

中图分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2023)01-0248-12

论文引用格式: Liu Y, Zhu W Y, Cheng J and Chen X. 2023. Multi-modal MR image super-resolution with residual dense attention network. Journal of Image and Graphics, 28(01):0248-0259 (刘羽, 朱文瑜, 成娟, 陈勋. 2023. 残差密集注意力网络多模态 MR 图像超分辨率重建. 中国图象图形学报, 28(01):0248-0259) [DOI:10.11834/jig.220286]

残差密集注意力网络多模态 MR 图像超分辨率重建

刘羽¹, 朱文瑜¹, 成娟¹, 陈勋^{2*}

1. 合肥工业大学生物医学工程系, 合肥 230009; 2. 中国科学技术大学电子工程与信息科学系, 合肥 230026

摘要: **目的** 现有医学图像超分辨率方法主要针对单一模态图像进行设计, 然而在磁共振成像 (magnetic resonance imaging, MRI) 技术的诸多应用场合, 往往需要采集不同成像参数下的多模态图像。针对单一模态的方法无法利用不同模态图像之间的关联信息, 很大程度上限制了重建性能。目前超分辨率网络模型参数量往往较大, 导致计算和存储代价较高。为此, 本文提出了一个轻量级残差密集注意力网络, 以一个统一的网络模型同时实现多模态 MR 图像的超分辨率重建。**方法** 首先将不同模态的 MR 图像堆叠后输入网络, 在低分辨率空间中提取共有特征, 之后采用设计的残差密集注意力模块进一步精炼特征, 再通过一个亚像素卷积层上采样到高分辨率空间, 最终分别重建出不同模态的高分辨率图像。**结果** 本文采用 MICCAI (medical image computing and computer assisted intervention) BraTS (brain tumor segmentation) 2019 数据集中的 T1 和 T2 加权 MR 图像对网络进行训练和测试, 并与 8 种代表性超分辨率方法进行对比。实验结果表明, 本文方法可以取得优于对比方法的重建性能。相对于对比方法中性能最优的方法, 本文方法在参数量低于对比模型的 10% 的情况下, 当放大因子为 2 时, T1 和 T2 加权模态的峰值信噪比 (peak signal to noise ratio, PSNR) 分别提高了 0.109 8 dB 和 0.415 5 dB; 当放大因子为 3 时, T2 加权模态的 PSNR 提高了 0.295 9 dB, T1 加权模态下降了 0.064 6 dB; 当放大因子为 4 时, T1 和 T2 加权模态的 PSNR 分别提高了 0.269 3 dB 和 0.042 9 dB。**结论** 本文提出的轻量级超分辨率网络能够充分利用不同模态 MR 图像之间的关联信息, 同时获得各输入模态图像高质量的超分辨率重建结果。

关键词: 图像超分辨率; 多模态 MR 图像; 卷积神经网络 (CNN); 残差学习; 密集连接; 注意力机制; 多模态信息融合

Multi-modal MR image super-resolution with residual dense attention network

Liu Yu¹, Zhu Wenyu¹, Cheng Juan¹, Chen Xun^{2*}

1. Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China;

2. Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230026, China

Abstract: **Objective** Image super-resolution (SR) is a sort of procedure that aims to reconstruct high resolution (HR) images from a given single or a set of low resolution (LR) images. This cost effective medical-related technique can improve the spatial resolution of images in terms of image processing algorithms. However, most of medical image super-resolution methods are focused on a single-modal super-resolution design. Current magnetic resonance imaging based (MRI-based) clinical applications, Multiple modalities are obtained by different parameter settings. In this case, a single modality super-

收稿日期: 2022-03-23; 修回日期: 2022-04-22; 预印本日期: 2022-04-29

* 通信作者: 陈勋 xunchen@ustc.edu.cn

基金项目: 国家自然科学基金项目 (62176081, 61922075, 62171176)

Supported by: National Natural Science Foundation of China (62176081, 61922075, 62171176)

resolution method cannot take advantage of the correlation information amongst multiple modalities, which limits the super-resolution capability. In addition, most existing deep learning based (DL-based) super-resolution models have constrained a number of trainable parameters, higher computational cost and memory storage in practice. To strengthen multi-modalities correlation information for reconstruction, our research is focused on a lightweight DL model (i.e., residual dense attention network) for multi-modal MR image super-resolution. **Method** A residual dense attention network is developed for multi-modal MR image super-resolution. Our network is composed of three parts: 1) shallow feature extraction, 2) feature refinement and 3) image reconstruction. Two of multi-modal MR images are input to the network after stacking. First, a 3×3 convolutional layer in the shallow feature extraction part is used to extract the initial feature maps in the low-resolution space. Next, the feature refinement part is mainly composed of several residual dense attention blocks. Each residual dense attention block consists of a residual dense block and an efficient channel attention module. Third, dense connection and local residual learning are adopted to improve the representation capability of the network in the residual dense block. The efficient channel attention module is facilitated the network to adaptively identify the feature maps that are more crucial for reconstruction. The outputs of all residual dense attention modules are stacked together and fed into two convolutional layers. These convolution layers are employed to reduce the channels of the feature maps as well as for feature fusion. After that, a global residual learning strategy is implemented to optimize the information flow further. The initial feature maps are added to the last layer through a skip connection. Finally, the obtained low-resolution feature maps in the image reconstruction part are up-scaled to the high-resolution space by a sub-pixel convolutional layer. Additionally, two symmetric branches are used to reconstruct the super-resolution results of the different modalities. To reconstruct the residual maps of the two modalities, each branch of them consists of two 3×3 convolutional layers. To obtain the final super-resolution results, the residual maps are linked to the interpolated low-resolution images. The popular L1 loss is used to optimize the network parameters. **Result** In the experiments, to verify the effectiveness of the proposed method, the MR images of two modalities (i.e., T1-weighted and T2-weighted) from the medical image computing and computer assisted intervention (MICCAI) brain tumor segmentation (BraTS) 2019 are adopted. The original MRI scans are split and segmented into a training set, a validation set and a testing set. To verify the effect of the multi-modal image super-resolution manner and the efficient channel attention module, two sets of ablation experiments are designed. The results show that these two of components can optimize the super-resolution performance more. Furthermore, eight representative image super-resolution methods are used for comparative analysis of performance in the experiments. Experimental results demonstrate that our method can improve these reference methods in terms of both of the objective evaluation and visual quality. Specifically, our method can obtain more competitive results as mentioned below: 1) when the up-scale factor is 2, the peak signal to noise ratio (PSNR) of the T1-weighted and T2-weighted modalities improve by 0.109 8 dB and 0.415 5 dB, respectively; 2) when the up-scale factor is 3, the PSNR of the T2-weighted modality improves by 0.295 9 dB while the T1-weight modality decreases by 0.064 6 dB; 3) when the up-scale factor is 4, the PSNR of the T1-weighted and T2-weighted modalities improve by 0.269 3 dB and 0.042 9 dB, respectively. It is worth noting that our network has a more than 10 times reduction in terms of network parameters compared to the popular reference method. **Conclusion** The correlation information of different modalities between MR images is beneficial to image super-resolution. Our multi-modal MR image super-resolution method can achieve high-quality super-resolution results of two modalities simultaneously in an integrated correlation information-based network. It can obtain more competitive performance than the state-of-the-art super-resolution methods with a relative lightweight model.

Key words: image super-resolution; multi-modal MR images; convolutional neural networks (CNN); residual learning; dense connection; attention mechanism; multi-modal information fusion

0 引言

影像技术在现代医疗系统中发挥着越来越重要

的作用,不同模态的医学图像能够表达不同的生理信息。磁共振成像(magnetic resonance imaging, MRI)由于技术的安全性和反映信息的全面性在临床医学中得到了非常广泛的应用。MRI系统可以

通过使用不同的成像参数生成多种模态的图像,如 T1 加权、T2 加权和质子密度加权等。不同模态的 MR 图像能够反映不同的生理信息,例如 T1 加权图像能够较好地反映解剖结构信息,而 T2 加权图像通常适用于体现组织的病变信息。空间分辨率是影响 MR 图像质量的一个重要因素,高分辨率(high-resolution, HR)的 MR 图像可以更加清晰准确地反映出结构和病变信息。然而,由于硬件配置、扫描时间和信噪比等因素的限制,原始 MR 图像的分辨率可能无法满足相关应用的需求。

图像超分辨率(重建)技术旨在从一幅或多幅低分辨率(low-resolution, LR)图像重建得到一幅高分辨率图像。该技术通过图像处理算法实现图像分辨率的提升,相对于依靠硬件的方式成本大幅降低,在医学图像处理领域具有重要的应用价值。众所周知,图像超分辨率实际上是一个高度病态的逆问题,也是图像处理领域的一个难点问题。在很多应用场合,往往输入只有一幅低分辨率图像,进一步增加了问题的不稳定性。近年来,基于单幅图像的超分辨率方法研究受到广泛关注,多种多样的方法不断提出,这些方法通常分为基于插值、基于重构和基于学习的方法 3 类。基于最近邻、双线性和双三次等插值算法的超分辨率方法具有运行速度很快、实现简单等优点,但这类方法结果中容易出现块效应和边缘模糊等问题;基于重构的方法通过构建包含图像先验的优化模型,提高超分辨率质量。Shi 等人(2015)提出了一种低秩全变分(low-rank and total variation, LRTV)模型用于 MR 图像的超分辨率重建,是一种典型的基于模型的医学图像超分辨率方法;基于学习的方法通过大量匹配的 HR 和 LR 样本来学习低分辨率图像到高分辨率图像的映射关系,从而实现图像的超分辨率。常用的机器学习模型包括稀疏表示、随机森林等。这类方法往往可以取得质量更高的重建结果,但是算法的计算效率有时较低。

随着深度学习理论技术的快速发展,基于深度学习的超分辨率方法相继提出。Dong 等人(2014)首次提出一种用于自然图像超分辨率的 3 层全卷积神经网络模型(super resolution with convolutional neural network, SRCNN),相对于传统方法取得了明显的性能提升。随后,越来越多基于深度学习的超分辨率方法陆续提出,并应用于医学图像超分辨率问

题。Shi 等人(2019)提出一种基于跳跃连接的宽残差网络(fixed skip connection-based wide residual network, FSCWRN)模型用于 MR 图像超分辨率。Chen 等人(2020)在反馈网络(Li 等, 2019)的启发下,提出一种结合密集块和反馈机制的网络框架(feedback adaptive weighted dense network, FAWDN)用于医学图像超分辨率。

值得指出的是,当前医学图像超分辨率研究还存在一些不足之处。首先,现有超分辨率方法主要针对单一模态的医学图像进行设计。事实上,在磁共振成像技术的大多数应用场合,通常需要采集不同成像参数下的多模态 MR 图像,如 T1 加权、T2 加权等。此时,针对单一模态的超分辨率方法没有利用不同模态 MR 图像之间的关联信息,很大程度上限制了重建性能,且对不同模态图像的超分辨率任务各自进行模型训练也会导致计算资源与存储资源的浪费。其次,现有大多数超分辨率方法通常采用加深网络深度的策略来获得较好的重建结果,随着网络的加深,网络需要学习的参数量也随之增加,导致计算效率降低和存储消耗增加,影响方法的实用性。针对上述问题,本文提出一种基于轻量级残差密集卷积网络的多模态 MR 图像超分辨率方法,主要创新点包括:1)与现有单一模态医学图像超分辨率方法不同,本文方法以一个统一的网络模型同时实现两种模态 MR 图像的超分辨率重建,能够充分利用不同模态图像之间的关联信息,提升超分辨率质量;2)提出了一个轻量级的超分辨率网络模型,与现有性能相当的网络相比,参数量低于对比模型的 10%,方法实用性更强。

1 相关工作

卷积神经网络由于强大的特征提取与表达能力,在图像超分辨率这个不稳定性问题中获得了十分优异的结果。Dong 等人(2014)首次将卷积神经网络引入图像超分辨率领域,提出一个端到端的 3 层卷积神经网络 SRCNN,该网络参考稀疏表示的原理,通过学习低分辨率图像块与高分辨率图像块之间的映射关系,获得了超越当时其他大多数方法的优异效果。之后, Kim 等人(2016a)提出了一个深度更大的超分辨率网络(very deep convolutional networks for super-resolution, VDSR)模型,通过加深网

络深度以获得更加优秀的重建结果,并利用残差学习和梯度裁剪来解决梯度爆炸和梯度消失的问题,使深度网络变得易于训练。随后,进一步结合VDSR的优势提出一种基于深度递归卷积神经网络(deeply-recursive convolutional network, DRCN)的超分辨率模型(Kim等,2016b),采用递归监督的学习策略成功减少了模型参数,大幅提升了网络的计算速度,节省了计算和存储资源。这些方法虽然相较于以往的超分辨率方法在重建性能方面都有一定提升,但是都使用插值后的LR图像作为输入图像,不仅会因为在高分辨率空间中进行多次卷积操作而增加大量的计算量,而且也无法真正建立从原始LR图像到HR图像的映射。

为了解决这一问题,Dong等人(2016)利用转置卷积作为上采样,提出一个快速超分辨率卷积神经网络(fast super-resolution convolutional neural network, FSRCNN)模型,采用原始LR图像作为整个网络的输入,在不降低重建质量的情况下,显著缩小了网络的参数量并提升了网络的计算速度。Shi等人(2016)提出高效亚像素卷积神经网络(efficient sub-pixel convolutional network, ESPCN)模型,采用一种高效亚像素卷积层作为上采样,与转置卷积相比具有精度更高的优势。该卷积层在后续许多工作中都得以应用,例如改良型深度超分辨率网络(enhanced deep super-resolution, EDSR)(Lim等,2017)、基于特征精炼模块的超分辨率网络(coarse-to-fine super-resolution, CFSRCNN)(Tian等,2021)等。这些上采样方式使超分辨过程中的特征提取和加工步骤在LR空间中进行,不仅节省了计算资源和训练时间,使网络更加容易收敛,而且真正建立了原始LR图像与HR图像的之间端到端的映射关系。之后,Tong等人(2017)提出一种用于超分辨率的密集网络(super-resolution via dense network, SRDenseNet),采用跳跃连接和密集块来达到重建良好质量的高分辨率图像。在此基础上,Zhang等人(2018)提出了残差密集网络(residual dense network, RDN)模型,通过级联多个残差密集块的方式重建出了高质量的高分辨率图像。在RDN启发下,应自炉和龙祥(2019)为充分提取LR图像的特征信息,采用不同大小的卷积核对LR图像进行卷积,提出一种多尺度密集残差网络。程德强等人(2021)在RDN的基础上,引入多通道交叉学习策略获得较为理想的结

果。沈明玉等人(2019)为克服链式结构网络容易丢失低频信息的问题,提出一种多阶段融合网络并成功取得良好结果。佟骏超等人(2019)采用递归学习的策略,提出了递归式多阶段融合网络,使网络在不断加深的同时获得质量优异的超分辨率图像。随着注意力机制在计算机视觉任务中的广泛应用,雷鹏程等人(2020)融合空间注意力机制和分层特征融合机制,提出一种轻量级的超分辨率网络,在使用较少参数情况下获得了理想的重建结果。周波等人(2021)提出一种改良的通道注意力机制,并将其应用于深度残差密集网络,在重建性能方面有所提升。

近年来,医学图像在临床实践中大量应用。但高分辨率的医学图像由于成像设备的限制常常难以获得,于是许多超分辨方法应用于医学图像领域。Xue等人(2020)提出一种渐进子频段残差超分辨率网络(progressive sub-band residual super-resolution network, PSR-SRN),使重建出的高分辨率MR图像具有更多的细节信息。Chen等人(2020)基于递归神经网络(recursive neural network, RNN)的原理,利用反馈机制提出用于医学图像的超分辨模型,该网络在多次循环中不断修正之前步骤产生的误差。Zeng等人(2018)提出一种结合T1加权和T2加权两种模态MR图像关联信息的超分辨率网络,但其中T2加权图像只是用于辅助T1加权图像的超分辨率重建。上述算法虽然取得了良好的重建效果,但依然存在不足。首先,目前医学图像超分辨率方法通常只针对单一模态的图像进行重建,忽视了不同模态图像之间存在的关联信息,不利于充分提升超分辨率质量;其次,网络结构复杂且深度通常较大,导致需要消耗较多的计算与存储资源,影响方法的实用性。针对这些问题,本文提出了一种基于轻量级密集残差注意力网络的多模态MR图像超分辨方法。

2 方 法

2.1 网络整体结构

图1是本文提出的多模态MR图像超分辨率方法的网络结构示意图,该网络主要由浅层特征提取、特征精炼和图像重建3部分构成。网络以低分辨率的两种模态MR图像对 I_{LR1} 和 I_{LR2} 作为输入,最终输出相应的超分辨率图像 I_{SR1} 和 I_{SR2} 。

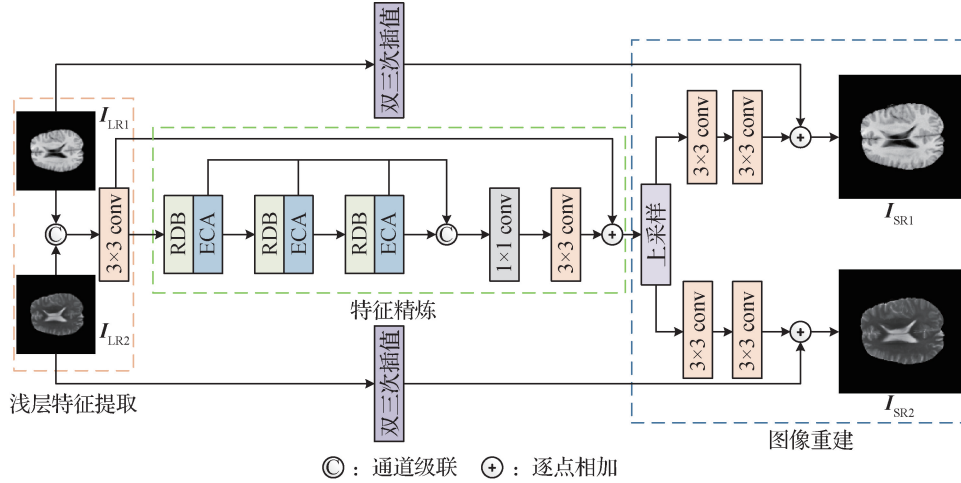


图1 本文提出的多模态 MR 图像超分辨率方法的网络结构示意图

Fig. 1 The network architecture of the proposed multi-modal MR image super-resolution method

2.1.1 浅层特征提取

为了利用不同模态图像之间的关联信息,将两个模态的 MR 图像堆叠在一起后,输入一个卷积核大小为 3×3 的卷积层提取浅层特征,该过程可表达为

$$\mathbf{F}_{\text{init}} = f_{\text{relu}}(\text{Conv}(\text{Cat}(\mathbf{I}_{\text{LR1}}, \mathbf{I}_{\text{LR2}}))) \quad (1)$$

式中, Cat 表示堆叠操作, Conv 表示卷积核大小为 3×3 的卷积层, $f_{\text{relu}}(\cdot)$ 表示 ReLU 激活函数, \mathbf{F}_{init} 表示提取到的浅层特征图,通道数默认设置为 64。

2.1.2 特征精炼

在特征精炼部分,浅层特征被输入到 K 个级联的残差密集注意力模块进行进一步的特征加工。残差密集注意力模块由 1 个残差密集块 (residual dense block, RDB) 和 1 个高效通道注意力 (efficient channel attention, ECA) 模块构成。每个残差注意力模块均输出 64 通道的特征图,将 K 个残差密集注意力模块的输出堆叠获得 $64 \times K$ 通道的特征图后进行全局特征融合。全局特征融合由 1 个卷积核大小为 1×1 的卷积层和 1 个卷积核大小为 3×3 的卷积层完成,前者用于将 $64 \times K$ 通道特征降低为 64 通道特征图并进行初步特征融合,后者用来进一步融合特征。获得融合特征后,再将浅层特征与融合特征相加,得到最终的低分辨率空间特征,形成全局残差学习,该过程可表达为

$$\mathbf{F}_{\text{RDA}}^1 = f_{\text{RDA}}^1(\mathbf{F}_{\text{init}}) \quad (2)$$

$$\mathbf{F}_{\text{RDA}}^k = f_{\text{RDA}}^k(\mathbf{F}_{\text{RDA}}^{k-1}), k \geq 2 \quad (3)$$

$$\mathbf{F}_{\text{C}} = \text{Cat}(\mathbf{F}_{\text{RDA}}^1, \dots, \mathbf{F}_{\text{RDA}}^k) \quad (4)$$

$$\mathbf{F}_{\text{R}} = f_{\text{relu}}(\text{Conv}(f_{\text{relu}}(\text{Conv}_{1 \times 1}(\mathbf{F}_{\text{C}})))) + \mathbf{F}_{\text{init}} \quad (5)$$

式中, $f_{\text{RDA}}^1(\cdot)$ 和 $f_{\text{RDA}}^k(\cdot)$ 分别表示第 1 个和第 k 个残差密集注意力模块的映射函数, $\text{Conv}_{1 \times 1}$ 表示卷积核大小为 1×1 的卷积层, $\mathbf{F}_{\text{RDA}}^1$ 和 $\mathbf{F}_{\text{RDA}}^k$ 分别表示第 1 个和第 k 个残差密集注意力模块输出的特征图, \mathbf{F}_{R} 表示精炼后的特征图。

2.1.3 图像重建

图像重建部分的输入是特征精炼部分得到的低分辨率特征图 \mathbf{F}_{R} 。首先,使用亚像素卷积层将特征图上采样到高分辨率空间。亚像素卷积层是一种在图像超分辨领域广泛使用的上采样层,与同样具有上采样功能的转置卷积相比,亚像素卷积层不引入待学习参数,仅通过重新排列不同通道的特征图来达到上采样的目的,且在上采样前不需进行零填充。然后,将上采样后的特征图分别输入到两个对称的高分辨率特征重建支路,每条支路包含 2 个卷积核大小为 3×3 的卷积层,前者输出 64 通道的特征图,后者重建得到不同模态的单通道残差图。最后,将两个模态的残差图分别与对应的低分辨率图像双三次插值结果相加,得到超分辨率图像。上述过程可表达为

$$\mathbf{I}_{\text{SR1}} = f_{\text{bi}}(\mathbf{I}_{\text{LR1}}) + \text{Conv}(f_{\text{relu}}(\text{Conv}(f_{\text{up}}(\mathbf{F}_{\text{R}})))) \quad (6)$$

$$\mathbf{I}_{\text{SR2}} = f_{\text{bi}}(\mathbf{I}_{\text{LR2}}) + \text{Conv}(f_{\text{relu}}(\text{Conv}(f_{\text{up}}(\mathbf{F}_{\text{R}})))) \quad (7)$$

式中, $f_{\text{up}}(\cdot)$ 表示亚像素卷积操作, $f_{\text{bi}}(\cdot)$ 表示双三次插值操作。

2.2 残差密集注意力模块

残差密集注意力模块是网络特征精炼部分的核心单元,由1个残差密集块(RDB)和1个高效通道注意力(ECA)模块级联而成。

2.2.1 残差密集块

RDB的网络结构如图2所示。1个RDB由若干个卷积核大小为 3×3 的卷积层和1个卷积核大小为 1×1 的卷积层构成。在 3×3 卷积层之间采用密集连接,允许浅层特征不断传输到更深层网络中,使浅层特征能够不断复用。这里,每个 3×3 卷积层输出特征图的通道数称为增长率,即每次级联后特征图增加的数量。卷积层数量和增长率越大,网络

的表示能力就越强。本文方法中,每个RDB中默认包含6个 3×3 卷积层,增长率设置为32,由于输入特征通道数量为64,故6个级联操作后的特征图的数量分别为96、128、160、192、224和256。 1×1 卷积层有两个作用,一是降低特征图的通道数(降为输入特征通道数量64),减少计算与存储代价;二是将密集连接产生的特征进行融合。最后,输入的特征图通过一个跳跃连接与获得的特征图相加,形成一个局部的残差学习机制,能够将浅层特征信息传递到深层网络,可以有效避免网络在不断提取深度特征时丢失浅层特征,同时加速梯度流动,避免梯度消失导致网络难以训练。

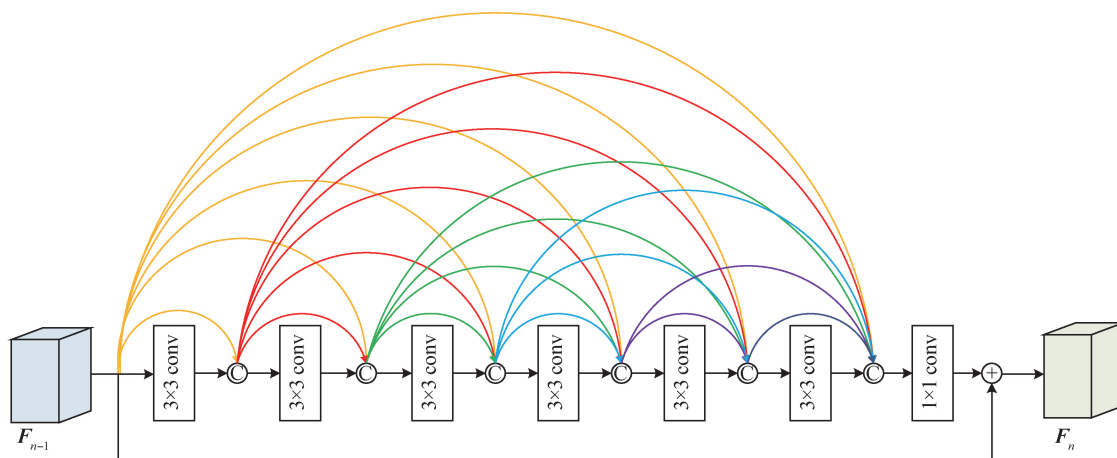


图2 残差密集块结构示意图

Fig. 2 The architecture of the residual dense block (RDB) module

2.2.2 高效通道注意力模块

在每个RDB之后,采用ECA模块(Wang等, 2020)对RDB得到的特征进一步精炼。ECA模块的结构如图3所示。该模块首先对输入的特征图进行一次全局平均池化(global average pooling, GAP),综合全局信息,然后将获得的特征向量输入到一个

卷积核大小可以自适应调节的1维卷积层中,得到的特征向量再经过softmax操作获得权重向量,最后将该权重向量与输入特征图相乘,得到加权后的输出特征图。该注意力机制模块使网络能够自适应地对不同通道特征图的权重进行调节,增大对重建效果有重要影响通道的权重,提升网络的特征

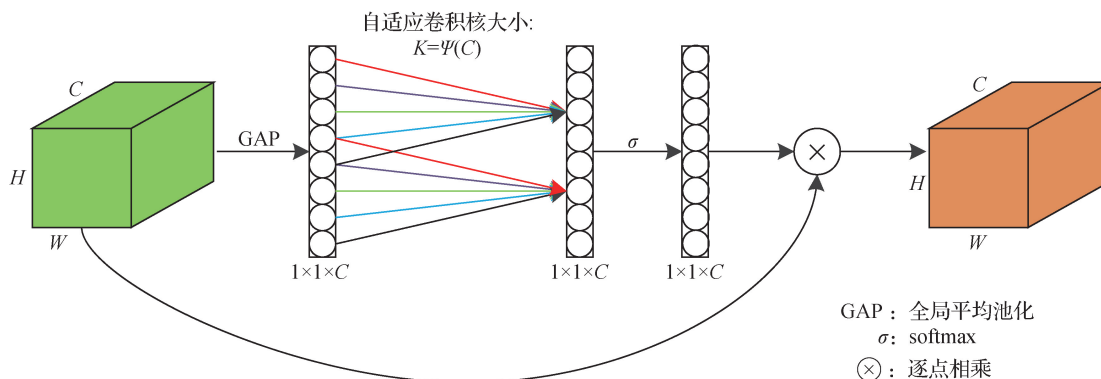


图3 高效通道注意力模块结构示意图

Fig. 3 The architecture of the efficient channel attention (ECA) module

学习能力。此外,采用可自适应调节大小的卷积核有助于提高模型的计算效率。

2.3 损失函数

基于 L1 范数的损失函数在超分辨率问题中广泛使用,具有形式简单和效果良好等优点。本文同样采用 L1 范数定义损失函数来实现网络参数优化。对于训练样本集中一批包含 M 个图像对的子集 $\{I_{LR1}^i, I_{LR2}^i, I_{GT1}^i, I_{GT2}^i\}_{i=1}^M$, 损失函数包含对称的两部分,分别对应两种模态图像的损失,具体表达为

$$L = \frac{1}{M} \sum_{i=1}^M \|I_{GT1}^i - I_{SR1}^i\|_1 + \frac{1}{M} \sum_{i=1}^M \|I_{GT2}^i - I_{SR2}^i\|_1 \quad (8)$$

式中, I_{GT1} 和 I_{GT2} 分别表示两种模态的参考图像 (ground truth, GT), 即原始高分辨率图像, 网络输入低分辨率图像 I_{LR1} 和 I_{LR2} 由其双三次插值下采样得到。 I_{SR1} 和 I_{SR2} 分别表示模型重建得到的超分辨率图像。

3 实验

3.1 实验设置

3.1.1 数据集

实验采用 MICCAI (medical image computing and computer assisted intervention) BraTS (brain tumor segmentation) 2019 数据集的 T1 加权和 T2 加权 MR 图像进行方法的有效性验证。该数据集包含 457 组大小为 $240 \times 240 \times 155$ 的 T1 加权和 T2 加权 MR 图像对, 实验时, 将其随机分为 3 份, 训练集 417 组、验证集 20 组、测试集 20 组。从每组 3D 图像数据中选取 10 幅切片 (从第 60 个切片开始每隔 5 个切片选取一幅) 作为原始高分辨率图像, 共获得 4 570 对 T1 加权和 T2 加权 MR 图像。其中, 训练集 4 170 对、验证集 200 对、测试集 200 对。

3.1.2 训练细节

为了控制网络的参数量, 在特征精炼部分采用 3 个残差密集注意力模块, 即参数 K 设置为 3。实验中, 训练 3 个不同放大因子的网络 ($\times 2$ 、 $\times 3$ 、 $\times 4$)。网络训练的总轮次设置为 300 轮, 学习率采用动态调整策略, 前 100 轮学习率为 1×10^{-4} , 之后降为 1×10^{-5} , 使用 Adam 优化器实现网络参数优化。实验在 PyTorch 深度学习框架下实现, GPU 为 NVIDIA GeForce GTX TITAN, 网络训练大约需要

9 h。

3.1.3 性能评价

使用峰值信噪比 (peak signal to noise ratio, PSNR) 和结构相似性 (structural similarity, SSIM) 作为客观评价指标。PSNR 是超分辨率研究中最常用的性能评价指标, 该指标数值越大, 表明图像重建质量越好, 但某些情况下 PSNR 并不能完全反映图像的重建质量。SSIM 是用来评价两幅图像结构相似性的指标, 图像重建质量越高, SSIM 值越接近于 1。

3.2 消融实验

为了验证网络各主要部分的作用, 设计了两组消融实验。1) 为了验证不同模态 MR 图像之间的关联信息对超分辨率重建性能的提升作用, 设计了两个针对单一模态图像的超分辨率网络, 分别命名为 T1-Net 和 T2-Net, 其网络结构与完整版网络 (Full-Net) 的不同之处仅在于只输入单一模态的 MR 图像 (T1 加权或 T2 加权), 并且在图像重建模块仅有一个重建分支; 2) 为了验证高效通道注意力模块对网络性能的提升作用, 设计了一个取消该注意力模块的网络模型, 命名为 NA-Net。上述消融实验均在相同实验设置下进行训练, 不同放大因子下的客观评价结果如表 1—表 3 所示。可以发现, 在各个放大因子 ($\times 2$, $\times 3$, $\times 4$) 下, 完整网络 Full-Net 在 PSNR 和 SSIM 上都明显优于仅针对单一模态的超分辨率模型, 体现出不同模态图像之间的关联信息对超分辨率重建是有益的。同时, 注意力机制模块的使用也可以显著提升超分辨率质量, 有助于在不过度增加网络深度的情况下获得较好的性能。需要指出的是, 注意力机制模块的影响相对更为明显, 这可能是由于网络模型的深度较小, 导致增加注意力机制模块能够使网络表示能力获得显著提升。

表 1 放大因子为 $\times 2$ 时的消融实验结果

Table 1 The results of ablation study when upscale factor is $\times 2$

模型	PSNR-T1/dB	SSIM-T1	PSNR-T2/dB	SSIM-T2
T1-Net	40.418 4	0.986 6	-	-
T2-Net	-	-	39.678 2	0.988 2
NA-Net	37.490 2	0.979 5	37.546 5	0.982 7
Full-Net	42.794 2	0.990 4	41.473 2	0.989 6

注: 加粗字体表示各列最优结果, “-”表示缺省数据。

表2 放大因子为×3时的消融实验结果
Table 2 The results of ablation study when
upscale factor is ×3

模型	PSNR-T1/dB	SSIM-T1	PSNR-T2/dB	SSIM-T2
T1-Net	34.583 4	0.967 3	-	-
T2-Net	-	-	35.187 9	0.973 0
NA-Net	33.323 6	0.952 7	33.308 5	0.955 3
Full-Net	37.498 2	0.978 5	36.694 3	0.976 2

注:加粗字体表示各列最优结果,“-”表示缺省数据。

表3 放大因子为×4时的消融实验结果
Table 3 The results of ablation study when
upscale factor is ×4

模型	PSNR-T1/dB	SSIM-T1	PSNR-T2/dB	SSIM-T2
T1-Net	32.830 1	0.952 1	-	-
T2-Net	-	-	32.664 4	0.953 1
NA-Net	31.132 1	0.923 6	31.066 8	0.923 9
Full-Net	34.770 0	0.963 5	34.016 8	0.958 9

注:加粗字体表示各列最优结果,“-”表示缺省数据。

3.3 对比实验

为了验证本文方法的性能,与双三次插值法(Bicubic)、LRTV(Shi等,2015)、EDSR(Lim等,2017)、VDSR(Kim等,2016a)、FSRCNN(Dong等,2016)、FAWDN(Chen等,2020)、CFSRCNN(Tian等,2021)和RDN(Zhang等,2018)等有代表性的超分辨率方法进行对比。公平起见,所有针对自然图像的超分辨率方法都在本文使用的数据集上分别对T1加权和T2加权两种模态进行重新训练,保证其针对医学图像的性能。

表4—表6给出了不同放大因子下各种方法的客观评价结果。由表4可以看出,放大因子为×2时,本文方法的T1加权和T2加权两种模态超分辨率结果的PSNR均为最优,比次优方法RDN分别提高了0.109 8 dB和0.415 5 dB;本文方法的T1加权超分辨率结果的SSIM与RDN相同,均为最优,T2加权超分辨率结果的SSIM为最优,与次优的RDN相比,提高了0.000 4。由表5可以看出,放大因子为×3时,本文方法的T1加权图像超分辨率结果的PSNR和SSIM均为次优,与最优的RDN相比,仅分别下降0.064 6 dB和0.000 3,T2加权图像超分辨率结果的PSNR和SSIM则均为最优,比次优的RDN

分别提高0.295 9 dB和0.000 6。由表6可以看出,放大因子为×4时,本文方法的T1加权和T2加权两种模态超分辨率结果的PSNR均为最优,比次优方法RDN分别提高0.269 3 dB和0.042 9 dB;本文方法的T1加权图像超分辨率结果的SSIM为次优,与最优的RDN相比,下降了0.000 4,T2加权图像超分辨率结果的SSIM为最优,与次优的RDN相比提高了0.000 9。总体来看,在3种放大因子下,本文方法的客观评价结果均优于对比方法。与传统方法Bicubic和LRTV相比,优势非常明显;与除RDN外的深度学习方法相比,在PSNR和SSIM

表4 放大因子为×2时不同方法的客观评价结果
Table 4 Objective evaluation results of different
methods when upscale factor is ×2

方法	PSNR-T1/dB	SSIM-T1	PSNR-T2/dB	SSIM-T2
Bicubic	34.775 1	0.976 3	33.961 2	0.971 4
LRTV	35.139 6	0.970 6	34.362 8	0.966 0
EDSR	35.245 1	0.981 8	36.394 4	0.981 1
FSRCNN	41.243 5	0.988 9	37.307 4	0.986 3
VDSR	36.845 7	0.984 3	36.633 7	0.982 9
FAWDN	38.507 3	0.986 4	37.973 9	0.984 5
CFSRCNN	42.559 4	0.990 2	40.754 9	0.988 9
RDN	<u>42.684 4</u>	0.990 4	<u>41.057 7</u>	<u>0.989 2</u>
本文	42.794 2	0.990 4	41.473 2	0.989 6

注:加粗字体表示各列最优结果,下划线字体表示各列次优结果。

表5 放大因子为×3时不同方法的客观评价结果
Table 5 Objective evaluation results of different
methods when upscale factor is ×3

方法	PSNR-T1/dB	SSIM-T1	PSNR-T2/dB	SSIM-T2
Bicubic	32.186 6	0.953 6	31.338 4	0.944 2
LRTV	33.069 6	0.957 7	32.200 7	0.950 0
EDSR	31.981 9	0.961 6	33.105 0	0.959 0
FSRCNN	36.287 4	0.974 1	34.534 3	0.969 4
VDSR	33.968 8	0.968 1	33.528 0	0.964 4
FAWDN	34.862 8	0.969 1	34.092 0	0.963 8
CFSRCNN	37.191 9	0.977 9	36.131 1	0.974 2
RDN	37.562 8	0.978 8	<u>36.398 4</u>	<u>0.975 6</u>
本文	<u>37.498 2</u>	<u>0.978 5</u>	36.694 3	0.976 2

注:加粗字体表示各列最优结果,下划线字体表示各列次优结果。

表 6 放大因子为 $\times 4$ 时不同方法的客观评价结果
Table 6 Objective evaluation results of different methods when upscale factor is $\times 4$

方法	PSNR-T1/dB	SSIM-T1	PSNR-T2/dB	SSIM-T2
Bicubic	30.350 6	0.927 8	29.585 1	0.914 3
LRTV	30.715 2	0.934 9	29.954 6	0.923 3
EDSR	29.645 6	0.931 1	30.665 7	0.928 3
FSRCNN	33.731 8	0.955 5	32.746	0.949
VDSR	31.812 8	0.944 2	31.411 5	0.938 2
FAWDN	32.735 9	0.947 5	31.862 2	0.938 1
CFSRCNN	<u>34.727 1</u>	0.963 1	33.543 9	0.956 5
RDN	34.698 7	0.963 9	<u>33.747 5</u>	<u>0.958</u>
本文	34.77	<u>0.963 5</u>	34.016 8	0.958 9

注:加粗字体表示各列最优结果,下划线字体表示各列次优结果。

指标上也明显占优;与 RDN 方法的性能较为接近(本文方法占优情形相对更多一些),但本文网络模型参数量少得多。

图 4—图 6 分别展示了不同放大因子下各种方法的超分辨率重建结果,每幅图的第 1 行是 T1 加权 MR 图像,第 2 行是 T2 加权 MR 图像。为便于对比,每个结果中均给出了两个局部放大区域。相对于其他方法,本文方法可以更好地保持图像细节信息。例如,观察图 4 中的放大区域,可以发现本文方法能够获得与参考图像(GT)更为接近的重建效果,双三次插值法在 T1 加权和 T2 加权的重建图像中均丢失边缘细节信息,其他对比方法中,重建效果较好的 RDN 和 CFSRCNN 对 T1 加权模型的超分辨率效果较好,但在 T2 加权超分辨率图像中丢失了一些细节信息。

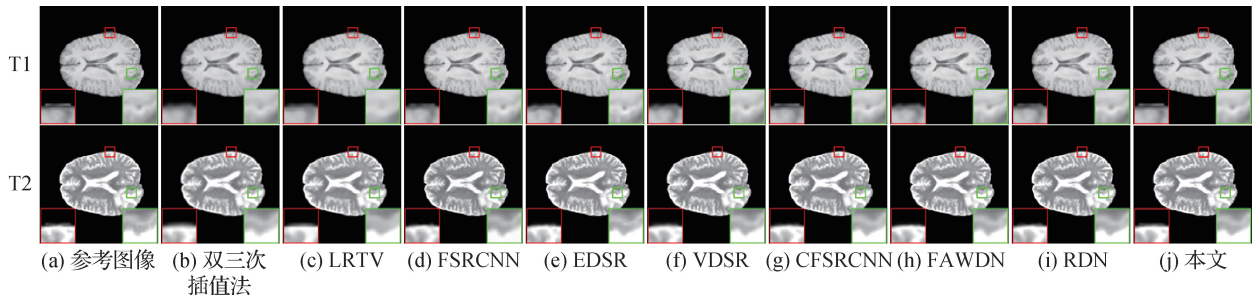


图 4 放大因子为 $\times 2$ 时不同方法的一组超分辨率结果图

Fig. 4 A set of super-resolution results obtained by different methods when the scale factor is $\times 2$

((a) GT; (b) bicubic; (c) LRTV; (d) FSRCNN; (e) EDSR; (f) VDSR; (g) CFSRCNN; (h) FAWDN; (i) RDN; (j) ours)

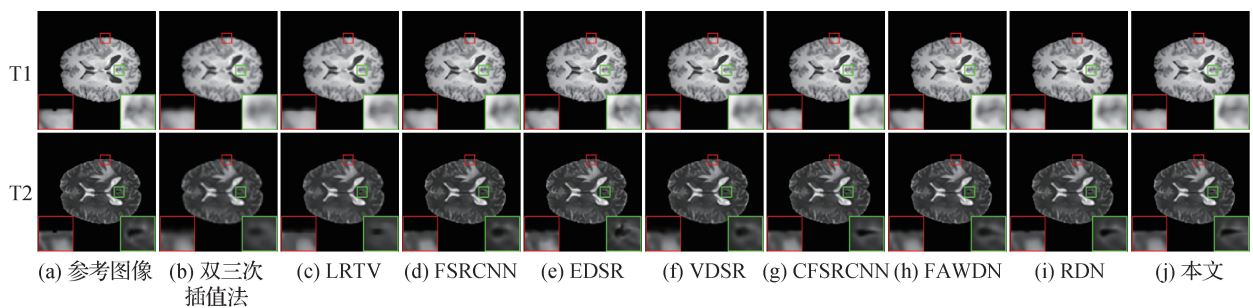


图 5 放大因子为 $\times 3$ 时不同方法的一组超分辨率结果图

Fig. 5 A set of super-resolution results obtained by different methods when the scale factor is $\times 3$

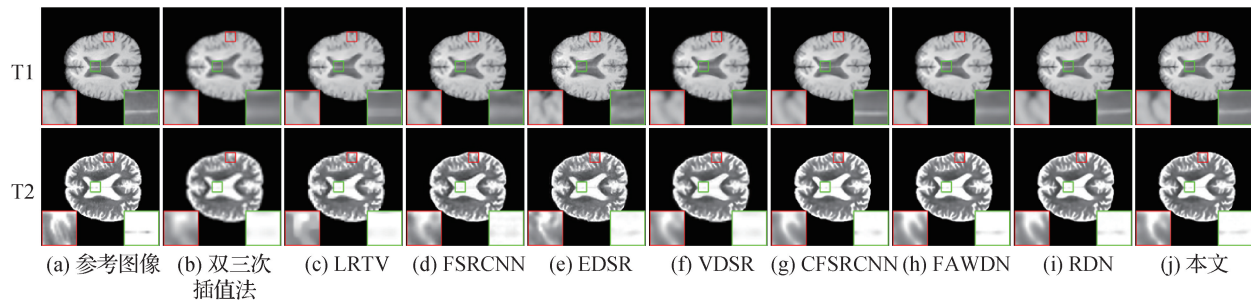
((a) GT; (b) bicubic; (c) LRTV; (d) FSRCNN; (e) EDSR; (f) VDSR; (g) CFSRCNN; (h) FAWDN; (i) RDN; (j) ours)

3.4 网络模型参数量与运行时间比较

一般而言,网络模型的计算时间和存储消耗与网络的参数量成正相关,故本文通过比较不同网络中的可训练参数量来衡量网络的复杂度。

图 7 展示了深度学习对比方法的网络模型参数

量。可以看出,FSRCNN(Dong 等,2016)因为网络结构过于简单,含有参数量最少,但重建出的高分辨率图像质量也较低。VDSR(Kim 等,2016a)也只含有较少参数量,但重建图像质量与本文方法相比也较低。CFSRCNN(Tian 等,2021)与本文方法含有的

图 6 放大因子为 $\times 4$ 时不同方法的一组超分辨率结果图Fig. 6 A set of super-resolution results obtained by different methods when the scale factor is $\times 4$

((a) GT; (b) bicubic; (c) LRTV; (d) FSRCNN; (e) EDSR; (f) VDSR; (g) CFSRCNN; (h) FAWDN; (i) RDN; (j) ours)

参数量相当,但超分辨率性能低于本文方法。RDN (Zhang 等,2018)的超分辨率重建质量与本文方法接近,但参数量约为本文方法的 17 倍。

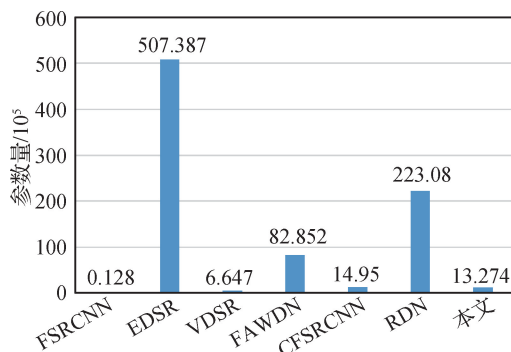


图 7 不同超分辨率网络的可训练参数量

Fig. 7 The numbers of trainable parameters of different super-resolution networks

表 7 给出了不同深度学习超分辨率方法重建一幅图像的平均运行时间(通过测试 200 组图像取平均得到)。可以看出,网络运行时间与参数量基本成正相关关系。需要指出的是,虽然 FAWDN 参数量不大但运行时间最长,这是由于该网络采用反馈设计思路获得,需要多次重建中间图像。RDN 在参数量上少于 EDSR,但该网络大量使用密集连接的方式提取特征,使得网络的宽度较大,导致运行时间较长。本文方法能够获得较高的计算效率,运行时间仅为 RDN 方法的 1/16 左右。综上,本文方法可以在参数量较低和运行时间较少的情况下获得较为理想的超分辨率性能。

4 结 论

针对当前医学图像超分辨率方法主要面向单一

表 7 不同超分辨率网络重建单幅图像的平均运行时间

Table 7 The average running time of different super-resolution networks for reconstructing a single image

方法	运行时间/ms
EDSR	5.9
FSRCNN	1.5
VDSR	2.2
FAWDN	87.3
CFSRCNN	4.8
RDN	38.5
本文	2.3

模态图像设计的问题,本文提出了一种基于轻量级残差密集卷积网络的多模态 MR 图像超分辨率方法。该方法能够在统一的网络模型下同时实现两种模态图像的超分辨率重建,充分利用不同模态图像之间的关联信息,帮助提升超分辨率重建质量。在网络结构中,主要利用残差密集块和高效通道注意力模块组成的残差密集注意力模块实现多模态特征的提取与精炼,再通过亚像素卷积层进行上采样,最终分别重建不同模态的高分辨率图像。

实验结果表明,本文提出的超分辨率方法在主观视觉效果方面能够重建出具有较多细节信息、较少伪影的高分辨率图像;在客观评价方面,本文方法在 PSNR 和 SSIM 两种常用客观评价指标上能够取得优于 8 种代表性对比方法的性能。此外,本文网络具有较小的参数规模与较高的计算效率,与性能接近的残差密集网络(RDN)方法相比,网络参数量低于 RDN 的 10%。

需要指出的是,受计算与存储代价等因素的限制,本文方法目前是针对 2D 图像切片进行设计。

考虑到现实中 MR 图像具有 3D 体数据结构,采用逐切片的方式进行 3D 图像超分辨率势必会丢失切片之间的关联信息,影响重建质量。因此,未来拟研究面向 3D 体数据的多模态医学图像超分辨率方法,进一步提升方法性能。

参考文献 (References)

- Chen L H, Yang X M, Jeon G, Anisetti M and Liu K. 2020. A trusted medical image super-resolution method based on feedback adaptive weighted dense network. *Artificial Intelligence in Medicine*, 106: #101857 [DOI: 10.1016/j.artmed.2020.101857]
- Cheng D Q, Guo X, Chen L L, Kou Q Q, Zhao K and Gao R. 2021. Image super-resolution reconstruction from multi-channel recursive residual network. *Journal of Image and Graphics*, 26(3): 605-618 (程德强, 郭昕, 陈亮亮, 寇旗旗, 赵凯, 高蕊. 2021. 多通道递归残差网络的图像超分辨率重建. *中国图象图形学报*, 26(3): 605-618) [DOI: 10.11834/jig.200108]
- Dong C, Loy C C, He K M and Tang X O. 2014. Learning a deep convolutional network for image super-resolution//*Proceedings of the 13th European Conference on Computer Vision*. Zurich, Switzerland; Springer: 184-199 [DOI: 10.1007/978-3-319-10593-2_13]
- Dong C, Loy C C and Tang X O. 2016. Accelerating the super-resolution convolutional neural network//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, the Netherlands; Springer: 391-407 [DOI: 10.1007/978-3-319-46475-6_25]
- Kim J, Lee J K and Lee K M. 2016a. Accurate image super-resolution using very deep convolutional networks//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA; IEEE: 1646-1654 [DOI: 10.1109/CVPR.2016.182]
- Kim J, Lee J K and Lee K M. 2016b. Deeply-recursive convolutional network for image super-resolution//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA; IEEE: 1637-1645 [DOI: 10.1109/CVPR.2016.181]
- Lei P C, Liu C, Tang J G and Peng D L. 2020. Hierarchical feature fusion attention network for image super-resolution reconstruction. *Journal of Image and Graphics*, 25(9): 1773-1786 (雷鹏程, 刘丛, 唐坚刚, 彭敦陆. 2020. 分层特征融合注意力网络图像超分辨率重建. *中国图象图形学报*, 25(9): 1773-1786) [DOI: 10.11834/jig.190607]
- Li Z, Yang J L, Liu Z, Yang X M, Jeon G and Wu W. 2019. Feedback network for image super-resolution//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA; IEEE: 3862-3871 [DOI: 10.1109/CVPR.2019.00399]
- Lim B, Son S, Kim H, Nah S and Lee K M. 2017. Enhanced deep residual networks for single image super-resolution//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition workshops (CVPRW)*. Honolulu, USA; IEEE: 1132-1140 [DOI: 10.1109/CVPRW.2017.151]
- Shen M Y, Yu P F, Wang R G, Yang J and Xue L X. 2019. Image super-resolution reconstruction via deep network based on multi-staged fusion. *Journal of Image and Graphics*, 24(8): 1258-1269 (沈明玉, 俞鹏飞, 汪荣贵, 杨娟, 薛丽霞. 2019. 多阶段融合网络的图像超分辨率重建. *中国图象图形学报*, 24(8): 1258-269) [DOI: 10.11834/jig.180619]
- Shi F, Cheng J, Wang L, Yap P T and Shen D G. 2015. LRTV: MR image super-resolution with low-rank and total variation regularizations. *IEEE Transactions on Medical Imaging*, 34(12): 2459-2466 [DOI: 10.1109/TMI.2015.2437894]
- Shi J, Li Z, Ying S H, Wang C F, Liu Q P, Zhang Q and Yan P K. 2019. MR image super-resolution via wide residual networks with fixed skip connection. *IEEE Journal of Biomedical and Health Informatics*, 23(3): 1129-1140 [DOI: 10.1109/JBHI.2018.2843819]
- Shi W Z, Caballero J, Huszár F, Totz J, Aitken A P, Bishop R, Rueckert D and Wang Z H. 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA; IEEE: 1874-1883 [DOI: 10.1109/CVPR.2016.207]
- Tian C W, Xu Y, Zuo W M, Zhang B, Fei L K and Lin C W. 2021. Coarse-to-fine CNN for image super-resolution. *IEEE Transactions on Multimedia*, 23: 1489-1502 [DOI: 10.1109/TMM.2020.2999182]
- Tong J C, Fei J L, Chen J S, Li H and Ding D D. 2019. Multi-level feature fusion image super-resolution algorithm with recursive neural network. *Journal of Image and Graphics*, 24(2): 302-312 (佟骏超, 费加罗, 陈靖森, 李恒, 丁丹丹. 2019. 递归式多阶特征融合图像超分辨率算法. *中国图象图形学报*, 24(2): 302-312) [DOI: 10.11834/jig.180410]
- Tong T, Li G, Liu X J and Gao Q Q. 2017. Image super-resolution using dense skip connections//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy; IEEE: 4809-4817 [DOI: 10.1109/ICCV.2017.514]
- Wang Q L, Wu B G, Zhu P F, Li P H, Zuo W M and Hu Q H. 2020. ECA-Net: efficient channel attention for deep convolutional neural networks [EB/OL]. [2022-01-18]. <https://arxiv.org/pdf/1910.03151v3.pdf>
- Xue X T, Wang Y, Li J, Jiao Z C, Ren Z Q and Gao X B. 2020. Progressive sub-band residual-learning network for MR image super-resolution. *IEEE Journal of Biomedical and Health Informatics*, 24(2): 377-386 [DOI: 10.1109/JBHI.2019.2945373]
- Ying Z L and Long X. 2019. Single-image super-resolution construction based on multi-scale dense residual network. *Journal of Image and Graphics*, 24(3): 410-419 (应自炉, 龙祥. 2019. 多尺度密集

- 残差网络的单幅图像超分辨率重建. 中国图象图形学报, 24(3): 410-419 [DOI: 10.11834/jig.180431]
- Zeng K, Zheng H, Cai C B, Yang Y, Zhang K H and Chen Z. 2018. Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network. *Computers in Biology and Medicine*, 99: 133-141 [DOI: 10.1016/j.compbiomed.2018.06.010]
- Zhang Y L, Tian Y P, Kong Y, Zhong B N and Fu Y. 2018. Residual dense network for image super-resolution//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 2472-2481 [DOI: 10.1109/CVPR.2018.00262]
- Zhou B, Li C H and Chen W. 2021. Region-level channel attention for single image super-resolution combining high frequency loss. *Journal of Image and Graphics*, 26(12): 2836-2847 (周波, 李成华, 陈伟. 2021. 区域级通道注意力融合高频损失的图像超分辨率重建. 中国图象图形学报, 26(12): 2836-2847) [DOI: 10.11834/jig.200582]

作者简介

刘羽,男,副教授,主要研究方向为图像处理、医学图像分析、机器学习、信息融合。E-mail: yuliu@hfut.edu.cn

陈勋,通信作者,男,教授,主要研究方向为医学人工智能、人机交互、移动健康监护。E-mail: xunchen@ustc.edu.cn

朱文瑜,男,硕士研究生,主要研究方向为图像超分辨率重建。E-mail: wyzhu@mail.hfut.edu.cn

成娟,女,副教授,主要研究方向为医学信号与图像处理。

E-mail: chengjuan@hfut.edu.cn