

# 基于注意力门 UNet 网络的 CT 金属伪影去除方法

师晓宇, 王斌

(中北大学 信息与通信工程学院, 太原 030051)

**摘要:** 目前, UNet 基本模型对带有金属伪影的 CT 图像的去除能力无法有效满足需求, UNet 的结构简单无法提取出足够精确的有效结构和细节信息, 并且深层卷积对低级特征的信息利用不够充分; 针对上述问题, 提出了一个基于注意力门的 UNet 金属伪影去除网络, 该网络采用了注意力门对低层级和高层级的信息进行注意力权重处理, 并利用跳跃连接机制到特征解码结构以提高生成 CT 图像的质量, 通过多层级的编解码结构得到最终的去除金属伪影 CT 图像; 实验结果表明, 该方法在视觉上取得了更好的条状和带状伪影去除效果的 CT 图像, 并在 PSNR 指标上取得了 35.5913, 在 FSIM 指标上取得了 0.9613, 在 SSIM 指标上取得了 0.9288 的成绩; 与 ADN、cGANMAR、UNet、CNNMAR、CycleGAN 等目前已有的方法相比, 该方法在诸多方面均取得了显著的优势。

**关键词:** 金属伪影去除; 注意力机制; 卷积神经网络; UNet; 编解码

## Metal Artifact Reduction for CT Method Based on Attention Gate UNet

SHI Xiaoyu, WANG Bin

(School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

**Abstract:** Currently, the basic UNet model cannot effectively meet the demand for CT images with metal artifact reduction, the simple structure of the UNet cannot precisely extract the accurate information on the effective structure and details, and deep convolutional neural network does not sufficiently use the information of low-level features. Based on the above problems, a metal artifact removal network with attention gates based on the UNet is proposed. The network adopts the attention gates to process the attention weights of the information at low and high levels, The jump connection mechanism and feature decoding structure are used to improve the quality of the generated CT images, the final CT images with metal artifact reduction are obtained through the multilevel encoding and decoding structure. The experimental results show that the proposed method achieves better visual effects in removing stripe and banding artifacts in CT images, with a PSNR of 35.5913, FSIM of 0.9613, and SSIM of 0.9288. Compared to existing methods such as the ADN, cGANMAR, UNet, CNNMAR, and CycleGAN, the proposed method has significant advantages in multiple aspects.

**Keywords:** metal artifact reduction; attention mechanism; convolutional neural network; UNet; encoder-decoder

## 0 引言

金属伪影是计算机断层扫描 (CT, computed tomography) 中常见的问题之一。当患者携带金属植入物 (例如牙科填充物和髓关节假体) 时就会出现这种情况。与身体组织相比, 金属材料在光谱上会造成显著的 X 射线衰减, 导致 X 射线投影不一致。不匹配的投影将会在重建得到的 CT 图像中引入十分明显的条纹和阴影伪影, 这会显著降低图像质量并影响医学图像分析工作以及随后的医疗服务。这些金属伪影会掩盖植入物周围组织的重要诊断信息, 严重影响医疗效果, 比如导致医生难以对靶区进行精准勾画, 进而对临床治疗效果产生诸多负面影响<sup>[1]</sup>。因此, 金属伪影去除 (MAR, metal artifact reduction) 对提高临床诊断的准确性有着重要意义。

近年来, 许多传统方法提出了 MAR 任务, 主要可分为三类, 即迭代重建、正弦图域 MAR 和图像域 MAR。迭代

算法旨在设计一些手工制作的正则化器, 例如总变分<sup>[1]</sup>和小波域<sup>[2]</sup>中的稀疏约束, 并将它们表示为算法优化以约束解空间。由于主观的先验假设, 这些方法无法精细地表示临床应用中复杂多样的金属制品。基于正弦图域的方法将受金属影响的区域 (即正弦图中的金属迹线) 视为缺失数据, 并通过线性插值<sup>[3]</sup>或先前图像的正向投影<sup>[4]</sup>来填充这些区域。然而, 金属痕迹中的这些替代数据通常不能正确满足 CT 成像几何约束, 这会导致重建 CT 图像中出现由于金属植入物影响的二次伪影。齐宏亮等人<sup>[5]</sup>提出了一种基于先验图像的 CT 插值校正算法, 其中对原始投影数据进行插值后使用滤波反投影算法进行图像域重建后得到去伪影后的图像, 为了避免初步校正产生的次级伪影, 又提出了一种全新的滤波算法。这种算法能够得到较好的重建图像, 但是由于先验图像的质量限制以及滤波算法无法进行自适应迭代, 相对于深度学习去伪影模型鲁棒性并不突出。

收稿日期: 2023-09-05; 修回日期: 2023-10-14。

基金项目: 山西省基础研究计划项目 (202203021211100, 202103021224204); 山西省专利转化专项计划项目 (202302006)。

作者简介: 师晓宇 (1998-), 男, 硕士研究生。

引用格式: 师晓宇, 王斌. 基于注意力门 UNet 网络的 CT 金属伪影去除方法[J]. 计算机测量与控制, 2024, 32(4): 219-225.

随着近年深度学习 (DL, deep learning) 在医学图像处理领域的快速发展, 基于深度学习的金属伪影校正方法也相继提出并取得了良好的临床效果。卷积神经网络 (CNN, convolutional neural networks) 在公开数据集的图像分析性能表现优于传统方法<sup>[4-7]</sup>, 同时比图形切割和多图集分割技术<sup>[8]</sup>快一个数量级。U-Net<sup>[9]</sup>, DeepMedic<sup>[10]</sup> 和整体嵌套网络<sup>[11-12]</sup>等全卷积网络 (FCN, fully convolutional networks)<sup>[13]</sup>已被证明可以在包括心脏磁共振在内的各种任务中实现稳健和准确的性能。UNet 是使用最广泛的深度 CNN 之一, 特别是在医学影像方面体现除了良好的效果<sup>[14]</sup>。Xu 等人<sup>[15-16]</sup>使用 VGG (visual geometry group) 网络, Wang 等人<sup>[17]</sup>用条件生成对抗网络 (cGAN, conditional generative adversarial network), Ding 等人<sup>[18]</sup>使用 FCN<sup>[13]</sup>分别进行了图像域的 MAR 方法研究? 这类方法在对射线状伪影及椒盐噪声时展现了出色的效果, 但对于带状伪影性能表现较差。史再峰等人<sup>[19]</sup>提出了多任务学习的生成对抗网络, 利用连续伪影图像的空间相关性和解剖组织相似性并行处理学习, 训练共享编码器和多个解码器重建出高质量的 CT 图像。Liao 等人<sup>[20]</sup>提出无监督的伪影解纠缠网络 (ADN, artifact disentanglement network), 对图像域中正常组织与金属伪影进行分离以达到去除伪影的效果? 但是由于金属伪影构成的复杂性, 该方法在分离组织与金属伪影的过程中性能不稳定, 加之在该方法中存在网络结构繁纷复杂, 后续优化难度较大, 提升效果有限。由于深度学习到的鲁棒特征表示, 基于 DL 的 MAR 技术通常优于基于手工特征的传统方法。然而, 现有的基于 DL 的 MAR 方法仍然存在着一定的局限性。他们中的大多数将 MAR 视作一般的图像恢复问题, 较少强调在整个学习过程中嵌入的物理几何约束。大多数现有方法依赖于现成的 DL 工具包来构建不同的网络架构, 对于特定的 MAR 任务缺乏足够的主要信息关注度, 在恢复图像细节过程中造成混入背景信息以及组织保留程度不足的问题。

为了解决网络去金属伪影中图像恢复信息不足, 背景

信息与主要信息混杂导致组织信息保留不足的问题。本文提出了一种注意力门的 UNet 网络的医学 CT 图像金属伪影去除模型, 能够在去除复杂 CT 图像金属伪影的同时有效保留金属植入物周围的解剖组织结构。

### 1 网络框架

基于 UNet 的 CT 图像金属伪影去除网络总体框架如图 1 所示, 网络整体分为两个模块, 分别为特征提取模块和强化特征解码模块。输入图像为大小  $256 \times 256$  的 CT 灰度图像。在特征编码模块设置为 4 层, 每一层分别有两组卷积层, 每层卷积层由大小  $3 \times 3$  的卷积核和线性整流函数 (ReLU, rectified linear unit) 激活函数构成。在经过每一层后都由大小  $2 \times 2$  的上采样进行特征扩充后进入下一层重复操作到最深层。强化特征解码模块与特征提取模块相似, 在接受深层提取信息的过程前会经过相同层级特征提取模块的注意力信息进行叠操作以强化信息提取编码的准确度, 实现更有效的特征提取。

UNet 网络结构作为经典的编解码器结构, 编解码信息利用跳跃连接进行沟通从而避免大量空间的精准信息丢失的问题, 直接将编码器中提取的特征合并至解码器相对应的层中。但是其中存在着初始编码器低阶层中由于提取的特征不精确, 从而导致相应叠加的解码器层中存在很多的冗余信息, 降低了网络的效果。为了避免这样的情况出现, 在 UNet 的基础上加入了软注意力结构<sup>[21]</sup>的注意力门模块。意在通过注意力门模块抑制无关区域像素的干扰, 突出特定局部区域的显著特征。使用软注意力结构代替硬注意力通过神经网络可以计算梯度并且前向传播和后向反馈来学习得到注意力的权重。软注意力结构的另一个好处在于其集成到标准 UNet 网络结构中时要简单方便、计算开销小, 更加可贵的是可以提高模型的灵敏度和预测的精度。解码器经过上采样恢复出位置细节, 但是上采样同样会导致边缘细节过平滑和位置关系不准确的问题。现有的一些工作<sup>[9]</sup>使用跳跃连接机制将底层特征与高层特征连接以补充位置信息。由于低层特征提取包含了无用的背景信息, 此

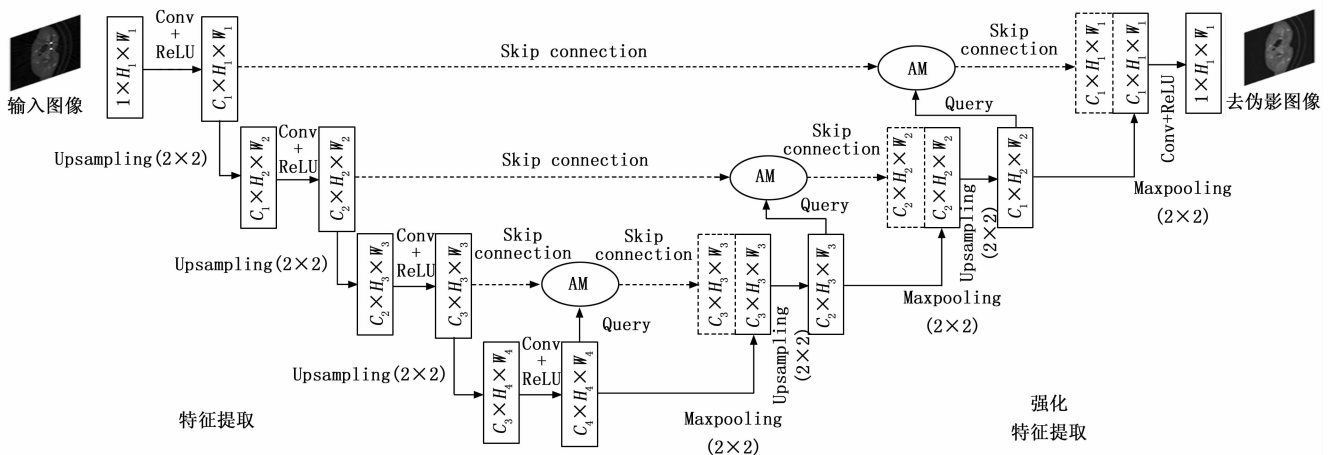


图 1 网络整体框架

信息反而会导致目标对象的提取精度受到影响。为了加强模型对目标的提取效果, 设计了一种为捕获高级语义信息并强调目标特征的注意力门模块。

注意力门控机制是一种通过动态调整模型对输入信息的关注度来提升模型性能和效率的方法。在深度学习中, 注意力门控机制常用于处理序列数据或图像数据, 帮助模型在处理和理解数据时, 更加关注输入的重要部分, 并忽略不重要的部分。注意力门控机制的核心思想是通过一个称为“门控”的机制来控制信息的流动。在最常见的形式中, 门控机制采用自注意力机制, 通过计算输入信息中每个位置的表示, 并将这些表示应用于输入信息中的每个位置, 以获得新的、经过注意力处理的信息。具体来说, 自注意力机制首先对输入信息进行线性变换, 得到一个表示输入信息的矩阵。接着, 这个矩阵被分成多个独立的头, 每个头都会独立地计算输入信息的注意力表示。最后, 这些注意力表示被加权求和, 得到最终的注意力表示。这个表示被应用于输入图像的每个像素, 以产生新的、经过注意力的信息。

## 2 注意力门

### 2.1 参数更新

为了捕获足够大的感受野并因此提取附近组织信息, 在 UNet 架构中逐渐对特征映射网络进行下采样。这样, 粗空间网格级模型的特征就可以定位与背景与组织的关系。通过将注意力门 (AG, attention gate) 加入 UNet 模型中同样可以实现定位组织与背景, 并且可以加强分割效果, 其中注意力门避免了训练多种模型以及由此带来的参数量膨胀的问题。AG 能够在训练中降低对图像中不相关的背景响应, 而避免了手动裁剪感兴趣区域 (ROI, region of interest)。

从粗尺度提取的信息用于 AG, 以消除跳跃连接中不相关和嘈杂的区域, 在跳跃连接操作之前执行仅合并相关的注意力区域, 并且在前向传播和后向反馈来学习得到注意力的权重, 这使得较低层中的模型参数能够基于 ROI 进行学习更新。

第  $l-1$  层中参数的更新由式 (1) 所示:

$$\frac{\partial (\hat{x}_i^l)}{\partial (\Phi^{l-1})} = \alpha_i^l \frac{\partial [f(x_i^{l-1}; \Phi^{l-1})]}{\partial (\Phi^{l-1})} + \frac{\partial (\alpha_i^l)}{\partial (\Phi^{l-1})} x_i^l \quad (1)$$

其中:  $x_i$  为图像像素矢量,  $\Phi$  为  $1 \times 1$  卷积,  $\alpha_i^l$  为第  $l$  层第  $i$  个注意力矢量。

在多层 AG 结构中,  $\alpha_i^l$  对应于像素的矢量。在每个子 AG 中, 经过池化操作后通过通道注意力向量和  $1 \times 1$  的卷积分别提取高层和底层特征, 再进行相加以融合信息。使用  $1 \times 1$  的卷积进行线性变换减少训练参数量, 降低 AG 的计算复杂性, 输入的特征映射适配输入向量的大小, 类似于非本地块 (Non-Local Block)。相应的线性变换对特征映射进行解码, 并将其映射到上一层进行门操作。最低层次的特征映射, 即第一个跳跃连接由于其中没有高层级的特征信息, 不用于门控功能, 使用深度监督令中间的特征图

在每个图像尺度上具有组织与背景高维信息的区分性。这有助于确保不同尺度的 AG 模块能够正确识别大范围图像背景。由此可以避免密集的预测结果从过小的跳跃连接中影响导致特征信息提取不足的问题出现。

每个通道对应于特定的语义响应。金属伪影和人体组织通常涉及不同的通道。AG 模块因此对语义依赖关系进行特征提取以强调目标通道。AG 利用深层中的抽象特征和浅层中的全局特征以编码依赖关系。深层特征映射含有丰富的语义信息, 可用于指导浅层特征映射以选择重要的位置详细信息。此外, 浅层映射利用整体图像信息对不同通道的语义关系编码, 可以实现过滤干扰信息。通过使用语义关系信息, AG 模块可以强调目标区域并改进特征表示。AG 模块如图 2 所示。

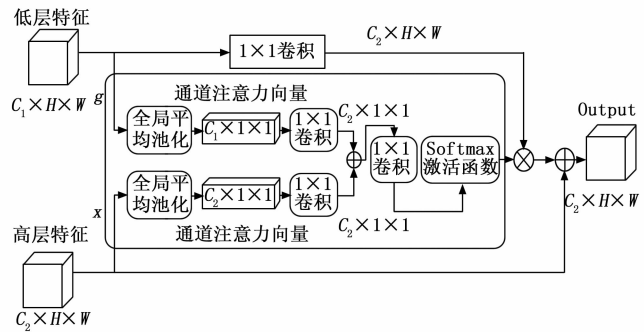


图 2 注意力门模块

### 2.2 池化操作

高级特征映射和低级特征映射执行全局平均池化。其将整体信息压缩成一个具有信息权重的注意力向量以强调关键特征并过滤干扰背景。生成注意力矢量如下:

$$F_a(x, y) = \delta_1 [W_a g(x) + b_a] + \delta_1 [W_b g(y) + b_b] \quad (2)$$

$$A_c = \delta_2 [W_\varphi F_a(h, l) + b_\varphi] \quad (3)$$

其中:  $h$  和  $l$  分别为深层和浅层映射。  $g$  为全局平均池化。  $\delta_1$  为线性整流 (ReLU, rectified linear unit) 激活函数,  $\delta_2$  为 softmax 函数。  $W_a, W_b, W_\varphi$  是指  $1 \times 1$  卷积的参数。  $b_a, b_b, b_\varphi$  为偏差。

$$g(x_n) = \frac{1}{W \times H} \sum_{i=1}^H \sum_{j=1}^W x_n(i, j) \quad (4)$$

其中:  $n = 1, 2, \dots, c$  和  $x = [x_1, x_2, \dots, x_c]$ ,  $W, H$  为图像宽和高。

我们对矢量进行  $1 \times 1$  卷积, 以实现特征依赖关系的进一步确定。具体而言,  $1 \times 1$  卷积核可以在空间维度上将输入特征映射的尺寸减小, 同时保持深度维度不变。这有助于在不改变特征通道数的情况下, 对特征图的感受野进行调整和改变。接着使用 softmax 函数对矢量进行归一化激活。softmax 函数可以将输入值映射为概率分布, 使得输出值的和为 1。这使得模型能够将输入特征映射为相对权重的表示, 进而实现特征的合理融合。然后将浅层特征映射与注意力向量相乘以生成注意特征映射。具体而言, 我们将注意力向量视为权重系数, 对浅层特征映射进行加权求

和。这种操作可以使得模型能够根据浅层特征映射中的不同特征，自主地选择关注哪些特征并抑制其他不重要的特征。最后通过添加深层特征映射来校准所关注的关键特征。这一步骤有助于将浅层特征映射中未考虑到的信息，如空间信息等，融入深层特征映射中。同时，也有助于提高模型的表达能力和泛化性能。此外还使用了全局平均池化和  $1 \times 1$  卷积来实现该模块。全局平均池化可以将输入特征图的尺寸减小至  $1 \times 1$ ，并将每个像素点的值压缩为一个标量。这种操作可以有效地降低计算复杂度，并且能够避免过拟合现象的产生。而  $1 \times 1$  卷积则可以在不改变特征通道数的情况下，实现特征的重新组合和利用。相较于其他模块，该模块并没有添加大量额外参数，因此不会导致模型复杂度的显著增加。同时，由于全局平均池化的使用，该模块的计算成本也得到了大大降低。此外，该模块还具有较好的泛化性能和表达能力，能够有效地提升模型的性能和效果。

### 3 训练过程与损失函数

网络训练过程采用 Pytorch 框架，迭代 epoch 为 500 次，每次 epoch 迭代 300 轮。批处理尺寸为 4，使用 Adam 优化网络， $\beta_1 = 0.6$ ， $\beta_2 = 0.999$ ，初始学习率为 0.000 01，每训练 20 个 epoch 学习率下降为原来的 1/2。实验环境：操作系统为 Ubuntu20.04，Python 版本 3.7，使用 Pytorch 框架，硬件采用 GPU：NVIDIA GeForce RTX2070 8 GB。

为了得到更加贴近主观视觉的 MAR 图像，设计了基于图像主观视觉的损失函数，在保证图像还原质量情况下充分保留原本的组织细节。该损失函数由灰度损失和总变分损失。

灰度损失  $L_{MSE}$ ，为生成样本与真实样本之间的均方误差，可以使生成样本图像尽可能贴近真实样本，如式 (5) 所示：

$$L_{MSE} = \frac{1}{N^2} \|X_{free} - Y\|_2^2 \quad (5)$$

其中： $N$  为训练集中样本图像对数； $X_{free}$  为生成的无伪影图像， $Y$  为真实无伪影图像。

总变分损失  $L_{TV}$ ，由式 (6) 表示：

$$L_{TV} = \frac{1}{H \times W} \|\nabla_x(x) + \nabla_y(x)\|_2 \quad (6)$$

式中， $H$ 、 $W$  分别为图像的高度和宽度； $\nabla_x$ 、 $\nabla_y$  分别为图像在横纵坐标上的变分。

总损失函数  $L_{total}$  由式 (9) 表示：

$$L_{total} = L_{MSE} + \alpha L_{TV} \quad (7)$$

其中： $\alpha$  为权重用于调节效果，初始取值为 0.2。

### 4 实验结果与分析

本文采用 DeepLesion 数据集<sup>[22]</sup>生成的图像验证所提方法的正确性和有效性，将所提方法与 CGANMAR<sup>[23]</sup>，CycleGAN<sup>[24]</sup>，CNNMAR<sup>[2]</sup>，UNet<sup>[9]</sup>，ADN<sup>[20]</sup> 进行对比实验。

#### 4.1 数据集准备

利用文献[2]的方法，基于 DeepLesion 医学公开数据集<sup>[25]</sup>选取生成了 3 040 对金属伪影对照图像用于网络训练和测试，选取 300 对进行测试，图像为大小  $256 \times 256$  的灰度图像。在对比实验中对选取的 300 对待测试灰度图像进行统一测试，得到结果图像进行标准客观标准定量对比，结果图像统一窗宽在  $-800 \sim 1\,000$  HU 间，以方便观察比较。

#### 4.2 评价指标

数据集中包含配对数据，可进行定量和定性评估。为了定量分析 CT 图像的金属伪影校正指标，采用峰值信噪比 (PSNR, peak signal to noise ratio)、结构相似性指数 (SSIM, structural similarity index measurement)、空域视觉信息保真度 (VIFs, spatial domain visual information fidelity)、特征相似度 (FSIM, feature similarity index measurement) 作为定量评估指标。

PSNR 是衡量图像失真程度或噪声水平的定量指标，使用 PSNR 对网络去除 CT 金属伪影效果进行评估，数值越高说明效果越好。MSE 是待测图像  $x$  与标签图像  $y$  的均方误差，分别为：

$$MSE = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} [x(i,j) - y(i,j)]^2 \quad (8)$$

$$PSNR = 10 \cdot \log_{10} \left( \frac{(2^n - 1)^2}{MSE} \right) \quad (9)$$

式中， $i, j$  为像素值， $W, H$  为图像的宽和高； $n$  为图像比特数。

SSIM 是用来衡量两幅图像的结构相似程度，越接近 1 说明结构相似度越高，计算方式如式 (10)：

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (10)$$

式中， $\mu_x, \mu_y$  为图像亮度均值， $\sigma_x, \sigma_y$  为  $\mu_x, \mu_y$  的标准差， $C_1, C_2$  为对比度。

FSIM 是用来衡量两幅图像特征相似度，其更加关注图片中界定物体的结构的边缘像素而降低了背景区域的像素的影响，更有助于反映主体的图像质量，其指标越接近 1 表明两幅图像的特征相似度越高，其计算方式如式 (11)：

$$FSIM = \frac{\sum_{x \in I} S_L(x) PC_m(x, y)}{\sum_{x \in I} PC_m(x, y)} \quad (11)$$

其中： $S_L(x)$  为相位一致性特征相似度与梯度一致性特征相似度融合的相似度， $PC(x, y)$  为相位两张图像的相位一致性特征相似度， $I$  为完整图像像素域。

VIFs 数值越高，结果图像和目标图像之间相关性就越强，指标越接近 1 效果越好，具体公式参考 Bovik 的计算方法<sup>[26]</sup>。

#### 4.3 实验结果分析

##### 4.3.1 定性分析

为了展示本文方法处理伪影图像、去除图像金属伪影恢复图像细节的性能，随机从测试集中选取了 3 组测试对

比结果图。能够从图 3 中观察到, CycleGAN<sup>[24]</sup> 作为一种无监督方法在无监督领域表现突出, 但是这种模型还要求伪影校正输出能够转换回原始伪影影响图像, 这虽然有助于保留内容信息但是同样鼓励了模型保留图像中的金属伪影结构, 影响了模型最终的重建图像质量。ADN 作为一种比 CycleGAN 先进的无监督学习方法, 其生成的图像质量高于 CycleGAN 但是相比有监督方法仍然在伪影的去除方面存在着去除条纹不均匀, 细节保留不充分的不足, 如图 4、5 所示, 伪影没有被有效地去除或减少。CNNMAR 是基于投影插值的方法。CNNMAR 能够较好地保留条纹但是结构性信息恢复严重不足。我们还发现, UNet 的效果接近于 cGANMAR 的效果, cGANMAR 在其后端使用类似 UNet 的体系结构。由于使用 GAN 结构网络, cGANMAR 产生了比 UNet 更锐利的输出图像。UNet 和 cGANMAR 在图像中都显示出了良好的效果, UNet 产生的图像伪影去除效果较好但是结构信息保留细节不够; cGANMAR 在感官上生成图像比较锐利并较好地去除了伪影亮条纹, 但是在处理暗条纹情况仍然有所欠缺。作为改进的 UNet 模型, 在单独 UNet 的基础上加入了注意力门模块, 从表中可以观察到在结构指标以及峰值信噪比等均优于 UNet。相比于 UNet 输出图像的整体清晰度下降的结果得到较高的 PSNR, 在图 3 在注意力门模块的加入后, 本文方法得到了相比 UNet 更加清晰且对比度更高的去金属伪影图像。从图 4、5 中可以观察到, 在有复杂金属植入物的情况下, 上述几种方法的性能有了比较明显的下降, 两种无监督方法 CycleGAN 和 ADN 去除条纹伪影的效果较差, CycleGAN 还产生了新的暗条纹伪影, CNNMAR 由于受到先验图像训练效果的制约, 恢复了一定的组织结构但伪影没有能够很好地去除且过平滑问题较为严重, 影响了图像质量。CGANMAR 得到了对比度较高的图像, 但是同样没有很好地去除金条纹伪影。UNet 去除伪影的效果较为明显但组织结构的信息保留并不充分, 在图像的部分区域出现了过平滑的问题。由于注意力门模块的加入, 增强了 UNet 网络提取主干信息的能力, 通过较粗粒度的特征图, 获得语义上下文信息, 进而对来自同层的 encoder 的特征图中不相关的特征进行抑制, 提高了模型对目标区域的灵敏度和准确性。本文方法在所示图像中均显示了优于上述方法图像的性能, 能够在较好去除金属伪影的同时保留组织细节结构。在图 4 所示的金属伪影原始图像中可以看出, 本图像中的金属植入物严重影响了 CT 扫描获得的图像质量, 对于细节恢复工作提出了很大的挑战, CycleGAN 恢复出的图像带有明显的金属条纹伪影, CNNMAR 恢复出的图像出现了严重的过平滑问题, 导致结构细节损失严重, 严重影响了重建图像的可用性, cGANMAR 在图像对比度方面取得了较好的视觉效果, 但是在图 4 所示的 ROI 区域内, 出现了结构丢失的问题, 影响了图像的质量, UNet 在 HU 值较低的区域丢失了较多的细节, 导致重建图像的质量下降, 这可能与 UNet 在设计损

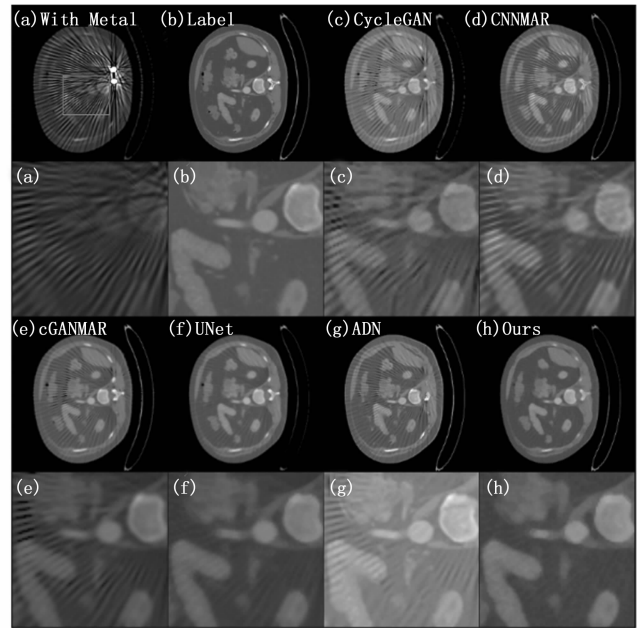


图 3 腹部 CT 图像 ROI 区域金属伪影去除结果

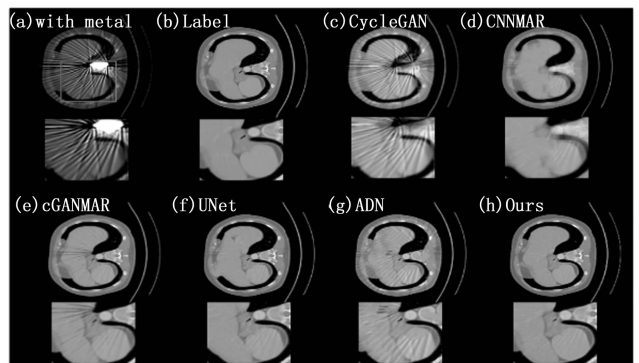


图 4 腹部 CT 图像金属伪影去除结果

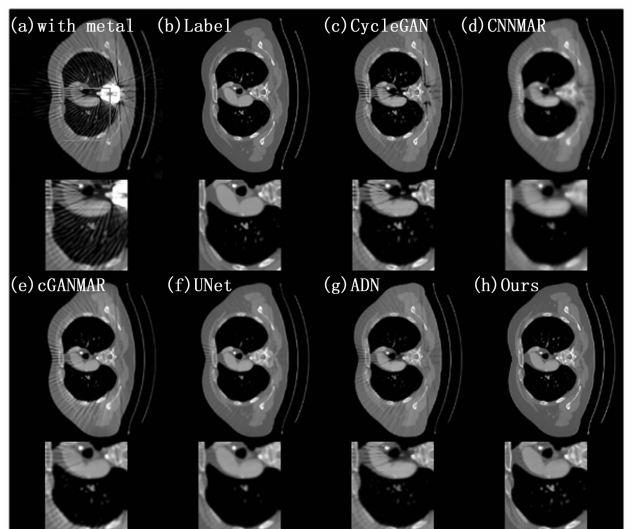


图 5 胸部 CT 图像金属伪影去除结果

表 1 测试集金属伪影去除结果平均指标对比

客观指标	校正前	cGANMAR	CNNMAR	CycleGAN	UNet	ADN	Ours
SSIM	0.495	0.817 2	0.845 9	0.876 2	0.909 2	0.885 4	0.928 8
FSIM	0.716 1	0.915 3	0.899 4	0.919 9	0.947 2	0.942 9	0.961 3
VIFs	0.230 4	0.544 6	0.575 3	0.628 7	0.662 6	0.735 9	0.698 2
PSNR/dB	21.217 8	32.604 4	31.136 8	31.832 0	35.223 8	33.198 1	35.591 3

失函数中只使用了均方差损失有关,均方差损失在最小化误差的训练过程中会忽略标签与输入较小误差的像素,从而鼓励平滑输出结果;在 ADN 取得的结果图中可以观察到其结果相对于 CycleGAN 的结果有一定的提高,但是由于无监督学习的不确定性,产生的图像质量仍然存在着不足。在图 5 所示的模型结果对比图中可以看到,由于金属植入物的位置 CT 细节受到严重干扰,产生了亮度值过大的区域,对于该区域的细节恢复同样是一大挑战。与上图类似,CycleGAN 保留了细节的重建图像对比度较高但是存在比较明显的金属伪影,CNNMAR 在恢复复杂图像的问题上平滑效果明显影响了图像可用度,cGANMAR 在 HU 值较低的位置去伪影效果较好,捕捉了图像细节,但是在组织结构明显的区域内没有取得理想的效果,UNet 仍然存在组织细节丢失与组织平滑的问题,此图中,ADN 在较低 HU 值时取得了较好的质量,同样留存了明显的条纹伪影,本文提出的方法在视觉效果上均超越了前述方法,取得了较为可用的重建图像。

#### 4.3.2 定量分析

表 1 展示了前述方法在选取的图像评价指标中所取得的结果。所示指标中 UNet 由于监督学习的学习方式,在优良先验图像和选取的合适的目标函数共同校正模型训练情况下,取得了一定的成绩,基于 UNet 模型和注意力门控模块的本模型同样取得了优良的成绩,在 PSNR 上取得了高于 UNet 的 35.591 3 的成绩,在 PSNR 的评价标准下已经达到了人眼分辨困难的程度。在 SSIM 和 FSIM 两种结构相似性评价指标中分别取得了 0.928 8 和 0.961 3 的成绩,在总体结构性和主要结构部分相似性得到了很好的效果。而其余方法在评价指标中并不理想,与前述定性分析结论相一致。

## 5 结束语

本文提出一种基于 UNet 的带有注意力门架构的 MAR 的模型。模型充分利用注意力门结构强大的细节提取能力进行 CT 图像的组织结构细节提取和恢复。在验证实验过程中,定性评估结果与定量结果指标均表明基于此监督数据集训练出的注意力 UNet 模型可以成功抑制金属伪影,并在重建图像中能够有效保留金属周围的组织结构和细节,与同类方法相比有着更好的性能。

#### 参考文献:

[1] GJESTEBY L, DE MAN B, JIN Y, et al. Metal artifact reduction in CT: where are we after four decades? [J]. IEEE Access, 2016, 4: 5826 - 5849.

[2] ZHANG Y, YU H. Convolutional neural network based metal artifact reduction in x-ray computed tomography [J]. IEEE Transactions on Medical Imaging, 2018, 37 (6): 1370 - 1381.

[3] KALENDER W A, HEBEL R, EBERSBERGER J. Reduction of CT artifacts caused by metallic implants [J]. Radiology, 1987, 164 (2): 576 - 577.

[4] ZHANG H, WANG L, LI L, et al. Iterative metal artifact reduction for x-ray computed tomography using unmatched projector/ backprojector pairs [J]. Medical Physics, 2016, 43 (6Part1): 3019 - 3033.

[5] 齐宏亮, 凌庆庆, 李翰威, 等. 基于先验图像的 CT 图像金属伪影校正算法 [J]. 北京生物医学工程, 2023, 42 (4): 384 - 389.

[6] KHENED M, KOLLERATHU V A, KRISHNAMURTHI G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers [J]. Medical Image Analysis, 2019, 51: 21 - 45.

[7] LIAO F, LIANG M, LI Z, et al. Evaluate the malignancy of pulmonary nodules using the 3-Deep leaky noisy-or network [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30 (11): 3484 - 3495.

[8] WOLZ R, CHU C, MISAWA K, et al. Automated abdominal multi-organ segmentation with subject-specific atlas generation [J]. IEEE Transactions on Medical Imaging, 2013, 32 (9): 1723 - 1730.

[9] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C] //Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, Proceedings, Part III 18. Springer International Publishing, 2015: 234 - 241.

[10] KAMNITSAS K, LEDIG C, NEWCOMBE V F J, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation [J]. Medical Image Analysis, 2017, 36: 61 - 78.

[11] LEE C Y, XIE S, GALLAGHER P, et al. Deeply supervised nets [C] //Artificial Intelligence and Statistics. PMLR, 2015: 562 - 570.

[12] XIE S, TU Z. Holistically-nested edge detection [C] //Proceedings of the IEEE International Conference on Computer Vision, 2015: 1395 - 1403.

[13] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3431 - 3440.

- [14] BAI W, SINCLAIR M, TARRONI G, et al. Human-level CMR image analysis with deep fully convolutional networks [J]. Arxiv Preprint ArXiv: 1710.09289.
- [15] XU S, DANG H. Deep residual learning enabled metal artifact reduction in CT [C] //Medical Imaging 2018: Physics of Medical Imaging. SPIE, 2018, 10573: 950-955.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Arxiv Preprint ArXiv: 1409.1556, 2014.
- [17] WANG J, ZHAO Y, Noble J H, et al. Conditional generative adversarial networks for metal artifact reduction in CT images of the ear [C] //Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference, Granada, Spain, Proceedings, Part I. Springer International Publishing, 2018: 3-11.
- [18] DING X, WANG Y, XU Z, et al. CcGAN: Continuous conditional generative adversarial networks for image generation [C] //International Conference on Learning Representations, 2021.
- [19] 史再峰, 王仲琦, 王 宁, 等. 基于对抗式多任务学习的医学 CT 图像金属伪影去除方法 [J]. 南京大学学报 (自然科学), 2021, 57 (3): 482-492.
- [20] LIAO H, LIN W A, ZHOU S K, et al. ADN: Artifact disentanglement network for unsupervised metal artifact reduction [J]. IEEE Transactions on Medical Imaging, 2019, 39 (3): 634-643.
- [21] SCHLEMPER J, OKTAY O, SCHAAP M, et al. Attention gated networks: learning to leverage salient regions in medical images [J]. Medical Image Analysis, 2019, 53: 197-207.
- [22] YAN K, WANG X, LU L, et al. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning [J]. Journal of Medical Imaging, 2018, 5 (3): 036501.
- [23] WANG J, ZHAO Y, NOBLE J H, et al. Conditional generative adversarial networks for metal artifact reduction in CT images of the ear [C] //Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference, Granada, Spain, Proceedings, Part I. Springer International Publishing, 2018: 3-11.
- [24] IKUTA M, ZHANG J. A deep recurrent neural network with primal-dual optimization for CT metal artifact reduction [C] //Medical Imaging 2022: Image Processing. SPIE, 2022, 12032: 408-414.
- [25] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C] //Proceedings of the IEEE International Conference on Computer Vision, 2017: 2223-2232.
- [26] SHEIKH H R, BOVIK A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15 (2): 430-444.
- [15] 曾德松, 高 琛, 谭北海, 等. 自动驾驶仿真测试场景库构建研究 [J]. 现代计算机, 2021, 27 (31): 1-12.
- [16] 王 可, 陆 键, 蒋思明. 基于车辆行驶轨迹的道路不良驾驶行为谱构建与特征值计算方法 [J]. 交通运输工程学报, 2020, 20 (6): 236-249.
- [17] 陆 键, 王 可, 蒋思明. 基于车辆行驶轨迹的道路不良驾驶行为实时辨识方法 [J]. 交通运输工程学报, 2020, 20 (6): 227-235.
- [18] 何青俊, 涂辉招, 魏海然. 基于驾驶模拟器的跟驰行为谱解析 [J]. 武汉理工大学学报 (交通科学与工程版), 2021, 45 (3): 419-423.
- [19] KIM S, COIFMAN B. Assessing the performance of Speedlnfo radar traffic sensors [J]. ITS Journal, 2017, 21 (1/2/3/4/5/6): 179-189.
- [20] KONG K, KAVALER R, RAJAGOPAL R, et al. Real-time measurement of link vehicle count and travel time in a road network [J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 11 (4): 814-825.
- [21] 刘元昊, 张 昱, 刘杲朋, 等. 基于雷视一体机的交通流数据采集系统设计 [J]. 计算机测量与控制, 2022, 30 (3): 161-167.
- [22] 朱红赛. 基于 ARM-Linux 的无人驾驶车载传感器数据采集与处理系统 [D]. 南京: 南京理工大学, 2021.
- [23] 张 艺, 严翌瑄, 李 静. 基于多传感器融合的交通数据采集系统概述 [J]. 物联网技术, 2021, 11 (2): 15-18.
- [24] 陈骏生. 自动驾驶电动车辆数据采集方法 [J]. 机电技术, 2021 (2): 22-23.
- [8] 冯 屹, 王 兆. 自动驾驶测试场景技术发展与应用 [J]. 道路交通管理, 2021 (3): 90.
- [9] 郑建明, 张建军, 覃 斌, 等. 自动驾驶场景大数据高精度采集技术研究 [C] //2021 中国汽车工程学会年会论文集 (1). 上海, 2021: 8.
- [10] LI L, HUANG W L, LIU Y, et al. Intelligence testing for autonomous vehicles: a new approach [J]. IEEE Transactions on Intelligent Vehicles, 2017, 1 (2): 158-166.
- [11] RSENER C, FAHRENKROG F, UHLIG A, et al. A scenario-based assessment approach for automated driving by using time series classification of human-driving behaviour [C] //IEEE. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). New York: IEEE, 2016: 1360-1365.
- [12] GAO F, DUAN J, HE Y, et al. A test scenario automatic generation strategy for intelligent driving systems [J]. Mathematical Problems in Engineering, 2019: 1-10.
- [13] ULBRICH S, MENZEL T, RESCHKA A, et al. Defining and substantiating the terms scene, situation, and scenario for automated driving [C] //IEEE 18th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2015: 982-988.
- [14] BACH J, OTTEN S, SAX E. Model based scenario specification for development and test of automated driving functions [C] //Intelligent Vehicles Symposium. IEEE, 2016: 1149-1155.