

文章编号: 2095-2163(2023)05-0082-09

中图分类号: TP391

文献标志码: A

基于改进 pix2pixHD 模型的 COVID-19 CT 图像生成方法研究

高志军, 冀远明, 史二美

(黑龙江科技大学 计算机与信息工程学院, 哈尔滨 150022)

摘要: 当前 COVID-19 病毒流行, 肺部 CT 图像已成为医生对 COVID-19 患者进行准确诊断和跟踪治疗的重要辅助工具之一, 但公开的患者数据集较少及数据获取较为困难的问题, 导致不易获得好的深度学习模型对肺部 CT 图像精准筛查和诊断。针对于此, 本文提出了一种基于 pix2pixHD 的深度网络图像生成模型, 用以提高 COVID-19 CT 图像的生成质量。采用 pix2pixHD 框架, 通过对模型上采样的过程中加入 SPADE 残差块, 并对模型的多尺度鉴别器增加 1/8 尺度鉴别器, 使模型可以生成病灶边界更为清晰的 COVID-19 患者肺部 CT 图像。在公开数据集上与现有模型进行的实验比较结果表明, 本文提出的模型在 COVID-19 CT 图像上峰值信噪比与结构相似性分别为 12.46 和 0.327, Fr'echet Inception 距离 (*FID*) 为 57.4, 较现有模型具有生成质量更高、生成细节更好、收敛速度更快的优势。

关键词: 新型冠状病毒; 电子计算机断层扫描; 改进 pix2pixHD 模型

Generation method of the COVID-19 CT image based on improved pix2pixHD model

GAO Zhijun, JI Yuanming, SHI Ermei

(School of Computer and Information Engineering, Heilongjiang University of Science and Technology, Harbin 150022, China)

[Abstract] At present, COVID-19 virus is pandemic all over the world. Lung CT images have become one of the important auxiliary tools for doctors to accurately diagnose and follow up the treatment of COVID-19 patients. However, it is difficult to obtain a good deep learning model for accurate screening and diagnosis of lung CT images due to the small number of published patient data sets and difficult data acquisition. We propose a depth network image generation model based on pix2pixHD to improve the generation quality of COVID-19 CT images. By adding spade residual block in the process of upsampling the model and adding 1/8 scale discriminator to the multi-scale discriminator of the model, the improved pix2pixhd model can generate CT images of COVID-19 patients with clearer lesion boundaries. We performed experimental comparisons with existing models on open datasets. The peak signal-to-noise ratio was 12.46. The structural similarity was 0.327. The Fr'echet Inception Distance (*FID*) is 57.4. The experimental results show that this model has the advantages of higher generation quality, better generation details and faster convergence than the existing models.

[Key words] COVID-19; electronic computed tomography; improved pix2pixHD.

0 引言

2019 新型冠状病毒 (Corona Virus Disease 2019, COVID-19) 是一个大型病毒家族, 具有高传染性。人感染了冠状病毒后常见体征有呼吸道症状、发热、咳嗽、气促和呼吸困难等。在较严重病例中, 感染可导致肺炎、严重急性呼吸综合征、肾衰竭, 甚至死亡。

电子计算机断层扫描 (Computed Tomography, CT) 是利用 X 线束对人体的某一部位进行连续的断

面扫描, CT 检查与以往的 X 线相比, 具有分辨率高、检查速度快、检查安全等特点, 已成为医生对 COVID-19 患者进行准确诊断和跟踪治疗的重要辅助工具之一。为了准确和快速地实现对 COVID-19 患者的智能检测, 基于深度学习和计算机视觉的 COVID-19 患者 CT 图像病变检测已成为相关学者的研究热点。

对于深度学习网络模型来说, 数据集的规模和数目将很大程度影响最终训练效果的好坏^[1]。但

基金项目: 黑龙江省省属高等学校基本科研业务费科研项目 (Hkdqg201911)。

作者简介: 高志军 (1980-), 男, 博士, 讲师, 硕士生导师, 主要研究方向: 医学图像处理和机器学习; 冀远明 (1993-), 男, 硕士研究生, 主要研究方向: 医学图像处理; 史二美 (1997-), 女, 硕士研究生, 主要研究方向: 医学图像处理。

通讯作者: 高志军 Email: zhagao@usth.edu.cn

收稿日期: 2022-08-10

是,在计算机辅助诊断领域中,由于涉及到患者隐私,研究者很难获取到大量的医学图像数据集,而且由于疾病的高传染性,医务人员在收集 COVID-19 CT 数据时面临着高风险^[2]。其次,因缺乏可用于数据标记的专家,对采用监督训练方法提出了另一个挑战。增加训练样本的传统方法包括缩放、旋转、翻转、平移和弹性变形,然而这些转换并未考虑由不同成像协议或序列引起的变化,更不用说尺寸、形状、位置的变化和特定病理的出现。

近年来,基于深度学习的计算机视觉方法在医学图像生成中显示出巨大的应用前景,许多学者提出了具有潜力的新模型。如:Jiang Y 等^[3]提出的具有条件生成对抗网络的 COVID-19 CT 图像生成;Zhou L 等^[4]提出的肺部自动交换编码器,对 COVID-19 CT 图像进行生成等。

但是,目前提出的 COVID-19 CT 图像生成模型均存在生成质量不高、病变区域不明显、模型泛化能力不强等问题。为此,本文提出了一种基于 pix2pixHD^[5] 的 CT 图像生成方法,该方法可以生成高质量的 COVID-19 CT 图像,在性能上优于现有的 COVID-19 CT 图像生成方法。可以达到扩充数据集,提升病变识别模型的准确性和泛化能力的作用。

1 方法

1.1 pix2pixHD 模型

pix2pixHD 是一个 CGAN^[6] 框架,由生成器(G)和鉴别器(D)组成,用于图像到图像的翻译。生成器的目标是将标签映射到真实的图像上,而鉴别器的目标是将真实图像和翻译后的图像区分开来。

pix2pixHD 采用了多尺度鉴别器,多尺度鉴别器由原始图像、原始图像的 1/2 下采样和原始图像的 1/4 下采样 3 个尺度组成。其多尺度的鉴别器目标函数可表示为

$$L_{\text{multi}} = \min_G \max_{D_1, D_2, D_3} \sum_{k=1,2,3} L_{\text{GAN}}(G, D_k) \quad (1)$$

具体来说, pix2pixHD 从鉴别器的多个层中提取特征,并学习从真实图像和生成图像中匹配这些中间表示。为了便于表示,本文将鉴别器 D_k 的第 i 层特征提取器表示为 $D_k^{(i)}$ (从输入到 D_k 的第 i 层)。pix2pixHD 通过极大极小博弈,对给定标签映射的真实图像的条件分布进行建模,其特征匹配损失 $L_{\text{FM}}(G, D_k)$ 计算为

$$L_{\text{FM}}(G, D_k) = E_{(s,x)} \sum_{i=1}^T \frac{1}{N_i} [\| D_k^{(i)}(s, x) - D_k^{(i)}(s, \quad$$

$$G(s) \|_1] \quad (2)$$

其中, T 是总层数; N_i 表示每层中的元素数; s 表示标签映射; x 表示真实图像。

最终目标函数将 GAN 损失和特征匹配损失结合为

$$\text{Loss} = \min_G \left(\max_{D_1, D_2, D_3} \sum_{k=1,2,3} L_{\text{GAN}}(G, D_k) + \lambda \sum_{k=1,2,3} L_{\text{FM}}(G, D_k) \right) \quad (3)$$

1.2 空间自适应归一化模型

空间自适应归一化 (Spatially - Adaptive Normalization, SPADE)^[7] 模型是一个条件归一化层,其通过空间自适应及学习的变换,使用输入语义布局来调制激活,并可以在整个网络中有效地传播语义信息。此方法中首先生成一系列学习好的数据分布,然后通过一层一层的 SPADE ResBlk 堆叠而成, feature map 尺寸由小到大,通道数由大到小来生成最终的真实图像。而在每一层 SPADE ResBlk 中,不断地加入语义分割图片来进行干预,使网络在每一层都能学习到多尺度的语义信息。

SPADE 残差块是一个类似于 ResNet^[8] 残差块的结构,除了具有 ReLU 激活函数和 3×3 的卷积以外,还会用标签信息和 SPADE 方式取代一次卷积,提升了残差块的运行速度,也可以时刻保持与真实标签的距离,提升模型的准确度。

1.3 改进的 pix2pixHD 模型

COVID-19 CT 图像生成的主要目标是要获得高质量、高分辨率的 COVID-19 CT 图像。为了提升图像质量和分辨率,本文对 pix2pixHD 网络进行了两点改进。首先对 pix2pixHD 网络引入了更多尺度的判别器进行均值判别,使生成图像的边界更为清晰,之后在生成器的上采样阶段,引入 SPADE 残差块,生成器包含一系列带有上采样层的 SPADE 残差块,以获得更好的性能。改进的 pix2pixHD 模型主要流程如图 1 所示。

1.3.1 SPADE 残差块

受 SPADE 模型的启发,本文在上采样层加入 SPADE 残差块,能够更好的提升图像质量,使用 SPADE 学习所有归一化层的调制参数。由于每个残差块以不同的比例运行,因此本文对语义掩码进行下采样,以匹配空间分辨率。如图 2 所示,在上采样阶段加入 SPADE 残差块,在不改变损失函数的情况下,使图像质量有所提升。其可以更好地保留针对常见归一化层的语义信息。

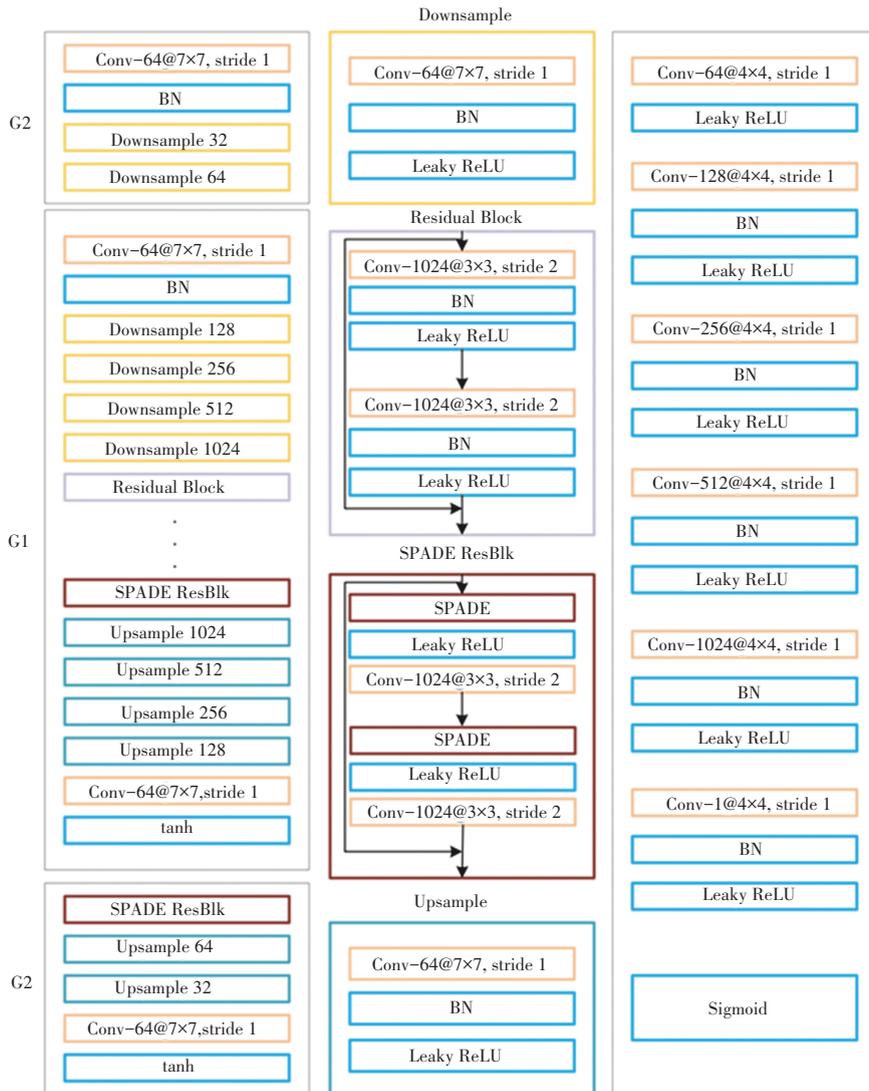


图 1 改进的 pix2pixHD 模型的主要网络框架

Fig. 1 Overall framework of the improved pix2pixHD

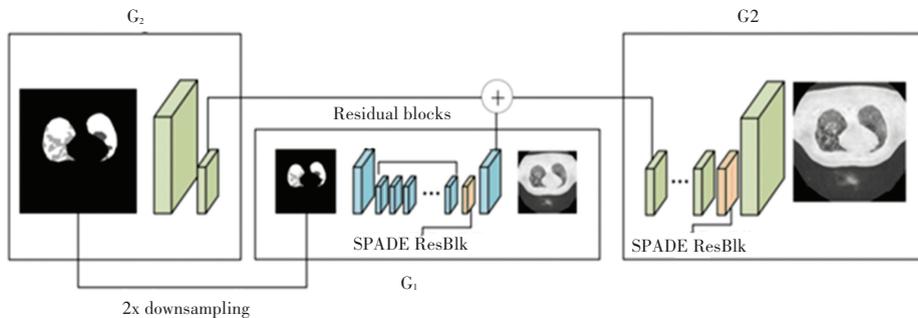


图 2 改进的 pix2pixHD 模型的生成器网络结构

Fig. 2 Generator framework of the improved pix2pixHD

同时,本文使用 LReLU 激活函数替换了原本的 ReLU 激活函数,主要是为了进一步缓解梯度消失的问题,SPAE 残差块具体流程如图 3 所示。

1.3.2 引入多尺度判别器

鉴别器结构采用了 pix2pixHD 的多尺度判别器。虽然鉴别器的结构是相同的,但尺度最大的鉴

别器接受域最大,其具有更全局的图像视图,可以指导生成器生成全局一致的图像,而尺度最小的判别器鼓励生成器生成更细的细节。本文引入 1/8 下采样鉴别器,使图像细节更为清晰,生成图像质量更好。新的多尺度的鉴别器损失函数可表示为

$$L_{mult} = \min_G \max_{D_1, D_2, D_3, D_4} \sum_{k=1,2,3,4} L_{GAN}(G, D_k) \quad (4)$$

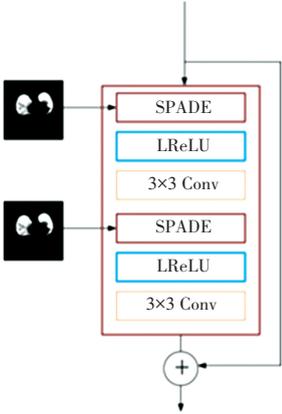


图 3 SPADE 残差块结构

Fig. 3 The framework of the SPADE residual block

本文完整目标函数将 GAN 损失和特征匹配损失结合为

$$Loss = \min_G \left(\max_{D_1, D_2, D_3, D_4} \sum_{k=1,2,3,4} L_{GAN}(G, D_k) + \lambda \sum_{k=1,2,3,4} L_{FM}(G, D_k) \right) \quad (5)$$

其中, λ 作为控制这两项的重要性的参数。对于特征匹配损失 L_{FM} , D_k 仅用作特征提取器, 不会最大化损失 L_{FM} 。

输入层为卷积核数目为 64、大小为 4×4 、步长为 2 的卷积层, 卷积操作后使用 Leaky ReLU 激活函数。输出层为卷积核数目为 1、大小为 4×4 、步长为 1 的卷积层, 卷积操作后使用 Sigmoid 激活函数。中间包括 4 个下采样操作, 每次操作之后通道数目会加倍, 下采样操作的卷积核为 4×4 , 步长为 2 的卷积层, 卷积操作后使用 BN 层加快学习速率, 激活函数采用 Leaky ReLU 函数。如图 4 所示。

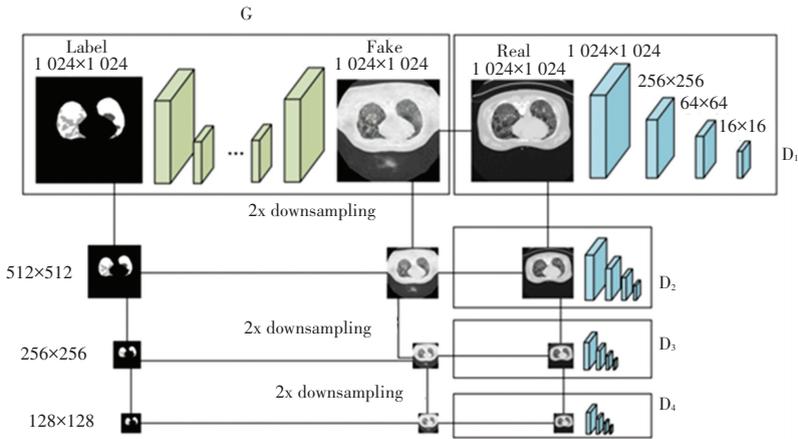


图 4 多尺度鉴别器结构

Fig. 4 Multi-scale discriminator framework of the improved pix2pixHD

综上所述, 本文所提出的算法流程如下:

算法 本文提出的算法流程

for the number of training iterations do

label $\rightarrow G_1$ downsample % 输入

G_2 downsample

G_2 upsample

G_1 upsample \rightarrow fake image % 输出

for $i = 0; i < 4; i++$ do

fake image downsample i times

real image downsample i times

Satisfied $Loss = \min_G \left(\max_{D_1, D_2, D_3, D_4} \sum_{k=1,2,3,4} L_{GAN}(G,$

$D_k) + \lambda \sum_{k=1,2,3,4} L_{FM}(G, D_k) \right)$

convolution $4 \times 4 \rightarrow$ update generator % 更新生

成器

end for

end for

2 实验及结果

2.1 数据集

本文使用的数据集是 20 例诊断为 COVID-19 患者的 CT 扫描图像, 且专家对肺部和感染区域进行了分割标注^[9]。大小为 $630 \times 630 \times 310$, 有病图占正常图比例的 52.86%。该数据集是公开数据集中少有的带有专家标注的数据集, 附有 Ma Jun、Ge Cheng、Wang Yixin、An Xingle 等专家的标注。

2.2 实验设置

本文使用了 cycleGAN^[10]、pix2pix^[11]、styleGAN^[12]

和 pix2pixHD 作为本方法的对比方法,将数据集分割为 630×630 的二维图像,去除边界没有肺部轮廓的图像后,得到 517 张图像。在训练过程中,本文将 450 张图像作为训练集,50 张作为验证集,剩余的 17 张作为测试集,并将图像分辨率扩大到 1 024×1 024,通过对测试集进行评价得出实验结论。

pix2pixHD 模型在 Pytorch 框架上使用 Adam 优化器^[13]进行训练,初始学习率为 0.000 2,前 100 次周期的学习率相同,在接下来的 100 次周期里线性下降到 0。权值初始化为均值为 0,标准差为 0.02 的高斯分布。

本文的实验平台为 NVIDIA Tesla V100 16 GB 及 32 GB 内存的服务器, Linux 系统, python 版本为 3.8, pytorch 版本为 1.11, CUDA 版本为 11.3。

2.3 评价指标

本文采用常见的评估方法,对每个模型生成的结果使用同一个分割模型 DeepLabV2^[14],并比较预测的分割区域与专家标注的匹配程度。如果输出图像是真实的,则训练的语义分割模型应该能够预测专家标注结果。实验中使用平均交并比 (*MIoU*) 和像素精度 (*Accu*) 去评价分割区域与专家标注的匹配程度。除 *MIoU* 和 *Accu* 分割性能指标外,还使用 Fréchet 初始距离^[15] (*FID*) 来测量生成结果分布与真实图像分布之间的距离,使用峰值信噪比 (*PSNR*) 和结构相似性 (*SSIM*) 来评价图像的生成质量。

2.3.1 峰值信噪比 (*PSNR*)

峰值信噪比是用来衡量两张图像差异的指标,一般通过均方误差 (*MSE*) 进行定义。若有两个 $m \times n$ 单色图像 I 和 K , 两张图像噪声近似的情况下,则两者的均方误差定义为

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (6)$$

峰值信噪比则定义为

$$PSNR = 10 \times \log_{10} \frac{\text{MAX}_I^2}{MSE} \quad (7)$$

其中, MAX_I 表示的是图像点颜色的最大数值, *PSNR* 越大则表示图像越清晰,噪声越小。

2.3.2 结构相似性 (*SSIM*)

结构相似性是一种衡量两幅图像相似度的指标,相似性越高则表明图像生成与原图越接近。给定两个图像,其结构相似性可表示为

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

其中, μ_x, μ_y 分别代表图像 x, y 的像素灰度平均值; σ_x^2, σ_y^2 分别代表图像 x, y 的像素方差; $\sigma_x\sigma_y$ 表示图像 x, y 的协方差; c_1, c_2 是用来维持稳定的常数。

2.3.3 Fréchet 初始距离 (*FID*)

Fréchet Inception 距离 (*FID*) 是评估生成图像质量的度量标准,专门用于评估生成对抗网络的性能。该分数作为对已有 Inception 分数 (*IS*) 的改进而被提出。由于 Inception 分数缺少生成图像与真实图像的比较,而研发 *FID* 分数的目的是基于一组生成图像的统计量与来自目标域的真实图像的统计量进行比较,实现对生成图像的评估。*FID* 值越低,图像质量越好;反之,得分越高,质量越差,两者关系应该是线性的。计算公式可表示为

$$FID(x,g) = \|\mu_x - \mu_g\|_2^2 + \text{tr}(\sigma_x + \sigma_g - 2\sqrt{(\sigma_x\sigma_g)}) \quad (9)$$

其中, x 表示真实图像; g 表示生成图像; tr 表示矩阵对角线上元素总和; μ 和 σ 的含义为用 Inception V3 来提取中间层的特征,然后使用一个均值为 μ , 方差为 σ 的正态分布去模拟这些特征的分布。较低的 *FID* 意味着生成样本和真实样本的相关性越高,即图像质量也会越高。

DeepLabV2^[15] 提出使用空洞卷积进行密集采样,在不增加参数量或计算量的情况下有效地扩大感受野,然后扩展空洞卷积提出了 ASPP 来捕获多个尺度的上下文,最后利用条件随机场 (*CRF*) 进行后处理以提高边缘定位精度。

DeepLabV2 是现在较为准确且流行的分割算法,大量的 GAN 模型使用其作为评价指标的一环,将生成结果分割后进行后续的平均交并比和像素精度评价具有很好的评价能力。

2.3.4 平均交并比 (*MIoU*)

平均交并比 (*MIoU*) 是衡量图像分割精度的重要指标,其计算两个集合的交集和并集之比,在语义分割的问题中,这两个集合为真实值和预测值。计算公式可表示为

$$MIoU = \frac{1}{k} \sum_{i=1}^k \frac{P \cap G}{P \cup G} \quad (10)$$

其中, P 代表预测值; G 代表真实值; *MIoU* 越高则意味着分割越准确。

2.3.5 像素精度 (*Accu*)

像素精度即预测正确的像素占有所有像素的比例。像素精度是图像分割的最简单指标,其是正确分类的总像素除以总像素,可以理解为图像中正确

分类像素的百分比。计算公式可表示为

$$Accu = \frac{R}{A} \quad (11)$$

其中, R 代表预测准确的像素数, 而 A 代表图像中全部的像素数。同样, $Accu$ 越高则意味着分割结果越准确。

2.4 实验结果

在同一数据集上, 利用本文方法与 pix2pixHD 模型、pix2pix 模型、cycleGAN 和 styleGAN 等 4 个语义图像生成模型进行实验比较, 得出的峰值信噪比、结构相似性、 FID 、平均交并比和像素精度指标结果见表 1。

表 1 各模型图像评价指标对比表

Tab. 1 Comparison of image evaluation indexes of each model

Method	$PSNR(\uparrow)$	$SSIM(\uparrow)$	$MIoU(\uparrow)$	$Accu(\uparrow)$	$FID(\downarrow)$
cycleGAN	9.27	0.214	12.58	57.69	289.7
pix2pix	11.04	0.257	30.16	70.62	227.6
styleGAN	9.14	0.302	-	-	214.2
pix2pixHD	11.41	0.296	29.48	74.22	107.3
本文方法	12.46	0.327	31.79	77.68	57.4

从表 1 中可知, 本文方法在 COVID-19 数据集上生成的图像具有更高的峰值信噪比和结构相似性, 分别达到了 12.46 和 0.327, 说明该模型具有更高的图像质量; $MIoU$ 达到了 31.79, $Accu$ 达到了 77.68, 高于其它对比模型(由于 styleGAN 是无标签生成, 所以不适用此分析)。可见, 分割后依然能与标签取得较好的匹配, 说明实验结果取得了更好的

分割精度, 对后续训练分割或分类模型可以更好的起到扩充数据集, 提升模型准确度的作用。同时在 FID 上, 本文模型达到了 57.4, 较其它方法有显著性的提升, 具有更好的图像多样性。

在生成的图像细节与特点上, 本文提出的方法也能有效提升图像细节, 生成的图像结果如图 5 所示:

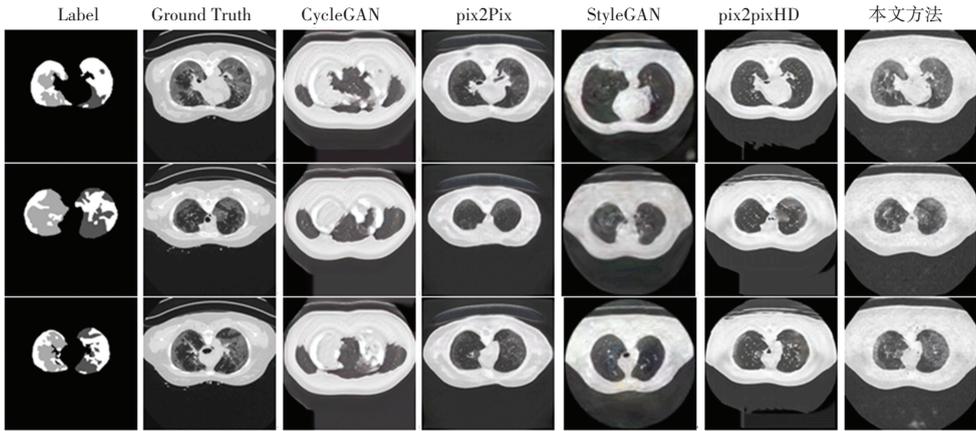


图 5 本文方法与其他对比方法实验结果图

Fig. 5 Experimental results of the improved pix2pixHD method and other comparative methods

由图 5 中可以发现, cycleGAN 生成的图像边界混乱, 没有现实意义; pix2pix 和 pix2pixHD 模型虽然具有清晰的肺部轮廓, 但是对于病灶区域的生成并不明显; styleGAN 模型生成的图像虽然较为模糊, 分辨率明显较低; 而本文方法生成的图像边界清晰, 毛玻璃区域明显且准确, 具有良好的现实意义, 可以起到扩充数据集, 提升分割模型的训练精度, 最终达到辅助医生诊断的效果。

2.4.1 图像质量评价

本文将通过峰值信噪比和结构相似性对生成图像的质量进行评价, 通过对本文方法和对比模型在每 20 次迭代时, 对测试集 $PSNR$ 进行统计, 结果如图 6 所示。

从图 6 中可知, 提出的模型在峰值信噪比上有较大提升, 不仅在质量上高于其他模型, 其收敛速度也有显著提升。本方法在该数据集上生成的图像峰

值信噪比可以达到 12.46, 高于 pix2pixHD、pix2pix 等模型。

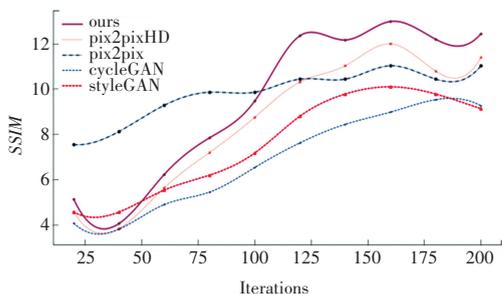


图 6 各模型的峰值信噪比迭代对比图

Fig. 6 Iterative comparison diagram of peak signal-to-noise ratio of each model

本文对图像的结构相似性 (SSIM) 进行了评价。通过对本文方法和对比模型在每 20 次迭代对测试集 SSIM 进行统计, 结果如图 7 所示:

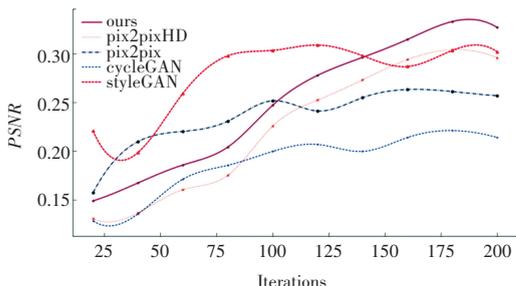


图 7 各模型的结构相似性迭代对比图

Fig. 7 Iterative comparison diagram of structural similarity of each model

从图 7 中可知, 本文所提出的模型对比其它方法具有更高的结构相似性, 说明生成的图像对比真实图像具有较高的相似性。对比 pix2pixHD 模型失真程度更小, 可以更好的保留图像的像素特点。

2.4.2 相关性和多样性评估

本文使用 Fr'echet Inception 距离, 对生成图像和真实图像进行相关性和多样性的定量评估, 实验结果如图 8 所示。

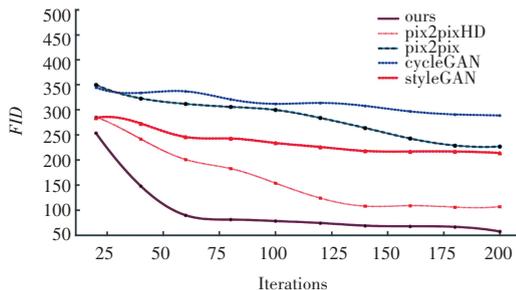


图 8 模型的 Fr'echet Inception 距离迭代对比图

Fig. 8 Iterative comparison diagram of FID of each model

通过图 8 可以看出: 提出的模型具有更低的 FID, 而且可以更快的达到收敛。本文模型取得了更好的相关性和多样性。对比 pix2pixHD 有所提高, 对比其它方法也具有明显的优势。

2.4.3 图像细节分析

除了在图像量化指标上的提高, 本方法生成的图像细节也有所提升, 图像细节对比如图 9 所示。

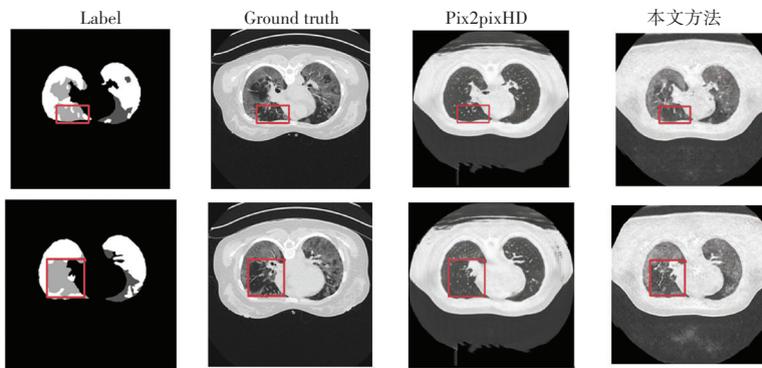


图 9 生成图像细节分析图

Fig. 9 Detail analysis diagram of the generated image

从图 9 中可以明显发现, 本文方法生成的图片可以在病变区域与正常区域看到明显的边界, 而 pix2pixHD 模型边界较为模糊, 几乎看不出病变区域边界, 使得生成的图像现实意义不大, 在后续的分割、分类任务中无法起到提升泛化能力的作用。

2.4.4 改进消融实验

为了证实加入 SPADE 残差块和 1/8 尺度鉴别器的效果, 本文将模型与 pix2pixHD、仅加入残差块和仅加入鉴别器的模型进行消融实验, 在各项指标上进行比较, 实验结果见表 2。

表 2 改进消融实验对比表

Tab. 2 Comparison table of improvement experiment

Method	PSNR(↑)	SSIM(↑)	FID(↓)	MIoU(↑)	Accu(↑)
pix2pixHD	11.41	0.296	107.3	29.48	74.22
pix2pixHD+SPADE block	12.34	0.316	87.3	29.89	74.68
pix2pixHD+1/8 scale discriminator	11.38	0.298	70.8	31.23	77.17
本文方法	12.46	0.327	57.4	31.79	77.68

从表 2 中可知, PSNR 和 SSIM 在加入 SPADE 残差块后有显著升高, 表现了 SPADE 残差块在 pix2pixHD 模型中提升图像质量的显著作用; 但在 MIoU 和 Accu 上并没有显著提升, 说明 SPADE 残差块对图像细节提升并不明显, DeepLabV2 模型并没有办法更准确的分割出病灶区域。但加入 1/8 尺度鉴别器可以提升图像细节部分, 使 MIoU 与 acc 指标得到提升, 与 SPADE 残差块形成很好的补充, 最终使得生成的图像与细节同时得到提升。

本文同时对加入每个模块的 PSNR 和 SSIM 进行了迭代统计, 结果如图 10、图 11 所示。加入了 SPADE 残差块的 pix2pixHD 模型, 对模型收敛速度也有着显著提升, 而未使用 SPADE 残差块的情况, 提前约 20 轮达到收敛。

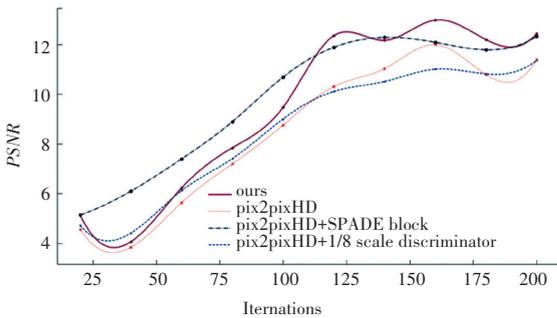


图 10 消融实验的峰值信噪比迭代对比图

Fig. 10 Iterative comparison diagram of the peak signal-to-noise ratio of improved experiment

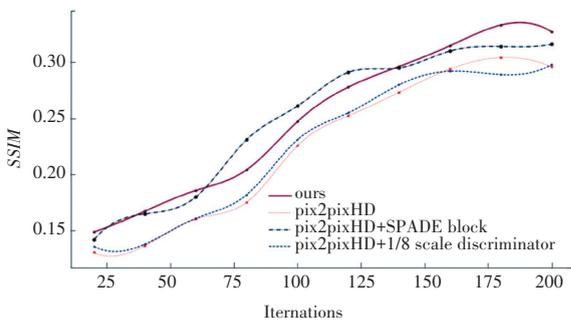


图 11 消融实验的结构相似性迭代对比图

Fig. 11 Iterative comparison diagram of structural similarity of improved experiment

3 结束语

为了解决 COVID-19 CT 图像生成质量不高, 边界不清晰, 病灶区域不明显的问题, 本文提出了一个改进的 pix2pixHD 模型。该网络通过改进鉴别器数量和生成器残差块的方式, 提升了生成图像的质量与细节。与 pix2pixHD 相比, 利用 SPADE 残差块补充了上采样过程中的信息损失, 减少了图像失真。实验结果表明, 文中提出的方法在 COVID-19 CT 图像数据集上对于生成图像的质量、多样性、相似性、匹配度上都有显著性提升, 可以有效解决 COVID-19 CT 图像较少的情况。

在后续的研究中, 模型还有进一步提高的潜力, 可在生成器部分使用更多样的特征提取方式或引入更高效鉴别器, 以提升模型的性能。同时研究高效的分割模型, 最终使生成的图像能够提升模型性能, 使医生可以得到合理的诊断建议。

参考文献

- [1] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016; 23.
- [2] 陈锬, 乔沁, 宋志坚. 生成对抗网络在医学图像处理中的应用[J]. 生命科学仪器, 2018, 16(4):11.
- [3] ZHOU L, BAE J, LIU H, et al. Lung swapping autoencoder: Learning a disentangled structure-texture representation of chest radiographs[J]. arXiv preprint arXiv:2201.07344, 2022.
- [4] JIANG Y, CHEN H, LOEW M, et al. COVID-19 CT image synthesis with a conditional generative adversarial network[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25(2): 441-452.
- [5] WANG T C, LIU M Y, ZHU J Y, et al. High-resolution image synthesis and semantic manipulation with conditional gans[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8798-8807.
- [6] Mirza M, Osindero S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.
- [7] PARK T, LIU M Y, WANG T C, et al. Semantic image synthesis with spatially-adaptive normalization[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 2337-2346.