

基于小波域的复数卷积和复数Transformer的 轻量级MR图像重建方法

张晓华^{1,2}, 练秋生^{1,3*}

(1. 燕山大学信息科学与工程学院, 河北秦皇岛 066000; 2. 河北科技师范学院数学与信息科技学院, 河北秦皇岛 066000;
3. 河北省信息传输与信号处理重点实验室, 河北秦皇岛 066000)

摘要: 卷积神经网络能够从大规模数据中学习图像先验信息, 在图像处理领域具有优异表现, 但局部感受野使其难以捕捉像素间的远程依赖关系. Transformer网络架构具有全局感受野, 在自然语言和高级视觉问题上表现出色, 但其计算复杂度与图像尺寸的平方成正比, 限制了其在高分辨图像处理任务中的应用. 此外, 许多MR (Magnetic Resonance) 图像重建算法仅使用幅值数据或将实部和虚部分离到两个独立的通道作为网络输入, 忽略了复值图像实部和虚部之间的相关性. 本文提出基于复数卷积和复数Transformer的混合模块, 既能利用卷积神经网络提取的高分辨率空间信息恢复MR图像细节, 又能通过自注意力模块获取的全局上下文信息捕获远程特征. 基于混合模块, 结合小波变换进一步提出基于小波域的复数卷积和复数Transformer的轻量级MR图像重建算法. 在Calgary-Campinas和fastMRI两个数据集上的实验结果表明, 所提出的模型与四种具有代表性的MR图像重建算法相比, 具有更高的重建性能和更少的资源消耗. 源代码公开于<https://github.com/zhangxh-qhd/WCCTNet>.

关键词: MR图像重建; 小波变换; 轻量级网络; 复数卷积; 复数Transformer; 感受野

基金项目: 河北省自然科学基金(No.F2022203030)

中图分类号: TN911.73; TP391.41

文献标识码: A

文章编号: 0372-2112(2025)04-1221-11

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20241058

Lightweight MR Image Reconstruction Network Based on Wavelet Domain Complex Convolution and Complex Transformer

ZHANG Xiao-hua^{1,2}, LIAN Qiu-sheng^{1,3*}

(1. School of Information & Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066000, China;

2. School of Mathematics and Information Science & Technology, Hebei Normal University of Science & Technology,
Qinhuangdao, Hebei 066000, China;

3. Hebei Key Laboratory of Information Transmission and Signal Processing, Qinhuangdao, Hebei 066000, China)

Abstract: Convolutional neural networks (CNNs) have demonstrated remarkable capabilities in learning image priors from large-scale datasets, achieving exceptional performance across various image processing tasks. However, the local receptive field inherently limit their ability to capture long-range dependencies between pixels. In contrast, the transformer architecture, renowned for its global receptive field, has exhibited outstanding performance in natural language processing and high-level vision tasks. Nevertheless, its computational complexity, which scales quadratically with image size, poses significant challenges for high-resolution image processing applications. Furthermore, many magnetic resonance (MR) reconstruction algorithms exhibit limitations by either relying exclusively on magnitude data or processing real and imaginary components as separate channels, thereby failing to account for the intrinsic correlations within complex-valued images. By integrating complex convolution and complex transformer, an innovative hybrid module is introduced, which leverages the high-resolution spatial information extracted by CNNs to enhance the details of MR images and capture long-range features through global contextual information obtained by the self-attention module. Building on this hybrid module and wavelet transform, a lightweight MR image reconstruction method using complex convolution and complex transformer in the wave-

let domain is further proposed. Experimental results on the Calgary-Campinas and fastMRI datasets demonstrate that the proposed model achieves superior reconstruction performance and while maintaining lower resource consumption compared to four representative MR image reconstruction algorithms. The source code is available at <https://github.com/zhangxh-qhd/WCCTNet>.

Key words: MR image reconstruction; wavelet transform; lightweight network; complex convolution; complex transformer; receptive field

Foundation Item(s): Natural Science Foundation of Hebei Province (No.F2022203030)

1 引言

核磁共振成像(Magnetic Resonance Imaging, MRI)作为一种重要的医学成像技术,利用磁场和计算机生成人体器官和组织结构的图像^[1],能够提供非侵入性、无电离辐射、高对比度、高分辨率的软组织空间图像,是一种常用的医学诊疗手段。但获取全采样MR(Magnetic Resonance)图像通常需要较长的扫描时间,在扫描过程中,患者须保持静止不动,其自主运动或生理性活动(如心跳等)易引发运动伪影,影响成像质量。为了解决这一挑战,通常采用的方法是在K空间进行欠采样缩短扫描时间,实现MRI的加速重建^[2]。MRI加速重建方法可以分为基于压缩感知的方法^[3]和基于深度学习^[4]的方法。

压缩感知(Compressed Sensing, CS)技术通过选择某种先验约束解空间,并利用优化迭代方法求解MR图像重建问题。目前学者们已经提出了许多基于CS的方法,包括全变差^[5]、低秩分解^[6]、BM3D去噪先验^[7]等。这些迭代优化方法计算时间长,难以满足临床应用的需求。此外,为了获得高质量的重建效果,在处理不同样本时需要手动调整多个超参数,增加了操作复杂度。

近年来,深度学习技术在医学成像领域取得了显著进展,尤其在图像分割^[8,9]、配准^[10]和快速重建^[11]等任务中展现出强大的潜力。Wang等人^[12]将深度学习方法引入MR图像重建领域,利用卷积神经网络(Convolutional Neural Network, CNN)建立了欠采样数据与目标图像之间的非线性映射关系。Lee等人^[13]提出了一种基于残差学习的快速MR图像重建技术。Schlemper等人^[14]针对二维心脏MR图像的动态序列快速重建问题,采用级联CNN架构和生成对抗网络从欠采样K空间数据中恢复高质量的图像,取得了显著效果。Wang等人^[15]提出一种基于复数卷积的残差网络用于并行MR图像重建,提高了MR图像的重建质量。Cole等人^[16]研究了复数卷积在MR图像重建和相位信息处理中的应用,并通过实验验证了复数卷积神经网络在脂肪-水分分离和血流量化等基于相位的应用中的潜力。尽管卷积神经网络在MR图像重建方面较传统方法展现出明显优势,但卷积算子的局部感受野限制了对像素间远程依赖关系的捕捉能力。虽然增加网络深度或卷积核的尺寸可扩大感受野并提升重建性能,但这无疑会增加网络的

计算成本。

与CNN相比,Transformer^[17-20]具有出色的全局建模能力,在自然语言任务和高级视觉问题中展现出显著优势。Transformer模型的核心是自注意力机制^[17],它通过对输入序列中所有位置的加权和计算特定位置的响应,能捕捉图像中任意两个像素之间的依赖关系,更好地理解图像的全局结构和上下文信息。尽管自注意力机制在捕获远距离像素相互作用方面非常有效,但其计算复杂度随着图像尺寸的增加呈二次方增长,限制了其在高分辨率图像处理任务中的应用。为降低运算复杂性,研究者们采取了多种策略。Liang等人^[21]提出了基于局部窗口的注意力机制,将图像划分为互不重叠的图像块,对每个图像块单独计算注意力,实现了性能、计算复杂性和内存消耗之间的平衡,扩展了Transformer在视觉任务中的应用。Zamir等人^[22]提出了多头转置注意力机制(Multi-head Transpose Attention, MDTA),在通道维度上应用自注意力机制,以减轻处理高分辨图像时的计算复杂性。为了提升性能,Restormer在每一级使用多个自注意力模块,导致在处理高分辨率图像时仍面临较高的内存需求,如在fastMRI数据集的单线圈膝关节训练集上训练Restormer(约5 M参数)模型,需要消耗大约48 GB的GPU显存,庞大的内存消耗主要源于Restormer模型中包含了44个Transformer模块。如何在确保重建质量不受影响的前提下,最大限度地减少自注意力模块数量是降低资源消耗的关键挑战。此外,利用Transformer实现图像重建的方法主要针对实值图像^[21,22],对复值数据通常用初始重建图像的幅值^[18,23,24]或将其实部分量和虚部分量拼接成双通道图像^[19]作为网络输入,忽略了复值数据实部和虚部之间的相关性。因此,以较少的资源实现复值MR图像的加速重建是亟待解决的问题。本文针对复值图像提出了一种基于小波域的复数卷积和复数Transformer的MR图像重建算法,主要贡献包括以下三点:

(1)将Transformer架构推广到复数域,并在通道维度上应用自注意力机制,提出了CTransformer模块,旨在实现复值图像的高效重建。本文探究了将Transformer架构应用于复值图像处理问题的潜力。

(2)设计了基于复数卷积和复数Transformer的混合模块(Complex Convolution and complex Transformer

波变换后的特征 $DWT(X) \in \mathbb{C}^{4 \times \frac{H}{2} \times \frac{W}{2}}$. 然后利用一个 3×3 复数卷积层提取浅层特征 $F_0 \in \mathbb{C}^{C \times \frac{H}{2} \times \frac{W}{2}}$. 浅层特征 F_0 通过三级编码器结构转换为隐层特征 $F_1 \in \mathbb{C}^{4C \times \frac{H}{8} \times \frac{W}{8}}$. 隐层特征 F_1 通过两级非对称的解码器结构得到重建特征表示 $F_r \in \mathbb{C}^{C \times \frac{H}{2} \times \frac{W}{2}}$. 重建特征 F_r 经过一个 1×1 的复数卷积层降维得到 4 个空间分辨率为 $\frac{H}{2} \times \frac{W}{2}$ 的复值图像, 对其实部和虚部分别进行逆离散小波变换后得到残差图像 $R \in \mathbb{C}^{1 \times H \times W}$. 将残差图像与 X 相加得到重建图像 $\hat{X} = X + R$. 其过程可描述为

$$F_0 = W_1 DWT(X) \quad (4)$$

$$F_1 = \text{Encoder}(F_0) \quad (5)$$

$$F_r = \text{Decoder}(F_1) \quad (6)$$

$$\hat{X} = X + \text{IDWT}(W_2 F_r) \quad (7)$$

其中, $\text{Encoder}(\cdot)$ 和 $\text{Decoder}(\cdot)$ 分别表示编码器和解码器模块, W_1 和 W_2 分别表示 3×3 和 1×1 的复数卷积, $DWT(\cdot)$ 和 $\text{IDWT}(\cdot)$ 分别表示离散小波变换和逆离散小波变换.

(1) 离散小波变换模块和逆离散小波变换模块

由于 MR 图像通常具有较高的空间分辨率, 直接提取特征对硬件资源的需求较高. 传统的下采样方法或者带步长的卷积运算虽然能降低图像的分辨率, 但会导致图像部分信息丢失. 离散小波变换 (Discrete Wavelet Transform, DWT) 是一种有效的解决方案^[29], 将其嵌入到网络中, 在确保图像质量不受影响的前提下, 可有效减少对硬件资源的依赖. 小波变换模块利用二维离散 Haar 小波变换 (discrete Haar wavelet) 将输入图像分解为 4 个子带, 即 LL、LH、HL 和 HH. Haar 小波变换的滤波器定义为

$$\begin{aligned} f_{LL} &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, & f_{LH} &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \\ f_{HL} &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}, & f_{HH} &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{aligned} \quad (8)$$

其中, f_{LL} 表示低通滤波器, f_{LH} 、 f_{HL} 和 f_{HH} 分别对应三个带通滤波器.

Haar 小波变换的滤波器是正交的, 使用相同的滤波器执行逆离散小波变换能够有效地将图像恢复至原始空间分辨率.

(2) CCTransBlock 模块

CCTransBlock 是 WCCTNet 的核心模块, 主要由 CCFBlock 和 CTransformer 模块构成. CCFBlock 包含两条并行支路和一个复数卷积层. 这两条并行支路分别用于空间特征和频域特征的提取, 复数卷积层则用于对这两个域的特征进行融合. 空间特征提取支路由两个级联的 3×3 复数卷积层组成, 两个卷积层中间加

CRELU 激活函数; 频域特征提取支路由傅立叶变换 (Fast Fourier Transform, FFT)、特征提取和傅立叶逆变换 (Inverse Fast Fourier Transform, IFFT) 三部分构成. 首先, 通过 FFT 将输入特征转换到频域. 然后, 使用两个 1×1 的复数卷积层, 中间加 CRELU 激活函数提取频域特征. 最后, 将提取到的频域特征通过 IFFT 变换到空间域. 两个域提取到的特征按通道拼接后, 再经过一个 3×3 的复数卷积层得到最终特征.

CTransformer 模块由复数转置注意力 (Complex Transposed Attention, CTA) 和复数门控前馈网络 (Complex Gated Feed-forward Network, CGFN) 构成, 如图 1 右下角所示. CTA 模块在通道维度上应用自注意力机制^[22], 通过计算不同特征图之间的相互关系动态调整各通道的权重, 实现对不同通道特征的自适应加权. 对复值数据的实部和虚部分别进行层归一化 (LayerNorm, LN)^[30] 可得到复数层归一化操作 (Complex LayerNorm, CLN). 首先, 对输入特征 $F_i \in \mathbb{C}^{C \times H \times W}$ 通过复数层归一化得到 $F_{CLN} \in \mathbb{C}^{C \times H \times W}$. 然后, 使用卷积核大小为 3×3 的复数深度卷积对 F_{CLN} 的每个通道进行分组卷积, 提取各通道的特征并改变其形状, 得到查询矩阵 Q 、键矩阵 K 和值矩阵 V . 最后, 通过 Q 、 K 和 V 计算全局注意力图, 进而得到输出特征 F_o . 其过程可表示为

$$\begin{aligned} Q &= \text{reshape}(W_Q \text{CLN}(F_i)) \\ K &= \text{reshape}(W_K \text{CLN}(F_i)) \end{aligned} \quad (9)$$

$$V = \text{reshape}(W_V \text{CLN}(F_i))$$

$$F_o = W(V \cdot \text{Softmax}(K \cdot Q/\alpha)) + F_i \quad (10)$$

其中, CLN 表示复数层归一化操作, reshape 表示形状重塑操作, 得到矩阵 $Q \in \mathbb{C}^{HW \times C}$ 、 $K \in \mathbb{C}^{C \times HW}$ 和 $V \in \mathbb{C}^{HW \times C}$, W_Q 、 W_K 和 W_V 表示 3×3 的复数深度卷积, W 表示 1×1 的复数卷积, α 表示可学习的尺度参数. 矩阵 K 和 Q 相乘得到的注意力图的维度为 $C \times C$, 远小于传统 Transformer 架构中图像尺寸的乘积 $HW \times HW$, 从而降低了运算复杂性.

传统 Transformer 架构中前馈网络 (Feed Forward Network, FFN) 通过全连接层实现^[21]. Restormer 模型提出了包含门控机制和深度卷积的模块 (Gated-Dconv Feed-forward Network, GDFN), 以增强局部特征提取能力^[22]. 本文进一步将 GDFN 推展至复数域, 提出了复数门控前馈网络 CGFN. CGFN 利用复数门控机制和复数深度卷积来增强网络的建模能力. 复数门控机制通过计算两条并行路径所捕获特征的逐元素乘积实现. 其中一条路径使用 CGeLU 非线性激活函数, 以自适应地调整特征图的权重, 强化重要特征并抑制非相关特征. 此外, CGFN 中使用复数深度卷积对空间上相邻的像素信息进行编码, 使模型能有效捕捉图像局部特征和空

间上下文信息,提高对图像细节的感知能力.对输入特征 $F_i \in \mathbb{C}^{C \times H \times W}$,CGFN 模块可表示为

$$F_o = W_3 \left(\left(\phi(W_1 \text{CLN}(F_i)) \right) \odot W_2 \text{CLN}(F_i) \right) + F_i \quad (11)$$

其中, W_1 和 W_2 表示 3×3 的复数深度卷积, W_3 表示 1×1 的复数卷积,CLN 表示复数层归一化操作, \odot 表示逐元素相乘, ϕ 表示 CGeLU 非线性激活函数.

(3) 特征融合模块(FusionBlock)

特征融合模块由两个复数卷积和一个 CTransformer 模块构成,用于融合编码器和对应解码器模块的特征.给定两个输入特征 $F_1 \in \mathbb{C}^{C \times H \times W}$ 和 $F_2 \in \mathbb{C}^{C \times H \times W}$,首先将它们沿通道维度拼接,得到特征 $F_c \in \mathbb{C}^{2C \times H \times W}$;然后,使用一个 1×1 的复数卷积层对 F_c 降维;最后,依次通过一个 CTransformer 模块和一个 1×1 的复数卷积层,得到最终输出特征 F_o . FusionBlock 表示为

$$F_o = W_2 \left(\text{CGFN} \left(\text{CTA} \left(W_1 \left(\text{concat}(F_1, F_2) \right) \right) \right) \right) \quad (12)$$

其中, W_1 和 W_2 表示 1×1 的复数卷积,concat 表示拼接算子,CTA 和 CGFN 分别表示 CTransformer 模块中的 CTA 模块和 CGFN 模块.

需要特别注意的是,WCCTNet 每一级包含 1 或 2 个 CCTransBlock 模块,每个 CCTransBlock 中仅包含 1 个 CCFBlock 和 CTransformer 模块,2 个融合模块中各包含 1 个 CTransformer 模块,模型中总共仅有 8 个 CTransformer 模块.

3 实验结果及分析

为评价所提出的 WCCTNet 的性能,本文采用峰值信噪比(Peak Signal to Noise Ratio, PSNR)、结构相似度(Structural Similarity, SSIM)、标准化均方误差(Normalized Root Mean Squared Error, NRMSE)、LPIPS(Learned Perceptual Image Patch Similarity)和 FID(Fréchet Inception Distance)作为定量评价指标,对所提出的算法进行评估.

3.1 数据集及实验设置

(1) 数据集

本文在两个公开的数据集,即 Calgary-Campinas^[31](简称 CC)单线圈数据集和 fastMRI^[32]单线圈数据集,使用不同的采样模式和采样率进行模型的训练与测试. CC 单线圈数据集包含 35 名受试者的复值单线圈数据,每个受试者包含大约 170 张大小为 256×256 的切片,选取其中 25 个对象共 4 254 张切片作为训练集,从剩余的 10 个对象中选取中间的 100 张切片,共 1 000 张二维图像作为测试集.

fastMRI 数据集包含 1 500 个复值单线圈冠状位质子密度加权膝关节 MRI 数据.每个 MRI 数据提供约 35 张膝关节冠状面图像,矩阵大小被中心裁剪为 320×320 .数据集被划分为训练集、验证集和测试集,由于测试集

中不含全采样的测量数据,实验中选取 fastMRI 验证集用于测试,即训练集包含 34 742 张图像,测试集包含 7 135 张图像.

(2) 采样策略

为获取欠采样的 K 空间图像,对全采样的 K 空间数据应用不同的欠采样掩膜.在 CC 数据集上使用欠采样率分别为 10% 和 30% (对应加速因子分别为 10 和 3.3) 的笛卡尔(Cartesian)采样掩膜,以及欠采样率均为 10% 的径向(Radial)采样掩膜和螺旋(Spiral)采样掩膜.图 2 展示了欠采样率为 10% 的笛卡尔掩膜、径向掩膜和螺旋掩膜.在 fastMRI 数据集上采用加速因子分别为 4 和 8 的笛卡尔欠采样掩膜.

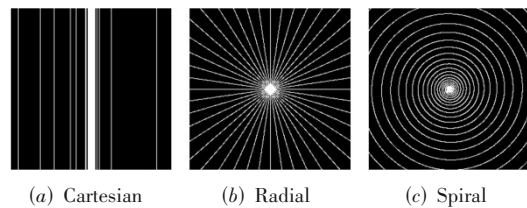


图 2 欠采样率为 10% 采样掩膜

(3) 实验参数设置

采用基于 Pytorch 的深度学习框架^[33]训练和测试网络性能,在两个数据集上均采用动量参数 $\beta_1=0.9$ 和 $\beta_2=0.999$ 的 Adam 优化器对损失函数进行最小化优化,初始学习率设置为 2×10^{-4} ,学习率采用余弦衰减策略下降到 1×10^{-6} ,批量大小设置为 8,浅层特征提取对应的特征通道数为 32.由于 CC 数据集和 fastMRI 数据集上训练样本数目不同,在两个数据集上设置不同的训练轮次,在 CC 数据集上训练 200 个轮次,在 fastMRI 数据集上训练 30 个轮次.网络的训练和测试均在配置为 Intel 至强 8336C CPU 和单个 NVIDIA RTX A6000 GPU 的 Ubuntu 20.04 操作系统上进行.将欠采样的测量数据进行傅里叶逆变换得到的初始重建图像作为网络输入,在 Calgary-Campinas(CC)数据集上训练 WCCTNet 的每个模型大约需要 14 h,而在 fastMRI 数据集上则大约需要 23 h.

3.2 实验结果及分析

3.2.1 在 CC 数据集上的性能分析

在 CC 数据集上,采用欠采样率为 10% 和 30% 的笛卡尔采样模式,以及欠采样率均为 10% 的径向采样模式和螺旋采样模式.为评估 WCCTNet 的性能,将其与几种基于深度学习的 MR 图像重建算法进行了比较分析,这些算法包括 SwinMR^[23]、KTMR^[18]、SwinGAN^[24]和 ReconFormer^[19].表 1 给出了几种重建算法在 CC 测试集上 1 000 张二维切片的平均 PSNR、SSIM、NRMSE、LPIPS 和 FID 值,表中性能最优的值用字体加粗标记.

从表 1 可以看出,WCCTNet 在不同采样模式及不

表1 不同重建算法在CC测试集上的性能评价

重建方法	采样模式	欠采样率	PSNR/dB	SSIM	NRMSE	LPIPS	FID
ZeroFilling	Cartesian	10%	22.29	0.623 1	0.077 3	0.367 8	330.87
SwinMR ^[23]			27.95	0.821 6	0.040 3	0.143 3	62.76
KTMR ^[18]			27.97	0.819 8	0.040 2	0.149 6	79.83
SwinGAN ^[24]			26.56	0.783 1	0.047 4	0.157 5	113.08
ReconFormer ^[19]			29.44	0.858 2	0.034 0	0.133 8	50.67
WCCTNet			29.92	0.870 1	0.032 2	0.113 3	47.92
ZeroFilling	Cartesian	30%	27.62	0.805 0	0.041 9	0.214 3	162.55
SwinMR ^[23]			33.92	0.932 2	0.020 3	0.056 8	22.47
KTMR ^[18]			33.65	0.929 6	0.021 0	0.067 8	32.29
SwinGAN ^[24]			32.92	0.920 5	0.022 8	0.087 8	28.53
ReconFormer ^[19]			36.28	0.955 1	0.015 5	0.059 0	20.78
WCCTNet			36.82	0.959 0	0.014 6	0.052 1	19.43
ZeroFilling	Radial	10%	23.17	0.529 4	0.069 8	0.393 0	311.46
SwinMR ^[23]			28.11	0.810 1	0.039 6	0.139 6	58.79
KTMR ^[18]			27.75	0.800 4	0.041 3	0.149 7	76.50
SwinGAN ^[24]			27.43	0.793 9	0.042 9	0.154 2	82.31
ReconFormer ^[19]			30.72	0.873 9	0.029 4	0.117 4	47.05
WCCTNet			31.28	0.887 5	0.027 6	0.127 7	45.63
ZeroFilling	Spiral	10%	22.81	0.578 5	0.072 8	0.374 1	343.80
SwinMR ^[23]			29.79	0.853 2	0.032 7	0.112 1	50.94
KTMR ^[18]			29.00	0.832 4	0.035 8	0.125 7	74.54
SwinGAN ^[24]			29.28	0.843 3	0.034 6	0.134 2	63.91
ReconFormer ^[19]			34.43	0.928 8	0.019 2	0.084 2	34.19
WCCTNet			34.80	0.933 6	0.018 3	0.083 3	32.11

同欠采样率下均展现了卓越的重建性能. 与 ReconFormer 相比, WCCTNet 在欠采样率为 10% 的笛卡尔采样模式、径向采样模式和螺旋采样模式下, 平均 PSNR 值分别提升了 0.48、0.56 和 0.37 dB, 平均 SSIM 值分别提升了 0.003 9、0.013 6 和 0.004 8. 在这些采样模式中, 径向采样模式下性能提升最显著, 螺旋采样模式下性能提升相对较小. 在相同采样模式但不同欠采样率(欠采样率分别为 10% 和 30%) 的笛卡尔采样模式下, WCCTNet 同样展现出最优的性能. 相较于 ReconFormer, 当欠采样率为 30% 时, WCCTNet 的 PSNR 和 SSIM 分别提高了 0.54 dB 和 0.003 9.

图 3~图 6 分别展示了不同重建算法在不同采样模式下得到的 MR 图像重建结果的直观对比结果. 为便于对重建图像的细节进行比较分析, 在每张图像中选取一个区域进行放大展示, 并提供相应的热力图以直观反映重建图像与真实图像之间的误差差异. 通过观察图 3~图 6, 可以发现 WCCTNet 和 ReconFormer 在整体重建质量和边缘信息方面更具有优势. 当欠采样率较低时, 即欠采样率为 10% 时, SwinGAN 的重建性能相对较差, 重建图像中存在大量伪影, 无法准确捕捉图像的细节信息. 相比之下, SwinMR 和 KTMR 能够重建图像

的大部分信息, 但仍有部分细节信息未能准确重建. 而 ReconFormer 和 WCCTNet 在重建图像中包含了更多的细节和边缘信息, 重建结果更接近于真实图像.

3.2.2 在 fastMRI 数据集上的性能分析

为进一步验证算法的有效性, 在 fastMRI 数据集上采用加速因子分别为 4 和 8 的笛卡尔采样模式进行 MR 图像重建. 由于 SwinGAN 公开的代码中要求输入图像空间分辨率为 256×256 , 而 fastMRI 数据集中图像的空间分辨率为 320×320 , 为公平起见, 仅对 WCCTNet 算法与 SwinMR^[23]、KTMR^[18]、ReconFormer^[19] 算法进行了比较. 表 2 给出几种算法在 fastMRI 测试集中 7 135 张二维切片的平均 PSNR、SSIM 和 NRMSE 值. 从表 2 的结果可以看出, 与 CC 数据集上的结果相似, WCCTNet 和 ReconFormer 的重建性能优于 SwinMR 和 KTMR 方法. 当加速因子为 4 时, WCCTNet 的 PSNR 与 ReconFormer 持平, 但其 SSIM 略高于 ReconFormer. 值得注意的是, 当加速因子增加至 8 时, WCCTNet 方法展现出较大的性能提升, 其平均 PSNR 值相较于 ReconFormer、SwinMR、KTMR 分别提高了 0.25、2.38 和 3.28 dB, SSIM 值分别提高了 0.008 9、0.031 9 和 0.052 3. 实验结果表明, 在高加速因子下, WCCTNet 能实现显著的性能提升.

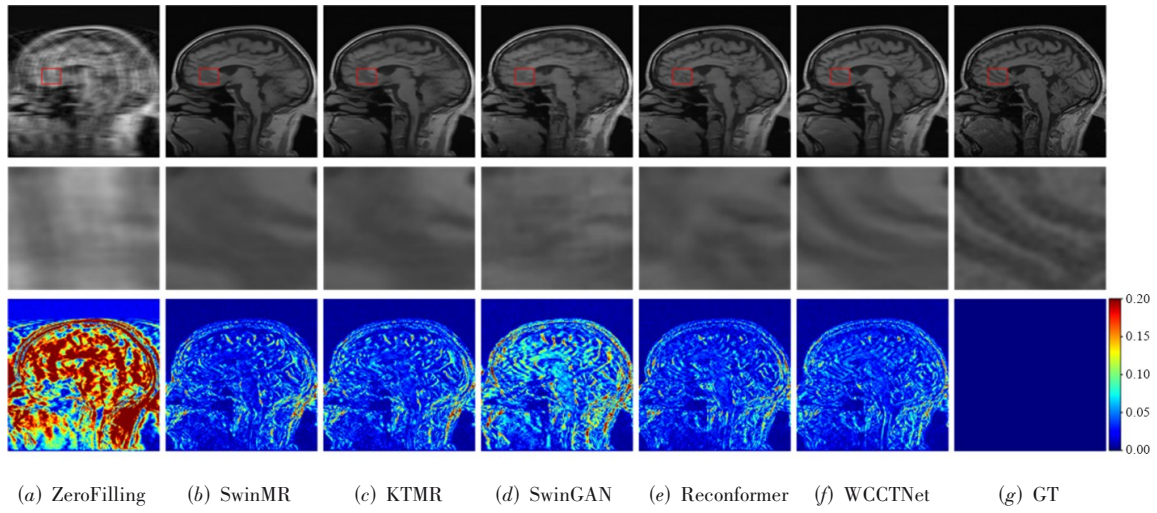


图3 不同算法在欠采样率为10%的笛卡尔掩膜下的重建图像

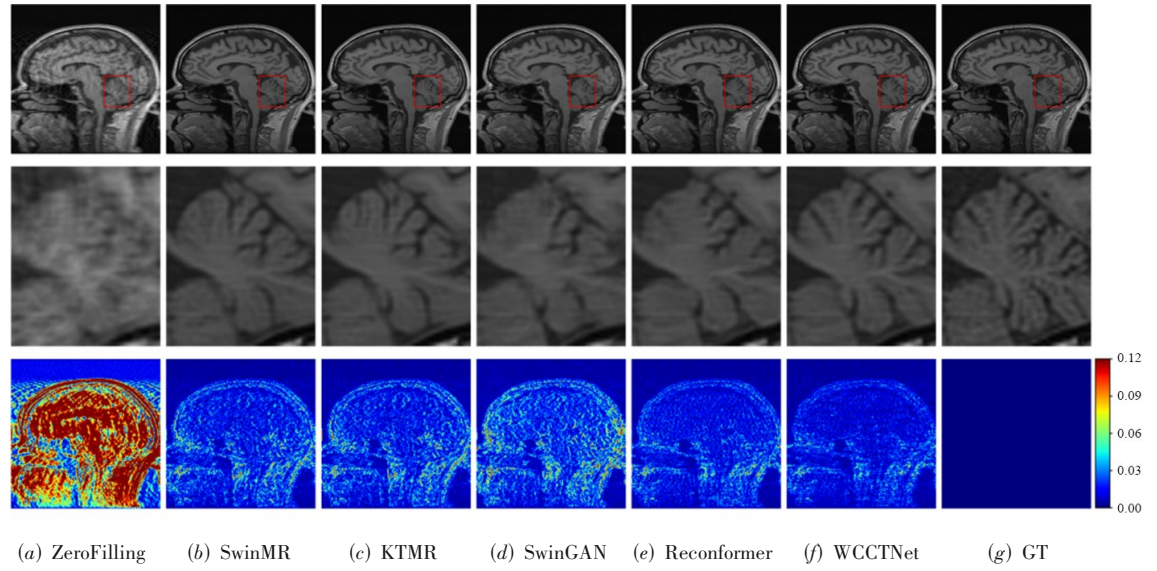


图4 不同算法在欠采样率为30%的笛卡尔掩膜下的重建图像

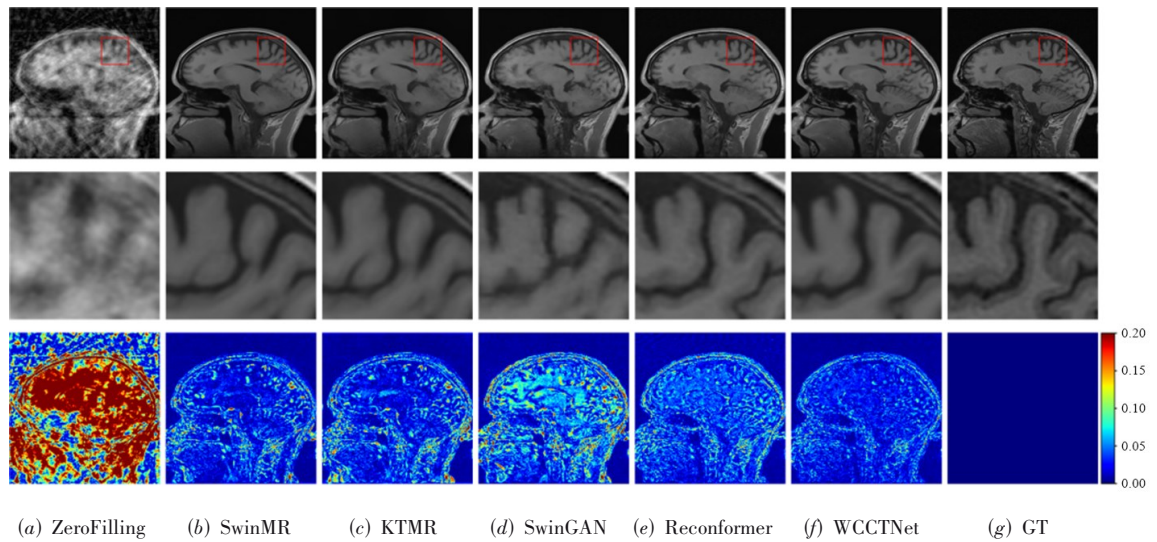


图5 不同算法在欠采样率为10%的径向掩膜下的重建图像

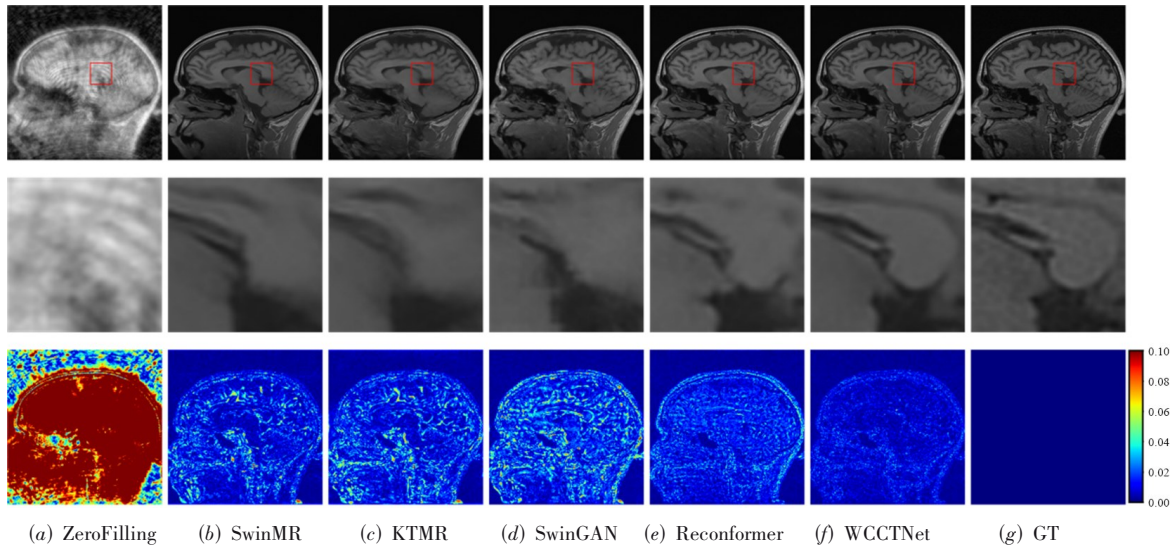


图6 不同算法在欠采样率为10%的螺旋掩膜下的重建图像

表2 不同算法在fastMRI数据集采用笛卡尔采样模式的重建结果

重建方法	加速因子	PSNR/dB	SSIM	NRMSE
ZeroFilling	4	29.50	0.654 9	0.052 7
SwinMR		31.77	0.720 4	0.034 9
KTMR		31.34	0.715 9	0.038 7
ReconFormer		32.73	0.738 3	0.031 0
WCCTNet		32.73	0.740 9	0.030 8
ZeroFilling		8	26.84	0.550 5
SwinMR	28.76		0.637 8	0.058 7
KTMR	27.86		0.626 3	0.069 1
ReconFormer	30.89		0.669 7	0.042 9
WCCTNet	31.14		0.678 6	0.040 4

图7和图8分别展示了在fastMRI数据集上,不同重建算法在加速因子为4和8的笛卡尔采样模式下,对

两个不同切片的重建效果,以及重建图像与真实图像之间误差分布的热力图.为便于对图像细节信息进行比较,在每个图像中选择一个区域并将其放大显示.可以看出,与其他方法相比,WCCTNet在重建质量上表现更佳,尤其在感知上显示出更好的伪影抑制能力,能够提供更明显的组织对比度和更清晰的视觉效果.

3.2.3 算法参数和运行时间分析

表3列出了各对比算法的参数数量以及在CC测试集上对1000张大小为 256×256 的二维切片进行GPU重建所需的平均时间.从表3数据可以看出,ReconFormer拥有最少的参数数量,但其重建时间相对较长.SwinGAN具有最短的重建时间,但其模型参数数量显著高于其他对比算法.相比之下,WCCTNet在保持较少参数的同时,显著提升了MR图像的重建效率.

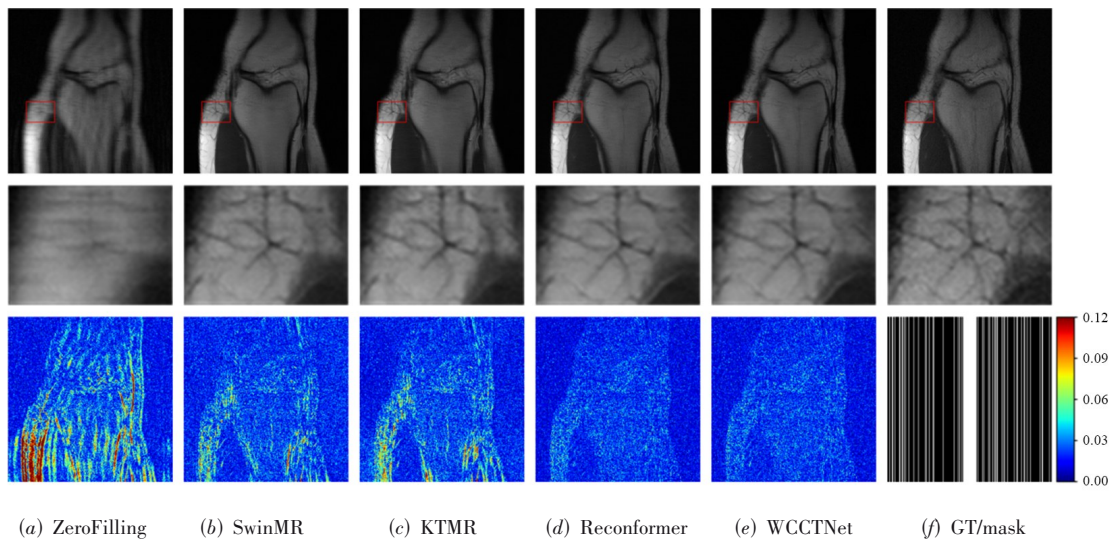


图7 不同算法在加速因子为4的笛卡尔采样模式下的重建图像

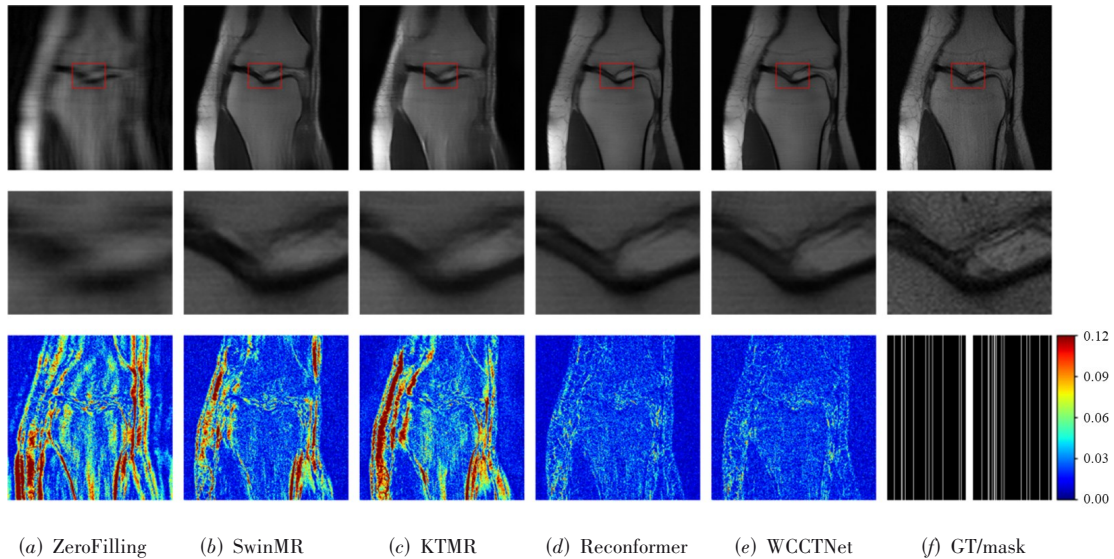


图8 不同算法在加速因子为8的笛卡尔采样模式下的重建图像

表3 不同重建算法的参数数量和GPU上的重建时间

重建方法	SwinMR	KTMR	SwinGAN	ReconFormer	WCCTNet
参数数量/M	11.50	11.60	82.50	1.14	4.52
重建时间/s	0.853	0.067	0.029	0.198	0.032

4 消融实验

为了深入探究影响 WCCTNet 网络性能的因素,通过消融实验分析网络中不同模块的有效性及其对网络性能的影响. 主要包括:(1)CCTransBlock 由两个 CCF-Block 构成;(2)CCTransBlock 由两个 CTransformer 模块构成;(3)CCTransBlock 由四个 CTransformer 模块构成;(4)移除 CCFBlock 模块中的频域分支;(5)用通道的拼接运算代替 FusionBlock;(6)将所有复数卷积替换为实数卷积,并保持浅层特征提取通道数不变;(7)将所有复数卷积替换为实数卷积,并将浅层特征提取的通道数翻倍. 为便于描述,将修改对应模块后的网络分别命名为 WC2Net、WCT2Net、WCT4Net、w/o FFT、w/o fuse、WRRTNetB 和 WRRTNetL. 所有消融实验均在 CC 数据

集上进行,采用欠采样率为 10% 的笛卡尔采样掩膜. 不同网络架构对应的重建结果如表 4 所示. 从表 4 的实验数据可以看出,复数卷积和 CTransformer 模块对 WCCTNet 的重建性能有重要的影响. 当完全移除 CTransformer 模块,仅使用卷积神经网络模块时,模型参数数量增加,性能有所下降,PSNR 值下降了 0.46 dB;当每个 CCTransBlock 模块中仅包含 2 个 CTransformer 模块时,参数数量大约是 WCCTNet 的一半,推理时间略有增加,性能显著下降,PSNR 值下降了 1.32 dB;当 CTransformer 模块数量增加至 4 时,参数数量与 WCCTNet 相当,推理时间翻倍,PSNR 值比 WCCTNet 下降了 0.92 dB. 移除 CCFBlock 中的频域分支时,性能下降较为明显. 去掉融合模块时,性能也有所降低. 将复数卷积替换为实数卷积时,性能下降最为显著. 尽管将浅层特征提取的通道数增加 1 倍显著提升了模型性能,但其重建结果仍不及 WCCTNet. 消融实验证实,在小波域将复数卷积和复数 Transformer 相结合的轻量级网络 WCCTNet 在降低资源需求、提高速度的同时,具有较优的重建性能.

表4 不同网络架构的重建性能比较

网络架构	PSNR/dB	SSIM	LPIPS	FID	参数数量/M	重建时间/s
WC2Net	29.46	0.856 9	0.127 2	53.37	7.094	0.029 7
WCT2Net	28.60	0.837 9	0.129 9	65.11	2.974	0.035 9
WCT4Net	29.00	0.848 3	0.126 7	64.80	4.562	0.066 4
w/o FFT	28.48	0.833 8	0.129 4	67.68	3.505	0.023 9
w/o fuse	29.49	0.858 6	0.122 5	54.67	5.237	0.026 6
WRRTNetB	28.40	0.830 2	0.146 6	70.19	2.262	0.029 7
WRRTNetL	29.15	0.848 7	0.131 8	60.89	8.949	0.035 9
WCCTNet	29.92	0.870 1	0.113 3	47.92	4.515	0.032 3

5 结论

本文将自注意力机制从实数域扩展到复数域,提出了一种新的模块,称为CTransformer. 基于CTransformer,结合复数卷积神经网络构建了一种基于小波域的混合轻量级网络WCCTNet,旨在实现复值MR图像的加速重建. WCCTNet采用了非对称的三级U型结构,每一级仅包含1或2个自注意力模块,有效减轻了传统Transformer架构高资源需求的问题. 在Calgary-Campinas数据集和fastMRI数据集上的实验结果表明,WCCTNet在保持重建性能的同时降低了对资源的需求. WCCTNet模型结构简洁、运算速度较快,展现出在多种应用场景中作为轻量级网络部署的潜力.

参考文献

- [1] MADL J E M, NIETO ALVAREZ I, AMFT O, et al. The psychological, physiological, and behavioral responses of patients to magnetic resonance imaging (MRI): A systematic review and meta-analysis[J]. *Journal of Magnetic Resonance Imaging*, 2024, 59(2): 675-687.
- [2] YOU X X, CAO N, WANG W. An MTLITV non-convex regularization model for MR Image reconstruction using the alternating direction method of multipliers[J]. *Electronic Research Archive*, 2024, 32(5): 3433-3456.
- [3] GENG C H, JIANG M F, FANG X, et al. HFIST-Net: High-throughput fast iterative shrinkage thresholding network for accelerating MR image reconstruction[J]. *Computer Methods and Programs in Biomedicine*, 2023, 232: 107440.
- [4] CHANDRA S S, BRAN LORENZANA M, LIU X W, et al. Deep learning in magnetic resonance image reconstruction[J]. *Journal of Medical Imaging and Radiation Oncology*, 2021, 65(5): 564-577.
- [5] GE Y M, XUE W, XU Y, et al. Magnetic resonance image denoising based on Laplacian prior sparsity constraint and nonconvex second-order TV penalty[J]. *Image Analysis & Stereology*, 2023, 42(2): 119-132.
- [6] LI Y Y, GAO L, HU S G, et al. Nonlocal low-rank plus deep denoising prior for robust image compressed sensing reconstruction[J]. *Expert Systems with Applications*, 2023, 228: 120456.
- [7] EKSIIOGLU E M. Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI[J]. *Journal of Mathematical Imaging and Vision*, 2016, 56(3): 430-440.
- [8] SOOMRO T A, ZHENG L H, AFIFI A J, et al. Image segmentation for MR brain tumor detection using machine learning: A review[J]. *IEEE Reviews in Biomedical Engineering*, 2023, 16: 70-90.
- [9] DAS N, DAS S. Attention-UNet architectures with pre-trained backbones for multi-class cardiac MR image segmentation[J]. *Current Problems in Cardiology*, 2024, 49(1): 102129.
- [10] HU Y X, LAFCI B, LUZGIN A, et al. Deep learning facilitates fully automated brain image registration of opto-acoustic tomography and magnetic resonance imaging[J]. *Biomedical Optics Express*, 2022, 13(9): 4817-4833.
- [11] LIANG D, CHENG J, KE Z W, et al. Deep magnetic resonance image reconstruction: Inverse problems meet neural networks[J]. *IEEE Signal Processing Magazine*, 2020, 37(1): 141-151.
- [12] WANG S S, SU Z H, YING L, et al. Accelerating magnetic resonance imaging via deep learning[C]//2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). Piscataway: IEEE, 2016: 514-517.
- [13] LEE D, YOO J, TAK S, et al. Deep residual learning for accelerated MRI using magnitude and phase networks[J]. *IEEE Transactions on Bio-Medical Engineering*, 2018, 65(9): 1985-1995.
- [14] SCHLEMPER J, CABALLERO J, HAJNAL J V, et al. A Deep Cascade of Convolutional Neural Networks for MR Image Reconstruction[M]//*Information Processing in Medical Imaging*. Cham: Springer International Publishing, 2017: 647-658.
- [15] WANG S S, CHENG H T, YING L, et al. Deep complex-MRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution[J]. *Magnetic Resonance Imaging*, 2020, 68: 136-147.
- [16] COLE E, CHENG J, PAULY J, et al. Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications[J]. *Magnetic Resonance in Medicine*, 2021, 86(2): 1093-1109.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//*Advances in Neural Information Processing Systems*. Long Beach: Curran Associates, 2017: 5998-6008.
- [18] WU Z L, LIAO W B, YAN C, et al. Deep learning based MRI reconstruction with transformer[J]. *Computer Methods and Programs in Biomedicine*, 2023, 233: 107452.
- [19] GUO P F, MEI Y Q, ZHOU J Y, et al. ReconFormer: Accelerated MRI reconstruction using recurrent transformer[J]. *IEEE Transactions on Medical Imaging*, 2024, 43(1): 582-593.

- [20] HAN K, WANG Y H, CHEN H T, et al. A survey on vision transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 87-110.
- [21] LIANG J Y, CAO J Z, SUN G L, et al. SwinIR: Image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Piscataway: IEEE, 2021: 1833-1844.
- [22] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: Efficient transformer for high-resolution image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 5718-5729.
- [23] HUANG J H, FANG Y Y, WU Y Z, et al. Swin transformer for fast MRI[J]. Neurocomputing, 2022, 493: 281-304.
- [24] ZHAO X, YANG T J, LI B J, et al. SwinGAN: A dual-domain Swin Transformer-based generative adversarial network for MRI reconstruction[J]. Computers in Biology and Medicine, 2023, 153: 106513.
- [25] 曹洁, 王乔, 梁浩鹏, 等. 深度复数轴向自注意力卷积循环网络的语音增强[J]. 计算机系统应用, 2024, 33(4): 60-68. CAO J, WANG Q, LIANG H P, et al. Speech enhancement based on deep complex axial self-attention convolutional recurrent network[J]. Computer Systems and Applications, 2024, 33(4): 60-68. (in Chinese)
- [26] EL-REWAIDY H, NEISIUS U, MANCIO J, et al. Deep complex convolutional network for fast reconstruction of 3D late gadolinium enhancement cardiac MRI[J]. NMR in Biomedicine, 2020, 33(7): e4312.
- [27] LEE C Y, HASEGAWA H, GAO S C. Complex-valued neural networks: A comprehensive survey[J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(8): 1406-1426.
- [28] LEE M. Mathematical analysis and performance evaluation of the GELU activation function in deep learning[J]. Journal of Mathematics, 2023, 2023: 4229924.
- [29] CHANG C L, GIROD B. Direction-adaptive discrete wavelet transform for image compression[J]. IEEE Transactions on Image Processing, 2007, 16(5): 1289-1302.
- [30] MIN T D, MANCINI M, ALAHARI K, et al. On the effectiveness of LayerNorm tuning for continual learning in vision transformers[C]//2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Piscataway: IEEE, 2023: 3577-3586.
- [31] 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.
- [32] ZBONTAR J, KNOLL F, SRIRAM A, et al. fastMRI: An open dataset and benchmarks for accelerated MRI[EB/OL]. (2018-11-21)[2025-04-01]. <https://arxiv.org/abs/1811.08839v2>.
- [33] SUBRAMANIAN V. Deep Learning with PyTorch: A Practical Approach to Building Neural Network Models Using PyTorch[M]. Birmingham, UK: Packt Publishing, 2018: 33-262.

作者简介



张晓华 女, 1979年12月出生于宁夏回族自治区青铜峡市. 现为燕山大学信息科学与工程学院博士生, 河北科技师范学院数学与信息科技学院讲师. 主要研究方向为医学图像重建、深度学习.

E-mail: zhang.xiaohua@126.com



练秋生 男, 1969年8月出生于江西遂川县. 现为燕山大学信息科学与工程学院教授, 博士生导师. 主要研究方向为图像处理、稀疏表示、压缩感知及深度学习等. 中国电子学会会员编号: E190026832M.

E-mail: lianqs@ysu.edu.cn