图

表 3 采样结构对网络效果的影响				
结构	CUpSample	BUpSample	PSNR (dB)	SSIM
Exp1	—	Branch1+Branch2	22.71	0.779
Exp2	Branch1+Branch2	—	22.83	0.814
Exp3	Branch1	Branch2	23.54	0.836
Exp4	Branch2	Branch1	23.15	0.822

表 4 损失函数对网络模型的影响					
结构	L1	SmoothL1	感知损失	PSNR (dB)	SSIM
Exp1	$\checkmark$	_	—	20.35	0.788
Exp2	—	$\checkmark$	—	21.19	0.763
Exp3	_	—	$\checkmark$	22.30	0.809
Exp4	$\checkmark$	—	$\checkmark$	21.03	0.756
Exp5	_	$\checkmark$	V	23.54	0.836
Exp6	$\checkmark$	V	V	23.14	0.823
		100			

	表5 名	<b>各分支对网</b>	络模型的影响	
结构	CNN	IEM	PSNR (dB)	SSIM
Exp1		_	18.34	0.682
Exp2	_	$\checkmark$	22.51	0.791
Exp3	$\checkmark$	$\checkmark$	23.54	0.836

## 4 结束语

本文提出了一种复合主干融合的网络结构,从特 征提取的方向提升了图像亮度增强的效果,为图像保 留了更多的细节特征使暗光图像能够更加清晰.在损

200 软件技术•算法 Software Technique•Algorithm

失函数中使用 SmoothL1 损失和感知损失相结合, 使图 像更加平滑, 增强图像的视觉效果. 在 LOL-V1 和 LOL-V2两个数据集上的实验结果表明,复合主干融合的网 络结构可以有效地增强暗光图像质量,优于全部对比 的算法.由于复合主干融合网络的模型复杂程度高,参 数量大,训练时间长,在未来的工作中,将继续寻找新 的轻型网络结构的组合模式,降低模型复杂度,提升网 络性能.

#### 参考文献

- 1 王利娟, 常霞, 张君. 基于 Retinex 的彩色图像增强方法综述. 计算机系统应用, 2020, 29(6): 13-21. [doi: 10.15888/j.cnki. csa.007430]
- 2 Wei C, Wang WJ, Yang WH, et al. Deep retinex decomposition for low-light enhancement. Proceedings of the 2018 British Machine Vision Conference. Newcastle: BMVC, 2018.
- 3 Zhang YH, Zhang JW, Guo XJ. Kindling the darkness: A practical low-light image enhancer. Proceedings of the 27th ACM International Conference on Multimedia. Nice: ACM, 2019. 1632-1640.
- 4 Zhang YH, Guo XJ, Ma JY, et al. Beyond brightening lowlight images. International Journal of Computer Vision, 2021, 129(4): 1013–1037. [doi: 10.1007/s11263-020-01407-x]

- 5 Guo CL, Li CY, Guo JC, et al. Zero-reference deep curve estimation for low-light image enhancement. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 1777-1786.
- 6 Li CY, Guo CL, Loy CC. Learning to enhance low-light image via zero-reference deep curve estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(8): 4225-4238.
- 7 Zeng H, Cai JR, Li LD, et al. Learning image-adaptive 3D lookup tables for high performance photo enhancement in real-time. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 2058-2073. [doi: 10.1109/ TPAMI.2020.3026740]
- 8 Jiang YF, Gong XY, Liu D, et al. EnlightenGAN: Deep light enhancement without paired supervision. IEEE Transactions on Image Processing, 2021, 30: 2340-2349. [doi: 10.1109/ tip.2021.3051462]
- 9 Wang ZD, Cun XD, Bao JM, et al. Uformer: A general Ushaped Transformer for image restoration. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022. 17662-17672.
- 10 Lim S, Kim W. DSLR: Deep stacked Laplacian restorer for low-light image enhancement. IEEE Transactions on Multimedia, 2021, 23: 4272-4284. [doi: 10.1109/TMM.2020. 3039361]

- 11 Chen HT, Wang YH, Guo TY, et al. Pre-trained image processing Transformer. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 12294-12305.
- 12 Kim H, Choi SM, Kim CS, et al. Representative color transform for image enhancement. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 4439-4448.
- 13 Tu ZZ, Talebi, Zhang H, et al. MAXIM: Multi-axis MLP for image processing. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 5759-5770.
- 14 Li CY, Guo CL, Zhou M, et al. Embedding Fourier for ultrahigh-definition low-light image enhancement. Proceedings of the 11th International Conference on Learning Representations. Kigali: ICLR, 2023.
- 15 Cui ZT, Li KC, Gu L, et al. You only need 90k parameters to adapt light: A light weight Transformer for image enhancement and exposure correction. Proceedings of the 33rd British Machine Vision Conference. London: BMVC, 2022.238.
- 16 Lei T, Sun R, Wang X, et al. CiT-Net: Convolutional neural networks hand in hand with vision Transformers for medical image segmentation. Proceedings of the 32nd International