

基于注意力机制结合残差神经网络的胃癌图像预测方法

郑子龙, 曾安

(广东工业大学计算机学院, 广东 广州 510006)

摘要: 随着深度学习的发展, 单图像超分辨率技术对医学图像的分析成为研究热点。传统方法对胃癌图像细节重建不充分, 丢失了图像的细节信息。提出一种新型高效的注意力机制结合残差神经网络的胃癌图像预测方法。对编码器输出的特征进行细化, 并根据改进的融合注意方法对空间注意模块和通道注意模块进行重组, 为像素融合提供更好的输入特征, 达到重建解码器的效果。最后, 通过添加像素融合后的输入特征, 从而加速网络高频残差的收敛, 提高图像重建效果。在重建放大倍数为2、3、4倍的条件下, 进行实验验证。实验结果表明, 在较低模型参数下, 方法具有最佳的PSNR和SSIM值。

关键词: 超分辨率重建; 注意力模块; 像素融合

中图分类号: TP18; TP391.41 文献标识码: A 文章编号: 1003-7241(2025)02-0136-05

Gastric Cancer Image Prediction Method Based on Attention Mechanism and Residual Neural Network

ZHENG Zi-long, ZENG An

(School of Computer Science, Guangdong University of Technology, Guangzhou 510006 China)

Abstract: With the development of deep learning, the analysis of medical images by single image super-resolution technology becomes a research hotspot. The traditional method is not enough to reconstruct the details of gastric cancer image, and the details of the image are lost. This paper presents a new and efficient method of gastric cancer image prediction based on attention mechanism and residual neural network. The output features of the encoder are refined, and the spatial attention module and channel attention module are reorganized according to the improved fusion attention method to provide better input features for pixel fusion, so as to achieve the effect of reconstructing the decoder. Finally, the input features after pixel fusion are added to accelerate the convergence of network high-frequency residuals and improve the image reconstruction effect. Experimental verification is carried out under the condition of reconstruction magnification of 2, 3 and 4 times. The experimental results show that the method has the best PSNR and SSIM values at lower model parameters.

Keywords: super resolution reconstruction; attention module; pixel fusion

0 引言

计算机视觉中的超分辨率技术是将低分辨率图像转换为高分辨率图像; 超分辨率重建是指在提高数字图像空间分辨率的方法。在诸多图像应用中起到关键作用^[1], 例如自然图像^[2]、卫星图像^[3]或医学图像^[4]。超分辨率重建算法可以根据过程中涉及的输入和输出图像的数量进行分类, 是医学图像预测技术的核心。

对于医学成像而言, 重建、去噪、超分辨率、分割对医学辅助诊断起到了巨大作用^[5]。Dong等人首次提出使用深度学习方法解决自然图像集上的超分辨率问题^[6]。近年来, 强化学习方法在自然图像超分辨率重建中的应用得到了广泛研究^[7]。强化学习方法也被提出用于医学图像中的超分辨率重建问题, 为辅助医学诊断提供了良好

条件^[8]。但是, 与自然图像超分辨率重建相比, 医学图像超分辨率重建技术在特定应用中需要额外的先验信息^[9]。由于医学成像中的随机共振技术通常在分割或诊断后进行, 因此增强感兴趣的结构并保留敏感信息是非常具有挑战性的。此外, 尤其是对于临床医学图像高分辨率和低分辨率图像的数据集相对较小且难以收集。

为了解决上述问题, 提出了一种高效的基于注意力机制结合残差神经网络的胃癌图像预测方法, 以增强超分辨率重建网络捕捉长距离上下文依赖的能力, 同时显著降低GPU内存成本。使用专门设计的高效多注意力机制来减少GPU内存消耗。

1 算法

1.1 模型总体架构

本文使用深度可分离卷积从输入图像中提取图像的

*基金项目: 广东省省级科技计划项目 (2019A050510041)

收稿日期: 2023-10-09

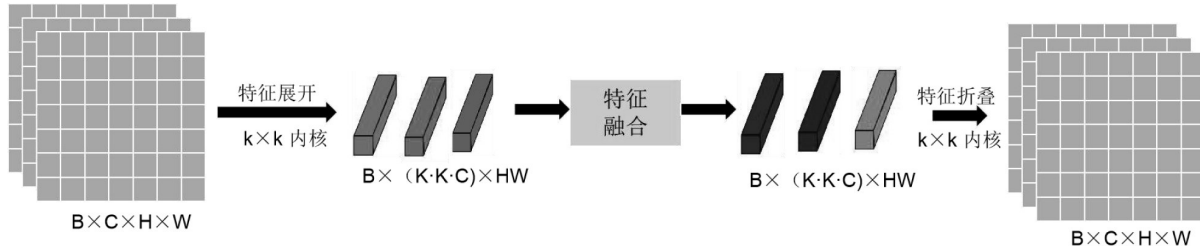


图4 轻量型融合特征网络模型

和 λ_x 分别是两条路径的两个自适应权重。

运算使用11个卷积层来改变通道数,以实现缩减和扩展功能。同时,将两个RU的输出串联起来,然后是一个11层卷积层,以充分利用分层特征。最后,采用3-3卷积层来减少特征映射的通道,并从融合的特征中提取有效信息。

图像中的相似图像块可以用作彼此的参考图像,因此可以参考其他图像块恢复当前图像块的纹理细节,适合使用网络迁移。然而,以前传统的视觉转换模型变体通常需要大量的GPU内存成本。本文提出了一种轻量型融合特征网络模型,可以以较低的计算成本捕获图像中相似局部区域的长期依赖性,如图4所示。

2 试验及结果讨论

2.1 实验设置

本文使用Python 2.7作为软硬件环境,操作系统平台为Ubuntu 20.04,提出的模型运行在两个NVIDIA 2080Ti GPU上。实验使用通用高质量图像数据集DIV2K进行训练和验证。在模型完成DIV2K的训练和验证后,使用公共数据集(Set5、Set14、B100、Urban100和Manga109)进行模型性能评估。批大小设置为64,迭代规模设置为500,一个迭代200次。重建比例设置为2。在此基础上,将3和4的比例应用于重建图像,与其他主流超分辨率重建算法进行比较。为了便于与其他先进模型(双三次插值、EDSR、MCRAN+)进行比较,本文在参考文献的基础上,选择峰值信噪比(PSNR)和结构相似性(SSIM)作为图像的质量评价指标。其中,峰值信噪比由均方误差(MSE)确定。峰值信噪比和SSIM值越高,图像重建效果越好。MSE、PSNR和SSIM的方程如下:

$$MSE = \frac{1}{H * W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} [x(i, j) - y(i, j)]^2 \quad (3)$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_x^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_x}{\sqrt{MSE}} \right) \quad (4)$$

$$SSIM(X, Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (5)$$

2.2 实施细节

为了验证模型中高频残差模块的有效性,设计了一组比较实验,以比较有无最近样本模块的性能。对于高

频残余模块的消融实验,比例为2。图6显示了在具有和不具有高频残差模块的DIV2K数据集上训练后的精度变化。从图5可以看出,当迭代为390时,使用高频残差模块的网络趋于平衡,不使用高频残差模块的网络趋于震荡波动且信PSNR较低。图6中的结果证明,添加最近的上采样模块可以改善模型的收敛特性。为了验证SFAN的有效性,对DIV2K验证集上的SFAN进行了定量评估,并与最近的模型Bicubic、MCRANC和EDSR进行了比较,如表1所示。结果表明,对于DIV2K数据集,在比例(2,3,4)的条件下,SFAN的峰值信噪比大于MCRANC,即SFAN对图像超分割重建的效果良好。在DIV2K数据集上训练网络后,模型的精度曲线如图6所示。

表1 Div2K验证的比较结果

重建倍数		DIV2K 验证集			
		双三次插值	EDSR	MCRAN+	SFAN
×2	PSNR	31.35	31.23	31.11	30.99
	SSIM	0.907 6	0.787 6	0.667 6	0.547 6
×3	PSNR	28.49	28.37	28.25	28.13
	SSIM	0.833 9	0.713 9	0.593 9	0.473 9
×4	PSNR	26.92	26.8	26.68	26.56
	SSIM	0.777 4	0.657 4	0.537 4	0.417 4

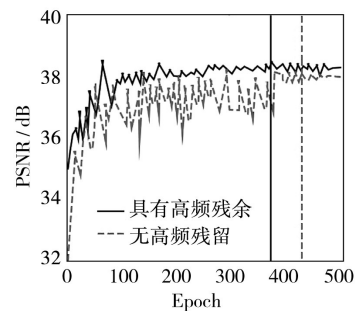


图5 消融研究高频残余模块

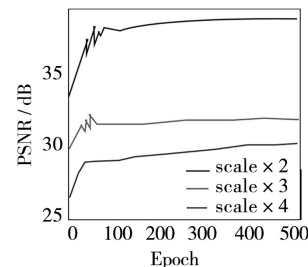


图6 模型训练收敛分析

表2 SISR模型的定量比较最佳结果和次佳结果

方法	重建倍数	参数	Set5	Set14	BSD100	Urban100	Manga100
			PSNR /SSIM	PSNR /SSIM	PSNR /SSIM	PSNR /SSIM	PSNR /SSIM
VDSR		666K	34.09/0.924 8	29.77/0.831 4	28.82/0.797 6	27.14/0.827 9	32.01/0.934 0
MemNet		678K	34.37/0.927 0	30.00/0.835 0	28.96/0.800 1	27.56/0.837 6	32.51/0.936 9
EDSR		1 555K	34.12/0.925 4	30.28/0.841 7	29.09/0.805 2	28.15/0.852 7	33.45/0.943 9
SRMDNF	×3	1 528K	34.29/0.925 5	30.04/0.838 2	28.97/0.802 5	27.57/0.839 8	33.00/0.940 3
CARN		1 592K	34.36/0.927 0	30.29/0.840 7	29.06/0.803 4	28.06/0.849 3	33.50/0.944 0
LatticeNet		765K	34.53/0.928 1	29.77/0.831 4	28.96/0.800 1	27.14/0.827 9	28.83/0.887 0
本文方法		770K	34.42/0.926 8	30.00/0.835 0	29.09/0.805 2	27.56/0.837 6	29.42/0.894 2
VDSR		666K	31.35/0.883 8	30.28/0.841 7	28.97/0.802 5	28.15/0.852 7	30.35/0.906 7
MemNet		678K	31.74/0.889 3	30.04/0.838 2	29.06/0.803 4	27.57/0.839 8	30.09/0.902 4
EDSR		1 518K	32.09/0.893 8	30.29/0.840 7	27.29/0.725 1	28.06/0.849 3	28.83/0.887 0
SRMDNF	×4	1 552K	31.96/0.892 5	28.01/0.767 4	27.40/0.728 1	25.18/0.752 4	30.47/0.908 4
CARN		1 592K	31.35/0.883 8	28.26/0.772 3	27.57/0.735 7	25.50/0.763 0	30.45/0.907 5
LatticeNet		777K	32.13/0.893 7	28.26/0.772 3	27.57/0.735 7	25.50/0.763 0	30.61/0.909 6
本文方法		751K	32.19/0.894 7	28.58/0.781 3	27.49/0.733 7	26.04/0.784 9	30.47/0.908 4

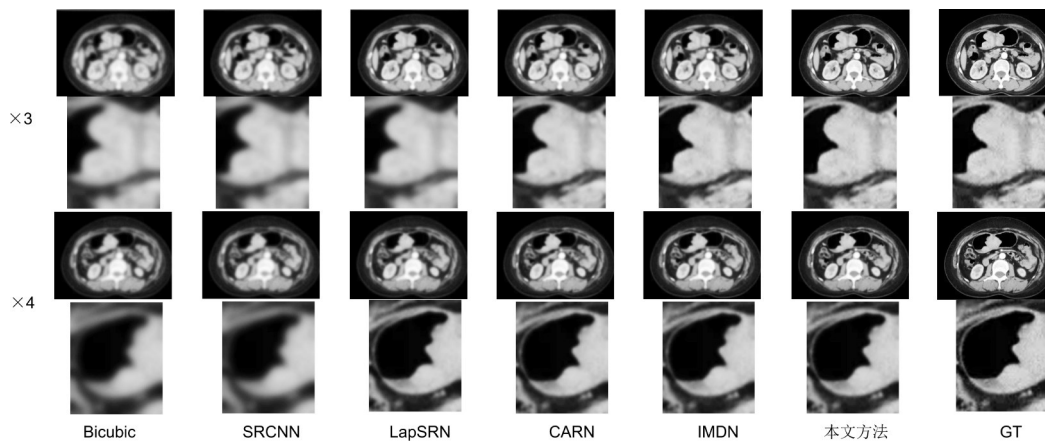


图7 对比模型的视觉可视化效果比较

在表2中,将本文提出方法与其他先进的超分辨率重建模型进行了比较。显然,本文提出方法在所有比例因子下都取得了有竞争力的结果。可以看出,虽然EDSR基线的性能接近本文提出方法,但其参数几乎是本文提出方法的两倍。此外,可以观察到,本文提出方法在Urban100上的性能比其他模型要好得多。这是因为该数据集的每个图像中都有许多相似的补丁。因此,在本文提出方法中引入的注意力机制可以用来捕捉这些相似图像块之间的长期依赖关系,并学习它们的相关性,从而获得更好的结果。

在图7中,显示了本文方法和其他传统主流模型在放大倍数3和4上的视觉比较。显然,本文提出方法重建的胃癌预测图像包含更精确的纹理细节,尤其是边缘和线条。值得注意的是,在2个尺度中,本文提出方法与其他超分辨率模型重建可视化效果更优。

3 结束语

为了提高胃癌图像的分辨率以达到预测胃癌的目

标,针对图像超分辨率重建网络中传统方法重建效果不佳,细节缺失等问题,提出了一种基于残差通道结合注意力网络模型。通过设置不同的剩余注意力模块来提取更多的图像特征。通过一种新的通道融合方法来重新定义和增强空间注意力机制模块和通道注意力机制模块中的特征信息,并使用像素的上采样方法来实现重建解码器的功能。通过整体结构化残差设置,将解码器输出的图像残差添加到输入图像,加速网络高频残差的收敛,从而提高图像重建的重建效果,减少模型参数的数量。

参考文献:

- [1] 曹泽红,刘高平,张志强,石峰,张煜.基于多模态MRI脑影像的超分辨率重建[J].南方医科大学学报,2022,42(7):1019-1025.
- [2] 李靖宇,程卫月,李子翔,林克正.超分辨率重建的微小人脸识别算法[J].哈尔滨理工大学学报,2022,27(3):52-58.DOI:10.15938/j.jhust.2022.03.007.
- [3] 朱海琦,李宏,李定文,李富.基于生成对抗网络的单图像超分辨率重建[J].吉林大学学报(理学版),2021,59(6):1491-1498.DOI:

(下转第184页)