

小波分解U型Transformer加速MRI重构

熊承义^{ab}, 李帆^{ab*}, 高志荣^c, 孙清清^{ab}, 陈文旗^{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 视觉Transformer在提升图像恢复性能方面表现出了良好潜能. 研究了一种基于小波分解U型Transformer网络的加速磁共振图像(MRI)重建方法. 重建网络的核心单元结合了Swin Transformer与Unet结构, 通过融合图像的多尺度特征, 改进网络的学习能力, 以达到更好的重建性能. 采用小波变换对输入图像进行分解, 减少了Swin Transformer输入的特征维度, 从而有效降低了重构网络的计算复杂度. 通过引入小波域损失来约束网络的训练, 更好地恢复图像的结构纹理信息. 在Calgary-Campinas大脑MR数据集上进行实验比较, 结果验证了此方法在提升重构图像质量及控制系统计算复杂度方面的有效性.

关键词 磁共振图像重构; 深度学习; Swin Transformer模型; Unet网络; 小波变换

中图分类号 TP391.4 文献标志码 A 文章编号 1672-4321(2025)05-0695-08

doi:10.20056/j.cnki.ZNMDZK.20250515

Accelerated MRI reconstruction based on U-Transformer

XIONG Chengyi^{ab}, LI Fan^{ab*}, GAO Zhirong^c, SUN Qingqing^{ab}, CHEN Wenqi^{ab}

(South-Center 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 Transformer has shown great potential to improve image recovery performance. A reconstruction method of the accelerated magnetic resonance image (MRI) based on a wavelet-decomposed U-shaped Transformer network is studied. The core unit of the reconstruction network is designed based on combining the Swin Transformer with the Unet. By fusing multi-scale features of images, the learning ability of the network is improved and better reconstruction performance is obtained. By using wavelet transform to decompose the input image, the input feature dimension of Swin Transformer is reduced, so as to effectively reduce the computational complexity of the reconstruction network. The wavelet domain loss is adopted to constrain network training, with better recovering the structure and texture information of the image. The experimental results on Calgary-Campinas brain MR Dataset verify the effectiveness of the proposed method in improving the quality of reconstructed images and balancing the complexity of system.

Keywords magnetic resonance image reconstruction; deep learning; Swin Transformer; Unet; wavelet transform

磁共振成像(Magnetic Resonance Imaging, MRI)是临床中广泛应用的医学影像诊断技术之一, MRI扫描仪在空间频域(也称为k空间)中获得患者身体测量值, 对其进行傅里叶变换即可得到MR图像. 与X射线、CT等医学影像技术相比, MRI具有无辐射、无创伤、成像模态丰富、多对比度等特点.

为了得到高质量的MRI重构图像, 早期的传统

方法包括基于并行成像(Parallel Imaging, PI)^[1-5]与基于压缩感知(Compressed Sensing, CS)^[6-8]的方法. 随着深度学习(Deep Learning, DL)的发展, 数据驱动的重建算法^[9]利用大规模数据集提取更复杂、更精细的先验, 在性能上已经远远超过了传统方法. 其中, 基于卷积神经网络(Convolutional Neural Networks, CNN)的Unet^[10]模型, 因其可融合不同尺

收稿日期 2023-02-23 *通信作者 李帆, 研究方向: 磁共振图像重构, E-mail: 731369073@qq.com

作者简介 熊承义(1969-), 男, 教授, 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费专项资金资助项目(CZY21013)

度特征,且具有参数量少、结构简单等特点,已经被广泛应用于医学图像处理任务中.如 IKIK-net^[11]使用 4 个 Unet 网络进行级联,增加网络深度,提高模型的学习能力. DAGAN^[12]联合 Unet^[10]和对抗损失进行高质量 MR 图像重构.然而,基于卷积操作的 DL 网络模型缺乏对图像的全局理解,无法建模图像中的长距离依赖关系,从而限制了该类方法性能的提升.最近,一种基于自注意力机制的 Transformer 结构被提出,在自然语言处理领域中表现出色.受此启发, DOSOVITSKIY 等^[13]将 Transformer 引入图像处理领域,提出了具有里程碑式意义的视觉 Transformer (Vision Transformer, ViT),其利用 Transformer 的自注意力机制,对图像中的远距离依赖关系进行建模,有效弥补了 CNN 的不足.随后,Transformer 被迅速应用于 MRI 重建任务中^[14]. SwinMRI^[15]提出直接将现有的 Swin Transformer^[16] (SwinT) 图像复原网络模型用于 MRI 重建,在多种欠采样轨迹和不同噪声水平条件下其重构性能都得到了有效提升. HUMUS-Net^[17]将卷积的隐式偏置与 Transformer 相结合,通过卷积块提取高分辨率特征,并采用基于 Transformer 的多尺度特征提取器提取低分辨率特征,然后将提取的不同分辨率特征融合,重构出高质量的 MR 图像.目前,如何设计出更为有效的基于 Transformer 结构的 MRI 重构网络,已成为该领域的重要研究方向.

基于上述背景,本文研究了一种基于传统 U 型网络结构与视觉 Transformer 结合的 MRI 重构方法.设计的主要思想是将传统 Unet^[10]结构中的卷积层替换为 SwinT^[16]中的 Transformer 层,构建一个具有跳跃连接的 Transformer 编码器-解码器结构,通过引入自注意力机制获取多个尺度下的特征信息,以实现 MR 图像的高质量重建.为了降低网络复杂度,采用 Haar 小波变换对输入图像进行分解,减少网络输入的特征维度.同时,为了保留图像中更多的结构纹理信息,在训练阶段引入了一种小波损失函数,通过对 MR 图像的高频信息加以约束,来增强网络对高频细节信息的重构能力.实验结果表明,本文方法在重构性能改进与平衡系统复杂度方面取得了较好结果.

1 相关问题描述

1.1 MRI 重构

设 $x \in C^N$ 为待重构的复数值 MR 图像,图像尺

寸为 $\sqrt{N} \times \sqrt{N}$, $y \in C^M$ 表示欠采样的 k 空间观测值.重建问题为从部分采样数据 y 中重建出 x . k 空间测量值 y 表示为:

$$y = F_u x + \varepsilon, \quad (1)$$

其中, ε 为噪声项, F_u 为欠采样傅里叶编码矩阵.由于在欠采样的模式下,其采样率往往低于 Nyquist 采样频率 ($M \ll N$),这使得式(1)的逆问题成为病态问题.而解决这一问题往往需要利用数据本身具有的先验知识,因此图像 x 的重建可以转化为一个无约束优化问题:

$$\min_x \frac{1}{2} \|F_u x - y\|_2^2 + \lambda R(x), \quad (2)$$

其中, λ 为平衡数据保真度与先验知识的平衡因子, $R(\cdot)$ 为正则化项,在传统方法中, $R(\cdot)$ 通常是 x 在特定变换域中(如小波变换^[18]、全变分^[19]等)的 l_0 或 l_1 范数.这类方法计算复杂,需要手动设置参数,稀疏性先验不能完全拟合图像,且会在重构图像中引入阶梯伪影.

1.2 基于 CNN 的 MRI 重构

为了缓解手工正则化和传统模型重建时间长带来的影响,基于 CNN 的 CS-MRI 将 CNN 与 CS-MRI 结合^[20-22],通过优化神经网络的参数 θ 使得重建图像与对应的全采样图像尽可能接近.该方法表示为:

$$\min_x \frac{1}{2} \|F_u x - y\|_2^2 + \lambda \|f_{\text{cnn}}(x_u \theta) - x\|_2^2, \quad (3)$$

其中, f_{cnn} 由 θ 参数化的 CNN 前向映射,其输入为待重构的零填充图像 $x_u = F^H y$, F^H 为傅里叶逆变换,并直接产生重构结果作为输出.这类基于 CNN 的图像重构方法,通常在大量训练数据的监督下学习得到输入和输出图像之间的非线性映射 $x_{\text{cnn}} = f_{\text{cnn}}(x_u \theta)$.给定一个训练数据集 Ω ,它由采样值和原始图像对 (y, x) 组成,基于 CNN 的 CS-MRI 重建方法,主要任务在于构建网络模型,并通过最小化损失函数来优化网络参数,即:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \frac{1}{|\Omega|} \sum_{(x,y) \in \Omega} L(x, x_{\text{cnn}}), \quad (4)$$

其中, $L(\cdot)$ 是定义的损失函数.基于 CNN 的方法可自动探索图像的复杂特征,能够大大降低重构时间.然而,这类方法的性能会受到卷积局部敏感性的限制.

1.3 Swin Transformer

Swin Transformer^[16] (SwinT) 中提出的窗口多头自注意力 (Window Multi-head Self Attention, W-MSA) 和移位窗口多头自注意力 (Shifted Window Multi-

head Self Attention, SW-MSA)缓解了 Transformer 在计算机视觉领域应用的困难. 其核心模块如图 1 所示,由两个连接的 SwinT 层构成. 每个 SwinT 层都包含层归一化(Layer Normalization, LN)、多头自注意力模块、残差连接和多层感知机 (Multi-Layer Perceptron, MLP),其中多层感知机(MLP)由两个线性层及高斯误差线性单元(Gaussian Error Linear Unit, GELU)组成. 基于窗口的多头自注意力(W-MSA)和移位窗口的多头自注意力(SW-MSA)模块分别应用于两个连续的 SwinT 层. 整体计算过程如下:

$$\hat{z}^l = W - \text{MSA}(\text{LN}(z^{l-1})) + z^{l-1}, \quad (5)$$

$$z^l = \text{MLP}(\text{LN}(\hat{z}^l)) + \hat{z}^l, \quad (6)$$

$$\hat{z}^{l+1} = \text{SW} - \text{MSA}(\text{LN}(z^l)) + z^l, \quad (7)$$

$$z^{l+1} = \text{MLP}(\text{LN}(\hat{z}^{l+1})) + \hat{z}^{l+1}, \quad (8)$$

其中, z^l 表示第 l 个模块经 MLP 和残差连接后输出的特征, \hat{z}^l 表示经 W-MSA 或者 SW-MSA 和残差连接后输出的特征.

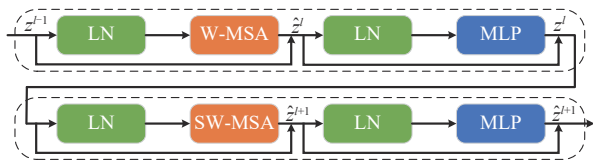


图 1 Swin Transformer 层

Fig. 1 Swin Transformer layer

多头自注意力中每个头的注意力计算如下:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}} + B\right)V, \quad (9)$$

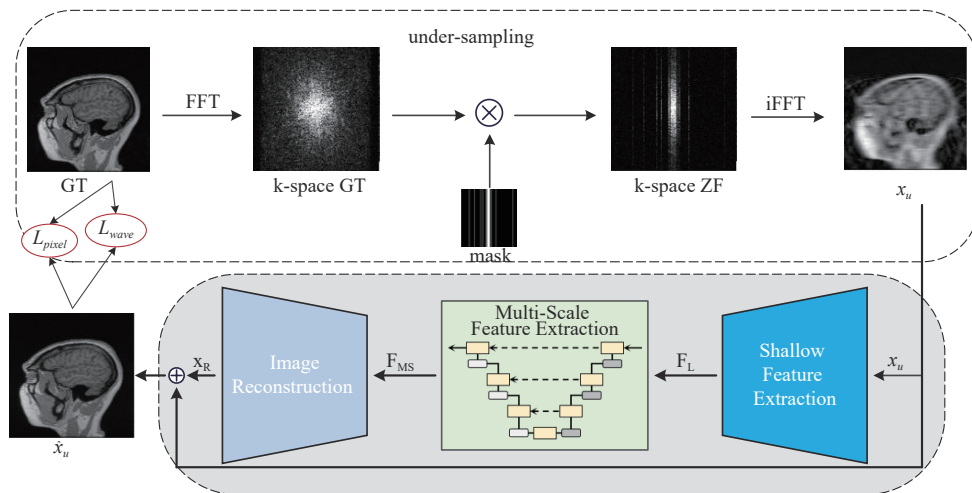


图 2 系统的整体结构框图

Fig. 2 The overall structure of the system

$$\hat{x}_u = f(x_u | \theta) + x_u, \quad (10)$$

其中, f 表示提出的网络, θ 为需要学习的网络参数.

其中 $Q, K, V \in R^{M^2 \times d}$ 分别表示查询(Query)、键(Key)和值(Value)矩阵, M^2 表示一个窗口里图像块(patch)的数量, d 表示序列特征的维度; 相对位置编码矩阵 $B \in M^2 \times M^2$ 的值来自于偏置矩阵 $\hat{B} \in R^{(2M-1) \times (2M-1)}$.

2 提出的方法

本文提出了一种 SwinT^[16]与 Unet^[10]结合的加速 MRI 重构网络, 整个网络为 U 型对称结构. 输入的 MR 图像被分割成不重叠的图像块, 每个图像块都被视为一个序列并输入至基于 Swin Transformer 的编码器中学习多尺度特征. 提取出的深度特征再使用解码器进行上采样, 并通过跳跃连接与来自编码器的多尺度特征进行融合, 改进网络学习能力, 以恢复更多的图像细节.

2.1 网络结构

MRI 采样与重构的系统结构框图如图 2 所示. 图像获取建模为原始图像经过快速傅里叶变换(Fast Fourier Transform, FFT)得到其 k 空间的欠采样表示. 重构过程首先对给定的 k 空间欠采样表示进行快速傅里叶逆变换(inverse Fast Fourier Transform, iFFT), 还原真实 MR 图像(Ground Truth, GT) x 的欠采样零填充重构 MR 图像 x_u . 重构网络的任务是利用低质量的 x_u 生成高质量的重构 MR \hat{x}_u . 并使用残差连接^[23]来加快和稳定训练过程. 重构过程可表示为:

重构网络主要由三部分组成: 浅层特征提取模块(Shallow Feature Extraction, SFE)、多尺度特征提

取模块(Multi-Scale Feature Extraction, MSFE)和图像重构模块(Image Reconstruction, IR).

浅层特征提取模块(SFE)如图 3(a)所示. 首先利用小波变换层(Discrete Wavelet Transform, DWT)对输入图像 $x_u \in R^{h \times w \times 2}$ (h 为图像高度, w 为图像宽度, 2 为图像的通道数)进行分解, 使其在空间维度上减小一半, 通道维度上增加 4 倍. 随后利用 3×3 卷积层将经过分解处理的特征映射至 d 维空间. 该过程可表示为:

$$F_L = H_L(\text{DWT}(x_u)), \quad (11)$$

其中, $\text{DWT}(\cdot)$ 为小波变换层, $H_L(\cdot)$ 为 3×3 卷积层, $F_L \in R^{\frac{h}{2} \times \frac{w}{2} \times d}$ 为得到的浅层特征.

多尺度特征提取模块(MSFE), 如图 4 所示, 由基于SwinT的编码器(Encoder)、瓶颈层(Bottleneck)、

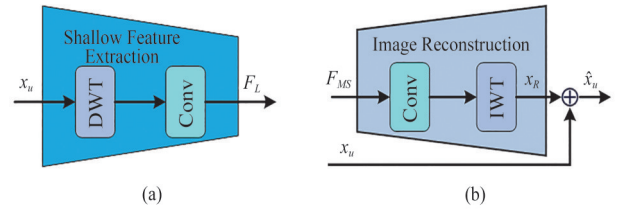


图 3 浅层特征提取模块(a)和图像重构模块(b)

Fig. 3 (a) Shallow feature extraction module; (b) Image reconstruction module

解码器(Decoder)和跳跃连接组成. F_L 输入至编码器中提取多尺度纹理特征, 然后在解码器中进行上采样, 并通过跳跃连接与来自编码器的多尺度特征融合. 该过程可表示为:

$$F_{MS} = H_{MSFE}(F_L), \quad (12)$$

其中, $H_{MSFE}(\cdot)$ 表示的特定架构将在 2.2 小节中详细介绍, $F_{MS} \in R^{\frac{h}{2} \times \frac{w}{2} \times d}$ 为得到的多尺度特征.

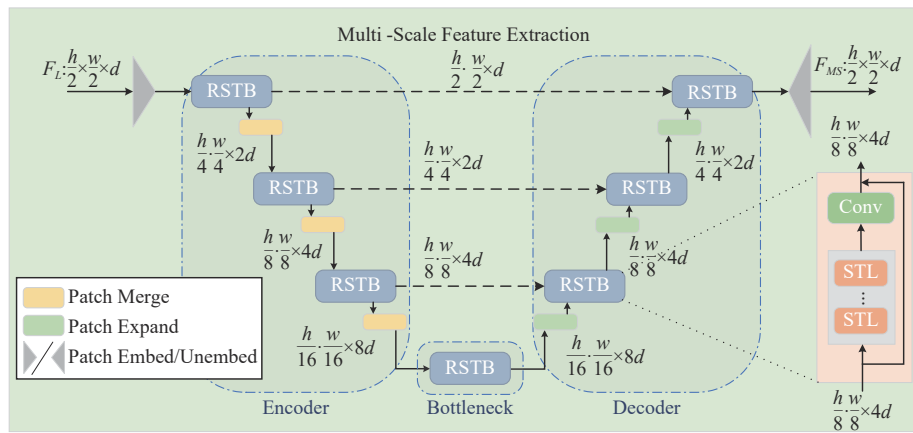


图 4 多尺度特征提取模块

Fig. 4 Multi-Scale Feature Extraction module

图像重构模块(IR), 如图 3(b)所示, 由 3×3 卷积层和逆小波变换(Inverse Wavelet Transformer, IWT)层构成, IWT使得特征图在空间维度上增加一半, 通道维度上减少 4 倍, 将其扩展至输入分辨率大小, 并重构出高分辨率残差图像 x_R , 同时引入残差连接, 得到最终的重构图像 \hat{x}_u :

$$x_R = \text{IWT}(H_R(F_{MS})), \quad (13)$$

$$\hat{x}_u = x_u + x_R, \quad (14)$$

其中, $H_R(\cdot)$ 表示 3×3 卷积层, $\text{IWT}(\cdot)$ 表示逆小波变换层.

2.2 多尺度深层特征提取模块

多尺度特征提取模块(MSFE)的整体结构如图 4 所示, 这是一个多尺度分层编码器-解码器结构, 用于提取多尺度特征. 首先, 通过 Patch Embed 操作

将形状为 $\frac{h}{2} \times \frac{w}{2} \times d$ 的 F_L 转换为维度为 d 的 $\frac{h}{2} \cdot \frac{w}{2}$ tokens. 随后, 将 tokens 合并输入至编码器学习多尺度特征. 最后, 由解码器通过跳跃连接融合来自瓶颈层路径的深层特征和编码器路径的多尺度特征, 并逐步扩展. 最终通过 Patch Unembed 得到输出为 $\frac{h}{2} \times \frac{w}{2} \times d$ 的多尺度特征图 F_{MS} .

其设计灵感来自于残差 SwinT 块(Residual Swin Transformer Blocks, RSTB)在图像去噪和超分辨率任务^[24]中的成功. RSTB 由堆叠的 Swin Transformer 层(STL)组成, 如图 4 右侧部分所示. 然而, RSTB 只能在单一尺度上操作, 不适合应用于分层编码器-解码器结构. 因此, 引入了块合并层(Patch Merge, PM)和块扩展层(Patch Expand, PE)分

别对中间的 tokens 进行下采样和上采样,以促进多尺度操作. 具体来说,PM类似于卷积神经网络中的池化层(Pooling),但不同的是,使用池化操作会丢失图像信息. 而PM则是通过线性方式将对应于 2×2 非重叠图像块的 tokens 进行合并,从而使得 tokens 的数量减少4倍,通道数增加4倍,并避免了图像信息丢失等问题. 然后,通过一个全连接层将通道数调整为输入通道数的2倍,以实现图像分辨率的2倍下采样. PE则为PM的逆操作,用于实现图像分辨率的2倍上采样.

2.2.1 编码器(Encoder)

Encoder用于学习多尺度特征信息. 将大小为 $\frac{h}{2} \cdot \frac{w}{2} \times d$ 的 tokens 输入至 Encoder 中并使用3个RSTB块进行特征学习,其中每两个RSTB块之间通过块合并层(PM)进行下采样操作,减少 tokens 数量,并增加特征的通道维度. 同时通过跳跃连接输出合并前的高维特征至解码器中,最终 Encoder 的输出大小为 $\frac{h}{16} \cdot \frac{w}{16} \times 8d$.

2.2.2 瓶颈层(Bottleneck)

由于 Encoder 中构建的网络层数较深,会导致网络训练时难以收敛^[25],因此使用1个RSTB块构建瓶颈层,用于学习深层特征,其特征维度和分辨率保持不变.

2.2.3 解码器(Decoder)

解码器是与编码器相对称的结构,用于融合来自编码器路径的多尺度特征和来自瓶颈层的深度特征. 将大小为 $\frac{h}{16} \cdot \frac{w}{16} \times 8d$ 的 tokens 输入至 Decoder 中,首先通过块扩展层(PE)对其进行上采样增加 tokens 数量,并相应地减少特征的通道维度,随后通过跳跃连接将来自编码器的多尺度特征与经过块扩展层(PE)扩展的特征进行融合,缓解下采样导致的空间信息损失,最后通过RSTB块合并增强. 重复上述过程3次,最终 Decoder 的输出大小为 $\frac{h}{2} \cdot \frac{w}{2} \times d$.

2.3 损失函数

本文定义网络训练的损失函数包含像素级损失和小波损失两个部分. 像素级损失计算重构图像与原始图像在像素域的偏差,与广泛使用的L2损失相比,Charbonnier Loss^[26]具有更好的鲁棒性,能够更好地处理异常值. 因此像素级损失采用 Charbonnier Loss:

$$L_{\text{pixel}} = \frac{1}{n} \sum_{i=1}^n \sqrt{\|x^i - \hat{x}_u^i\|_2^2} + \xi, \quad (15)$$

其中, ξ 是一个常数,一般设置为 10^{-9} , x 和 \hat{x}_u 分别表示原始的无损图像和最终的重构图像.

定义小波损失为重构图像与原始图像在小波域的高频分量的偏差,用于约束图像高频信息以恢复更多的纹理细节. 具体公式为:

$$L_{\text{wave}} = 1 - \text{SSIM}\left(H_{\text{cat}}(\text{DWT}(x)), H_{\text{cat}}(\text{DWT}(\hat{x}_u))\right). \quad (16)$$

图像经过 $\text{DWT}(\cdot)$ 被分解为4个子带 LL、LH、HL 和 HH. 其中 LL 对应图像的低频信息, LH 对应图像水平方向的高频信息, HL 对应图像垂直方向的高频信息, HH 对应图像对角线的高频信息. 先利用 $H_{\text{cat}}(\cdot)$ 将3个高频子带沿通道维度进行合并,再计算 SSIM (Structural Similarity) 损失,以保留更多的高频纹理信息,提高网络重构性能.

最终,网络的总损失函数 L_{total} 为:

$$L_{\text{total}} = L_{\text{pixel}} + \lambda L_{\text{wave}}, \quad (17)$$

其中, L_{pixel} 为像素级损失, L_{wave} 为小波损失, λ 为小波损失项的权重系数.

3 实验

3.1 实验设置和训练数据

采用 Calgary-Campinas-359 Brain (CC359-Brain) 单线圈重构数据集^[27], CC359-Brain 数据是临床上使用商业 MR 扫描仪 (Discovery MR750, GE) 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 . 训练数据集使用23个样本包含4140张切片,验证集使用的4个样本包含720张切片,测试集使用8个样本包含1440张切片. 在训练与测试实验中,为了避免引入噪声数据,对于每个样本只选用100个中心切片,同时为了提高模型的收敛速度和精度,本文对数据进行了归一化处理. BatchSize 设置为8,训练200个周期. 学习率初始值设为 4×10^{-4} ,使用余弦退火函数 (CosineAnnealingLR) 调整学习率,周期 T_{max} 为40. 损失函数权重系数 $\lambda = 0.3$. 所有实验使用 Python 语言在 Pytorch 深度学习框架下完成,使用 Adam 优化器优化所有网络参数. 在 Ubuntu20.04 系统上使用两张 24 GB 的 NVIDIA RTX 3090 显卡完成训练和测试.

3.2 欠采样模式和评价指标

为了验证本文算法在不同的欠采样模式和采样率下的性能,进行了一系列实验.如图5所示,分别为高斯欠采样(Gaussian 1D)、径向欠采样(Radial)和螺旋欠采样(Spiral).

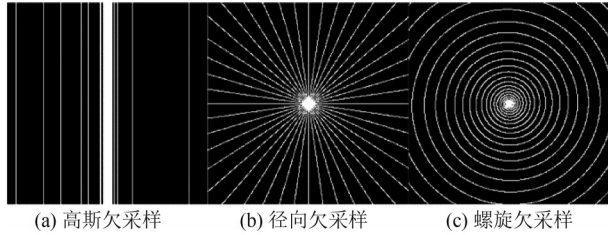


图5 不同的欠采样模式

Fig. 5 Different under-sampling masks

采用了三种通用的图像质量评价指标来评估模型性能.分别为峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似度(Structural Similarity, SSIM)和标准化均方误差(Normalized Mean Squared Error, NMSE). PSNR是最为普遍使用的一种图像客

观评价指标,它是基于对应像素点之间的误差. SSIM是一种全参考的图像质量评价指标.其分别从亮度、对比度、结构三方面度量图像相似性. NMSE是基于误差敏感的图像质量评价,强调整体重建精度,其数值越低表示重构图像与参考图像越接近.

3.3 实验结果

将本文模型与其他基准MRI重建方法进行比较,定量结果如表1所示,最好结果用加粗字体标出.在不同欠采样模式和采样率下,提出的方法均取得了最好的重构质量,其PSNR、SSIM、NMSE指标对比SwinMRI均有显著提升.如在较高采样率如Gaussian 1D 50%时,本文算法的PSNR和SSIM指标分别比SwinMRI高3.75dB和0.025,比U-Net高6.6 dB和0.059, NMSE指标比SwinMRI低0.0024,比U-Net低0.0077;在较低采样率如Spiral 10%时,本文算法的PSNR和SSIM指标分别比SwinMRI高1.81 dB和0.004,比U-Net^[10]高3.21 dB和0.057, NMSE指标比SwinMRI低0.0027,比U-Net低0.0175.

表1 不同欠采样率下不同模式的重建精度

Tab. 1 Reconstruction accuracy of different mask under different under-sampling rate

Mask	ZF			U-Net ^[10]			SwinMRI ^[15]			This paper		
	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE
G1D10%	23.72	0.5746	0.0918	25.64	0.7453	0.0447	27.85	0.7816	0.0355	28.88	0.8176	0.0277
G1D30%	27.81	0.8211	0.0271	31.48	0.8762	0.0141	33.06	0.9041	0.0116	35.83	0.9420	0.0056
G1D50%	30.49	0.8654	0.0112	34.42	0.9145	0.0084	37.27	0.9488	0.0041	41.02	0.9739	0.0017
R10%	24.18	0.6105	0.0749	26.64	0.7510	0.0456	28.38	0.7862	0.0337	30.06	0.8340	0.0212
S10%	24.58	0.6248	0.0813	27.02	0.7763	0.0381	28.42	0.8295	0.0233	30.23	0.8334	0.0206

为了比较不同重构算法的主观视觉效果,本文给出了在Gaussian 1D 30%(G1D30%)、Gaussian 1D 50%(G1D50%)、Radial 10%(R10%)欠采样模式下的重建图像,以及重构图像局部2.5倍放大图,如图6所示.观察重构结果可知,本文算法重构图像的结构纹理更加清晰,产生的伪影较少.

3.4 消融实验及分析

为了验证MSFE模块中U型架构的有效性,在Gaussian 1D 30%采样率下进行消融实验.如表2所示,其中线性结构_a表示使用线性结构的网络,其输入通道与本文所提网络一致;线性结构_b表示增大使用线性结构的网络的输入通道数,使其参数量与本文所提网络相当.实验结果表明,当MSFE模块采用线性架构时,网络的重构性能明显下降.这是因为线性架构只能捕捉单一尺度的特征信息,而无法获取多尺度结构信息.相比之下,U型架构则克服了这一缺陷,通过多级残差连接可以减轻网络的训

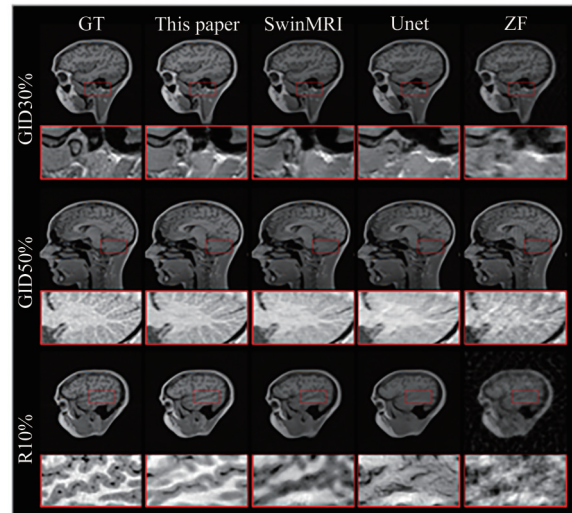


图6 不同采样率下各重构算法的主观视觉效果

Fig. 6 The subjective visual quality of various reconstruction algorithms at different sampling rates

练难度,提高网络的综合实力,从而明显提高了重构性能.

表2 MSFE中U型架构的消融实验

	参数量/M	计算量/GMac	PSNR	SSIM
线性架构_a	0.798	12.54	33.42	0.897
线性架构_b	12.16	198.33	35.32	0.935
U型架构	12.52	13.61	35.83	0.9420

为了验证使用小波变换(DWT)层对于减少网络计算复杂度的有效性,本文在 Gaussian 1D 30% 采样率下进行了消融实验,具体结果如表3所示. 实验表明,使用小波变换(DWT)对输入图像进行分解,减少了输入的特征维度,从而显著降低了网络计算量,改善了网络的复杂度. 同时,使用一级小波变换时,对重构网络性能影响甚微. 然而,当使用二级或三级小波变换时,尽管网络的计算量进一步降低,但由于使用多级小波变换,导致网络中间层的特征图尺寸过小,影响了网络的重构质量. 基于此,本文采用一级小波变换层来平衡网络的计算复杂度与性能.

表3 DWT的消融实验

	参数量/M	计算量/GMac	PSNR	SSIM
without DWT	12.52	54.46	35.87	0.9421
with DWT	12.52	13.61	35.83	0.9420
with DWT*2	12.52	3.4	34.62	0.9285
with DWT*3	12.52	0.85	33.14	0.9017

为了验证小波损失的有效性,尝试采用不同的损失权重,并在 Gaussian 1D 30% 采样率下,评估小波损失对网络性能的影响,结果如表4所示. 从表中可以看出,相较于仅使用单一像素及损失($\lambda = 0$),增加小波损失可以明显提升重构质量,当权重系数 $\lambda = 0.3$ 时,可获得最佳的重构效果. 这是因为在小波域增加了约束后,网络能够从中学习更多高频特征,为重构过程提供额外的结构信息,有助于恢复高质量的MR图像.

表4 小波损失的消融实验

	$\lambda = 0$	$\lambda = 0.01$	$\lambda = 0.1$	$\lambda = 0.3$	$\lambda = 0.5$
PSNR	34.60	35.19	35.22	35.41	34.93
SSIM	0.9299	0.9363	0.9356	0.9388	0.9327
NMSE	0.0074	0.0065	0.0064	0.0060	0.0069

4 结语

研究了一种基于传统U型网络结构与 Swin Transformer 结合的MRI重构方法,并提出了一种U型Transformer的加速MRI重构网络. 网络为端到端

的U型对称结构,可以同时捕捉多尺度特征信息和降低网络的参数和计算复杂度. 引入小波损失约束MR图像的高频信息,以达到保留更多图像的结构纹理信息,提高重构性能的目的. 在不同采样轨迹的 Calgary-Campinas 大脑MR数据集上进行了大量实验,结果证明了本文提出的网络可有效提升MR图像重构质量,并在模型复杂度和性能之间取得了较好的平衡.

参考文献

- [1] BLAIMER M, BREUER F, MUELLER M, et al. SMASH, SENSE, PILS, GRAPPA: How to choose the optimal method [J]. *Topics in Magnetic Resonance Imaging*, 2004, 15(4): 223-236.
- [2] 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.
- [3] PRUESSMANN K P, WEIGER M, BORNERT P, et al. Advances in sensitivity encoding with arbitrary k-Space trajectories [J]. *Magnetic Resonance in Medicine*, 2001, 46: 638-651.
- [4] 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.
- [5] LUSTIG M, DONOHO D L, SANTOS J M, et al. Compressed sensing MRI [J]. *IEEE Signal Processing Magazine*, 2008, 25(2): 72-82.
- [6] LIANG D, LIU B, WANG J, et al. Accelerating SENSE using compressed sensing [J]. *Magnetic Resonance in Medicine*, 2009, 62(6): 1574-1584.
- [7] 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.
- [8] HONG M, YU Y, WANG H, et al. Compressed sensing MRI with singular value decomposition-based sparsity basis [J]. *Physics in Medicine & Biology*, 2011, 56(19): 6311.
- [9] WANG S, XIAO T, LIU Q, et al. Deep learning for fast MR imaging: A review for learning reconstruction from incomplete k-space data [J]. *Biomedical Signal Processing and Control*, 2021, 68: 102579.
- [10] 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.

- Switzerland: Springer, 2015: 234-241.
- [11] SOUZA R, BENTO M, NOGOVITSYN N, et al. Dual-domain cascade of U-nets for multi-channel magnetic resonance image reconstruction [J]. *Magnetic resonance imaging*, 2020, 71: 140-153.
- [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] DOSOVITSKIY A, Beyer L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale [C]//International Conference on Learning Representations. Vienna: ICLR, 2021: 1-14.
- [14] SHAMSHAD F, KHAN S, ZAMIR S W, et al. Transformers in medical imaging: A survey [J]. *arXiv: 2022, 2201.09873*.
- [15] HUANG J, FANG Y, WU Y, et al. Swin transformer for fast MRI [J]. *Neurocomputing*, 2022, 493: 281-304.
- [16] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]//International Conference on Computer Vision. Montreal: IEEE, 2021: 10012-10022.
- [17] FABIAN Z, TINAZ B, SOLTANOLKOTABI M. HUMUS-net: Hybrid unrolled multi-scale network architecture for accelerated MRI reconstruction [C]//Neural Information Processing System. San Diego: NeurIPS, 2022.
- [18] QU X, GUO D, NING B, et al. Undersampled MRI reconstruction with patch-based directional wavelets [J]. *Magnetic Resonance Imaging*, 2012, 30(7): 964-977.
- [19] BLOCK K T, UECKER M, FRAHM J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint [J]. *Magnetic Resonance in Medicine*, 2007, 57(6): 1086-1098.
- [20] HAN Y, SUMWOO L, Ye C J. *k*-Space deep learning for accelerated MRI [J]. *IEEE Transactions on Medical Imaging*, 2020, 39(2): 377-386.
- [21] RAN M, XIA W, HUANG Y, et al. MD-Recon-Net: A parallel dual-domain convolutional neural network for compressed sensing MRI [J]. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2021, 5(1): 120-135.
- [22] LEE D, YOO J, TAK S, et al. Deep residual learning for accelerated MRI using magnitude and phase networks [J]. *IEEE Transactions on Biomedical Engineering*, 2018, 65(9): 1985-1995.
- [23] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [24] LIANG J, CAO J, SUN G, et al. Swinir: Image restoration using swin transformer [C]//International Conference on Computer Vision. Montreal: IEEE, 2021: 1833-1844.
- [25] TOUBRON H, Cord M, SABLAYROLLES A, et al. Going deeper with image transformers [C]//International Conference on Computer Vision. Montreal: IEEE, 2021: 32-42.
- [26] 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.
- [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.

(责编&校对 姚春娜)