

基于生成对抗网络的图像重建方法改进

史雪杰, 王同喜, 向 华

(长江大学 计算机科学学院, 湖北 荆州 434023)

摘要:针对 SRGAN 图像超分辨率重建算法中存在的生成图像边缘模糊、纹理丢失以及无法有效捕捉不同尺度下细节信息的问题,提出一种改进的多尺度感知三重密集生成对抗网络(MSRTD-GAN)。该算法在生成器部分使用三重密集块(RRTA)替代原有的残差块,增强模型对图像关键细节的提取能力;引入感受野机制(RFB),提高模型对图像不同尺度下细节信息的敏感度;在上采样过程中,采用深度可分离卷积与像素重排技术结合的方法,减少计算成本并有效地保留细节和纹理信息。生成器的目标函数结合五种损失函数,以提高模型的细节恢复能力和生成图像的整体质量。此外,在判别器部分引入金字塔结构,使模型能够在多个尺度上分析图像,更全面地评估图像质量。实验结果表明,MSRTD-GAN 算法重建图像的 PSNR 和 SSIM 指标分别达到 29.36 dB 和 0.863,相比 SRGAN 分别提高 0.68 dB 和 0.015,且重建后的图像边缘轮廓更加明显,纹理细节更为清晰。这些结果证明了该改进算法在提高图像重建的整体质量方面的有效性。

关键词:生成对抗网络;金字塔结构;超分辨率重建;感受野机制;密集残差块;三重注意力机制。

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2025)06-0062-08

doi:10.20165/j.cnki.ISSN1673-629X.2025.0013

Improvement of Image Reconstruction Method Based on Generative Adversarial Network

SHI Xue-jie, WANG Tong-xi, XIANG Hua

(School of Computer Science, Yangtze University, Jingzhou 434023, China)

Abstract: Aiming to address the problems of fuzzy edges, texture loss, and the inability to effectively capture detailed information at different scales in the SRGAN image super-resolution reconstruction algorithm, we propose an improved multi-scale-aware triple dense generative adversarial network (MSRTD-GAN). The algorithm replaces the original residual block in the generator with a triple dense block (RRTA), enhancing the model's ability to extract key image details. It also introduces the receptive field block (RFB), improving the model's sensitivity to multi-scale details. Additionally, in the up-sampling process, a combination of depthwise separable convolutions and pixel rearrangement techniques is adopted to reduce computational cost while effectively pre-serving detail and texture information. The generator's objective function combines five loss functions to enhance detail recovery and overall image quality. Additionally, a pyramid structure is introduced in the discriminator, enabling the model to analyze images at multiple scales and evaluate image quality more comprehensively. Experimental results show that the PSNR and SSIM metrics of the reconstructed images produced by the MSRTD-GAN algorithm reach 29.36 dB and 0.863, respectively, representing improvements of 0.68 dB and 0.015 over SRGAN. The reconstructed images also exhibit more pronounced edge contours and clearer texture details. These results demonstrate the effectiveness of the proposed algorithm in improving overall image reconstruction quality.

Key words: generating adversarial networks; pyramidal structures; super-resolution reconstruction; receptive field block; residual in residual dense block; triple attention

0 引言

在数字图像处理领域,单帧图像超分辨率(SISR)重建技术是将低分辨率图像转换成高分辨率图像的关键方法,并被广泛应用于卫星成像、医疗成像、视频监控等领域。传统的超分辨率方法主要包括插值

法^[1-2]、重建法^[3-4]和基于学习的方法。插值法简单高效,但通常无法恢复高频细节;重建法通过引入图像先验知识来增强生成图像细节,然而其性能极大受限于先验知识的准确性;随着深度学习的迅速发展,基于深度学习的超分辨率方法成为研究热点。超分辨率生成

收稿日期:2024-10-12

修回日期:2025-02-13

作者简介:史雪杰(2000-),女,硕士生,研究方向为图像处理;向 华(1980-),男,讲师,硕导,研究方向为大数据分析、软件架构、演化学习;通信作者:王同喜(1971-),男,副教授,硕导,研究方向为软件架构与组织、数据科学与大数据分析技术。

对抗网络(SRGAN)^[5]作为一种基于残差网络和生成对抗网络(GAN)的深度学习方法,在提升视觉质量方面取得突破。

胡国星^[6]提出的密集残差块(RRDB)改进生成对抗网络,移除批量归一化层以加快计算速度,并通过层间密集连接增强特征提取效果。Xu等人^[7]提出的TE-SAGAN采用联合损失函数,结合L1范数、感知损失和纹理损失,以优化训练并去除伪影,改善图像的边缘和纹理细节。漆梓渊等人^[8]提出的增强型图像轮廓超分辨率生成对抗网络,通过多级残差模块改进,提高图像边缘修复能力和消除伪影的效果。Li等人^[9]提出基于多尺度双注意力的残差密集生成对抗网络(MARDGAN)并用于图像超分辨率重构,该网络通过多分支路径提取图像特征,并结合通道和空间注意力块增强特征重用,提高图像细节与真实感。Li等人^[10]在其研究中引入Charbonnier损失函数替代传统的均方误差(MSE),以减少大误差的不利影响,增强网络训练的稳定性并有效减少超分辨率图像中的伪影。彭晏飞等人^[11]提出的生成网络以金字塔方差池化模块为核心,其通过不同级别的方差池化提取低分辨率图像中的特征,并结合金字塔结构获取多尺度和子区域的上下文信息,从而丰富特征信息量。Xie等人^[12]提出采用RMSProp优化器替代传统基于动量的梯度优化算法,显著提高训练过程的稳定性,并有效解决深度网络训练中的梯度消失问题。

为解决SRGAN在重建具有复杂结构的图像时难以有效恢复图像的高频细节,生成图像不够自然,且训

练过程困难,模型难以收敛的问题,该文提出一种改进的多尺度感知三重密集生成对抗网络(MSRTD-GAN)。该算法在生成器中集成三重注意力机制^[13]、残差密集块(RRDB)^[14]和感受野机制(RFB)^[15],使得模型能够更加聚焦于图像的关键区域,以增强对图像关键细节的捕捉和恢复能力。在上采样过程中结合深度可分离卷积和像素重排技术,有效保持图像的高清晰度细节。在判别器中引入金字塔结构,使模型能够在不同尺度上进行特征分析,以更准确地恢复细节并提高图像质量。此外,通过整合五种损失函数对生成器的目标函数进行优化,以生成更逼真的图像。

为验证该算法的有效性,在标准数据集Set5、Set14和Urban100上进行验证,并将实验结果与现有的几种超分辨率方法进行对比,包括RCAN、RealESRGAN、SRResNet和SRGAN方法。

1 SRGAN网络模型改进

1.1 超分辨率生成对抗网络

超分辨率生成对抗网络(SRGAN)通过生成器和判别器构成的对抗框架,实现图像的超分辨率重构。生成器的任务是将低分辨率图像转换为高分辨率图像,而判别器则用于评估生成图像的逼真程度,以促使生成器产生更高质量的图像。该对抗训练的关键在于实现生成器与判别器之间的纳什平衡^[16],当判别器无法区分生成图像与真实图像时,模型达到最佳状态。SRGAN网络结构如图1所示。

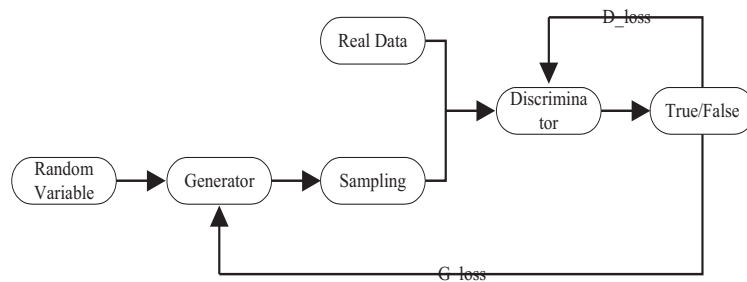


图1 SRGAN网络结构

针对SRGAN算法在重建图像时存在的网络训练困难、边缘模糊、纹理丢失等问题,在SRGAN的基础上进行改进。由于深层网络和多级连接能够增强网络的性能,在生成器中,将传统的残差块替换为残差密集块(RRDB)并集成三重注意力机制,精确调节图像的通道维度 C 和两个空间维度(高度 H 与宽度 W)之间的相互作用,以增强模型对各区域细节的提取能力。同时,引入感受野机制(RFB),增强模型在处理图像的全局特征和局部特征时的能力。在上采样过程中结合深度可分离卷积和PixelShuffle像素重排技术,在减

少计算成本的同时保留图像的纹理信息。在判别器中引入金字塔结构,使判别器能在多个尺度上分析图像,从而更精确地评估图像的真实性和质量。在网络训练方面,使用RMSprop优化器替代原来的Adam优化器,并结合余弦衰减策略,以确保网络训练过程的稳定性。

此外,通过融合内容损失、感知损失、总方差损失、对抗损失和纹理损失,使得训练过程更加全面,能够综合关注到图像的全局一致性、局部细节、纹理质感和整体的视觉效果。

1.2 网络结构改进

1.2.1 生成网络

该文提出的生成网络结构基于 SRGAN 生成器框架。改进包括将 PReLU 激活函数替换为 SiLU 激活函数^[17], 引入 RFB 感受野机制, 将原始的 16 个残差块替

换为 12 个三重密集块 (RRTA), 并采用结合深度可分离卷积和 PixelShuffle 像素重排技术的上采样部分。该结构主要由全局特征提取、局部特征提取和图像重构三部分组成, MSRTD-GAN 生成网络结构如图 2 所示。

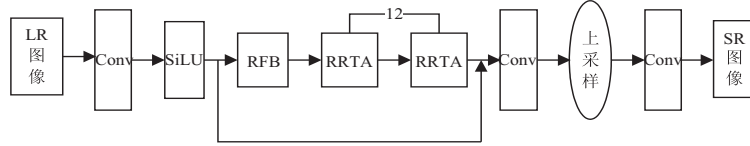


图 2 MSRTD-GAN 生成网络结构

生成网络获取 LR 图像后, 首先通过一个 9×9 卷积层 Conv 提取全局特征, 采用结合 ReLU (Rectified Linear Unit) 和 Sigmoid 函数特点的 SiLU 激活函数以允许网络根据特征的具体需求自动调节激活强度, 接着通过一个 RFB 感受野机制和 12 个 RRRTA 块提取局部特征, 最后使用结合深度可分离卷积和 PixelShuffle 像素重排技术^[18]的上采样机制实现图像重构来生成 SR 图像。整个生成网络采用 SiLU 激活函数替代传统的 ReLU 函数, 为细节恢复和增强提供更加自适应的非线性激活方式。

为降低生成图像中的伪像并提高泛化性, 移除生成网络中的所有批归一化层 (BN)^[19]。引入的感受野机制 (RFB) 是一个多尺度特征提取机制, 其核心在于通过模拟不同感受野的卷积操作, 以增强模型对不同尺度特征的感知能力。RFB 包含多个分支, 每个分支使用不同的卷积核大小和填充策略来处理不同尺度的信息, 使模型能够同时处理图像的局部细节和全局特征, 提升生成器对图像整体结构和细微特征的感知能力, 从而在复杂场景下实现更精确的图像重建。RFB 结构如图 3 所示。

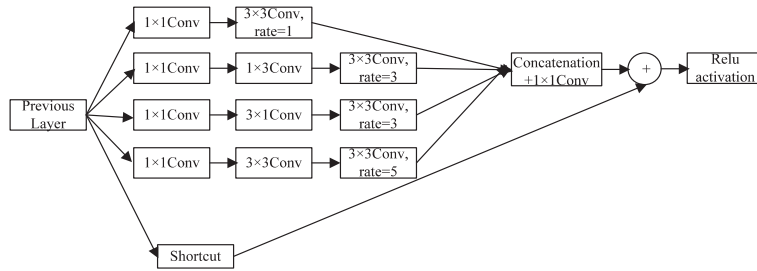


图 3 RFB 结构

每个 RRRTA 块由一个密集残差块 (Residual in Residual Dense Block) 结构和一个三重注意力机制 (Triplet Attention) 组成。RRDB 利用其密集连接^[20]来增强特征的传递能力, 同时通过残差连接^[21]防止梯度消失问题的发生。在 RRDB 中, 权重因子 β (默认值为 0.2) 用于调节残差操作的强度, 有助于控制信息流动的强度, 从而避免训练过程中出现数值不稳定的情况。Triplet Attention 结构由三个独立分支组成, 分别对输入张量 $\chi \in R^{C \times H \times W}$ 进行处理, 以优化输出特征。在第一个分支中, 张量 χ 沿高度轴 (H 轴) 逆时针旋转 90° , 旋转后张量 $\hat{\chi}_1$ 维度为 $(W \times H \times C)$, 经过 Z-pool 被缩减成 $\hat{\chi}_1^*$, 此时维度为 $(2 \times H \times C)$, 再经过 7×7 的卷积层和批量归一化处理输出维度为 $(1 \times H \times C)$, 再使用 sigmoid 激活函数生成注意力权重, 生成的注意力权重应用到 $\hat{\chi}_1$, 最后顺时针旋转 90° 得到 χ 原始输入维度。第二个分支的处理方式与第一个分支类似, 但沿宽度轴 (W 轴) 进行, 最后生成的注意力权重应用到 $\hat{\chi}_2$, 并沿着 W 轴顺时针旋转 90° 得到 χ 原始输

入维度。第三个分支则集中处理通道维度, 通过 Z-pool 和一个核大小为 k 的卷积层以及批量归一化, 同样使用 sigmoid 激活函数生成影响输出张量形状维度为 $(1 \times H \times C)$ 的注意力权重, 应用到 χ 。最后, 将三个分支的结果通过简单平均法聚合, 得到最终的输出张量。Z-pool 的数学公式如式 1 所示, 从输入张量的 Triplet Attention 中获得的细粒度注意力张量 y 的过程如式 2 所示。RRDB 结构如图 4 所示, Triplet Attention 结构如图 5 所示。

$$Z - pool(x) = [MaxPool_{0d}(x), AvgPool_{0d}(x)] \quad (1)$$

其中, $0d$ 是最大池化操作和平均池化操作的第 0 维度。

$$y = \frac{1}{3} (\hat{\chi}_1 \sigma(\psi_1(\hat{\chi}_1^*)) + \hat{\chi}_2 \sigma(\psi_2(\hat{\chi}_2^*)) + \chi \sigma(\psi_3(\hat{\chi}_3))) \quad (2)$$

其中, σ 表示 sigmoid, ψ_1, ψ_2, ψ_3 代表 triplet attention 中三个分支的卷积核为 k 的标准二维卷积层。

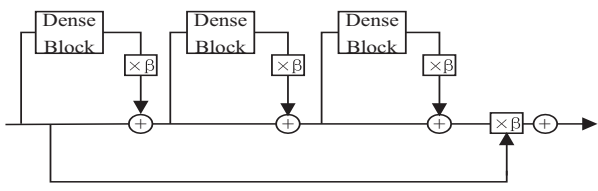


图 4 RRDB 结构

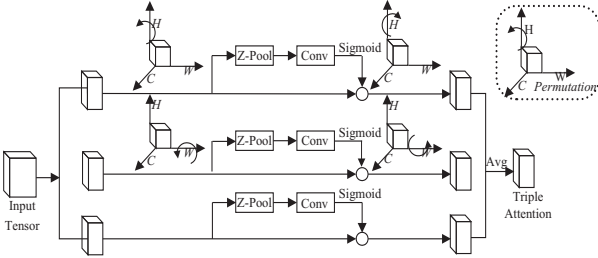


图 5 Triplet Attention 结构

上采样机制结合深度可分离卷积和像素重排技术 (PixelShuffle)。深度可分离卷积通过将标准卷积分解为深度卷积和点卷积,从而大幅降低计算量和参数数量;PixelShuffle 技术通过重排特征图的通道维度来实现高效的上采样。两者的结合使生成器可以在增加图像分辨率的同时保持细节和纹理的完整性。上采样结构如图 6 所示。

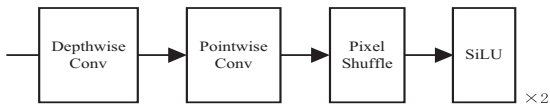


图 6 上采样结构

1.2.2 判别网络

金字塔结构是一种广泛应用于计算机视觉中的多

尺度图像处理方法,其通过对输入图像逐层降采样,生成一系列不同分辨率的图像。传统判别器仅依赖单一路径进行特征提取,难以充分捕捉多尺度图像结构的细节,导致细节信息融合不足,判别器的特征表达能力差。为解决这一问题,该文对判别器进行改进,采用金字塔结构并引入三个并行路径,以捕捉不同尺度的图像特征,并实现多尺度特征融合,增强判别器的特征判定能力。具体通过以下三个路径处理图像以实现多尺度分析:原始分辨率路径,先使用一个 3×3 卷积层和 LeakyReLU 激活函数提取初始特征,再经过两个 DownsampleBlock 结构实现连续降采样和特征提取。此过程中,采用 BatchNorm2d 进行批量归一化处理,以提升训练速度和网络稳定性,最后通过全局平均池化 (AdaptiveAvgPool2d) 将整个特征图缩减为单一特征向量;二分之一分辨率路径,通过平均池化减少至原始分辨率的二分之一,随后执行与第一条路径类似的特征提取过程,获取该尺度下的图像特征;四分之一分辨率路径,以平均池化减少至原始分辨率的四分之一作为起点,捕获更宏观的图像信息。三个路径的特征向量在经过全连接层处理后,通过 Sigmoid 激活函数映射为 0 到 1 之间的概率值,表示图像为真实图像的可能性。通过采用金字塔结构,使判别器能够在不同分辨率层次上捕捉图像特征,从局部细节到整体布局进行全面分析,从而显著提升对图像质量和真实性的判断能力,并增强模型的鲁棒性。MSRTD-GAN 判别网络结构如图 7 所示, DownsampleBlock 结构如图 8 所示。

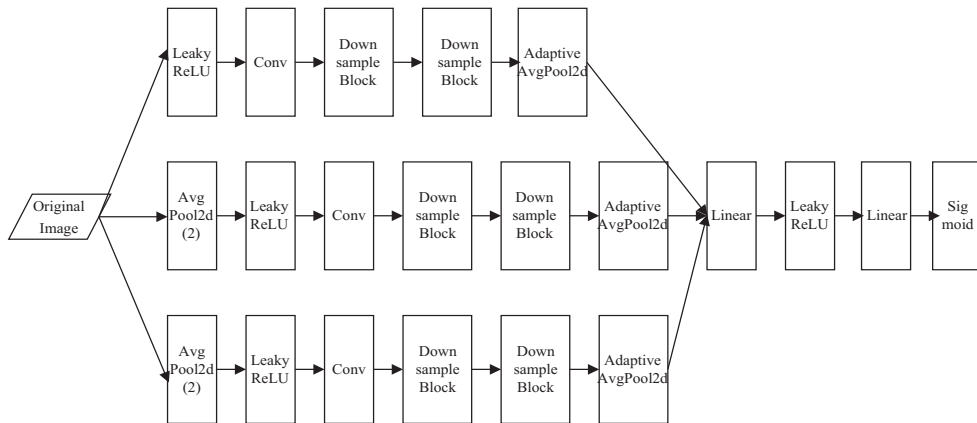


图 7 MSRTD-GAN 判别网络结构

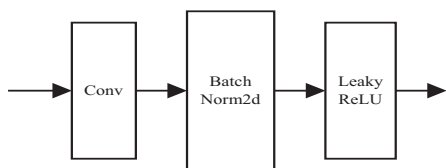


图 8 DownsampleBlock 结构

1.3 损失函数

该 MSRTD-GAN 网络的目标函数由两部分构成:

生成器损失和判别器损失。判别器损失如式 3 所示。

$$L_D = \frac{1}{2} (-\log \sigma(D(I^{HR})) - \log(1 - \sigma(D(G(I^R)))))) \quad (3)$$

生成器损失如式 4 所示。

$$L_G = 1.7 \times L_{con} + 1e - 2 \times L_{per} + 2e - 5 \times L_{tv} + 1e - 3 \times L_{adv} + 8e - 3 \times L_{tex} \quad (4)$$

其中,内容损失 L_{con} 、感知损失 L_{per} 、总方差损失 L_{tv} 和

纹理损失 L_{tex} 采用 Charbonnier loss^[22] 代替 MSE loss 作为损失函数。Charbonnier loss 是 L1 损失的一种变体, 其中增加一个正则化项 ϵ (设为 $1e-6$), 以确保数值稳

$$L_{\text{con}} = \sqrt{\frac{1}{CHW} \sum_{c=1}^C \sum_{h=1}^H \sum_{w=1}^W (I_{c,h,w}^{\text{HR}} - G(I^{\text{LR}})_{c,h,w})^2 + \epsilon^2} \quad (5)$$

其中, C 、 H 、 W 是图像的通道数、高度和宽度。通过应用 Charbonnier 损失, 减少生成图像与目标图像在像素级的偏差, 确保高频和低频信息的有效传递, 从而实

$$L_{\text{per}} = \sqrt{\frac{1}{CHW} \sum_{c=1}^C \sum_{h=1}^H \sum_{w=1}^W (F_{\text{VGG}}(I^{\text{HR}})_{c,h,w} - F_{\text{VGG}}(G(I^{\text{LR}}))_{c,h,w})^2 + \epsilon^2} \quad (6)$$

其中, F^{VGG} 表示通过 VGG 网络提取的特征, 感知损失通过利用 F^{VGG} 来评估图像, 侧重于较高层次的视觉特征, 关注于保持图像的整体结构和纹理信息, 主要影响结构相似度 (SSIM)。

$$L_{\text{iv}} = \sum_{i,j} \sqrt{(x_{i,j+1} - x_{i,j})^2 + (x_{i+1,j} - x_{i,j})^2} \quad (7)$$

其中, x 表示图像, 而 $x_{i,j}$ 表示图像在第 i 行和第 j 列的像素值。总变差损失帮助生成的图像保持结构上的连

$$L_{\text{tex}} = \sqrt{\| \text{Gram}(F_{\text{VGG}}(G(I^{\text{LR}}))) - \text{Gram}(F_{\text{VGG}}(I^{\text{HR}})) \|^2 + \epsilon^2} \quad (9)$$

其中, $\| \cdot \|^2$ 表示平方 L2 范数, Gram 矩阵的计算公式为: $\text{Gram}(F) = F \cdot F^T$, Gram 矩阵是特征向量的外积, 其每个元素反映不同特征通道之间的相关性, 从而有效捕捉图像的纹理信息。

2 实验与结果分析

2.1 实验设置

2.1.1 实验环境

实验环境为 NVIDIA 显卡 GeForce RTX 4060 Laptop GPU、13th Gen Intel(R) Core(TM) i5-13500H 2.60 GHz、16.0 GB RAM, Python 版本为 Python 3.8.18。本次训练使用 RMSprop 优化器对模型参数进行优化, $\alpha = 0.95$, $\text{eps} = 1e-8$, $\text{weight_decay} = 0$, $\text{momentum} = 0.1$, 训练过程中生成器学习率设置为 $5e-4$, 判别器学习率设置为 $3e-4$, 并使用余弦衰减策略逐步降低学习率, Batch_size 设置为 4, 总共进行 50 轮训练。

2.1.2 数据集

在四个数据集上进行实验, 包括 DIV2K、Set5、Set14 和 Urban100。DIV2K 数据集由 1 000 张 2K 分辨率的高清图像组成, 其中 800 张用于训练, 200 张用于验证。为增强训练集, 采用随机水平翻转、旋转和中心裁剪等数据扩增技术, 最终从中随机选取 6 271 张作为实验的训练样本。在常用的基准数据集 Set5、Set14 以及 Urban100 上进行测试实验, 设置高分辨率 (HR) 图像与低分辨率 (LR) 图像之间的下采样因子为 4, 高分辨率图像被裁剪为 96×96 像素的图像块进行处理。

定性。内容损失 L_{con} 、感知损失 L_{per} 、总方差损失 L_{iv} 、对抗损失 L_{adv} 、纹理损失 L_{tex} 分别如式 5~9 所示。

现在保证图像清晰度和纹理细节的同时提高重建图像的峰值信噪比 (PSNR)。

贯性和平滑性, 避免在增强过程中产生不自然的纹理和边缘。

$$L_{\text{adv}} = \sum_{n=1}^N -\log D(G(I^{\text{LR}})) \quad (8)$$

对抗损失通过优化生成器使其产生的图像在视觉上无法被判别器区分出与真实图像的差异, 不仅提升生成图像的逼真度, 还有助于提高 PSNR 和 SSIM。

2.1.3 评价指标

为评估改进算法 MSRTD-GAN 的性能, 使用峰值信噪比 (PSNR) 和结构相似度 (SSIM) 作为客观评价指标来衡量图像超分辨率重建的质量。PSNR 主要用于评估图像在像素级别上的失真程度, 而 SSIM 则从亮度、对比度和结构三个方面综合衡量图像的相似性。较高的 PSNR 和 SSIM 值表明图像重建的质量更接近于原始的高分辨率图像。相关的计算公式分别如式 10 和式 11 所示。

$$\text{PSNR} =$$

$$10 \log_{10} \frac{255^2 \times W \times H \times C}{\sum_{i=1}^W \sum_{j=1}^H \sum_{z=1}^C [\bar{x}(i,j) - x(i,j)]^2 + 1.0 \times 10^{-9}} \quad (10)$$

$$\text{SSIM}(X, Y) = \frac{(2\mu_X \mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (11)$$

2.2 实验结果

2.2.1 消融实验

为验证改进生成器和判别器对 SRGAN 的影响, 选择使用标准数据集 Set5 作为测试集设计以下消融实验, 实验结果如表 1 所示。

(1) 在生成器中引入 RRTA 和 RFB, 采用改进的 SRGAN 算法进行图像超分辨率重建。

(2) 在判别器中引入金字塔结构, 采用改进的 SRGAN 算法进行图像超分辨率重建。

(3) 在生成器中引入 RRTA 和 RFB, 在判别器中

引入金字塔结构,采用改进的 SRGAN 算法进行图像超分辨率重建。

表 1 测试集的 PSNR(dB)、SSIM 比较

Evaluation metric	SRGAN	SRGAN+RRTA +RFB	SRGAN+ Pyramid	MSRTD-GAN
PSNR	28.68	29.21	28.78	29.36
SSIM	0.848	0.859	0.853	0.863

从实验结果可以看出,在生成器中引入 RRTA 和 RFB 后,相比原始 SRGAN,在 PSNR 指标提高 0.53 dB,SSIM 指标提高 0.011。在判别器引入金字塔结构后,PSNR 提高 0.1 dB,SSIM 提高约 0.005。综合两种方法后,PSNR 和 SSIM 指标分别提高 0.68 dB 和 0.015,这表明生成图像在保留纹理细节和提高清晰度方面有明显的提升。

2.2.2 对比实验

为证明该文提出的方法的有效性,将 MSRTD-GAN 与 RCAN^[23]、RealESRGAN^[24]、SRResNet、SRGAN 算法在 Set5、Set14、Urban100 标准数据集上进行对比,结果如表 2、表 3 所示。

表 2 测试集的 PSNR 比较 dB

Test Set	RCAN	Real-ESRGAN	SRResNet	SRGAN	MSRTD-GAN
Set5	26.89	25.82	28.24	28.68	29.36
Set14	24.49	23.20	25.27	25.61	26.12
Urban100	22.09	20.61	22.69	23.22	23.63

表 3 测试集的 SSIM 比较

Test Set	RCAN	Real-ESRGAN	SRResNet	SRGAN	MSRTD-GAN
Set5	0.784	0.766	0.832	0.848	0.863
Set14	0.691	0.646	0.726	0.740	0.758
Urban100	0.667	0.643	0.703	0.731	0.749

从实验结果可以看出,MSRTD-GAN 算法在 Set5、Set14、Urban100 标准数据集上的 PSNR 和 SSIM 值均优于 RCAN、RealESRGAN、SRResNet、SRGAN 算法。这些结果证明,与其他算法相比 MSRTD-GAN 算法能更有效地实现图像的超分辨率,最终 PSNR 和 SSIM 值分别为 29.36 dB 和 0.863。

2.2.3 图像重建效果对比

接着,从 Set5 和 Set14 数据集中各挑选出一幅图像,评估 RCAN、RealESRGAN、SRResNet、SRGAN 和 MSRTD-GAN 五种算法在图像重建方面的性能。相关的对比结果如图 9 和图 10 所示。具体分析如下,RCAN 算法通过级联通道注意力模块增强图像重建能力,但对复杂场景中的高频细节仍显不足,生成图像在

边缘区域出现模糊,放大的局部细节也存在锐度不足的问题,对应图 9(b)的 PSNR 值、SSIM 值分别为 21.25 dB 0.718;RealESRGAN 网络结构虽然引入生成对抗的思路,但其网络深度有限,对细节的恢复能力不足。生成图像在复杂纹理区域显得不自然,无法准确反映原始高分辨率图像的真实特性,对应图 9(c)的 PSNR 值、SSIM 值分别为 19.82 dB、0.766;SRResNet 采用残差网络结构,在一定程度上提高细节恢复能力,但生成图像放大的局部细节图仍然表现出过渡平滑的现象,部分细节信息同样丢失,对应图 9(d)的 PSNR 值、SSIM 值分别为 22.93 dB、0.795;SRGAN 通过对抗训练生成高分辨率图像,但由于生成器中的传统残差块无法充分利用网络的层级信息,导致对复杂细节的捕捉能力不足,判别器也难以在多个尺度上全面评估图像的真实性和质量,导致生成图像在某些区域出现不自然的纹理和伪影,对应图 9(e)的 PSNR 值、SSIM 值分别为 23.62 dB、0.840;MSRTD-GAN 算法则通过在生成器中引入更深层次的网络结构和 RFB 感受野机制,更有效地提取图像的关键细节,并结合更有效的损失函数作为生成器的目标函数来提高模型细节恢复的精确性和图像质量,减少不自然的纹理。在判别器中引入金字塔结构,实现多尺度分析并缓解梯度消失,以增强模型对图像质量真实性的判断能力,并使用 RMSprop 优化器并结合余弦衰退策略调整学习率提供更稳定的训练过程,对应图 9(f)的 PSNR 值、SSIM 值分别为 25.08 dB、0.871。

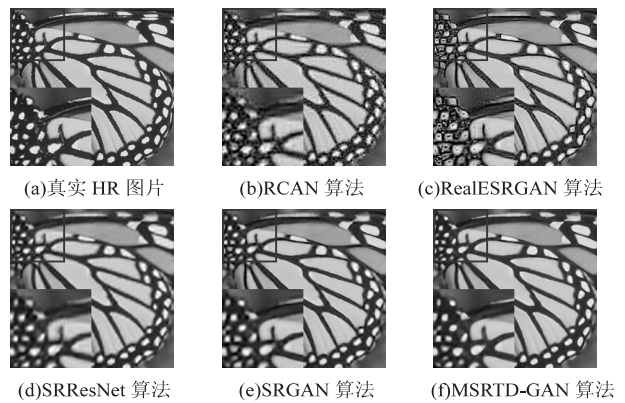


图 9 不同算法在 Set5 数据集上的图像重建效果

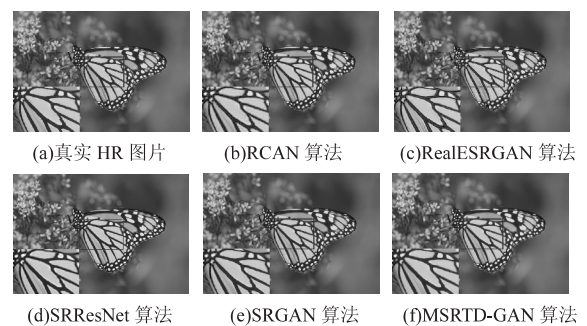


图 10 不同算法在 Set14 数据集上的图像重建效果

由图 9 可知,MSRTD-GAN 算法生成的图像边缘明显、细节纹理清晰,放大的细节图显示的蝴蝶纹理更加逼真。图 10 中的放大细节同样显示 MSRTD-GAN 算法在纹理细节和亮度信息的表现上更为精确。改进 MSRTD-GAN 算法在图像重建效果上优于 RCAN、RealESRGAN、SRResNet 和 SRGAN 算法。

2.3 实验分析

对于 MSRTD-GAN,其总目标函数是生成器和判别器之间的博弈过程,可以形式化如式 12 所示。

$$\min_c \max_D V(G, D) = E_{I^{HR} \sim p_{\text{real}}(I^{HR})} [\log D(I^{HR})] + E_{I^{LR} \sim p_c(I^{LR})} [\log(1 - D(G(I^{LR}))) \quad (12)$$

其中,判别器 D 的目标是最大化其对真实高分辨率图像 I^{HR} 的识别概率,以提高识别准确性,同时最小化对生成器 G 产生的图像 $G(I^{LR})$ 的误判概率。生成器 G 的目标是生成足够逼真的图像 $G(I^{LR})$,最小化判别器 D 在其生成图像上的判断准确性,使得 $\log(1 - D(G(I^{LR})))$ 尽可能大。

对生成器 G 的核心改进包括引入三重密集块 (RRTA)、感受野块 (RFB) 和结合深度可分离卷积与像素重排技术的上采样方法,以及调整生成器损失函数的权重比例。其中,RRTA 增强模型对图像关键细节的提取能力,优化感知损失 L_{per} 和纹理损失 L_{tex} ,直接影响生成图像的视觉质量和细节再现,提升 SSIM。RFB 通过调整其结构来适应图像的不同尺度,增强生成图像的尺度适应性和细节敏感度,影响感知损失 L_{per} 和纹理损失 L_{tex} ,提升生成图像的尺度适应性和纹理细节的准确性。此外,深度可分离卷积与像素重排技术的组合技术在上采样过程中有效减少计算成本,同时保留关键的细节和纹理信息,优化总变差损失 L_{tv} ,改善图像平滑度,间接提升 PSNR。对判别器 D 的核心改进是通过引入金字塔结构,使模型能在多个尺度上分析图像,从而在判定生成图像 $G(I^{LR})$ 的真实性时更为精确。又因为生成器 G 在对抗损失 L_{adv} 的驱动下,必须生成更高质量的图像以欺骗判别器,从而推动整个 MSRTD-GAN 的性能提升。

3 结束语

针对 SRGAN 重建算法存在生成图像边缘模糊和纹理丢失,以及无法有效捕捉不同尺度下细节信息的问题,提出一种改进算法 MSRTD-GAN。

通过在 Set5、Set14、Urban100 等标准数据集上的实验结果表明,MSRTD-GAN 算法在 PSNR 和 SSIM 指标上均高于 RCAN、RealESRGAN、SRResNet 和 SRGAN 算法,且重建图像在边缘清晰度和纹理细节方面也表现更优。具体得出以下结论:

(1)在生成器中引入感受野机制 (RFB),并在 RRDB 集成三重注意力机制,可以增强模型对图像的细节处理能力,提高重建图像的质量和真实感;

(2)引入金字塔结构的判别器,可以提高模型评估图像的准确性,从而提升整体图像重建效果;

(3)在训练过程中融合内容损失、感知损失、总方差损失、对抗损失、纹理损失,可以使网络能够在多个维度优化生成图像的质量。

实验结果验证了 MSRTD-GAN 算法在整体图像质量、边缘清晰度以及纹理细节恢复方面的有效性。

未来的研究工作将集中于以下两方面:

(1)进一步优化对抗训练的稳定性,减少训练过程中可能出现的不稳定现象;

(2)优化金字塔结构的判别器设计,以提高计算效率,降低资源消耗,同时保持高质量的图像生成效果。

参考文献:

- [1] KEYS R. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6):1153-1160.
- [2] ZHANG Lei, WU Xiaolin. An edge-guided image interpolation algorithm via directional filtering and data fusion[J]. IEEE Transactions on Image Processing, 2006, 15(8):2226-2238.
- [3] NAYAK R, HARSHAVARDHAN S, PATRA D. Morphology based iterative back-projection for super-resolution reconstruction of image [C]//2014 2nd international conference on emerging technology trends in electronics, communication and networking. Surat: IEEE, 2014:1-6.
- [4] STARK H, OSKOU P. High-resolution image recovery from image-plane arrays, using convex projections[J]. J Opt Soc Am A, 1989, 6(11):1715-1726.
- [5] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu: IEEE, 2017:105-114.
- [6] 胡国星. 一种基于图像超分辨率技术的图像质量提升技术 [J]. 中国科技信息, 2023(4):112-115.
- [7] XU Y Y, LUO W, HU A N, et al. TE-SAGAN: an improved generative adversarial network for remote sensing super-resolution images[J]. Remote Sensing, 2022, 14(10):2425.
- [8] 漆梓渊, 吴浩, 陈明举, 等. 基于聚合多尺度特征的图像轮廓增强超分辨率重建算法 [J]. 国外电子测量技术, 2023, 42(7):50-58.
- [9] LI H A, WANG D, ZHANG J, et al. Image super-resolution reconstruction based on multi-scale dual-attention [J]. Connection Science, 2023, 35(1):2182487.
- [10] LI W, LIU Z, DENG H. A self-attention based SRGAN for

- super-resolution of astronomical image[C]//2022 IEEE 8th international conference on computer and communications (ICCC). Chengdu:IEEE,2022:1977-1981.
- [11] 彭晏飞,李泳欣,孟欣,等. 金字塔方差池化网络的图像超分辨率重建[J]. 液晶与显示,2024,39(10):1380-1390.
- [12] XIE Z,ZHANG N. Improved algorithm for GAN super-resolution face image reconstruction[C]//2023 IEEE 5th international conference on civil aviation safety and information technology (ICCASIT). Dali:IEEE,2023:285-289.
- [13] MISRA D,NALAMADA T,ARASANIPALAI A U, et al. Rotate to attend;convolutional triplet attention module[C]//2021 IEEE winter conference on applications of computer vision (WACV). Waikoloa:IEEE,2021:3138-3147.
- [14] WANG X,YU K,WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks [C]//Computer vision - ECCV 2018 workshops. Cham:Springer,2019:63-79.
- [15] LIU S,HUANG D,WANG Y. Receptive field block net for accurate and fast object detection[C]//Computer vision - ECCV 2018. Cham:Springer,2018:404-419.
- [16] RATLIFF L J,BURDEN S A,SASTRY S S. Characterization and computation of local Nash equilibria in continuous games [C]//2013 51st annual Allerton conference on communication, control, and computing (Allerton). Monticello: IEEE, 2013:917-924.
- [17] RAMACHANDRAN P,ZOPH B,LE Q V J A N, et al. Swish: a self-gated activation function [J]. arXiv: 1710.05941,2017.
- [18] SHI W,CABALLERO J,HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas:IEEE,2016:1874-1883.
- [19] LIM B,SON S,KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW). Honolulu:IEEE,2017:1132-1140.
- [20] HUANG G,LIU Z,MAATEN L V D, et al. Densely connected convolutional networks[C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu:IEEE,2017:2261-2269.
- [21] HE K,ZHANG X,REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2016:770-778.
- [22] TIMOFTE R,AGUSTSSON E,GOOL L V, et al. NTIRE 2017 challenge on single image super-resolution: methods and results[C]//2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW). Honolulu: IEEE,2017:1110-1121.
- [23] ZHANG Y,LI K,LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//Computer vision - ECCV 2018. Cham:Springer,2018:294-310.
- [24] WANG X,XIE L,DONG C, et al. Real-ESRGAN: training real-world blind super-resolution with pure synthetic data [C]//2021 IEEE/CVF international conference on computer vision workshops (ICCVW). Montrea: IEEE, 2021: 1905 - 1914.