

## 生成对抗网络对 OCT 视网膜图像的超分辨率重建

柯舒婷<sup>1</sup>, 陈明惠<sup>1\*</sup>, 郑泽希<sup>2</sup>, 袁媛<sup>1</sup>, 王腾<sup>1</sup>, 何龙喜<sup>1</sup>, 吕林杰<sup>1</sup>, 孙好<sup>1</sup><sup>1</sup>上海理工大学健康科学与工程学院, 上海介入医疗器械工程技术研究中心, 教育部医学光学工程中心, 上海 200093;<sup>2</sup>上海理工大学机械工程学院, 上海 200093

**摘要** 光学相干层析成像(OCT)的质量通常会受到固有散斑噪声和低采样率的影响。为了在短扫描时间内获得高信噪比和高分辨率的 OCT 图像,本文提出了一种改进的 OCT 图像超分辨率重建网络模型 PPECA-SRGAN。该模型将生成对抗网络(GAN)作为基础结构,可以不依赖配对数据集进行训练。在该模型的生成器残差块之间添加了金字塔注意力模块 PANet,同时在判别器中加入了本文新提出的 PECA 模块,使其更加注重捕捉图像细节,提升模型对图像边缘纹理的重建能力。将所提 PPECA-SRGAN 模型在 OCT 图像数据集上进行实验,得到的峰值信噪比和结构相似性指标的平均值较当前三种经典模型的平均值分别约提高了 3.5% 和 5.6%。实验结果表明,所提模型在鲁棒性和 OCT 图像细节重建方面较经典模型有较大提升。

**关键词** 生物光学; 光学相干层析成像; 超分辨率; 生成对抗网络; 无配对图像

中图分类号 TP391

文献标志码 A

DOI: 10.3788/CJL202249.1507203

## 1 引言

光学相干层析成像(OCT)是获取眼部图像的主要技术手段之一,在临床诊断中得到了广泛应用。然而,光学相干断层扫描设备的成像能力会受到两种因素的影响:一是成像系统固有的干涉性质,它会降低图像的对比度,模糊其精细的结构特征;二是 OCT 图像采集过程中无意识的眼球运动、头部运动和呼吸等,都可能导致得到的图像缺乏高频纹理细节,在视觉质量方面达不到满意的效果。在临床诊断中,使用较低的空间采样率是实现大视野并减少无意识微颤动影响的一种方式,因此,大多数 OCT 图像在信噪比和空间采样率方面都不是最优的<sup>[1]</sup>。为了获得高信噪比和高分辨率的 B 扫描 OCT 图像,实现图像超分辨率重建是非常有必要的。

通常将低分辨率(LR)图像进行一系列变换得到高分辨率(HR)图像的过程称为图像超分辨率重建。基于深度学习的超分辨率(SR)重建是近年来的研究热点,越来越受到医学视觉界的广泛关注,并且在医学图像重建等领域具有广阔的应用前景<sup>[2]</sup>。Dong 等<sup>[3]</sup>在图像超分辨率方面提出了基于三层卷积的端到端的 SRCNN 模型。Ma 等<sup>[4]</sup>提出了一种边缘敏感的条件生成对抗网络(CGAN),并用该网络对 OCT 图像进行去噪处理,提高了图像的清晰度。Qiu 等<sup>[5]</sup>提出了一种具有感知敏感的损失函数的深度神经网络,并用其来降低 OCT 图像中的噪声。Huang 等<sup>[6]</sup>提出了一种生成对抗网络,并利用该网络同时实现了图像去噪和 OCT 图像超分辨率重建。上述这些基于监督学习的方法面临的巨大挑战之一是它们通常需要大规模成对的训练数据集。Ledig 等<sup>[7]</sup>提出了一种基于生成对抗网络超分辨率重建的经典 SRGAN 模型,该模型可以不使用成对的图像数据集<sup>[8]</sup>。最近,针对 OCT 图像的超分辨率重建,Das 等<sup>[9]</sup>提出了一种基于深度学习的方法——SDSR-OCT,该方法是在只有 10 对 LR-HR 图像的小数据集上训练的。实际上,眼疾患者视网膜层的病变区域与正常区域在纹理和亮度上存在显著差异,而且病变区域的形状和大小也有很大差异,上述模型无法可靠地重建病理性结构。

为了满足临床需求,解决图像超分辨率重建方法依赖配对数据集训练且重建图像细节不够清晰的问题,本文提出了一个改进的无配对图像<sup>[10]</sup>的超分辨率重建模型 PPECA-SRGAN。该模型将可以细化特征、丰富多尺度信息的注意力机制<sup>[11]</sup>集成到 SRGAN 架构中,并在生成器中间的两个残差块之间添加了金字塔注意力模块 PANet<sup>[12]</sup>,多尺度特征金字塔可以提取长距离与短距离的特征关系,捕获图像在不同尺寸中的信息。同时,本文还提出了一种改进的注意力机制模块 PECANet(用 ECANet<sup>[13]</sup>替换 PSANet 中的 SENet<sup>[14]</sup>

收稿日期: 2021-12-14; 修回日期: 2022-01-10; 录用日期: 2022-01-14

基金项目: 上海市科委产学研医项目(15DZ1940400)

通信作者: \*cmhui.43@163.com

得到),并将该模块添加到模型的判别器中,利用 ECANet 不降维的局部跨信道交互策略和自适应选择一维卷积核大小的性质学习有效的通道信息。最后,在两组 OCT 视网膜图像数据集上,将所提模型与其他模型重建图像的峰值信噪比(PSNR)和结构相似性(SSIM)指标进行了对比,并对它们重建图像的质量进行了评估。

## 2 实验方法与原理

本研究的 OCT 图像超分辨率重建分为 4 个步骤:

1) 图像数据集的处理。对数据集中的 OCT 图像进行初步筛选与裁剪,保留 OCT 图像中的主要特征信息。

2) 数据集划分,即将处理后的 OCT 图像划分为训练集与测试集。

3) 模型训练。把划分好的 OCT 图像训练集导入模型的预处理模块中扩充图像数据,再进行 4 倍下采

样处理得到 LR 图像,然后训练模块通过生成器将 LR 图像生成 SR 图像,并利用生成器和判别器的博弈来提升生成器的性能。

4) 超分辨率重建测试。利用训练好的生成器对测试集中的图像进行超分辨率重建,通过图像质量指标来评估模型的性能。

为了在 OCT 临床数据集上实现超分辨率重建,本文提出了 PPECA-SRGAN 网络。该网络模型以生成器和判别器组成的超分辨率生成对抗网络(SRGAN)为框架<sup>[15]</sup>。首先,在生成器中间的两个残差块之间添加 PANet 模块<sup>[16]</sup>,利用其细化特征的特点增加对 OCT 视网膜图像特征的提取能力;其次,在判别器中加入 PECA 模块,该模块是对 PSANet(Pyramid Split Attention Network)的改进,它能够充分捕获多尺度特征图的空间信息,并能实现图像跨维度通道特征的交互。本文网络架构如图 1 所示,生成器中的黄色模块为 PANet,判别器中的橘粉色模块为 PECA。

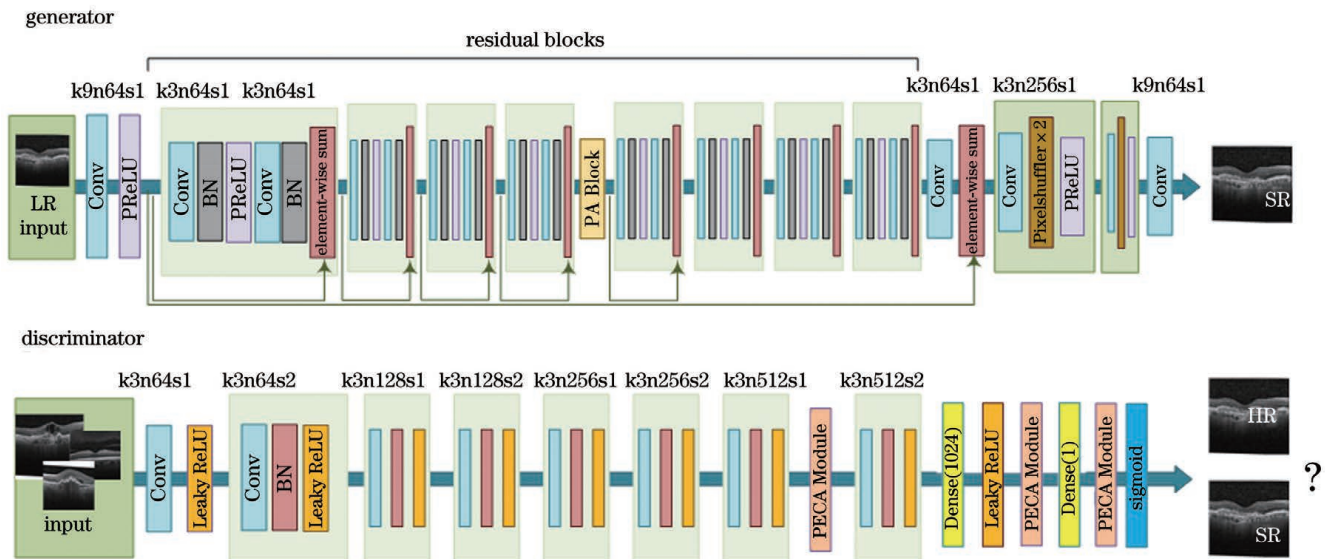


图 1 PPECA-SRGAN 网络结构图

Fig. 1 Structure of PPECA-SRGAN

在本网络中,LR 图像是生成器的输入。先用卷积层提取输入图像的特征,随后将特征映射输入到非线性映射的残差块模型中,最后通过上采样层和卷积层重建图像,得到 SR 图像。接下来,将生成的 SR 图像和原始图像(HR 图像)都输入到判别器中判断图像是 SR 图像还是 HR 图像,同时把 SR 图像和 HR 图像的差异反馈给生成器学习,尽可能地使生成的 SR 图像拟合 HR 图像数据的分布空间。利用这种 GAN 网络模型进行训练,直至判别器无法判断 SR 图像和 HR 图像的区别,此时生成器可以生成逼真的高分辨率图像。

### 2.1 多尺度特征提取的金字塔注意力

在本文的模型中,将 PANet 加入到生成器网络中间的两个残差块之间,这样能够从多尺度特征金字塔中提取到丰富的特征关系,生成更精细的图像。

PANet 的结构如图 2 所示。PANet 是一种自顶向下和自底向上的双向融合骨干网络模型,它包含自适应特征池化和全连接融合两个模块,其中:自适应特征池化模块可以用于聚合不同层之间的特征,保证特征的完整性和多样性;全连接融合模块可以得到更加准确的特征提取。由于降采样能够有效减小压缩伪影等图像噪声,因此,PANet 的金字塔结构可以利用不同采样倍数的特征图来相互传递注意力信号,以更灵活的方式得到不同尺寸特征之间的信息<sup>[17]</sup>。

PANet 分为两个部分,即金字塔采样和 S-A Attention。金字塔采样部分的作用是对图像进行降采样处理,S-A Attention 部分则是构建 O、P、Q 三种特征图,以捕获图像在不同尺寸下的信息,然后再将 O 特征图和 P 特征图作为卷积核与 Q 特征图进行卷积和反卷积操作。PANet 充分利用了特征融合,它的应

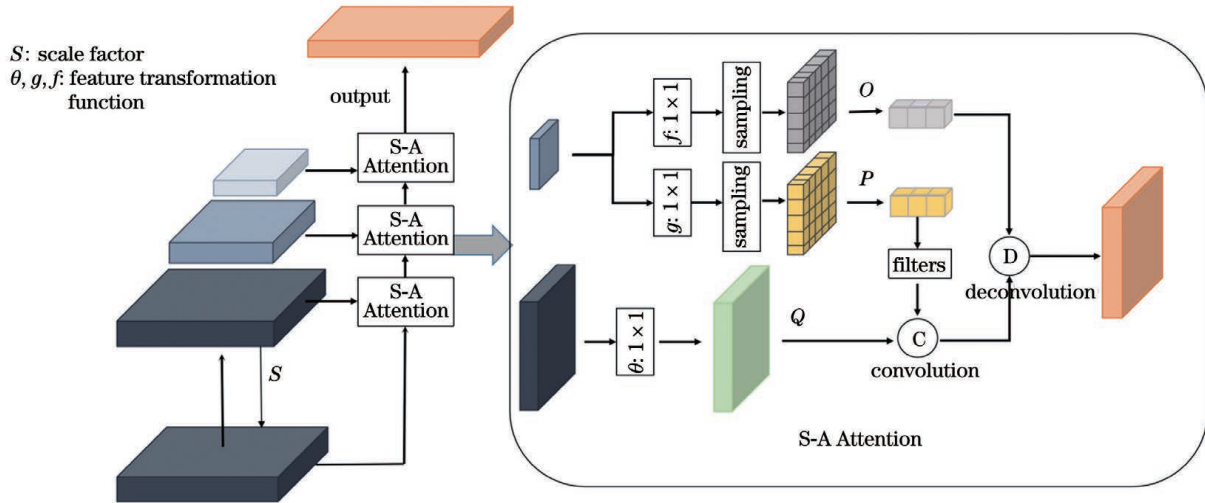


图 2 PANet 的结构  
Fig. 2 Structure of PANet

用可使本文所提网络在 OCT 图像超分辨率重建任务上取得不错的结果。

### 2.2 高效的金字塔拆分通道注意力(PECA)

本文新提出了一个 PECA 模块,它是以前 PANet 为基础进行改进的金字塔拆分通道注意力模块,如图 3 所示。本文用该模块来有效地建立多尺度通道注意力间的长期依赖关系,并将其分别加入到本网络判别器中的最后一个卷积块之前和两个密集块之后,处理多尺度的输入特征图的空间信息。与其他注意力机制相比,PECA 模块可以有效地学习更丰富的图像特征表示。同时,它是一种轻量且高效的注意力模块,能够自适应地对多维度的通道注意力权重进行特征重标定。PECA 模块主要通过 4 个步骤实现:

1)通过分割和融合模块(SPC 模块)对通道进行拆分,然后对每个通道特征图上的空间信息进行多尺度特征提取;2)用 ECA 模块替换原始 PSANet 模块中的 SENet,提取不同尺度特征图的通道注意力,这样可以在不降低通道维度的条件下学习有效的通道注意力,同时还能以极其轻量化的方式获取跨通道的交互信息<sup>[18]</sup>;3)利用 Softmax 对多尺度通道注意力向量的特征重新进行标定,得到新的多尺度通道交互之后的注意力权重;4)对重新校准的权重和相应的特征图按元素进行点乘操作,输出得到一个多尺度特征信息注意力加权之后的特征图(该特征图的信息更加丰富,有助于提高判别器模型判别图像真伪的能力)。

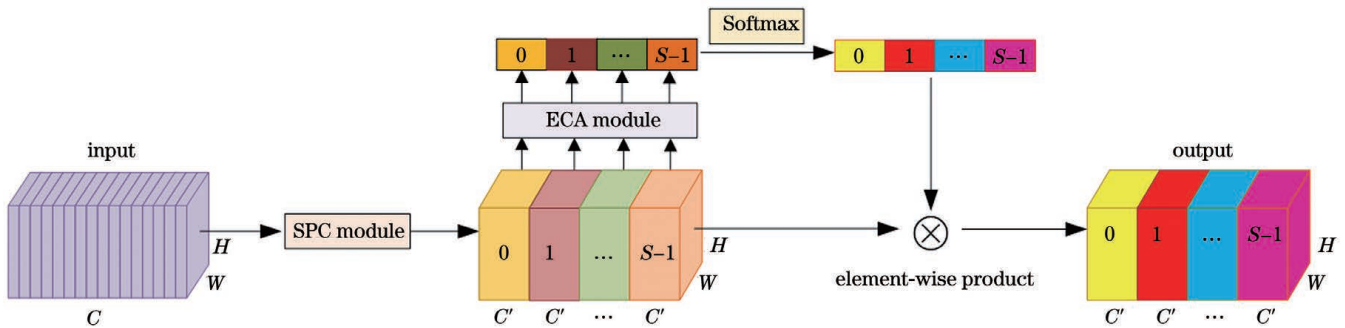


图 3 PECA 模块  
Fig. 3 PECA module

#### 2.2.1 分割和融合模块

在 PECA 模块中,实现判别器多尺度特征提取的重要模块是 SPC 模块<sup>[19]</sup>,如图 4 所示。SPC 模块可以获得不同尺度特征图的注意力权重,融合不同尺度的上下文信息,因此能产生更好的像素级注意力。SPC 模块的具体实现方式如下:

输入的特征图  $X$ ,它的宽和高分别为  $W$  和  $H$ 。首先将特征图  $X$  切分为  $S$  个部分,用  $[X_0, X_1, \dots, X_{S-1}]$  来表示,每个分割部分的通道数为

$C', C' = C/S$ 。分割之后的每个特征图都为  $X_i \in \mathbf{R}^{C' \times H \times W}, i = 0, 1, \dots, S-1$ 。然后用多尺度卷积的方式提取不同尺度特征图的空间信息。分组卷积可以减少参数量,这里主要用多尺度卷积核分组卷积,同时根据卷积核的大小来自适应选择组别的大小,组  $G'$  与内核大小  $K$  的关系式为  $G' = 2^{\frac{K-1}{2}}$ 。因此,多尺度特征提取过程的计算公式为

$$F_i = \text{Conv}(K_i \times K_i)(X_i), i = 0, 1, \dots, S-1, \quad (1)$$

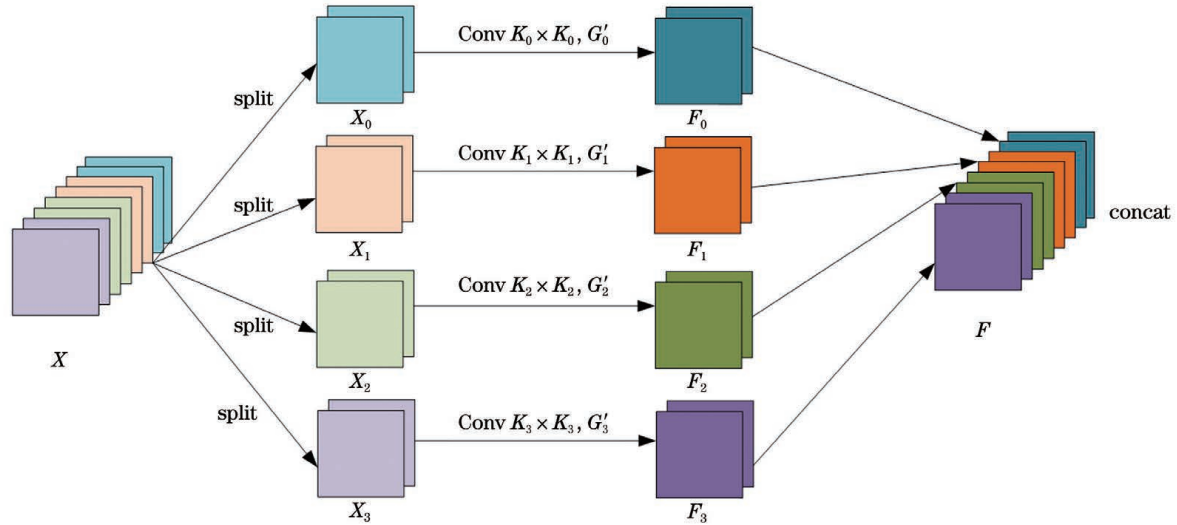


图 4 SPC 模块

Fig. 4 SPC module

式中:  $K_i = 2 \times (i + 1) + 1$ ,  $G'_i = 2^{\frac{K_i - 1}{2}}$ ,  $F_i \in \mathbf{R}^{C' \times H \times W}$ 。最后得到多尺度融合之后的特征图为

$$F_i = \text{Cat}([F_0, F_1, \dots, F_{S-1}]), F \in \mathbf{R}^{C \times H \times W}。 (2)$$

### 2.2.2 超强通道注意力(ECA)

本文使用的 PECA 模块通过使用不同大小的核卷积将输入特征图拆分成多尺度特征图,并且在不同尺度上分别使用超强通道注意力模块(ECA 模块)<sup>[20]</sup>,从而避免了维度缩减并有效捕获了跨通道交互特征。ECA 模块结构如图 5 所示。 $H$ 、 $W$ 、 $C$  分别

表示特征图的高、宽及通道数。该模块在进行逐通道全局平均池化之后,通过大小为  $k$  的快速一维卷积有效实现局部跨通道交互(内核大小  $k$  代表局部跨通道交互的覆盖范围)。该模块可以自适应确定  $k$  近邻通道,并完成通道注意力的计算,其中  $k$  与  $C$  成比例。ECA 模块与 SE 模块<sup>[21]</sup>的区别是 ECA 模块中的全连接层(FC 层)直接在全局平均池化(GAP)之后的特征上通过权重共享的一维卷积进行学习。ECA 模块是超轻量级的通道注意力模块,可以显著降低模型的复杂度并保持模型的性能。

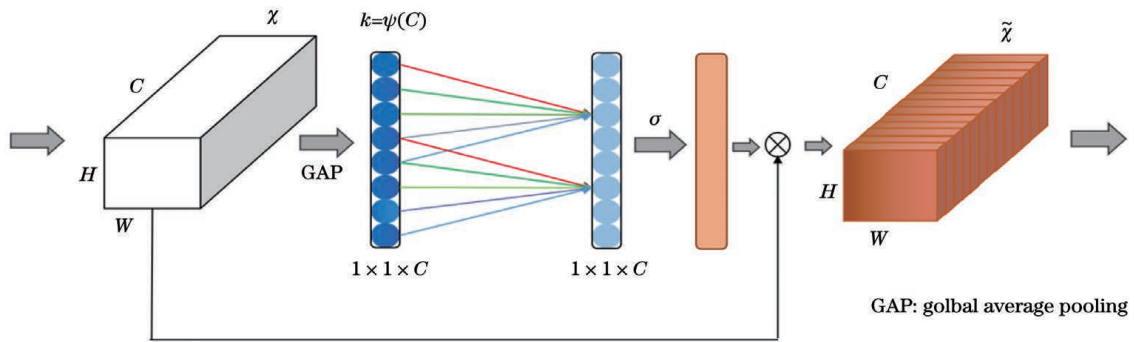


图 5 ECA 模块

Fig. 5 ECA module

## 2.3 损失函数

对于生成器网络来说,感知损失函数<sup>[22]</sup>的定义至关重要。传统的损失函数只是进行像素级别的比较,无法考虑结构上的特点。本文的图像超分辨率重建算法在感知特征上采用的损失函数包括基于 VGG 网络模型的内容损失和基于判别器的对抗损失。经过实验验证,损失函数的对抗损失权重取  $10^{-3}$  时,本文模型性能最佳,因此损失函数的计算公式为

$$L^{\text{SR}} = L^{\text{SR}}_{\text{VGG}/i,j} + 10^{-3} L^{\text{SR}}_{\text{G}}, (3)$$

式中: $L^{\text{SR}}$  表示本文算法的损失函数; $L^{\text{SR}}_{\text{VGG}/i,j}$  是内容损失函数,其中  $i$  表示 VGG 网络中第  $i$  个最大池化层, $j$  表示第  $i$  个最大池化层前的第  $j$  个卷积层; $L^{\text{SR}}_{\text{G}}$

是對抗损失函数。

### 2.3.1 内容损失函数

本文模型采用的内容损失函数如式(4)所示,就是将生成器生成的假高分辨率图像和真实的高分辨率图像送入 VGG19<sup>[23]</sup>网络中进行特征提取,然后计算提取的特征图中中像素值的均方根误差。这样做能够使生成的图像具有较好的结构相似性。内容损失函数的计算公式为

$$L^{\text{SR}}_{\text{VGG}/i,j} = \frac{1}{W_{i,j} H_{i,j}} \times$$

$$\sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} [\phi_{i,j}(I^{\text{HR}})_{x,y} - \phi_{i,j}(G_{\theta_G} I^{\text{LR}})_{x,y}]^2, (4)$$

式中:  $I^{HR}$  表示高分辨率图像;  $I^{LR}$  表示低分辨率图像;  $\theta$  为生成器参数;  $G$  表示生成器;  $W$ 、 $H$  分别为图像的宽和高;  $\phi_{i,j}$  是通过 VGG 网络中的池化层  $i$  和卷积层  $j$  得到的特征图;  $W_{i,j}$ 、 $H_{i,j}$  是特征图的维度。

### 2.3.2 对抗损失函数

为了提高生成器的性能,使生成的图像分辨率更高,本文模型引入了对抗损失函数。对抗损失函数使用负对数求和,其目的是使判别器误以为生成器生成的图像是真实的高分辨率图像<sup>[24]</sup>。其公式为

$$L_G^{SR} = \sum_{n=1}^N -\lg D_{\theta_D}(G_{\theta_G} I^{LR}), \quad (5)$$

式中:  $D$  表示判别器;  $N$  为重建图像数量。

由内容损失函数和对抗损失函数构成的感知损失函数与传统的均方误差 (MSE) 损失函数相比,能够保持更稳定的像素空间变化,提升重建图像的真实感。

## 3 实验结果与分析

### 3.1 数据集和实验设置

采用本文方法对眼科临床数据进行图像超分辨率重建处理,这里用到了两个数据集:数据集 1 (MICCAI RETOUCH, <https://retouch.grand-challenge.org/>)<sup>[25]</sup> 是第四届 MICCAI 眼科医学图像分析研讨会的 OCT 视网膜图像,该数据集中的图像是使用不同制造商的 OCT 成像设备获取的,其中 Cirrus (Zeiss Meditec) 进行了 128 次 B 扫描, T-1000 和 T-2000 (Topcon) 进行了 128 次 B 扫描, Spectralis (Heidelberg Engineering) 进行了 49 次 B 扫描;数据集 2 是由温州医科大学利用医用 OCT 设备采集的眼科视网膜图像的临床数据。这两个数据集均包含 1000 张图像的训练集和 50 张图像的测试集。

实验基于 Nvidia Tesla P100 显卡、内存为 16 GB 的公共 PyTorch 平台进行。在训练过程中,选取自适应矩估计 (Adam) 作为优化器,每批次训练数量大小为 4,生成器和判别器的初始学习率分别设置为 0.0002 和 0.00002,训练的最大周期为 100。训练集和测试集的图片尺寸均为 455 pixel × 455 pixel,训练得到生成器模型后,用图像质量指标评估该生成器对 OCT 图像超分辨率重建的效果。

### 3.2 比较方法与评价指标

将本文所提方法与一些最常见的超分辨率算法 (SRResNet 以及基于生成对抗网络的 SRGAN 和 ESRGAN<sup>[26]</sup>) 进行了比较。这些算法分别在两个数据集上进行训练和测试。为了评估网络的超分辨率重建性能,采用结构相似性 (SSIM) 和峰值信噪比 (PSNR) 来评估网络重建 OCT 图像的性能。

SSIM 是一种基于亮度、对比度和结构的主观度量,用于衡量生成的图像与原始图像的结构相似度。该值越大,代表重建图像与原始图像的误差越小。亮度  $l(x,y)$ 、对比度  $c(x,y)$  和结构  $s(x,y)$  的定义分别为

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}, \quad (6)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \quad (7)$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}, \quad (8)$$

式中:  $\mu_x$  为  $x$  的均值;  $\sigma_x$  为  $x$  的方差;  $\mu_y$  为  $y$  的均值;  $\sigma_y$  为  $y$  的方差;  $\sigma_x\sigma_y$  为  $x$  和  $y$  的协方差;  $c_1$  和  $c_2$  是为防止分母为 0 而设置的常数,  $c_1 = (k_1L)^2$ ,  $c_2 = (k_2L)^2$ ,  $L$  为像素值范围。SSIM 的计算公式为

$$S(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \quad (9)$$

PSNR 是用于衡量生成图像质量的指标,其计算公式为

$$P = 10 \cdot \lg_{10} \left( \frac{i_{MAX}^2}{f_{MSE}} \right), \quad (10)$$

式中:  $i_{MAX}$  是图像像素值的最大值,本实验中取 1;

$f_{MSE}$  是均方误差,其计算公式为  $f_{MSE} = \frac{1}{mn} \times$

$$\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2, I \text{ 和 } K \text{ 为 } m \times n \text{ 的单个图像。}$$

### 3.3 数据结果

采用本文提出的 PPECA-SRGAN 网络模型在两个 OCT 数据集上进行图像的超分辨率重建,在相同的训练集和测试集下,执行了三次训练和测试,将得到的图像的客观质量指标 SSIM 和 PSNR 的平均值作为最终的结果。将本文所提网络模型与 SRResNet 网络模型以及基于生成对抗网络的 SRGAN 和 ESRGAN 网络模型进行对比,以验证所提网络模型对 OCT 图像的处理效果。

表 1 列出了上述 4 种网络模型在 OCT 数据集 1 和数据集 2 上 4 倍重建后图像的 SSIM 指标。这些模型都使用相同的数据集进行训练,所得结果可以客观地反映本文模型与其他 3 种超分辨率重建模型的性能差异。从 SSIM 结果来看,基于生成对抗网络的 SRGAN 和 ESRGAN 的重建效果要好于传统的 SRResNet,且 ESRGAN 的最佳 SSIM 性能指标 (0.827) 仅次于 PPECA-SRGAN 的最佳 SSIM 指标

表 1 不同模型 4 倍重建后图像的 SSIM 指标对比

Table 1 Comparison of SSIM indexes of images reconstructed by 4 times of different models

Model	SSIM	
	Data set 1	Data set 2
SRResNet	0.764	0.733
SRGAN	0.813	0.809
ESRGAN	0.827	0.818
PPECA-SRGAN	0.842	0.835

(0.842)。总的来说,本文提出的 PPECA-SRGAN 网络模型获得了最好的 SSIM 指标。

表 2 列出了各网络模型在两个数据集上的 PSNR 指标。比较表中的数据可以发现,PSNR 指标平均值最高的是 PPECA-SRGAN 模型,达到了 32.59 dB,较

表 2 不同模型 4 倍重建后图像的 PSNR 指标对比

Table 2 Comparison of PSNR indexes of images reconstructed by 4 times of different models

Model	PSNR /dB	
	Data set 1	Data set 2
SRResNet	30.53	30.36
SRGAN	31.94	31.82
ESRGAN	32.18	32.07
PPECA-SRGAN	32.69	32.48

ESRGAN 网络模型提高了约 0.47 dB,比 SRGAN 和 SRResNet 模型分别提高了约 0.71 dB 和 2.15 dB。

使用不同的网络模型对测试集中 1 张正常图像和 1 张病变图像进行重建,结果如图 6 和图 7 所示,将图像中的局部区域(重建难度较大的细节部分)进行放大并放在图像的右下角。结合表 1、表 2 和图 6、图 7 可以发现 4 种网络模型在 OCT 图像数据集上的重建差异,并可以发现本文所提模型重建出的图像更接近原始 HR 图像,视觉效果最优,ESRGAN 和 SRGAN 网络模型的重建效果都优于 SRResNet 网络模型。

实验数据可以说明本文所提 PPECA-SRGAN 网络对于 OCT 图像超分辨率重建具有较好的效能,能解决临床诊断中图像模糊的问题,而且还可以有效地保留 OCT 图像中的重要信息,使图像的细节得以恢

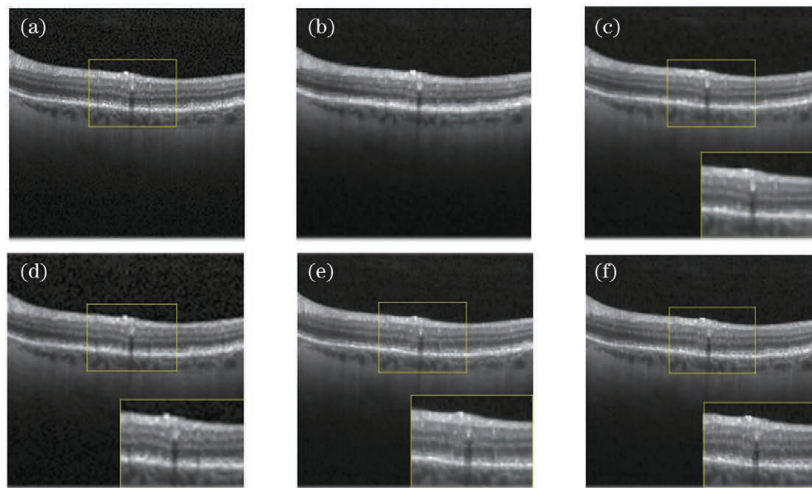


图 6 1 张正常图像的不同网络模型的超分辨率重建结果。(a)HR 图像;(b)LR 图像;(c)SRResNet 重建结果;(d)SRGAN 重建结果;(e)ESRGAN 重建结果;(f)PPECA-SRGAN 重建结果

Fig. 6 Super-resolution reconstruction of normal image by different models. (a) High resolution image; (b) low resolution image; (c) reconstructed image by SRResNet; (d) reconstructed image by SRGAN; (e) reconstructed image by ESRGAN; (f) reconstructed image by PPECA-SRGAN

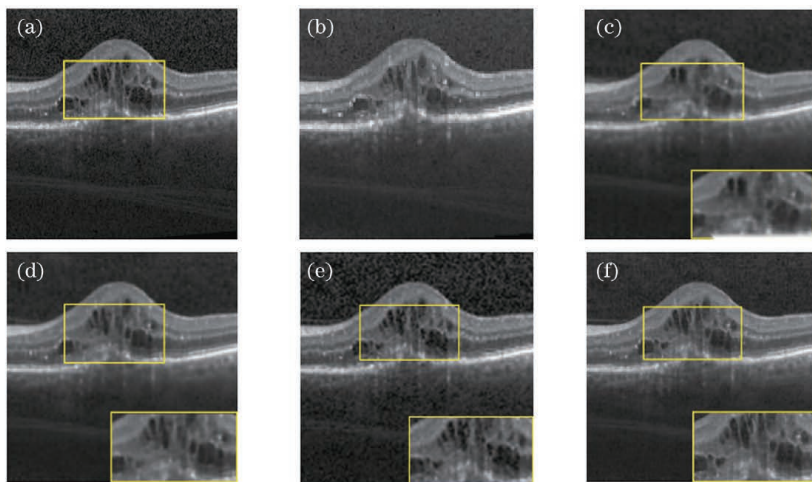


图 7 1 张病变图像的不同网络模型的超分辨率重建结果。(a)HR 图像;(b)LR 图像;(c)SRResNet 重建结果;(d)SRGAN 重建结果;(e)ESRGAN 重建结果;(f)PPECA-SRGAN 重建结果

Fig. 7 Super-resolution reconstruction of pathological image by different models. (a) High resolution image; (b) low resolution image; (c) reconstructed image by SRResNet; (d) reconstructed image by SRGAN; (e) reconstructed image by ESRGAN; (f) reconstructed image by PPECA-SRGAN

复,尤其是对病变区域的重建效果较为清晰。此外,本文所提网络模型的实验周期较短,性能稳定,能在比较短的训练时间内达到预期的效果。

### 3.4 分析与讨论

通过实验将 PPECA-SRGAN 与 SRResNet、SRGAN、ESRGAN 的超分辨率重建结果进行了对比。总的来看,生成对抗网络模型(SRGAN 和 ESRGAN)的重建效果要比 SRResNet 好,但对于图像细节的恢复,本文所提 PPECA-SRGAN 网络模型重建的图像质量具有更高的视觉质量,其 SSIM 指标平均值比 SRResNet、SRGAN、ESRGAN 分别高约 0.090、0.028、0.016,PSNR 指标平均值分别提高了 2.15 dB、0.71 dB、0.47 dB。PPECA-SRGAN 良好的重建效果得益于添加了 PANet 以及新提出的 PECANet 注意力机制,这两个模块增强了模型对 OCT 视网膜图像特征的捕获能力以及对细节信息的重构能力。PECA 模块包含金字塔拆分提取特征且通过 ECANet 注意力机制融合多尺度信息,而 PANet 可以有效减少压缩伪影等图像噪声,从而使得所提网络模型的重建效果优于 SRGAN 网络和其他的传统网络。

## 4 结 论

本文提出的 PPECA-SRGAN 是一个改进的 SRGAN 网络模型,用它对 OCT 视网膜图像进行超分辨率重建,可以解决光学相干断层扫描成像设备采集的图像清晰度不高、细节较少的问题。由于医学图像与自然图像在重建上有所区别,因此本文模型利用改进的生成对抗网络<sup>[27-28]</sup>生成高分辨率 OCT 图像,并在 MICCAI RETOUCH 数据集和温州医科大学采集的数据上进行了训练和验证。所提模型首先在 SRGAN 网络生成器的残差块之间添加了 PANet 模块,利用金字塔结构提取多尺度特征关系,同时抑制不必要的运动伪影;然后将 PECA 模块插入判别器中,通过有效地结合空间和通道注意力来学习更多的图像细节,使判别器能够得到更丰富的图像特征信息。实验证明,本文所提模型可以有效地提高 OCT 视网膜图像的重建分辨率,与 SRResNet、SRGAN、ESRGAN 重建图像的 PSNR 和 SSIM 指标的平均值相比分别提高了约 3.5% 和 5.6%。在临床诊断中使用所提网络模型可以克服低分辨率成像系统固有的局限性,重建图像获取过程中丢失的各种细节,且易于集成和实现。未来若可以获得更高质量的数据集和更轻量化的算法,本文所提网络模型有可能进一步提高医学图像的分辨率,使超分辨率重建技术更好地应用于临床实践中。

### 参 考 文 献

- [1] Khan A, Sohail A, Zahoora U, et al. A survey of the recent architectures of deep convolutional neural networks [J]. *Artificial Intelligence Review*, 2020, 53(8): 5455-5516.
- [2] 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.
- [3] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [4] Ma Y H, Chen X J, Zhu W F, et al. Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN [J]. *Biomedical Optics Express*, 2018, 9(11): 5129-5146.
- [5] Qiu B, Huang Z Y, Liu X, et al. Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function [J]. *Biomedical Optics Express*, 2020, 11(2): 817-830.
- [6] Huang Y Q, Lu Z X, Shao Z M, et al. Simultaneous denoising and super-resolution of optical coherence tomography images based on generative adversarial network [J]. *Optics Express*, 2019, 27(9): 12289-12307.
- [7] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [8] Bing X Y, Zhang W W, Zheng L Y, et al. Medical image super-resolution using improved generative adversarial networks [J]. *IEEE Access*, 2019, 7: 145030-145038.
- [9] Das V, Dandapat S, Bora P K. A diagnostic information based framework for super-resolution and quality assessment of retinal OCT images [J]. *Computerized Medical Imaging and Graphics*, 2021, 94: 101997.
- [10] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C] // 2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.
- [11] 尹海涛, 邓皓. 基于混合注意力的对偶残差去噪网络 [J]. *激光与光电子学进展*, 2021, 58(14): 1410002.
- [12] Yin H T, Deng H. Dual residual denoising network based on hybrid attention [J]. *Laser & Optoelectronics Progress*, 2021, 58(14): 1410002.
- [13] Fu J, Li W S, Du J, et al. A multiscale residual pyramid attention network for medical image fusion [J]. *Biomedical Signal Processing and Control*, 2021, 66: 102488.
- [14] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [15] Cheng X, Li X, Yang J, et al. SESR: single image super-resolution with recursive squeeze and excitation networks [C] // 2018 24th International Conference on Pattern Recognition (ICPR), August 20-24, 2018, Beijing, China. New York: IEEE Press, 2018: 147-152.
- [16] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. (2015-11-19) [2021-05-04]. <https://arxiv.org/abs/1511.06434>.
- [17] Wang H, Wu C D, Chi J N, et al. Image super-resolution using multi-granularity perception and pyramid attention networks [J]. *Neurocomputing*, 2021, 443: 247-261.
- [18] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 8759-8768.
- [19] Qing Y H, Liu W Y. Hyperspectral image classification based on multi-scale residual network with attention mechanism [J].

- Remote Sensing, 2021, 13(3): 335.
- [19] Zhang H, Zu K K, Lu J, et al. EPSANet: an efficient pyramid split attention block on convolutional neural network[EB/OL]. (2021-05-30)[2021-05-07]. <https://arxiv.org/abs/2105.14447>.
- [20] Lin X, Huang Q, Huang W, et al. Single image deraining via detail-guided efficient channel attention network [J]. Computers & Graphics, 2021, 97: 117-125.
- [21] Cao Y, Xu J R, Lin S, et al. GCNet: non-local networks meet squeeze-excitation networks and beyond[C] // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), October 27-28, 2019, Seoul, Korea. New York: IEEE Press, 2019: 1971-1980.
- [22] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[M] // Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016: 694-711.
- [23] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2021-05-07]. <https://arxiv.org/abs/1409.1556>.
- [24] Li J C, Wu L M, Wang S M, et al. Super-resolution image reconstruction of textile based on SRGAN [C] // 2019 IEEE International Conference on Smart Internet of Things (SmartIoT), August 9-11, 2019, Tianjin, China. New York: IEEE Press, 2019: 436-439.
- [25] Bogunović H, Venhuizen F, Klimescha S, et al. RETOUCH: the retinal OCT fluid detection and segmentation benchmark and challenge[J]. IEEE Transactions on Medical Imaging, 2019, 38(8): 1858-1874.
- [26] Wang X T, Yu K, Wu S X, et al. ESRGAN: enhanced super-resolution generative adversarial networks [M] // Leal-Taixé L, Roth S. Computer vision-ECCV 2018 workshops. Lecture notes in computer science. Cham: Springer, 2019, 11133: 63-79.
- [27] Zhang W L, Liu Y H, Dong C, et al. RankSRGAN: generative adversarial networks with ranker for image super-resolution [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea. New York: IEEE Press, 2019: 3096-3105.
- [28] 朱斯琪, 王珏, 蔡玉芳. 基于改进型循环一致性生成对抗网络的低剂量CT去噪算法[J]. 光学学报, 2020, 40(22): 2210002. Zhu S Q, Wang J, Cai Y F. Low-dose CT denoising algorithm based on improved cycle GAN[J]. Acta Optica Sinica, 2020, 40(22): 2210002.

## Super-Resolution Reconstruction of Optical Coherence Tomography Retinal Images by Generating Adversarial Network

Ke Shuting<sup>1</sup>, Chen Minghui<sup>1\*</sup>, Zheng Zexi<sup>2</sup>, Yuan Yuan<sup>1</sup>, Wang Teng<sup>1</sup>, He Longxi<sup>1</sup>, Lü Linjie<sup>1</sup>, Sun Hao<sup>1</sup>

<sup>1</sup> Shanghai Engineering Research Center of Interventional Medical Device, the Ministry of Education of Medical Optical Engineering Center, School of Health Sciences and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;

<sup>2</sup> School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

### Abstract

**Objective** Optical coherence tomography (OCT) imaging shows great potential in clinical practice because of its noninvasive nature. However, two critical issues affect the diagnostic capability of OCT imaging. The first problem is that the interferential nature of OCT imaging produces interference noise, which reduces contrast and obfuscates fine structural features. The second problem is caused by the low spatial sampling rate of OCT. In fact, in clinical diagnosis, the use of a lower spatial sampling rate is a method to achieve a wide field of vision and reduce the impact of unconscious movement. Therefore, most OCT images obtained in reality are not optimal in terms of signal-to-noise ratio and spatial sampling rate. There are significant differences in the texture and brightness of the retinal layer in patients, as well as in the shape and size of the lesion area, so traditional models may not be able to reliably reconstruct the pathological structure. To obtain high peak signal-to-noise ratio (PSNR) and high-resolution B-scan OCT images, it is necessary to develop sufficient methods for super-resolution reconstruction of OCT images. In this paper, an improved OCT super-resolution image reconstruction network structure (PPECA-SRGAN) was proposed.

**Methods** In this paper, a PPECA-SRGAN network based on generative adversarial network (GAN) was proposed. The network model includes a generator and a discriminator. A PA module was added between the residual blocks of the generator to increase the feature extraction capability of OCT retinal image reconstruction. In addition, a PECA module was added to the discriminator, which is an improvement of the pyramid split attention network (PSANet) and can fully capture the spatial information of multi-scale feature maps. First, we used two data sets to test a training set of 1000 images and a test set of 50 images, respectively. The data set was imported into the preprocessing module, and the low-resolution image was obtained through four down-sampling processes. Then, the generator was used to train the model to generate high-resolution images from low-resolution images. When the discriminator could not distinguish the authenticity of the images, it indicated that the generation network generated high-resolution images. Finally, the image quality was evaluated using the structural similarity index measure (SSIM) and PSNR.



**Results and Discussions** The super-resolution index evaluation results of PPECA-SRGAN and the other three models were compared, as well as the final reconstruction effect images. In general, PPECA-SRGAN's reconstruction effect was better than SRResNet; however, for the restoration of the image details, the image quality of the PPECA-SRGAN network reconstruction was more in line with the satisfaction degree of human vision. Compared with SRResNet, SRGAN, and ESRGAN, the SSIM indexes of PPECA-SRGAN were 0.090, 0.028, and 0.016 higher and the PSNR indexes were 2.15 dB, 0.71 dB, and 0.47 dB higher, respectively. The good reconstruction effect of PPECA-SRGAN was due to the addition of the attention mechanism called path aggregation network (PANet) and the proposed attention mechanism named PECA, both enhancing the capture of OCT retinal image features and the reconstruction of details. The PECA module was composed of pyramid splitting and extracting features, with the use of ECANet to fuse multi-scale information. PANet can effectively reduce image noise, such as compression artifacts. This makes our model better than the SRGAN network and other traditional networks. Therefore, the application of the proposed model in OCT image super-resolution reconstruction has been verified, and its performance has been improved compared with other reconstruction algorithms.

**Conclusions** The PPECA-SRGAN network structure proposed in this paper is an improved model of the SRGAN network for super-resolution reconstruction of retinal OCT B-scan images. We conducted training and verification on MICCAI RETOUCH data set and data collected by Wenzhou Medical University to solve the problems of low-resolution and few details of images collected by OCT. We used advanced GAN to improve the super-resolution reconstruction of OCT images, and the SRGAN network was improved due to the difference in reconstruction between medical images and natural images. Firstly, a PANet module was added between the residual blocks of the generator to extract multi-scale feature relations by pyramid structure and suppress unnecessary artifacts. Then, the PECA module was inserted into the discriminator to effectively combine spatial and channel attentions to learn more image details for the discriminator and obtain richer image pair feature information. The experimental results show that this model is effective and stable in improving the resolution of medical images. Compared with SRResNet, SRGAN, and ESRGAN, the PSNR and SSIM indexes of the reconstructed images were improved by about 3.5% and 5.6%, respectively. In clinical diagnosis, the proposed algorithm can overcome the inherent limitations of low-resolution imaging systems and reconstruct various details lost in the process of image acquisition; the algorithm is easy to integrate and implement. In the future, if higher-quality data sets and lighter algorithms can be obtained, it is possible to further improve the quality of super-resolution reconstruction medical images and make them more applicable in clinical practice.

**Key words** biotechnology; optical coherence tomography; super-resolution; generative adversarial network; unpaired image