

DOI: 10. 19650/j. cnki. cjsi. J2412698

# 多尺度注意力融合与视觉 Transformer 方法优化的 电阻抗层析成像深度学习方法\*

王 琦<sup>1,2</sup>, 张 涛<sup>1,2</sup>, 徐超炜<sup>1</sup>, 卢梦凡<sup>1</sup>, 王子辰<sup>1,2</sup>

(1. 天津工业大学电子与信息工程学院 天津 300387; 2. 天津市光电检测技术与系统重点实验室 天津 300387)

**摘要:**电阻抗层析成像 (EIT) 具有显著的可视化和非侵入性等特点, 在工业和生物医学工程领域展现了其广阔的应用潜力。由于其逆问题存在高度非线性和病态性特点, 导致了数值成像方法在空间分辨率上的局限性, 尤其是在多相介质分布情况下, 现有 EIT 技术在成像过程中出现边界失真和电导率误差, 从而影响最终的成像精度。本文提出了一种基于卷积注意力机制的 U 型深度成像方法——MAT-UNet, 将卷积块注意力模块 (CBAM) 与 U-Net 结构相结合, 在特征提取与融合过程中嵌入卷积块注意力模块, 以增强模型的注意力定向和特征表征能力, 同时跳跃连接引入了压缩-激励 (SE) 注意力机制与视觉 Transformer (ViT) 来优化全局特征的学习, 使用多头交叉注意力模块 (MHCA) 实现编码器与解码器的多尺度信息融合。MAT-UNet 通过大量的仿真数据训练获得最优模型参数, 并在多样化复杂形状和肺部仿真模型进行了实验验证。定量评估指标表明, 该方法在重建图像中的均方根误差 (RMSE) 结果为 2.3156, 结构相似性指数 (SSIM) 结果为 0.9437, 可视化结果与真实分布和边界具有很好的 consistency。实验结果表明, 本文提出的 MAT-UNet 模型展现出良好的鲁棒性和泛化能力, 相较于传统的单一卷积结构, 集成 Transformer 结构提供了更精准的 EIT 图像重建效果, 在无损测量与检测应用中存在很大的潜力和价值。

**关键词:** 电阻抗层析成像; 卷积注意力机制; SE-ViT 连接; 多头交叉注意力模块; U 型卷积网络; 无损测量

**中图分类号:** R318 TH701 **文献标识码:** A **国家标准学科分类代码:** 460.40

## Optimized learning method for electrical impedance tomography with multi-scale attention fusion and vision transformer

Wang Qi<sup>1,2</sup>, Zhang Tao<sup>1,2</sup>, Xu Chaowei<sup>1</sup>, Lu Mengfan<sup>1</sup>, Wang Zichen<sup>1,2</sup>

(1. School of Electronics & Information Engineering, Tiangong University, Tianjin 300387, China;

2. Tianjin Key Laboratory of Optoelectronic Detection Technology and System, Tianjin 300387, China)

**Abstract:** Considering the advantage of visualization and non-invasiveness of electrical impedance tomography (EIT), it's broadly applied in industrial and biomedical fields. However, due to the highly nonlinear and ill-posed nature of inverse problem, numerical imaging methods face spatial resolution limitations. These limitations are especially evident in multiphase media distributions, where current EIT technology encounters boundary distortions and conductivity errors, thereby affecting the final imaging accuracy. To address mentioned issues, this paper introduces a learning-based model for EIT reconstruction, referred as MAT-UNet, which is mainly composed of U-shaped backbone and optimized multi-head attention block. The MAT-UNet integrates convolutional block attention module (CBAM) in the Encoders for feature extraction to construct the spatial and channel latent feature. In addition, the Squeeze-and-Excitation Vision Transformer (SE-ViT) is introduced in the skip connection between Encoder and Decoder, which optimizes the global feature learning. Also, the Multi-Head Cross-Attention (MHCA) module facilitates multi-scale information fusion between the encoder and decoder. MAT-UNet is trained on extensive simulation data to obtain optimal model parameters and is experimentally validated on diverse complex shapes and lung simulation models. The quantitative evaluation metrics indicate that this method achieves a Root Mean Square Error (RMSE) of 2.3156 and a Structural Similarity Index (SSIM) of 0.9437 in reconstructed images. The visualized results closely match the true distribution and boundaries. Experimental outcomes demonstrate that the proposed MAT-UNet model exhibits robust performance and generalization capability. Compared to traditional single convolution structures, the integration of the Transformer

收稿日期: 2024-04-07 Received Date: 2024-04-07

\* 基金项目: 国家自然科学基金 (62072335, 62071328, 61872269, 61903273) 项目资助

structure provides more accurate EIT image reconstruction, presenting significant potential and value in non-destructive measurement and detection applications.

**Keywords:** electrical impedance tomography; convolutional attention mechanism; SE-ViT connection; multi-head cross-attention mechanism; U-shaped convolutional network; non-destructive measurement

## 0 引言

在现代检测技术领域,过程参数可视化是一项非常关键的任务。电阻抗层析成像(electrical impedance tomography, EIT)因其非侵入性和无辐射的特点,在工业动态过程监测和医学生理参数的可视化方面显示出巨大的应用潜力<sup>[1]</sup>。但是,EIT测量只能利用有限的传感器来获取感应信号,且其调制电场的“软场”介质参数分布重建)的高度病态性和非线性,这使得重建图像尤其在复杂边界的表示和电导率参数的准确度明显降低。EIT反问题通过线性近似等方法构建灵敏度矩阵,来模拟边界测量电压与电场内参数分布之间的映射关系<sup>[2]</sup>。近十年的研究表明,改进反投影算法<sup>[3]</sup>、自适应混合总变差正则化<sup>[4]</sup>、组稀疏优化<sup>[5]</sup>和D-Bar方法<sup>[6]</sup>等优化算法,以及基于形状的重建框架,超形状方法<sup>[7]</sup>和Bayesian拓扑形状构建方法<sup>[8]</sup>,在处理EIT图像的复杂形状表示上取得了一定进步。然而,这些方法通常依赖于正问题模型的精确性和先验知识,并且迭代过程对噪声干扰十分敏感,因此,现有的方法在准确重建复杂电导率分布和内含物形状方面存在局限。

近年来,深度学习在图像重建领域尤其是医学图像重建方面取得了显著成就<sup>[9]</sup>。目前,电学层析成像的深度学习框架大多基于卷积神经网络(convolutional neural networks, CNNs),通过引入新的信息传递机制以增强图像边界并抑制伪影。卷积神经网络,凭借其自动特征提取和参数重构能力,以及巧妙融合空间先验信息的优势,在电学层析成像领域的图像重建任务中获得广泛应用<sup>[10]</sup>。例如,王子辰等<sup>[11]</sup>通过结合迭代展开的预重建模块和改进的注意力深度U形CNN,有效地解决了电学层析重建中由“软场”特性和逆问题病态性导致的边界伪影和低空间分辨率问题。Cen等<sup>[12]</sup>提出了一种名为Deep Calderón的方法,通过Calderón预重建方法获取电导率的初始分布,再利用U-Net锐化感兴趣区域(region of interest, ROI)的边界,以提高成像质量。尽管卷积神经网络在重建图像形状方面取得了显著进展,但其使用固定卷积核构建的感受野有时会忽视测量区域内部的空间位置特征和电学参数的分布。这种设计限制了CNNs在处理复杂的多相电导率分布时的能力,特别是在重建具有低电导率对比度的内含物分布方面,仍然面临着显著的挑战。与此同时,Transformer技术,在视觉应用领域

得到广泛运用,其核心在于采用自注意力机制来构建特征间的长距离依赖关系<sup>[13]</sup>。与传统的CNNs相比,Transformer减少了对局部性的依赖,增强了对全局特征交互的处理能力<sup>[14]</sup>。例如,TransUNet首次将Transformer应用于医学图像分割,通过结合卷积运算和Transformer,实现了对输入图像块的有效特征映射<sup>[15]</sup>。Zomir等<sup>[16]</sup>提出了一种高效的Transformer模型,设计的多头注意力和前馈网络,使其能够捕捉长距离像素交互,同时适用于大尺寸图像。冷冰等<sup>[17]</sup>提出基于DETR的Transformer模型T-DETR,有效提高了外周血白细胞检测的准确性和效率,减轻了人工镜检的压力。DS-TransUNet增加了一个新的编码器来处理多尺度输入,并通过自注意力机制在不同尺度的特征之间建立全局依赖关系<sup>[18]</sup>。类似地,熊巍等<sup>[19]</sup>提出了一种融合通道注意力(coordinate attention, CA)特征的Transformer模型CAET,有效提升了图像超分辨率重构的性能,同时降低了模型复杂度,使其更适合低成本和实时应用的需求。Guo等<sup>[20]</sup>提出ReconFormer,融合了循环卷积金字塔模块用于局部特征提取和Transformer全局特征捕捉能力,实现8×加速MRI成像。虽然Transformer在图像处理任务中性能突出,但是单独使用Transformer结构并不能完全发挥其长距离特征依赖的优势,尤其是在需要与卷积特征精确耦合的场景下,即使结合多个Transformer也不能达到满意效果。

针对上述问题,本文提出了使用卷积块注意力模块和基于视觉Transformer(vision transformer, ViT)的压缩-激励(squeeze-and-excitation, SE)注意力运算(SE-ViT)来改进单一前馈信息传递的U-Net模型,以优化EIT的特征表达和重建精度。通过在U-Net的每个卷积层后引入卷积块注意力,不仅提升了特征图的空间和通道注意力,而且在下采样过程中有效地保留了对重建任务有益的关键特征。同时,SE-ViT模块的加入优化了传统ViT的参数效率,强化了全局信息特征的捕捉能力,并通过简化的跳跃连接实现了特征的有效传递。在编码器与解码器之间采用了多头交叉注意力模块(multi-head cross-attention module, MHCA),实现了多尺度特征的融合,通过全局相似分数调整模型参数,进一步提高了重建图像的质量。MAT-UNet模型显著提升了EIT在表达边界特征和多相电导率分布方面的表现,展现了优于传统单一卷积网络结构的重建性能,在具有复杂的非线性和病态性的软场成像领域验证了该方法的鲁棒性和准确性。

## 1 电阻抗层析成像原理

### 1.1 测量原理与数学模型

对于本文研究的二维观测区域,在边界  $\partial\Omega$  均匀添加  $l$  个电极(本文测量阵列采用十六电极)记作  $\{e_l\}_{l=1}^L$ ,选择两个相邻电极施加一定频率的交变电流  $I_l$ ,作为初始激励注入观测区域  $\Omega$  内,产生感应电势  $u(\mathbf{r})$ ,测量剩余传感器的响应电压,共测量  $L \times (L - 3)$  个有效电压(共计 208 个数据点),表征该截面完整的一帧数据,其原理如图 1 所示。

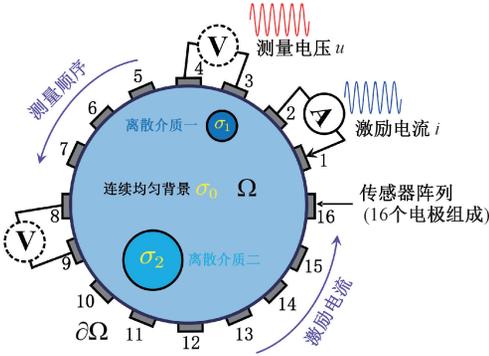


图 1 EIT 测量的仿真模型

Fig. 1 Simulation model for EIT measurement

在 EIT 的正问题中,使用的是全电极模型(complete electrode model, CEM)建模<sup>[21]</sup>,其中电导率拉普拉斯方程:

$$\nabla \cdot (\sigma(\mathbf{r}) \nabla u(\mathbf{r})) = 0, \mathbf{r} \in \Omega \quad (1)$$

边界条件:

$$u(\mathbf{r}) + z_l \sigma(\mathbf{r}) \frac{\partial u(\mathbf{r})}{\partial n} = U_l, \mathbf{r} \in e_l, l = 1, 2, \dots, L, \quad (2)$$

$$\sigma(\mathbf{r}) \frac{\partial u(\mathbf{r})}{\partial n} = 0, \mathbf{r} \in \partial\Omega \setminus \bigcup_{l=1}^L e_l \quad (3)$$

$$\int_{e_l} \sigma(\mathbf{r}) \frac{\partial u(\mathbf{r})}{\partial n} dS = I_l, l = 1, 2, \dots, L, \quad (4)$$

为确保解的存在性和唯一性,在 CEM 中加入了电流守恒和电压守恒方程

$$\sum_{l=1}^L I_l = 0, \sum_{l=1}^L U_l = 0 \quad (5)$$

式中:观测区域记为  $\Omega$ ,  $\partial\Omega$  为对应边界,  $n$  为边界  $\partial\Omega$  的外法向矢量,  $\mathbf{r}$  为空间位置,  $\sigma(\mathbf{r})$  为内部电导率分布,  $u(\mathbf{r})$  是感应区域内的电势,  $z_l$  为第  $l$  个电极的接触阻抗,  $I_l, U_l$  是第  $l$  个电极的注入电流和边界测量电压。式(1)~(5)通常使用有限元方法(finite element method, FEM)求解,考虑到加性测量高斯噪声  $\varepsilon$ , 观测

模型表示为:

$$\mathbf{V} = U(\boldsymbol{\sigma}) + \varepsilon \quad (6)$$

式中:  $\mathbf{V} \in \mathbb{R}^{L \times (L-3)}$  为标准测量电压向量,  $\boldsymbol{\sigma} \in \mathbb{R}^{m \times 1}$  为场域内的离散电导率向量,  $U$  为正向问题算子(forward operator)。

### 1.2 EIT 逆问题

本文使用的是差分方法成像,使用不同时刻的测量电压和参考状态下的电位差重建电导率变化的图像。差分成像可以更加准确地捕捉电导率分布的微小变化,有助于提高图像的空间分辨率和对比度。由于 EIT 逆问题存在非线性和病态性,通常采用一阶近似的方法将其转化为一个优化问题。EIT 逆问题的优化模型可以表示为:

$$\min_{\sigma(\mathbf{r})} \mathcal{L}(\mathbf{V}^{meas}, \mathbf{V}^{cal}(\boldsymbol{\sigma})) + \alpha R(\boldsymbol{\sigma}) \quad (7)$$

式中:  $\mathcal{L}$  为损失函数(loss function),它衡量了实测电压  $\mathbf{V}^{meas}$  与计算电压  $\mathbf{V}^{cal}(\boldsymbol{\sigma})$  之间的差异,  $R(\boldsymbol{\sigma})$  为正则化项,用于引入先验知识,如电导率的平滑性或解的稀疏性,减小问题的不适定性,参数  $\alpha$  用于平衡数据拟合与正则化的影响。

为了增强逆问题求解的稳定性和鲁棒性,在研究中使用了卷积神经网络或 Transformer 的模型,学习从边界数据到电导率分布的映射,基于深度学习的 EIT 逆问题数学模型表述为:

$$f_{EIT} = \underset{\mathbf{R}, \theta}{\operatorname{argmin}} \sum_{n=1}^N \mathcal{L}(\mathbf{x}_n^{\text{Recon}}, \mathbf{y}_n^{\text{Label}}) + R(\boldsymbol{\theta}) \quad (8)$$

式中:  $\{\mathbf{x}_n^{\text{Recon}}, \mathbf{y}_n^{\text{Label}}\}_{n=1}^N$  为训练样本,  $\underset{\mathbf{R}, \theta}{\operatorname{argmin}}$  为训练网络模型,  $R(\boldsymbol{\theta})$  为正则项。

## 2 MAT-UNet 模型

### 2.1 方法概述

MAT-UNet 的结构如图 2 所示,利用一步正则化重建方法,将边界电压转化为描述电导率分布的图像矩阵,作为编码器的输入。该网络模型以 U-Net 为骨干网络,主要由改进的编码器、跳跃连接和解码器 3 个部分构成。在编码阶段,通过 4 步下采样过程配合 5 个卷积块注意力模块(convolutional block attention module, CBAM)来提炼关键特征。CBAM 模块的设计,包含通道注意力(channel attention mechanism, CAM)和空间注意力(spatial attention mechanism, SAM)子模块,专门用于关注和强调图像的重要特征区域,有效地增强了网络在多种特征尺度上的表征能力;相比基于特征拼接或加性求和的传递方式, MAT-UNet 设计了一种基于全局特征表达的跳跃连接(SE-ViT),弥补了卷积运算中,仅对于局部特征表达的不足;在解码阶段, MAT-UNet 采用了 MHCA,这

些模块结合上采样操作,负责将编码阶段压缩后的粗略特征图逐步重建为细致的多尺度特征图。MHCA 模块利用编码器传递的跳跃连接特征以及解码器自身采样的特征,从而在局部和全局层面上实现了对重要特征的精确恢复。跳跃连接不仅保持了在网络深层中的特征完整性,而且通过 SE-ViT 的优化设计,结合了模型的全局和局部特征捕捉能力,对于准确重建尖锐边界和复杂多相电导率参数至关重要。

### 2.2 卷积块注意力模块 (CBAM)

编码器部分(即 MAT-UNet 的特征压缩路径,如图 2

所示)将待重建图像的分辨率从  $256 \times 256$  被全局压缩至  $16 \times 16$ ,同时特征图的通道数目从 3 扩增至 64。压缩路径由 CBAM 组成,以优化特征提取和映射过程,如图 3 所示。该模块主要通过 CAM 和 SAM 两个环节来提炼关键特征。CAM 侧重于评估图像中每个通道的显著性,捕获重要特征,而 SAM 则专注于图像的特定空间区域,突出有价值的局部细节。这种策略使网络能够自适应地精炼特征,并在通道和空间两个层面上进行优化。

在卷积块注意力模块中,通道注意力模块利用全局最大池化和全局平均池化策略,产生两种不同的  $1 \times 1 \times C$

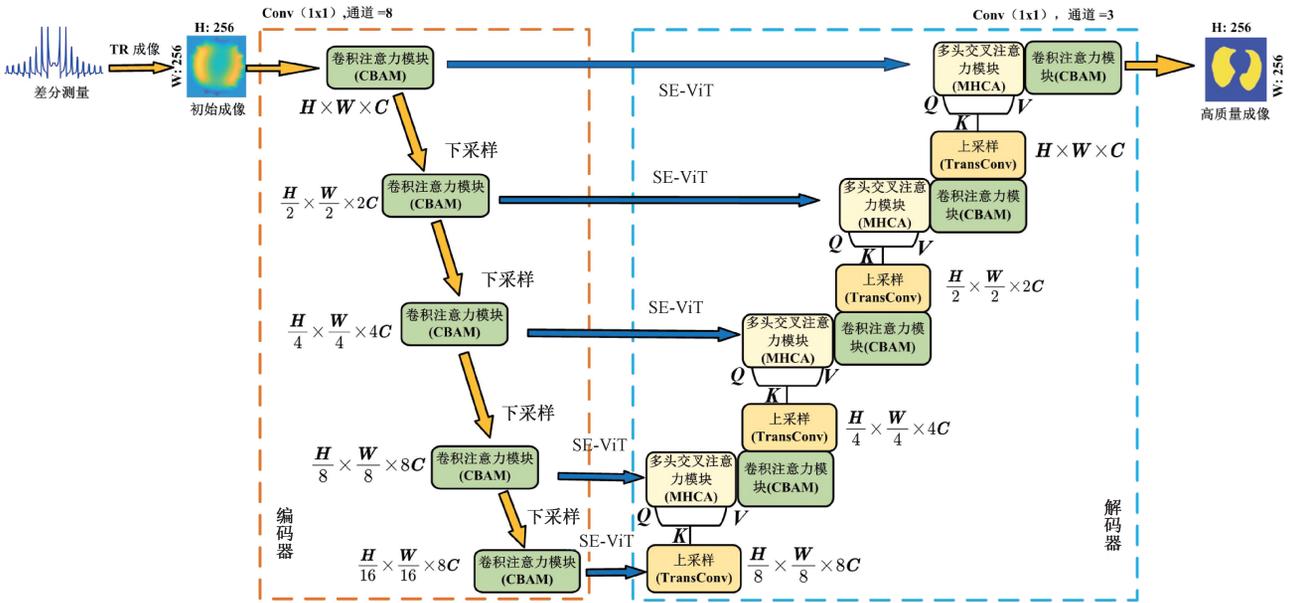


图 2 MAT-UNet 模型架构图

Fig. 2 MAT-UNet model architecture diagram

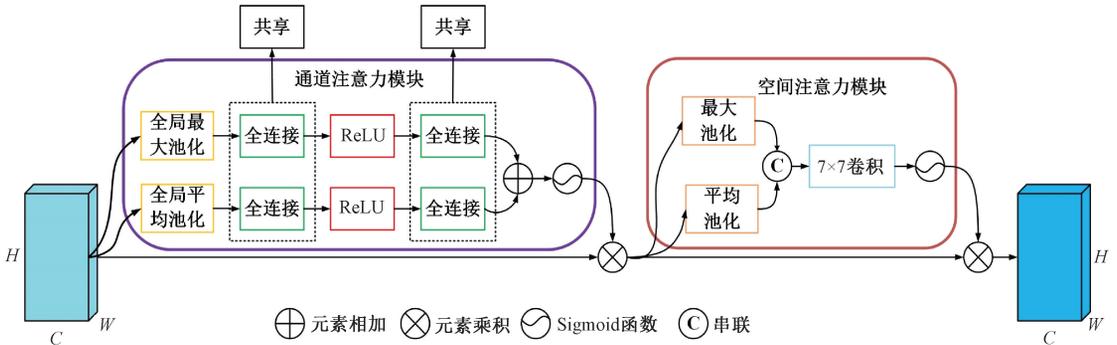


图 3 卷积快注意力模块 CBAM

Fig. 3 Convolutional block attention module (CBAM)

的最大池化和平均池化特征图。然后这些信息经由一个共享的多层感知器 (multilayer perceptron, MLP) 处理,以学习每个通道的重要权重,用来增强关键通道的信息。而空间注意力模块则接收通道注意力处理后的特征图,

通过最大池化和平均池化生成新的特征图,并在通道上合并成  $H \times W \times 2C$  的特征图,计算过程如下:

$$N_c(V) = \delta(MLP(AvgPool(V))) + \delta(MLP(MaxPool(V))) \quad (9)$$

$$N_s(V) = \delta(g^{7 \times 7} [AvgPool(V); Max(Pool(V))]) \quad (10)$$

式中:  $N_c(V)$  为通道注意力模块输出,  $N_s(V)$  为空间注意力模块输出,  $V$  为特征图,  $\delta$  为激活函数,  $[*; *]$  表示池化后的特征图在通道维度进行拼接的操作,  $g^{7 \times 7}$  是将这两个特征图拼接后通过一个  $7 \times 7$  的卷积操作进行处理。最终得到的特征图在空间上被加强, 从而在模型中实现对重要特征的更有效聚焦。在 U-Net 的编码器中嵌入 CBAM, 不仅提高了模型对 EIT 数据的全局和局部特征捕捉能力, 还增强了对边缘和内部结构的重建精度。这种方法使得模型在处理 EIT 图像时, 能够更有效地解决复杂电导率分布的问题, 进而提高成像的准确性和分辨率。

### 2.3 SE-ViT 跳跃连接

在 MAT-UNet 中, SE-注意力机制通过 ViT 结构的优化<sup>[22]</sup>, 使得新的注意力模块拥有更强的建模能力, 能够捕捉到更为复杂的通道间关系。由于该机制能够为卷积滤波器所对应的每一个重构信息分配适当的注意力权重, 所以有效减少了图像重构过程中的误差, 其结构如图 4 所示。

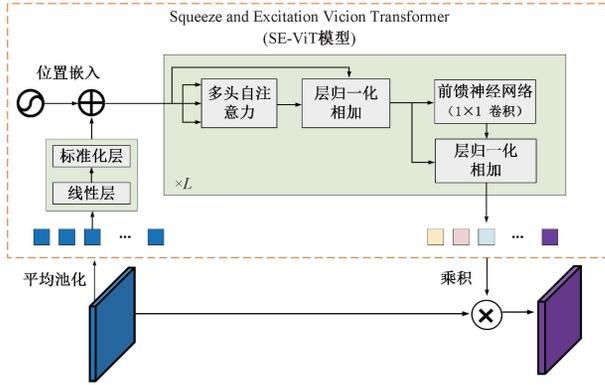


图4 SE-ViT 跳跃连接模块

Fig. 4 SE-ViT skip connection module

SE-ViT 通过全局平均池化操作将输入的特征图转换为一系列像素块, 然后利用可训练的线性投影将这些量化的图像块映射到一个潜在的高维嵌入空间。为了能够编码图像块的空间信息, SE-ViT 学习并加入了特定的位置嵌入到每一个图像块嵌入中, 以此来保存空间位置信息, 位置嵌入方式为:

$$U_0 = [x_p^1 P, x_p^2 P, \dots, x_p^N P] + P_{position} \quad (11)$$

式中:  $P \in \mathbb{R}^{C_k \times D}$  为图像块嵌入的投影,  $P_{position} \in \mathbb{R}^{N \times D}$  表示位置嵌入。这种方式不仅保留了空间信息, 而且丰富了模型对全局上下文的理解, 使得模型在进行电导率图像重构时, 能够更加精确地恢复图像细节和电导率分布。Transformer 的设计目标是基于通道上下文信息对每个图

像块进行编码, 以捕获所有像素块之间的交互作用。这一机制通过多头自注意力 (multi-head self-attention, MHSA) 和前馈网络来实现。具体的数学表达式如下:

$$u'_l = \text{MHSA}(\text{LN}(u_{l-1})) + u_{l-1} \quad (12)$$

$$u_l = \text{Conv}_{1 \times 1}(\text{LN}(u'_l)) + u'_l \quad (13)$$

其中, LN 表示层归一化运算,  $u'_l$  表示第  $l-1$  层输出与自注意力权重累加和,  $u_l$  表示第  $l$  层的 Transformer 输出。经过  $L$  个 Transformer 模块的处理之后, 得到不同通道的参数权重, 利用 SE 注意力机制, 把这些权重赋予到输入特征图的每一个通道, 在特征融合和信息流动过程中实现更为精细的特征表示, 其运算方式为:

$$\mathbf{X}_{out} \in \mathbb{R}^{\frac{H}{2} \times \frac{W}{2} \times C_i} = \mathbf{X}_{in} \otimes U_L \quad (14)$$

### 2.4 MHCA

MHCA 通过抑制不必要的信息和突出关键区域来增强特征图中的目标区域。其结构如图 5 所示。

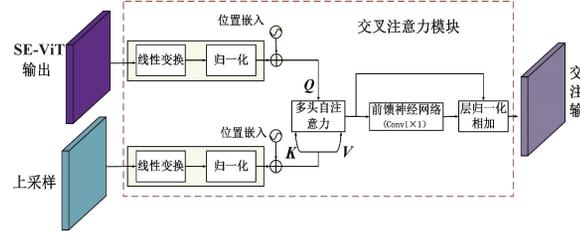


图5 交叉多尺度信息融合连接

Fig. 5 Cross-scale information fusion connection

首先对 SE-ViT 和上采样 (使用  $3 \times 3$  的转置卷积实现) 的特征图进行分块和位置嵌入, 如式 (11) 所示, 从跳跃连接的输出特征图中获得查询向量  $Q_i = X_{SE-ViT} \cdot W^{Q_i}$ , 从上采样的特征图中获得键向量和值向量分别为  $K_i = X_{Up} \cdot W^{K_i}$ ,  $V_i = X_{Up} \cdot W^{V_i}$ , 其中,  $W^{K_i} \in \mathbb{R}^{d_i \times d_k}$  和  $W^{V_i} \in \mathbb{R}^{d_i \times d_v}$  是通过学习获得的权重矩阵, 而且  $X_{SE-ViT}$  和  $X_{Up}$  具有相同的维度。注意力模块的输出特征  $G_i \in \mathbb{R}^{c \times d_v}$  可以由式 (15) 计算得到:

$$G_i = \text{softmax}\left(\frac{Q_i \cdot K_i^T}{\sqrt{d_q}}\right) \cdot V \quad (15)$$

MHCA 将获得的注意力特征  $G_i$  经过  $1 \times 1$  卷积的 FFN 和特征连接后, 获得融合了编码模块和上采样恢复的多尺度特征, 并再次使用残差卷积模块对局部特征进行构建, 作为下一个 MHCA 的输入。

### 2.5 损失函数

MAT-UNet 的损失函数定义为:

$$Loss = \mathcal{L}_{MSE} + \lambda \cdot \mathcal{R}(\theta) = \sum_{(\hat{x}, y) \in \mathcal{D}} \|\hat{x} - f_{MAT-UNet}(y | \theta)\|_2^2 + \lambda \cdot \|\theta\|_2^2 \quad (16)$$

式中:  $\hat{x}$  为待重建的目标电导率图像,  $y$  为输入的初始化

电导率分布,  $\mathcal{D}$ 为给定的训练样本,  $\theta$  为 DHU-Net 待训练的参数。损失函数的第 1 项 ( $\mathcal{L}_{MSE}$ ) 用于确保模型的输出尽可能与标签分布地接近, 第 2 项  $\mathcal{R}(\cdot)$  为损失函数的正则项, 用于避免训练过程中出现过拟合的问题, 对应的  $\lambda$  为正则项参数, 在本文中设定为 0.01。通过持续监控损失函数的变化, 调整网络参数得到最优的模型参数  $\theta^*$ , 使得网络可以重建与目标最为一致的结果  $\mathbf{x}^*$ 。

### 3 实验设置

#### 3.1 数据集制备

为了提高 EIT 的泛化能力, 并适应多样的成像任务, 研究中采用了 2 种仿真数据集:

1) 多相介质分布模拟数据: 仿真平台使用 COMSOL Multiphysics Livelink Matlab。设置测量区域是一个直径 0.2 m 的圆形来模拟多相流体分布管道, 在测量区域边界均匀放置 16 个钛电极形成传感器阵列。其中自来水作为空场中的均匀分布介质, 设置电导率为 0.6 S/m, 非均匀场使用电导率为 0.06~1 S/m 的圆形介质模拟不同条件下的介质分布, 不同的介质使用随机的尺寸(几何尺寸范围是 0.02~0.08 m)和位置进行模拟, 并且介质之间没有重合。本研究共模拟了 42 830 组数据样本, 其中 85% 的样本用作训练数据集, 10% 的样本用作验证数据集, 剩余 5% 的样本用作测试数据集。为提升图像重建的空间分辨率, 将逆问题的网格细化到 256×256(网格尺寸为  $7.8 \times 10^{-4}$  m)。

2) 肺部仿真数据集的构建: 肺部仿真模型的构建包含 3 个关键步骤: 选取适当的 CT 图像、确立肺部区域以及模型的进一步扩充。本文选用的肺部 CT 图像来自联盟影像集合(LIDC-IDRI)的胸部薄层 CT 图库<sup>[23]</sup>。从库中随机挑选了 80 名患者的图像资料, 并将这些资料分为 2 批, 一批 70 名患者的图像用于网络的训练, 另一批 10 名患者的图像留作测试用例。每位患者的图像选取中心位于乳头下方第 5 肋骨附近, 进一步从中心图像上下各取 8 cm 处获取 5 张 CT 切片, 以此构成 350 张训练图像与 50 张测试图像。在模拟健康与病变肺组织时, 电导率的设定考虑到了不同组织类型, 将皮下组织电导率设为 0.037 S/m, 正常肺组织为 0.1 S/m, 心脏为 0.25 S/m。为了模拟肺部病变, 模拟肺部疾病的电导率范围则设定为 0.1~0.142 S/m。为了提升数据集的质量并模拟实际的测量噪声, 研究过程中在正问题求解得到的电压数据中加入随机噪声, 以此方法扩增数据集, 最终形成含有 9 100 个训练样本与 1 000 个测试样本的综合肺部仿真数据集, 用作深度学习模型训练与评估。

数据集集中的每个样本都包含了差分电压数据和对应的电导率分布标签, 采用线性归一化将电导率值限制在

0~1 范围内, 降低训练过程的复杂度, 增加模型训练的稳定性。标签向量和预测的介质分布是损失函数的自变量, 用于监督和约束深度学习的训练过程。

#### 3.2 模型训练

在 MAT-UNet 模型的训练流程中, 首先, 将由传感器阵列收集的边界电压数据输入模型, 并进行前向传播以预测电导率分布。模型通过损失函数评估预测介质分布与实际分布的差异, 并通过梯度下降法反向传播更新权重, 以此迭代优化模型参数。该过程持续进行, 直至模型性能达到预定的收敛标准。

在本次实验中, 利用 Pytorch 框架进行模型训练, 所有计算过程均在装配有 Intel Core i7-9700K CPU, Nvidia RTX 3060 的计算机上执行, 整个训练耗时约 20 小时。具体的训练参数配置如下: 模型中 Transformer 模块的层数分别为 4, 8, 8, 8 层, SE-ViT 和 MHCA 设置注意力头的数量则设定为 6 个。在模型结构中, 下采样部分采用  $2 \times 2$  的平均池化, 而上采样部分则采用  $3 \times 3$  的转置卷积。使用小批量随机梯度下降法对 MAT-UNet 进行训练, 每 128 个样本进行一次训练迭代, 总迭代次数设定为 500 次。初始学习率设置为 0.1, 并随着训练进度参数每更新一次, 按照式(17)的方式进行参数调整:

$$\eta_t = \eta_{t-1} \cdot \left(1 - \frac{t}{T}\right)^{0.9} \quad (17)$$

式中:  $t$  表示当前训练的步数,  $\eta_t, \eta_{t-1}$  表示第  $t$  次、第  $t-1$  次的学习率参数,  $T$  即为总训练步数。模型训练前, 权重参数随机设置为均值为 0, 标准差为 0.01 的高斯分布随机数。

### 4 实验结果分析

在本节中, 进行了仿真和实验成像, 以证实 MAT-UNet 模型的有效性。为了对模型重建图像与真实分布进行定量比较, 研究中采用了均方根误差 (root mean square error, RMSE)、结构相似性指数 (structural similarity, SSIM) 量化评价指标, 评估图像重建的精度和质量, 评价指标定义如下:

$$\mathbf{x}_{REC\_normalized} = \frac{\mathbf{x}_{REC} - \mathbf{x}_{REC\_min}}{\mathbf{x}_{REC\_max} - \mathbf{x}_{REC\_min}} \quad (18)$$

$$\mathbf{x}_{label\_normalized} = \frac{\mathbf{x}_{label} - \mathbf{x}_{label\_min}}{\mathbf{x}_{label\_max} - \mathbf{x}_{label\_min}} \quad (19)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{x}_{REC\_normalized} - \mathbf{x}_{label\_normalized})^2} \quad (20)$$

$$SSIM = \frac{(2\mu_{Rec}\mu_{label} + C_1)(2\Sigma_{Rec, label} + C_2)}{(\mu_{Rec}^2 + \mu_{label}^2 + C_1)(\Sigma_{Rec}^2 + \Sigma_{label}^2 + C_2)} \quad (21)$$

其中,  $\mathbf{x}_{REC\_min}$  和  $\mathbf{x}_{REC\_max}$  分别是重建电导率值的最小值和最大值,  $\mathbf{x}_{label\_min}$  和  $\mathbf{x}_{label\_max}$  分别是真实电导率值的最小值和最大值.  $Rec$  表示重建图像,  $label$  表示真实分布图像,  $N$  为图像的总像素数,  $\mu$  和  $\Sigma$  分别表示均值和协方差矩阵.

### 4.1 鲁棒性测试

如图 6 所示, 在仿真模拟实验数据中, 为了评估 MAT-UNet 对噪声的抵抗力, 本文引入了不同信噪比水平 (20~60 dB) 的高斯白噪声. 同时, 本文也选取了 VD-Net<sup>[24]</sup>、VA-Net<sup>[25]</sup>、以及 TransUNet<sup>[15]</sup> 作为对照组, 对比不同的信噪比水平下, 这 4 种深度学习成像算法在量化指标 RMSE 与 SSIM 上的差异.

MAT-UNet 有更好的噪声鲁棒性. 当 SNR 水平低于

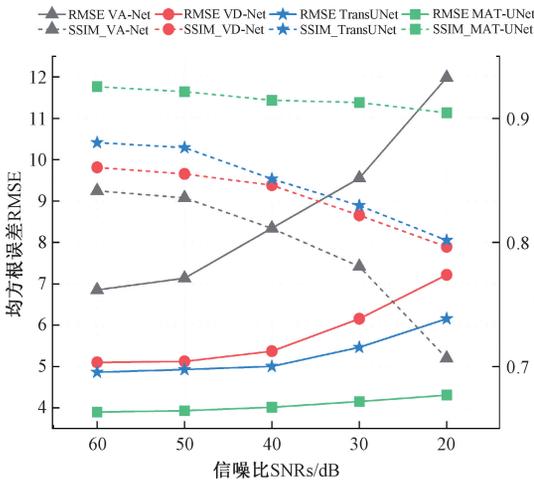


图 6 不同信噪比水平下的成像结果量化指标

Fig. 6 Quantitative imaging metrics at different signal-to-noise

40 dB 时, RMSE 能够维持在 5.00 以下, SSIM 高于 0.88, 这表明在复杂边界的形状重建任务中, MAT-UNet 模型能够获得比较精确的成像结果. 与此相反, 另外 3 种比较成像模型的结果显示, 在扰动增强的情况下, 成像质量有显著的下降, 尤其是在多相电导率介质场景及形状重建任务中, 成像效果明显变差. 此次实验结果进一步印证了基于多尺度 Transformer 的成像模型在抗噪声方面的性能更加突出.

### 4.2 多相电导率重建测试

在验证 MAT-UNet 模型的广泛适用性方面, 本文构建了多相电导率分布模型进行验证. 如图 7 所示, 实验中设定了一个背景电导率为 0.6 S/m 的模拟场景, 并在其中放置了具有不同电导率的物体以模拟多样的介质分布情况. 通过 MAT-UNet 模型进行重建后得到的图像结果清晰, 展示了其成像能力. 与此同时, 实验也采用了 RMSE 和 SSIM 这两项量化指标来评估模型的性能, 具体数值如表 1 所示.

表 1 分析了不同成像算法在 4 种不同测试场景下的性能对比, 通过 RMSE 和 SSIM 两个指标来评估. 其中, 基于灵敏度矩阵的成像方法可以重建出与背景电导率相差比较大的介质, 但是伪影干扰非常明显. 基于深度学习的成像方法能够更好的表示出内含物的形状与分布. 然而, 仅简单的卷积特征融合方法在描绘全局特征时显得不足, 在执行多电导率重建任务时不能精确地恢复出电导率分布的细节. 本文提出的 MAT-UNet 模型在所有测试场景下均显示出色的性能, 尤其在模型三中, 比现有的 ResAttU-Net 方法 RMSE 降低了 8.8%, SSIM 提升了 2.5%.

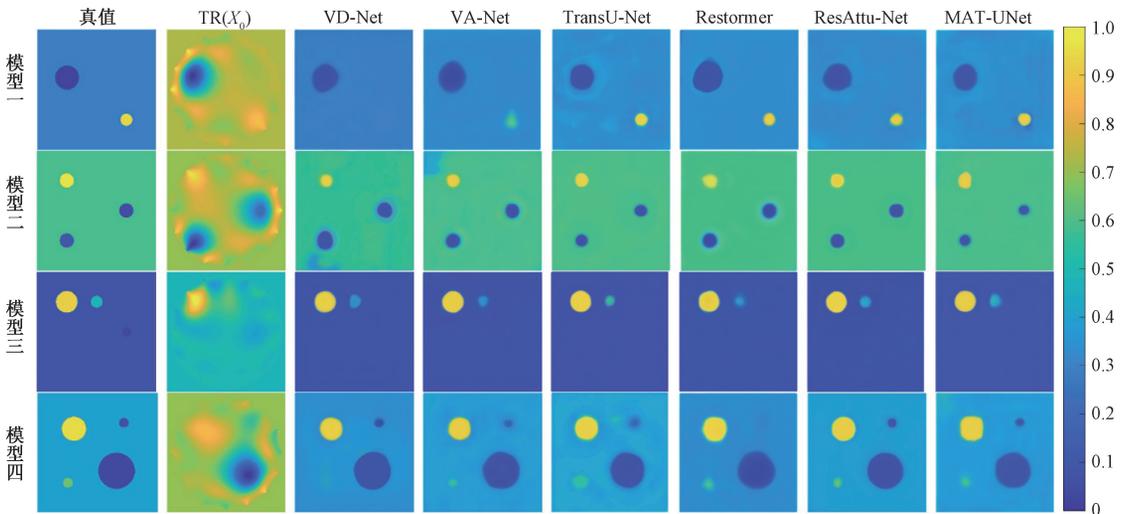


图 7 多相电导率重建实验结果

Fig. 7 Experimental results of multiphase conductivity reconstruction

表 1 多相电导率介质分布重建指标 (RMSE 和 SSIM)

Table 1 Reconstruction metrics for multiphase conductivity media distribution (RMSE and SSIM)

模型指标	模型一		模型二		模型三		模型四	
	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM	RMSE	SSIM
TR	14.194 5	0.654 3	16.908 4	0.521 6	17.827 1	0.628 6	15.109 8	0.643 0
VD-Net	6.924 9	0.716 5	7.568 0	0.684 4	7.489 4	0.718 0	7.241 5	0.713 1
VA-Net	4.254 7	0.768 9	7.507 6	0.725 9	5.628 9	0.751 6	4.986 1	0.742 8
TransUNet	4.213 5	0.864 0	6.465 2	0.851 7	4.916 8	0.854 6	4.712 4	0.857 3
Restormer	3.991 2	0.870 2	6.207 6	0.865 1	4.768 9	0.863 4	4.501 7	0.862 9
ResAttU-Net	3.979 8	0.879 7	5.973 4	0.869 0	4.521 7	0.868 7	4.259 0	0.869 4
<u>MAT-UNet</u>	<u>3.942 1</u>	<u>0.891 6</u>	<u>5.608 6</u>	<u>0.881 6</u>	<u>4.121 9</u>	<u>0.890 5</u>	<u>4.021 9</u>	<u>0.882 5</u>

4.3 肺部实验结果重建实验

实验采用了一个配备 16 个电极的 EIT 数据获取系统来收集数据。这些参数的设定与仿真环境相一致,其中水槽直径 0.19 m,以 0.06 S/m 电导率的水溶液作为均一背景(即基准测量状态),并用有机玻璃棒来模拟非均一介质(内含物电导率假定为  $10^{-12}$  S/m)。实验设备如图 8 所示,数据采集系统具有 70 dB 的信噪比和 14 位的比特分辨率,以及 4.5 mA 和 100 kHz 的激励电流。该系统由传感器阵列、FPGA 控制芯片、数字/模拟转换器、电压控制电流源 (voltage controlled current source, VCCS)、程序控制的可变增益放大器 (programmable gain amplifiers, PGA) 以及计算机组成。实验结果如图 9 所示。



图 8 实验设备

Fig. 8 Experimental equipment

在实验中,模拟了多相电导率的肺部场景,旨在测试 MAT-UNet 在实际应用中的重建效果。作为参照,本实验

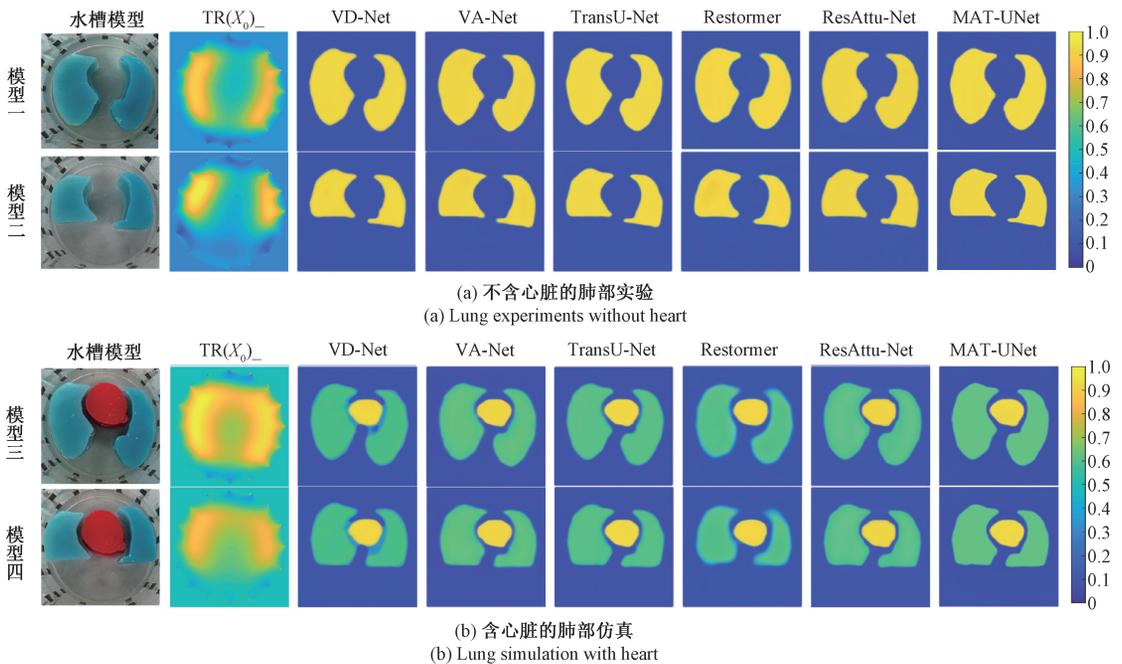


图 9 实验测量数据图像重建结果(所有结果均为归一化处理后的重建图像)

Fig. 9 Image reconstruction results from experimental measurement data (All the results are the reconstructed images after normalization processing)

还对比了 TR、VD-Net、VA-Net、TransUNet、Restormer<sup>[16]</sup>和 ResAttU-Net<sup>[11]</sup>等方法。图 9 (a) 中模型一和模型二展示了肺部形状重建场景;在图 9 (b) 中,模型三和四展示了包含肺部形状和心脏的多相介质重建场景。可以看出 MAT-UNet 在准确重建复杂的形状特征和多相电导率信息方面表现出色。该模型能够清晰地描绘肺部和心脏的电导率参数及其边界特征,几乎不受伪影的影响。相比之下,其他重建算法在处理形状信息和电导率变化复杂区域时,常出现伪影,或者将复杂图像边缘过度平滑,从而无法准确反映介质的分布特性。

为了进一步分析观察结果,对每种对比模型分别取 50 组数据进行测试。在相应测试数据集上得到的 RMSE 和 SSIM 的统计数据,如图 10 和 11 所示。尤其是,传统重建方法(tikhonov regularization, TR)在这两个指标上显示出较低的平均值和较大的方差,表明其成像结果易受噪声和场景复杂度的影响。

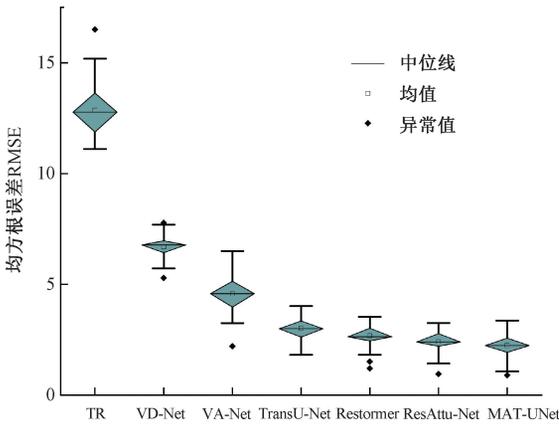


图 10 均方根误差 RMSE  
Fig. 10 Root mean square error

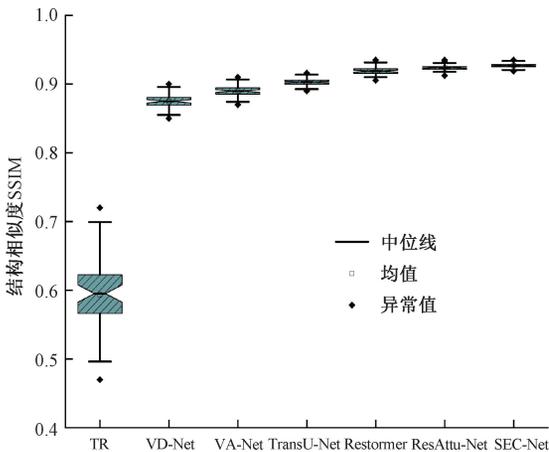


图 11 结构相似度 SSIM  
Fig. 11 Structural similarity index

尽管稠密连接、注意力跳跃连接和 TransUNet 等方法有所改善,但它们在平均 RMSE 和 SSIM 方面仍然表现出较大的方差,说明这些模型的泛化能力较弱,在复杂场景下难以达到理想效果。相比之下,本研究提出的 MAT-UNet 模型在 RMSE 和 SSIM 上均表现出优异的数值指标,且方差分布相对集中,这表明该方法具有较强的泛化能力和鲁棒性。

4.4 TR 参数分析

在研究中,为了分析正则化参数  $\alpha$  对 MAT-UNet 的影响效果,我们采用两种电导率分布模型,选取不同的  $\alpha$  参数值(0.1~1.0)进行 TR 重建实验,结果如图 12 所示。将 TR 重建结果为模型的输入,以验证不同正则化参数对模型成像效果的影响,结果如图 13 所示。实验结果表明,在不同的 TR 参数下,即使在 TR 方法重建效果不佳时,本文提出的 MAT-UNet 仍然可以获得较为满意的成像结果。如图 14 所示,这一结论的量化指标中得到了一致体现。

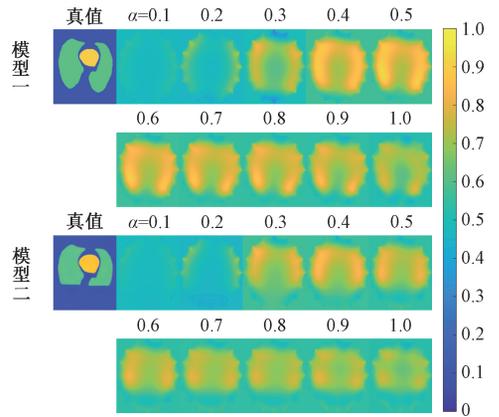


图 12 TR 图像重建效果

Fig. 12 Reconstruction results of tikhonov regularization

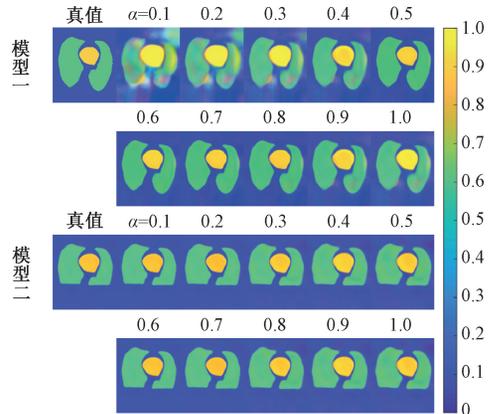


图 13 MAT-UNet 模型成像效果

Fig. 13 Performance of the MAT-UNet model

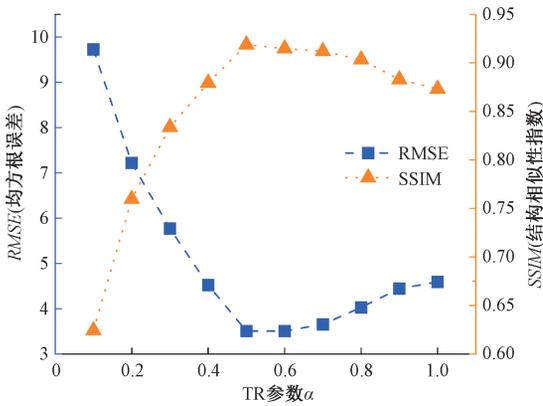


图 14 MAT-UNet 成像的 RMSE 和 SSIM

Fig. 14 RMSE and SSIM of MAT-UNet imaging

通过图 14 量化指标表明,  $\alpha = 0.5$  时 SSIM 最优, RMSE 最小, 因此在实验中选取正则化参数为 0.5。

#### 4.5 消融实验

MAT-UNet 方法是以 U-Net 网络为基础, 嵌入了 CBAM, SE 网络, 以及 MHCA 的 EIT 深度学习模型。为了检验不同机制在 U-Net 架构中的作用, 设计了消融实验, 评估不同组件或者特性对最终模型的性能贡献。其结果如表 2 所示, 加入的卷积注意力模块和 SE 连接和多头交叉注意力模块均可以有效的提高模型处理图像的性能。在 U-Net 模型中加入 CBAM, SE 和 MHCA 可以显著提高 EIT 的性能, 根据消融实验结果表明比基础的 U-Net 网络, RMSE 降低了 67.05%, SSIM 增加了 35.82%。

表 2 消融实验在测试集的表现

Table 2 Performance of ablation experiments on the testing set

模型				量化指标	
U-Net	SE	CBAM	MHCA	RMSE	SSIM
√				12.227 2	0.645 4
√		√		9.398 0	0.743 3
√	√	√		6.030 5	0.825 0
√	√	√	√	4.028 2	0.876 6

#### 4.6 重建时间分析

为了分析不同算法之间的训练时间和图像重建时间, 测试实验是重建一幅空间分辨率为  $256 \times 256$  的图像, 使用的设备是 Intel Core i7-9700K CPU, Nvidia RTX 3060 的显卡, 重建时间取 50 幅测试集图像重建时间的均值, 如表 3 所示。在上表中可以看到, 本文提到的 MAT-UNet 模型训练时间比较长, 但这一时间在同类高精度模型如 TransU-Net 和 Restormer 中是相对较短的, 显著提高了开

发效率, 并加快了模型的迭代速度。而且 MAT-UNet 方法在图像重建时间彰显了它的优势, 使其在实际应用中显示出了它的高效率和适用性。

表 3 各算法重建时间

Table 3 Reconstruction Time of Various Algorithms

算法	训练时间/h	重建时间/s
TR	-	0.003 94
VD-Net	12	0.065 01
VA-Net	10	0.043 61
TransU-Net	48	0.160 01
Restormer	50	0.155 99
ResAttU-Net	26	0.125 00
MAT-UNet	28	0.104 95

## 5 结 论

本文研究了一种创新的多尺度 Transformer 为基础的 EIT 深度成像技术, 称为 MAT-UNet。该方法采用端到端的训练策略, 显著提升了成像的空间分辨率, 并能更精确地描绘内含物的形态特征。MAT-UNet 结合了 SE-ViT 架构和卷积操作, 分别用于捕捉全局电学参数和局部形状信息。此外, 通过多头交叉注意力模块, 该方法能够有效地融合多尺度特征, 以恢复高分辨率的图像。实验结果证明, 这一深度成像网络能够更准确地重构被测区域内的介质分布。然而, 该方法在重建电导率对比度较低的介质方面表现不足, 未来的工作将探讨融合物理先验信息, 以提升模型在低对比度成像任务中的性能。

### 参考文献

[ 1 ] BRAZEY B, HADDAB Y, ZEMITI N. Robust imaging using electrical impedance tomography: Review of current tools [ J ]. Proceedings of the Royal Society A, 2022, 478(2258): 20210713.

[ 2 ] ADLER A, BOYLE A. Electrical impedance tomography: Tissue properties to image measures [ J ]. IEEE Transactions on Biomedical Engineering, 2017, 64(11): 2494-2504.

[ 3 ] GAO X, CAO ZH, TIAN Y, et al. Inverse radon method based on electrical field lines for dual-modality electrical tomography [ J ]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(10): 8250-8260.

[ 4 ] SHI Y Y, ZHANG X, WANG M, et al. An adaptive non-convex hybrid total variation regularization method for

- image reconstruction in electrical impedance tomography[J]. *Flow Measurement and Instrumentation*, 2021, 79: 101937-101945.
- [ 5 ] SUO P, SUN J T, ZHANG X K, et al. Adaptive group-based sparse representation for image reconstruction in electrical capacitance tomography[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 4504509.
- [ 6 ] SANTOS T B R, NAKANISHI R M, CAMARGO E D L B D, et al. Improved resolution of D-bar images of ventilation using a schur complement property and an anatomical atlas[J]. *Medical physics*, 2022, 49(7): 4653-4670.
- [ 7 ] GU D P, LIU D, SMYL D, et al. Supershape recovery from electrical impedance tomography data [ J ]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 4503711.
- [ 8 ] WU Y, ZHOU T, CHEN B, et al. Bayesian shape reconstruction using B-spline level set in electrical impedance tomography [ J ]. *IEEE Sensors Journal*, 2022, 22(19): 19010-19019.
- [ 9 ] 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.
- [ 10 ] ZHANG T, TIAN X, LIU X C, et al. Advances of deep learning in electrical impedance tomography image reconstruction [ J ]. *Frontiers in Bioengineering and Biotechnology*, 2022, 10: 1019531.
- [ 11 ] 王子辰, 陈晓艳, 王倩, 等. 基于残差自注意力连接的深度电学层析成像方法[J]. *仪器仪表学报*, 2023, 44(5): 288-301.  
WANG Z CH, CHEN X Y, WANG Q, et al. Electrical tomography imaging method based on deep CNN with residual self-attention skip connection [ J ]. *Chinese Journal of Scientific Instrument*, 2023, 44 ( 5 ): 288-301.
- [ 12 ] CEN S, JIN B, SHIN K, et al. Electrical impedance tomography with deep Calderón method[J]. *Journal of Computational Physics*, 2023, 493: 112427.
- [ 13 ] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale [ J ]. *ArXiv preprint*, 2020, arXiv: 11929, 2010.
- [ 14 ] ZHOU H Y, LU CH X, YANG S B, et al. Convnets vs. transformers: Whose visual representations are more transferable? [ C ]. 2021 IEEE/CVF International Conference on Computer Vision Workshops, 2021: 2230-2238.
- [ 15 ] CHEN J, LU Y, YU Q, et al. Transunet: Transformers make strong encoders for medical image segmentation [ J ]. *ArXiv preprint arXiv*: 2102.04306, 2021.
- [ 16 ] 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, 2022: 5728-5739.
- [ 17 ] 冷冰, 冷敏, 常智敏, 等. 基于Transformer结构的深度学习模型用于外周血白细胞检测[J]. *仪器仪表学报*, 2023, 44(5): 113-120.  
LENG B, LENG M, CHANG ZH M, et al. Deep learning model based on Transformer architecture for peripheral blood leukocyte detection [ J ]. *Chinese Journal of Scientific Instrument*, 2023, 44(5): 113-120.
- [ 18 ] LIN AI L, CHEN B ZH, XU J Y, et al. Ds-transunet: Dual swin transformer U-net for medical image segmentation[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 4005615.
- [ 19 ] 熊巍, 熊承义, 高志荣, 等. 通道注意力嵌入的Transformer图像超分辨率重构[J]. *中国图象图形学报*, 2023, 28(12): 3744-3757.  
XIONG W, XIONG CH Y, GAO ZH R, et al. Image super-resolution with channel-attention-embedded [ J ]. *Transformer. Journal of Image and Graphics*, 2023, 28(12): 3744-3757.
- [ 20 ] 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.
- [ 21 ] VAUHKONEN P, VAUHKONEN M, SAVOLAINEN T, et al. Three-dimensional electrical impedance tomography based on the complete electrode model [ J ]. *IEEE Transactions on Biomedical Engineerin*, 1999, 46(9): 1150-1160.
- [ 22 ] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [ 23 ] CLARK K, VENDT B, SMITH K, et al. The cancer imaging archive (TCIA): Maintaining and operating a public information repository [ J ]. *Journal of Digital Imaging*, 2013, 26: 1045-1057.
- [ 24 ] LI F, TAN CH, DONG F. Electrical resistance

tomography image reconstruction with densely connected convolutional neural network [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 4500811.

- [25] WANG Z CH, ZHANG X Y, FU R, et al. Electrical impedance tomography image reconstruction with attention-based deep convolutional neural network [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 5011318.

## 作者简介



王琦(通信作者),2012年于天津大学获得博士学位,现为天津工业大学教授,主要研究方向为过程层析、图像重建和智能信息处理。

E-mail: wangqitju@163.com

**Wang Qi** (Corresponding author) received her Ph. D. degree from Tianjin University in 2012. She is currently a professor at Tiangong University. Her research interests include process tomography, image reconstruction and intelligent information processing.



张涛,2023年于山东理工大学获得学士学位,现为天津工业大学硕士研究生。主要研究方向图像处理、凸优化、张量计算,多模态成像技术等。

E-mail: zhangtao@tiangong.edu.cn

**Zhang Tao** received his B. Sc. degree from Shandong University of Technology in 2023. He is currently a master student in Tiangong University. His main research interests include image processing, convex optimization, tensor computing, and multimodal imaging techniques.