

# 基于KAN网络的U型医学图像分割方法

陈楚琪, 张圣涛

河北工业大学理学院, 天津

收稿日期: 2025年4月15日; 录用日期: 2025年5月16日; 发布日期: 2025年5月26日

## 摘要

随着深度学习的持续发展, 基于人工智能的计算机辅助诊断技术在医学领域也有了許多进展, 尤其是在医学图像分割领域。本文针对当前医学图像分割中网络复杂度高, 深度学习网络可解释性差问题, 提出了一种基于KAN网络的U型医学图像分割网络, KAU-Net网络。通过提出KAN通道注意力机制与KAN空间注意力机制, 分别改进了U型网络中的解码器部分与跳跃连接部分, 使分割网络能够在降低模型计算复杂度, 提高深度学习网络可解释性。实验表明, KAU-Net网络能有效降低计算复杂度, 同时提高对目标图像的分割精度, 解决了由于注意力机制的加入, 分割网络中计算量大大增加的问题。

## 关键词

U型分割网络, KAN网络, 医学图像分割, 注意力机制

# U-Shaped Medical Image Segmentation Method Based on KAN Network

Chuqi Chen, Shengtao Zhang

College of Science, Hebei University of Technology, Tianjin

Received: Apr. 15<sup>th</sup>, 2025; accepted: May 16<sup>th</sup>, 2025; published: May 26<sup>th</sup>, 2025

## Abstract

With the continuous development of deep learning, computer-aided diagnosis technology based on artificial intelligence has also made many advancements in the medical field, especially in the area of medical image. This paper addresses the issues of high network complexity and poor interpretability of deep learning networks in current medical image segmentation, and proposes a U-shaped medical image segmentation network based on KAN network, KAU-Net. By introducing KAN channel attention mechanism and KAN spatial attention mechanism, the decoder part and skip connection part of the U—are respectively improved, enabling the segmentation network to

reduce the computational complexity of the model and enhance the interpretability of deep learning networks. Experiments show that the KAU-Net network can effectively reduce computational complexity while improving the segmentation accuracy of target images, solving the problem of significantly increased computational load due to the addition of attention mechanisms in the segmentation network.

## Keywords

U-Shaped Segmentation Network, KAN Network, Medical Image Segmentation, Attention Mechanisms

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

近年来,随着深度学习持续发展,基于人工智能的计算机辅助诊断技术在医学领域也有了許多显著进展,为疾病的诊断与治疗争取时间[1]。其中基于深度学习对医学图像进行分割的许多模型也被提出。2015年 Long 等人[2]提出了全卷积神经网络(FCN),它使用反卷积取代了之前网络中的全连接层,并对最后一层特征图进行上采样操作使其恢复到原始大小,最终实现了逐个像素精准分类预测的目的。同年, Ronneberger 等人[3]提出了 U-Net 网络。U-Net 网络由通过独特的 U 型对称结构和跳跃连接部分组成,能够准确捕捉到可用图像中的特征信息。U-Net 模型由于其灵活性、优化的模块化设计成为了最广泛应用的医学图像分割基准模型,许多研究者也在 U-Net 网络模型的基础上提出了一系列改进模型,如 Ibtehaz 等人[4]将 MultiRes 模块融入卷积,提出了 MultiResUNet, Zhou 等人[5]首次对跳跃连接部分进行改进,提出了 UNet++, Qin 等人[6]将 U 型结构模型进行两层嵌套改进网络模型,提出了 U2-Net 网络。

纯卷积网络,其重点关注局部细节处的特征,对全局整体的关注度不够,全局特征建模能力不足。2017年,Google 机器翻译团队发表《Attention is all you need》[7],该文章中完全使用自注意力机制来学习文本表示,该结构又被称之为 Transformer,此后开起了注意力机制的研究热潮。2018年, Oktay 等人[8]首次将注意力机制与 U-Net 模型相结合,提出 Attention U-Net 模型。2021年, Chen 等人[9]引入 ViT 模块(Vision Transformer),提出了 TransUNet 网络。2023年, Coa 等人[10]将 Swin Transformer 作为 U 型网络主干,提出了 Swin-Unet 网络。Wang 等人[11]在跳跃连接处加入 Swin Transformer,提出了 CA-UNet 网络。

注意力机制的加入,一定程度上能提升模型的精度,但加也会致使分割网络中参数量与计算量大大增加,对算力、显存等硬件资源的要求也急剧提高,对于硬件设施的要求也逐渐提高,这与其在基层医疗场景的部署需求形成显著矛盾。因此,亟需提供一种更为轻量化的医学图像分割方法。

本文提出了一种基于 KAN 网络的 U 型医学图像分割网络,KAU-Net 网络。通过提出 KAN 通道注意力机制与 KAN 空间注意力机制,分别改进了 U 型网络中的解码器部分与跳跃连接部分,使分割网络能够在降低模型计算复杂度,提高深度学习网络可解释性的同时提高对目标图像的分割精度,解决了由于注意力机制的加入,分割网络中计算量大大增加的问题。

## 2. 模型构建

本文提出的基于 KAN 网络的 U 型医学图像分割模型,KAU-Net 网络的整体模型结构如图 1 所示。该网络模型依旧沿用 U 型网络结构,由编码器、解码器、跳跃连接三部分组成。编码器部分采用 ConvNeXt

网络来提取图像特征, 采用 ConvNeXt Block 模块代替原 U-Net 网络编码器中用于提取图像特征的  $3 \times 3$  卷积核, 选用每层模块数分别为 3, 3, 9, 3 的结构; 在解码器部分采用深度可分离卷积模块代替 U-Net 网络解码器中  $3 \times 3$  的常规卷积核, 并在每层解码器后加入 KAN 通道注意力模块; 在跳跃连接部分对每层编码器提取出的特征通过 KAN 空间注意模块后与对应层的解码器部分特征合并。

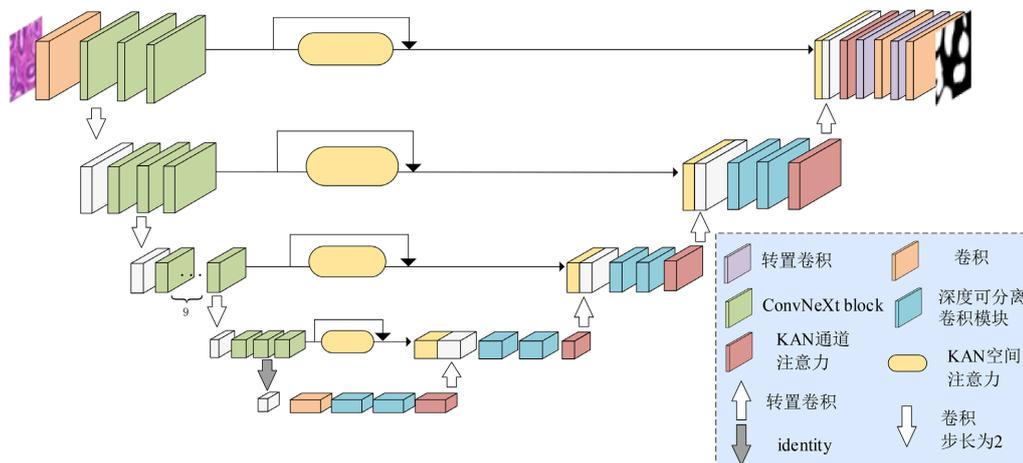


Figure 1. U-shaped medical image segmentation network based on KAN  
图 1. 基于 KAN 网络的 U 型医学图像分割模型

### 2.1. 编码器

编码器部分采用 ConvNeXt-T 网络来提取图像特征, ConvNeXt-T 是 ConvNeXt 系列中的轻量级版本, 该网络是在 ResNet 网络模型的基础上仿照 swin Transformer 结构进行改进而得到的纯卷积网络, 通过引入 Transformer 的设计特性, 提升了模型提取特征的性能, 同时因为没有引入自注意力机制, 硬件兼容性得到提升, 参数量与计算量也有所降低。其核心模块主要采用 ConvNeXt Block, 其具体结构如图 2 所示, 主要由深度可分离卷积、层归一化(Layer Norm)、倒置瓶颈层和残差连接四部分组成。

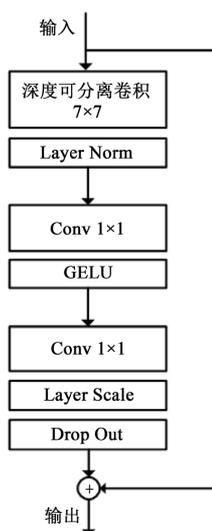


Figure 2. The specific structure of the ConvNeXt Block  
图 2. ConvNeXt Block 具体结构

## 2.2. KAN 空间注意力机制

传统 U 型分割网络模型一般采用纯卷积网络, 如 U-Net 网络, UNet++网络, U<sup>2</sup>-Net 网络等, 但纯卷积网络重点关注局部细节处的特征, 对全局整体的关注度不够, 全局特征建模能力不足。因此, 在分割网络中引入注意力机制, 注意力机制能提高网络对整体的关注度, 如 TransUNet 网络, Swin-Unet 网络, CA-UNet 网络等。但引入注意力机制后, 模型中参数量与计算量大大增加, 对算力、显存等硬件资源的要求也急剧提高, 不利于模型工作推广。同时, 深度学习的可解释性较差。因此, 为减小模型的复杂度并提升模型一定的可解释性, 本文提出了 KAN 空间注意力机制模块, 并将该模块插入跳跃连接部分。

KAN (Kolmogorov-Arnold Networks) [12]是引入了一种基于 Kolmogorov-Arnold 表示定理的新型神经网络架构, 与传统的多层感知器(Multilayer Perceptron, MLP)不同, 它们通过使用可学习的函数替代固定的激活函数, 从而消除了对线性权重矩阵的依赖。KAN 的理论基础, 源自苏联数学家 Vladimir Arnold 和 Andrey Kolmogorov 的研究成果, 他们理论研究基于多变量连续函数的概念。

$$f(x) = f(x_1, \dots, x_n) = \sum_{q=0}^{2n+1} \Phi_q \left( \sum_{p=1}^n \phi_{q,p}(x_p) \right). \quad (1)$$

如公式 2.1 所示, KAN 定理可将复杂的高维函数分解为简单的一维函数组合。通过集中优化这些一维函数而非整个多变量空间, KAN 大幅降低了实现精确建模所需的复杂性和参数数量。此外, 由于这些函数较为简单, KAN 网络也会使模型简单且易于解释。同时, KAN 能够将网络修剪成更小的子网, 并在节点级别对 KAN 进行稀疏化, 对不重要的神经元进行修剪。这些操作都增强了 KAN 模型的可解释性。

在 KAN 中, 传统的权重参数在网络的边缘被单变量函数参数所取代, 每个节点汇总这些函数的输出时不进行任何非线性变换, 这与 MLP 中的做法, 线性变换后再通过非线性激活函数形成了鲜明对比, 从以下两公式中可看出两者区别。

$$KAN(x) = (\Phi_L \circ \Phi_{L-1} \circ \dots \circ \Phi_1 \circ \Phi_0), \quad (2)$$

$$MLP(x) = (W_L \circ \sigma \circ W_{L-1} \circ \sigma \circ \dots \circ W_1 \circ \sigma \circ W_0 \circ \sigma), \quad (3)$$

其中,  $\circ$  代表两个映射的乘法,  $W_L$  表示第  $L$  层的线性权重参数,  $\sigma$  表示非线性激活函数。

样条是 KAN 学习机制的核心, 它们取代了神经网络中通常使用的传统权重参数。样条是一种数学函数, 用于通过一组控制点创建光滑和灵活的曲线或曲面。在数学术语中, 样条是一个分段多项式函数, 它在多项式块相交的地方保持高度平滑。KAN 网络中采用  $b$  样条为通用公式:

$$spline(x) = \sum_i c_i B_i(x), \quad (4)$$

其中,  $spline(x)$  表示样条函数,  $c_i$  代表训练期间优化的系数,  $B_i(x)$  是定义在网格上的  $b$  样条基函数。

KAN 空间注意力模块具体结构如图 3 所示。与 ViT 模型(Vision Transformer)类似, 其先将输入特征复制为三份  $X$ ,  $Q$ ,  $V$  三份输入的特征尺寸均为  $C \times H \times W$ 。对于输入特征  $Q$ , 先通过一个全局池化操作得到  $C \times 1 \times 1$  的特征矩阵, 紧接着在空间范围内对其压缩, 将特征尺寸变为  $C \times 1$ , 使其尺寸满足 KAN 网络的输入; 之后通过两层 KAN 网络, 使输出的特征通道数变为  $\frac{C}{2}$ , KAN 网络对特征向量进行非线性变换, 从而引入更多的非线性特征, 增加了模型的非线性能力, 从而可以提高特征提取的效果; 最后通过激活函数 Softmax, 得到输入特征  $Q$  的最终输出特征, 其尺寸为  $1 \times \frac{C}{2}$ 。对于输入特征  $V$ , 先经过  $1 \times 1$  的卷积核后尺寸变为  $\frac{C}{2} \times H \times W$ ; 之后变化特征维度, 尺寸变为  $\frac{C}{2} \times HW$ 。通过对  $Q$  与  $V$  的上述操作, 最终

分别得到了  $1 \times \frac{C}{2}$  尺寸的特征和  $\frac{C}{2} \times HW$  尺寸的特征, 将这两尺寸特征通过矩阵乘法, 得到尺寸为  $1 \times HW$  的矩阵, 并对该矩阵进行维度变化与 Sigmoid 激活函数处理后得到  $1 \times H \times W$  的空间权重矩阵, 识别图像中不同局部特征信息的重要性, 最后通过将该空间权重矩阵与输入特征  $X$  进行逐元素相乘, 使注意力能够集中于图像空间上需要重点关注的目标区域, 并得到最终的输出特征  $C \times H \times W$ 。

KAN 空间注意力模块中, 由于注意力模块的引入, 使得分割网络能重点关注重要区域的特征, 并从空间角度增加了对全局的关注度和全局特征的建模能力, 从而可以有效提高网络分割性能。同时, 由于 KAN 网络的引入降低了一般注意力机制中使用 MLP 的计算复杂度, 使计算量与参数量降低, 减少了算力成本, 并且 KAN 网络的非线性函数嵌套与剪枝处理等操作使得模型的可解释性增强。

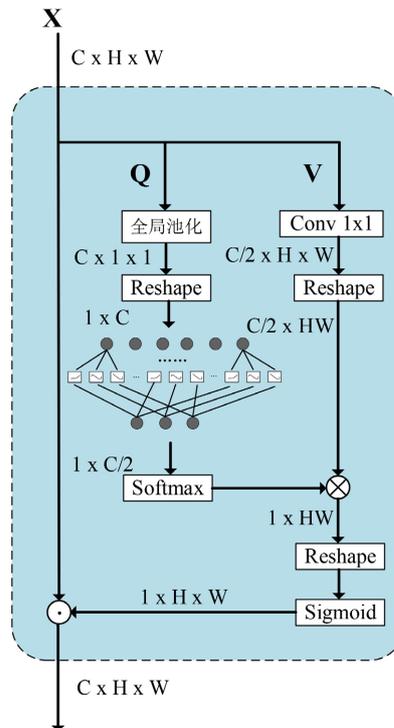


Figure 3. KAN spatial attention structure  
图 3. KAN 空间注意力机制结构

### 2.3. KAN 通道注意力机制

传统的 U-Net 网络在分割图像过程中重点捕获局部特征, 并且对于所有的局部特征都等量的看待, 对重要区域没有重点关注, 导致预测结果易出现错误分割, 或者过度分割。因此, CA-UNet 网络在解码器中引入了借助 MLP 实现的通道注意力机制, 用于减少模型的错误分割与过度分割, 但同时也导致了模型中参数量与计算量的大大增加。因此, 本文提出了 KAN 通道注意力机制, 放置于每层解码器处, 用于减少模型的错误分割, 并能有效减少模型的参数量与计算量。

KAN 通道注意力模块结构如图 4 所示, 其以通道注意力模块(Channle Attention Module, CAM) [13]为基础, 与 CAM 不同的是, 将其中 MLP 部分换为了 KAN 网络, 来减少模块的参数量, 同时增强了可解释性。其具体操作从图中可看到, 将输入特征经过最大池化操作和平均池化操作后, 分别获取每个通道的全局统计信息, 之后通过两层 KAN 网络的非线性函数的嵌套来学习通道的权重, 然后将处理后产生的

两个结果通过矩阵元素加法进行相加, 并使用 Sigmoid 激活函数将权重归一化 0 到 1 之间得到最终通道权重矩阵, 最后, 将通道权重矩阵与输入特征相乘得出最后输出特征。

KAN 通道注意力机制, 从通道角度关注重点特征, 提高对全局的特征提取能力, 从而用于解决解码器中错误分割与过度分割的问题。并且 KAN 网络相比传统的 MLP, 能减少模型的计算量与参数量, 从数学上更具有可解释性。

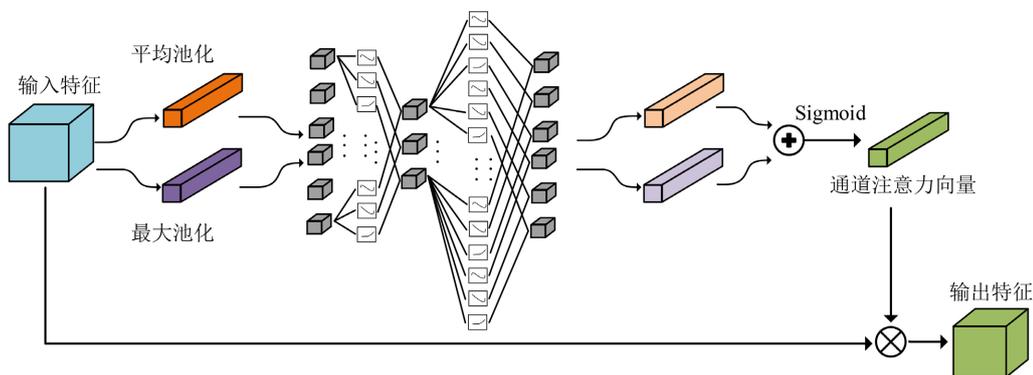


Figure 4. KAN channel attention structure

图 4. KAN 通道注意力机制结构

## 2.4. 损失函数

根据医学图像分割对每一像素二分类的特性, 选取二分类交叉熵损失函数, 其具体公式如下:

$$Loss = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log p_i + (1 - y_i) \cdot \log(1 - p_i)] \quad (5)$$

交叉熵能衡量真实值和预测值分布之间的信息差异。式中,  $N$  是总样本数, 即像素的数量,  $y_i$  是第  $i$  个样本的所属类别, 医学图像分割为二分类,  $p_i$  是第  $i$  个样本的预测值, 一般来说, 它是一个概率值, 该式直接测量两个样本的相似度。采用交叉熵损失函数作为网络训练的损失函数, 可以有效避免梯度小时神经网络无法继续训练的问题, 使神经网络的权重在训练过程中有效收敛到真实值。

## 3. 实验结果与分析

### 3.1. 数据预处理

为了评估本实施例提出的医学图像分割网络, 选用 GLaS 数据集[14]来展示其实验结果。GLaS 数据集是一组用于结直肠腺体组织的分割任务的数据集, 最初由 2015 年 MICCAI 挑战赛发布。该数据集包含 165 张图像, 四种不同分辨率, 每张图像都对应各自标签。这些图像来源于 16 个 H&E 染色的组织学切片, 这些切片代表了 T3 或 T4 阶段的结直肠腺癌。

在训练过程中, 为便于实验, 将数据集中每个图片分辨率大小都进行统一, GLaS 数据集提供了四种图像分辨率样本, 将其像素大小修改为  $512 \times 512$ , 医学图像数据量少, 样本宝贵, 为了充分利用数据, 对训练样本及对应的标签进行了数据增强操作, 例如: 随机旋转、水平翻转、平移缩放和随机裁剪等。并对每个数据集连同对应的标签样本, 按照 4:1 比例进行训练集和测试集划分。

### 3.2. 实验环境

实验使用的设备为 CPU 型号为 I7-11700K, GPU 型号为 NVIDIA GeForce RTX 3090 的笔记本, 其

中显存为 24G, 内存为 32G, 操作系统为 Windows10。实验采用 PyTorch 框架搭建实验网络模型, 将迭代次数 epoch 设置为 200 代, 批次大小 batch size 设置为 32。优化器选择 Adam, 并将优化器的学习率设定为  $1e-3$ , 衰减系数将其设定为  $1e-4$ 。

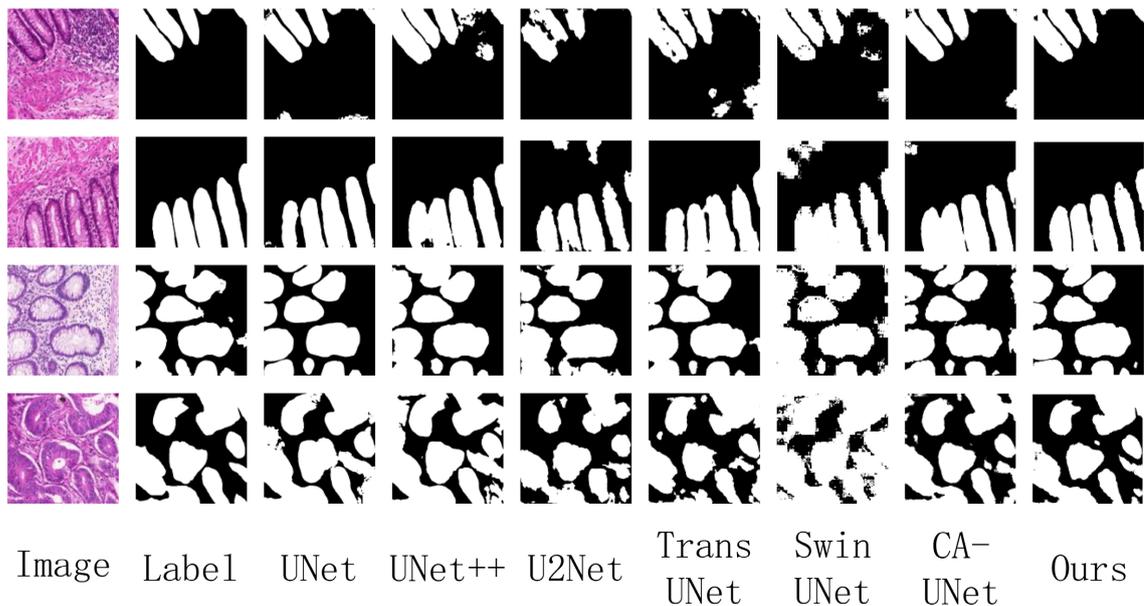
### 3.3. 对比实验

为了验证本文提出方法的效果与性能, 实验验证中使用以下三种统计指标来评价网络的分割性能, 分别为平均交并比(Mean Intersection over Union, IoU)、Dice 相似系数(Dice Similarity Coefficient, DSC)和召回率(Recall), 下面给出了三种统计指标的具体公式。使用模型参数量(Params)和训练计算量(FLOPs)两种指标分别从空间复杂度和时间复杂度两方面来评价模型复杂度。

$$mIoU = \frac{1}{K} \sum_{i=1}^K IoU_i = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FP_i + FN_i}. \quad (6)$$

$$DSC = \frac{2TP}{FP + 2TP + FN}. \quad (7)$$

$$Recall = \frac{TP}{TP + TN}. \quad (8)$$



**Figure 5.** GLaS dataset segmentation results and comparison results  
**图 5.** GLaS 数据集分割结果及其对比结果

为了验证本文所提出的方法的有效性, 与不同医学图像分割模型进行比较分割性能, 本章将所提出的模型与 U-Net、U-Net++、U2-Net、TransUNet、Swin-UNet 和 CA-UNet 这六种模型进行实验性能对比。在 GLaS 数据集上的分割结果如图 5 所示, 对于选用的 6 种对比模型, U-Net++ 在跳跃连接部分采用密集连接, U2-Net 采用嵌套 U 型结构, 这两种网络结构对特征的提取能力均优于 U-Net, 但由于采用纯卷积的网络结构, 对整体关注度不够, 容易过度分割, 因此分割结果仍然不算精确。TransU-Net 引入了 Transformer 模块, 而 Transformer 模块参数复杂, 对于训练样本少且样本之间差异大的数据集, 泛化性较差, 并且对局部关注度不足。Swin-UNet 引入 Swin-Transformer 模块, 同 TransU-Net 的问题类似, 都是

对于局部细节和边缘部分的把控能力不足, 对于腺体组织边缘的分割结果零散, 并且注意力网络模块的参数量与计算量都远超过纯卷积网络结构, 在数据量小时, 注意力网络的性能优势不足以发挥出来。CA-UNet 模型引入了 Swin-Transformer 模块和通道注意力两种注意力模块, 使得模型提取特征能力增强, 分割结果相比前几种网络模型稍佳, 但是对于靠近图像边框的部分分割不够精确。相比以上的 6 种对比模型本文提出的 KAU-Net 网络模型引入 KAN 空间注意力机制与 KAN 通道注意力机制, 使多尺度的图像特征在空间和通道维度上得到增强, 将注意力机制与卷积网络相结合, 既能关注图像整体特征, 也能注意局部与细节部分, 并且由于医学图像数据集样本量较小, 两者结合能更好发挥出优势。分割结果图中可以直观看到所提出的网络不仅对全局有所关注, 没有分割出目标的情况降低, 在局部细节上过度分割的情况也明显减少, 分割效果更好。

下表 1 为所提出本方案网络与其他网络对比的具体衡量指标, 从具体指标上也能看到, 本方案网络在平均交并比(mIoU)、Dice 相似系数(DSC)和召回率(Recall)这三类评价指标上表现都优于其他网络, 相比 6 种网络分割性能有明显提升, mIoU 和 DSC 相比 U-Net++分别提升了 0.7%和 1.2%, Recall 相比 CA-UNet 提升了 2.6%。U-Net 网络, U2-Net 网络和 UNet++网络均为纯卷积结构, 本文提出的网络由于引入注意力结构因此模型参数量上有所增加, 但在训练计算量(FLOPs)上因采用 KAN 网络代替 MLP 结构, KAN 网络可将网络修剪成更小的子网, 在节点级别对 KAN 进行稀疏化, 将部分不重要的神经元进行修剪, 从而使其参数量与训练速度低于使用 MLP 的常规注意力模块, 以及采用了相比常规卷积参数量更少的深度可分离卷积, 从而使其训练计算量远低于其他网络。因此, 所提出网络总体降低了模型复杂度, 并提升了分割性能。

**Table 1.** GLaS dataset statistical result

**表 1.** GLaS 数据集统计指标

模型	mIoU (%)	DSC (%)	Recall (%)	Params	FLOPs
U-Net	68.02	79.67	79.50	31.04M	193.30G/s
UNet++	70.35	80.92	77.12	47.19M	543.59G/s
U2-Net	65.96	79.06	76.42	44.00M	151.00G/s
TransUNet	65.25	77.74	79.21	105.91M	168.73G/s
Swin-Unet	66.84	79.53	78.38	149.25M	162.07G/s
CA-UNet	68.92	80.13	80.51	146.45M	260.54G/s
ours	70.84	81.88	82.62	93.19M	50.17G/s

## 4. 结语

本文提出了一种基于 KAN 网络的 U 型医学图像分割网络, KAU-Net 网络, 其能有效减少模型复杂度, 提高模型可解释性, 同时提升分割网络性能。该分割网络模型由编码器, 解码器与跳跃连接三个部分组成, 编码器部分采用 ConvNeXt-T 网络来提取图像特征; 在跳跃连接部分设计了 KAN 空间注意力机制来减少模型复杂度; 解码器部分采用深度可分离卷积替代常规卷积, 并设计了 KAN 通道注意力机制减少对图像的过度分割。最后在医学图像数据集 GLaS 上进行实验验证, 实验结果表明, 相比选用的六种对比模型, 本文提出的 KAU-Net 网络在 mIoU, DSC 和 Recall 上优于其他对比模型, 在 FLOPs 上远小于其他对比模型。

本研究的不足之处在于医学图像数据量少, 专业标注数据成本高昂, 但常用的分割模型多采用有监督学习, 需要大量数据。未来可结合半监督学习方式或自监督学习方式, 利用无标签数据进一步提升网络的分割性能。

## 参考文献

- [1] Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T. and Yang, X. (2020) Deep Learning in Medical Image Registration: A Review. *Physics in Medicine & Biology*, **65**, 20TR01. <https://doi.org/10.1088/1361-6560/ab843e>
- [2] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/cvpr.2015.7298965>
- [3] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [4] Ibtehaz, N. and Rahman, M.S. (2020) MultiResUNet: Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation. *Neural Networks*, **121**, 74-87. <https://doi.org/10.1016/j.neunet.2019.08.025>
- [5] Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. and Liang, J. (2020) UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Transactions on Medical Imaging*, **39**, 1856-1867. <https://doi.org/10.1109/tmi.2019.2959609>
- [6] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O.R. and Jagersand, M. (2020) U2-Net: Going Deeper with Nested U-Structure for Salient Object Detection. *Pattern Recognition*, **106**, Article ID: 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- [7] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [8] Oktay, O., Schlemper, J., Folgoc, L.L., et al. (2018) Attention U-Net: Learning Where to Look for the Pancreas. arXiv: 1804.03999.
- [9] Chen, J., Lu, Y., Yu, Q., et al. (2021) TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv: 2102.04306.
- [10] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., et al. (2023) Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. In: Karlinsky, L., Michaeli, T. and Nishino, K., Eds., *Computer Vision—ECCV 2022 Workshops*, Springer, 205-218. [https://doi.org/10.1007/978-3-031-25066-8\\_9](https://doi.org/10.1007/978-3-031-25066-8_9)
- [11] Wang, T., Wu, F., Lu, H. and Xu, S. (2023) CA-UNet: Convolution and Attention Fusion for Lung Nodule Segmentation. *International Journal of Imaging Systems and Technology*, **33**, 1469-1479. <https://doi.org/10.1002/ima.22878>
- [12] Liu, Z., Wang, Y., Vaidya, S., et al. (2024) KAN: Kolmogorov-Arnold Networks. arXiv: 2404.19756.
- [13] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 3-19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)
- [14] Sirinukunwattana, K., Pluim, J.P.W., Chen, H., Qi, X., Heng, P., Guo, Y.B., et al. (2017) Gland Segmentation in Colon Histology Images: The Glas Challenge Contest. *Medical Image Analysis*, **35**, 489-502. <https://doi.org/10.1016/j.media.2016.08.008>