

doi:10.6041/j.issn.1000-1298.2023.04.024

基于改进 RDN 网络的无人机茶叶图像超分辨率重建

鲍文霞¹ 吴育桢¹ 胡根生¹ 杨先军² 汪振宇¹

(1. 安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心, 合肥 230601;
2. 中国科学院合肥物质科学研究院, 合肥 230031)

摘要: 针对无人机搭建可见光传感器进行茶叶长势、病害等监测中因飞行高度影响图像分辨率的问题,本文提出了一种改进的残差密集网络(Residual dense network, RDN)用于无人机茶叶图像超分辨率重建。针对无人机茶叶图像纹理复杂的特点,以 RDN 为基线网络,在其结构中引入了残差组(Residual group, RG)模块,将多个残差通道注意力模块(Residual channel attention block, RCAB)组合在一起,通过引入注意力机制来区别对待不同的通道,关注无人机茶叶图像高频细节信息,从而提高网络的表征能力;同时设计了一个卷积长跳跃结构,利用带有卷积的远程跳跃连接,动态调整经过残差密集块(Residual dense block, RDB)后特征的权重,更好地利用无人机茶叶图像的分层特征信息,从而提升超分辨率重建图像的质量。实验结果表明,本文改进的 RDN 网络在无人机茶叶图像测试集上相较于其他算法表现更优,超分辨率重建后的图像具有更高的峰值信噪比和结构相似度,在 4 倍超分的情况下分别达到 36.03 dB 和 0.913 2,能够为茶叶智能化监测研究提供支持。

关键词: 茶叶; 图像重建; 超分辨率; 残差组模块; 卷积长跳跃结构

中图分类号: S127 文献标识码: A 文章编号: 1000-1298(2023)04-0241-09 **OSID:** 

Super-resolution Reconstruction of Unmanned Aerial Vehicle Tea Images Based on Improved RDN Network

BAO Wenxia¹ WU Yu'an¹ HU Gensheng¹ YANG Xianjun² WANG Zhenyu¹

(1. National Engineering Research Center for Agro-ecological Big Data Analysis and Application, Anhui University, Hefei 230601, China
2. Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

Abstract: It is a relatively economical, flexible and time-effective method to build a visible light sensor for monitoring of tea growth and diseases, but the resolution of the image will be affected by the flying height of the UAV. Therefore, an improved residual dense network (RDN) for super-resolution reconstruction of UAV tea images was proposed. Specifically, in view of the complex texture of UAV tea images, taking RDN as the baseline network, residual group (RG) was introduced into its structure, combining multiple residual channel attention modules were combined together to treat different channels differently by introducing an attention mechanism, and paying attention to the high-frequency detail information of UAV tea images, thereby improving the representation ability of the network; at the same time, a convolutional long jump structure was designed, using the long-range skip connection with convolution, to dynamically adjust the weight of the feature after passing through the residual dense block (RDB), and making better use of the hierarchical feature information of the UAV tea image, thereby improving the super-resolution of the quality of reconstructed image. The experimental results showed that the improved RDN network performed better than other algorithms on the test set of UAV tea images, and the super-resolution reconstructed images had higher peak signal-to-noise ratio and structural similarity. In the case of quadruple super resolution, it can reach 36.03 dB and 0.913 2, respectively, which can provide support for the follow-up research of tea intelligent monitoring.

Key words: tea; image reconstruction; super-resolution; residual group; convolutional long jump structure

收稿日期: 2022-06-24 修回日期: 2022-08-25
基金项目: 安徽省自然科学基金项目(2208085MC60)、国家自然科学基金项目(62273001)、安徽省高等学校自然科学研究重大项目(KJ2020ZD03)、安徽省中央引导地方科技发展专项(202107d06020001)和安徽省高校研究生科学研究项目(YJS20210013)
作者简介: 鲍文霞(1980—),女,教授,博士,主要从事农业与生态视觉分析与模式识别研究,E-mail: bwxia@ahu.edu.cn
通信作者: 胡根生(1971—),男,教授,博士,主要从事农业信息处理与应用研究,E-mail: hugs2906@sina.com

0 引言

为保证茶叶产量,需要实时关注茶树长势,并且对茶叶病虫害进行有效监测,同时也要避免大面积喷洒出现过度污染问题。如果仅仅依靠人工目测的方法,具有较强的主观性,且耗时又费力^[1]。采用无人机搭载可见光传感器对茶叶进行监测不仅效率高而且成本低。然而无人机图像由于飞行高度的影响导致获取的图像分辨率受到限制,虽然获取的茶叶图像整体区域较大,但是分配到某块小区域上难以清晰地呈现其细节,不利于进行后续的茶叶病害精确检测等研究。因此,无人机茶叶图像的超分辨率重建具有重要意义。

图像超分辨率重建^[2] (Super-resolution reconstruction, SRR) 是一种提高图像分辨率的计算机视觉方法,成本极低,可应用于各种领域。图像超分辨率重建方法主要分为基于插值^[3-4]的方法、基于重建^[5]的方法和基于学习^[6-10]的方法 3 类。其中插值法通过像素点之间的空间位置关系推导出缩小和放大后对应点的像素值,虽然能够简单并快速地重建图像,但是超分辨率重建的能力有限,重建的图像并不理想。基于重建的算法是通过图像的先验知识结合数学模型求解高分辨率 (High resolution, HR) 图像的特征表达,从而重建出高分辨率图像。基于机器学习的超分辨率算法相比于基于重建的算法在视觉上效果更好,因为此算法利用大量的样本能够直接学习低分辨率 (Low resolution, LR) 图像和高分辨率图像之间的映射关系。一般基于学习的超分辨率算法包括基于邻域嵌入^[11]的方法、基于样本^[12]的方法和基于深度学习^[13-17]的方法等。DONG 等^[18]使用卷积神经网络来实现图像超分辨率重建,提出了采用端到端训练方式的 SRCNN 算法,展现了良好的重建效果,为深度学习在超分辨率重建领域的发展奠定了基础。HE 等^[19]提出了残差卷积神经网络解决随着网络加深导致网络训练出现梯度消失的问题,大大提高了网络泛化性能。KIM 等^[20]将残差思想用在了图像超分辨问题上,提出了深度超分辨 (Super-resolution using very deep convolution network, VDSR) 模型,在解决网络训练难以收敛的同时取得了更好的重建效果。随着卷积神经网络在图像超分辨领域中的不断发展,更深的网络使得每个卷积层具有不同层级的感受野,但大多数基于卷积神经网络的超分辨模型并没有充分利用每个卷积层的信息。

ZHANG 等^[21]提出了一种残差密集网络 (Residual dense network, RDN), 该网络通过残差密

集块 (Residual dense block, RDB) 充分利用待重建图像中各个卷积层的分层特征,并利用全局特征融合方式获取待重建图像中的全局密集特征,提升了图像重建性能。无人机茶叶图像相较于地面拍摄的图像更为模糊,因此超分辨率重建更困难。另外,由于图像中茶叶叶片具有不同的尺寸、视角和高宽比,采用深度卷积神经网络进行超分辨率重建时分层特征更加重要,在重构时能提供更多的线索,因此选用 RDN 作为基线网络,利用残差密集的连接方式充分提取茶叶图像丰富的局部特征,然后通过全局特征融合的方式学习全局特征。同时,无人机茶叶图像具有复杂的纹理信息,因此超分辨率重建过程中需要关注有价值的高频信息才能充分恢复图像纹理细节。在深度学习领域中,大部分图像超分辨率重建算法都是平等对待通道特征,但各个特征的重要程度都不相同,从而导致网络在不重要的特征上浪费了很多计算资源。Senet 网络^[22]构建了 Squeeze 和 Excitation 结构来对特征图进行全局信息集成和自适应校准,使得模型更加关注重要的特征信息,其通道注意力思想在图像处理领域也发挥了重要的作用。ZHANG 等^[23]首次将通道注意力机制引入到图像超分辨率重建领域,提出残差通道注意力网络 (Residual channel attention network, RCAN), 在公开数据集上取得了良好的效果。受 RCAN 启发,本文在 RDN 网络基础上,引入残差组 (Residual group, RG) 模块,将通道注意力机制与残差思想相结合,利用注意力机制来区别对待不同的通道,从而提高网络对无人机茶叶图像的超分辨率重建能力,尽可能地恢复图像复杂的纹理信息。同时还设计一种卷积长跳跃结构,通过 1×1 卷积层将不同 RDB 模块的权重进行动态调整,充分并有效利用中间层的特征来提高最后的特征,以期提高网络性能。

1 研究方法

1.1 数据来源及预处理

本文的实验数据采集于安徽省巢湖市坝镇 (北纬 $31^{\circ}14'$, 东经 $117^{\circ}36'$, 海拔 40 m), 数据获取区域如图 1 所示。考虑到风力以及光照强度对图像质量的影响,选择数据的采集时间为 2021 年 5 月 3—4 日的 14:00 到 17:00, 气象条件良好,微风,少云。图像获取设备为 DJI Mavic Air2 型无人机,配备一台 4 800 万像素可见光摄像机,摄像机型号为 DJI FC3170,光圈值为 $f/2.8$,曝光时间为 $1/60\text{ s}$,ISO 速度为 ISO - 100,焦距 4 mm,最大续航时间为 34 min, 最远续航里程为 18.5 km,最大飞行速度为 19 m/s,

抗风等级 5 级, 拍摄图像尺寸为 8 000 像素 × 6 000 像素。无人机的飞行速度和方向由人工控制, 摄像机与地面垂直。无人机具有精确的 GPS 和 GLONASS 定位能力, 详细记录每幅图像的位置和高

度。无人机飞行高度是 5 m, 图 2 展示了部分无人机拍摄的茶叶图像, 其分辨率为 8 000 像素 × 6 000 像素, 将其裁剪成 935 幅 256 像素 × 256 像素的图像作为数据集, 部分裁剪图像如图 3 所示。

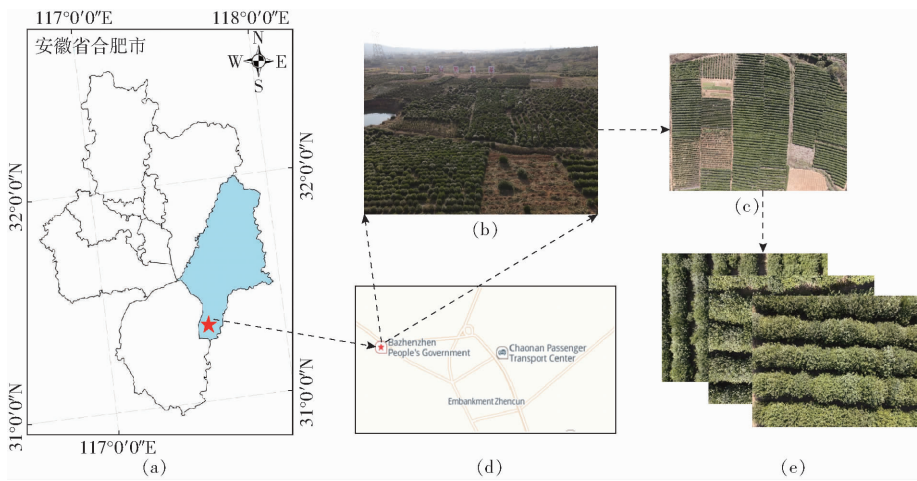


图 1 数据获取区域
Fig. 1 Data acquisition area



图 2 部分无人机拍摄的茶叶图像
Fig. 2 Partial drone image of tea leaves

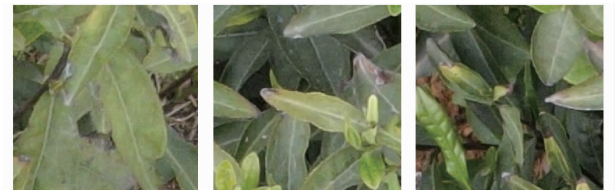


图 3 部分裁剪图像
Fig. 3 Partially cropped image

1.2 改进的 RDN 网络

1.2.1 网络总体框架

本文改进的 RDN 网络模型主要包括浅层特征提取网络、多重残差密集块 (Redidual dense blocks, RDBs)、密集特征融合以及上采样网络 4 个模块, 网络总体结构如图 4 所示。在浅层特征提取网络模块中, 和 RDN 相同, 由 2 个卷积层组成, 用于从输入的低分辨率图像中提取低水平特征; 在多重残差密集块中, 对于已经获取到的低分辨率图像的特征图, 需要把这些特征映射到高维上继续提取有用的信息, 这一部分除了原有的多个相同结构的 RDB 模块, 本文添加了残差组 RG 模块和卷积长跳跃网络, 目的是使网络专注于更加有用的信息并且增强辨别学习能力。另外在密集特征融合模块中, 包含全局特征

融合 (Global feature fusion, GFF) 和全局残差学习 (Global residual learning, GRL) 两部分, 其中 GFF 用于将所有的密集残差块提取到的特征融合在一起, 得到全局的特征, GRL 主要是利用了上面浅层特征和最终提取到的特征进行残差学习, 整个 GFF 模块充分利用了上述各层的特征, 使后面重建效果有了很大的提升。在上采样网络中, 采用上采样加卷积的操作, 重建高分辨率图像。

1.2.2 残差组

在自然场景中捕捉到的茶叶图像具有复杂的纹理信息, 因此图像超分辨的过程中必须注意到更为有价值的高频信息, 才能充分恢复茶叶图像复杂的纹理细节。一般来说, 图像的关键信息在通道中是分布不均匀的, 普通的卷积神经网络无法轻易地区分出关键的信息, 因此可以引入注意力机制来解决这个问题。本文引入的注意力机制通过对特征通道之间的相互依赖型建模来自适应地重新缩放每个通道的特征, 使网络能够专注于更有用处的信息并且提高判别学习的能力。RG 模块是由多个带有短跳跃连接的残差通道注意块 RCAB 组成的残差组, 模块结构如图 5 所示。这种残差结构能够促进卷积神经网络进行更高性能的图像超分辨训练。

在 RG 模块中, 引入短跳跃连接来获得输出 F_g , 该特征提取可表示为

$$F_g = F_{g-1} + W_g F_{g,B} \tag{1}$$

式中 F_{g-1} ——RG 模块输入
 W_g ——RG 尾部卷积层权重
 $F_{g,B}$ ——最后一个 RCAB 块输出

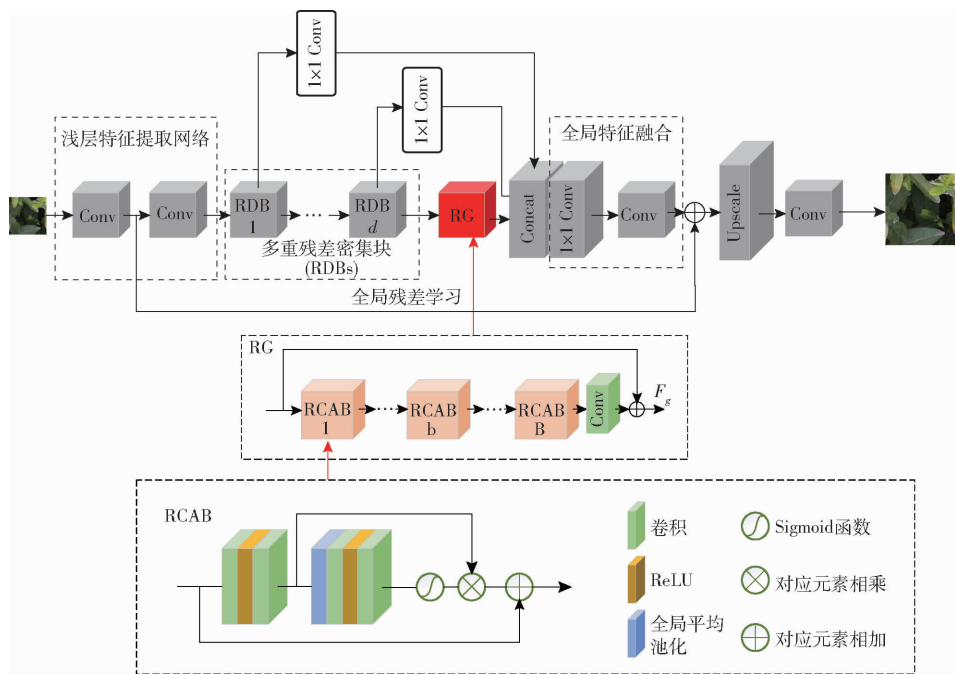


图 4 改进的 RDN 网络总体框架

Fig.4 Improved overall framework of RDN network

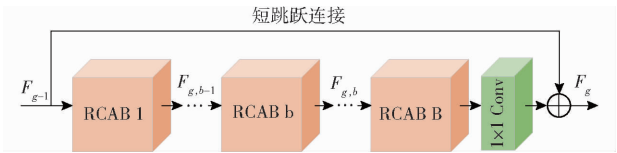


图 5 RG 结构

Fig.5 RG structure

图 5 中使用短跳跃连接可以使得网络更具有区分性的学习,从而更加关注通道与关键信息的相关性,重建出更高的图像分辨率。

在图 6 中,通过将注意力与残差思想相结合,将通道注意力集成到残差块当中,即 RCAB 模块。对于第 b 个残差块

$$F_{g,b} = F_{g,b-1} + R_{g,b}(X_{g,b})X_{g,b} \tag{2}$$

式中 $F_{g,b-1}$ ——RCAB 模块的输入

$F_{g,b}$ ——RCAB 模块的输出

$R_{g,b}$ ——通道注意力

$X_{g,b}$ ——经过卷积和 ReLU 操作

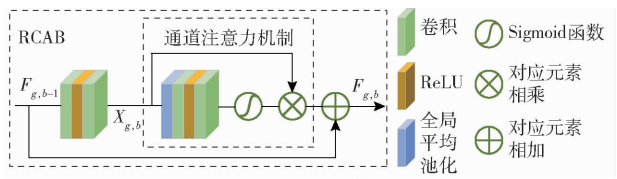


图 6 RCAB 结构

Fig.6 RCAB structure

首先输入经过卷积和 ReLU 操作得到 $X_{g,b}$,然后 $X_{g,b}$ 再经过一个通道注意力机制,得到的结果与输入进行相加得到 RCAB 的输出。其中卷积操作都采用 3×3 的卷积核。

在图 6 中,由于卷积核进行卷积操作时,只能提

取感受野内的信息,因此使用通道注意力机制可以对输入特征的每一个通道进行全局平均池化操作,将全局空间内的信息生成通道描述符,最终通道注意力机制给各个通道设置相应的权重来表示通道和关键信息的相关程度,若相关程度越大,则此通道的信息越重要,输出的特征也就越多,从而超分辨率重建出来的图像质量更高。

1.2.3 卷积长跳跃结构

在多重残差密集块中,主要由密集残差块和引入的 RG 模块组成,并且包含了卷积长跳跃结构。其中 RDB 模块如图 7 所示,包含密集连接层、局部特征融合和局部残差学习,从而形成了一种连续内存机制,将前一层 RDB 的状态传递到当前 RDB 的每一层;局部特征融合也就是图中的 1×1 Conv 部分,自适应融合之前 RDB 和当前 RDB 中整个卷积层的状态,起到一个降维的作用,其表达式为

$$F_{d,LF} = H_{LF}^d([F_{d-1}, F_{d,1}, \dots, F_{d,c}, \dots, F_{d,C}]) \tag{3}$$

式中 H_{LF}^d ——密集连接及局部融合

$F_{d,LF}$ ——降维后的输出

其中,局部残差学习 (Local residual learning) 用于将 F_{d-1} 和 $F_{d,LF}$ 的特征进行融合。表达式为

$$F_d = F_{d-1} + F_{d,LF} \tag{4}$$

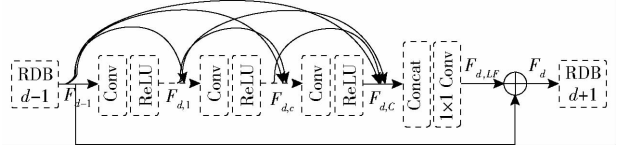


图 7 RDB 结构

Fig.7 RDB structure

式中 F_d ——经过第 d 个 RDB 后的输出

在每个 RDB 模块上,通过远程跳跃连接的操作有效促进了梯度的反向传播,充分利用了中间层的特征来提高最后的特征,卷积长跳跃结构如图 8 所示。但由于不同的 RDB 残差块的输出对于图像特征的影响程度是不同的,因此在面对复杂的茶叶图像时,单纯通过远程跳跃连接操作无法很好地恢复其高频的纹理信息,所以本文对每个远程跳跃连接操作都添加了 1×1 卷积层来动态调整每一部分的权重。后续实验也充分证明了此改进提升了网络的超分辨率重建效果。

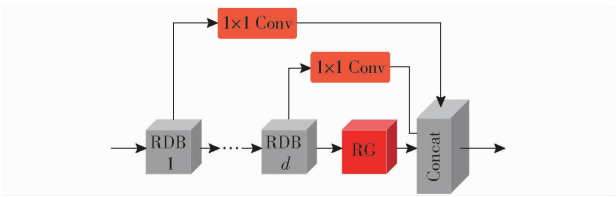


图 8 卷积长跳跃结构
Fig. 8 Convolutional long-skip structure

1.3 算法流程

由于拍摄设备为无人机,图像尺寸较大,为适应卷积神经网络模型输入尺寸,将采集的遥感图像裁剪为 256 像素 $\times$ 256 像素,并将裁剪后的图像构建训练集和测试集,采取旋转和翻转方式对训练样本进行增广,提高模型的泛化性能。使用增广后的样本训练改进后的 RDN 模型。将测试集中的样本输

入训练好的模型,输出超分辨率重建后图像以及评价指标,用于评估模型超分辨率重建的效果。由于图像超分辨率重建大多是使用下采样的方式构建成的数据集进行训练,本文使用插值算法进行下采样构建训练数据集。目前双三次插值^[24]是二维空间中最常用的插值方法,它相较于其他插值算法虽然复杂,但更为有效,能精准地插补图形,故本文采用双三次插值的下采样方式构建成的数据集进行训练。

无人机茶叶图像超分辨率重建流程如图 9 所示,算法的具体步骤如下:

- (1) 将原始遥感图像裁剪成 256 像素 $\times$ 256 像素,并将其分为训练集和测试集。
- (2) 将训练集中每幅图像进一步裁剪,即在 4 个角和中心处各截取 1 幅尺寸为 128 像素 $\times$ 128 像素的图像,并将此图像作为图像的原始高分辨率图像。由于超分辨率网络需要低分辨率图像作为输入,因此,将这些截取过后的训练样本通过双三次插值算法进行下采样,从而得到相应倍数的低分辨率图像。
- (3) 将得到的成对的高-低分辨率图像通过水平、竖直翻转和旋转等操作随机进行数据增广。
- (4) 增广后的数据用于训练改进的 RDN 模型。
- (5) 测试集中的图像由经过训练的模型进行超分辨率重建,根据指标评估模型超分辨率重建的效果。

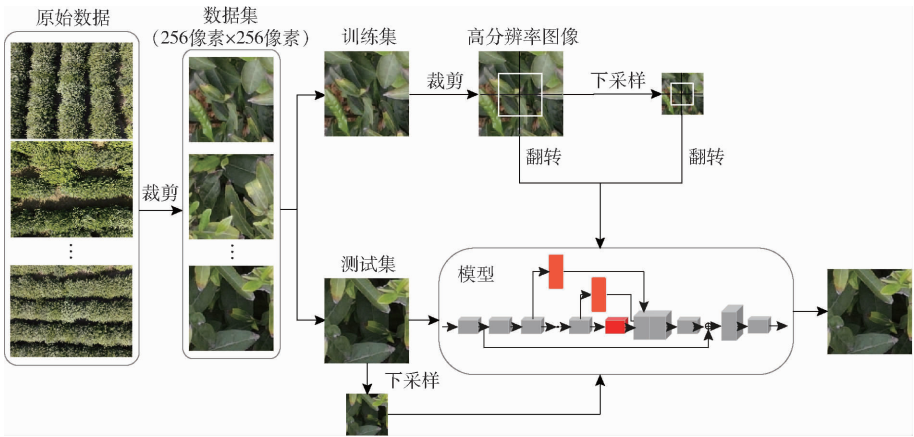


图 9 无人机茶叶图像超分辨率重建流程图
Fig. 9 Flowchart of UAV tea image super-resolution reconstruction

2 实验与结果分析

2.1 实验设置

实验以 Python 为编程语言,Pytorch 为深度学习框架,在 Intel 处理器以及 NVIDIA GPU 上运行。硬件和软件配置如表 1 所示。

在图像超分辨率重建领域中普遍是研究 2 倍到 4 倍的重建,本文研究同样的重建倍数方便与之比

表 1 硬件、软件环境配置

Tab. 1 Hardware and software environment configuration

平台	配置
操作系统	Ubuntu 18. 04 LTS 64-bits
CPU	Intel(R) Core(TM) i7-9700
GPU	NVIDIA GeForce RTX 2080Ti
GPU accelerator	CUDA 10. 1 & cuDNN 7. 6. 5
深度学习框架	PyTorch 1. 5
开发环境	PyCharm & Anaconda
编程语言	Python 3. 7

较。在每个训练批中,随机抽取 14 个低分辨率图像作为输入,并且随机将每个图像块水平或垂直翻转和旋转 90°。网络使用 ADAM^[25] 优化器,优化参数为 $\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8}$, 设置初始学习率为 10^{-4} , 并采用 L1 损失函数作为 ADAM 的优化目标, 计算式为

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|X_{HR}^i - D(X_{LR}^i)\|_1 \quad (5)$$

式中 $L(\theta)$ ——平均绝对误差
 X_{HR}^i ——低分辨率图像
 N ——训练对的数量
 $D(X_{LR}^i)$ ——重建低分辨率图像 X_{LR}^i

其余训练参数见表 2。

表 2 训练参数配置

Tab.2 Training parameters configuration

参数	数值
输入图像尺寸/(像素×像素×通道)	256×256×3
批量大小	20
训练迭代次数	500

2.2 评价指标

本文采用峰值信噪比^[26] (Peak signal-to-noise ratio, PSNR) 和结构相似度 (Structural similarity index, SSIM)^[27] 两种评价指标对重建图像的质量进行评估。PSNR 通过计算对应像素点之间的差值来衡量图像是否失真,其差值越小则 PSNR 的值越大,表明图像的超分效果越好。其计算公式为

$$V_{PSNR} = 10\lg \frac{255 \times 255}{L_{MSE}} \quad (6)$$

式中 L_{MSE} ——均方根误差
 V_{PSNR} ——峰值信噪比, dB

结构相似度更多的是体现结构上的差异, 计算式为

$$V_{SSIM} = \frac{(2\mu_a\mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)} \quad (7)$$

式中 a ——原始高分辨率图像
 b ——重建后的超分辨率图像
 μ_a, μ_b —— a, b 的灰度平均值
 σ_a^2, σ_b^2 —— a, b 的方差
 σ_{ab} —— a 和 b 的协方差
 C_1, C_2 ——常数 V_{SSIM} ——结构相似度

其中,结构相似度在 0~1 之间取值,结果越接近 1,表示重建后的超分辨率图像越接近原始高分辨率图像,效果越好。

2.3 实验对比分析

2.3.1 残差注意力块数量对模型性能的影响

为了验证 RG 注意力组中残差注意力块数量对

模型性能的影响,分别采用不同数量的 RCAB 模块进行茶叶图像超分辨率重建实验,实验结果如表 3 所示。

表 3 RCAB 模块数量对模型性能的影响

Tab.3 Effect of number of RCAB modules on model performance

RCAB 模块数量	PSNR/dB	SSIM
15	35.91	0.911 2
20	36.00	0.912 5
25	36.01	0.912 6
30	36.01	0.912 7

通过表 3 可以得出,当增加 RCAB 残差单元的数量时,模型的非线性表达能力得到提升,并且在数量达到 20 之后开始慢慢趋于稳定,当模块数量达到 25~30 时,相比较数量为 20 的网络仅仅提高 0.01 dB。因此考虑到模型的轻量化以及保证模型的性能,选取 RCAB 模块数量为 20 或 25 个。

上述实验是没有嵌入卷积长跳跃结构时的结果,为了验证两者共同作用下对模型性能的影响,嵌入卷积长跳跃结构,保证其他条件一致,进行茶叶图像超分辨率重建实验,结果如表 4 所示。从实验结果可以得知,当 RCAB 残差模块数量不超过 20 时,两者能够起到一个共同促进的作用,但超过一定数量时,卷积长跳跃结构并不能发挥其最佳效果,因此本文选择 RCAB 模块数量为 20。

表 4 卷积长跳跃结构与 RCAB 模块共同作用对模型的影响

Tab.4 Influence of convolutional long-skip structure and RCAB module on model

RCAB 模块数量	PSNR/dB	SSIM
15	35.94	0.911 6
20	36.03	0.913 7
25	36.00	0.912 4
30	35.99	0.912 4

2.3.2 不同模块对模型性能的影响

本实验以未做任何改进的 RDN 作为基线网络,对比了带有 1×1 卷积的长跳跃网络以及 RG 残差注意力组对于无人机茶叶图像超分辨率重建性能的影响,消融实验结果如表 5 所示。首先验证 1×1 卷积层在远程跳跃连接中的必要性。在没有 1×1 卷积的情况下,来自不同残块的特征对最终特征的贡献是相同的,不管是高级特征还是低级特征。从表 5 可以看出,带有 1×1 卷积的远程跳跃连接网络重建结果 PSNR 提升 0.22 dB,这也验证了 1×1 卷积的长跳连接的有效性;同样,在加入 RG 模块

后,PSNR 与 SSIM 的提升更为显著,充分说明了在针对无人机茶叶图像的超分辨率重建上,引入注意力的思想能够很好地提升茶叶图像超分辨率重建的效果;当两模块共同作用时,PSNR 和 SSIM 分别达到 36.03 dB 和 0.913 2,说明两模块相互作用可以有效提升模型对无人机茶叶图像的超分辨率重建性能。

表 5 各模块对模型性能的影响

Tab.5 Effect of each module on performance				
基线网络	RG	Skip + Conv	PSNR/dB	SSIM
RDN	×	×	35.75	0.910 2
	×	√	35.97	0.911 8
	√	×	36.00	0.912 5
	√	√	36.03	0.913 2

注: × 表示不使用该模块;√表示使用该模块。

2.3.3 注意力机制模块对模型性能的影响

为了进一步验证所提模型中使用的 RG 模块的优势,在相同的实验条件下,将 RG 模块替换成通道注意力(Squeeze and excitation,SE)以及卷积块注意力模块^[28](Convolutional block attention module,CBAM)并进行比较,对比实验结果如表 6 所示。由表 6 可以看出,使用 SE 模块后,模型 PSNR 提升 0.20 dB,说明通道注意力机制在自适应学习和调整各通道的特征图权重的同时,能够适当提高模型重建的效果;使用 CBAM 的模型 PSNR 比使用 SE 的模型略高一些,因为在空间上不同位置的纹理细节有所区别,所以空间上不同位置的重要性也不同,因此融入 CBAM 模块相对来说更有利于茶叶图像的恢复;而本文引入的 RG 模块 PSNR 提高 0.25 dB,SSIM 也取得了最优,效果相对明显。因此,通过引入具有残差思想的注意力机制能够使模型超分辨率重建的效果更佳,从而更加高效地恢复茶叶图像的纹理细节。

表 6 不同注意力模块对模型性能的影响

Tab.6 Performance impact of different attention modules on model		
模型	PSNR/dB	SSIM
RDN	35.75	0.910 2
RDN + SE	35.95	0.911 7
RDN + CBAM	35.96	0.912 0
RDN + RG	36.00	0.912 5

2.3.4 与其他模型性能对比

为了进一步验证所改进的模型对于无人机茶叶图像超分辨率重建的效果,在相同的实验条件下将其与 SRCNN^[18]、VDSR^[20]、RDN^[21]、RCAN^[23]等主

流图像超分辨率重建算法在不同超分辨率尺度上进行比较,见表 7。从表 7 可以看出,在不同尺度上,本文改进的 RDN 算法 PSNR 和 SSIM 均优于其他算法(PSNR 和 SSIM 越大,图像重建效果越好)。在放大倍数较低时,大多数超分算法都能够达到不错的效果,这是由于 2 倍的下采样得到的低分辨率图像保留了很多信息,使得超分辨率重建的难度相对较低,所以为了从定量和定性的角度证明本文改进算法的有效性,此处选择难度较高的 4 倍超分辨率对测试集的部分图像进行可视化重建,并与其他方法的重建图像进行比较,如图 10 所示。

表 7 本文算法与主流算法对比结果

Tab.7 Results of comparison between proposed algorithm and mainstream algorithm						
模型	PSNR (×2)/dB	SSIM (×2)	PSNR (×3)/dB	SSIM (×3)	PSNR (×4)/dB	SSIM (×4)
Bicubic	37.74	0.957 2	33.17	0.891 7	30.18	0.805 7
SRCNN	41.36	0.977 3	36.50	0.935 5	32.69	0.865 5
FSRCNN	41.65	0.978 5	37.01	0.939 7	33.27	0.875 8
VDSR	41.66	0.978 3	37.11	0.941 0	33.44	0.880 0
RCAN	41.90	0.977 9	37.47	0.940 2	34.94	0.896 6
RDN	42.48	0.980 0	37.66	0.941 8	35.75	0.910 2
本文算法	42.63	0.980 6	37.70	0.942 1	36.03	0.913 2

注:(×2)、(×3)、(×4)表示 2 倍、3 倍、4 倍超分辨率重建。

为了更直观地观察出无人机茶叶图像超分辨率重建的效果,图 10 用红色小框标注了观察区域。从放大之后的图像可以很明显看出,本文改进的 RDN 算法重建出来的图像无论是从纹理细节还是从整体上都更为清晰,与 HR 更为接近,而大多数的重建算法茶叶的纹理基本模糊,特别是 Bicubic 算法比较模糊。通过定量和可视化结果可知,针对无人机茶叶图像,本文改进的 RDN 算法相对现有的超分辨率重建算法有了一定提高。

3 结束语

本文提出一种改进的 RDN 网络用于无人机茶叶图像的超分辨率重建。该网络设计了一个卷积长跳跃结构并引入了残差注意力模块,不仅充分并有效地利用中间层的特征,还使得网络能够根据权重精准分配计算资源,通过与大多数经典的超分网络在 3 种不同超分辨率重建尺度上的对比,改进的模型在无人机茶叶数据集上取得了最优的效果,在 4 倍超分的情况下 PSNR 和 SSIM 分别达到 36.03 dB 和 0.913 2,实验充分证明了本文提出的超分辨率重建算法不仅在评价指标上有所提高,同时在视觉上

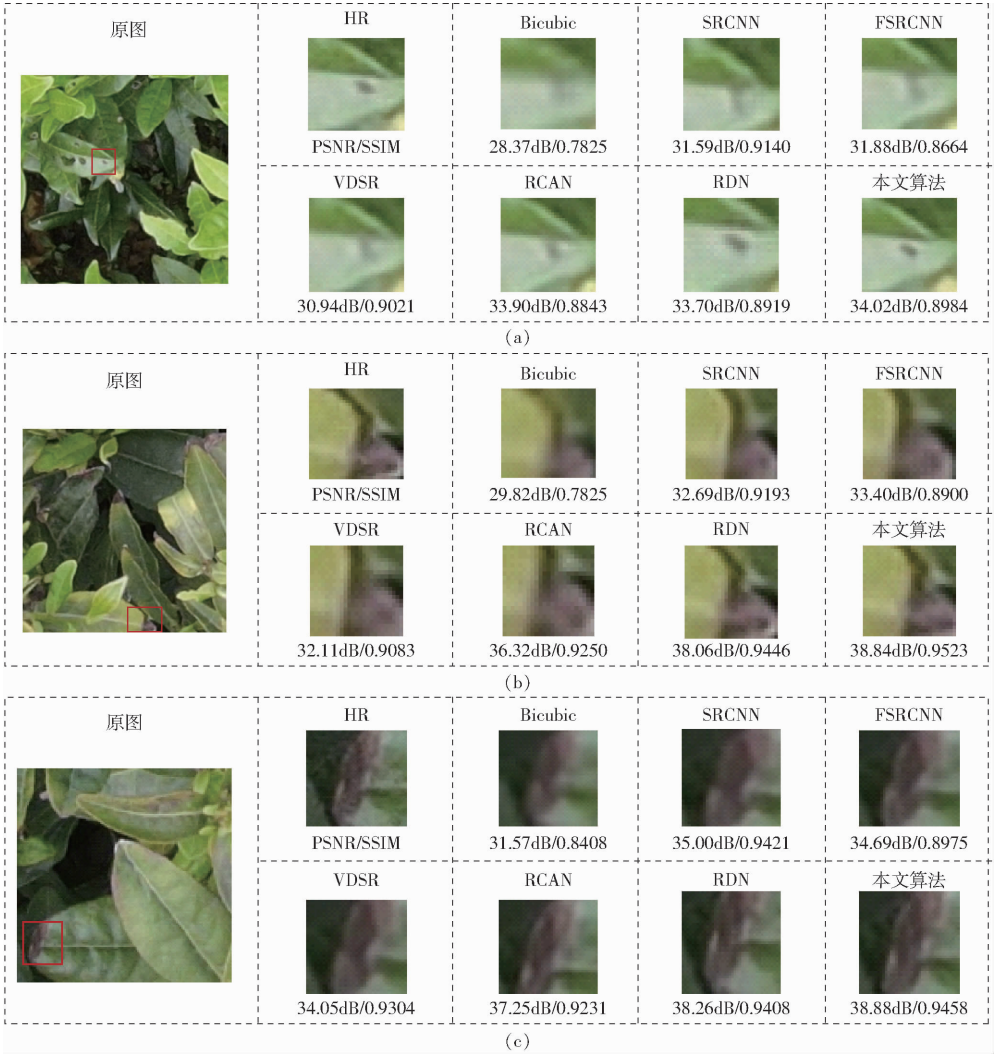


图 10 主流深度学习超分辨重建算法在测试集上的可视化结果对比

Fig. 10 Comparisons of visualization results of mainstream deep learning super-resolution reconstruction algorithms on test set

也取得了不错的效果。通过这种方式能够有效提升无人机遥感图像的分辨率,从而广泛应用到搭配无人机的农业领域当中,在解决无人机硬件成本问题的同时,不仅能够有助于观察茶树的长势,做好茶叶的产量估计,还能起到预防茶叶病害,减少茶叶生产损失等作用。

参 考 文 献

[1] 杨蜀秦,刘江川,徐可可,等. 基于改进 CenterNet 的玉米雄蕊无人机遥感图像识别[J]. 农业机械学报, 2021,52(9): 206-212.
YANG Shuqin, LIU Jiangchuan, XU Keke, et al. Improved CenterNet based maize tassel recognition for UAV remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 206-212. (in Chinese)

[2] 唐艳秋,潘泓,朱亚平,等. 图像超分辨率重建研究综述[J]. 电子学报,2020,48(7):1407-1420.
TANG Yanqiu, PAN Hong, ZHU Yaping, et al. A survey of image super-resolution reconstruction[J]. Acta Electronica Sinica, 2020, 48(7): 1407-1420. (in Chinese)

[3] WANG L, XIANG S, MENG G et al. Edge-directed single-image super-resolution via adaptive gradient magnitude self-interpolation[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2013, 23(8):1289-1299.

[4] RUANGSANG W, ARAMVITH S. Efficient super-resolution algorithm using overlapping bicubic interpolation[C]// Consumer Electronics. IEEE, 2017:1-2.

[5] DUN Y, DA Z, YANG S, et al. Kernel-attended residual network for single image super-resolution[J]. Knowledge-Based Systems, 2021, 213: 106663.

[6] 李佳星,赵勇先,王京华. 基于深度学习的单幅图像超分辨率重建算法综述[J]. 自动化学报, 2021, 47(10): 2341-2363.
LI Jiaxing, ZHAO Yongxian, WANG Jinghua. A review of single image super-resolution reconstruction algorithm based on deep

- learning[J]. *Acta Automatica Sinica*, 2021, 47(10): 2341–2363. (in Chinese)
- [7] SHAH Z H, MÜLLER M, WANG T C, et al. Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images[J]. *Photonics Research*, 2021, 9(5): B168–B181.
- [8] ZHU X B, LI S S, WANG L. A brief survey on deep learning based image super-resolution[J]. *High-tech Communication*, 2021, 27(3): 294–302.
- [9] LIU Z, MAO X, HUANG J, et al. Stratified attention dense network for image super-resolution[J]. *Signal, Image and Video Processing*, 2022, 16(3): 715–722.
- [10] 孙旭, 李晓光, 李嘉锋, 等. 基于深度学习的图像超分辨率复原研究进展[J]. *自动化学报*, 2017, 43(5): 697–709.
SUN Xu, LI Xiaoguang, LI Jiafeng, et al. Review on deep learning based image super-resolution restoration algorithm[J]. *Acta Automatica Sinica*, 2017, 43(5): 697–709. (in Chinese)
- [11] CHANG H, YEUNG D Y, XIONG Y. Super-resolution through neighbor embedding[C] // *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004.
- [12] 李晓光, 沈兰荪, LAM Kin Man, 等. 一种基于自样本学习的人脸图像超分辨率复原算法[J]. *高技术通讯*, 2009, 19(4): 377–381.
LI Xiaoguang, SHEN Lansun, LAM Kin Man, et al. An algorithm for super-resolution restoration of human face based on self-sample learning[J]. *High Technology Letters*, 2009, 19(4): 377–381. (in Chinese)
- [13] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 4681–4690.
- [14] REN H, EL-KHAMY M, LEE J. Image super resolution based on fusing multiple convolution neural networks [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017: 54–61.
- [15] YOUM G Y, BAE S H, KIM M. Image super-resolution based on convolution neural networks using multi-channel input [C] // *2016 IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. IEEE, 2016: 1–5.
- [16] WANG J, WU Y, WANG L, et al. Lightweight feedback convolution neural network for remote sensing images super-resolution [J]. *IEEE Access*, 2021, 9: 15992–16003.
- [17] CHEN J, HE X, CHEN H, et al. Single image super-resolution based on deep learning and gradient transformation [C] // *2016 IEEE 13th International Conference on Signal Processing (ICSP)*. IEEE, 2016: 663–667.
- [18] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution [C] // *European Conference on Computer Vision*. Springer, Cham, 2014: 184–199.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778.
- [20] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1646–1654.
- [21] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 2472–2481.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132–7141.
- [23] ZHANG Y, LI K. Image super-resolution using very deep residual channel attention networks [C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 286–301.
- [24] ZHANG X, ZHENG Z, ASANUMA I, et al. A new kind of super-resolution reconstruction algorithm based on the ICM and the bicubic interpolation [J]. *Information*, 2008, 16(11): 8027–8036.
- [25] SINGARIMBUN R N, NABABAN E B, SITOMPUL O S. Adaptive moment estimation to minimize square error in backpropagation algorithm [C] // *2019 International Conference of Computer Science and Information Technology (ICoSNKOM)*. IEEE, 2019: 1–7.
- [26] HUYNH-THU Q, GHANBARI M. Scope of validity of PSNR in image/video quality assessment [J]. *Electronics Letters*, 2008, 44(13): 800–801.
- [27] HORE A, ZIOU D. Image quality metrics: PSNR vs SSIM [C] // *2010 20th International Conference on Pattern Recognition*. IEEE, 2010: 2366–2369.
- [28] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module [C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3–19.