

基于改进 GAN 的猕猴桃低光照花朵图像增强模型

杜金芝, 李书琴*

(西北农林科技大学信息工程学院, 杨凌 712100)

摘要: 针对自然条件下采集的猕猴桃花朵图像由于相互遮挡导致的图像亮度低的问题, 提出一种基于改进 GAN (generative adversarial network) 的猕猴桃低光照花朵图像增强模型。首先, 设计一种基于混合注意力的生成器, 捕获低光照花朵图像亮度分布特征, 强化生成器对图像亮度分布的关注。其次, 引入 Swin Transformer Block 对图像特征进行全局建模, 减少生成器在下采样过程中的特征丢失。最后, 采用双判别器对生成图像进行全局和局部判别, 提升生成器的细节感知力和评估精度。该方法与 RetinexNet、EnlightenGAN、Zero-DEC、Cycle-GAN 四种方法进行比较, 在峰值信噪比 (peak signal to noise ratio, PSNR) 方面比 EnlightenGAN 提高了 0.11db; 在 NIQE (naturalness image quality evaluator) 方面, 比 Zero-DCE 降低 12.41%。该方法不仅实现了低光照猕猴桃花朵图像地亮度增强, 同时很好地抑制了噪声和伪影地出现, 可为低光照花朵图像增强提供参考。

关键词: 猕猴桃; 花朵; 图像增强; 低光照; GAN; Transformer; 注意力机制

doi: 10.11975/j.issn.1002-6819.202405156

中图分类号: S24; TP391.4

文献标志码: A

文章编号: 1002-6819(2024)-24-0165-07

杜金芝, 李书琴. 基于改进 GAN 的猕猴桃低光照花朵图像增强模型[J]. 农业工程学报, 2024, 40(24): 165-171. doi: 10.11975/j.issn.1002-6819.202405156 <http://www.tcsae.org>

DU Jinzhi, LI Shuqin. Model for enhancing low-light kiwifruit flower images based on improved GAN[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(24): 165-171. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202405156 <http://www.tcsae.org>

0 引言

中国是猕猴桃生产大国^[1], 其猕猴桃产业对国家农业经济的发展具有重要贡献。然而, 在自然环境中, 由于猕猴桃花朵间的相互遮挡, 所采集的图像存在亮度不足、噪声等问题。这些复杂因素使后续视觉任务产生了识别率下降、准确性降低等负面影响, 因此提升猕猴桃低光照花朵图像质量具有重要意义。

传统低光照图像增强算法通过建立和优化数学模型对图像进行像素层面的修正^[2], 主要包括基于直方图的低光照图像增强方法和基于 Retinex 的低光照图像增强方法。例如, 李金朋等^[3]采用限制对比度自适应直方图均衡化算法对 HSV 颜色空间中饱和度 (S) 和亮度 (V) 分量进行增强, 取得较好的增强效果。WEN 等^[4]根据 RGB (red green blue) 3 个颜色通道将水下彩色图像进行直方图均衡化, 最终输出的图像有较高对比度。上述基于直方图的低光照图像增强方法通过限制对比度来减少图像失真现象, 只考虑图像的亮度和对比度改变, 忽略了图像纹理信息。而在猕猴桃低光照花朵图像增强中, 纹理信息的恢复对于提高后续检测模型的精度、准确率等性能至关重要。针对基于直方图方法造成的纹理信息的丢失, 部分学者采用基于 Retinex 的方法, 在处理图像

时人为设计先验函数。例如, 张军国等^[5]通过对照度分量图像的自适应估计与校正对 Retinex 进行优化, 实现不同光照条件下野生动物监测图像的自适应增强。姬伟等^[6]采用引导滤波对 Retinex 算法进行改进, 实现果实边缘信息保存。郑涤尘等^[7]提出了一种基于 Retinex 理论的低照度图像自适应增强算法, 对原有图像亮度提升的同时真实自然地还原了图像。尽管基于 Retinex 的方法很好地保留了图像纹理等细节信息, 但由于其严重依赖成对数据集, 导致泛化能力差。且基于 Retinex 模型假设的基本场景是环境光照均匀的情况, 因此该模型对局部光照不均匀的图像进行增强时性能明显下降。

相比之下, 基于深度学习的低光照图像增强方法通过有监督/无监督学习, 有效解决了泛化能力差的问题。例如, 在有监督学习领域, YANG 等^[8]提出两个 CNN 网络结合的方式用于低光照图像增强, 实现对图像的矫正和重建。WU 等^[9]通过使用迭代优化控制图像噪声生成, 提升图像视觉质量。上述有监督学习在低光照图像增强去噪方面表现出色, 但严重依赖成对数据集进行训练。然而收集庞大且高质量的成对图像比较困难, 且在成对数据上训练的模型泛化能力较差。针对此问题, 无监督学习通过挖掘正常光照图像与低光照图像之间的内在联系增强图像。例如, ZHU 等^[10]提出 RRDNet 网络模型, 扩大图像对比度的同时并抑制噪声产生。WANG 等^[11]利用混合注意力机制对 WAGAN (Wasserstein GAN) 进行改进, 能有效地去噪。尽管无监督学习方法解决了模型对成对数据集的严重依赖, 但在处理亮度分布不均

收稿日期: 2024-05-21 修订日期: 2024-11-16

基金项目: 高端外国专家引进计划 (G20200027087)

作者简介: 杜金芝, 研究方向为目标检测。Email: dujinzhi1999@163.com

*通信作者: 李书琴, 教授, 博士生导师, 研究方向为农业信息化与智能信息系统。Email: lsq_cie@nwsuaf.edu.cn

的图像时仍存在一定的色彩偏差问题。

因此,针对无监督学习在增强时出现的亮度和色彩偏差问题,本文结合猕猴桃低光照花朵图像,提出一种基于改进 GAN^[12] 的猕猴桃低光照花朵图像增强模型 EKIGAN (enhanced kiwiflower image generator)。模型通过改进 GAN 的生成器和判别器,从而实现对自然环境下猕猴桃低光照花朵图像的增强。

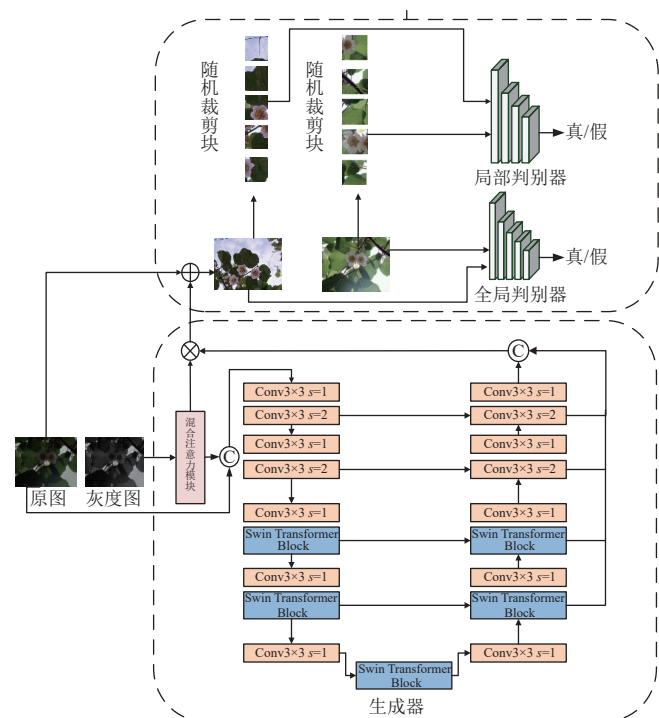
1 猕猴桃低光照花朵图像增强模型构建

GAN^[12] 是 Goodfellow 于 2014 年提出的无监督模型,其核心思想是通过生成器和判别器之间的对抗过程,来学习数据的真实分布并生成新数据

GAN 的训练过程如下,首先初始化生成器和判别器参数,生成器生成 n 个假样本。接着固定生成器,从训练集中随机选取 n 个真实样本,将真实和假样本输入判别器进行训练,以区分真假。经过 D 轮判别器更新后,更新生成器,使生成的假样本更真。这一过程不断增强两个网络,提高生成器生成质量和判别器识别能力。尽管 GAN 生成力强,但其训练易发生模式坍塌现象,导致生成图像存在过曝、增强不足或噪声的问题。因此需要对 GAN 进行改进。

1.1 基于改进 GAN 的猕猴桃低光照花朵图像增强模型

本文提出的猕猴桃低光照花朵图像增强模型其网络结构如图 1 所示。



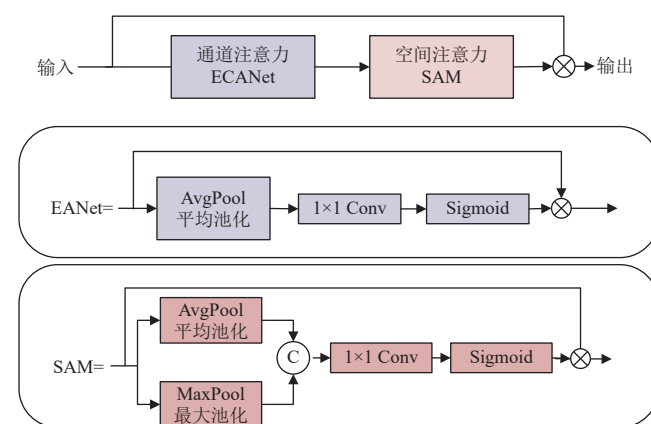
注: Swin Transformer Block 代表移位窗口 transformer 模块; s 表步长; \oplus 表示通道拼接; \otimes 表示按元素求和; \otimes 表示按元素相乘。下同
Note: Swin Transformer Block is shifted window transformer block; s is stride; \oplus represents channel splicing; \otimes represents element-wise sum; \otimes represents multiplication by element. The same as below.

图 1 基于改进 GAN 的猕猴桃低光照花朵图像增强模型结构图
Fig.1 Structure diagram of kiwifruit low light flower image enhancement model based on improved GAN

模型以 GAN 基础进行优化改进,生成器部分引入混合注意力模块和 Swin Transformer Block^[13],以更好的捕捉图像亮度和特征信息;在判别器部分使用 PatchGAN^[14] 对生成器生成的图像进行全局和局部判别,以提高生成器生成图像的质量和清晰度。

1.1.1 生成器网络

在原始 GAN 模型中,使用 U-Net^[15] 为生成器,其编码结构将输入图像的特征信息进行压缩,再经由解码结构恢复原始图像分辨率^[16],因此难以有效获取全局或广泛范围的亮度变化,导致生成图像易出现局部增强不足或过亮的问题。混合注意力模块使用残差结构将通道注意力模块与空间注意力模块结合,捕捉图像中的亮度分布情况,其网络结构如图 2 所示。



注: 1×1 Conv 表示 1×1 的常规卷积; sigmoid 代表激活函数。
Note: 1×1 Conv is 1×1 convolution; sigmoid representing activation function.

图 2 混合注意力模块

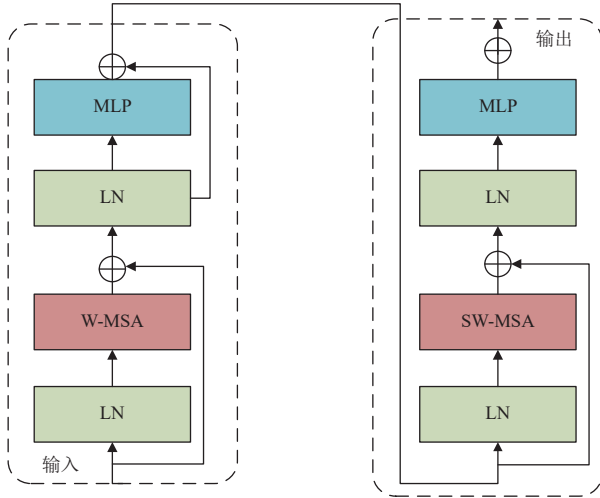
Fig.2 Mixed attention module

通道注意力模块采用 ECANet^[17] (efficient channel attention module) 对图像特征进行提取,其结构如图 2 所示。输入特征图 $X \in R^{H \times W \times C}$, (C 是输入通道数, H 特征图的高, W 是特征图的宽,下同) 首先进行全局平均池化操作,计算每个通道上的特征平均值,得到一个长度等于通道数量的一维向量。然后将该向量输入到 Sigmoid 函数中得到代表各通道权重的向量,最后将原特征图与权重相乘。ECANet 在 SENet^[18] 架构基础上改进,摒弃了降维策略,利用局部跨通道的交互方式辅助网络更精确地捕捉图像中特定区域的亮度分布。

空间注意力模块采用 SAM^[19] 对图像特征进行提取,其主要结构如图 2 所示。对输入尺寸为 $X \in R^{H \times W \times C}$ 的特征图先分别进行全局最大池化和全局平均池化,得到两个 $H \times W \times 1$ 的特征图。然后将两个 $H \times W \times 1$ 的特征图,按照通道拼接的方式得到尺寸为 $H \times W \times 2$ 的特征图,随后对拼接结果进行卷积和 Sigmoid 激活函数操作得到空间注意力权重矩阵 M ,最后将原特征图与权重进行相乘得到输出结果。SAM 的全局池化操作,用于提取特征图的全局亮度信息;SAM 的最大池化操作用于保留每个窗口最显著的特征,即亮度高的图像信息。同时 SAM 利用轻量级的计算操作,使模型聚焦于图像中的特定区域,实现局部关注。

混合注意力模块利用通道注意力和空间注意力相结合的方式，实现对猕猴桃低光照花朵图像特征中的通道相关性和空间相关性信息提取，突出有利的光照特征信息，去除无效信息。

GAN 生成器在进行下采样过程中，为减少计算复杂度，增加感受野，图像空间分辨率会逐渐减小，导致图像中的细节信息丢失。因此本文在生成器的下采样和上采样阶段之间引入 Swin Transformer Block。Swin Transformer Block 利用 Transformer 的全局建模能力有效建立图像中的长距离依赖关系，提高生成器对图像细节特征的提取。其结构如图 3 所示。



注：LN 表示层归一化；W-MSA 和 SW-MSA 分别表示具有常规窗口配置和移位窗口配置的多头自注意力模块；MLP 表示多层感知器。
Note: LN represents layer normalization; W-MSA and SW-MSA represent multi-head self attention module with regular window configuration and shifted window configuration, respectively; MLP represents multi-layer perceptron.

图 3 两个连续的 Swin Transformer Block

Fig.3 Two consecutive Swin Transformer Block structures

Swin Transformer Block 是基于原始 Transformer 中的标准多头自注意力演变而来的。通过对输入图像进行拆分计算图像各模块间的依赖关系。

具体来说，对于输入尺寸为 $X \in \mathbb{R}^{H \times W \times C}$ 的猕猴桃低光照花朵图像，先根据图像尺寸将其拆分为若干个 $S \times S$ (S 为窗口大小) 的局部窗口并重塑为 $HW/S^2 \times S^2 \times C$ 。随后对窗口进行重塑以适应后续的计算。对于局部窗口特征 $P \in \mathbb{R}^{S^2}$ ，在计算标准的自注意力时，通过构建 Q 、 K 和 V 矩阵来捕捉窗口内的特征依赖关系，其计算见式 (1)。

$$Q = PI_Q, K = PI_K, V = PI_V \quad (1)$$

式中 I_Q 、 I_K 和 I_V 是投影矩阵，参数由不同的窗口相互共用，一般来说 $Q, K, V \in \mathbb{R}^{S^2 \times d}$ 。每个独立窗口内的注意力矩阵计算过程见式 (2)。

$$\text{Attention}(Q, K, V) = \text{SoftMax}(QK^T / \sqrt{d} + B)V \quad (2)$$

式中 B 表示可学习的相对位置编码， d 表示 K 的维度。

两个连续的 Swin Transformer Block 结构由 W-STB 和 SW-STB 组成，其运行流程参考文献 [13]。经过改进后的生成器既能利用混合注意力机制捕获图像亮度信息，

还具有 Swin Transformer Block 全局特征提取的能力，这极大的改进了生成器的生成能力。

1.1.2 全局和局部判别器

在处理低光照条件下猕猴桃花朵图像时，若背景昏暗而花朵区域明亮，单判别器会遇到模式崩溃的问题。这是因为单判别器难以提供生成器所需的自适应调节能力，其反馈给生成器的信息也不够全面。为自适应地增强局部区域以及改善全局光线，本文使用双判别器机制增强对图像细节的感知能力，提升对生成图像评估的准确性。其具体实现细节如下。

当输入图像进入全局判别器时，首先被切割成多个固定大小的局部区域，这些局部区域称为 Patch。接着通过卷积操作提取每个局部区域的细节和纹理特征。随后每个区域特征表示被用于进行真假分类。PatchGAN 独立地评估每个区域，并输出其真实性概率。最终这些局部区域的真实性概率被汇总为整体真实性评分，该评分用于计算 GAN 的损失函数并指导生成器的训练。

进入局部判别器的图像首先也会被切割成 Patch。局部判别器从生成器输出图像及正常光照图像中，随机抽取 Patch 进行判别，并通过损失函数的约束，学习判断这些局部区域的真实性。通过模拟正常光照图像的局部特征来增强网络对于局部细节和亮度分布，从而生成更为自然的增强结果。局部判别器的结构与全局判别器相似，随机抽取的每个 Patch 都经过真假判别，得出其真实性概率。不同的是局部判别器专注于图像的局部区域或小块，而全局判别器关注整个图像的信息。最终这些概率相加求平均，作为该局部区域的判别结果。

1.1.3 损失函数

为保障训练过程的稳定，增强网络的学习能力，该模型损失函数的参数参考文献 [20-24]，权重设置为 1，训练模型的整体损失函数为

$$\text{Loss} = L_{SFP}^{Global} + L_{SFP}^{Local} + L_G^{Global} + L_D^{Local} \quad (3)$$

式中 Loss 代表模型的整体损失， L_{SFP}^{Global} 代表全局自特征保留损失， L_{SFP}^{Local} 代表局部自特征保留损失， L_G^{Global} 代表全局对抗性损失， L_D^{Local} 代表局部对抗性损失。

1.2 试验环境及参数设置

本文试验在 Ubuntu 20.04 系统上进行，GPU 采用 NVIDIA GeForce RTX3090，CUDA 版本为 11.6，深度学习框架为 Pytorch1.8.0。模型在训练时 epoch 为 200，学习率为 0.000 1，使用 Adam 作为训练时的优化器，BatchSize 为 32。

1.3 评价指标

下面本文所用到的评价指标进行介绍。

(1) PSNR^[25] (peak signal to noise ratio)

对于一个尺寸位 $m \times n$ 的无噪图像 I 以及相同尺寸的含噪图像 T ，通过均方误差 MSE (mean squared error) 量化两幅图像之间的差异程度，如式 (4) 所示：

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - T(i, j)]^2 \quad (4)$$

式中 $I(i, j)$ 和 $T(i, j)$ 分别表示原始图像和处理后的图像在 (i, j) 位置处的像素值。

PSNR 定义如式 (5):

$$PSNR = 10 \lg \frac{255^2}{MSE} \quad (5)$$

PSNR 的取值范围为 $[0, 100]$, 两幅图的 PSNR 值越大代表两幅图越相近, PSNR 为 100 时表示两幅图完全相同。通常情况下, PSNR 大于 30 时, 可以认为图像质量达到了较高的标准。

(2) NIQE

NIQE^[26] (natural image quality evaluator) 是一种无参考图像质量评估方法, 用于在没有原始参考图像的情况下评估图像的质量。NIQE 的设计基于自然场景统计 (natural scene statistics, NSS) 模型, 旨在使用一个多元线性回归模型模拟人类视觉系统对图像质量的感知。计算如式 (6):

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{\left((v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2) \right)} \quad (6)$$

式中 v_1 、 v_2 表示图像多元高斯模型的均值向量, Σ_1 、 Σ_2 代表图像的模型协方差矩阵, D 为 NIQE 的评估结果, NIQE 值越小代表图像越自然。

2 数据采集及处理

2.1 数据采集

本文以不同开放状态的猕猴桃花朵作为研究对象。图像采集工作在西北农林科技大学眉县猕猴桃试验站内完成, 使用数码相机 (Canon EOS 40D) 和安卓手机作为图像的采集设备, 共收集 1 280 张自然条件下低光照的猕猴桃花朵图像和对应的正常光照图像, 如图 4 所示。所采集的低光照的猕猴桃花朵图像和正常光照的猕猴桃花朵图像为同一天内傍晚、中午、破晓等时段内收集而成, 包含猕猴桃花朵开放的不同状态, 经过简单裁剪处理后, 将所有图像的分辨率调整到为 1280×960 像素。

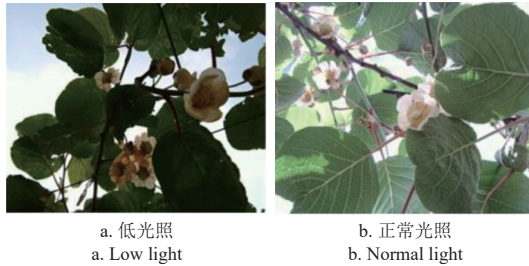


图 4 不同光照强度图像采集
Fig.4 Different light intensity image acquisition

2.2 数据处理

为适应模型对输入图像尺寸的统一要求, 本文采用双线性插值法来调整猕猴桃花朵图像的大小。利用双线性插值法将猕猴桃花朵原始图像调整为 640×400 像素。

3 结果与分析

3.1 可视化结果对比与分析

为了证明本文方法的可行性, 将提出的方法与现有的低光照图像增强方法 Retinex-Net^[27]、Cycle-GAN^[28]、EnlightenGan^[20]、Zero-DCE^[29] 比较。将数据集中相同的图像输入到各个模型中, 结果如图 5 所示。

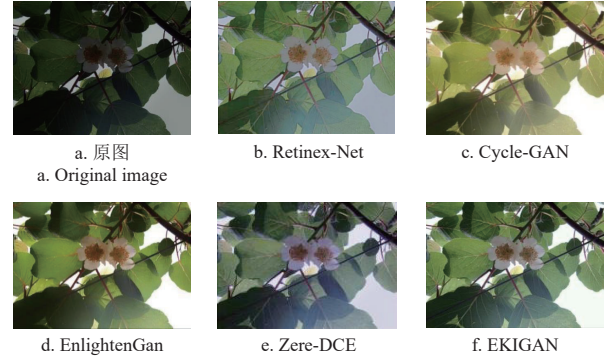


图 5 不同图像增强模型可视化对比
Fig.5 Visual comparison of different image enhancement models

从图 5 中可以看到, 在猕猴桃低光照花朵图像增强方法中, Retinex-Net、Cycle-GAN、EnlightenGan 和 Zero-DCE 等方法存在着不同程度的不足。经过 Retinex-Net 处理后的图像色彩严重失真, 花朵和叶片的颜色过于鲜艳, 同时纹理信息丢失, 与实际情况不符。Cycle-GAN 方法的增强结果树叶细节出现丢失现象, 整体与原图像颜色存在较大差异。EnlightenGan 方法的增强结果在图像左下方的树叶部分出现一定程度的噪音和伪影, 影响了整体图像质量, 树叶细节处理不够理想。Zero-DCE 增强后的图像整体色彩偏灰, 增强亮度不够突出, 但细节保留较多且未出现明显的伪影现象。相比之下, 本文提出的方法 EKIGAN 在猕猴桃低光照花朵图像增强方面表现出色。增强后的图像亮度适中, 花朵和叶片分别得到了适度的增强, 颜色自然且未出现明显的伪影和噪音现象, 细节信息得到了较好地保留。这表明本文方法在处理猕猴桃低光照花朵图像时, 能够有效改善图像质量, 保持自然感和细节清晰度。

3.2 图像增强效果与分析

为比较本文提出的模型 EKIGAN 对低光照图像增强的有效性, 选取 2 幅低光照猕猴桃花朵图像, 对比分析使用本文模型进行增强前后的在细节和颜色方面的效果。试验中采用测试集中正常光照的图片作为衡量指标的参考图像, 结果如图 6 所示。

图 6 a 中左下角区域表现极暗, 细节丢失明显, 图像亮度也比较低, 但是经过增强后的图 6 c 比较完整地恢复了细节信息, 亮度也得到比较明显的提升。图 6 b 中的低光照图像对比度不明显, 叶片与叶片之间的边界模糊, 无法看清, 经过增强后的图 6 d 对比度得到明显的提升, 叶片与叶片之间, 以及叶片和花朵之间边界明显。

图 6 中图像增强前后的 PSNR 和 NIQE 见表 1。可看出与原始图像相比, 增强图像的 PSNR 值均得到了明

显的提升, NIQE 值降低, 图像质量变好, 说明本文提出的方法有效。

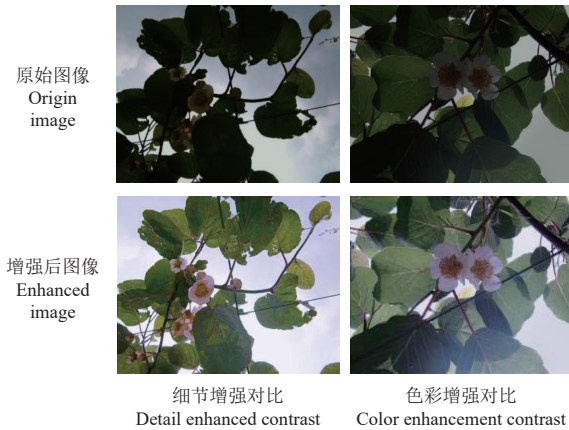


图 6 本文方法增强效果

Fig.6 The method in this article enhances the effectiveness

表 1 本文方法的图像增强评价结果

Table 1 The evaluation results of image enhancement using the method described in this article

增强效果 Enhance effectiveness	PSNR/db		NIQE	
	原始图像 Original image	增强图像 Enhanced image	原始图像 Original image	增强图像 Enhanced image
细节增强 Detail enhancement	5.87	6.46	11.572	10.654
色彩增强 Color enhancement	6.67	6.83	11.469	10.789

为进一步验证 EKIGAN 的有效性, 本文选取 5 种主流的图像增强技术来进行研究与比较, 试验结果如表 2 所示, EKIGAN 在 PSNR 上的指标为 7.09 db, NIQE 为 10.36。Retinex-Net 在数据集上的 PSNR 指标为 6.77, 低于 EKIGAN, 说明 EKIGAN 增强的图像失真效果小。Zero-DCE 方法在数据集上的 PSNR 指标为 6.64 db, 比本文方法少 0.45 db。EnlightenGan 方法在数据集上的 PSNR 指标为 6.98 db, 比 EKIGAN 少 0.11 db。Cycle-GAN 方法在数据集上的 NIQE 指标为 11.02, 比 EKIGAN 高 0.66, 说明 Cycle-GAN 增强后的图像自然度不如本文方法, 色彩偏差大。

表 2 不同方法在自建数据集上的评价结果

Table 2 Evaluation results on self built datasets using different methods

方法 Methed	PSNR/db	NIQE
Retinex-Net	6.77	10.57
EnlightenGan	6.98	10.48
Cycle-GAN	6.75	11.02
Zero-DCE	6.64	11.83
Diffusion-Low-Light ^[30]	7.05	10.47
EKIGAN	7.09	10.36

Zero-DCE 方法在数据集上的 NIQE 指标为 11.83, EKIGAN 为 10.36, 比 Zero-DCE 低 12.41%。除此之外, Diffusion-Low-Light 方法在本文自建数据集上面也表现良好, PSNR 为 7.05 db 与本文所提出方法差距很小, NIQE 也达到 10.36, 与 EKIGAN 仅有 0.11 的差距。综上所述,

EKIGAN 与其他的图像增强方法相比在 PSNR 与 NIQE 指标上均达到了最好, 说明 EKIGAN 增强后的图像质量好, 失真程度小, 图像自然度好, 色彩偏差小。

3.3 消融试验

为了探究本文改进方式给 GAN 模型带来的性能提升, 进行消融试验, EKIGAN 模型在 DICM 数据集^[31]上的消融试验结果如表 3 所示。

表 3 EKIGAN 模型消融试验

Table 3 Ablation experiment of EKIGAN model

混合注意力模块 Mixed attention module	移位窗口 transformer 模块 Swin Transformer block	局部判别器 Local discriminator	自特征损失函数 Self-characteristic retention loss	NIQE
×	√	√	√	9.84
√	×	√	√	9.78
√	√	×	√	9.88
√	√	√	×	9.98
√	√	√	√	9.32

注: √为采用该模块, ×为未采用该模块

NOTE: √ represents the adoption of this module, and × represents did not use this module

试验采用了逐步排除的方法, 具体做法是逐一不采用某个模块, 观察模型处理图像能力和输出结果的变化。通过每次只改变一个模块的状态, 评估每个模块对整体性能的影响。表 3 展示了分别不采用某个模块或损失函数输出结果的 NIQE 指标。从表中可以看出, 当不采用混合注意力模块时, NIQE 值比采用该模型上升 0.52, 这说明混合注意力模块可以提取到图像亮度信息。当不采用 Swin Transformer 模块时 NIQE 值为 9.78, 上升 0.46, 这说明 Swin Transformer 模块对生成器提取全局特征信息有帮助。当进不采用局部判别器时, NIQE 为 9.88, 当不采用自特征保留损失时, NIQE 为 9.98。然而当将上述改进均加入时, NIQE 为 9.32, 试验证明改进均对模型整体具有良好的作用, 本文算法的改进有效。

4 结 论

针对无监督学习在增强时出现的亮度和色彩偏差问题, 提出一种基于改进 GAN 的猕猴桃低光照花朵图像增强模型 EKIGAN。首先, 使用混合注意力模块提取图像亮度分布信息, 引导生成器根据图像亮度强度分布生成图像; 其次, 在生成器网络中引入 Swin Transformer, 对图像特征信息进行全局建模, 提升模型提取图像特征信息的能力, 避免出现图像内容特征丢失的问题; 最后, 为进一步避免图像出现局部过度增强以及局部增强不足的情况, 引用双判别器对生成图像进行判别。研究结果表明, EKIGAN 在自建的猕猴桃低光照图像数据集上, 峰值信噪比达到 7.09 db, 超过一众主流低光照图像增强模型。在 NIQE 方面达到 10.36, 较 Diffusion-Low-Light 小 0.11, 达到不错的自然图像质量。改进模型具有出色的图像增强效果, 可应用于实际生产中, 为智慧农业方面的应用带来便利。

[参 考 文 献]

[1] 潘时佳, 吴津乐, 程梅, 等. 基于改进 CNN 的猕猴桃根区土

- 壤含水率反演方法[J]. *农业工程学报*, 2024, 40(11): 85-91.
- PAN Shijia, WU Jinle, CHENG Mei, et al. Inversion method for root soil water content using improved CNN[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2024, 40(11): 85-91. (in Chinese with English abstract)
- [2] 郑建华, 杨高林, 刘双印, 等. 双背景光自适应融合与透射图精准估计水下图像复原[J]. *农业工程学报*, 2022, 38(14): 174-182.
- ZHENG Jianhua, YANG Gaolin, LIU Shuangyin, et al. Accurate estimation of underwater image restoration based on dual-background light adaptive fusion and transmission maps[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2022, 38(14): 174-182. (in Chinese with English abstract)
- [3] 李金朋, 冯帅, 杨鑫, 等. 融合 CLAHE-SV 增强 Lab 颜色特征的水稻覆盖度提取[J]. *农业工程学报*, 2023, 39(24): 195-206.
- LI Jinpeng, FENG Shuai, YANG Xin, et al. Unsupervised extraction of rice coverage with incorporating CLAHE-SV enhanced Lab color features[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2023, 39(24): 195-206. (in Chinese with English abstract)
- [4] WEN Y, TIAN W, ZHENG B, et al. New approach for underwater imaging and processing[J]. *Proceedings of Spie*, 2014, 9100: 91000S-8.
- [5] 张军国, 程浙安, 胡春鹤, 等. 野生动物监测光照自适应 Retinex 图像增强算法[J]. *农业工程学报*, 2018, 34(15): 183-189.
- ZHANG Junguo, CHENG Zhe'an, HU Chunhe, et al. Adaptive image enhancement algorithm for wild animal monitoring based on Retinex theory[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2018, 34(15): 183-189. (in Chinese with English abstract)
- [6] 姬伟, 吕兴琴, 赵德安, 等. 苹果采摘机器人夜间图像边缘保持的 Retinex 增强算法[J]. *农业工程学报*, 2016, 32(6): 189-196.
- JI Wei, LV Xingqin, ZHAO De'an, et al. Edge-preserving Retinex enhancement algorithm of night vision image for apple harvesting robot[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2016, 32(6): 189-196. (in Chinese with English abstract)
- [7] 郑涤尘, 何继开, 刘艺, 等. 基于 Retinex 理论的低照度图像自适应增强算法[J]. *计算机科学*, 2024, 51(10):1-10.
- ZHENG Dichen, HE Jikai, LIU Yi, et al. Low light image adaptive enhancement algorithm based on Retinex theory[J]. *Computer Science*, 2024, 51(10):1-10. (in Chinese with English abstract)
- [8] YANG X, XU K, SONG Y, et al. Image correction via deep reciprocating HDR transformation [C]// *Proc of IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, Salt Lake City, Utah, USA: IEEE, 2018: 1798-1807.
- [9] WU W, WENG J, ZHANG P, et al. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, Louisiana, USA: IEEE, 2022: 5901-5910.
- [10] ZHU A, ZHANG L, SHEN Y, et al. Zero-shot restoration of underexposed images via robust retinex decomposition [C]// *Proc of IEEE International Conference on Multimedia and Expo*. Piscataway, London: IEEE, 2020: 1-6.
- [11] WANG R, JIANG B, YANG C, et al. MAGAN: Unsupervised low-light image enhancement guided by mixed-attention[J]. *Big Data Mining and Analytics*, 2022, 5(2): 110-119.
- [12] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// *Proceedings of the 27th International conference on Neural Information Processing Systems*. Cambridge, UK: MIT Press, 2014: 2672-2680.
- [13] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Montreal, Quebec, Canada: IEEE, 2021: 10012-10022.
- [14] HEPBURN A, LAPARRA V, MCCONVILLE R, et al. Enforcing perceptual consistency on generative adversarial networks by using the normalised laplacian pyramid distance[EB/OL].(2020-06)[2024-11-26]. <http://dx.doi.org/10.7557/18.5124>.
- [15] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//*Medical Image Computing and Computer-assisted Intervention*. Munich, Germany:Springer International Publishing, 2015: 234-241.
- [16] 何银银, 胡静, 陈志泊, 等. 融合门控变换机制和 GAN 的低光照图像增强方法[J]. *计算机工程*, 2024, 50(2): 247-255.
- HE Yinyin, HU Jing, CHEN Zhibo, et al. Low-light image enhancement method combining gated transformation mechanism and GAN[J]. *Computer Engineering*, 2024, 50(2): 247-255. (in Chinese with English abstract)
- [17] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE, 2020: 11534-11542.
- [18] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE, 2018: 7132-7141.
- [19] ZHU X, CHENG D, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Seoul, Korea, 2019: 6688-6697.
- [20] JIANGY, GONG X, LIU D, et al. Enlightengan: Deep light enhancement without paired supervision[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [21] LIU X, TAN H, TONG X, et al. Feature preserving GAN and multi-scale feature enhancement for domain adaption person re-identification[J]. *Neurocomputing*, 2019, 364: 108-118.
- [22] FU Y, HONG Y, CHEN L, et al. LE-GAN: Unsupervised low-light image enhancement network using attention module and identity invariant loss[J]. *Knowledge-Based Systems*, 2022, 240: 108010.

- [23] ZHANG Y, WAN Z, WANG D, et al. Multi-scale feature aggregation and fusion network with self-supervised multi-level perceptual loss for textures preserving low-dose CT denoising[J]. *Physics in Medicine & Biology*, 2024, 69(10): 105003.
- [24] DING Y, ZHANG Z, ZHAO X, et al. Self-supervised locality preserving low-pass graph convolutional embedding for large-scale hyperspectral image clustering[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16.
- [25] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th International Conference on Pattern Recognition. Istanbul, Turkey:IEEE, 2010: 2366-2369.
- [26] MOORTHY A K, BOVIK A C. Blind image quality assessment: From natural scene statistics to perceptual quality[J]. *IEEE Transactions on Image Processing*, 2011, 20(12): 3350-3364.
- [27] LIANG J, XU Y, QUAN Y, et al. Deep bilateral retinex for low-light image enhancement[EB/OL]. (2020-07)[2024-05-21]. <https://arxiv.org/pdf/2007.02018v1>.
- [28] ZHU J, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2223-2232.
- [29] GUO C, LI C, GUO J, et al. Zero-reference deep curve estimation for low-light image enhancement[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, Tennessee, USA: IEEE, 2020: 1780-1789.
- [30] LAN G, MA Q, YANG Y, et al. Towards Flexible and efficient diffusion low light enhancer[EB/OL].(2024-07)[2024-05-21]. <https://arxiv.org/html/2410.12346v1>.
- [31] LEE C, LEE C, KIM C S. Contrast enhancement based on layered difference representation[C]//2012 19th IEEE International Conference on Image Processing, Lake Buena Vista, Orlando, USA: IEEE, 2012: 965-968.

Model for enhancing low-light kiwifruit flower images based on improved GAN

DU Jinzhi , LI Shuqin^{*}

(College of Information Engineering, Northwest A & F University, Yangling 712100, China)

Abstract: Uneven brightness and noise images have posed a great challenge on the recognition rate and accuracy of visual tasks, due mainly to the mutual occlusion of kiwifruit flowers in natural environments. In this article, the image enhancement was proposed to detect the kiwifruit flower under low light using improved generative adversarial network (GAN). The research object was taken as the flowers at the Meixian Kiwi Experimental Station of Northwest Agriculture and Forestry University, Shaanxi Province, China. Digital cameras and mobile phones were used to capture the images of flowers in different open states at some periods. All uniform pixel images were selected to create a dataset. Three steps were divided in the improved model of image enhancement. Firstly, GAN was used to optimize and improve the model. The channel and spatial attention were combined to perform the residual connections in the generator section. Among them, ECANet was used to more accurately extract the image features under the brightness distribution of specific areas in the channel attention. SAM was used to focus the specific regions in the image using lightweight computational operations in the spatial attention, in order to achieve the local attention. As such, the generator was obtained to extract the brightness distribution of kiwifruit low-light flower images. Secondly, a Swin Transformer block was added at the connection between the upsampling and downsampling of the generator. The Swin Transformer block was used as the standard multi-head self-attention evolution of the original Transformer. Therefore, the global modelling of the Transformer was also utilized to split the input image. The dependency relationship among each module of the image was then calculated to improve the global modelling and feature extraction of image details in the generator. Finally, the pattern collapse that encountered by single discriminators was reduced to adaptively enhance the local regions and global lighting. The generator was then provided with the required adaptive adjustment. A dual discriminator mechanism was utilized to enhance the perception of image details. The accuracy of image evaluation was improved to generate the clearer images in the generator network. On the self-built dataset, the PSNR and NIQE of the improved model were 7.09 and 10.36, respectively. Meanwhile, a comparison was then made on the five models, such as RetinexNet, Enlighten GAN, Zero DEC, Cycle GAN, and Diffusion Low Light. The peak signal-to-noise ratio of the improved model was 0.11db higher than that of EnlightenGan. The quality assessment of natural image was 12.41% lower than that of Zero DCE. Therefore, the low-light image enhancement model was performed the better quality, lower distortion, better naturalness, and smaller color deviation after enhancement. This improved model can also be applied in the practical production to develop the computer vision in smart agriculture.

Keywords: kiwifruit; flowers; image enhancement; low light; GAN; transformer; attention mechanism