

局部特征增强的磁共振图像Transformer重构

熊承义^{ab}, 陈文旗^{ab*}, 高志荣^c, 马帅^{ab}, 李帆^{ab}

(中南民族大学 a. 电子信息工程学院; b. 智能无线通信湖北省重点实验室; c. 计算机科学学院, 武汉 430074)

摘要 研究了一种基于多头自注意力与卷积特征融合的磁共振图像Transformer重构方法. 采用U型网络结构, 通过学习图像的多尺度特征以提升重构性能. 采用深度分离卷积与多头自注意力融合的Swin Transformer结构, 改善网络的特征学习能力. 在CC359-Brain数据集下基于多种采样模式进行仿真实验, 结果证明了该方法在提升磁共振图像重构质量与降低系统复杂度方面的有效性.

关键词 磁共振成像; 卷积神经网络; 变换器; 深度学习

中图分类号 TP391.4 **文献标志码** A **文章编号** 1672-4321(2025)01-0050-08

doi: 10.20056/j.cnki.ZNMDZK.20250107

MRI reconstruction based on Transformer with local feature enhancement

XIONG Chengyi^{ab}, CHEN Wenqi^{ab*}, GAO Zhirong^c, MA Shuai^{ab}, LI Fan^{ab}

(1 South-Central Minzu University, a. College of Electronic and Information Engineering; b. Hubei Key Lab of Intelligent Wireless Communication; c. College of Computer Science, Wuhan 430074, China)

Abstract A Transformer reconstruction method of Magnetic Resonance Image (MRI) based on multi-head self-attention and convolution feature fusion is studied. The U-shaped network structure is adopted to improve the reconstruction performance by learning the multi-scale features of the image. The Swin Transformer structure of deep separation convolution and multi-head self-attention fusion is adopted to improve the feature learning ability of the network. Simulation experiments based on multiple sampling modes are carried out under CC359 Brain dataset, and the results show that the proposed method is effective in improving the quality of MRI reconstruction and reducing the system complexity.

Keywords Magnetic Resonance Image (MRI); Convolution Neural Network; Transformer; deep learning

磁共振成像(Magnetic Resonance Imaging, MRI)作为一种非侵入式的诊断工具,具有安全无辐射、软组织对比度高以及成像清晰度高等特点,可应用于诊断中枢神经系统、头颈部、呼吸系统、乳腺等部位的疾病.然而超长的扫描时间不仅阻碍诊断效率,也增加患者的不舒适感.为了减少扫描时间,提出了并行成像(Parallel Imaging, PI)和压缩感知^[1](Compressed Sensing, CS).应用于临床的并行成像技术,通过联合多个线圈进行并行采样,之后利用特定的重构算法(SENSE^[2]、GRAPPA^[3]、SPIRiT^[4]等)

对采样图像进行重构.广泛应用于图像处理^[5]的压缩感知^[1]技术,利用MR图像的稀疏性,以低于奈奎斯特采样率的方式,对K空间数据进行欠采样,利用Fourier Transform(FT)、Total Variation (TV)、小波变换等算法进行重建^[6-8].但是当进一步加速采样时,两者方法的性能都会退化,不仅图像的重构时间增加,而且还会产生大量的图像伪影.

近年来,深度学习的方法在图像处理领域取得巨大的成功,尤其是卷积神经网络(Convolution Neural Network, CNN).在MR重构中,深度学习网络

收稿日期 2022-10-10 * **通信作者** 陈文旗,研究方向:医学图像重构,E-mail: 2085142515@qq.com

作者简介 熊承义(1969-),男,教授,博士,研究方向:图像处理与模式识别,E-mail: xiongcy@mail.scuec.edu.cn

基金项目 多谱信息处理技术国家重点实验室基金资助项目(6142113210303);中央高校基本科研业务专项资金资助项目(CZY21013);中南民族大学研究生课程示范课程资助项目(YJS22039)

大概可以分成两类,一类是模型驱动型网络^[9-11],通过与传统CS算法进行结合,使得原本需要手工设计的正则化项,直接由网络从数据中学习,降低了网络训练的复杂度,但是图像的重构时间太长.另一类是数据驱动型网络^[12-14],这些网络以端到端的方式学习配对数据之间的映射关系,无需多次迭代,直接生成重构图像,并且在性能和运行速度上明显优于基于模型和传统的重构算法.在CNN中,DAGAN^[12]联合U-Net^[15]和对抗损失进行高质量MR图像重构.但是由于卷积结构具有天然狭窄的感受野结构,尽管CNN对局部特征的学习存在优势,但是无法直接捕获全局特征.除了CNN之外,基于Transformer的网络也开始应用于MR图像重构,例如SwinMR^[14].Transformer通过多头自注意力结构对远距离依赖信息进行建模,可以感受全局信息,抓取全局相关性,在图像处理任务上优于CNN.Transformer架构虽然擅长捕获全局特征,但是缺乏对局部信息的感知.

由以上背景,本文提出了一种局部特征增强的磁共振图像Transformer重构方法.首先,为了增强Transformer对局部特征的学习能力,本文通过融合多头自注意力和卷积结构提出了Swin-Conv模块(Swin-Conv Block, SCB),SCB通过深度可分离卷积(Deep-Wise Convolution, DW)增强Swin Transformer^[16]中W-MSA/SW-MSA结构的局部信息.其次,基于SCB,本文构造了Swin-Conv Transformer Layer (SCTL)作为网络的基本特征学习模块,SCTL使用通道注意力进一步增强网络对重点特征的关注.最后,为了降低网络复杂性,本文将网络设计为U型结构,获取多个尺度下的特征信息进行学习,实现MR图像高质量重建.在性能指标和网络复杂度上,本文提出的网络明显优于其他重构网络.

1 相关工作

1.1 基于CNN的磁共振图像重构

在图像处理中,基于深度学习的方法已经得到广泛应用.其中具有平移不变性和局部特性的CNN流行在各种视觉任务中,包含图像分割^[17]、超分辨率^[18]和图像分类^[19]等.WANG等^[13]最开始把CNN应用到MR图像重构中,直接使用多个卷积层堆叠而成的CNN,学习欠采样K空间数据所逆傅里叶变换得

到的初始重构图像与全采样图像之间的映射关系.之后,应用于图像分割领域的U-Net^[15]也引入到MR图像重构中,通过逐层下采样的编码结构和对应的解码结构进行MR图像的恢复.同时,生成对抗网络(Generative Adversarial Network, GAN)也被引入到MR图像重构中,通过引入生成器与鉴别器之间的竞争机制,来改善图像质量.DAGAN^[12]设计一种改进学习策略用来稳定基于UNet^[15]的生成器,同时在网络的损失函数上,使用内容和感知损失保持MR图像的细节信息,最后联合对抗损失生成高质量的MR图像.QUAN等^[20]进一步采用更深层次的CNN结构和循环损失,来改善MR图像质量.基于卷积结构的CNN尽管擅长捕获局部特征信息,但是缺乏感知全局的能力.虽然可以通过堆叠更深层次的CNN进行缓解,但也会带来高昂的训练成本以及网络复杂度.

1.2 基于Transformer的磁共振图像重构

在网络的架构上,相比较卷积结构的局部感受野特性,通过自注意力机制捕获全局依赖信息的Transformer也在视觉领域迅速流行.FENG等^[21]使用任务型Transformer架构,共享MR重构任务和MR超分任务之间的关键特征表示.在基于多模态的任务中,FENG等^[22]使用融合交叉注意力机制的Transformer架构迁移备用模态的特征到目标模态上,进行多模MR重构.但是由于自注意力机制通过计算每个像素点与其他像素点的相关性,计算复杂度与图像大小的二次方成正比,具有很高的计算代价.Swin Transformer^[16]通过基于平移窗口的多头注意力(Shifted Windows based Multi-Head Self-Attention, W-MSA/SW-MSA)结构,将注意力计算限制在了一个个非重叠窗口内,将计算复杂度从输入尺寸的二次方降低为线性,有效地降低了计算代价.YAN等^[23]基于Swin Transformer提出了用于MR图像超分辨率的Transformer网络.SwinMR^[14]通过将SwinIR^[24]引入到MR图像重构任务,在重构图像的质量和泛化性上都优于基于卷积的网络,但是SwinMR^[14]的计算复杂度远高于CNN.SDAUT^[25]将Swin Transformer^[16]与U-Net^[15]结合,同时结合可变形的注意力提升重构质量,在一定程度上降低网络的复杂性和计算代价.但是这些网络主要利用了自注意力结构进行特征学习,缺乏对局部特征的感知,阻碍了MR图像精细结构的恢复.

2 方法

本文提出了一种局部特征增强的磁共振图像 Transformer 重构网络. 整个网络为 U 型对称结构, 使用多次下采样和上采样来获取不同尺度下的特征信息. 使用 SCB 增强网络局部特征的学习能力, 同

时构造 SCTL 来学习不同尺度下的特征, 改善 MR 图像重构质量.

2.1 网络结构

本文提出的网络如图 1 所示. 通过输入欠采样的伪影图像 X_u , 网络输出为去除伪影的图像 X , 网络使用残差连接来加快训练以及稳定训练过程.

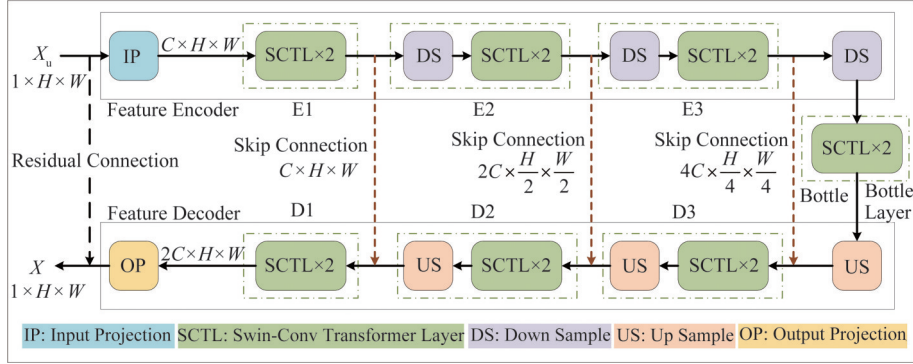


图 1 提出网络的结构图

Fig. 1 Structure diagram of the proposed network

整个处理流程可以表示为:

$$X = f(X_u | \theta) + X_u, \quad (1)$$

其中: f 表示提出的网络, θ 为需要学习的网络参数.

整个网络分为三部分, 分别是特征编码层 (Feature Encoder)、瓶颈层 (Bottle Layer) 和特征解码层 (Feature Decoder). 在特征编码层阶段, 网络会提取多尺度特征进行学习. 具体来说, 伪影图像 X_u 会使用 3×3 大小的卷积层进行非重叠映射 (Input Projection, IP), 再送入两层的 Swin-Conv Transformer Layer (SCTL) 进行特征学习, 得到第一层特征 F_{encoder_1} , 可表示为:

$$F_{\text{encoder}_1} = E1(\text{IP}(X_u)). \quad (2)$$

之后再使用卷积核为 4×4 大小、步长为 2 的重叠下采样层 (Down Sample, DS) 进行下采样, 每个下采样后面跟随两层 SCTL 进行特征学习. 值得注意的是, 在每次下采样后, 特征图的大小减少一半, 但通道数量会增加一倍. 因此在经历两次下采样后, 会得到不同尺度下的特征信息, 可表示为:

$$F_{\text{encoder}_2} = E2(F_{\text{encoder}_1}), \quad (3)$$

$$F_{\text{encoder}_3} = E3(F_{\text{encoder}_2}). \quad (4)$$

在网络的瓶颈层阶段, 在进行特征学习之前, 还会进行一次下采样, 此时的特征图仅为刚开始的 $1/8$, 包含了图像的全局特征信息和丰富的结构信息. 本文同样使用两层的 SCTL 对特征进行学习, 得到瓶颈层的特征信息 F_{Bottle} , 可表示为:

$$F_{\text{Bottle}} = \text{Bottle}(\text{DS}(F_{\text{encoder}_3})). \quad (5)$$

接下来为网络的特征解码层阶段, 网络通过跳跃连接 (Skip Connection) 联合编码层的特征信息进行重构. 网络的特征解码层模块由 SCTL 和上采样层 (Up Sample, US) 构成, 上采样层使用 2×2 大小的转置卷积, 使用的步长为 2, 在增加特征图尺寸的同时, 保持低尺度的特征信息不会损失. 但是与下采样层相反, 每次上采样过后, 特征图的大小增加一倍, 通道数减少 4 倍, 除了第一次上采样的通道数只会减少一半. 由此, 在网络特征解码端的具体流程可表示为:

$$F_{\text{decoder}_3} = D3(\text{Concat}(F_{\text{encoder}_3}, \text{US}(F_{\text{Bottle}}))), \quad (6)$$

$$F_{\text{decoder}_2} = D2(\text{Concat}(F_{\text{encoder}_2}, F_{\text{decoder}_3})), \quad (7)$$

$$F_{\text{decoder}_1} = D1(\text{Concat}(F_{\text{encoder}_1}, F_{\text{decoder}_2})). \quad (8)$$

在得到解码端的输出 F_{decoder_1} 后, 网络使用一个输出映射层 (Output Projection, OP), 具体为一个 3×3 的卷积层结构, 重构出图像 F_{op} , 可表示为:

$$F_{\text{op}} = \text{OP}(F_{\text{decoder}_1}). \quad (9)$$

通过残差连接加上最开始的伪影图像 X_u 得到最终的重构图像 X :

$$X = F_{\text{op}} + X_u. \quad (10)$$

2.2 Swin-Conv Transformer Layer

之前提到, 具有局部归纳特性的卷积结构以及擅长捕获全局依赖特性的自注意力结构已经应用于各种视觉任务, 但是很少有工作联合两者的特

性,尤其是在磁共振图像重构任务中.为此,本文通过增强原有 Swin Transformer^[16] W-MSA 结构中的局部信息为切入点,使用两层且卷积核大小为 3×3 的深度分离卷积提取特征图像中的局部信息,再与原有 Swin Transformer^[16] 中 W-MSA 结构获取到的全局信息进行融合,以此丰富 MR 图像的局部特征信息.但在信息融合方式上,针对并行和串行两种方式,本文采取并行方式构成 SCTL(3.5 节可见消融实验).

具体而言,网络的每层结构都由连续的两个 SCTL 构成,分别对应 Swin Transformer^[16] 中的 W-MSA 结构和 SW-MSA 结构.由图 2 可见:SCTL 的主要模块主要由三部分组成,包括 Swin-Conv 模块(Swin-Conv Block, SCB)、局部增强的前馈神经网络(Locally-Enhanced Feed-Forward Network, LeFF)和通道注意力(Channel Attention, CA).其余的层归一化(Layer Normalization, LN)和残差连接(Residual Connection)被用来稳定训练及加速网络收敛,每个 DW 卷积或者 PW 卷积后面会跟随一个 GELU 激活函数,输出层的卷积除外.

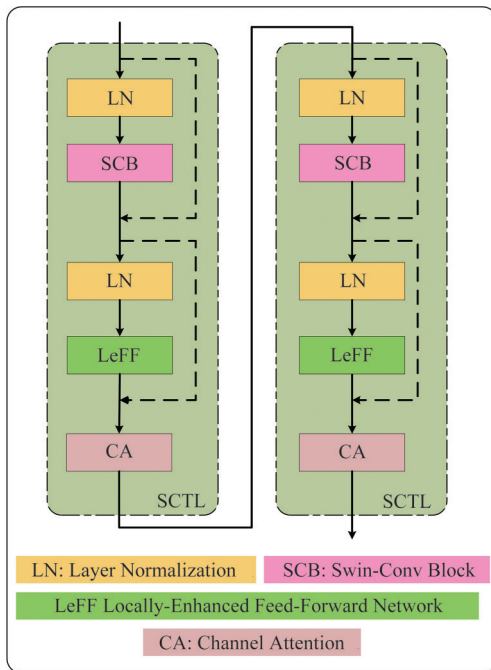


图 2 双层 SCTL 的结构图

Fig. 2 Structure diagram of double layer SCTL

SCB 的详细结构如图 3(a) 所示,在原有 W-MSA 结构的基础上,本文引入了基于卷积结构的分支,利用深度分离卷积结构,在没有增加太多的参数和网络复杂度的前提下,进行逐个通道的局部特征提取,之后再使用 1×1 卷积(Point-Wise Convolution,

PW)对两个分支上的通道特征进行融合,以此增强 MR 图像的局部特征信息.具体来说,给定输入的特征张量 $Z^l \in \mathbf{R}^{C \times H \times W}$,送入到 SCB 的两个分支中,对于 W-MSA 分支,首先会对输入进行窗口偏移和划分(Window Partition),得到划分后的窗口特征 $Y^l \in \mathbf{R}^{N \times T^2 \times C}$, N 和 T 分别为划分窗口的数量和大小,再送入到 W-MSA 分支中,如图 3(a) 左侧所示,对窗口特征 Y^l 进行线性映射分别得到 $Q \in \mathbf{R}^{N \times H \times T^2 \times \frac{C}{H}}$ 、 $K \in \mathbf{R}^{N \times H \times T^2 \times \frac{C}{H}}$ 和 $V \in \mathbf{R}^{N \times H \times T^2 \times \frac{C}{H}}$ 三个映射张量, H 为头的数量.以单头为例,计算过程与 Swin Transformer^[16] 中一致,加上相对位置编码信息 $B \in \mathbf{R}^{N \times T^2 \times T^2}$,得到相应注意力输出 $\text{Attention}(Q, K, V) \in \mathbf{R}^{N \times H \times T^2 \times \frac{C}{H}}$:

$$\text{Attention}(Q, K, V) = \left(\text{SoftMax} \left(\frac{QK'}{\sqrt{C}} \right) + B \right) V, \quad (11)$$

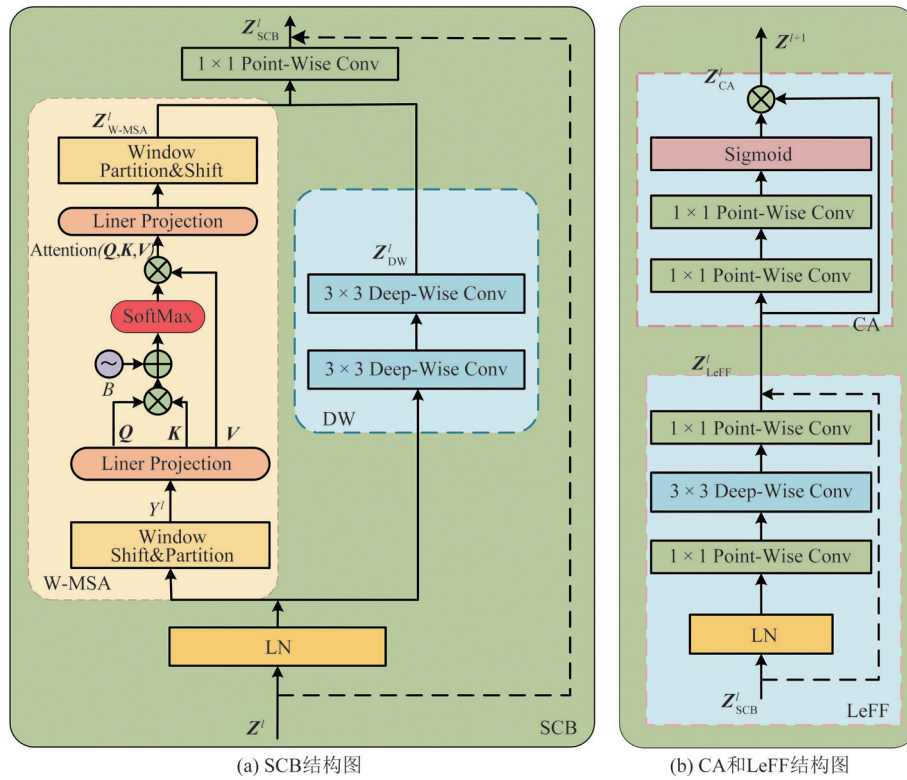
其中 K' 为 K 的转置.通过联合所有多头自注意力的输出,再经过一个线性映射层以及窗口划分和偏移,恢复出原本的特征图结构,即为 W-MSA 分支的输出 $Z_{W-MSA}^l \in \mathbf{R}^{C \times H \times W}$,由于包含连续两层的 SCTL 层,在下一层得到 SW-MSA 分支的输出 $Z_{SW-MSA}^l \in \mathbf{R}^{C \times H \times W}$,而在 DW 分支中,如图 3(a) 右边所示,使用两层卷积核大小为 3×3 的 DW 卷积后,会得到 DW 分支的输出 $Z_{DW}^l \in \mathbf{R}^{C \times H \times W}$,最后经过 PW 卷积进行每个通道上的信息融合 $Z_{PW}^l \in \mathbf{R}^{C \times H \times W}$,加上来自输入的残差连接,即为 SCB 模块的输出 $Z_{SCB}^l \in \mathbf{R}^{C \times H \times W}$.

在图 3(b) 中,对于 Swin Transformer^[16] 中前馈神经网络,本文利用来自 Uformer^[26] 中的 LeFF,因为 LeFF 在前馈神经网络中引入了 DW 卷积,能更有效地进行跨通道信息交流,有利于不同通道之间特征信息的融合,可见消融实验的对比结果.但在代码实现方式上,与 Uformer^[26] 中不同的是,本文没有使用线性层进行扩张或者压缩通道,而使用 PW 卷积实现,由于线性层的输入数据形式与卷积层的输入数据形式不一致,以上实现方式可以避免对数据进行操作.具体流程为:

$$Z_{LeFF}^l = \text{PW}_2 \left(\text{DW} \left(\text{PW}_1 \left(\text{LN} \left(Z^l \right) \right) \right) \right) + Z^l, \quad (12)$$

其中 $Z_{LeFF}^l \in \mathbf{R}^{C \times H \times W}$ 为 LeFF 的输出.如图 3(b) 中所示,在 SCTL 的尾部,本文使用通道注意力,对重点特征信息进行判别,使得网络更多关注重点信息,提升网络学习能力.流程为:

$$Z_{CA}^l = \text{Sigmoid} \left(\text{PW}_2 \left(\text{PW}_1 \left(Z_{LeFF}^l \right) \right) \right) \times Z_{LeFF}^l \quad (13).$$



(a) SCB结构图

(b) CA和LeFF结构图

图3 SCTL主要模块的详细结构

Fig. 3 Detailed structure of main modules of SCTL

最后得到 SCTL 的输出 $Z^{l+1} \in \mathbb{R}^{C \times H \times W}$, 整个 SCTL 的流程为:

$$Z_{SCB}^l = \text{PW} \left(\text{Concat} \left(W - \text{MSA} \left(\text{LN} \left(Z^l \right) \right), \right. \right. \\ \left. \left. \text{DW} \left(\text{LN} \left(Z^l \right) \right) \right) \right) + Z^l, \quad (14)$$

$$Z_{LeFF}^l = \text{LeFF} \left(\text{LN} \left(Z_{SCB}^l \right) \right) + Z_{SCB}^l, \quad (15)$$

$$Z^{l+1} = \text{CA} \left(Z_{LeFF}^l \right). \quad (16)$$

3 实验

3.1 实验数据集

本文采用 Calgary-Campinas-359 Brain (CC359-Brain) 单线圈重构数据集^[27], 其中, CC359-Brain 数据是临床上使用商业 MR 扫描仪 (Discovery MR750, GE Healthcare, Waukesha) 进行采集得到, 为全采样 T1 加权的大脑扫描数据, 采样矩阵大小为 256×256 . 遵循原有的数据集划分格式, 25 个训练集样本包含 4527 张大脑矢状面切片以及由 1700 张切片组成 10 个测试集样本. 在训练与测试实验中, 为了避免的引入噪声数据, 对于每个样本只选用 100 个中心切片.

3.2 损失函数

在常用图像的损失函数中, 相比较均方误差损失 (Mean Square Error, MSE) 会对图像产生光滑的重

构效果, 使用 Charbonnier^[28] 损失作为本文的损失函数. 对于网络恢复的 MR 图像 X 和全采样的 MR 图像 \hat{X} , 损失函数的公式定义为:

$$\text{Loss}(\theta) = \frac{1}{S} \sum_{i=1}^S \sqrt{\left(\|X^i - \hat{X}^i\|_2^2 + \varepsilon^2 \right)}, \quad (17)$$

其中本文设置 ε 的值为 10^{-9} .

3.3 实验细节与设置

网络由 Pytorch 实现, 为了监控网络的训练过程, 本文在 pytorch-lightning 框架上使用默认的 Adam 优化器完成网络参数更新. 在网络的结构设置中, 网络的深度为 3, 每层 SCTL 的个数为 2, 窗口划分大小为 8 以及通道数默认为 32. 每个 SCB 中头的数量为默认通道数基于当前通道数的倍数. 使用水平翻转来随机增加训练样本, 并将图像随机旋转 90° 、 180° 或 270° , 训练和测试的图片大小分别为 128 和 256. 学习率的设置采用余弦衰减策略将学习率降低到 10^{-6} , 初始学习率为 0.0002. 每次使用 8 张切片以及 500 周期完成网络的训练. 硬件条件为 Ubuntu 18.04 系统和一张 12 GB 的 Nvidia GTX-1080Ti 图形计算卡.

本文使用峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似性 (Structural Similarity Index, SSIM) 作为评价指标. 基于图像像素点之差的

PSNR 通过利用最大信号功率与信号噪声之比来衡量图像的精细程度.而 SSIM 从图像的亮度、对比度和结构方面来测量图像的相似度.在网络的复杂度和运行速度方面,本文在给定输入数据尺寸为 $1 \times 1 \times 256 \times 256$ 以及不包含鉴别器参数的条件下,统计重构网络的参数量和计算量 (Floating Point Operations, Flops),并且在所有测试切片上,计算所有网络的平均测试时间.

3.4 实验结果

在与 MR 基准重建方法比较中,本文选取了 U-Net^[15]、DAGAN^[12]、SwinMR^[14]、SDAUT^[25]、真实图像 (Ground Truth, GT) 和欠采样的零填充图像 (Zero-

Filled, ZF) 在采样率为 30% 的一维高斯 (Gaussian 1D 30%, G1D30%) 采样轨迹下进行比较.结果见表 1 和图 4.

表 1 在一维高斯采样轨迹 30% 上的比较结果
Tab. 1 Comparison results on G1D30% mask

| 方法 | 参数量/M | 计算量/GMac | 测试时间/s | PSNR | SSIM |
|--------|-------|----------|--------|--------------|---------------|
| ZF | — | — | — | 27.61 | 0.8857 |
| U-Net | 7.8 | 13.66 | 0.019 | 31.13 | 0.9387 |
| DAGAN | 98.59 | 33.97 | 0.021 | 31.64 | 0.9454 |
| SwinMR | 11.4 | 800.73 | 0.958 | <u>33.72</u> | <u>0.9636</u> |
| SDAUT | 66.4 | 293.02 | 0.672 | 33.24 | 0.9603 |
| 本文 | 6.1 | 39.63 | 0.108 | 34.06 | 0.9661 |

注:最好的结果由黑体加粗标出,次好的由下划线标出.

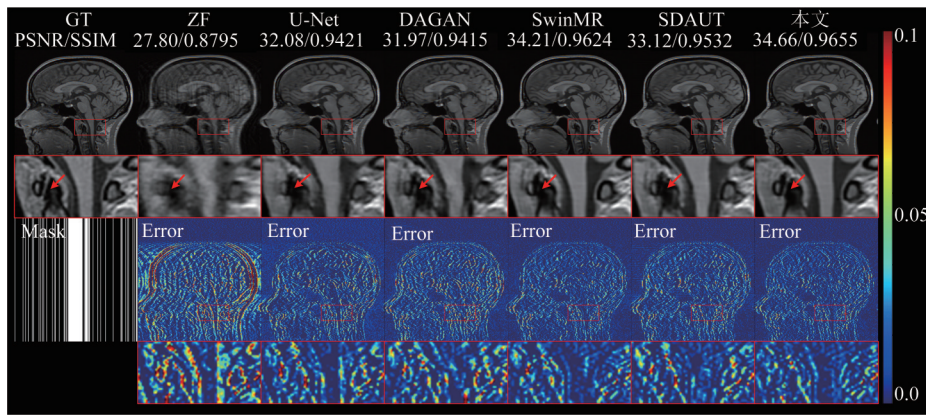


图 4 在一维高斯采样轨迹 30% 上的重构结果
Fig. 4 Reconstructed results on G1D30% mask

在表 1 中可见,在评价指标上,本文提出的网络明显优于对比网络.与 DAGAN^[12]、SDAUT^[25] 和 SwinMR^[14] 相比,在 PSNR 值上,本文网络分别高 2.42、1.32、0.34 dB.在参数量上,本文网络远低于对比网络.值得注意的是,SwinMR^[14] 的计算量以及测试时间分别是本文的 20 倍和 9 倍左右,与本文同样 U 型结构的 SDAUT^[25],计算代价也远高于本文网络.

图 4 展示了不同网络重构出 MR 图像的效果图.由红色箭头所指可见:相比较 SwinMR^[14]、DAGAN^[12] 以及 SDAUT^[25],本文提出的网络能恢复出更多细节信息,这也证明了卷积结构可以增强网

络对局部特征的感知能力,以此改进网络对精细结构的恢复.图 4 还展示了不同网络重构结果与真实图像之间放大 10 倍的误差图,可见本文网络所持有的最大误差和连续误差明显少于对比网络.

本文也研究了低采样轨迹下的性能表现,包含一维高斯 10% (Gaussian 1D 10%, G1D10%)、二维高斯 10% (Gaussian 2D 10%, G2D10%)、径向 10% (Radial 10%, R10%) 以及螺旋 10% (Spiral 10%, S10%).对比的网络包含 U-Net^[15]、DAGAN^[12]、SwinMR^[14] 和 SDAUT^[25],结果见表 2.由表 2 可见:在不同的低采样轨迹下,本文网络在 PSNR 和 SSIM 上都明显优于对比网络.

表 2 在不同低采样轨迹上的比较结果
Tab. 2 Comparison results of using different low sampling masks

| Mask | ZF | | U-Net | | DAGAN | | SwinMR | | SDAUT | | 本文 | |
|--------|-------|--------|-------|--------|-------|--------|--------------|---------------|-------|--------|--------------|---------------|
| | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| G1D10% | 22.29 | 0.7458 | 26.49 | 0.8669 | 25.64 | 0.8434 | <u>27.90</u> | <u>0.8931</u> | 26.84 | 0.8732 | 28.08 | 0.8965 |
| G2D10% | 24.93 | 0.7855 | 28.94 | 0.9064 | 28.58 | 0.8972 | <u>29.97</u> | <u>0.9215</u> | 29.70 | 0.9180 | 30.20 | 0.9256 |
| R10% | 23.17 | 0.7030 | 27.03 | 0.8687 | 26.27 | 0.8464 | <u>28.00</u> | <u>0.8874</u> | 27.29 | 0.8732 | 28.14 | 0.8914 |
| S10% | 22.81 | 0.7264 | 28.43 | 0.8979 | 28.04 | 0.8843 | <u>29.62</u> | <u>0.9157</u> | 29.26 | 0.9101 | 29.83 | 0.9193 |

注:最好的结果由黑体加粗标出,次好的由下划线标出.

3.5 消融实验及分析

在 SCB 中, 针对于 W-MSA 和 DW 模块的融合方式, 本文提出了三种结构, 如图 5 所示, 分别是两种串联结构 Case 1 和 Case 2, 一种并行结构 Case 3. 由表 3 数据可知: 在增加少量参数与计算量的情况下, 并行结构 Case 3 有着更好的特征增强能力.

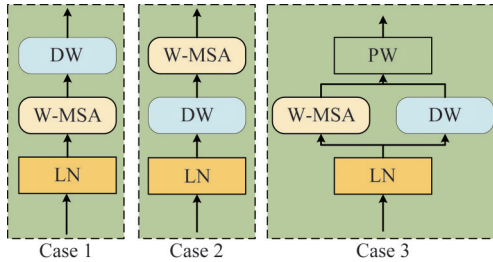


图 5 W-MSA 和 DW 的三种融合结构

Fig. 5 Three fusion structures of W-MSA and DW

表 3 SCB 上 W-MSA 和 DW 的融合实验

Tab. 3 Fusion study of W-MSA and DW on SCB

| 融合方式 | 参数量/M | 计算量/GMac | PSNR | SSIM |
|--------|-------|----------|--------------|---------------|
| Case 1 | 5.4 | 35.36 | <u>33.99</u> | <u>0.9657</u> |
| Case 2 | 5.4 | 35.36 | 32.34 | 0.9528 |
| Case 3 | 6.1 | 39.63 | 34.06 | 0.9661 |

注: 最好的结果由黑体加粗标出, 次好的由下划线标出.

为了验证不同组件对网络性能的影响, 在 G1D30% 的采样轨迹下, 本文对网络架构进行消融实验, 如表 4 所示, 当本文引入 DW 卷积来改善网络对局部信息的学习能力时, 在增加少量参数和计算量的情况下, 网络的 PSNR 与 SSIM 分别增加 0.34 dB 和 0.0024. 在本文替换 MLP 为 UFormer^[26] 提出的 LeFF 后, 网络的性能得到进一步改善, 表明 LeFF 相比 MLP, 能有效改进网络的学习能力. 同时证明了 CA 可以帮助网络关注重点信息, 对主要特征进行学习增强. 以上的消融实验验证了所有网络模块的有效性.

表 4 不同模块的消融实验

Tab. 4 Ablation study of different modules

| W-MSA | MLP | DW | LeFF | CA | 参数量/M | 计算量/GMac | PSNR | SSIM |
|-------|-----|----|------|----|-------|----------|--------------|---------------|
| √ | √ | | | | 5.1 | 33.05 | 33.51 | 0.9623 |
| √ | √ | √ | | | 5.8 | 37.75 | 33.85 | 0.9647 |
| √ | | √ | √ | | 5.9 | 38.56 | <u>34.00</u> | <u>0.9657</u> |
| √ | | √ | √ | √ | 6.1 | 39.63 | 34.06 | 0.9661 |

注: 最好的结果由黑体加粗标出, 次好的由下划线标出.

4 结语

本文研究了一种基于多头自注意力与卷积特

征融合的磁共振图像 Transformer 重构方法, 提出了一个局部特征增强的磁共振图像 Transformer 重构网络. 网络为端到端的 U 型结构, 可以同时捕获多尺度特征信息和降低网络的参数和计算复杂度. 本文使用深度分离卷积增强多头自注意力结构的局部特征学习能力, 改善网络的重构性能. 在不同采样轨迹的 CC359-Brain 数据集上进行了大量实验, 结果证明了本文网络既能提升磁共振图像重构质量, 也能降低网络的参数和复杂度, 具有一定的应用前景.

参考文献

- [1] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [2] PRUESSMANN K P, WEIGER M, BÖRNERT P, et al. Advances in sensitivity encoding with arbitrary k -space trajectories[J]. Magnetic Resonance in Medicine, 2001, 46(4): 638-651.
- [3] GRISWOLD M A, JAKOB P M, HEIDEMANN R M, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA)[J]. Magnetic Resonance in Medicine, 2002, 47(6): 1202-1210.
- [4] LUSTIG M, PAULY J M. SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k -space[J]. Magnetic Resonance in Medicine, 2010, 64(2): 457-471.
- [5] 熊承义, 李治邦, 高志荣, 等. 基于小波变换的图像压缩感知深度重构网络[J]. 中南民族大学学报(自然科学版), 2020, 39(4): 397-404.
- [6] LUSTIG M, DONOHO D L, SANTOS J M, et al. Compressed sensing MRI [J]. IEEE Signal Processing Magazine, 2008, 25(2): 72-82.
- [7] LIANG D, LIU B, WANG J, et al. Accelerating SENSE using compressed sensing [J]. Magnetic Resonance in Medicine, 2009, 62(6): 1574-1584.
- [8] LUSTIG M, DONOHO D, PAULY J M. Sparse MRI: The application of compressed sensing for rapid MR imaging[J]. Magnetic Resonance in Medicine, 2007, 58(6): 1182-1195.
- [9] YANG Y, SUN J, LI H, et al. ADMM-CSNet: A deep learning approach for image compressive sensing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 42(3): 521-538.
- [10] 张登强, 刘霄汉, 庞彦伟. 基于双域并行编解码网络的磁共振图像重建[J]. 激光与光电子学进展, 2022, 59(12): 172-179.
- [11] 杜年茂, 宋威. 深度迭代卷积神经网络的快速脑部 MRI 重建算法[J]. 计算机科学与探索, 2020, 14(12):

- 2150-2160.
- [12] YANG G, YU S, DONG H, et al. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction [J]. *IEEE transactions on medical imaging*, 2017, 37(6): 1310-1321.
- [13] WANG S, SU Z, YING L, et al. Accelerating magnetic resonance imaging via deep learning [C]//2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). Prague: IEEE, 2016: 514-517.
- [14] HUANG J, FANG Y, WU Y, et al. Swin transformer for fast MRI[J]. *Neurocomputing*, 2022, 493: 281-304.
- [15] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015: 234-241.
- [16] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 9992-10022.
- [17] 李丰卓,符玲玲,郭金鑫,等.应用对抗神经网络的胶质瘤MR图像分割[J].*中南民族大学学报(自然科学版)*, 2022, 41(1): 64-70.
- [18] 王孝天,卢紫微,张燕.基于多尺度融合的图像超分辨率重建[J].*控制工程*, 2022, 29(9): 1573-1579.
- [19] 王文涛,柳鸣,赵志伟,等.改进卷积神经网络的苹果叶分类方法[J].*中南民族大学学报(自然科学版)*, 2022, 41(1): 71-78.
- [20] QUAN T M, NGUYEN-DUC T, JEONG W K. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss [J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1488-1497.
- [21] FENG C M, YAN Y, FU H, et al. Task transformer network for joint MRI reconstruction and super-resolution [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Strasbourg: Springer, 2021: 307-317.
- [22] FENG C M, YAN Y, CHEN G, et al. Accelerated multi-modal MR imaging with transformers [J]. *arXiv Preprint arXiv: 2106.14248*, 2021, 3.
- [23] YAN C, SHI G, WU Z. SMIR: A transformer-based model for MRI super-resolution reconstruction [C]//2021 IEEE International Conference on Medical Imaging Physics and Engineering (ICMIPE). Hefei: IEEE, 2021: 1-6.
- [24] LIANG J, CAO J, SUN G, et al. SwinIR: Image restoration using swin transformer [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 1833-1844.
- [25] HUANG J, XING X, GAO Z, et al. Swin Deformable Attention U-Net Transformer (SDAUT) for explainable fast MRI [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Singapore: Springer, 2022: 538-548.
- [26] WANG Z, CUN X, BAO J, et al. Uformer: A general u-shaped transformer for image restoration [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 17683-17693.
- [27] SOUZA R, LUCENA O, GARRAFA J, et al. An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement [J]. *NeuroImage*, 2018, 170: 482-494.
- [28] LAI W S, HUANG J B, AHUJA N, et al. Fast and accurate image super-resolution with deep Laplacian Pyramid Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 41(11): 2599-2613.

(责编&校对 刘钊)