

基于极坐标拉伸的CT胃癌关键特征增广方法

张贻钦¹, 陈庆奎^{1*}, 黄 陈², 张正杰³, 陈美玲¹, 付直兵¹

¹上海理工大学光电信息与计算机工程学院, 上海

²上海市第一人民医院, 胃肠外科, 上海

³复旦大学附属华山医院, 放射科, 上海

收稿日期: 2025年3月22日; 录用日期: 2025年4月15日; 发布日期: 2025年4月23日

摘 要

目的: 缓解CT影像的切面空间定义域与神经网络特征提取范围不匹配的问题。资料与方法: 在临床中收集了7年跨度内被诊断为胃癌的病人在诊疗过程中生成的895例CT影像资料, 清洗出689个高质量序列, 由第一名医师确定最大癌症负载横断面并进行标注, 并由另一名医师进行审核, 组建实验数据集。引入一种基于拉伸的图像增广方法, 对齐CT横断面定义域与神经网络特征提取范围, 最大化利用CT影像序列的定义域外空间进行数据增广, 提高训练效率。结果: 在使用所提出的增广方法后, 神经元前向响应与反向梯度的均匀度得到提升。同等复杂度下, 更少的迭代轮次即能达到更优的效果。结论: 所提出的增强方法可以提高神经网络在CT模态上的计算效率, 在临床识别任务中有一定应用潜力。

关键词

数据增广, 胃癌成像, 神经网络, 数据效率

Enhancing Key Features of Gastric Cancer in CT Modality Using Polar Coordinate Stretching

Yiqin Zhang¹, Qingkui Chen^{1*}, Chen Huang², Zhengjie Zhang³, Meiling Chen¹, Zhibing Fu¹

¹School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

²Department of Gastrointestinal Surgery, Shanghai General Hospital, Shanghai

³Department of Radiology, Huashan Hospital, Fudan University, Shanghai

Received: Mar. 22nd, 2025; accepted: Apr. 15th, 2025; published: Apr. 23rd, 2025

Abstract

Purpose: To address the mismatch between the slice spatial domain of CT images and the feature

*通讯作者。

文章引用: 张贻钦, 陈庆奎, 黄陈, 张正杰, 陈美玲, 付直兵. 基于极坐标拉伸的 CT 胃癌关键特征增广方法[J]. 建模与仿真, 2025, 14(4): 736-748. DOI: 10.12677/mos.2025.144325

extraction range of neural networks. **Materials and Methods:** A total of 895 CT imaging scans from patients diagnosed with gastric cancer over a span of 7 years were collected in clinical practice. After data cleaning, 689 high-quality sequences were obtained. The maximum cancer load slice was determined and annotated by the first physician and reviewed by another physician to form the experimental dataset. A stretching-based image augmentation method was introduced to align the slice domain of CT with the feature extraction range of neural networks, maximizing the use of the extra-domain space of CT image sequences for augmentation and improving training efficiency. **Results:** After applying the proposed augmentation method, the uniformity of neuronal forward responses and backward gradients was enhanced. Under the same complexity, fewer iterations were required to achieve better results. **Conclusion:** The proposed augmentation method can improve the computational efficiency of neural networks on CT modality and has potential for application in clinical cancer detection tasks.

Keywords

Data Preprocessing, Gastric Cancer Imaging, Neural Networks, Data Efficiency

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

深度学习模型在医疗领域取得了显著进步，尤其是自动癌症病变分割与识别技术的广泛应用，这不仅减轻了放射科医生的工作压力，还提升了诊断的准确性[1]-[3]。然而，最新型的数据驱动模型需要大量数据才能达到理想效果，这在医学研究中可能导致高昂的成本[4]。在计算机视觉领域，主流的研发范式都会采用多种数据预处理手段对样本进行增强，帮助神经网络更高效地学习[5]。

在胃癌 CT 影像分析领域，常用的预处理手段包括：自由形变[6]与仿射[7]、擦除、曝光调整、多窗口控制[8] [9]、前后景替换[10]。这些方法主要对 CT 重建序列中每个横断切面(Slice)的 Hounsfield (HU) 值[11]进行调整，以此生成新的增强样本，并配合相同的标签进行训练[12]。这种做法能够改善神经元对于同一个样本的特征提取深度，但往往需要将同一个源样本变换后分多次输入给神经网络用于训练，因此这些方法可被认为是对数据集的扩充[13] [14]。在相同的数据集遍历次数下，神经网络的训练用时和数据集的样本数成正比，因此这些方法总体上是使用时间来换取模型精度。

本文着眼于利用 CT 模态中横断面影像的独特空间定义域，对影像进行基于拉伸的增广，使处理后的图像及其特征能更均匀地覆盖神经网络的特征提取计算范围，提高训练效率。

2. CT 数据与增广理论

2.1. 数据收集

胃癌 CT 影像序列数据集从上海市第一人民医院胃肠外科的临床诊疗实践中收集。一名胃肠外科的医生为每位患者的影像序列标记了包含最大胃癌病灶面积的切面，另一名影像科医师在该切片上进行癌症区域注释，形成像素级分割基准。胃肠外科医生对标注结果进行审查，并进行修改。这意味着不论 CT 扫描序列有多少个 Slice，总是只有一个 Slice 被标注。在标注过程中，影像科医师和胃肠外科医生对所有的扫描序列和分割结果都最终达成了一致，最终在 895 个病例中选出了 689 例用于进一步的研究，序列清洗标准[15]如下：

- 伪影严重，或是其他成像效果不满意的情形
- 对应的病理结果不可用，无法确定病灶性质的
- 未完整覆盖所有病灶组织的
- 病灶区域小，无法判定最大截面积的
- 患者或家属在本研究的任一阶段提出异议的
- Dicom 元数据不可用的

在这样的构建流程下，所得到的图像层数和标注层数严重不平衡，数据集是少样本的，绝大多数研究在该种情况下均会采用多样化的增广手段加入预处理过程。存储方面，影像序列以 DICOM 文件形式存储，每个序列均带有规范化的元数据。医师的标注结果则以 nrrd 格式独立存储，附加记录了切面原点坐标，使用该坐标来和 dicom 序列进行对应，确定该标注所属的截面。所收集的 895 例影像序列包含 0.625 mm、1 mm、1.25 mm、5 mm 三种重建层厚，切片 spacing 则没有统一的标准，所有序列都以 $2\text{mm} \times 2\text{mm} \times 2\text{mm}$ 的体素尺寸被重采样，并进行后续实验处理。

被纳入数据集的病人在 2014~2021 年间曾在或仍在上海市第一人民医院胃肠外科就诊。就诊过程中，在院内使用了相同的一台 GE Revolution CT 机对患者进行扫描，成像区域包括：胸部、下腹部、全腹部，形成了 120 层~750 层不等的序列扫描结果。在 pTNM 分期上，有 1a 期占 8%、1b 期占 7%、2a 期占 9%、2b 期占 15%、3a 期占 24%、3b 期占 20%、3c 期占 17%、4 期患者仅有 2 例。首次诊断时，腺癌占绝大多数(74%)。这一比例来自于其在接受 CT 扫描之后的第一次诊断分期，不代表其在接受 CT 扫描时的分期。所有病例中，男性占 72%，年龄平均数为 67 岁，中位数为 68 岁。

本项研究中所使用的所有样本都进行了匿名化处理并遵从了患者意愿。该研究已获得上海市第一人民医院人体试验伦理审查委员会的批准(批准编号[2024] 032)。

2.2. 特征提取计算和 CT 横断面的定义域不对齐问题

设 CT 横断面图像为一个二维矩阵 $I \in \mathbb{R}^{w,h}$ ，其中 w, h 分别为横断面的宽度与高度，其定义域为 $A = \{(x, y) | x^2 + y^2 \leq r^2\}$ ，其中 r 为内切圆半径。DICOM 文件使用方形矩阵来存储该数据，其定义域为 $B = \{(x, y) | 0 \leq x \leq w, 0 \leq y \leq h\}$ 。在特征提取过程中，整个定义域 B 会被计算，这一过程可用公式 1 表达：

$$F(i, j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} I(i+m, j+n) \cdot K(m, n) \quad (1)$$

其中 K 表示特征提取泛函。当 $(i+m, j+n) \notin A$ 时， $I(i+m, j+n)$ 是无效值，对应的特征提取结果 $F(i+m, j+n)$ 也是无效的，这部分无意义的计算限制了模型的计算效率，在图 1 中，HU 域中的圆外黑色部分表示无效值域，也代表着无效特征提取发生的范围。

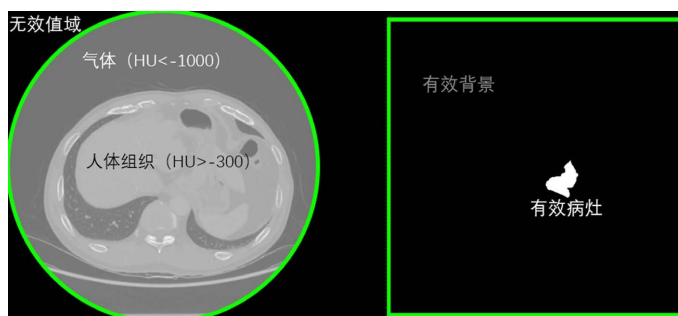


Figure 1. The definition domain of the HU domain (circle) and the annotation domain (square) do not match

图 1. HU 域(圆形)和标注域(方形)的定义域不匹配

需要注意的是,虽然下文提出的方法主要针对于卷积网络进行设计[16],但以 ViT 为设计范式的所有 Transformer 视觉神经网络也有类似的问题[12]。在这一类网络中,图像会被切分为多个 patch,若 patch 中仅包含无意义值时,任何对该 patch 进行的计算都被认为是无效的,不过注意力机制使其能降低模型对于这些 patch 的关注度。因此在精度上,它们受到的影响不会像卷积网络受到的那么多,但其计算性能依旧会被降低。

2.3. 基于仿射拉伸的 CT 横断切面增广

为了解决上述问题,本文引入了一种针对 CT Slice 特点设计的仿射映射方式(如图 2),其将存储着 Slice 信息的正方形数据矩阵进行变换,使正方形矩阵的最大内切圆外的区域也能够存储有效的 HU 值。这样,底层的卷积网络遍历到曾经无效的数值区域时,也能进行有意义的计算,进而提升了在整个神经网络处理过程中,模型进行神经运算的有效率,并最终体现在提升模型的精度指标上。

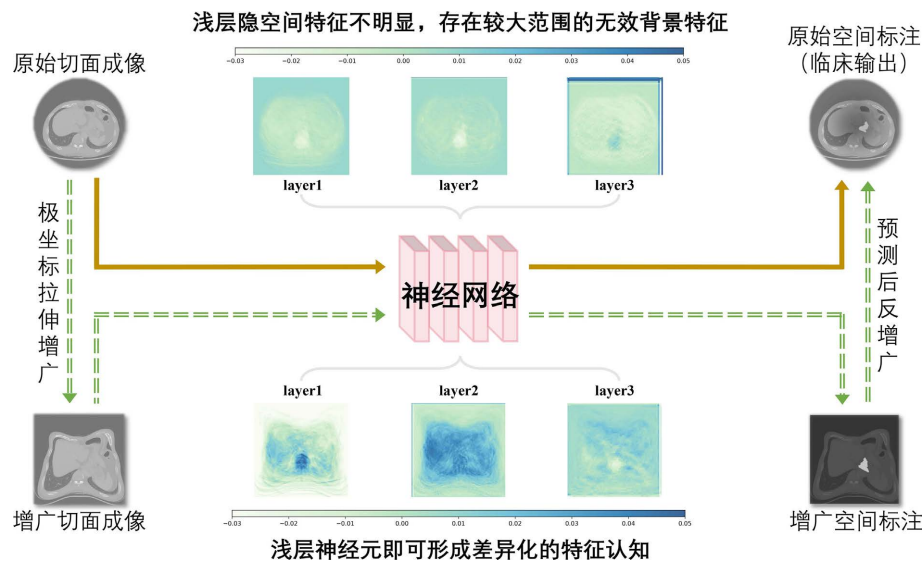


Figure 2. An overview of the proposed stretching augmentation
图 2. 本文提出的拉伸增广的概览

在这一过程中,原始重建 Slice 在形态学上被不可避免地破坏了,模型也将输出基于变换后的体素位置的预测。对于临床场景而言,这一情形是不可接受的,因为医生通常只能基于标准的人体 CT 扫描结果进行分析,最多也只会进行窗宽窗位限制,形态学上的变化将使专业医师无法根据既有经验和直觉作出准确诊断,并且也会大大限制该方法在医疗情形下的可信度。因此一种仿射反变换重采样方式被引入,它将对模型的预测输出进行处理,使其恢复到原始 CT Slice 被重建的形态学结构上,并将处理后的预测结果和原始 Slice 进行叠加,形成最后的可视化结果。在进行反变换后,医师对神经网络推理中涉及的仿射变换和反变换是不可感知的,所有模块被封装在一个完整的服务中,无需外部干预。这样,所提出的方法只会在后端进行特征提取时生效,临床使用者无需进行任何额外的训练,或是实践上的额外学习,即可轻便部署所提出的方法。

从数学角度讲,首先定义 CT Slice 中心点为 (y_c, x_c) , 其对应的极坐标系表达为 $(r(x, y), \theta(x, y))$ 。针对任意一个 Slice 像素,在极坐标系中使用公式(1), (2)确定径向拉伸系数,对其进行坐标映射。

$$\theta' = \frac{\pi}{8} [1 - \cos(4\theta)] \quad (1)$$

$$f(\theta) = \frac{1}{[\cos(\theta')]^\alpha} \cdot \beta \quad (2)$$

其中 $\alpha \geq 0$ 为边角拉伸强度系数, $\beta \geq 1$ 为全局缩放系数。当 $\alpha = 0$ 时, $f(\theta) = \beta$, 退化为均匀缩放。当 α 增大时, 边缘区域 $\left(\theta \rightarrow \frac{k\pi}{4}, k \in \mathbb{Z}^+\right)$ 的 $\cos(\theta')$ 值减小, 导致 $f(\theta)$ 增大, 对应方向上的所有像素被重映射至远离极点的位置, 填充到背景像素区域, 并以此降低在后续神经网络计算中神经元的无效计算。当需要执行反变换时, 使用 $f^{-1}(\theta)$ 作为径向拉伸因子再次执行拉伸映射即可。此时 $0 \leq f^{-1}(\theta) \leq 1$, 故在直观上反变换过程呈现为四角内缩的趋势。当 α 或 β 被设为一个过大的值时, 重映射后的像素位置可能位于矩阵以外, 而无法被记录。因此, 在确定参数时, 应当确保 $f\left(\frac{\pi}{4}\right) \leq \sqrt{2}$ 。

在信息论上, 这样做具有一定的理论基础。在 CT 图像中, 边缘区域的信息熵较低, 因为边缘有效像素的分布相对稀疏。在使用所提出的方法后, 可以增加边缘区域的信息熵, 使图像的整体信息分布更加均匀。这有助于提高神经网络对图像特征的提取能力, 因为在同一层图像或特征图上的多次卷积操作能接收到更均匀的信息密度, 在反向计算时避免梯度被稀释。

2.4. 癌组织定位与关键小样本生成

在临床中, 胃癌的病灶多数聚集在胃壁上, CT 横断面视图可见胃壁异常增生[9][17][18], 病灶所占横断面的比例通常较低, 这是因为人体内部与胃相连的还有肺、脾、胰、肝等器官, 会占据横断面上更大的面积[11]。因此, 胃癌的相对特征较弱, 直接拟合时, 损失值易被背景区域稀释, 导致拟合效率差。为了解决上述问题, 在使用 2.3 节的方法对 CT 横断面进行增广后, 本文进一步引入了关键特征定位算法。

具体来说, 首先基于标注信息定位每个胃癌病灶的最小内接正方形, 将此正方形的边长扩大 10%, 以避免在卷积计算时边缘关键信息无法被充分提取。此时, 所得正方形范围内的特征主体为癌症病灶, 以此区域作为基本单位进行训练时, 损失函数会分配更多强度在前景上。综合 2.3 与 2.4 节的方法, 可以生成多样化的增强小样本, 图 3 展示了在不同的增广强度下, 使用关键小样本生成的病灶伪样本示例。在典型的深度学习分割任务的训练范式中, 使用预测特征图和目标图计算所得的二维损失值矩阵会被直接求平均, 这意味着所有的像素都以相同的权重参与到神经网络的梯度计算中去, 优化器在此基础上进行对神经网络进行参数更新。在使用 2.3 节提出的方法后, 越靠近 $\theta = k\pi/4, k \in \mathbb{Z}^+$ 方向的有效图像特征会被越显著地稀疏化, 这意味着图像中该方向上的语义信息熵较低。这说明典型的实现方法并不很好地匹配使用增广之后的图像语义。从最优化学习效率的角度讲, 应当按照图像中的语义密集度来分配参数更新权重。训练时, 本文在切面维度上引入了基于拉伸强度的损失权重分配算法, 其在样本增广时记录拉伸参数属性, 并配套生成一个非均匀权重图, 权重图在极坐标系上遵循 $\mathbf{W}_\theta = f^{-1}(\theta)$ 。该权重图会与损失值矩阵进行逐元素点乘, 而后对损失值矩阵取平均并执行反向传播。此时, 由于梯度和神经网络响应具有正比关系, 反向传播得到的梯度已综合了逐元素的先验语义, 能够差异性地赋予不同极角方向以不同的拟合强度。

这种方法能够生成密集的前景特征, 在预测过程中, 神经网络同样需要能够有效地识别背景特征, 否则模型将更容易产生伪正判断。因此, 需要在横断面上进行随机子区域提取, 并作为独立的样本加入训练。在选定随机子区域时, 正方形形状相同于该横断面上所确定的胃癌最小内接正方形(图 4)。

显然, 生成的小样本具有和原始 CT 切面不一致的分辨率, 若将其与横断面图像直接混合将导致尺度不统一, 破坏分布一致假设, 打乱模型对于特征尺度的学习过程。因此, 本文将关键小样本作为前置

预训练环节的输入,目的是使模型能够先“预习”胃癌的基本形态学特征、HU 强度特征等。在“再训练”阶段,模型在已有一定基础知识把握的情况下对 CT 横断切面的全局语义进行学习。

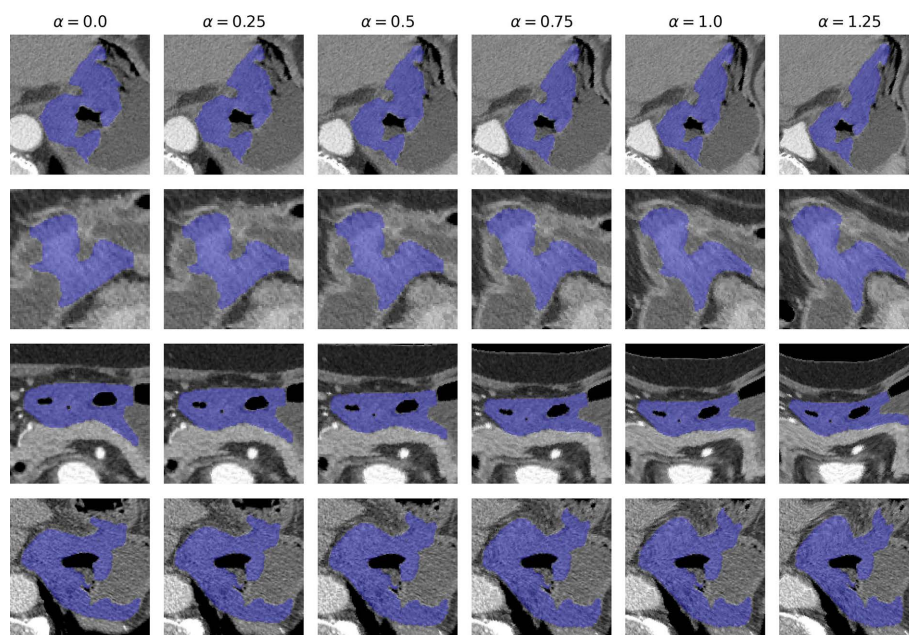


Figure 3. The augmented gastric cancer lesions

图 3. 增广后的胃癌病灶

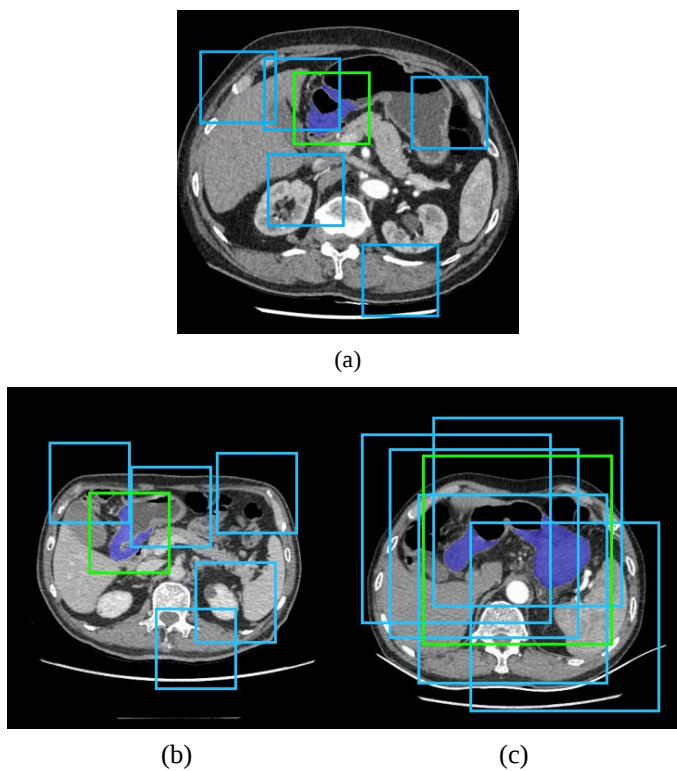


Figure 4. Examples of selected cross-sectional sub-regions (green indicates key features)

图 4. 选定的横断面子区域示例(绿色为关键特征)

3. 胃癌识别与分析的建模

为了应用本文所提出的方法，首先需要在临床中收集 CT 重建序列 V_i ，由医师标注得到最大胃癌负载横断面 S_i ，在该横断面建立极坐标系，用 $S_i(r, \theta)$ 可表示该横断面上任意一点。使用公式(3)所描述的过程对坐标系进行映射，得到 S'_i 。

$$\mathcal{F}_{\alpha, \beta}(\theta) = \frac{\beta}{\left[\cos\left(\frac{\pi}{8} [1 - \cos(1 - \cos(4\theta))]\right) \right]^\alpha} \quad (3)$$

$$S'_i(r', \theta) = S_i(r \cdot \mathcal{F}_{\alpha, \beta}(\theta), \theta)$$

公式(3)定义了本文使用的映射方法连续域上的表达形式，由于 CT 重建序列时在空间上是离散采样的，且以像素点存储，因此需要使用离散重采样方法来生成新的横断面图像。需要保证标注采样过程不会引入原始数据集中不存在的新值，因此对图像域和标注域的离散重采样方法是不一致的，如公式(4)所述：

$$S'_{img,i}(r', \theta) = \sum_{m=-1}^1 \sum_{n=-1}^1 S_{img,i}(r \cdot \mathcal{F}(\theta) + m, \theta + n) \cdot W(m, n) \quad (4)$$

$$S'_{ann,i}(r', \theta) = S_{ann,i}(r \cdot \mathcal{F}(\theta), \theta)$$

其中， img 代表 HU 域， ann 代表标注域， W 代表采样权重。

拉伸增广数据集可使用 $D = \{S'_i | i \in [0, \text{Samples}]\}$ 表示。对 D 中的每个样本，使用公式(4)定位癌症最小内接正方形，并裁剪，得出关键胃癌特征小样本 m_i ， L_m 代表样本 m 的最小内接正方形边长。

$$H = \left\{ \left(\frac{m_{hmin} + m_{hmax}}{2}, \frac{m_{hmin} + m_{hmax}}{2}, 1.1 \cdot L_h \right) \middle| h \in D \right\} \quad (4)$$

在得到了增广数据集 D 和关键小样本数据集 H 后，使用算法 1 所述的过程进行训练和推理。

算法 1. 神经网络的训练与推理过程

```

训练：
初始模型  $\mathcal{M} \in \{\text{SegNeXt}, \text{MedNeXt}, \text{DaTransUNet}\}$ 
While True:
 $\mathcal{M}_{pre} = \arg \min_{\mathcal{M}} L(\mathcal{M}, H)$ 
 $\mathcal{M}_{re}(t) = \arg \min_{\mathcal{M}} L(\mathcal{M}_{pre}, D)$ 
if (for  $t > 10$ :  $\text{Acc}(\mathcal{M}_{re}(t)) \leq \text{Acc}(\mathcal{M}_{re}(t-1))$ ):
break
推理：
 $P_i^{avg}(r', \theta, \alpha_k) = \frac{1}{K} \sum_{k=1}^K \mathcal{M}_{re}(\{Aug(S_i, \alpha_k) | i \in \{1, 2, \dots, K\}\})$ 

```

算法 1 得到的 P_i^{avg} 为模型对输入 CT 序列 S_i 的胃癌识别结果，且两者具有相同的形状。

在实验部分中，不仅需要对端到端任务的分割精度进行分析，同时也需要对神经隐空间进行分析，从定性的角度分析所提出的方法是否改善了特征提取效率(图 5)。目前在分割领域主要有两种编解码器范式，分别是对称型和非对称型，两者的编码器具有相似的层级结构，但解码器差异较大。具体来说，第 l

层的编码器隐空间可以使用公式(5)表示:

$$h^{(l)} = \mathcal{M}_{enc}^{(l)}(h^{(l-1)}) \quad (5)$$

编码器的隐空间特征 $h^{(L)}$ 通过解码器 $\mathcal{M}_{dec}$ 映射到最终的分割结果 P_i :

$$P_i = \mathcal{M}_{dec}(h^{(L)}) \quad (6)$$

因此, 编码器的隐空间特征与输出的二类分割结果之间存在联系。

在反向传播中, 反向梯度 g 是通过对所有神经隐层的梯度乘积计算得来的:

$$g_{i,j} = \frac{\partial L}{\partial S_{i,j}} = \mathcal{M}_{(qi,uj)}^T * \frac{\partial L}{\partial P_{q,u}} \quad (7)$$

其中, S 是输入横断面数据, P 是对应横断面的标注, i, j 是输入数据的二维索引, q, u 则是标注域的二维索引。

从该公式中可以得出, 梯度图 g 中每个索引 (i, j) 与损失图在 (q, u) 位置的值存在联系, 关联性由神经网络具体参数决定。也就是说, $g_{i,j}$ 体现了输入中的哪些神经计算对模型的预测输出有所贡献, 这为本文 4.4 节奠定了理论基础。图 5 介绍了所述的两种不同的编解码器架构范式, 梯度观察发生在从编码器至解码器的各层残差连接中。

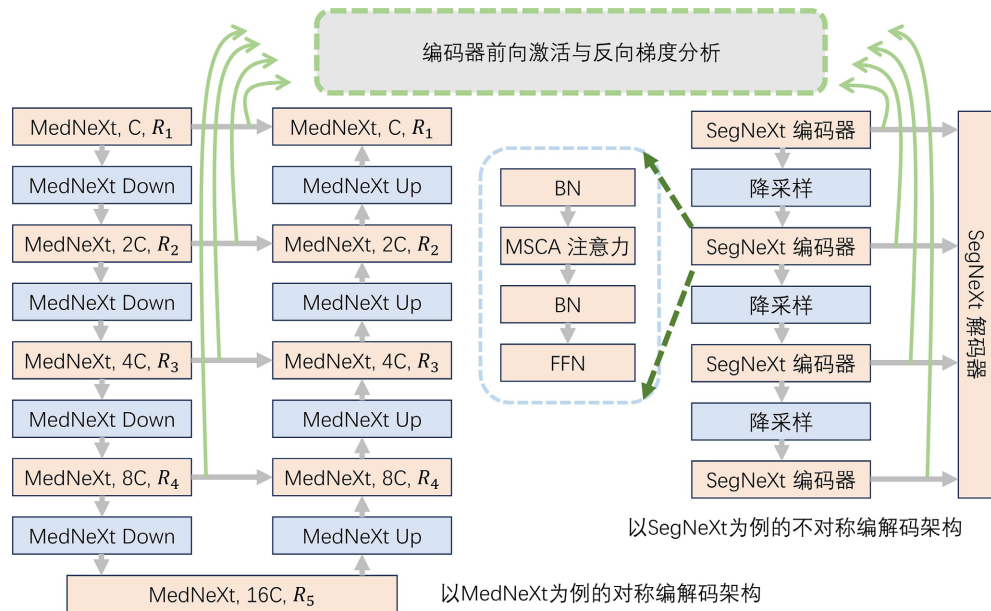


Figure 5. The framework to analyze neural implicit space

图 5. 神经隐空间分析架构

4. 实验与结果

4.1. 神经网络部署

本文所提出的增强方法涉及端到端胃癌病灶识别神经网络的前置处理与后置处理, 共计部署了 3 种主流的视觉神经网络, 用于验证它们是否受益于本文的方法。SegNeXt [1] 是著名的轻量级卷积分割网络, 它以较低的计算复杂度达成了较高的分割性能, 由于 CT 扫描序列通常拥有数百个 Slice, 分割计算量大,

因此该网络特别适合于临床快速辅助诊断；DATransUNet [19]则是一种先进的由 Transformer 启发的卷积分割网络，在多种领域的任务上具有较高的通用性。该两种模型具有较高的代表性，目前多数前沿的算法均基于类似于这两者的神经架构范式——编解码器设计的卷积或 Transformer 网络；MedNeXt [20]为 nnUNet 团队开源的新一代医学卷积模型，用于为不同的医学分割任务提供有竞争力的实验基准。

所有实验由 OpenMM 框架搭配 PyTorch 2.6.0 实现，以确保数据流一致性。学习率设定为 0.0001，batchsize 为 8，嵌入维度遵从神经网络的默认设定。当连续 10 个 Epoch 的验证集精度不下降时停止训练，并取停止时的权重对模型进行测试集推理评估。图 6 展示了本文的整体数据流。

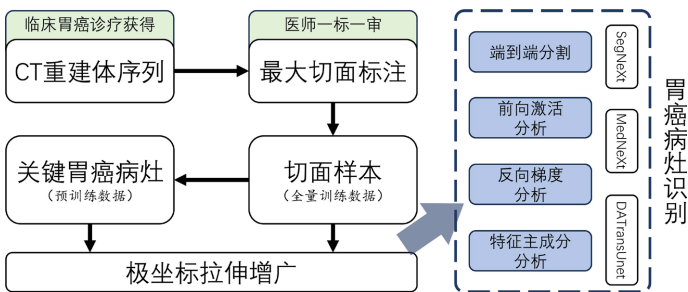


Figure 6. The data flow for training
图 6. 训练数据流

4.2. 端到端分割训练的量化分析

从表 1 中可以分析得知，所提出的方法能够稳定地提高 3 个模型的分割性能。平均来看，病灶交并比提升了 4.36，Dice 系数提升了 2.69，病灶识别敏感度(Recall)提升了 3.39，病灶识别特异性(Precision)提升了 1.94。SegNeXt 模型是轻量模型，自身参数量较少，在标注稀缺的训练情形下，难以拟合，在实验中因梯度不稳定而无法训练成功。在使用所提出的增广方法和前置预训练后，其能够成功拟合，并且略微超越 DATransUNet 模型。

Table 1. Segmentation accuracy of gastric cancer lesions
表 1. 胃癌病灶的分割精度

拉伸	模型	mIoU	mDice	mRecall	mPrecision
×	SegNeXt	相同参数下，训练因梯度爆炸而失败			
	MedNeXt	77.93	87.60	85.57	89.73
	DATransUNet	71.47	83.36	81.89	84.89
√	SegNeXt	79.91	88.84	87.91	89.78
	MedNeXt	88.73	94.03	93.76	94.30
	DATransUNet	77.79	87.51	87.15	87.87

由于所提出的方法能够借助 CPU 实现异步处理，对 GPU 计算的影响很小，而目前的分割框架的性能瓶颈主要围绕在 GPU 上。因此，所获得的精度提升并不需要更多训练卡时。在临床中，这意味着医师能够在实时分割的应用场景下，获得更准确的分割预测结果。

4.3. 基于前向激活图的特征分辨能力的定性分析

上一节所述内容能够定量地确认所提出的方法是否满足了实际临床任务的要求，本节进一步从定性

的角度解构训练流程中神经网络前向激活值。具体来说,需要观察编码器靠近数据输入端的层激活情况,来确认神经网络进行特征提取时,哪些计算对训练和预测具有显著的贡献。前向激活图能够显示神经元对数据输入的响应程度,如果观察到激活图中具有强烈的特征,说明神经元对图像中不同部分的响应是有差异性的、有着重点的[21][22]。

图 7 展示了在编码器靠近嵌入层的一端的前向激活响应分布图,前者的颜色映射显示了完整的响应值域,后者则映射了 65%到 95%分位数范围内的响应值,以改善直观度。数据为测试集所有样本上的前向响应的平均值,以此来准确体现模型在预测场景下的特性。

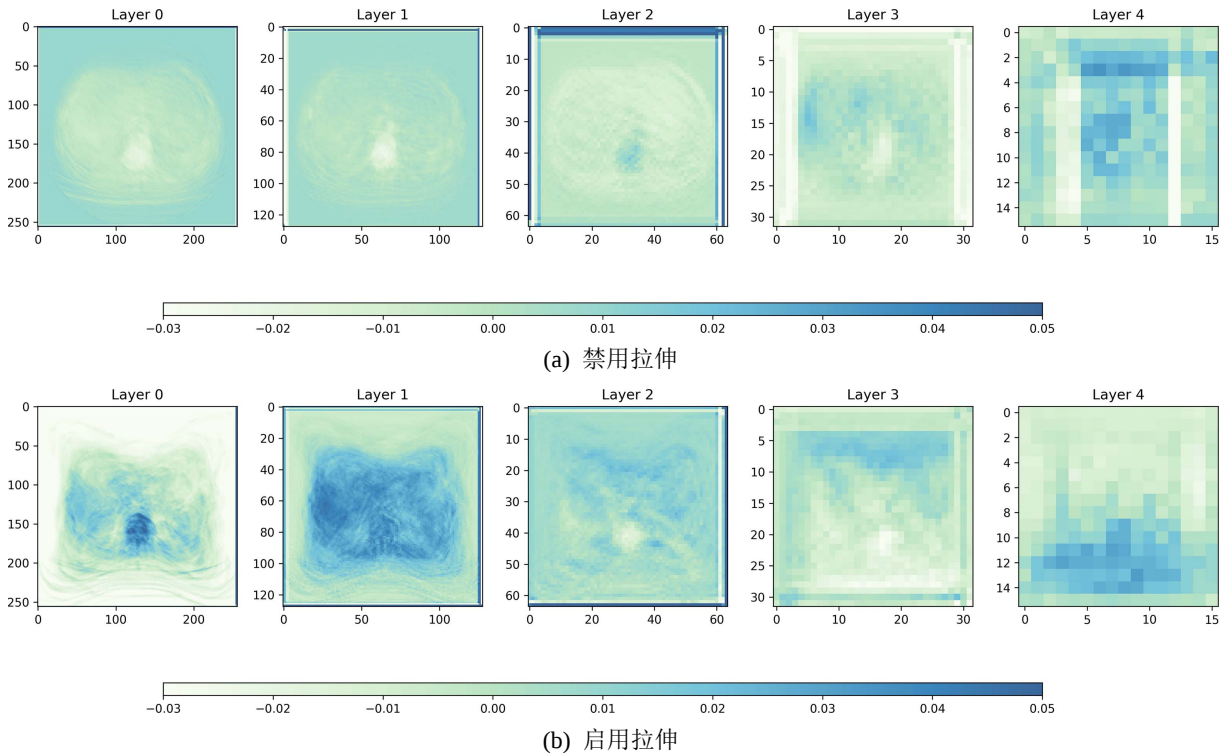


Figure 7. Encoder forward activation map
图 7. 编码器前向激活图

激活图观察显示,在不使用拉伸增广时,编码器的有效信号占比明显比使用拉伸增广时的要低。浅层的编码器倾向于输出一致性的信号,对特征的提取有限,仅在编码器深层有一定大尺度特征体现。即使映射重点激活值域,其提取的特征区分度也依旧不理想。另一边,当使用拉伸增广时,编码器的第一层即能够差异性地分辨前后景(前景为绿色的较高值,后景为较低的深蓝色值)。

4.4. 基于像素梯度分布的神经计算有效性分析

在本文第 3 节中已经提及,反向梯度图能够展示训练过程中 CT 切面的不同部分是否对损失函数的计算具有影响力。如果梯度仅聚集在较小的范围内,则提示神经网络关注的特征尺度小,容易引发梯度不稳定,同时也意味着大多数神经元计算并不为最终的参数更新做出显著贡献[23]。

本节通过在 PyTorch 的 forward 函数过程中执行反向梯度注入(Backward Hook),可视化编码器梯度,分析来自于目标损失函数的梯度在传至神经网络数据输入端时的像素尺度梯度分布情况。从图 8 中可知,在不使用拉伸映射时,仅靠近横断面中央的部分像素拥有梯度,而当使用增广后,梯度分散到了更

大的区域中。特别地，在靠近右上角的区域有显著的增强，而这正对应着大多数患者的胃癌 CT 病灶在最大病灶横断面上的位置。

梯度不显著的像素意味着在其位置进行的神经计算对最终的模型预测的贡献是相对微弱的，即有大量的可被剪枝的神经元计算操作。本文所提出的方法让神经网络的有效计算操作的像素相对范围变高了，分布更广了，即提高了神经计算效率。

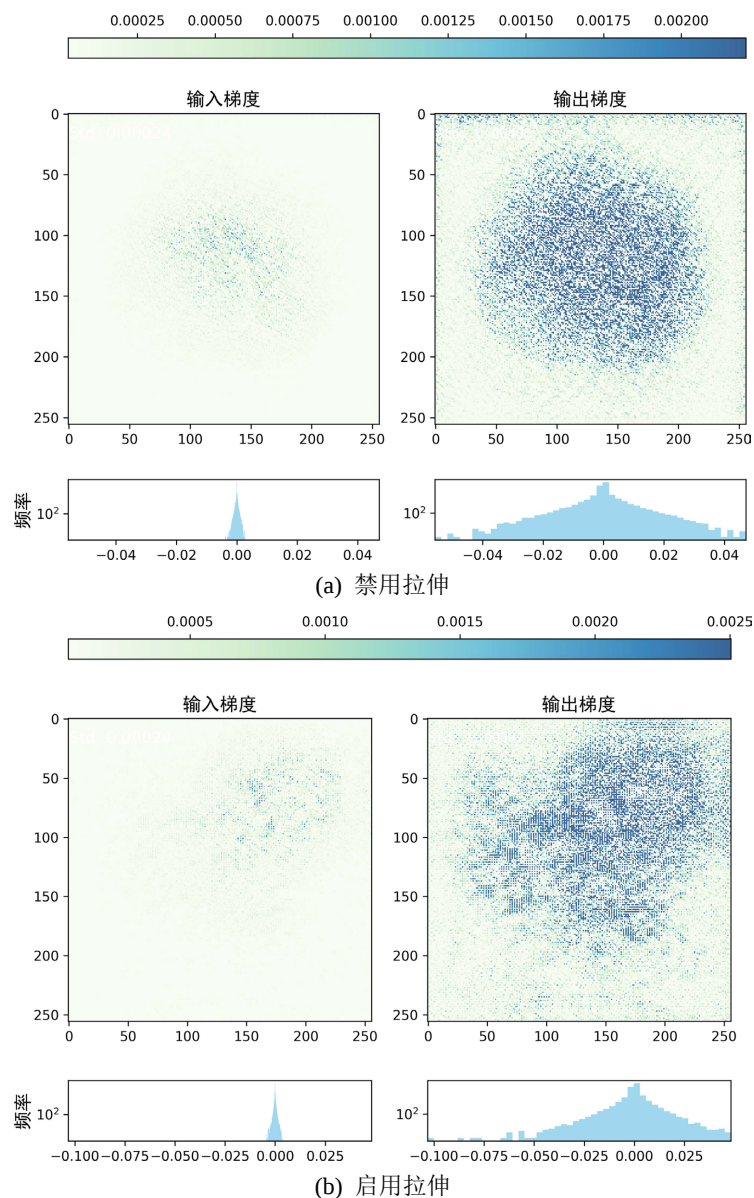


图 8. 编码器梯度 $\nabla_{\text{color}} \in [Q_{65\%}(\nabla), Q_{95\%}(\nabla)]$

Figure 8. Encoder gradients $\nabla_{\text{color}} \in [Q_{65\%}(\nabla), Q_{95\%}(\nabla)]$

4.5. 神经隐空间的特征主成分分析

端到端分割任务能够体现神经网络在目标分割任务上的定位、分类能力，但不能体现不同的策略

对神经隐空间的影响。虽然分析神经网络作出判断的具体依据至今依旧是难点,但由于在胃癌分割中只涉及两个类(正常组织区域、病灶组织区域),所以我们能够使用主成分分析来尝试解析出在高维隐空间中是否存在两个互相分离的主成分特征(PC)。本节引入假设:主成分分析(PCA)所解析出的两个成分与胃癌的二值分类有关。

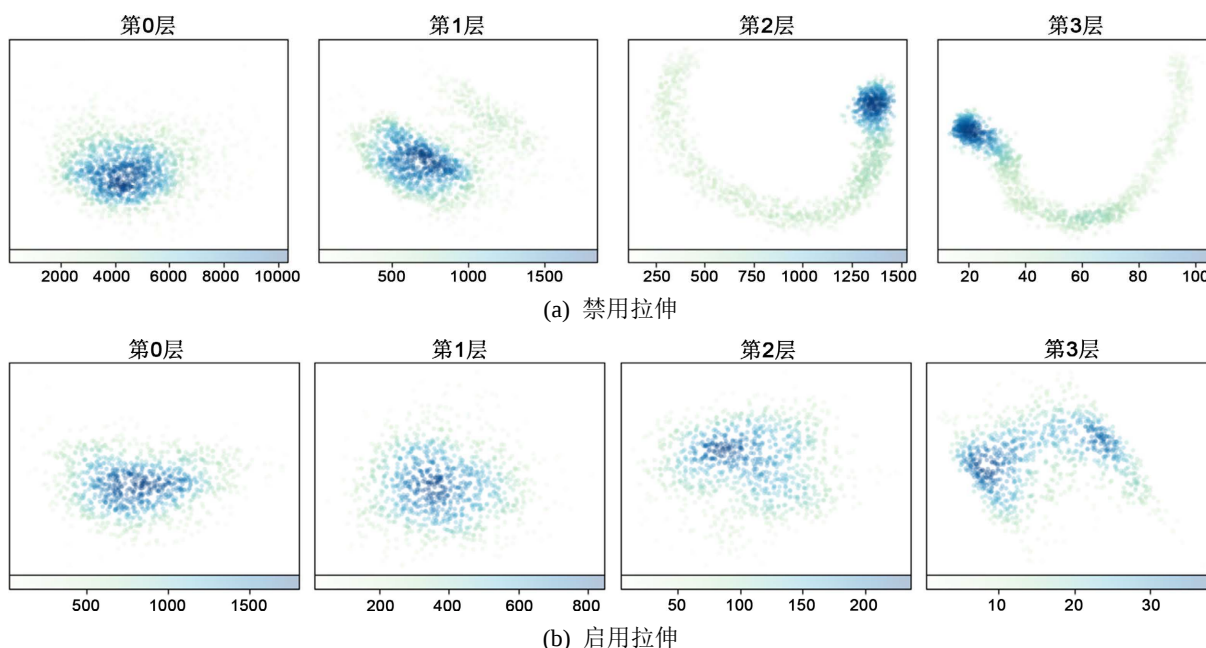


Figure 9. Binary principal component analysis results of the neural implicit space

图 9. 神经隐空间二元主成分分析

从图 9 中可知,在禁用拉伸时,在第 1 层的隐空间可见轻微的主成分分离趋势,而在其他层中,隐空间的分离趋势均不明显,这意味着编码器尚未拥有较强的胃癌分辨能力。在使用本文提出的方法之后,深层(第 3 层)隐空间呈现出两条平行的渐近线,这代表着隐空间中存在着两个重点特征,能够被 PCA 所解析。

需要注意的是,图 9 中的隐空间为 MedNeXt 模型的编码器向解码器传输的空间,采样自模型拟合完成后在 89 个样本上进行推理时产生的高维特征的平均值。该隐空间的特征会被解码器进一步处理,以生成最终的胃癌分类判断。即使在禁用拉伸时隐空间的特征分离程度不明显,在经过解码器进一步特征映射后,依旧可能得到可用的分类结果,这体现在定量分析部分(3.2 节)的精度并不为 0。

5. 结论

本研究针对卷积神经网络在 CT 胃癌影像分析中存在的卷积感受野与 CT 重建域不一致问题,提出了一种基于仿射拉伸的截面增强方法,旨在通过优化数据分布特征减少神经元的冗余计算。实验结果表明,该方法在多个主流分割模型上显著提升了胃癌病灶的识别效率与精度,为临床快速辅助诊断提供了新的技术思考。本文使用了临床数据源进行再分析,更进一步的研究应当对这类方法在临床场景下医师的接受程度和认可程度作出细致评估。

致 谢

本研究受到了国家自然科学基金项目(61572325);上海市重点科技项目(19DZ1208903)的资助,本文

作者团队对上述项目的支持表示感谢。

基金项目

国家自然科学基金项目(61572325); 上海市重点科技项目(19DZ1208903)。

参考文献

- [1] 祝蕾, 朱坤福. 人工智能在医学影像中的应用及面对的机遇和挑战[J]. 智慧健康, 2021, 7(30): 1-3.
- [2] 周翔, 王培军. 中国医学影像人工智能发展现状及展望[J]. 同济大学学报(医学版), 2025, 46(1): 1-7.
- [3] 陈冲, 陈俊, 夏黎明. 人工智能促进医学影像临床应用与研究[J]. 放射学实践, 2024, 39(1): 12-16.
- [4] 曹玉红, 徐海, 刘荪傲, 等. 基于深度学习的医学影像分割研究综述[J]. 计算机应用, 2021, 41(8): 2273-2287.
- [5] 魏小娜, 李英豪, 王振宇, 等. 医学影像人工智能辅助诊断的样本增广方法[J]. 计算机应用, 2019, 39(9): 2558-2567.
- [6] 杭利华, 蒋佩钊, 邓冲. 基于仿射变换与B样条自由形变的医学图像配准[J]. 兰州交通大学学报, 2013, 32(3): 44-48+65.
- [7] Zhang, Z., Yang, H., Guo, Y., Bolo, N.R., Keshavan, M., DeRosa, E., *et al.* (2023) Affine Image Registration of Arterial Spin Labeling MRI Using Deep Learning Networks. *NeuroImage*, **279**, Article 120303. <https://doi.org/10.1016/j.neuroimage.2023.120303>
- [8] 张颖麟, 胡衍, 东田理沙, 等. 生成对抗式网络及其医学影像应用研究综述[J]. 中国图象图形学报, 2022, 27(3): 687-703.
- [9] 周俊林, 洪楠, 张惠茅, 等. 结直肠癌 CT 影像数据库构建及质量控制专家共识[J]. 中国医学影像学杂志, 2025, 33(1): 1-9.
- [10] Liu, X., Ono, K. and Bise, R. (2023) Mixing Data Augmentation with Preserving Foreground Regions in Medical Image Segmentation. 2023 *IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Cartagena, 18-21 April 2023, 1-5.
- [11] 徐承辉, 张永生. 双源 64 层 CT 的质量控制检测分析[J]. 中国医疗器械信息, 2024, 30(18): 164-166.
- [12] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6.
- [13] El Jiani, L., El Filali, S. and Benlahmer, E.H. (2022). Overcome Medical Image Data Scarcity by Data Augmentation Techniques: A Review. 2022 *International Conference on Microelectronics (ICM)*, Casablanca, 4-7 December 2022, 21-24. <https://doi.org/10.1109/icm56065.2022.10005544>
- [14] Garcea, F., Serra, A., Lamberti, F. and Morra, L. (2023) Data Augmentation for Medical Imaging: A Systematic Literature Review. *Computers in Biology and Medicine*, **152**, Article 106391. <https://doi.org/10.1016/j.compbiomed.2022.106391>
- [15] 唐翠松, 李励, 邱裕友, 等. 放射影像检查中的医学伦理困境及对策分析[J]. 医学与哲学, 2022, 43(1): 34-37.
- [16] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [17] 孟玫, 鲍伟宇, 尤云峰. 磁共振成像与多层螺旋 CT 诊断对胃癌术前 N 分期及淋巴结转移的诊断效能及影像学特点分析[J]. 实用医学影像杂志, 2024, 25(1): 29-32.
- [18] 朱艳琳, 黄陈恕, 孙宗琼. CT 灌注成像在胃癌诊断中的临床应用[J]. 医学影像学杂志, 2015, 25(7): 1194-1198.
- [19] Sun, G., Pan, Y., Kong, W., Xu, Z., Ma, J., Racharak, T., *et al.* (2024) Da-Transunet: Integrating Spatial and Channel Dual Attention with Transformer U-Net for Medical Image Segmentation. *Frontiers in Bioengineering and Biotechnology*, **12**, Article 1398237. <https://doi.org/10.3389/fbioe.2024.1398237>
- [20] Roy, S., Koehler, G., Ulrich, C., Baumgartner, M., Petersen, J., Isensee, F., *et al.* (2023) Mednext: Transformer-Driven Scaling of Convnets for Medical Image Segmentation. In: Greenspan, H., *et al.*, Eds., *Lecture Notes in Computer Science*, Springer, 405-415. https://doi.org/10.1007/978-3-031-43901-8_39
- [21] 司念文, 张文林, 屈丹, 等. 卷积神经网络表征可视化研究综述[J]. 自动化学报, 2022, 48(8): 1890-1920.
- [22] 尚骏远, 杨乐涵, 何琨. 基于特征可视化分析深度神经网络的内部表征[J]. 计算机科学, 2020, 47(5): 190-197.
- [23] 白林亭, 海钰琳. 基于梯度分析的卷积神经网络可视化方法[J]. 信息技术与信息化, 2021(4): 61-63.