

# 基于多通道交叉卷积 UCTransNet 的双能 CT 基材料分解方法

吴凡<sup>1</sup>, 金潼<sup>1</sup>, 詹郭睿<sup>1</sup>, 解晶晶<sup>1</sup>, 刘进<sup>1,2\*</sup>, 张谊坤<sup>2,3</sup>

<sup>1</sup>安徽工程大学计算机与信息学院, 安徽 芜湖 241000;

<sup>2</sup>计算机网络和信息集成教育部重点实验室(东南大学), 江苏 南京 210096;

<sup>3</sup>东南大学影像科学与技术实验室, 江苏 南京 210096

**摘要** 提出一种基于多通道交叉卷积 UCTransNet (MC-UCTransNet) 的图像域双材料分解方法。该网络以 UCTransNet 为基础架构, 采用通道交叉融合转换器和通道交叉注意模块来提高基材料分解性能, 实现双输入双输出的端到端映射。网络中通道交叉融合模块和通道交叉注意模块可更好地捕捉复杂的通道信号相关性, 以更充分地进行特征提取与融合, 实现基材料生成路径之间的信息交换。为进一步提高模型的拟合性能, 网络训练时采用混合损失及 Sigmoid 函数的归一化方法。实验结果表明, 在骨骼基材料及软组织碘基材料分解任务中, 所提方法能获得优质的基材料图像, 与对比方法相比, 其分解后的基材料图像在准确度及噪声伪影抑制上表现更好。

**关键词** 机器视觉; 双能计算机断层成像; 基材料分解; 多通道交叉卷积; 注意力; 噪声抑制

中图分类号 TP391.7

文献标志码 A

DOI: 10.3788/AOS231715

## 1 引言

目前, 双能计算机断层成像 (DECT)<sup>[1]</sup> 已经被广泛应用于肾结石表征<sup>[2]</sup>、碘定量检查<sup>[3-4]</sup>等临床诊断中, 可以实现基材料分解以及能量选择性成像, 具有重要的临床诊断意义<sup>[5]</sup>。现有的 DECT 基材料分解方法可分为 3 类: 直接重建分解法、基于投影域的分解方法和基于图像域的分解方法<sup>[6]</sup>。直接重建分解方法是直接从两个 X 射线能谱上收集数据来重建基图像, 但该类方法的计算复杂度高, 且分解结果受参数的影响大。基于投影域的分解方法能够较好地补偿投影域材料分解过程中的非线性效应, 并能较好地修正射线束硬化伪影, 但高度依赖于测量投影的空间几何一致性, 这对双源和快速切换 ( $10^3 V_p$  量级, 其中  $V_p$  为峰值电压) DECT 等成像设备来说是一个挑战。基于图像域的分解方法首先从两个测量数据集单独或联合重建出双能量图像, 然后通过重建图像的线性组合形成基图像, 具有灵活性高的优势。当前 DECT 技术中普遍存在的问题是: 受高低能 CT 图像的噪声影响<sup>[7]</sup>, 分解后的基材料图像噪声会进一步放大, 导致分解精度下降, 基材料图像中出现与噪声相关的伪影, 从而影响后续诊断任务。

基于图像域的基材料分解中, 传统的矩阵反演算

法对于输入的重建图像噪声敏感, 会显著降低基材料图像的信噪比<sup>[8-9]</sup>。为此, Zhao 等<sup>[10]</sup>通过将边缘保持滤波器与局部高分辨率约束反投影重建相结合, 间接降低了分解图像的噪声水平。Heo 等<sup>[11]</sup>使用块匹配和 3D 滤波去噪算法实现对分解图像的降噪。Lee 等<sup>[12]</sup>将全变分去噪方法应用于分解的预处理和后处理过程, 以达到噪声抑制的效果。Xue 等<sup>[13]</sup>通过引入负对数似然项和边缘保持正则化来提升基材料分解后的图像质量。Harms 等<sup>[14]</sup>利用图像像素之间的相似性正则化来平滑图像并抑制噪声, 并通过保留更均匀的噪声功率谱来提高分解效果。Li 等<sup>[15]</sup>通过预先学习不同材料的共享特征和交叉依赖关系来提升基材料分解质量。降俊汝等<sup>[16]</sup>利用字典学习技术来挖掘基图像之间的稀疏性和关联性, 以进一步降低基材料的噪声和伪影。Ding 等<sup>[17]</sup>考虑到图像域不同材料图像具有共同或互补的边缘先验信息, 提出一种稀疏正则化方法, 用于对材料组成进行约束。总的来说, 该类基于迭代正则化的方法在分解过程的迭代时间过长<sup>[18]</sup>, 分解效率低, 难以准确描述基材料图像之间的映射关系。卷积神经网络的发展为 DECT 材料分解研究提供了新的途径。Lantz 等<sup>[19]</sup>提出一种基于交替方向乘子法 (ADMM) 的迭代展开方法, 将神经网络作为初始化和正则化器, 获得了较高的基材料分解精度。王冲旭

收稿日期: 2023-10-31; 修回日期: 2023-12-11; 录用日期: 2023-12-14; 网络首发日期: 2023-12-23

基金项目: 国家自然科学基金(61801003)、安徽省高等学校科学研究项目(2022AH050968)

通信作者: \*liujin@ahpu.edu.cn

等<sup>[20]</sup>在传统迭代分解的基础上,使用卷积神经网络(CNN)代替正则化项,可有效抑制分解后图像的噪声。Li等<sup>[21]</sup>将基材料重建得到的低能图像与原始低能图像的差作为先验项,可对先前的输出进行反馈和调整,该方法有效减少了材料边缘伪影,但存在超参数难调的问题。Xu等<sup>[22]</sup>提出一种由全卷积网络和全连接层组成的端到端分解双网络,可实现对高低能图像信息的分别提取,但在部分边缘结构的保持方面效果不佳。Zhang等<sup>[23]</sup>提出一种蝴蝶型网络结构,该网络是一种交叉结构,通过残差学习的交叉架构实现特征共享,可显著提升分解的基材料图像质量。Kawahara等<sup>[24]</sup>利用从 DECT 图像中重建的等效 kV 量级 CT 单能图像代替 DECT 图像作为生成对抗网络的输入来完成基材料分解。受蝴蝶型网络的启发,Shi等<sup>[25]</sup>提出一种 DIWGAN 模型,该模型使用两个 U-Net 作为生成器来分别提取特征,并在生成器的解码层以残差学习方式实现多层次特征融合,但该模型在软组织分解图中仍然存在少量的骨骼伪影。

受上述深度学习的基材料分解模型设计思想的启发,本文提出一种基于多通道交叉卷积 UCTransNet (MC-UCTransNet) 的图像域双材料分解方法。该网络采用双输入双输出架构,可实现一次性获得不同的基材料图像。所提网络以 UCTransNet<sup>[26]</sup>为基础架构,采用通道交叉融合 Transformer 模块和通道交叉注意模块来捕捉更复杂的通道信号相关性,以更好地进行特征提取与融合,实现基材料生成路径之间的信息交换。与此同时,为了更好地适应 CT 图像数据的特殊性,在网络训练时使用基于 Sigmoid 函数的归一化方法和混合损失函数来提高模型收敛速度及稳定性。研究表明,所提方法可以完成 DECT 的基材料分解任务,并能有效抑制基图像噪声和伪影,提高材料分解精度,与对比方法相比具有一定的优势。

## 2 方 法

### 2.1 图像域材料分解

在 DECT 中,CT 图像的衰减系数可由基材料 CT 图像衰减系数的线性组合来表征<sup>[27]</sup>,这时基材料图像分解模型可以表述为

$$\begin{pmatrix} \mu_L \\ \mu_H \end{pmatrix} = \begin{pmatrix} \mu_{1L} & \mu_{2L} \\ \mu_{1H} & \mu_{2H} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (1)$$

式中:变量的下标 L 和 H 分别表示低能谱和高能谱数据,下标 1 和 2 表示两种不同的基材料; $\mu_{ij}$  为材料  $i \in \{1, 2\}$  在能谱  $j \in \{L, H\}$  下的线性衰减系数; $x_1, x_2$  为两个基材料在图像中相同位置的归一化体积分数; $\mu_L, \mu_H$  为 DECT 图像的像素对。设  $\mathbf{A}$  表示维数为  $2N \times 2N$  ( $N$  为图像中的像素总数)的材料分解矩阵,则

$$\mathbf{A} = \begin{pmatrix} \mu_{1L} \mathbf{I} & \mu_{2L} \mathbf{I} \\ \mu_{1H} \mathbf{I} & \mu_{2H} \mathbf{I} \end{pmatrix}, \quad (2)$$

式中: $\mathbf{I}$  表示维数为  $N \times N$  的单位矩阵。式(1)可表述为  $\boldsymbol{\mu} = \mathbf{A}\mathbf{x}$ , 其中  $\boldsymbol{\mu}$  表示由高低能图像  $\mu_L, \mu_H$  组成的  $2N$  向量。 $\mathbf{x}$  为分解后基材料图像  $\mathbf{x}_1, \mathbf{x}_2$  共同组成的  $2N$  向量。这时,可以通过矩阵反演方法从 DECT 图像中得到基材料图像。通常材料分解的过程被表述为矩阵反演的不定问题的求解过程,这是因为在诊断 X 射线能量范围内,两种基材料的双能比  $\mu_{1H}/\mu_{1L}$  没有显著差异,导致线性矩阵  $\mathbf{A}$  的条件数很大,直接求逆时所生成的基材料图像与原始 CT 图像相比信噪比低,图像严重退化<sup>[28-31]</sup>。此外,这一分解过程受原 DECT 图像噪声的影响较大。为了降低这一影响,常采用正则化的约束方法来抑制分解后的基材料图像中的噪声。此时,图像域的基材料分解问题可表示为如下目标函数:

$$\min_{\mathbf{x}} (\mathbf{x} - \mathbf{A}^{-1}\boldsymbol{\mu})^T (\mathbf{x} - \mathbf{A}^{-1}\boldsymbol{\mu}) + \lambda \cdot R(\mathbf{x}), \quad (3)$$

式中: $R(\mathbf{x})$  为先验约束项,用于约束分解后基材料图像的特定信息。 $R(\mathbf{x})$  的设计在图像基材料分解领域中至关重要。虽然正则化约束类分解方法能获得高质量的基材料图像,但其超参数多、算法复杂度高、处理时间长和数据不相容导致的稳定性差等问题,限制了其在临床上的应用,难以充分发挥其价值。

为实现从 DECT 图像到基材料图像这一分解过程,即对目标函数式(3)的求解,深度学习(DL)方法可以将其简化为求解如下目标函数:

$$\min_M \|\mathbf{x} - \mathbf{M}(\boldsymbol{\mu})\|_2^2, \quad (4)$$

式中:函数  $\mathbf{M}(\cdot)$  表示从 DECT 图像到基材料图像的映射。这时函数  $\mathbf{M}(\cdot)$  在某种意义上为逆线性矩阵  $\mathbf{A}^{-1}$ 。为此,基于深度学习的 DECT 材料分解任务可以表述为通过数据驱动的方式寻找一个近似的逆线性矩阵函数。近年来,深度学习已成为解决医学成像中逆问题的一种重要方法,与传统方法相比具有更优异的性能和更快的处理速度,已经受到越来越多的关注。在 DECT 图像分解领域,深度学习模型可通过捕获图像高层特征来显示其学习不确定噪声分布的能力,以有效地适应任意噪声类型,显著提高基材料分解能力及分解质量<sup>[32-33]</sup>。

### 2.2 MC-UCTransNet

以 UCTransNet 为基础网络模型,以软阈值分割结果为标签图像,将重建得到的高低能图像拼接成多通道数据作为网络输入,使用交叉卷积完成 DECT 图像特征信息交互。相较于设计独立的两个网络来分别提取高低能 CT 图像特征的方法,所提方法能够更好地利用双能信息,从而得到更为准确可靠的结果。UCTransNet 使用编码器-解码器架构,其中编码器生成低分辨率图像特征,解码器将特征上采样到基材料图像。为了防止编码器中提取的丰富特征信息丢失,在编码器和解码器之间增加通道 Transformer 结构,重复  $L$  ( $L=4$ ) 次以构建  $L$  层



Transformer, 最终第  $L$  层的 4 个输出通过上采样分别与编码器特征进行残差连接, 被重构为  $O_1$ 、 $O_2$ 、 $O_3$  和  $O_4$ , 再通过与解码器特征  $D_1$ 、 $D_2$ 、 $D_3$  和  $D_4$  计算尺度权重, 对其进行加权得到  $\hat{O}_1$ 、 $\hat{O}_2$ 、 $\hat{O}_3$  和  $\hat{O}_4$ , 最后与解码器特征  $D_1$ 、 $D_2$ 、 $D_3$  和  $D_4$  进行级联, 用于实现编码器特征和解码器特征之间的特征融合。如图 1 所示, UCTransNet 是在 U-Net 的基础上改进得到的, U-Net 具有捕获多层次特征信息的优势, 但其编码器

和解码器之间存在语义鸿沟, 通过简单跳跃连接会影响模型的最最终性能。UCTransNet 使用通道 Transformer (CTrans) 思想来取代 U-Net 中的跳跃连接, 以更好地融合编码器特征。通道 Transformer 由 2 个模块组成: 通道交叉融合 Transformer (CCT) 模块, 用于多尺度编码器特征融合; 通道交叉注意 (CCA) 模块, 用于解码器特征和增强 CCT 特征的融合。

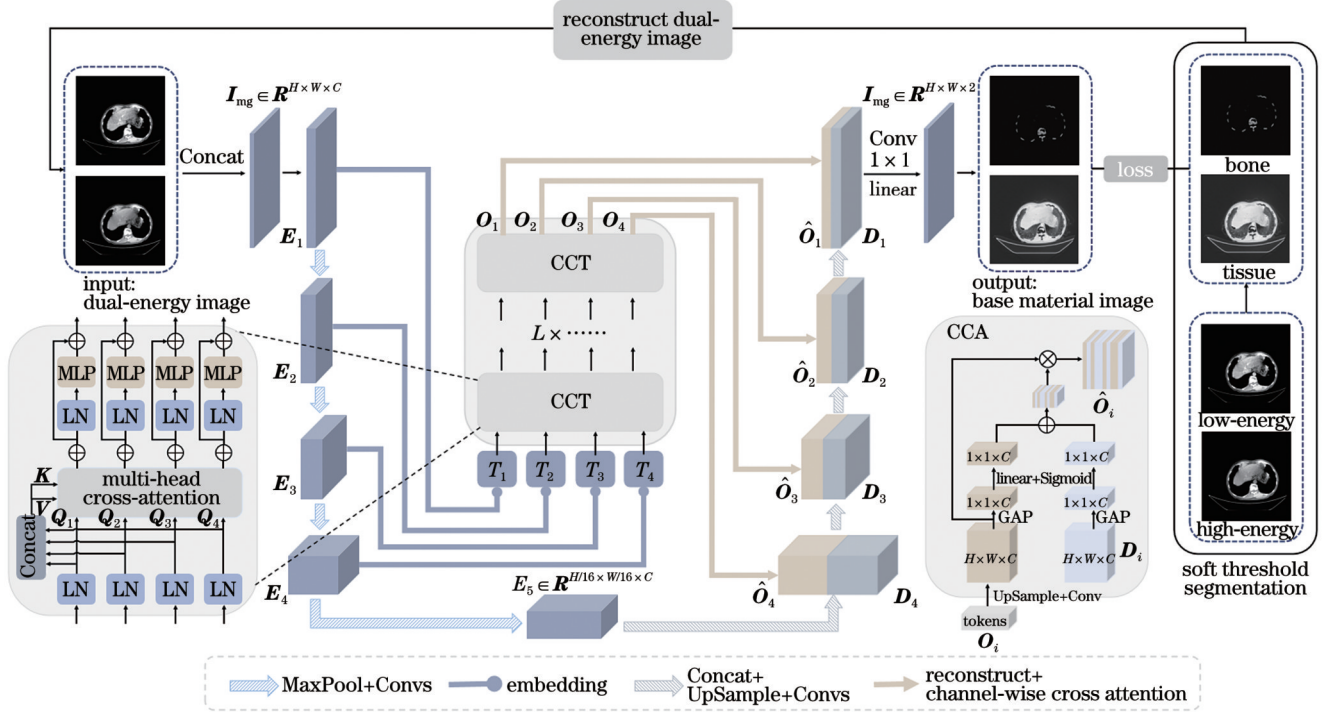


图 1 MC-UCTransNet 结构图  
Fig. 1 Architecture of MC-UCTransNet

为实现 DECT 图像的有效分解, 所使用的 UCTransNet 包括 4 个不同尺度: 第 1 个尺度的特征图  $E_1$  尺寸为  $512 \text{ pixel} \times 512 \text{ pixel}$ , 通道数为 64; 第 2 个尺度的特征图  $E_2$  尺寸减小为原来的 1/2, 通道数增加为原来的 2 倍; 第 3 个尺度的特征图  $E_3$  尺寸进一步减小为原来的 1/4, 通道数增加为原来的 4 倍; 第 4 个尺度的特征图  $E_4$  尺寸减小为原来的 1/8, 通道数增加为原来的 8 倍。在编码器端, 不同尺度特征图采用 CCT 模块来捕获局部跨通道信息, 实现多尺度上下文与交叉注意力融合, 减缓可能存在的尺度语义差距。在解码器端, 使用 CCA 模块融合多尺度特征和解码器阶段的特征, 以解决语义层次不一致的问题, 提高网络对语义层次不一致的处理能力和对目标的感知能力。

1) CCT 模块结构

所提网络使用多尺度特征嵌入 (MFE) 获取图像在不同尺度上的特征表示  $T_i (i = 1, 2, 3, 4)$ , 通过层归一化 (LN) 处理将特征进行归一化。层归一化是通过计算每一层的输入数据的均值和方差进行归一化的, 并对归一化后的数据进行缩放和平移, 以恢复数据的

原始分布, 从而提高模型的稳定性、加快其收敛速度, 再使用多头通道交叉注意 (MCA) 进行自注意力计算。在 MCA 中, 特征图被分为 4 个头 (head), 每个头都有自己的查询 ( $Q_i$ )、键 ( $K$ ) 和值 ( $V$ ) 矩阵。通过计算查询和键之间的相似度, 得到特征图的注意力权重, 其中权重表示每个通道对其他通道的重要性。将注意力权重应用于值矩阵, 以加权融合不同通道的特征表示, 从而捕捉不同通道之间的信号相关性, 最后使用多层感知器 (MLP) 学习一系列非线性映射。其中, 该 MLP 由一个展平层和一个全连接层组成, 可进一步提取和表达特征之间的复杂关系。

2) CCA 模块结构

所提网络将 Transformer 的输出  $O_i$  与解码器特征  $D_i$  作为 CCA 的输入, 再利用全局平均池化 (GAP) 生成尺度权重, 全局平均池化针对每幅输入特征图, 将其所有的像素值进行平均, 得到一个单一的值作为该特征图的池化输出。最后对不同尺度特征  $O_i$  进行加权得到  $\hat{O}_i$ , 进一步强调重要尺度信息的融合。

### 2.3 损失函数

受文献[34-36]的启发,网络训练时采用平均绝对误差(MAE)和结构相似性指数(SSIM)来构造适用于 DECT 基材料分解的混合损失函数。其中 MAE 损失的计算公式为

$$L^{\text{MAE}}(\mathbf{y}, \mathbf{x}) = \frac{1}{m} \sum_{i=1}^m |\mathbf{y}_i - \mathbf{x}_i|, \quad (5)$$

式中:  $\mathbf{x}$  和  $\mathbf{y}$  表示预测的基材料图像和对应的标签图像;  $m$  为像素的总数; 下标  $i$  表示第  $i$  个像素。SSIM 指数的计算公式为

$$E_{\text{SSIM}}(\mathbf{x}_1, \mathbf{x}_2) = \frac{(2\mu_{x_1}\mu_{x_2} + c_1)(2\sigma_{x_1, x_2} + c_2)}{(\mu_{x_1}^2 + \mu_{x_2}^2 + c_1)(\sigma_{x_1}^2 + \sigma_{x_2}^2 + c_2)}, \quad (6)$$

式中:  $\mu$  表示图像的均值;  $\sigma^2$  表示图像的方差;  $\sigma_{x_1, x_2}$  表示两幅图像的协方差;  $c_1$  和  $c_2$  表示与图像的值范围相关的常数。这时 SSIM 损失可表示为

$$L^{\text{SSIM}}(\mathbf{y}, \mathbf{x}) = 1 - E_{\text{SSIM}}(\mathbf{y}, \mathbf{x}). \quad (7)$$

对于整个网络,其混合损失函数设计为

$$L = \alpha \cdot L^{\text{MAE}} + \beta \cdot L^{\text{SSIM}}, \quad (8)$$

式中:  $\alpha$  和  $\beta$  为两个手动设置的超参数。

## 3 实验和结果

### 3.1 数据准备

#### 1) 训练标签

实验数据集由上海联影医疗科技股份有限公司提供,使用的医学成像设备型号为 uCT 960+, 采用轴向扫描方式进行图像采集,管电流大小为 533 mA,管电压大小分别为 80 kV 和 140 kV,切片厚度为 0.5 mm,源到探测器的距离为 1118.36 mm,源到旋转中心的距离为 600 mm,像素间距为 0.98 mm,重建图像尺寸为 512 pixel  $\times$  512 pixel。实验中对 80 kV 和 140 kV 管电压下大小为 512 pixel  $\times$  512 pixel 的 CT 图像进行加权软阈值分解来获得基材料图像。以制作骨头和组织基材料图像为例,加权软阈值算法的实施过程为:将低于某个阈值  $T_s$  的像素认为是软组织,高于某个阈值  $T_b$  的像素视为骨骼,而处于两个阈值中间的像素可以被视为两种材料的混合,根据该点像素值赋予其骨头和组织所占的权重。

#### 2) DECT 图像获取

为了模拟不同能谱的 CT 投影,需要知道 X 射线能谱和指定的能量段。实验中可进一步假设 X 射线能谱被划分成  $M$  个能量段,这时扇形射束扫描几何下的能谱 CT 投影是来自第  $m$  ( $m = 1, \dots, M$ ) 个能量窗口 ( $\Delta E_m = E_{m+1} - E_m$ ) 沿着 X 射线路径  $\ell$  ( $\ell = 1, \dots, L$ ) 的光子数,计算过程可表示为

$$z_{m\ell} = \int_{E_m}^{E_{m+1}} I_{m\ell}(E) \exp\left[-\int_{r \in \ell} \mu(E, r) dr\right] dE, \quad (9)$$

式中:  $\int_{E_m}^{E_{m+1}} dE$  和  $\int_{r \in \ell} dr$  表示在  $\Delta E_m$  范围内沿着 X 射线

路径  $\ell$  的积分; 衰减系数  $\mu(E, r)$  为能量  $E$  下位置  $r$  处的线性衰减系数;  $I_{m\ell}(E)$  表示从 X 射线源发射的能量  $E$  的原始光子强度。由于成像对象仅考虑两种材料(骨头和软组织),  $\mu(E, r)$  可以表示为

$$\mu(E, r) = \phi_b(E) \rho_b(r) + \phi_s(E) \rho_s(r), \quad (10)$$

式中:  $\phi_b(E)$  和  $\phi_s(E)$  分别为能量  $E$  下骨头和软组织的质量衰减系数,该系数可以通过检索美国国家标准与技术研究所(NIST)报告(Hubbell & Seltzer, 1995)<sup>[37]</sup>中的表格来确定;  $\rho_b(r)$  和  $\rho_s(r)$  分别表示骨头和软组织在位置  $r$  处的占比。X 射线源光子通量  $I_{m\ell}^{(0)}$  可以表示为

$$I_{m\ell}^{(0)} = \int_{E_m}^{E_{m+1}} I_{m\ell}(E) dE. \quad (11)$$

将式(10)和式(11)代入式(9),可以得到

$$e_{m\ell} = \int_{E_m}^{E_{m+1}} r_{m\ell}(E) \exp\left[-\int_{r \in \ell} \phi_b(E) \rho_b(r) + \phi_s(E) \rho_s(r) dr\right] dE, \quad (12)$$

式中:  $e_{m\ell} = z_{m\ell}/I_{m\ell}^{(0)}$ ,  $r_{m\ell}(E) = I_{m\ell}(E)/I_{m\ell}^{(0)}$  分别表示 X 射线强度和探测器灵敏度的归一化能谱分布。最后,通过对式(12)两侧进行对数运算,得到能谱 CT 的投影数据:

$$p_{m\ell} = \ln \int_{E_m}^{E_{m+1}} r_{m\ell}(E) \exp\left[-\int_{r \in \ell} \phi_b(E) \rho_b(r) + \phi_s(E) \rho_s(r) dr\right] dE. \quad (13)$$

目前,在给定 CT 图像数量的情况下,可以按照式(13)生成能谱 CT 投影,然后利用统计迭代重建技术(SIRT)重建图像。图 2 为能谱 CT 数据模拟的流程图。可以看到,生成流程主要可以分为 3 个阶段:基材料图像获取、能谱投影生成和能谱图像重建。实验中首先将 DECT 图像进行阈值分割以获得指定的基材料图像,如骨骼和软组织。然后,根据 X 射线发射光谱和材料衰减系数,生成能谱 CT 投影数据。最后,利用重建算法<sup>[38]</sup>对能谱 CT 投影数据进行重建,模拟得到不同扫描能量段的能谱图,如  $80 \times 10^3 \text{V}_p$ 、 $140 \times 10^3 \text{V}_p$  的 DECT 图像<sup>[39]</sup>。

### 3.2 数据预处理

考虑到 CT 图像的 HU 值范围跨度较大这一特殊性,使用一般的 Min-Max 归一化方法可能会导致归一化后的数据失去数据的分布特征,且其对输入数据的异常值非常敏感,会影响模型在训练和预测中的表现。因此,实验中采用 Sigmoid 函数进行数据的归一化处理,将数据映射到一个较小的范围内,这样可以降低数据的偏度和峰度,使数据更接近正态分布,从而减小数据的尺度差异和模型对输入数据的敏感度,提高模型的鲁棒性。归一化计算公式为

$$y = \frac{1}{\exp(-x/c)}, \quad (14)$$



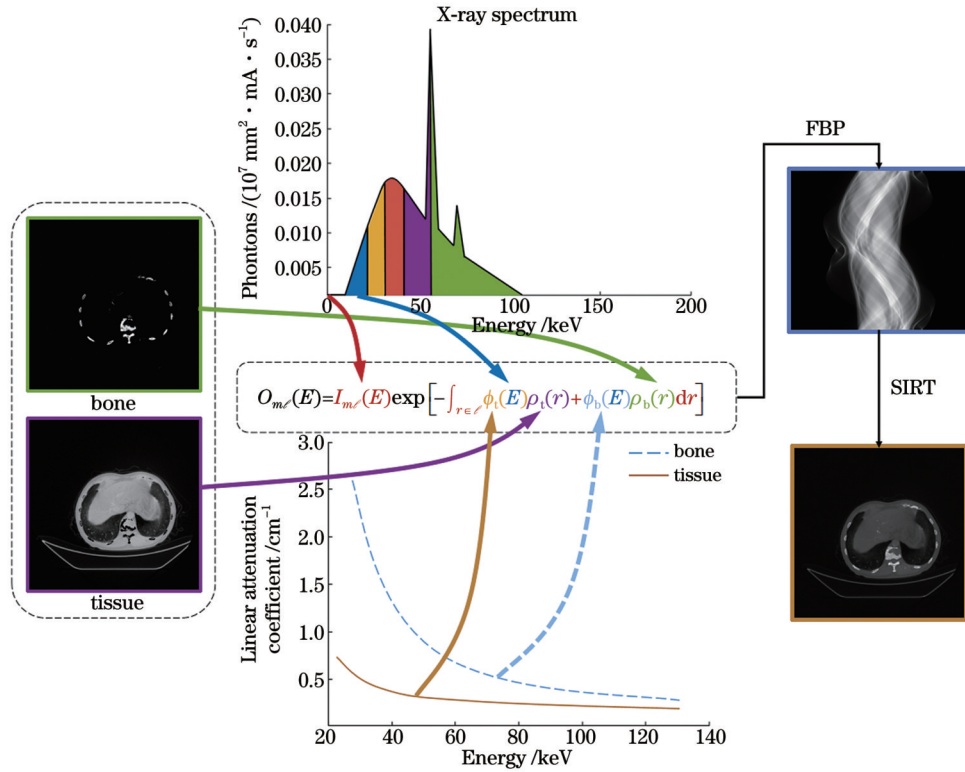


图 2 DECT 图像生成流程图

Fig. 2 Flow chart of the generation of DECT image

式中： $x$ 为原始图像数据； $y$ 为归一化后的图像数据； $c$ 为常数，实验中将其设置为 1000，用于缩小数据范围，减小数值溢出风险。

### 3.3 实验配置

实验是在 PC (12th Gen Intel(R) Core(TM) i5-12600X CPU(16CPUs), 32 GB RAM, NVIDIA GeForce RTX 4090) 上的 TensorFlow 2.6 和 python 3.8 环境中进行的，4000 幅分别包含 80 kV 和 140 kV 能量的 DECT 图像及其骨骼和软组织基材料标签图像用于网络训练，图像大小均为 512 pixel × 512 pixel。为加快训练速度，提高模型整体性能，实验采用 Adam 优化器实现参数更新，学习率采用 Keras 中的回调函数 ReduceLROnPlateau 自适应调整学习率，激活函数采用 ReLU 激活函数，批大小设为 2，训练过程共包含 20 个 epochs。实验采用 MAE、均方误差 (MSE)、峰值信噪比 (PSNR) 和 SSIM 对不同方法的性能进行定量评估。

MSE 用于全局评价基材料分解后的图像质量，其计算公式为

$$I_{MSE} = \frac{1}{m} \sum_{i=1}^m (y_i - x_i)^2, \quad (15)$$

式中： $y_i$  表示预测值； $x_i$  表示真实值； $m$  为分解图像中所有像素的数量。PSNR 用于衡量分解后图像失真程度及噪声水平，其计算公式为

$$I_{PSNR} = 10 \times \lg \frac{I_{MaxValue}^2}{I_{MSE}}, \quad (16)$$

式中： $I_{MaxValue}$  表示参考图像的最大值； $I_{MSE}$  表示分解图像和参考图像之间的均方误差。

### 3.4 基材料分解实验结果

#### 3.4.1 骨骼和软组织基材料分解实验结果

为了验证所设计的 MC-UCTransNet 在基材料分解任务中的有效性，实验选用矩阵反演 (Matrix Inversion)<sup>[22]</sup>、迭代分解 (Iterative Decomposition)<sup>[40]</sup>、全卷积网络 (FCN)<sup>[22]</sup>、蝴蝶型网络 (Butterfly-net)<sup>[23]</sup> 和 DIWGAN<sup>[25]</sup> 5 种方法进行对比，其中矩阵反演和迭代分解方法是传统的基材料分解方法，FCN、Butterfly-net 和 DIWGAN 是深度学习方法。为定性分析分解图像的质量，图 3 和图 4 中给出了不同方法对两个部位 (Group 1 和 Group 2) 的基材料分解结果，其中：第一列为骨骼和软组织基材料图像的真实标签，作为分解后的参考图像；第二列为直接矩阵反演方法生成的基材料图像，可以看出，生成的图像中噪声较为明显，这是因为该方法仅线性地组合了高低能图像，导致噪声水平被放大；第三列为迭代分解后的结果，可以看到，基材料图像中噪声和伪影得到了很好的抑制，但受到平滑正则化项的约束，分解后的组织基材料内部的纹理结构出现严重的丢失，从而出现过平滑现象；第四列为 FCN 方法生成的基材料图像，可以发现 FCN 方法虽然能够实现基材料的分解，但分解结果的边缘和纹理依然不够清晰，整个图像过于平滑；第五列为 Butterfly-net 的生成结果，其分解图像质量显著提高，骨骼和软组织的内部结构得到较好保留，但仍存在少

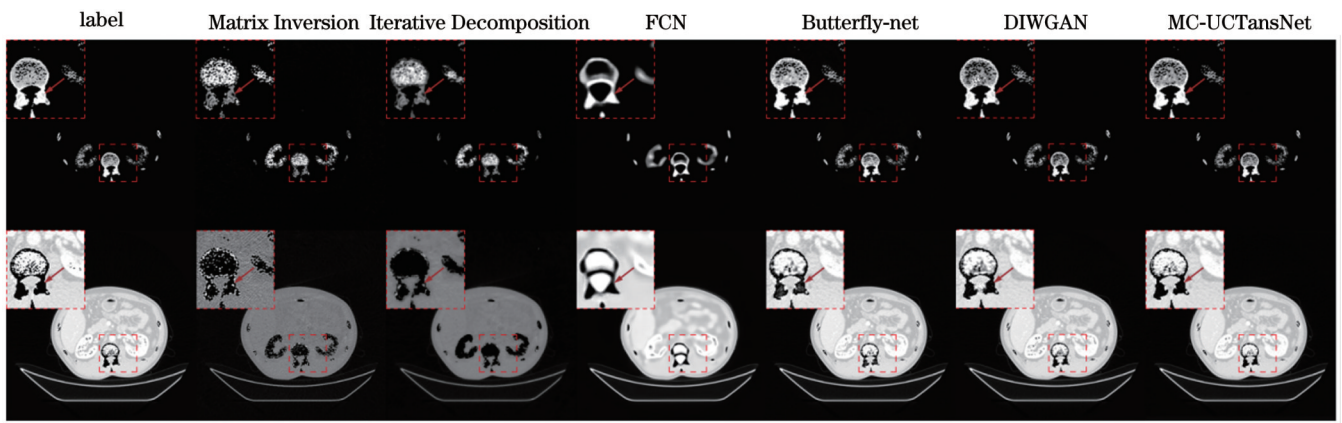


图 3 不同方法在 Group 1 测试集中的分解结果,其中第一行为骨骼基材料图,第二行为软组织基材料图

Fig. 3 Decomposition results of different methods in the Group 1 test set, where the first row represents images of bone-based material, and the second row represents images of soft tissue-based material

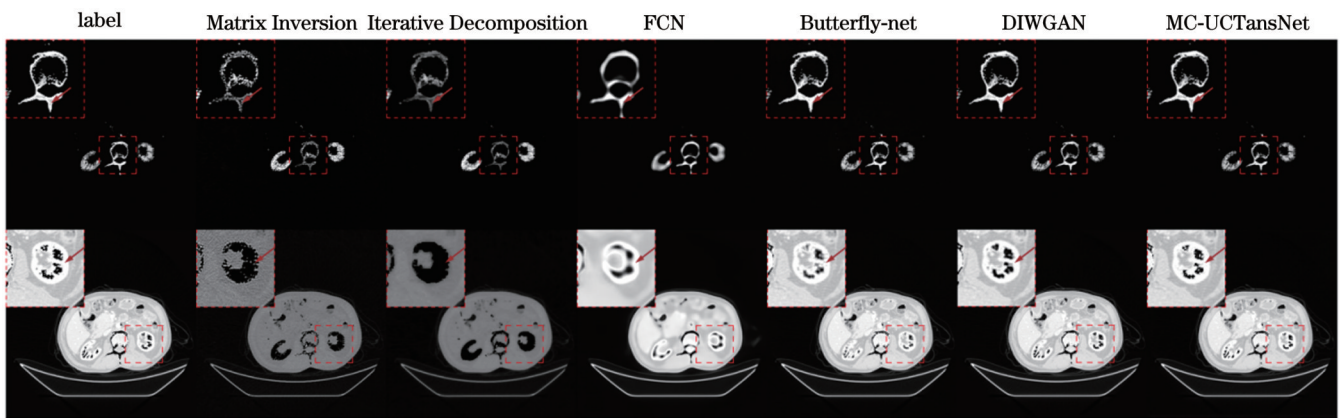


图 4 不同方法在 Group 2 测试集中的分解结果,其中第一行为骨骼基材料图,第二行为软组织基材料图

Fig. 4 Decomposition results of different methods in Group 2 test set, where the first row represents images of bone-based material, and the second row represents images of soft tissue-based material

量的细节结构丢失以及边缘模糊的现象;第六列为 DIWGAN 方法的分解结果,其分解质量相较于 Butterfly-net 有一定的提升,但同样存在边缘模糊和少量细节丢失的情况;第七列为所提方法的分解结果,可以看到,所生成的骨骼和软组织图像在视觉效果上与标准图像相似,与其他对比方法相比,所提方法分解后

的图像拥有更加清晰的材料图像边缘,组织纹理保持性能更理想,且噪声和伪影最少。

不同方法分解后的骨骼材料和软组织材料图像的量化结果如表 1、表 2 所示,数值的表示形式为(均值±方差)。可以看出:MC-UCTansNet 在 MAE、MSE、PSNR 和 SSIM 上的表现大部分优于其他 5 种方法,具

表 1 不同算法对骨骼基材料分解结果的定量评价

Table 1 Quantitative scores of bone-based material decomposition results with different methods

Method	Matrix Inversion	Iterative Decomposition	FCN	Butterfly-net	DIWGAN	MC-UCTansNet	unit: mean±variance
MAE	$0.7676 \times 10^{-2} \pm 4.13 \times 10^{-8}$	$0.7953 \times 10^{-2} \pm 2.83 \times 10^{-8}$	$0.2218 \times 10^{-2} \pm 2.40 \times 10^{-8}$	$0.1299 \times 10^{-2} \pm 0.88 \times 10^{-8}$	$0.1022 \times 10^{-2} \pm 0.65 \times 10^{-8}$	$0.0924 \times 10^{-2} \pm 0.59 \times 10^{-8}$	
MSE	$0.0479 \times 10^{-2} \pm 2.26 \times 10^{-9}$	$0.0434 \times 10^{-2} \pm 1.89 \times 10^{-9}$	$0.0366 \times 10^{-2} \pm 1.66 \times 10^{-9}$	$0.0115 \times 10^{-2} \pm 0.30 \times 10^{-9}$	$0.0108 \times 10^{-2} \pm 0.22 \times 10^{-9}$	$0.0098 \times 10^{-2} \pm 0.22 \times 10^{-9}$	
PSNR	$32.6585 \pm 0.21$	$33.0787 \pm 0.20$	$33.8317 \pm 0.24$	$38.8924 \pm 0.47$	$39.1308 \pm 0.35$	$39.5849 \pm 0.46$	
SSIM	$0.9701 \pm 4.02 \times 10^{-6}$	$0.9851 \pm 0.88 \times 10^{-6}$	$0.9807 \pm 0.94 \times 10^{-6}$	$0.9926 \pm 0.77 \times 10^{-6}$	$0.9934 \pm 0.54 \times 10^{-6}$	$0.9939 \pm 0.65 \times 10^{-6}$	



表 2 不同算法对软组织基材料分解结果的定量评价

Table 2 Quantitative scores of soft tissue-based material decomposition results with different methods

unit: mean  $\pm$  variance

Method	Matrix Inversion	Iterative Decomposition	FCN	Butterfly-net	DIWGAN	MC-UCTransNet
MAE	$2.7783 \times 10^{-2} \pm 2.05 \times 10^{-7}$	$2.4921 \times 10^{-2} \pm 2.30 \times 10^{-7}$	$0.5653 \times 10^{-2} \pm 0.15 \times 10^{-7}$	$0.2742 \times 10^{-2} \pm 0.06 \times 10^{-7}$	$0.2468 \times 10^{-2} \pm 0.04 \times 10^{-7}$	$0.2650 \times 10^{-2} \pm 0.04 \times 10^{-7}$
	$0.2662 \times 10^{-2} \pm 0.83 \times 10^{-8}$	$0.2719 \times 10^{-2} \pm 1.03 \times 10^{-8}$	$0.0315 \times 10^{-2} \pm 7.06 \times 10^{-8}$	$0.0116 \times 10^{-2} \pm 0.008 \times 10^{-8}$	$0.0089 \times 10^{-2} \pm 0.006 \times 10^{-8}$	$0.0062 \times 10^{-2} \pm 0.003 \times 10^{-8}$
PSNR	$23.6462 \pm 0.02$	$23.5547 \pm 0.03$	$32.9346 \pm 0.14$	$37.2569 \pm 0.10$	$38.4406 \pm 0.13$	$39.9546 \pm 0.13$
SSIM	$0.8589 \pm 1.41 \times 10^{-5}$	$0.9321 \pm 0.18 \times 10^{-5}$	$0.9574 \pm 0.13 \times 10^{-5}$	$0.9871 \pm 0.03 \times 10^{-5}$	$0.9888 \pm 0.03 \times 10^{-5}$	$0.9904 \pm 0.02 \times 10^{-5}$

有更高的分解精度和更好的图像质量；MC-UCTransNet在骨骼及软组织的分解结果中的量化方差也较小,表明所设计网络模型具有更好的稳定性。

图 5 所示分别为不同分解方法在不同测试集上的骨骼和软组织材料分解结果的 PSNR, Case 1、Case 2、

Case 3 和 Case 4 分别表示 4 个测试集。可以看出,大部分情况下,深度学习类分解方法相较于传统方法能获得更高的信噪比,而 MC-UCTransNet 在 4 个测试集中的结果均优于对比算法。这一量化比较进一步证明了所提方法具有较强的鲁棒性。

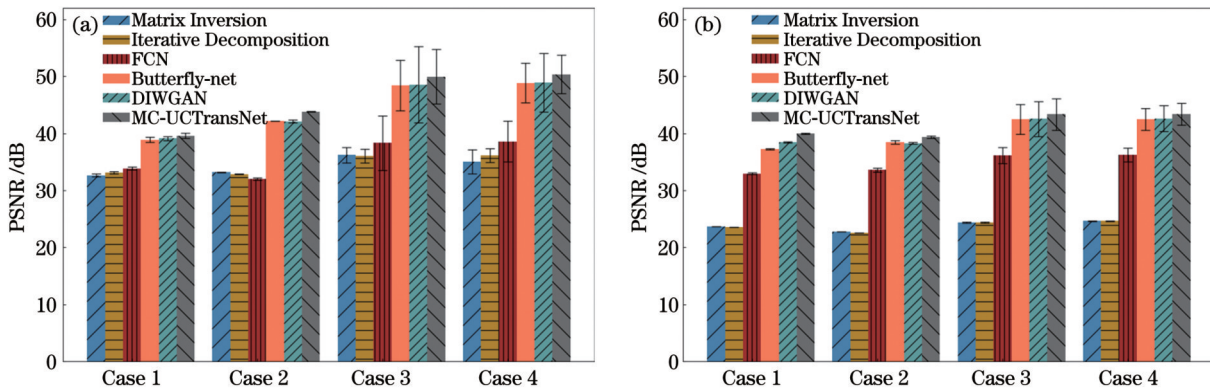


图 5 4 组测试集中不同方法的骨骼基材料和软组织基材料分解结果的 PSNR 值。(a) 骨骼基材料; (b) 软组织基材料

Fig. 5 PSNR values of bone-based materials and soft tissue-based materials decomposition results of different methods in four test sets.

(a) Bone-based materials; (b) soft tissue-based materials

图 6 所示分别为不同分解方法在 4 个测试集上对骨骼和软组织材料分解结果的 MAE、MSE、SSIM。通过对比可以看出,不同分解方法在骨骼基材料上的表现均优于在软组织基材料上的表现,迭代分解方法比矩阵反演和 FCN 方法具有更好的结构保持能力,与传统方法和 FCN 相比,Butterfly-net、DIWGAN 在骨骼和软组织分解图像上的 MAE、MSE 和 SSIM 均有显著提升,而 MC-UCTransNet 在软组织分解图像上取得了与 DIWGAN 相近的结果,在骨骼分解图像的表现要优于 Butterfly-net 方法。这也表明所提方法在基材料分解的质量及结构保持上具有较好的性能。

为进一步比较深度学习方法的分解精度,将不同深度学习方法分解得到的图像重建回 CT 图像,其结果如图 7 所示。图 7(a) 所示为原始低能 ( $80 \times 10^3 V_p$ ) CT 图像。图 7(b) 所示为 FCN 方法分解的基材料图像重建结果,可以看到,图像内部出现一些器官轮廓,但是图像过于平滑,在纹理和细节结构显示方面存在

不足。图 7(c) 所示为 Butterfly-net 分解的基材料图像重建结果,可以看到,图像有了清晰的纹理和细节,已经较为接近原始的低能 CT 图像。图 7(d) 所示为 DIWGAN 分解的基材料图像重建结果,其取得了与 Butterfly-net 方法相似的结果,但重建图中存在一些条状伪影。图 7(e) 所示为使用 MC-UCTransNet 分解并重建得到的图像,其整体图像质量优于 Butterfly-net 和 DIWGAN 方法的重建结果,重建图像中没有出现条状伪影,且图像细节保留较好。图 8(a)~(d) 为对应方法重建图像与原始低能图像之间的差异图。从差异图可以看到:FCN 方法对软组织材料的拟合效果较好,但对骨骼材料的拟合还有所欠缺;Butterfly-net 方法进一步改善了对骨骼和软组织的拟合效果,显著缩小了分解再重建图像与原始图像之间的差异;DIWGAN 方法在骨骼材料的一些小区域上存在明显差异,表明其在该区域的重建效果较差;MC-UCTransNet 取得最好的拟合结果,它的基材料重建图像与原始低能图像之间的差异最小。

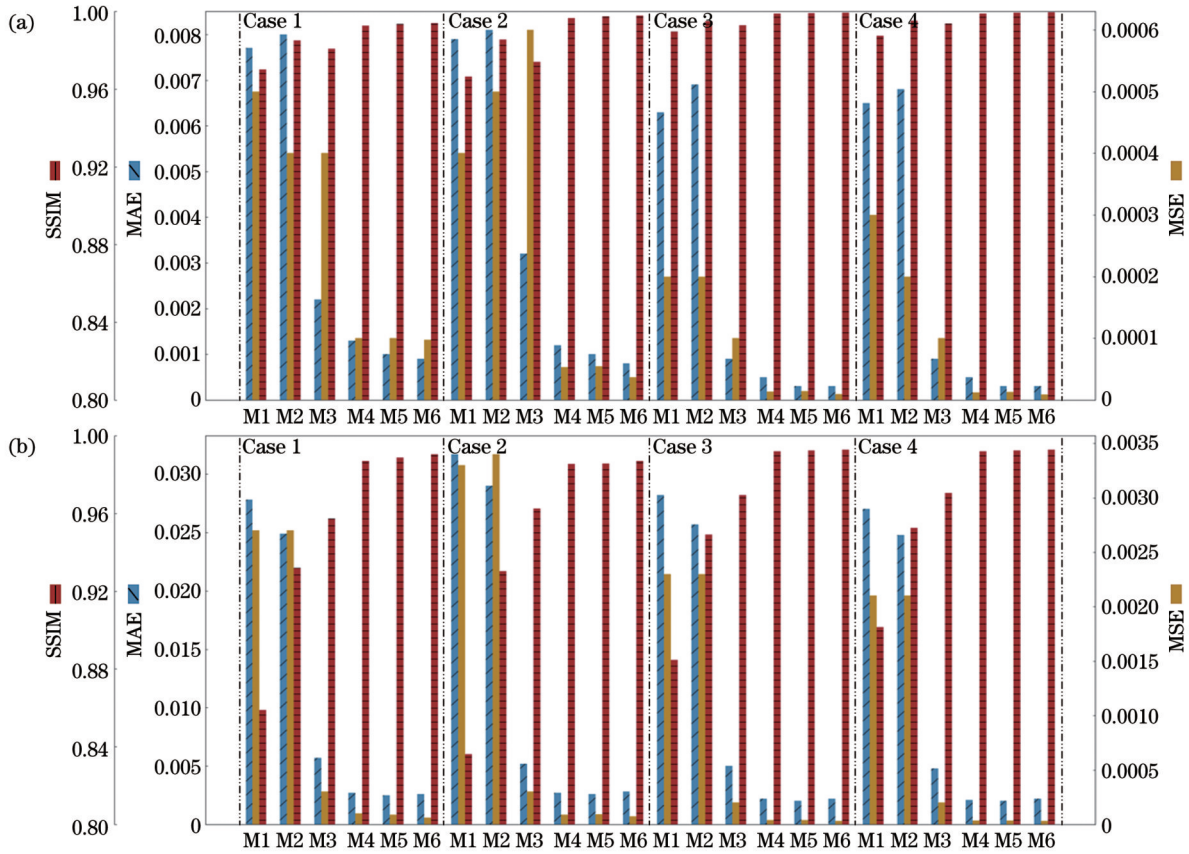


图 6 4组测试集中不同方法的骨骼基材料和软组织基材料分解结果的MAE、MSE、SSIM值,其中M1、M2、M3、M4、M5和M6分别代表Matrix Inversion、Iterative Decomposition、FCN、Butterfly-net、DIWGAN、MC-UCTransNet。(a)骨骼基材料;(b)软组织基材料

Fig. 6 MAE, MSE, and SSIM values of bone-based materials and soft tissue-based materials decomposition results of different methods in four test sets. Among them, M1, M2, M3, M4, M5, and M6 represent Matrix Inversion, Iterative Decomposition, FCN, Butterfly-net, DIWGAN, and MC-UCTransNet, respectively. (a) Bone-based materials; (b) soft tissue-based materials

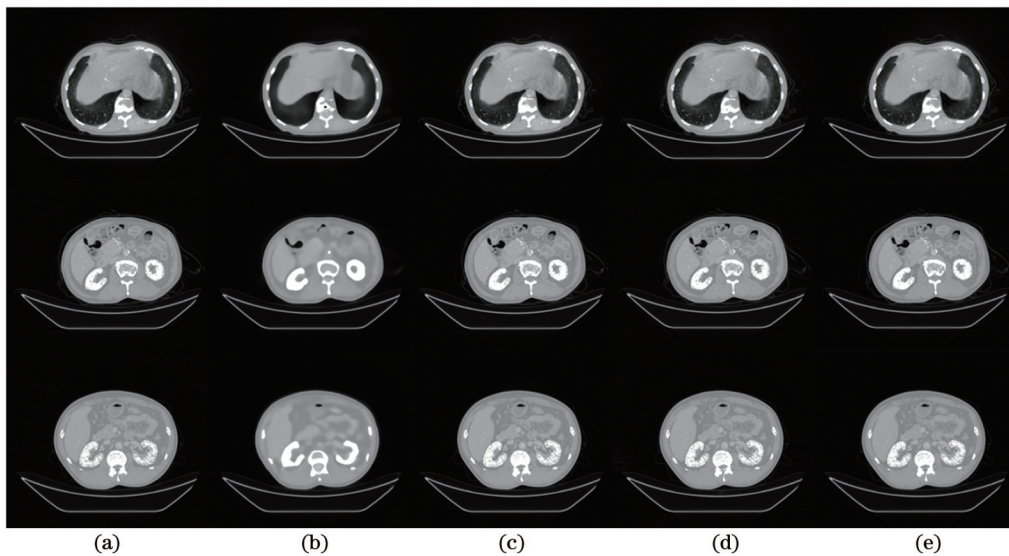


图 7 不同方法得到的分解结果的低能重建图。(a)参考的低能图像;(b)FCN的重建图像;(c)Butterfly-net的重建图像;(d)DIWGAN的重建图像;(e)MC-UCTransNet的重建图像

Fig. 7 Low-energy reconstruction images of decomposition results obtained by different methods. (a) Reference low-energy images; (b) reconstructed images of FCN; (c) reconstructed images of Butterfly-net; (d) reconstructed images of DIWGAN; (e) reconstructed images of MC-UCTransNet



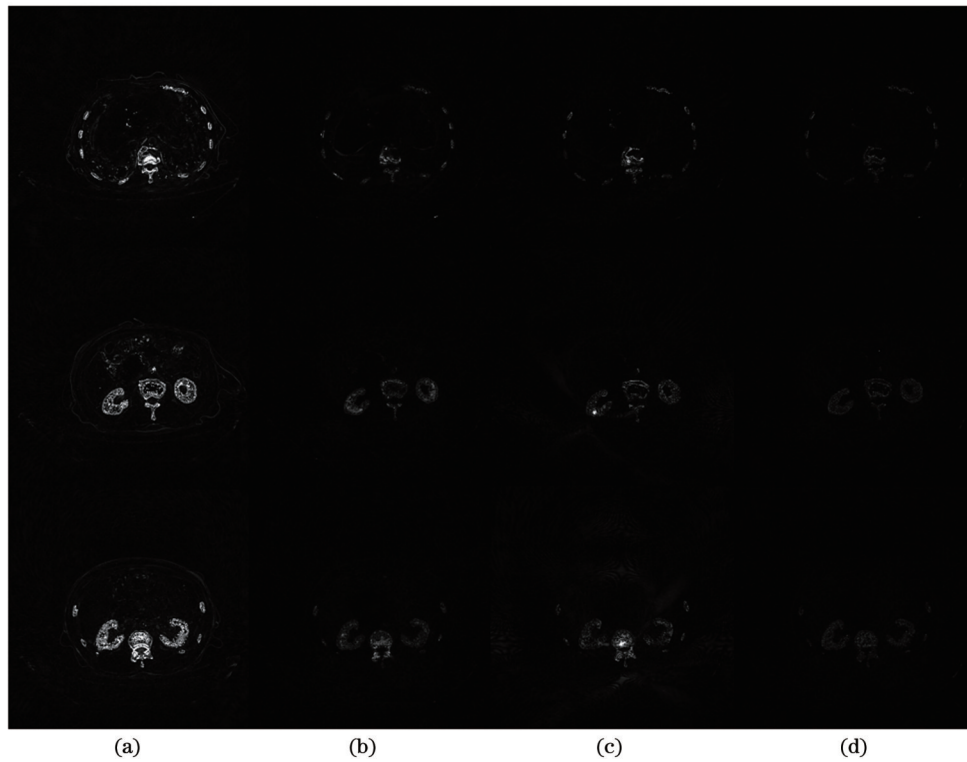


图 8 不同方法得到的分解结果的低能重建图与参考低能图像之间的差异图。(a) FCN; (b) Butterfly-net; (c) DIWGAN; (d) MC-UCTransNet

Fig. 8 Difference between reference low energy images and decomposition results obtained by different methods. (a) FCN; (b) Butterfly-net; (c) DIWGAN; (d) MC-UCTransNet

### 3.4.2 碘和软组织基材料分解实验结果

为验证所提方法在不同分解任务中的性能,进一步验证了其在胸部碘和软组织基材料图像分解中的效果。图 9 和图 10 所示为不同方法分解后的结果,其中:第一列为标签图像;第二、第三列分别为矩阵反演和迭代分解方法获得的碘和软组织基材料图像,可以发现,分解后的基材料图像出现明显的细节丢失,软组织材料中的感兴趣区域轮廓并不清晰;第四列为 FCN 方法的分解结果,虽然感兴趣区域的轮廓较为清

晰,但其碘基材料图像的骨头出现大面积丢失,同时,软组织基材料图像中也出现一些错误结构;第五列为 Butterfly-net 方法的分解结果,其分解图像质量较前 3 种方法有了明显的提升,但边缘仍存在一定程度的模糊;第六列为 DIWGAN 方法的分解结果,虽然相较于 Butterfly-net 方法的图像边缘更为清晰,但仍存在一些错误结构;第七列为 MC-UCTransNet 方法的分解结果,可以看到,分解后的基材料图像边缘结构较清晰,图像对比度高。

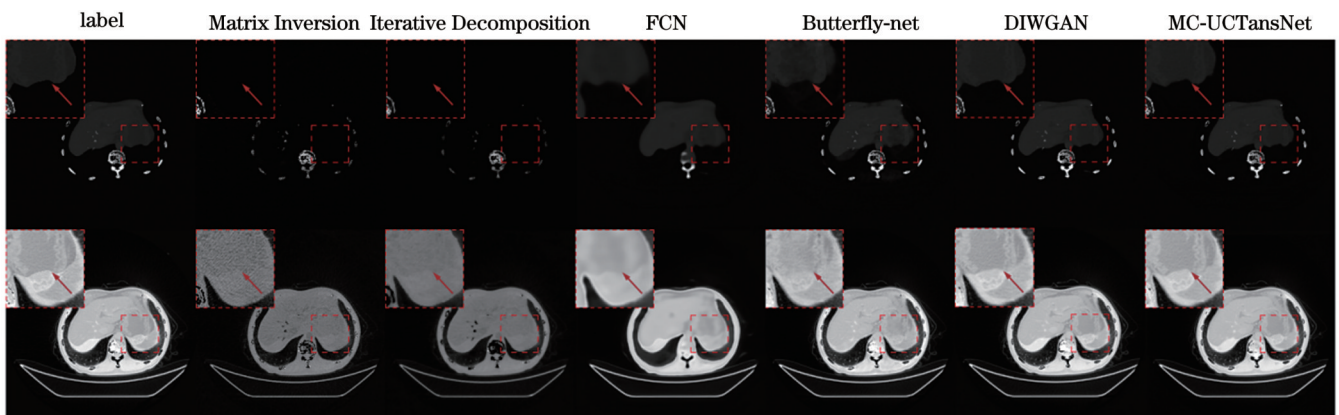


图 9 不同方法在 Group 1 测试集中的分解结果,其中第一行为碘基材料图,第二行为软组织基材料图

Fig. 9 Decomposition results of different methods in the Group 1 test set, where the first row represents images of iodine-based material, and the second row represents images of soft tissue-based material

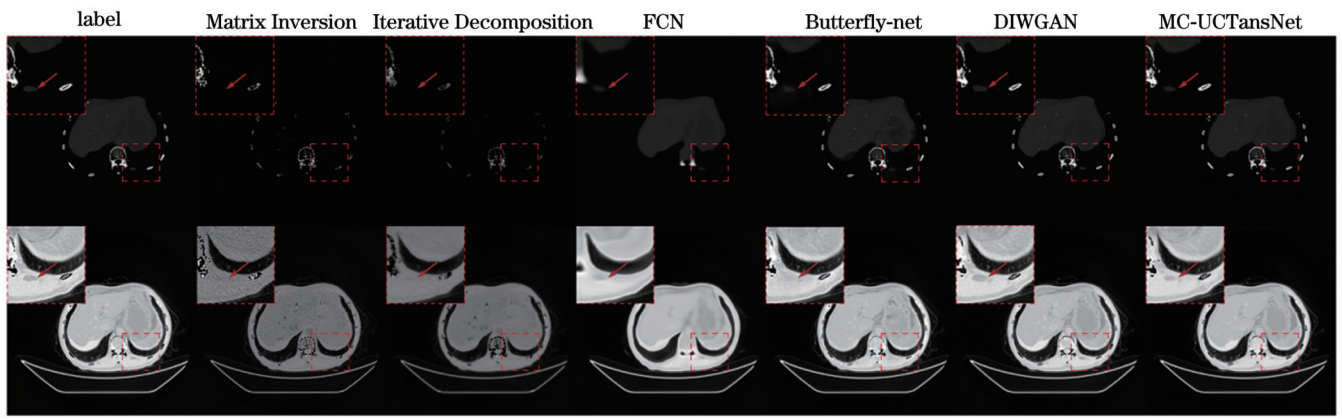


图 10 不同方法在 Group 2 测试集中的分解结果,其中第一行为碘基材料图,第二行为软组织基材料图

Fig. 10 Decomposition results of different methods in the Group 2 test set, where the first row represents images of iodine-based material, and the second row represents images of soft tissue-based material

不同方法分解后的碘材料和软组织材料图像的量化结果如表 3、表 4 所示。可以看到:MC-UCTransNet 除了在碘材料上的 SSIM 表现略低于 Butterfly-net 以及在 MAE 上的表现略低于 DIWGAN,在 MAE、

MSE、PSNR 和 SSIM 上的表现均优于其他 5 种方法,具有更高的分解精度和更好的图像质量;MC-UCTransNet 在碘材料及软组织的分解结果中量化方差也较小,表明所设计网络模型具有更好的稳定性。

表 3 不同方法对碘基材料分解结果的定量评价

Table 3 Quantitative scores of iodine-based material decomposition results of different methods

Method	Matrix Inversion	Iterative Decomposition	FCN	Butterfly-net	DIWGAN	MC-UCTransNet	unit: mean $\pm$ variance
MAE	$1.0684 \times 10^{-2} \pm 73.29 \times 10^{-8}$	$1.0729 \times 10^{-2} \pm 77.35 \times 10^{-8}$	$0.2807 \times 10^{-2} \pm 3.22 \times 10^{-8}$	$0.1634 \times 10^{-2} \pm 0.51 \times 10^{-8}$	$0.1157 \times 10^{-2} \pm 0.95 \times 10^{-8}$	$0.1427 \times 10^{-2} \pm 0.43 \times 10^{-8}$	
MSE	$0.0613 \times 10^{-2} \pm 3.26 \times 10^{-9}$	$0.0649 \times 10^{-2} \pm 3.87 \times 10^{-9}$	$0.0322 \times 10^{-2} \pm 1.04 \times 10^{-9}$	$0.0042 \times 10^{-2} \pm 0.04 \times 10^{-9}$	$0.0052 \times 10^{-2} \pm 0.06 \times 10^{-9}$	$0.0040 \times 10^{-2} \pm 0.01 \times 10^{-9}$	
PSNR	$31.6845 \pm 0.25$	$31.4407 \pm 0.26$	$34.4763 \pm 0.25$	$43.3787 \pm 0.34$	$42.3866 \pm 0.33$	$43.5526 \pm 0.13$	
SSIM	$0.9667 \pm 6.22 \times 10^{-6}$	$0.9793 \pm 4.38 \times 10^{-6}$	$0.9753 \pm 1.89 \times 10^{-6}$	$0.9948 \pm 0.10 \times 10^{-6}$	$0.9940 \pm 0.22 \times 10^{-6}$	$0.9947 \pm 0.14 \times 10^{-6}$	

表 4 不同方法对软组织基材料分解结果的定量评价

Table 4 Quantitative scores of soft tissue-based material decomposition results with different methods

Method	Matrix Inversion	Iterative Decomposition	FCN	Butterfly-net	DIWGAN	MC-UCTransNet	unit: mean $\pm$ variance
MAE	$2.3850 \times 10^{-2} \pm 6.30 \times 10^{-7}$	$2.2024 \times 10^{-2} \pm 5.85 \times 10^{-7}$	$0.6133 \times 10^{-2} \pm 0.43 \times 10^{-7}$	$0.4009 \times 10^{-2} \pm 0.17 \times 10^{-7}$	$0.3295 \times 10^{-2} \pm 0.24 \times 10^{-7}$	$0.3466 \times 10^{-2} \pm 0.16 \times 10^{-7}$	
MSE	$0.2148 \times 10^{-2} \pm 1.96 \times 10^{-8}$	$0.2202 \times 10^{-2} \pm 1.98 \times 10^{-8}$	$0.0313 \times 10^{-2} \pm 0.04 \times 10^{-8}$	$0.0115 \times 10^{-2} \pm 0.08 \times 10^{-8}$	$0.0125 \times 10^{-2} \pm 0.02 \times 10^{-8}$	$0.0112 \times 10^{-2} \pm 0.03 \times 10^{-8}$	
PSNR	$24.5465 \pm 0.08$	$24.4386 \pm 0.08$	$32.9122 \pm 0.08$	$36.3110 \pm 0.58$	$36.9219 \pm 0.22$	$37.3843 \pm 0.37$	
SSIM	$0.8732 \pm 2.19 \times 10^{-5}$	$0.9317 \pm 0.47 \times 10^{-5}$	$0.9539 \pm 0.23 \times 10^{-5}$	$0.9807 \pm 0.19 \times 10^{-5}$	$0.9814 \pm 0.11 \times 10^{-5}$	$0.9815 \pm 0.15 \times 10^{-5}$	

### 3.5 模型分析

#### 3.5.1 网络设计消融实验

为了验证所提方法的有效性,进一步对通道注意力机制、损失函数和归一化预处理 3 个模块进行了验证。实验采用无注意力的 U-Net 方法验证通道

注意力对于 DECT 基材料分解问题的适用性,使用仅包含 MAE 的损失函数验证结构相似性损失对分解结果的促进作用,并使用最大最小归一化验证所提出的归一化方法对于分解后图像质量的提升效果。



不同模块有效性的验证结果见表 5 和表 6。从表 5 和表 6 的第二列可以看出,当没有引入通道注意力机制时,模型的预测结果受限于特征提取和特征融合的能力,导致该模型在高低能图像之间的信息交换上表现不佳,影响最终分解图像的量化指数。引入通道注意力机制可以有效提取特征并进行特征融合,从而获得更准确的分解结果。进一步考虑加入 SSIM 的混合损失函数的有效性,通过比较表 5 和表 6 中的第三列数据,可以看出量化结果均有所下降,这验证了混合损

失函数在基材料分解任务中的有效性。通过结合 MAE 损失和 SSIM 损失,能够更好地优化模型,提高基材料分解结果的准确性。最后,实验验证最大最小归一化预处理方法的有效性。从表 5 和表 6 中第四列和第五列的量化结果可以看到,最大最小归一化预处理方法分解后的基材料图像的 PSNR 显著下降,这也验证了采用基于 Sigmoid 的归一化方法的有效性,通过这种预处理方法,可有效降低异常数据对分解结果的影响。

表 5 不同模型下骨骼基材料分解结果的定量评价

Table 5 Quantitative scores of decomposition results of bone-based material images under different models

unit: mean±variance

Model	(w/o) Attention	(w/o) SSIM	Max-Min normalization	MC-UCTransNet
MAE	$0.1271 \times 10^{-2} \pm 1.09 \times 10^{-8}$	$0.1772 \times 10^{-2} \pm 1.56 \times 10^{-8}$	$0.0933 \times 10^{-2} \pm 0.65 \times 10^{-8}$	$0.0924 \times 10^{-2} \pm 0.59 \times 10^{-8}$
MSE	$0.0129 \times 10^{-2} \pm 3.09 \times 10^{-10}$	$0.0123 \times 10^{-2} \pm 2.51 \times 10^{-10}$	$0.0085 \times 10^{-2} \pm 1.55 \times 10^{-10}$	$0.0098 \times 10^{-2} \pm 2.17 \times 10^{-10}$
PSNR /dB	$38.3890 \pm 0.37$	$38.5899 \pm 0.34$	$35.4132 \pm 0.50$	$39.5849 \pm 0.46$
SSIM	$0.9912 \pm 9.33 \times 10^{-7}$	$0.9917 \pm 8.66 \times 10^{-7}$	$0.9932 \pm 8.35 \times 10^{-7}$	$0.9939 \pm 6.53 \times 10^{-7}$

表 6 不同模型下软组织基材料分解结果的定量评价

Table 6 Quantitative scores of decomposition results of soft tissue-based material image under different models

unit: mean±variance

Model	(w/o) Attention	(w/o) SSIM	Max-Min normalization	MC-UCTransNet
MAE	$0.3258 \times 10^{-2} \pm 0.13 \times 10^{-7}$	$0.4766 \times 10^{-2} \pm 1.11 \times 10^{-7}$	$0.2369 \times 10^{-2} \pm 0.07 \times 10^{-7}$	$0.2650 \times 10^{-2} \pm 0.04 \times 10^{-7}$
MSE	$0.0159 \times 10^{-2} \pm 3.10 \times 10^{-10}$	$0.0278 \times 10^{-2} \pm 6.91 \times 10^{-10}$	$0.0069 \times 10^{-2} \pm 0.41 \times 10^{-10}$	$0.0061 \times 10^{-2} \pm 0.27 \times 10^{-10}$
PSNR /dB	$35.9114 \pm 0.22$	$33.4747 \pm 0.17$	$29.9182 \pm 0.16$	$39.9546 \pm 0.13$
SSIM	$0.9832 \pm 0.91 \times 10^{-6}$	$0.9826 \pm 0.92 \times 10^{-6}$	$0.9859 \pm 2.02 \times 10^{-6}$	$0.9904 \pm 0.15 \times 10^{-6}$

为了确定最佳的混合损失函数权重超参数组合,使用网格搜索法进行超参数检索,分别将  $\alpha$  和  $\beta$  作为 MAE 和 SSIM 的权重超参数。为了保证权重的相对均衡性、控制总体损失,将这两个权重超参数之和设置

为 1,同时根据经验,设置  $\alpha$  取值范围为  $[0.2, 1]$ ,  $\beta$  取值范围为  $[0, 0.8]$ 。图 11 所示为不同超参数组合下的测试集 PSNR 曲线,可以看出,当  $\alpha = 0.75, \beta = 0.25$  时,PSNR 和 SSIM 值较高。

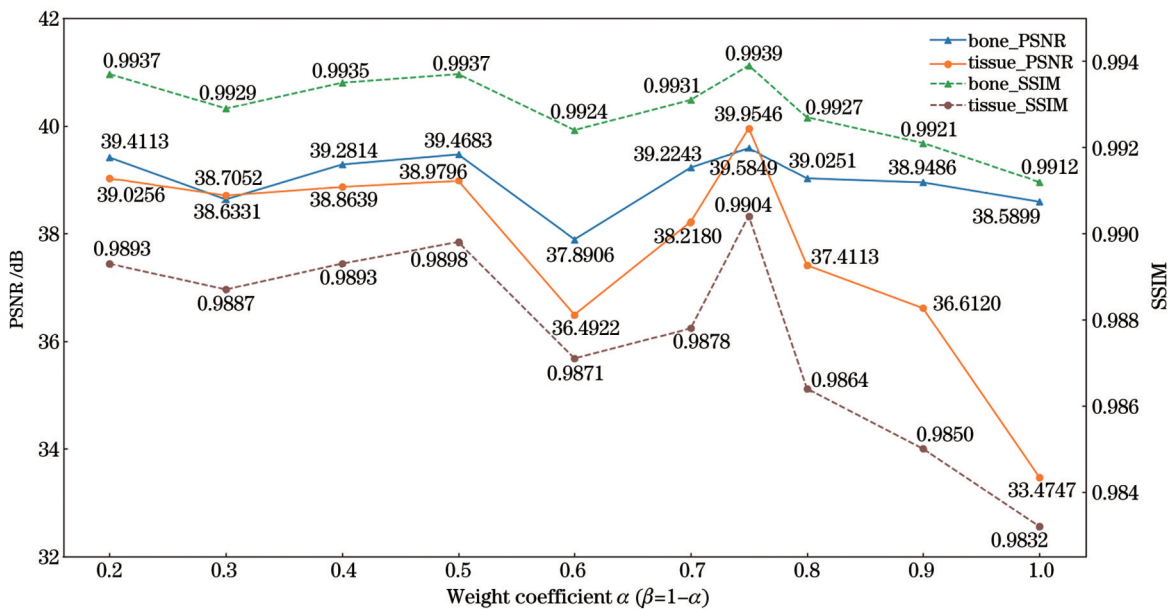


图 11 不同超参数组合下损失函数在测试集中的 PSNR 和 SSIM

Fig. 11 PSNR and SSIM in the test set for the loss function under different combinations of hyperparameters

为了确定最佳的通道注意力结构,使用不同组合的特征融合尺度对网络通道注意力模块进行测试,其量化结果如表 7 所示。可以看到,在使用所有的特征融合尺度(记为  $E_1 + E_2 + E_3 + E_4$ )的情况下可以获得最佳性能,其 PSNR 和 SSIM 分别为 39.5849 dB 和 0.9939(骨骼)、39.9546 dB 和 0.9904(软组织),也是所有模型中的最高值。当除去最低尺度的特征融合( $E_1$ )时,性能稍微下降。仅使用第二、第三与第四特征融合尺度(记为  $E_2 + E_3 + E_4$ )时,PSNR 和 SSIM 分别为 38.0509 dB 和 0.9922(骨

骼)、35.6849 dB 和 0.9857(软组织)。当仅使用最高尺度的特征融合(记为  $E_3 + E_4$ )时,量化指标仅略低于使用所有尺度的情况,其 PSNR 和 SSIM 分别达到 39.4329 dB 和 0.9934(骨骼)、38.8624 dB 和 0.9892(软组织)。当不使用通道注意力模块时,性能明显下降。在没有注意力模块的情况下,PSNR 和 SSIM 分别为 38.3890 dB 和 0.9912(骨骼)、35.9114 dB 和 0.9832(软组织)。上述实验结果表明,所使用的特征融合尺度并结合通道注意力模块可以获得更好的基材料分解性能。

表 7 不同尺度特征组合下的模型分解结果量化比较

Table 7 Quantitative comparison of model decomposition results under different scale feature combinations

Model	Bone		Soft tissue	
	PSNR /dB	SSIM	PSNR /dB	SSIM
$E_1 + E_2 + E_3 + E_4$	39.5849	0.9939	39.9546	0.9904
$E_2 + E_3 + E_4$	38.0509	0.9922	35.6849	0.9857
$E_3 + E_4$	39.4329	0.9934	38.8624	0.9892
(w/o) Attention	38.3890	0.9912	35.9114	0.9832

相关消融实验的结果表明,引入通道注意力机制能更有效地提取特征并进行特征融合,训练时混合损失函数在基材料分解任务中起到了重要作用,并且使用基于 Sigmoid 的归一化方法能够降低异常数据的影响,从而提升分解结果的准确性。这些实验结果进一步验证了所提方法的有效性。

3.5.2 网络收敛性与复杂性分析

图 12(a)所示为不同网络模型在训练集和验证集

的损失曲线。可以看到:在训练初始阶段,MC-UCTransNet 的初始损失值较低,而随着训练的进行,DIWGAN 的损失值下降最快;在收敛阶段,MC-UCTransNet 的最终收敛点损失值最小。整体来看,MC-UCTransNet 网络在训练过程中的损失曲线收敛较为平稳,具有较好的稳定性和鲁棒性,且 MC-UCTransNet 和 DIWGAN 网络的验证集损失曲线相较于其他两种网络的波动性较小,最终收敛点的损失

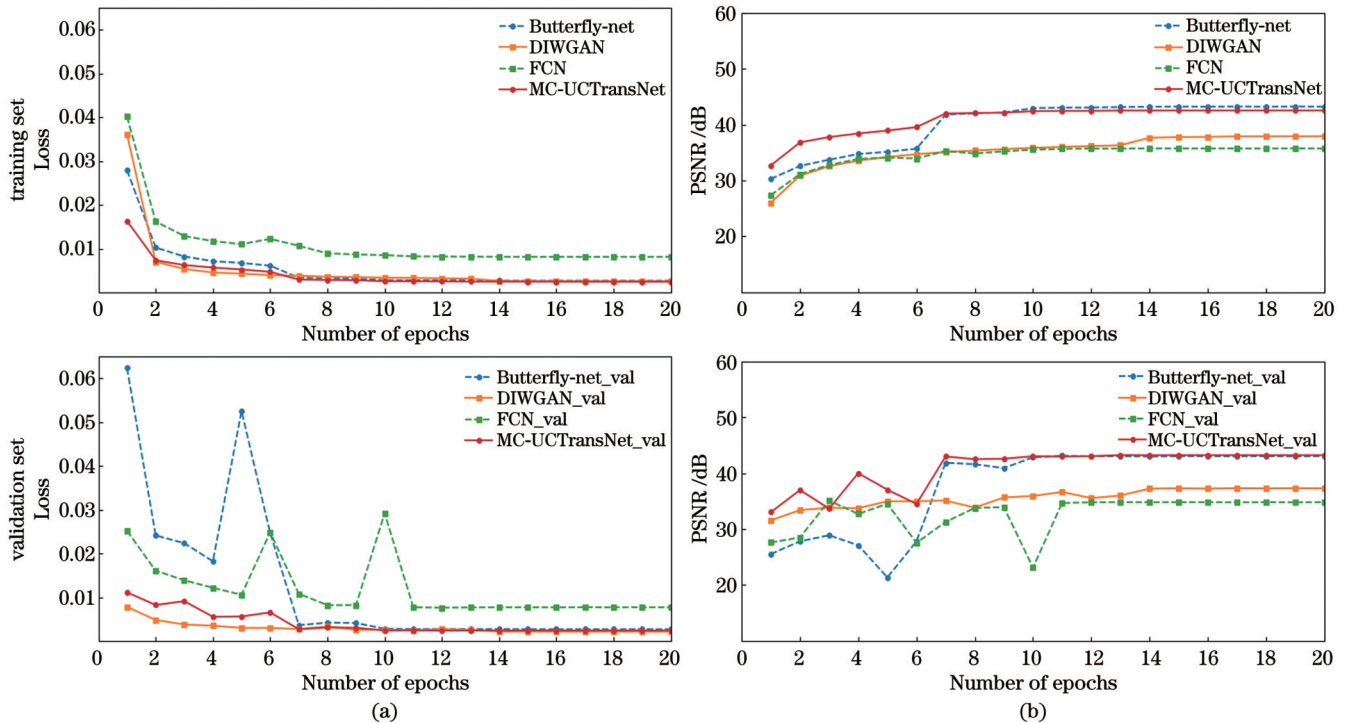


图 12 不同网络模型在训练集及验证集上的损失值和 PSNR 的收敛性情况。(a) 损失值; (b) PSNR

Fig. 12 Convergence of loss and PSNR on different networks in training and validation sets. (a) Loss; (b) PSNR



值最小。图 12(b)所示为不同网络模型在训练集和验证集的 PSNR 曲线。在训练集上, MC-UCTransNet 网络的初始 PSNR 较高, 曲线波动较小, 且收敛速度最快, 而 Butterfly-net 的最终收敛点 PSNR 值最大, 但在验证集上, MC-UCTransNet 网络的最终收敛点 PSNR

值优于 Butterfly-net。表 8 为不同方法的复杂度对比。可以看到, 虽然所提方法的参数量较大, 但其计算量仅稍微增加, 训练时间与对比算法相当。从测试计算时间来看, 所提方法的时间较长, 后续将进一步研究提高其测试效率的方法。

表 8 不同方法的复杂情况对比

Table 8 Comparison of the complexity of different methods

Method	Number of parameters	FLOP / 10 <sup>9</sup>	Train time / min	Test time / s
FCN	49	448	92	24
Butterfly-net	3	1702	106	20
DIWGAN	19	1905	128	43
MC-UCTransNet	67	354	94	110

## 4 结 论

针对 DECT 基材料分解时存在的噪声伪影大、精度低等问题, 提出一种基于多通道交叉卷积的 UCTransNet 模型, 从而优化了 DECT 基材料图像的分解效果。DECT 图像具有能谱相关性和空间相关性, 使用多通道交叉卷积方式进行特征提取, 能获得更丰富、更准确的特征表示。另外, 借助通道交叉融合变换和通道交叉注意模块, 可进一步从特征通道信息维度上提高 DECT 中的信息交互程度。相较于使用多个网络进行高低能信息特征提取与特征共享的方式, 所提网络具有更好的简洁性和交互性。与此同时, 在网络训练时, 采用混合损失函数及使用基于 Sigmoid 归一化的数据预处理方法, 可促进分解图像中的噪声伪影抑制和组织结构细节信息的保持。对比实验和消融实验的结果表明, 所提方法在噪声抑制、定量和定性评价方面具有普遍优势。后续研究可进一步提高分解效率, 同时可考虑扩展为多材料分解模型, 以扩大其应用面。

### 参 考 文 献

- [1] 王毅忠, 蔡爱龙, 梁宁宇, 等. 双即插即用能谱计算机断层重建算法[J]. 光学学报, 2023, 43(14): 1434001.  
Wang Y Z, Cai A L, Liang N N, et al. Dual plug and play multi-energy computed tomography reconstruction algorithm[J]. Acta Optica Sinica, 2023, 43(14): 1434001.
- [2] Große Hokamp N, Lennartz S, Salem J, et al. Dose independent characterization of renal stones by means of dual energy computed tomography and machine learning: an *ex-vivo* study[J]. European Radiology, 2020, 30(3): 1397-1404.
- [3] Jacobsen M C, Cressman E N K, Tamm E P, et al. Dual-energy CT: lower limits of iodine detection and quantification[J]. Radiology, 2019, 292(2): 414-419.
- [4] Chandarana H, Megibow A J, Cohen B A, et al. Iodine quantification with dual-energy CT: phantom study and preliminary experience with renal masses[J]. American Journal of Roentgenology, 2011, 196(6): W693-W700.
- [5] Wang L, Wang Y, Bian Z, et al. A nonlocal spectral similarity-induced material decomposition method for noise reduction of dual-energy CT images[J]. Journal of Southern Medical University, 2022, 42(5): 724-732.
- [6] Su T, Sun X D, Yang J C, et al. DIRECT-Net: a unified

mutual-domain material decomposition network for quantitative dual-energy CT imaging[J]. Medical Physics, 2022, 49(2): 917-934.

- [7] 符慧娟, 席晓琦, 韩玉, 等. 基于深层残差编解码的显微 CT 图像去噪算法[J]. 激光与光电子学进展, 2023, 60(14): 1410014.  
Fu H J, Xi X Q, Han Y, et al. Micro-CT image denoising algorithm based on deep residual encoding-decoding[J]. Laser & Optoelectronics Progress, 2023, 60(14): 1410014.
- [8] 龙超, 金恒, 黎玲, 等. 基于特征融合的非局部均值 CT 图像降噪[J]. 光学学报, 2022, 42(11): 1134024.  
Long C, Jin H, Li L, et al. CT image denoising with non-local means based on feature fusion[J]. Acta Optica Sinica, 2022, 42(11): 1134024.
- [9] Kelcz F, Joseph P M, Hilal S K. Noise considerations in dual energy CT scanning[J]. Medical Physics, 1979, 6(5): 418-425.
- [10] Zhao W, Niu T Y, Xing L, et al. Using edge-preserving algorithm with non-local mean for significantly improved image-domain material decomposition in dual-energy CT[J]. Physics in Medicine and Biology, 2016, 61(3): 1332-1351.
- [11] Heo S Y, An B, Kim D, et al. Feasibility study of block-matching and 3D filtering denoising algorithm in multi-material decomposition technique for dual-energy computed tomography [J]. Journal of the Korean Physical Society, 2023, 82(3): 305-314.
- [12] Lee H, Kim H J, Lee D, et al. Improvement with the multi-material decomposition framework in dual-energy computed tomography: a phantom study[J]. Journal of the Korean Physical Society, 2020, 77(6): 515-523.
- [13] Xue Y, Ruan R S, Hu X H, et al. Statistical image-domain multimaterial decomposition for dual-energy CT[J]. Medical Physics, 2017, 44(3): 886-901.
- [14] Harms J, Wang T H, Petrongolo M, et al. Noise suppression for dual-energy CT via penalized weighted least-square optimization with similarity-based regularization[J]. Medical Physics, 2016, 43(5): 2676-2686.
- [15] Li Z P, Ravishankar S, Long Y, et al. DECT-MULTRA: dual-energy CT image decomposition with learned mixed material models and efficient clustering[J]. IEEE Transactions on Medical Imaging, 2020, 39(4): 1223-1234.
- [16] 降俊汝, 余海军, 龚长城, 等. 基于双能 CT 图像域的 DL-RTV 多材料分解研究[J]. 光学学报, 2020, 40(21): 2111004.  
Jiang J R, Yu H J, Gong C C, et al. Image-domain multimaterial decomposition for dual-energy CT based on dictionary learning and relative total variation[J]. Acta Optica Sinica, 2020, 40(21): 2111004.
- [17] Ding Q Q, Niu T Y, Zhang X Q, et al. Image-domain multimaterial decomposition for dual-energy CT based on prior information of material images[J]. Medical Physics, 2018, 45(8): 3614-3626.
- [18] 陈强华, 丁锦红, 周胜, 等. 一种指数型滤波反投影解析法与

- 迭代法相结合的层析图像重建算法[J]. 激光与光电子学进展, 2022, 59(23): 2310001.
- Chen Q H, Ding J H, Zhou S, et al. Tomographic image reconstruction method combining exponential filtering inverse projection reconstruction and iterative reconstruction algorithms [J]. *Laser & Optoelectronics Progress*, 2022, 59(23): 2310001.
- [19] Lantz M, Ongie G. Learning-based material decomposition in dual energy CT using an unrolled estimator[C]//2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), April 18–21, 2023, Cartagena, Colombia. New York: IEEE Press, 2023.
- [20] 王冲旭, 陈平, 潘晋孝, 等. 基于迭代残差网络的双能CT图像材料分解研究[J]. *CT理论与应用研究*, 2022, 31(1): 47-54.  
Wang C X, Chen P, Pan J X, et al. Research on material decomposition of dual-energy CT image based on iterative residual network[J]. *Computerized Tomography Theory and Applications*, 2022, 31(1): 47-54.
- [21] Li Z P, Long Y, Chun I Y. An improved iterative neural network for high-quality image-domain material decomposition in dual-energy CT[J]. *Medical Physics*, 2023, 50(4): 2195-2211.
- [22] Xu Y F, Yan B, Zhang J F, et al. Image decomposition algorithm for dual-energy computed tomography via fully convolutional network[J]. *Computational and Mathematical Methods in Medicine*, 2018, 2018: 2527516.
- [23] Zhang W K, Zhang H M, Wang L Y, et al. Image domain dual material decomposition for dual-energy CT using butterfly network[J]. *Medical Physics*, 2019, 46(5): 2037-2051.
- [24] Kawahara D, Saito A, Ozawa S, et al. Image synthesis with deep convolutional generative adversarial networks for material decomposition in dual-energy CT from a kilovoltage CT[J]. *Computers in Biology and Medicine*, 2021, 128: 104111.
- [25] Shi Z F, Li H L, Cao Q J, et al. A material decomposition method for dual-energy CT via dual interactive Wasserstein generative adversarial networks[J]. *Medical Physics*, 2021, 48(6): 2891-2905.
- [26] Wang H N, Cao P, Wang J Q, et al. UCTransNet: rethinking the skip connections in U-net from a channel-wise perspective with transformer[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(3): 2441-2449.
- [27] Szczykutowicz T P, Chen G H. Dual energy CT using slow kVp switching acquisition and prior image constrained compressed sensing[J]. *Physics in Medicine and Biology*, 2010, 55(21): 6411-6429.
- [28] Kalender W A, Klotz E, Kostaridou L. An algorithm for noise suppression in dual energy CT material density images[J]. *IEEE Transactions on Medical Imaging*, 1988, 7(3): 218-224.
- [29] Guimaraes L S, Fletcher J G, Harmsen W S, et al. Appropriate patient selection at abdominal dual-energy CT using 80 kV: relationship between patient size, image noise, and image quality [J]. *Radiology*, 2010, 257(3): 732-742.
- [30] Rutherford R A, Pullan B R, Isherwood I. Measurement of effective atomic number and electron density using an EMI scanner[J]. *Neuroradiology*, 1976, 11(1): 15-21.
- [31] Warp R J, Dobbins J T. Quantitative evaluation of noise reduction strategies in dual-energy imaging[J]. *Medical Physics*, 2003, 30(2): 190-198.
- [32] Duan J Y, Mou X Q. Image quality guided iterative reconstruction for low-dose CT based on CT image statistics[J]. *Physics in Medicine and Biology*, 2021, 66(18): 185018.
- [33] Wang G. A perspective on deep imaging[J]. *IEEE Access*, 2016, 4: 8914-8924.
- [34] Zhao H, Gallo O, Frosio I, et al. Loss functions for image restoration with neural networks[J]. *IEEE Transactions on Computational Imaging*, 2017, 3(1): 47-57.
- [35] Yang Q S, Yan P K, Zhang Y B, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1348-1357.
- [36] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9906: 694-711.
- [37] Hubbell J H, Seltzer S M. Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z=1 to 92 and 48 additional substances of dosimetric interest[EB/OL]. (1995-01-01)[2023-10-21]. <http://physics.nist.gov/PhysRefData/XrayMassCoef/cover.html>.
- [38] 郇江磊, 林俊成, 钟丽云, 等. 基于深度学习的稀疏或有限角度CT重建方法研究综述[J]. *激光与光电子学进展*, 2023, 60(8): 0811002.  
Di J L, Lin J C, Zhong L Y, et al. Review of sparse-view or limited-angle CT reconstruction based on deep learning[J]. *Laser & Optoelectronics Progress*, 2023, 60(8): 0811002.
- [39] Wu W W, Hu D L, Niu C, et al. Deep learning based spectral CT imaging[J]. *Neural Networks*, 2021, 144: 342-358.
- [40] Niu T Y, Dong X, Petrongolo M, et al. Iterative image-domain decomposition for dual-energy CT[J]. *Medical Physics*, 2014, 41(4): 041901.

## Dual-Energy CT Base Material Decomposition Method Based on Multi-Channel Cross-Convolution UCTransNet

Wu Fan<sup>1</sup>, Jin Tong<sup>1</sup>, Zhan Guorui<sup>1</sup>, Xie Jingjing<sup>1</sup>, Liu Jin<sup>1,2\*</sup>, Zhang Yikun<sup>2,3</sup>

<sup>1</sup>College of Computer and Information, Anhui Polytechnic University, Wuhu 241000, Anhui, China;

<sup>2</sup>Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing 210096, Jiangsu, China;

<sup>3</sup>Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, Jiangsu, China

### Abstract

**Objective** Dual-energy computed tomography (DECT) is a medical imaging technology that provides richer tissue contrast and material decomposition capabilities by simultaneously acquiring X-ray absorption information at two different

energy levels, and it is increasingly widely used. In DECT, based on the energy absorption differences of different materials, the scanning objects can be decomposed into different base material components, such as bone and soft tissue. However, accurate decomposition and reconstruction of base material images remain a challenging problem due to factors such as noise, artifacts, and overlap. Therefore, we aim to improve the quality and accuracy of base material decomposition in DECT imaging. Current base material decomposition methods may have some limitations in complex scenarios, such as the failure to accurately decompose overlapping materials, vulnerability to noise interference, and poor image quality. To solve these problems and improve the properties of base material decomposition, a new base material decomposition method is proposed in this study.

**Methods** We aim to improve the quality and accuracy of base material decomposition in DECT images. To achieve this goal, we propose a method based on the multi-channel cross-convolutional UCTransNet (MC-UCTransNet), which is performed by fitting the mapping function in DECT. The network is designed to be a double-in-double-out architecture based on UCTransNet. During training, with the real decomposition image as labels, a pair of double energy images as input, and its concatenating into the form of multi-channel, our multi-channel network structure aims to realize the information exchange between two material generation paths in the network. The channel cross-fusion converter and channel cross-attention module are used to improve the decomposition of base materials, realizing double-input-double-output and end-to-end mapping. Further, the channel cross-fusion module and the channel cross-attention module can better capture the complex channel correlation to more fully conduct feature extraction and fusion and realize the information exchange between the generation paths of base materials. To improve the model fitting performance, the network is trained using a hybrid loss. Meanwhile, in order to better adapt to the particularity of CT image data, the model uses the normalization method based on the Sigmoid function to preprocess the network input data and improve the model fitting performance.

**Results and Discussions** In order to verify the decomposition accuracy of each method, we not only compare the base material images decomposed by various methods but also reconstruct the base material images to the low energy image, and we compare them with the original low energy image. By obtaining the difference map to intuitively feel the decomposition effect of each method, the experimental results show that the proposed method is able to obtain images of water and soft tissue. Compared with the contrast method, the decomposed images perform better in accuracy and noise contrast suppression. Meanwhile, the results of the ablation experiments also demonstrate the attention mechanism, the mixed loss, and the effectiveness of the Sigmoid normalization method in this task. The introduction of the attention mechanism enables the network to better capture the information of key features in the image and improves the accuracy of decomposition. The mixed loss function of mean absolute error (MAE) and structural similarity index measure (SSIM) is used to improve the network decomposition effect and performance. In addition, the application of the Sigmoid normalization method can better adapt to the particularity of CT image data. On the premise of maintaining the distribution characteristics of the data, the interference of abnormal data to the model can be reduced, and the stability and accuracy of the model can be improved. The loss and peak signal-to-noise ratio (PSNR) values of the proposed method are superior in both the training and validation sets, with fast convergence and good stability, as well as a good decomposition effect on different test sets, showing strong generalization ability. This indicates that the dual energy-based MC-UCTransNet method has high utility in the base material decomposition task of DECT imaging.

**Conclusions** We aim to improve the quality and accuracy of base material decomposition in DECT, and remarkable progress is made by proposing a dual material decomposition method based on MC-UCTransNet. Our study innovatively adopts the MC-UCTransNet network to integrate multi-channel cross-convolution with cross-attention mechanism modules to better capture the correlation among complex channels and realize information exchange between generation pathways of base materials. Moreover, the multi-channel cross structure avoids the use of multi-network for high and low energy information extraction, which makes the network model more convenient. In addition, we further improve the fitting performance of the model by the use of mixed loss and normalization methods based on the Sigmoid function. The experimental results show that the proposed method can ensure a promising improvement in water bone-based material and soft tissue iodine-based material decomposition tasks.

**Key words** machine vision; dual-energy computed tomography; base material decomposition; multi-channel cross convolution; attention; noise suppression