

DynKANet: KA 表示定理启发的图像语义分割动态融合网络模型

梁 欣¹, 姜 林^{1,2*}, 彭 超¹, 刘志义¹, 黄依莎¹, 艾彦迪^{2,3}

(1. 湖南工商大学人工智能与先进计算学院, 湖南长沙 410000; 2. 湘江实验室, 湖南长沙 410000;
3. 湖南工商大学前沿交叉学院, 湖南长沙 410000)

摘 要: 针对现有语义分割任务在复杂场景下全局语义与局部细节特征融合时难以平衡的问题, 本文提出 KA 表示定理启发的图像语义分割动态融合网络 (Dynamic Kolmogorov-Arnold Network, DynKANet), 包含多级特征提取模块和动态特征融合模块. 在多级特征提取阶段, 设计基于残差连接的 U 型上下文增强模块和基于差异图的特征细化模块, 分别用于精准捕捉全局语义和局部细节特征, 以增强特征表达能力. 在上述特征提取基础上, 设计 KA 表示定理启发的动态特征融合模块, 采用内外两层函数对前序模块提取的全局语义与局部细节特征进行深度提取和精准拟合, 并结合自学习动态特征融合策略, 确保全局语义与局部细节信息的平衡互补, 有效缓解两类特征融合时难以平衡的问题. 此外, 设计 CE/TopK+Dice 动态联合损失函数, 并采用条件触发策略, 增强模型对困难样本的特征学习能力. 实验结果表明, 所提模型在涉及 5 个领域的 10 个公开数据集上均取得优异性能, 平均在每个数据集上性能提升 8%, 具备较强的泛化能力与应用潜力.

关键词: 语义分割; KA 表示定理; 动态特征融合; 全局语义; 局部细节

基金项目: 湘江实验室重大项目 (No.23XJ01003, No.23XJ01009); 湖南省教育厅科学研究重点项目 (No.22A0441); 国家级大学生创新创业训练计划项目 (No.S202410554037)

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2025)10-3671-21

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20250412

DynKANet: Dynamic Fusion Network for Real-Time Semantic Segmentation Based on Kolmogorov-Arnold Representation Theorem

LIANG Xin¹, JIANG Lin^{1,2*}, PENG Chao¹, LIU Zhi-yi¹, HUANG Yi-sha¹, AI Yan-di^{2,3}

(1. School of AI and Advanced Computing, Hunan University of Technology and Business, Changsha, Hunan 410000, China;
2. Xiangjiang Laboratory, Changsha, Hunan 410000, China;
3. School of Advanced Interdisciplinary, Hunan University of Technology and Business, Changsha, Hunan 410000, China)

Abstract: Semantic segmentation in complex scenes often suffers from the challenge of effectively balancing global semantic context and local fine-grained details during feature fusion. To address this issue, we propose dynamic Kolmogorov-Arnold network (DynKANet), a novel segmentation framework inspired by the Kolmogorov-Arnold representation theorem. The proposed architecture comprises a multi-level feature extraction module and a dynamic feature fusion module. Specifically, the feature extraction stage integrates a residual-connected U-shaped context enhancement module for robust global semantic modeling and a difference-map-based refinement module to enhance local detail representation. Building on these representations, we design a KA-inspired dynamic fusion module that decomposes the fusion process into nested inner and outer functions, enabling precise modeling of complex interactions between global and local features. A self-adaptive dynamic fusion strategy is incorporated to ensure complementary integration and mitigate the conflict between different feature types. Additionally, we introduce a dynamic compound loss function—CE/TopK+Dice—guided by a conditional trigger mechanism to strengthen the network's ability to learn from hard samples. Extensive experiments on 10 public datasets spanning 5 domains demonstrate that DynKANet achieves consistent improvements, with an average performance gain of

8% per dataset. These results highlight the strong generalization capability and practical potential of the proposed approach for real-world semantic segmentation tasks in challenging scenarios.

Key words: semantic segmentation; Kolmogorov-Arnold representation theorem; dynamic feature fusion; global semantics; local details

Foundation Item(s): Major Project of Xiangjiang Laboratory (No.23XJ01003, No.23XJ01009); Key Scientific Research Project of Hunan Provincial Department of Education (No.22A0441); National-Level Innovation and Entrepreneurship Training Program for College Students (No.S202410554037)

1 引言

语义分割旨在为图像中的每个像素分配相应语义标签,其不仅赋予机器理解视觉内容的能力,也是构建智能系统的重要基础.近年来,语义分割技术在自动驾驶^[1]、医疗图像^[2]、卫星遥感^[3]、城市场景^[4]和工业质量控制^[5]等领域得到广泛应用,具有重要学术价值和广阔应用前景.

编码器-解码器U型架构和多分支架构是两种主流的语义分割网络设计范式^[6].UNet^[7]作为U型架构的典型代表,采用卷积与池化提取特征,并通过转置卷积与跳跃连接恢复细节信息.另一类典型结构是多分支架构,其通过并行分支提取多路特征,提升分割精度.然而,各分支通常采用较浅的特征提取路径,表征能力有限,整体性能高度依赖分支间的深度融合,而简单地拼接或相加难以有效融合异构特征间的互补信息.因此,精确提取与高效融合特征,成为提升语义分割性能的关键^[8].

特征融合是语义分割的关键环节,旨在高效融合深层全局语义与浅层局部细节信息.而高质量的特征提取是实现高效融合的基础,因此优化特征提取过程至关重要.现有特征提取策略主要分为两类:其一是全局特征提取,侧重捕捉图像的整体结构与长距离依赖关系.代表性方法包括ACC-UNet^[9]中MLFC模块、DeepLabV3^[10]中ASPP结构,及PSPNet^[11]的PPM模块等,但上述方法在融合阶段侧重全局特征,易导致局部细节信息弱化;其二是局部特征提取,强调对边缘细节、纹理等局部区域的精细建模.代表性方法包括BiSeNetV2^[12]与PIDNet^[8],但其在融合策略上过度强化局部结构,导致高层语义信息表达受限,破坏语义一致性.

为兼顾全局语义理解与局部细节保留,部分研究尝试融合两类特征提取策略,力求在二者之间实现更优的协同融合效果.UNetV2^[13]利用跳跃连接和SDI模块,实现跨层特征融合;DDRNet^[14]则通过显式计算全局与局部特征相似性,指导融合过程.尽管上述方法采用多种策略整合全局与局部信息,但仍存在全局语义与局部细节特征难以平衡的问题^[8],难以兼顾大范围语义一致性与边界精细性^[15].此外,现有融合机制通常采

用线性加权范式,难以高效建模特征间复杂的非线性多变量关系,导致特征表达能力受限.尽管Fast-SCNN^[16]在特征融合后引入ReLU激活函数进行非线性变换,以增强特征表达,但本质上属于后置补偿机制,未能充分建模原始特征间的复杂非线性关系.

近年来,Kolmogorov-Arnold(KA)表示定理^[17]的理论框架为语义分割网络设计提供了新的理论支撑,尤其在缓解全局与局部特征失衡问题上展现出重要潜力.U-KAN^[18]通过将KAN替换U-Net中MLP模块,实现高效特征提取;KAT^[19]在计算机视觉任务中采用KAN替代Transformer的MLP模块,显著增强模型的特征表达能力和效率;Path-SAM2^[20]则设计KAN分类网络,有效增强病理特征编码与判别能力.这些最新进展表明,将KA表示定理应用于特征提取与网络结构设计,有望突破现有语义分割的性能瓶颈.KA表示定理将高维函数分解为单变量函数组合,深刻揭示了深度网络通过分层非线性变换与聚合操作逼近复杂映射的核心机制,这与特征融合中多尺度特征的层级组合具有理论同构性.尽管KA定理已在特征提取、增强及结构优化中展现应用价值,但其在特征融合阶段的潜力尚未被深入探索,相关研究仍待进一步拓展.

针对上述挑战,本文提出KA表示定理启发的图像语义分割动态融合网络模型,设计双层函数架构的动态特征融合模块.本文的主要贡献如下.

(1)构建多级特征提取框架:设计基于残差连接的U型上下文增强模块(ReSidual U-block Downsample, RSUD-ownsampling),实现全局语义特征的高效传递.提出基于差异图的特征细化模块(Spatial-Semantic Feature Enhancement, SSFE),利用二值化差异图映射自适应增强局部细节特征,强化局部细节信息的语义感知.

(2)在多级特征提取的基础上,提出KA表示定理启发的动态特征融合模块(Dynamic Kolmogorov-Arnold Fusion, DynKAFusion),构建双层嵌套函数融合机制,通过内外层函数实现特征深度提取和精准拟合,并结合可学习动态融合策略,有效提升全局语义与局部细节信息的协同融合效果.

(3)设计CE/Topk+Dice动态联合损失函数,并结合条件触发策略,进一步增强模型对困难样本的特征学习能力.

2 相关工作

2.1 图像语义分割

图像语义分割作为计算机视觉领域的核心任务之一,旨在实现像素级场景解析的同时满足高帧率推理需求. 早期 U-Net^[7]通过编码器-解码器架构及连接跳跃实现多级特征融合,验证跨层级特征融合在医学图像分割中的有效性. DeepLabV3^[10]结合多尺度特征融合机制,在边界保留与多尺度解析方面取得进展,但其对称编解码结构在高分辨率场景下易导致特征失配与计算冗余,促使研究者探索多分支异构架构. BiSeNet^[21]首创空间-上下文双路径并行框架,分别保留几何细节与捕获全局语义,然而其融合策略未能充分捕捉模态间协同效应. 为此, BiSeNetV2^[12]引入双边引导聚合两分支,提升跨模态信息融合效果,但在计算效率方面仍有提升空间. 为优化计算效率, Fast-SCNN^[16]提出共享低级特征提取模块,显著降低模型参数量,为嵌入式部署奠定基础. 近年来, PIDNet^[8]受控制理论启发构建三分支架构,缓解双分支模型的固有限制. Liu 等人^[22]与 Wang 等人^[23]引入 Mamba 解码器并设计跨模态注意力机制,在全局依赖建模与弱监督学习方面取得突破.

尽管图像语义分割在多分支架构、轻量化特征提取及高效融合策略方面取得进展,但在复杂场景下的表达能力仍需优化提升. 如何在保持高帧率推理的同时,精准建模全局语义与局部细节特征,仍是该领域的核心问题.

2.2 特征融合

在图像语义分割任务中,特征融合是提升模型性能的关键,其通过整合多来源、多层次的特征信息,增强模型对复杂场景的理解和识别能力. 该技术从早期的简单拼接演进至基于注意力的动态自适应融合,逐步实现由线性向非线性、由静态向动态的融合策略转变. 早期研究如特征金字塔^[24]和空间金字塔池化^[25]聚焦于多尺度特征提取与组合,但对局部细节处理不足,导致精细边界与小目标解析存在局限. 随着深度学习发展,残差连接^[26]有效缓解梯度消失问题,注意力机制的引入则进一步提升特征融合效果. Hu 等人^[27]通过学习通道间的依赖关系,对特征通道进行加权,增强模型对重要特征的专注. 全局上下文网络^[28]和双注意力网络^[29]则结合空间和通道维度的信息,实现全局与局部特征的自适应融合. 为应对多尺度和多模态特征融合需求,动态特征融合应运而生. Hu 等人^[30]结合特征相似性与空间连续性,实现无监督场景下的全局与局部特征融合. 为提升多尺度特征融合效果, Wang 等人^[31]构建渐进式金字塔网络,采用自适应空间融合操作,增强模型对跨尺度语义的整合能力. Qiao 等人^[32]则基于

多源特征选择性机制,提出自适应特征融合模型,实现对多源特征的动态筛选与整合. 进一步地, Xie 等人^[33]引入基于 Mamba 的动态特征增强模块,自适应补充高层语义中的细节纹理信息. Yang 等人^[34]则聚焦全局上下文建模,设计动态调控融合权重机制,提升模型在尺度变动场景下的鲁棒性. 此外, Wu 等人^[35]构建动态稀疏融合框架,兼具分割精度与计算效率.

尽管特征融合技术持续演进,当前研究仍面临全局语义与局部细节难以平衡、线性建模能力受限等挑战. 部分研究尝试引入非线性机制以弥补不足,但往往依赖额外网络层次或复杂参数,计算开销较大,限制其在大规模实时任务中的应用.

2.3 Kolmogorov-Arnold 表示定理

Kolmogorov-Arnold (KA) 表示定理由数学家 Arnold 和 Kolmogorov 提出. 该定理表明,任何定义在有界域上的多变量连续函数,均可表示为有限个单变量连续函数的两层嵌套组合形式. 1976 年,日本数学家志村五郎进一步探讨这一问题,相关研究最终推导出人工神经网络中通用近似定理,揭示人工神经网络可以任意精度逼近任意连续函数. 受 KA 表示定理启发, Kolmogorov-Arnold Network (KAN)^[17]被提出作为多层感知机的替代结构. 该网络在权重连接处引入可学习激活函数,节点仅进行加权求和,理论上仅需两层即可逼近任意有监督学习任务,展现出较强表达能力. 然而,其训练效率较低. 为提升模型参数效率与可扩展性, TaHT^[36]引入参数化样条函数并优化网络结构,在 MNIST 与 Fashion-MNIST 等任务中表现出良好性能.

得益于上述工作,KA 表示定理在语义分割任务中得到初步应用. 根据当前研究技术路径,可将相关工作归为以下三类:

(1) KAN 层与经典架构融合. Ma 等人^[37]首次将 KAN 模块引入语义分割任务中,采用 KAN 线性层替代传统 MLP,并结合深度特征精炼模块,在遥感图像中有效提升目标边缘与细节结构的恢复能力. Li 等人^[18]将 KAN 层集成于 U-Net 中间表示,通过卷积模块提取局部细节特征,再结合标记化 KAN 模块建模全局语义特征,有效提升整体分割性能.

(2) KAN 与注意力机制/多尺度结构集成. Yang 等人^[38]提出 KAN-Transformer 模块,结合 TransUNet 编码器中 Res2Net 结构与 SimAM 注意力机制,有效提升模型对复杂结构信息的感知能力,但其仅针对工业焊缝场景,未验证其他模态(如超声)的泛化性;针对超声图像低对比、边界模糊、ROI 尺度变化剧烈的挑战. Zhang 等人^[39]在 U 形网络中引入全阶段 Tok-KAN 块、FCSA 注意力与多尺度特征融合机制,显著增强超声图像中模糊边界与低对比区域的分割效果. Pang 等人^[40]则结合

Fourier-KAN层与自适应门控多尺度融合模块,实现多尺度语义的高效融合,但其依赖全监督学习,标注成本较高。

(3) KAN与任务优化策略结合. Li等人^[41]结合KAN与可变形卷积,显著增强对遥感影像小目标与边缘的分割能力. Lin等人^[42]设计Shift-Tokenized KAN块并引入自适应分组策略,优化医学图像中复杂病灶区域的特征提取效果. 虽引入AGS缓解通道分组不均问题,但未针对复杂病灶设计更强的特征交互机制. Li等人^[43]将KAN模块与半监督学习策略融合,通过设计U形网络加强非线性特征处理,但其对不规则病变识别精度有限. 针对高分辨率医学图像计算成本高、小病灶分割不准的问题, Yuan等人^[44]提出轻量化双路径结构,结合KA卷积与分支交互模块,实现跨层特征协同,虽在整体效率上具备优势,但在不规则病变识别上仍有改进空间。

尽管KA表示定理在语义分割任务中已展现初步潜力,现有研究仍主要集中于特征提取与结构设计层

面,而对其在特征融合阶段的系统性探索仍显不足. 为此,本文从特征融合角度出发,提出受KA定理启发的双层函数动态融合模块,将特征融合过程分解为内层非线性变换、动态加权融合及外层非线性聚合三阶段,提升融合表达能力。

3 本文方法

3.1 整体架构

针对当前语义分割任务在复杂场景下,全局语义信息与局部细节信息在特征融合过程中难以平衡的问题,提出KA表示定理启发的图像语义分割动态融合网络模型(DynKANet). 在双分支语义分割架构基础上,从特征提取(包含下采样和双分支特征提取)和特征融合两方面进行创新改进,提出基于残差连接的U型上下文增强模块(RSUDownsampling)、基于差异图的特征细化模块(SSFE)和受KA表示定理启发的动态特征融合模块(DynKAFusion)等创新模块,模型整体架构如图1所示。

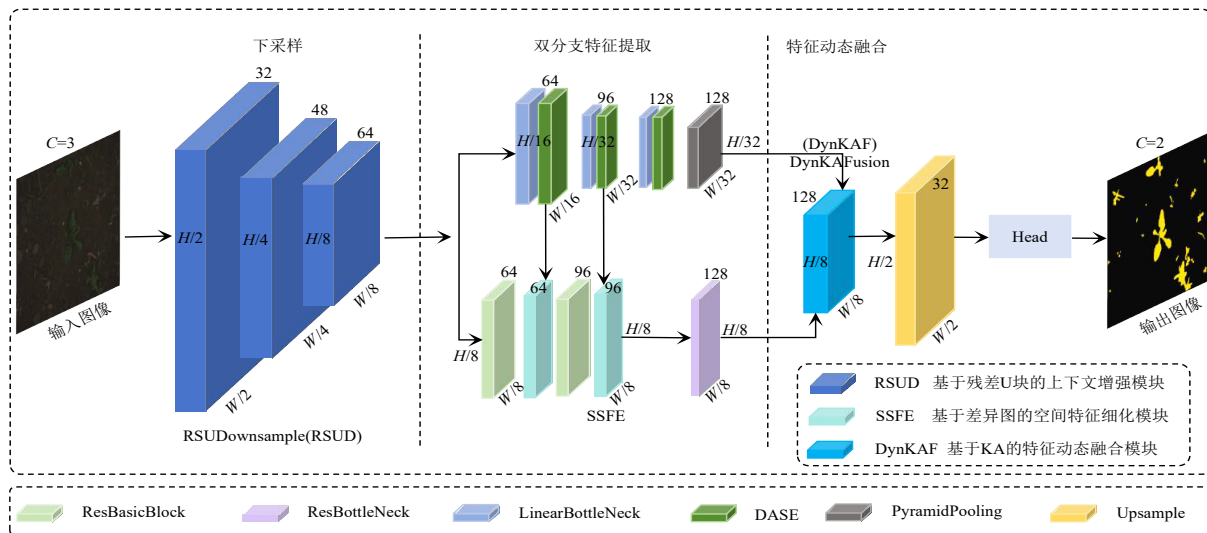


图1 DynKANet整体框架

3.2 RSUDownsampling

在经典语义分割框架中,传统下采样操作通常通过串联卷积层与池化层逐步扩大感受野,提取深层次语义信息. 然而,此类设计存在固有局限:单一尺度特征提取难以建立跨层次上下文关联,易导致全局语义建模能力不足. 为缓解上述问题,现有改进方法多引入金字塔结构,实现多尺度特征提取与融合^[45,46],但通常伴随参数冗余与计算开销增加,难以在特征表达与计算效率之间取得平衡。

为解决上述问题,本文提出基于残差连接的U型上下文增强模块(RSUDownsampling),利用RSU(Re-

sidual U-block)进行多尺度特征提取与融合,并结合深度不对称注意力机制(Depthwise Asymmetric Squeeze-and-Excitation, DASE)进行特征重组,在控制参数增量的同时,强化全局上下文建模能力。

如图2所示,RSUDownsampling模块借鉴U2-Net^[47]中RSU块的设计思想,采用对称的U-Net型编码器-解码器结构,并通过设置不同深度L的RSU单元,实现多层次特征提取. 具体而言,依次级联深度为7、6和5的多尺度RSU单元,并在其后接入步幅为2的最大池化层,以提升多尺度语义提取能力与上下文感知效果. 此外,为进一步降低计算开销,引入轻量化REBNCNConv结构

(ReLU BatchNorm Depthwise Separable Convolution), 以深度可分离卷积代替传统卷积, 显著降低计算复杂度。

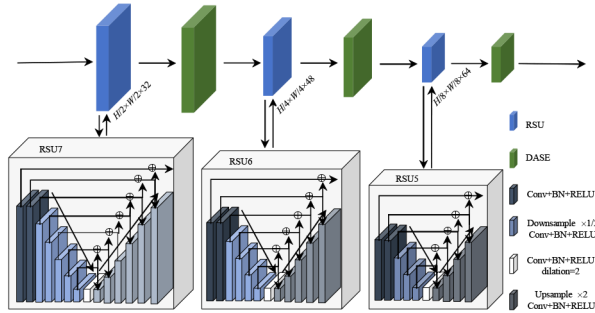


图2 RSUDownsample 模块结构

然而, 仅依靠多尺度特征提取难以有效建模全局语义信息. 为此, 本文进一步提出深度不对称注意力机制(DASE), 强化特征重组过程, 提升全局语义信息的建模能力. DASE 模块融合深度可分离卷积^[48]、不对称卷积^[49]和膨胀卷积^[50]的优势, 有效提升模型对多尺度上下文的建模能力, 兼具计算效率与特征表达能力. 具体而言: (1) 深度可分离卷积采用逐通道和逐点卷积, 在降低计算复杂度的同时保持通道特征独立性; (2) 不对称卷积将标准卷积分解为多方向低维卷积, 减少参数并增强方向感知; (3) 膨胀卷积则在不增加计算量的前提下扩展感受野, 提升全局上下文建模能力. 此外, 为聚焦关键特征, 引入多级 SE(Squeeze Excitation) 注意力机制, 并结合 Shuffle-Block 实现通道重排, 进一步增强全局特征表达能力。

3.3 SSFE

在语义分割任务中, 有效融合全局语义与局部细节特征是提升分割精度的关键. 然而, 现有特征融合方法^[51-53]往往难以在语义完整性与细节准确性之间取得平衡. 鉴于低分辨率特征富含全局语义, 而高分辨率特征聚焦局部细节, 且低分辨率向高分辨率传递计算更为高效, 本文提出基于差异图的特征细化模块(SSFE), 通过构建单向特征传递路径, 实现语义分支对细节分支的引导学习, 提升局部细节特征的语义感知能力. 在全局语义的指导下, 局部细节分支能够在特征学习时保持语义一致性, 聚焦关键区域并抑制冗余干扰, 显著提升模型泛化能力。

如图 3 所示, SSFE 模块通过语义分支引导细节分支, 提取语义增强的细节特征, 进而缓解局部建模中全局语义信息缺失的问题. 具体而言, 输入图像经主干网络共享下采样后, 分别提取局部细节分支特征图 P_1 与全局语义分支特征图 P_2 , 随后输入 SSFE 模块进行处理. 在 SSFE 模块中, 对齐后的特征图 P_1 与 P_2 通过逐元素相减并计算平方均值的方式, 量化生成差异图, 如

式(1)所示:

$$D = \frac{1}{C_h} \sum_{c=1}^{C_h} (H^{(c)} - L^{(c)})^2 \quad (1)$$

其中, $H, L \in \mathbf{R}^{C_h \times H \times W}$ 分别表示对齐后的高分辨率特征与低分辨率特征, C_h 表示高分辨率特征图的通道数, D 量化全局语义与局部细节特征的空间偏差, 高值区域表示两者语义不一致性显著。

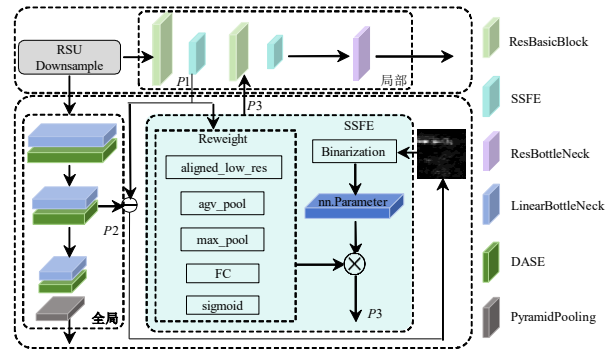


图3 SSFE 模块结构

差异图本质上用于量化全局语义与局部细节特征之间的差异, 反映高分辨率特征图与低分辨率特征图在语义上的偏差. 然而, 作为连续值标量图, 差异图的数值范围较宽且缺乏明确语义边界, 直接用于特征增强可能导致信息冗余, 且无法有效突出关键语义信息. 为此, 将计算得到的差异图进行二值化处理, 以强化语义边界并抑制冗余. 具体而言, 引入可学习阈值 $\tau \in \mathbf{R}$, 生成离散语义掩码 M , 通过硬判决函数实现二值化, 如式(2)所示:

$$M = I(D \geq \tau) \in \{0, 1\}^{1 \times H \times W} \quad (2)$$

其中, $I(\cdot)$ 为硬判决示性函数, 差异图阈值 τ 初始设为 0, 并在每次训练开始时加入微量高斯噪声 ($\varepsilon \sim N(0, \sigma^2)$, $\sigma = 1 \times 10^{-3}$), 以兼顾初始的中立性, 同时通过随机扰动加速收敛. 训练过程中, τ 通过反向传播进行动态优化, 实现语义显著与非显著差异区域的自适应区分。

考虑到将差异图与细节特征图直接逐元素融合可能削弱特征精度与关键区域响应, 本文引入重加权优化策略, 增强局部细节特征表达能力. 具体而言, 该策略采用双路池化注意力机制, 结合全局平均与最大池化提取特征统计量, 并通过多层感知机计算特征图的权重系数. 最终, 将二值化差异图与加权后的局部细节特征图进行逐元素相乘, 得到特征图 P_3 , 实现对语义显著区域的选择性增强。

3.4 DynKAFusion

本节首先回顾传统特征融合的限制性, 强调非线性融合的必要性, 并引入 KA 表示定理作为理论支撑. 基于此, 提出 DynKAFusion 模块, 并对其设计进行详细

阐述. 最后介绍模块的结构与实现流程.

3.4.1 传统特征融合

在图像语义分割任务中, 高效融合全局语义与局部细节是提升模型表达能力的关键^[8,54]. 然而, 现有特征融合方法面临两大关键挑战: (1) 特征融合过程中易出现尺度与语义失配问题, 导致全局语义与局部细节难以兼顾, 进而削弱语义一致性或丢失关键细节信息; (2) 传统特征融合方法多依赖线性加权机制, 难以建模特征间复杂的非线性关系, 限制融合表达能力. 后续部分研究尝试引入更复杂的结构提升融合能力, 但通常伴随计算成本显著上升, 限制其在实时语义分割场景中应用.

3.4.2 KA表示定理相关性阐述

近年来, KA表示定理为语义分割特征融合网络设计提供了全新理论启示. 该定理从数学上揭示复杂映射的“可分解性”, 即任意连续函数均可通过有限次单变量非线性映射逼近, 如式(3)所示:

$$f = \sum_{q=1}^{2n+1} \phi_q \left(\sum_{p=1}^n \psi_{q,p}(x_p) \right) \quad (3)$$

在多尺度特征融合过程中, 网络通过逐层特征提取构建层次化的特征表达结构. 其中, 内层函数 $\psi_{q,p}$ 负责对输入特征进行非线性变换, 可视为低层特征的深度抽象过程, 用于提取更具层次的全局语义与局部细节信息; 而外层函数 ϕ_q 则侧重于对前述深层特征进行非线性聚合, 可类比为高层特征融合操作, 旨在精准整合来自不同尺度与维度的关键信息. 二者通过层次化嵌套设计, 实现全局语义与局部细节特征的高效融合, 这一过程与KA表示定理中“先分解后合成”两阶段思想高度契合.

3.4.3 DynKAFusion模块设计

依据KA表示定理, 任意维度的连续函数理论上均可通过有限数量的“内层-外层”嵌套函数组合逼近. 传统方法常用B样条构建嵌套函数, 但其计算复杂度较高, 训练时间成本较大. 为优化计算效率, 本文借鉴Bodner等人^[55]提出的卷积神经网络架构, 在保留卷积神经网络局部感知与权重共享优势的基础上, 采用可学习样条插值核(KANConv)替代传统基函数, 构建具备更强非线性建模能力的卷积网络.

然而, 标准KAN卷积在增强非线性建模能力的同时, 通常伴随较大参数冗余与计算成本. 为此, 本文受Drokin等人^[56]提出的KAN卷积层参数高效设计理念启发, 并借鉴ResNet瓶颈结构, 进一步设计瓶颈式KAN卷积(BKAN-Convs). 该结构采用“1×1压缩→KAN核心运算→1×1扩展”三阶段变换架构, 能够在保持标准KAN卷积非线性表达优势的同时, 有效控制参数开销, 提升计算效率.

同时, 考虑到语义分割网络中不同层级对非线性表达的差异化需求, 本文在内外层函数设计中采用“标准KAN→瓶颈式KAN”的递进式配置策略, 以兼顾性能与效率. 具体而言, 浅层侧重结构细节与纹理边缘的建模, 特征变化剧烈, 因而更依赖强非线性表达能力, 适合保留标准KAN结构; 而深层聚焦高阶语义整合, 特征分布相对平稳, 更适合引入瓶颈KAN结构以提升整体运算效率.

在此基础上, 针对全局语义与局部细节难以平衡的问题, 本文进一步引入可学习动态加权系数, 构建包含两步高效函数变换的三阶段特征融合结构, 实现有限参数条件下对复杂融合的有效逼近. 具体而言, 将特征融合建模为内层非线性变换、动态融合及外层非线性聚合的协同过程: 首先通过内层非线性变换提升各路特征的表达能力, 随后引入动态加权机制实现自适应融合, 最终通过外层非线性聚合强化全局与局部特征间的协同表达.

为此, 提出特征融合模块DynKAFusion的具体设计, 模块原理如图4所示.

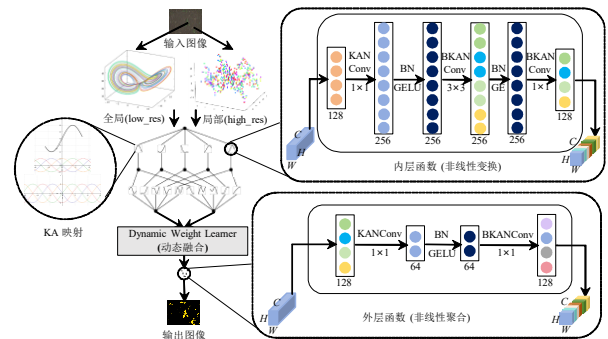


图4 DynKAFusion模块原理

(1) 内层函数. 针对高低分辨率输入特征, 分别引入基于KAN卷积的三阶段递进式变换结构: 首先采用标准KAN卷积模块, 结合标准卷积与可学习样条插值, 实现初始高维表达; 随后叠加两级瓶颈式KAN卷积模块, 先进行非线性耦合以增强特征表征能力, 再执行维度压缩以聚焦关键信息.

(2) 动态融合. 借助可学习加权系数动态调整两路特征流的融合比例, 自适应整合全局语义与局部细节信息在融合时的互补优势.

(3) 外层函数. 针对动态加权融合后的特征, 采用双阶段聚合结构, 实现深层非线性表征: 首先通过浅层标准KAN卷积进行维度收缩, 实现信息蒸馏与冗余抑制; 随后级联引入瓶颈式KAN卷积进行维度恢复, 实现特征重构与语义增强.

3.4.4 DynKAFusion模块实现

DynKAFusion实现由“内层非线性变换”与“外层动

态非线性聚合”两阶段构成. 对于输入的高分辨率特征 F_{high} 与低分辨率特征 F_{low} , DynKAFusion 模块整体流程为

$$F_{\text{DKAF}} = \text{KANConv}_{\phi} \left(\omega^h \text{KANConv}_{\psi} (F_{\text{high}}) + \omega^l \text{KANConv}_{\psi} (F_{\text{low}}) \right) \quad (4)$$

其中, $\text{KANConv}_{\psi}(\cdot)$ 为内层函数, $\text{KANConv}_{\phi}(\cdot)$ 为外层函数, ω^h 和 ω^l 分别为可学习动态融合权重.

在内层非线性变换阶段, 首先将输入特征通过标准 KAN 卷积映射至高维非线性空间 (128 \rightarrow 256), 同时施加 BatchNorm2D 稳定训练过程, 并利用 GELU 激活函数引入非线性表达, 增强特征细节建模能力, 为后续复杂变换提供充分表征空间. 具体实现如式 (5) 所示:

$$\text{KANConv}_{\psi_0}(x) = \sigma \left(\text{BN} \left(\text{KANConv}_{1 \times 1}^{256}(x) \right) \right) \quad (5)$$

随后, 通过 3×3 瓶颈式 KAN 卷积捕捉精细特征信息, 并再次施加 BatchNorm2D 和 GELU 激活. 该阶段借助 KAN 卷积的可学习样条插值特性, 增强对复杂特征关系的拟合能力. 具体实现如式 (6) 所示:

$$\text{KANConv}_{\psi_1}(x) = \sigma \left(\text{BN} \left(\text{BKANConv}_{3 \times 3}^{256} \left(\text{KANConv}_{\psi_0}(x) \right) \right) \right) \quad (6)$$

最后, 通过 1×1 瓶颈式 KAN 卷积将通道数降回 128. 该降维过程通过压缩机制引导模型关注关键信息, 促使其学习更具判别性的紧凑表征, 如式 (7) 所示:

$$\text{KANConv}_{\psi}(x) = \text{BKANConv}_{1 \times 1}^{128} \left(\text{KANConv}_{\psi_1}(x) \right) \quad (7)$$

其中, $\sigma(\cdot)$ 为 GELU 激活函数, $\text{KANConv}_{1 \times 1}^{256}$ 用于升维以提供充分表征空间, $\text{BKANConv}_{3 \times 3}^{256}$ 用于捕捉细致特征信息, $\text{BKANConv}_{1 \times 1}^{128}$ 则用于降维以聚焦关键特征信息. 具体而言, KAN 标准卷积 (KANConv) 包含标准卷积核 W 与可学习样条插值核 S , 并通过可学习缩放因子 α 实现动态融合, 最终输出为

$$\text{KANConv}(X) = \text{Conv}(X, W) + \alpha \cdot \text{SplineConv}(X, S) \quad (8)$$

其中, $X \in \mathbf{R}^{\beta \times C_m \times H \times W}$ 表示输入特征图, α 用于调节样条

分支权重, 提升模型对不同通道特征的重要性建模能力. 瓶颈式 KAN 卷积 (BKANConv) 在此基础上引入通道压缩与扩展机制, 提高模型的计算效率.

在外层动态非线性聚合阶段, 高低分辨率特征经内层变换后可与学习权重相乘, 实现动态融合. 随后, 外层函数对初步融合后特征进行非线性聚合. 该阶段主要采用由 1×1 标准及瓶颈版 KAN 卷积、BatchNorm2D 与 GELU 激活级联构成的非线性映射操作.

具体而言, 首先通过维度收缩实现信息蒸馏, 缓解动态融合过程中潜在的冗余信息及语义冲突. 该过程有助于抑制噪声, 引导模型聚焦于关键特征, 提升特征表征的紧凑性与判别性, 如式 (9) 所示:

$$\text{KANConv}_{\phi_0}(x) = \sigma \left(\text{BN} \left(\text{KANConv}_{1 \times 1}^{64}(x) \right) \right) \quad (9)$$

随后进行维度恢复与特征重构, 引导特征向语义一致与判别增强方向优化, 提升模型对复杂目标的刻画能力, 完成最终的特征增强, 如式 (10) 所示:

$$\text{KANConv}_{\phi}(x) = \text{BKANConv}_{1 \times 1}^{128} \left(\text{KANConv}_{\phi_0}(x) \right) \quad (10)$$

为直观呈现两路特征融合的整体流程与具体实现结构, 图 5 依次展示内层函数、动态加权学习器以及外层函数的层级衔接关系, 并将高分辨率 (局部) 与低分辨率 (全局) 特征的动态融合进行可视化呈现, 全面揭示 DynKAFusion 模块的设计实现.

3.5 动态损失函数设计

为增强模型对困难样本的表征能力, 本文采用 CE/TopK+Dice 动态联合损失函数 [见式 (11)] 监督网络训练, 并设计条件触发机制, 在交叉熵或 TopK 损失^[57] 与 Dice 损失之间动态切换, 从训练层面优化模型的泛化能力, 与 DynKAFusion 模块在结构层面引导的全局语义与局部细节融合形成良性协同. 该动态策略包含 2 个关键阶段, 其切换逻辑与 TopK 损失^[57] 的特性紧密

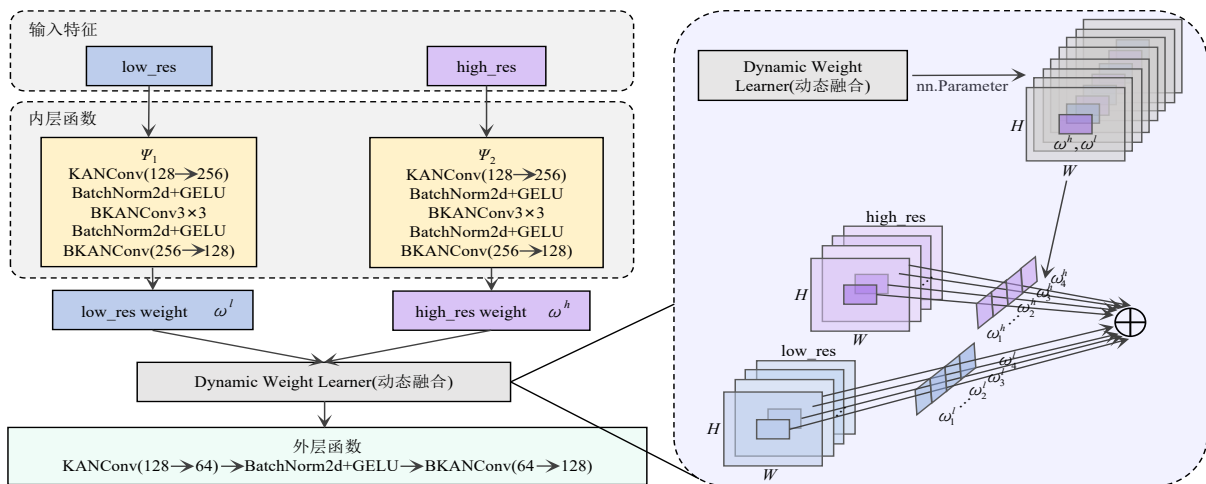


图 5 DynKAFusion 模块结构

相关。

第一阶段为基础特征学习阶段,此时TopK损失尚未触发,模型采用CE+Dice组合.CE损失[见式(12)]通过最小化预测与标签之间的差异,实现像素级多类分类;Dice损失[式(13)]则基于预测区域与真实区域的重叠程度,缓解类别不平衡及背景主导问题.两者联合有助于建立稳健的初始特征表示.

$$L = L_{CE}/L_{TopK} + L_{Dice} \quad (11)$$

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(p_{i,c}) \quad (12)$$

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N \sum_{c=1}^C p_{i,c} y_{i,c}}{\sum_{i=1}^N \sum_{c=1}^C p_{i,c}^2 + \sum_{i=1}^N \sum_{c=1}^C y_{i,c}^2} \quad (13)$$

$$L_{TopK} = \frac{1}{K} \sum_{i \in \Omega_K} L_{CE}^{(i)} \quad (14)$$

第二阶段为困难样本优化阶段,此时TopK损失[见式(14)]满足触发条件,损失函数切换为TopK+Dice组合.此阶段模型已具备基本判别能力,TopK通过截断梯度,引导训练聚焦于高损失的困难样本,提升决策边界的判别精度;Dice损失则持续约束整体区域结构,防止分割性能退化,增强模型鲁棒性.

基于此,为强化模型在难分类区域的表现,本文设计两种具有不同触发依据的动态损失调整策略,分别针对模型训练进度与性能变化的不同特性进行优化调控.

动态策略1基于模型性能实现自适应调整:考虑到训练过程中模型收敛速度存在波动,当训练集与验证集的mIoU在连续2个epoch内均无提升时,则认为模型对易分类样本的学习已趋近饱和,此时切换至TopK+Dice组合,以增强对困难样本的优化能力;反之,则保持原有组合,确保训练过程的稳定性.该自适应机制动态调节模型对不同难度样本的关注,有效提升整体分割性能与鲁棒性.

动态策略2基于训练进度设定切换机制:考虑到模型在训练早期侧重全局特征学习,后期逐步转向细粒度优化,因此当训练进度超过80%(即 $r=0.8$)时,将原先的CE+Dice损失切换为TopK+Dice组合.此阶段,TopK损失仅保留损失值排名前20%的像素(即 $K=0.2$)参与梯度更新,强化模型对高难度样本的关注,提升决策边界及复杂区域的分类精度.该切换机制有效引导训练重点的动态调整,促进模型性能提升与泛化能力增强.

4 实验结果及分析

4.1 数据集介绍

为系统评估所DynKANet模型在语义分割任务中

的性能及其跨场景适应性,如表1所示,本文选取10个具有代表性的公开数据集进行实验,涵盖钢材缺陷、医学影像、城市道路、月球景观与农作物种植等不同领域,场景特征差异显著.

表1 5个领域的10个语义分割数据集

数据集	年份	类别	规模	分辨率	类型
NEU ^[57]	2024	4	4430	200×200	缺陷检测
FSSD-12 ^[58]	2023	2	600	200×200	缺陷检测
Synapse ^[59]	2023	2	1 807	不固定	医学影像
ACDC ^[60]	2017	4	943	不固定	医学影像
KITTI ^[61]	2012	2	289	1 242×375	城市道路
TP ^[62]	2023	2	1 391	2 976×1984	城市道路
LuSNAR ^[63]	2024	5	2 196	1 024×1 024	月球景观
人工月球景观	2023	4	5656	720×480	月球景观
Sugar Beets ^[64]	2016	2	1 321	1 296×966	农作物种植
PlantSegV2 ^[65]	2024	2	7 909	448×448	农作物种植

注:人工月球数据集网址:<https://www.cvmart.net/dataSets/detail/733>.

不同任务场景对模型特征建模能力提出差异化需求.具体而言,钢材图像纹理细腻、对比度低,要求模型具备较高的局部细节捕捉能力;医学影像边界模糊、结构多变,强调对模糊区域的边缘感知与结构鲁棒性;城市道路图像尺度差异显著、遮挡频繁,要求模型强化对上下文信息的建模能力;月球图像纹理弱、噪声强,考验模型在复杂背景中提取全局语义与边界信息的能力;农作物病害图像病斑类型复杂、叶片遮挡明显,则更依赖细节保留与类别判别的协同建模.对于上述5类数据集,训练集、验证集和测试集的划分比例为6:2:2.

4.2 实验设置及评价指标

实验在Ubuntu 22.04操作系统上完成,使用24 GB RTX 4090D GPU处理器,搭载Python 3.10和PyTorch 2.1.2深度学习框架.训练共进行200轮,每批次包含8张图像.学习率设置为 1×10^{-6} ,并使用RMSprop优化器进行参数更新,权重衰减系数为 1×10^{-8} ,动量为0.999.此外,为适应模型输入要求,图像可按比例scale=0.5进行缩放处理.

采用平均交并比(mIoU)、F1分数、召回率(Recall)、精确率(Precision)、准确度(Accuracy)、模型参数量(Params)及计算复杂度(FLOPs)共7个指标对模型性能进行全面评估.其中,mIoU表示各类别预测区域与真实区域交并比的平均值[见式(15)],是衡量分割精度的核心指标.

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \left(\frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \right) \quad (15)$$

4.3 损失函数对比分析

为确定最终损失函数方案,本文在统一实验配置下进行 100 轮训练,基于 5 类数据集开展 mIoU 指标对比实验,统计各组合的平均值与标准差,以系统评估不同损失函数组合的综合表现. 鉴于动态策略 2 在不同数据集上对应的最优参数组合存在差异,统一设定参数为 $r=0.80$ 、 $K=0.20$,具体最优组合详见后续超参数分析. 实验结果(见表 2)表明,CE/TopK+Dice(动态 2)组合在五类数据集上平均性能优于其他组合,充分验证该损失函数组合的有效性,表中最优结果以加粗字体

表示.

进一步分析表明,CE/TopK+Dice(动态 2)在多方面展现出显著优势:(1)CE 损失有效捕捉类别判别信息,提升像素级分类精度;(2)Dice 损失缓解语义分割中类别不平衡问题,有助于保持整体区域结构一致性;(3)TopK 损失聚焦训练中困难样本,引导模型学习更具区分性的特征表示;(4)动态调整机制则增强损失函数对不同数据分布的适应性,提升训练过程的稳定性与灵活性. 综合来看,CE/TopK+Dice(动态 2)在提升语义分割性能的同时,还具备较强泛化能力,因而被确定为最终优化方案.

表 2 五类数据集下不同损失函数组合的语义分割平均精度结果(epoch:100)

损失函数	mIoU(mean±std)	F1(mean±std)	Recall(mean±std)	Precision(mean±std)	Accuracy(mean±std)
交叉熵(CE)	0.788 6±0.13	0.880 9±0.06	0.875 2±0.11	0.877 8±0.06	0.933 2±0.02
Dice Loss	0.629 4±0.22	0.713 2±0.24	0.698 4±0.21	0.712 8±0.24	0.916 3±0.06
TopK Loss ^[56]	0.598 8±0.36	0.697 3±0.28	0.929 7±0.04	0.644 0±0.34	0.849 5±0.22
CE+Dice Loss	0.781 9±0.10	0.873 2±0.07	0.859 6±0.13	0.885 3±0.15	0.919 5±0.06
CE/TopK+Dice(动态 1)	0.786 4±0.11	0.878 3±0.04	0.866 3±0.03	0.883 4±0.08	0.922 9±0.03
CE/TopK+Dice(动态 2)	0.801 9±0.12 ↑	0.888 2±0.07 ↑	0.892 1±0.09 ↑	0.879 6±0.05 ↑	0.929 9±0.01 ↑

注:动态 1 为基于 mIoU 趋势的自适应 TopK 损失切换策略;动态 2 为 $r=0.80$ 、 $K=0.20$ 的 TopK 损失条件策略.

4.4 超参数分析

为进一步验证 DynKANet 在语义分割任务中的稳定性与鲁棒性,本文在 5 类不同数据集上开展动态损失函数中关键超参数 r 与 K 的敏感性分析,训练轮数设为 100,评估其在不同场景下的性能影响.

在损失函数的设计与优化过程中,预设比例因子 K 作为 TopK Loss^[57] 的核心参数,直接影响损失计算过程中选取的损失值比例,进而在更大程度上决定损失函数的关注焦点与整体收敛特性. 相比之下,超参数 r 作为损失策略切换的阈值参数,主要调控从基础损失函数(CE+Dice)向联合损失函数(TopK+Dice)过渡的时机,侧重于动态训练阶段的策略调节. 考虑到 K 值对模型优化路径与梯度特性的基础性影响更为显著,因此本文优先固定 $K=0.2$,确保训练过程的收敛性与可控性;在此基础上,选取集合 {0.75, 0.80, 0.85, 0.90, 0.95} 内数值对 r 值进行系统性实验,探索不同 r 值对模型性能的影响,以确定策略切换的最优时机. 随后,在固定最优 r 值的前提下,对 K 值在 0.1~0.5 区间内以步长 0.05 进行精细化调节,进一步优化联合损失函数的表现. 上述 2 个超参数的优化过程均基于 mIoU 指标进行定量分析,以验证各参数配置下模型性能的最优组合.

实验结果如图 6、图 7 所示. 由图 6 可见,不同数据集在特定 r 值下均取得最优的 mIoU 性能:NEU_Seg、ACDC、KITTI、LuSNAR 和 PlantSegV2 数据集分别在 $r=0.80$ 、0.95、0.90、0.85 和 0.85 时达到峰值. 图 7 显示,当

预设比例 $K=0.2$ 时,NEU_Seg、ACDC、KITTI 和 LuSNAR 数据集的 mIoU 仍为最优值,而 PlantSegV2 数据集在 $K=0.3$ 时取得最佳性能. 由此可看出, $K=0.2$ 是 TopK Loss 的最佳比例损失值,能够在多数数据集上实现最优性能. 鉴于不同数据集在最优性能下对应的 r 和 K 值存在差异,后续对比实验均采用各自最优参数组合进行训练.

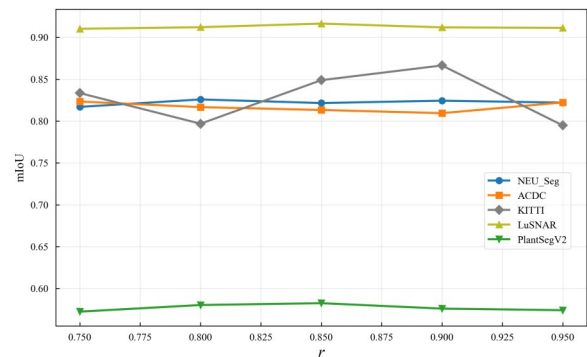


图 6 r 值超参数敏感性分析实验结果

为进一步评估超参数 r 和 K 对模型性能的敏感性,本文基于 mIoU 指标对不同参数设定下的波动幅度进行系统分析. 具体而言,NEU_Seg 数据集在 r 值与 K 值变化时,mIoU 波动幅度分别为 0.008 9、0.010 5,表明其对 r 、 K 两个参数的敏感性均较低;ACDC 数据集下 mIoU 的变化幅度分别为 0.013 9、0.018 1,表现出较高敏感性;KITTI 数据集的 mIoU 分别波动 0.071 6、0.084 0,反映出极高

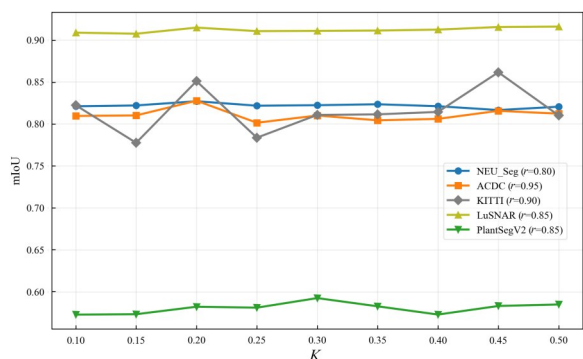


图7 K值超参数敏感性分析实验结果图

的参数敏感性; LuSNAR 数据集的变化幅度分别为 0.006 2、0.008 5, 敏感性整体偏低; PlantSegV2 数据集的变化幅度为 0.010 1、0.019 7, 显示出较高的敏感性。

综上所述, r 值和 K 值对模型性能具有显著影响, 且不同数据集在参数敏感性上存在明显差异. KITTI 数据集作为自动驾驶领域基准数据集, 因其复杂的场景结构、多变的视角与光照条件, 对 r 和 K 值表现出较高敏感性. 该现象表明, 合理设置 r 值和 K 值可有效调控损失函数的动态调整机制: 在训练初期引导模型聚焦全局特征提取, 而在模型收敛阶段加强目标区域的细粒度表达, 进而实现全局语义与局部细节的协同优化, 提升语义分割性能。

表3 5类数据集下不同特征融合模块关键性能指标对比

特征融合模块	FLOPs/G	Params/M	NEU_Seg	ACDC	KITTI	LuSNAR	PlantSegV2
			mIoU	mIoU	mIoU	mIoU	mIoU
简单求和	1.31	1.58	0.827 0	0.843 9	0.831 0	0.921 7	0.586 5
固定权重求和	1.31	1.58	0.830 4	0.840 3	0.801 4	0.923 1	0.585 1
自适应加权	1.31	1.58	0.835 0	0.835 8	0.847 5	0.921 1	0.581 8
注意力机制	1.35	1.61	0.839 9	0.829 6	0.810 8	0.924 3	0.584 2
DynKAFusion	1.60 ↑	2.26 ↑	0.861 6 ↑	0.875 6 ↑	0.879 8 ↑	0.959 4 ↑	0.618 2 ↑

从计算复杂度 (FLOPs) 与模型参数量 (Params) 来看, 简单求和、固定权重求和、自适应加权 3 种特征融合方法计算量较小, FLOPs 均为 1.31 G, 模型大小仅为 1.58 M, 但在性能表现上表现欠佳, mIoU 值普遍较低. 以 NEU_Seg 数据集为例, 三者的 mIoU 分别为 0.827 0、0.830 4、0.835 0; 在 LuSNAR 数据集上分别为 0.921 7、0.923 1、0.921 1, 表明线性融合策略仅能完成浅层特征叠加, 难以建模复杂的特征依赖关系, 充分验证 KA 表示定理关于非线性变换重要性的理论洞察. 引入注意力机制后, 模型对关键信息的关注度有所提升, 在 NEU_Seg 数据集上 mIoU 提升 1.29%, 但计算复杂度与模型参数量有所增加 (FLOPs 上升至 1.35 G, Params 增至 1.61 M), 且在 ACDC、KITTI 等数据集上性能提升不明显, 进一步说明缺乏 KA 式双层嵌套结构的注意力机

4.5 DynKAFusion 模块分析

本节基于 KA 表示定理的双层嵌套函数理论框架, 系统分析现有特征融合方法的理论局限性: (1) 简单求和与固定权重求和缺乏内外层非线性变换函数, 难以实现有效特征映射; (2) 自适应加权机制虽具动态性, 但本质仍属于线性插值, 难以刻画复杂的特征交互关系; (3) 传统注意力机制虽引入非线性注意力权重, 但缺乏对输入特征本身的非线性变换, 未能充分发挥 KA 定理强调的双层嵌套建模优势。

针对上述问题, DynKAFusion 模块严格遵循 KA 表示定理的设计范式: 首先通过内层函数对多路输入特征应用独立的非线性变换, 提升表达能力; 随后引入动态加权机制, 实现自适应融合; 最终由外层函数对融合结果进行非线性聚合, 构建统一表达. 在理论上, 该结构具备良好的融合表达潜力。

4.5.1 DynKAFusion 模块定量分析

为全面验证 DynKAFusion 模块在特征融合中的有效性, 在 5 类数据集上进行消融实验. 实验以 DynKANet 为基线模型, 通过替换不同的特征融合模块, 与 4 种典型融合策略 (简单求和、固定权重求和、自适应加权、注意力机制) 进行对比分析, 系统评估各特征融合方法在计算复杂度、模型参数量及语义分割性能方面的表现, 实验结果如表 3 所示, 表中最优结果以加粗字体表示。

制, 在特征交互建模方面仍存在瓶颈, 难以实现深层次的特征融合。

相较于上述融合机制, DynKAFusion 模块通过设计内外层函数实现非线性映射, 并引入动态融合策略, 进一步提升全局语义和局部细节特征交互度, 特征融合效果明显增强. 在计算资源方面, FLOPs 小幅上升至 1.60 G, 参数量微增至 2.26 M. 虽带来一定复杂度增长, 但在所有数据集上均取得最优性能 (在 KITTI 数据集上提升 4.88%, 于 PlantSegV2 数据集提升达 3.17%), 体现出良好的性能-效率平衡. 上述结果充分表明, KA 表示定理启发下的动态融合范式, 能够有效增强模型的结构语义感知与特征判别能力, 展现出较强的任务适应性与泛化能力。

综合上述实验结果, DynKAFusion 引入 KA 定理启

发的双层嵌套结构后, 相较其他融合策略在 mIoU 指标上表现出更优性能, 并实现了良好的参数-性能平衡, 充分验证其结构设计的合理性与应用潜力. 然而, 相较传统融合策略, 其参数规模与计算开销仍较高. 为进一步提升其实用性, 未来可结合模型剪枝等轻量化技术, 在保持精度的前提下有效降低计算成本, 增强模型的部署灵活性.

4.5.2 DynKAFusion 模块定性分析

为直观展示 DynKAFusion 的有效性, 在 NEU_Seg 数据集上, 将其与 4 种典型融合策略进行对比定性分析. 如图 8 所示, 通过特征热力图与类激活图, 展示各模块在训练过程中关注的关键特征及其空间分布. 其中, 红黄色区域表示网络关注区域.

从特征热力图 8(b) 来看, DynKAFusion 模块在目标定位上展现出明显优势. 在简单求和与固定权重求和的热力图中, 目标区域激活较为弥散, 存在较多非目标区域干扰; 相比之下, DynKAFusion 的热力图则清晰突出目标区域, 激活区域更为聚焦, 有效抑制背景噪声, 与目标语义区域高度吻合. 进一步对比自适应加权与注意力机制, 可见 DynKAFusion 热力图分布更为连贯, 有效避免因局部特征过度放大而导致的关键语义信息割裂现象, 体现出更强的结构一致性和语义完整性. 该优异表现主要得益于其通过内层函数实现非线性特征变换, 协同外层函数进行非线性聚合调控, 实现局部细节与全局语义的协同建模与动态融合, 有效提升特征表达的整体协调性与完整性.

进一步观察类激活图 8(c) 可知, DynKAFusion 在关键区域类别判别方面同样表现优越. 如图所示, DynKAFusion 类激活图在关键区域呈现出强烈且集中的响应信号, 激活边界清晰锐利, 与真实语义边界高度

契合, 体现出双层嵌套结构在增强特征判别性方面的结构优势, 进一步验证在目标类别特征判别中的显著优势. 相比之下, 简单求和与固定权重求和因缺乏有效的特征选择机制, 类激活图响应较弱且易受背景干扰; 而自适应加权与注意力机制虽有所提升, 但在多类别重叠区域仍易出现类别混淆与边界模糊, 限制其在精细语义分割中的表现.

综合图 8 来看, DynKAFusion 依托 KA 定理构建双层嵌套函数结构, 在特征融合与类别建模方面具备更强的表达能力与结构优势, 能够实现更精确的类别区分与空间定位, 显著提升了特征融合效果.

4.6 对比实验定量分析

为全面验证 DynKANet 模型的先进性, 将其与 16 种主流语义分割网络进行性能对比与分析, 涵盖 8 种经典模型、5 种近期提出的模型, 以及 3 种基于 KA 表示定理的语义分割算法. 实验涉及五大应用领域的 10 个数据集, 对比结果如表 4~8 所示. 其中, 加粗部分表示在对应指标下表现最佳的模型结果.

由表 4~8 可见, DynKANet 在涵盖五大领域的语义分割任务中, 绝大多数评价指标均实现有效提升, 充分体现出良好的跨领域泛化能力与适应性. 此外, DynKANet 在模型轻量化方面亦表现出色, 在大幅压缩参数规模的同时, 仍保持领先的分割性能, 展现出优异的性能与复杂度平衡. 这一优势得益于其多项结构设计优化: RSUD 模块引入深度可分离卷积替代标准卷积, SSFE 模块构建由低至高分辨率的单向传递路径, DynKAF 模块则基于瓶颈式 KAN 卷积, 构建双层函数嵌套的递进式非线性表达结构. 综合上述设计, DynKANet 参数量仅 2.26 M, FLOPs 仅为 1.60 G, 兼具计算效率与表达精度.

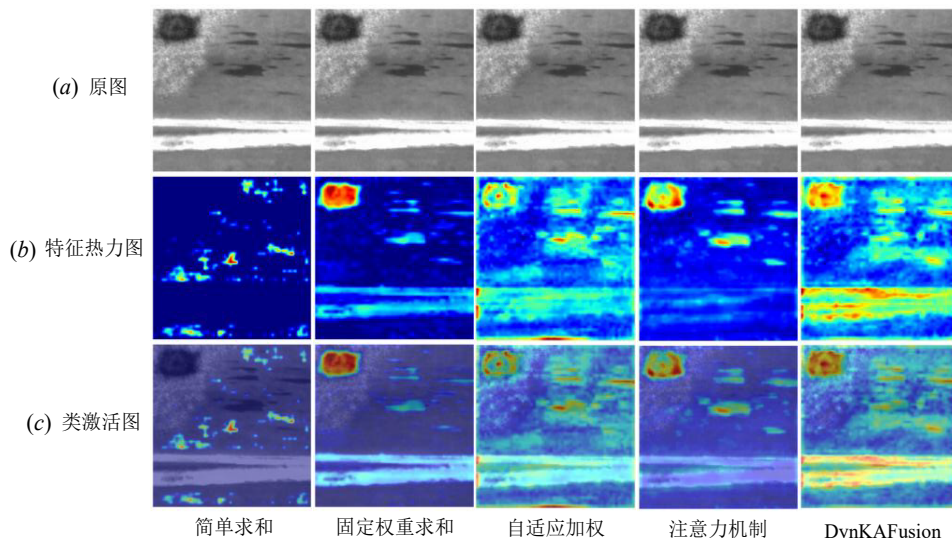


图 8 NEU_Seg 数据集下各特征融合模块原图、特征热力图和类激活图

表 4 NEU_Seg、FSSD-12 下各模型对比实验结果

Model	FLOPs/G	Params/M	NEU_Seg					FSSD-12				
			mIoU	F1	Rec	Pre	Acc	mIoU	F1	Rec	Pre	Acc
FCN ^[66] (CVPR 2015)	34.72	32.95	0.844	0.914	0.920	0.909	0.982	0.810	0.895	0.888	0.903	0.970
UNet ^[7] (MICCAI 2015)	40.19	7.76	0.823	0.902	0.895	0.910	0.906	0.796	0.886	0.850	0.926	0.850
DeepLabV3 ^[10] (TPAMI 2017)	20.76	54.60	0.850	0.918	0.920	0.916	0.983	0.804	0.891	0.874	0.909	0.969
SegNet ^[67] (TPAMI 2017)	40.17	29.45	0.822	0.901	0.907	0.896	0.979	0.806	0.893	0.873	0.914	0.970
ACC-UNet ^[9] (MICCAI 2017)	4.11	1.10	0.840	0.912	0.913	0.912	0.981	0.774	0.873	0.848	0.899	0.965
PSPNet ^[11] (CVPR 2017)	50.54	53.32	0.853	0.920	0.921	0.919	0.984	0.821	0.902	0.894	0.909	0.972
Fast-SCNN ^[16] (BMVC 2019)	0.22	1.14	0.792	0.883	0.874	0.891	0.888	0.758	0.862	0.847	0.878	0.847
DFANet ^[55] (CVPR 2019)	0.45	2.17	0.783	0.877	0.865	0.891	0.878	0.771	0.871	0.845	0.899	0.845
BiSeNetV2 ^[12] (IJCV 2021)	2.80	4.95	0.738	0.846	0.829	0.865	0.867	0.559	0.717	0.690	0.747	0.690
PIDNet-S ^[8] (CVPR 2023)	1.49	7.62	0.825	0.904	0.898	0.910	0.980	0.801	0.889	0.883	0.896	0.969
DDRNet23-S ^[14] (TITS 2023)	1.14	5.70	0.823	0.902	0.895	0.909	0.980	0.796	0.886	0.873	0.900	0.968
MEWUNet ^[68] (Mach Learn 2025)	40.52	140.27	0.840	0.913	0.910	0.915	0.981	0.814	0.898	0.882	0.914	0.971
UNetV2 ^[13] (ISBI 2025)	5.40	25.15	0.786	0.879	0.879	0.878	0.974	0.766	0.868	0.851	0.885	0.963
UKAN ^[18] (AAAI 2025)	6.89	25.36	0.843	0.914	0.912	0.916	0.923	0.717	0.824	0.782	0.875	0.925
KAT ^[19] (ICLR 2025)	20.13	31.34	0.756	0.859	0.863	0.858	0.864	0.720	0.837	0.824	0.851	0.824
Path-SAM2 ^[20] (arxiv 2024)	371.99	94.60	0.609	0.741	0.705	0.787	0.758	0.518	0.683	0.632	0.742	0.632
Ours(DynKANet)	1.60	2.26	0.862	0.938	0.932	0.930	0.989	0.833	0.913	0.899	0.919	0.986

表 5 Synapse、ACDC 下各模型对比实验结果

Model	FLOPs/G	Params/M	Synapse					ACDC				
			mIoU	F1	Rec	Pre	Acc	mIoU	F1	Rec	Pre	Acc
FCN ^[66] (CVPR 2015)	34.72	32.95	0.825	0.901	0.901	0.901	0.996	0.854	0.920	0.912	0.928	0.995
UNet ^[7] (MICCAI 2015)	40.19	7.76	0.782	0.870	0.862	0.880	0.941	0.844	0.914	0.906	0.923	0.009
DeepLabV3 ^[10] (TPAMI 2017)	20.76	54.60	0.833	0.907	0.891	0.925	0.997	0.857	0.922	0.921	0.922	0.995
SegNet ^[67] (TPAMI 2017)	40.17	29.45	0.728	0.825	0.782	0.889	0.996	0.838	0.910	0.907	0.913	0.994
ACC-UNet ^[9] (MICCAI 2017)	4.11	1.10	0.823	0.899	0.905	0.893	0.997	0.741	0.847	0.841	0.855	0.990
PSPNet ^[11] (CVPR 2017)	50.54	53.32	0.831	0.903	0.907	0.899	0.997	0.864	0.926	0.936	0.917	0.995
Fast-SCNN ^[16] (BMVC 2019)	0.22	1.14	0.719	0.836	0.813	0.861	0.813	0.804	0.890	0.879	0.900	0.884
DFANet ^[55] (CVPR 2019)	0.45	2.17	0.697	0.821	0.780	0.867	0.780	0.698	0.812	0.737	0.911	0.752
BiSeNetV2 ^[12] (IJCV 2021)	2.80	4.95	0.682	0.811	0.802	0.820	0.802	0.718	0.833	0.833	0.835	0.836
PIDNet-S ^[8] (CVPR 2023)	1.49	7.62	0.802	0.885	0.894	0.878	0.996	0.842	0.913	0.915	0.911	0.994
DDRNet23-S ^[14] (TITS 2023)	1.14	5.70	0.796	0.881	0.875	0.887	0.996	0.835	0.909	0.902	0.915	0.994
MEWUNet ^[68] (Mach Learn 2025)	40.52	140.27	0.817	0.894	0.876	0.913	0.997	0.816	0.897	0.904	0.890	0.993
UNetV2 ^[13] (ISBI 2025)	5.40	25.15	0.682	0.796	0.778	0.817	0.994	0.490	0.648	0.609	0.693	0.977
UKAN ^[18] (AAAI 2025)	6.89	25.36	0.781	0.877	0.842	0.915	0.968	0.868	0.928	0.928	0.928	0.995
KAT ^[19] (ICLR 2025)	20.13	31.34	0.639	0.780	0.748	0.815	0.748	0.702	0.820	0.835	0.806	0.845
Path-SAM2 ^[20] (arxiv 2024)	371.99	94.60	0.576	0.731	0.668	0.807	0.668	0.546	0.696	0.641	0.770	0.651
Ours(DynKANet)	1.60	2.26	0.845	0.914	0.888	0.929	0.998	0.876	0.934	0.939	0.937	0.997

在 5 类不同领域分割任务中, DynKANet 在医学影像与城市道路场景 2 类领域的性能提升最为显著. 具体而言, 在医学影像任务中, DynKANet 在息肉分割 (Synapse 数据集) 中 mIoU 为 0.845, 较其他模型平均提升 9.4%, 在 F1 与 Precision 上亦取得明显优势, 展现出对组织边界的精准刻画能力; 在心脏分割 (ACDC 数据集) 中, mIoU 平均提升 10.6%, 以 0.876 超越 PSPNet, 充

分验证其在复杂解剖结构建模中的适应性. 在城市道路场景下, DynKANet 在 KITTI 道路检测中 mIoU 平均提升 6.3%, 表现出优异的上下文建模能力. 尽管整体精度相较 PSPNet 略低, 但在边缘结构区域的预测更为稳定. 在 TP 盲道检测任务中, mIoU 平均提升高达 12.1%, 显著优于对比模型, 体现其在特殊道路结构场景中的出色语义分割能力. 但与 PSPNet 相比, 性能仍略低

表 6 KITTI, TP 下各模型对比实验结果

Model	FLOPs/G	Params/M	KITTI					TP				
			mIoU	F1	Rec	Pre	Acc	mIoU	F1	Rec	Pre	Acc
FCN ^[66] (CVPR 2015)	34.72	32.95	0.880	0.936	0.939	0.933	0.939	0.782	0.877	0.852	0.905	0.852
UNet ^[7] (MICCAI 2015)	40.19	7.76	0.817	0.899	0.907	0.892	0.907	0.851	0.919	0.922	0.916	0.922
DeepLabV3 ^[10] (TPAMI 2017)	20.76	54.60	0.879	0.936	0.947	0.925	0.947	0.830	0.907	0.875	0.941	0.875
SegNet ^[67] (TPAMI 2017)	40.17	29.45	0.841	0.914	0.901	0.926	0.901	0.603	0.752	0.655	0.883	0.655
ACC-UNet ^[9] (MICCAI 2017)	4.11	1.10	0.860	0.925	0.924	0.925	0.924	0.860	0.925	0.913	0.937	0.913
PSPNet ^[11] (CVPR 2017)	50.54	53.32	0.897	0.946	0.944	0.947	0.944	0.905	0.950	0.942	0.958	0.942
Fast-SCNN ^[16] (BMVC 2019)	0.22	1.14	0.819	0.901	0.922	0.881	0.922	0.735	0.847	0.799	0.902	0.799
DFANet ^[55] (CVPR 2019)	0.45	2.17	0.763	0.866	0.874	0.858	0.874	0.691	0.817	0.749	0.899	0.749
BiSeNetV2 ^[12] (IJCV 2021)	2.80	4.95	0.794	0.885	0.907	0.865	0.907	0.568	0.724	0.665	0.794	0.665
PIDNet-S ^[8] (CVPR 2023)	1.49	7.62	0.843	0.915	0.898	0.933	0.898	0.798	0.888	0.871	0.905	0.871
DDRNet23-S ^[14] (TITS 2023)	1.14	5.70	0.855	0.922	0.925	0.919	0.925	0.805	0.892	0.863	0.923	0.863
MEWUNet ^[68] (Mach Learn 2025)	40.52	140.27	0.855	0.922	0.943	0.902	0.943	0.810	0.895	0.902	0.888	0.902
UNetV2 ^[13] (ISBI 2025)	5.40	25.15	0.769	0.870	0.874	0.866	0.874	0.686	0.814	0.765	0.870	0.765
UKAN ^[17] (AAAI 2025)	6.89	25.36	0.848	0.918	0.927	0.909	0.927	0.875	0.933	0.921	0.946	0.921
KAT ^[18] (ICLR 2025)	20.13	31.34	0.747	0.855	0.888	0.825	0.888	0.595	0.746	0.753	0.739	0.753
Path-SAM2 ^[19] (arxiv 2024)	371.99	94.60	0.743	0.853	0.860	0.846	0.860	0.657	0.793	0.715	0.890	0.715
Ours(DynKANet)	1.60	2.26	0.889	0.949	0.948	0.948	0.947	0.874	0.955	0.945	0.963	0.947

表 7 人工月球景观、LuSNAR 下各模型对比实验结果

Model	FLOPs/G	Params/M	人工月球景观					LuSNAR				
			mIoU	F1	Rec	Pre	Acc	mIoU	F1	Rec	Pre	Acc
FCN ^[66] (CVPR 2015)	34.72	32.95	0.670	0.786	0.737	0.852	0.911	0.937	0.967	0.962	0.971	0.988
UNet ^[7] (MICCAI 2015)	40.19	7.76	0.710	0.819	0.776	0.872	0.922	0.926	0.961	0.955	0.966	0.986
DeepLabV3 ^[10] (TPAMI 2017)	20.76	54.60	0.685	0.798	0.748	0.863	0.914	0.937	0.967	0.965	0.969	0.989
SegNet ^[67] (TPAMI 2017)	40.17	29.45	0.630	0.751	0.698	0.822	0.964	0.869	0.927	0.917	0.937	0.988
ACC-UNet ^[9] (MICCAI 2017)	4.11	1.10	0.698	0.808	0.766	0.860	0.971	0.951	0.974	0.974	0.974	0.995
PSPNet ^[11] (CVPR 2017)	50.54	53.32	0.706	0.815	0.766	0.877	0.921	0.942	0.970	0.971	0.968	0.990
Fast-SCNN ^[16] (BMVC 2019)	0.22	1.14	0.609	0.731	0.657	0.850	0.888	0.891	0.948	0.936	0.946	0.978
DFANet ^[55] (CVPR 2019)	0.45	2.17	0.611	0.733	0.680	0.805	0.894	0.857	0.920	0.905	0.938	0.963
BiSeNetV2 ^[12] (IJCV 2021)	2.80	4.95	0.587	0.710	0.650	0.799	0.886	0.827	0.900	0.890	0.911	0.956
PIDNet-S ^[8] (CVPR 2023)	1.49	7.62	0.674	0.789	0.743	0.848	0.914	0.918	0.956	0.948	0.965	0.983
DDRNet23-S ^[14] (TITS 2023)	1.14	5.70	0.669	0.785	0.735	0.850	0.910	0.916	0.955	0.955	0.956	0.984
MEWUNet ^[68] (Mach Learn 2025)	40.52	140.27	0.696	0.807	0.762	0.863	0.970	0.930	0.963	0.955	0.971	0.994
UNetV2 ^[13] (ISBI 2025)	5.40	25.15	0.596	0.719	0.667	0.793	0.960	0.876	0.932	0.923	0.941	0.989
UKAN ^[17] (AAAI 2025)	6.89	25.36	0.717	0.824	0.782	0.875	0.925	0.941	0.969	0.969	0.970	0.990
KAT ^[18] (ICLR 2025)	20.13	31.34	0.579	0.701	0.650	0.772	0.879	0.928	0.962	0.960	0.964	0.985
Path-SAM2 ^[19] (arxiv 2024)	371.99	94.60	0.625	0.747	0.696	0.815	0.896	0.865	0.924	0.906	0.947	0.977
Ours(DynKANet)	1.60	2.26	0.736	0.839	0.813	0.875	0.981	0.959	0.982	0.981	0.979	0.998

3.1%, 进一步分析推测, 该差异源于 PSPNet 通过金字塔池化增强全局建模, 提升道路区域形状与宽度识别精度. 对此, 后续工作将围绕多尺度特征建模优化和场景针对性数据增强策略展开, 进一步提升模型在复杂场景下的适应性和精度.

4.7 对比实验定性分析

为直观验证本文算法的有效性, 将 DynKANet 与当前主流的 16 种语义分割方法在 10 个基准数据集上进行对比定性分析, 以全面考察 DynKANet 在不同应用场景下的表现, 结果如图 9 和图 10 所示.

表 8 2016 Sugar Beets、PlantSegV2 下各模型对比实验结果

Model	FLOPs/G	Params/M	2016 Sugar Beets					PlantSegV2				
			mIoU	F1	Rec	Pre	Acc	mIoU	F1	Rec	Pre	Acc
FCN ^[66] (CVPR 2015)	34.72	32.95	0.760	0.864	0.878	0.850	0.878	0.583	0.737	0.729	0.745	0.889
UNet ^[7] (MICCAI 2015)	40.19	7.76	0.870	0.930	0.930	0.931	0.930	0.532	0.695	0.687	0.703	0.687
DeepLabV3 ^[10] (TPAMI 2017)	20.76	54.60	0.760	0.863	0.847	0.880	0.847	0.574	0.730	0.727	0.732	0.885
SegNet ^[67] (TPAMI 2017)	40.17	29.45	0.759	0.863	0.860	0.866	0.977	0.552	0.711	0.686	0.739	0.881
ACC-UNet ^[9] (MICCAI 2017)	4.11	1.10	0.852	0.920	0.929	0.912	0.986	0.582	0.735	0.718	0.753	0.890
PSPNet ^[11] (CVPR 2017)	50.54	53.32	0.801	0.890	0.870	0.910	0.870	0.614	0.761	0.762	0.760	0.898
Fast-SCNN ^[16] (BMVC 2019)	0.22	1.14	0.750	0.857	0.857	0.871	0.843	0.539	0.700	0.688	0.713	0.875
DFANet ^[55] (CVPR 2019)	0.45	2.17	0.804	0.891	0.891	0.878	0.905	0.549	0.709	0.720	0.698	0.874
BiSeNetV2 ^[12] (IJCV 2021)	2.80	4.95	0.638	0.779	0.779	0.794	0.764	0.521	0.685	0.646	0.729	0.873
PIDNet-S ^[8] (CVPR 2023)	1.49	7.62	0.763	0.865	0.860	0.871	0.860	0.555	0.714	0.722	0.706	0.877
DDRNet23-S ^[14] (TITS 2023)	1.14	5.70	0.760	0.864	0.848	0.880	0.848	0.534	0.697	0.699	0.694	0.870
MEWUNet ^[68] (Mach Learn 2025)	40.52	140.27	0.843	0.915	0.905	0.925	0.986	0.596	0.747	0.750	0.743	0.892
UNetV2 ^[13] (ISBI 2025)	5.40	25.15	0.815	0.898	0.890	0.906	0.983	0.558	0.717	0.707	0.727	0.881
UKAN ^[18] (AAAI 2025)	6.89	25.36	0.874	0.932	0.924	0.942	0.924	0.601	0.751	0.735	0.767	0.896
KAT ^[19] (ICLR 2025)	20.13	31.34	0.702	0.825	0.847	0.804	0.847	0.508	0.628	0.671	0.591	0.834
Path-SAM2 ^[20] (arxiv 2024)	371.99	94.60	0.796	0.886	0.856	0.918	0.856	0.523	0.642	0.642	0.642	0.847
Ours(DynKANet)	1.60	2.26	0.882	0.935	0.936	0.949	0.991	0.622	0.771	0.787	0.744	0.917

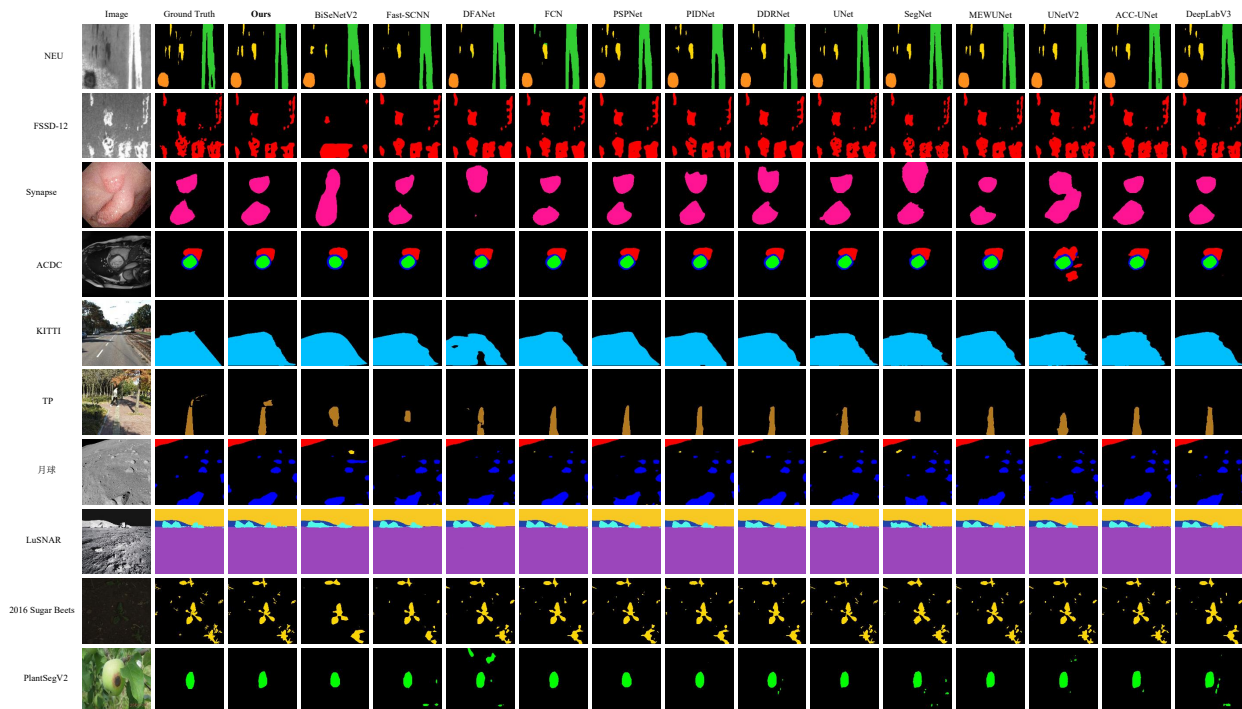


图 9 DynKANet 与 13 种语义分割方法在 10 个数据集上分割结果对比

在 NEU 和 FSSD-12 数据集上, DynKANet 在微小划痕、凹坑等细粒度缺陷的边界勾勒方面表现更为精准, 有效降低误检率与漏检率. 相较于 Path-SAM2 等模型在细节区域响应不足的表现, DynKANet 展示出更高的缺陷识别敏感性与边界表达能力, 进一步体现其在特征判别性方面的结构优势. 在 ACDC 和 Synapse 医学图

像数据集中, DynKANet 准确分割复杂医学组织边界, 有效缓解 KAT 等模型存在的边界粘连问题. 该实验现象表明, 基于可学习样条插值核卷积的双层嵌套结构能够有效促进特征融合, 增强模型在复杂医学场景下的分割能力; 在 TP 数据集中, DynKANet 相较 KAT 等方法, 在盲道等形态细长且纹理复杂的区域表现出更优

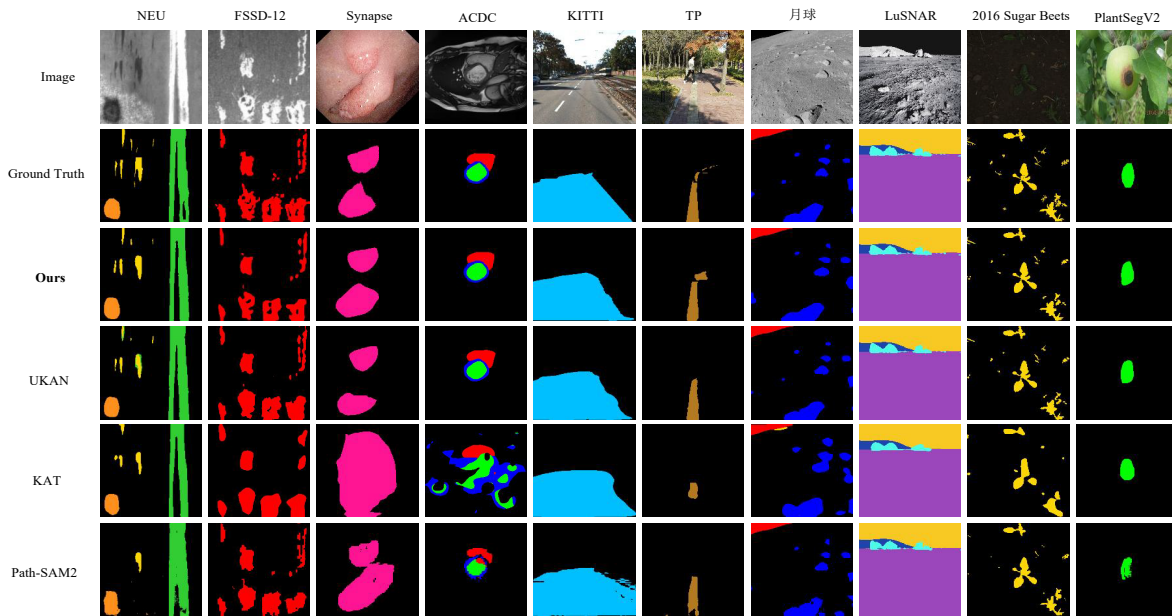


图 10 DynKANet 与 3 种基于 KA 表示定理的语义分割算法在 10 个数据集上分割结果对比

的结构感知与边界保持能力,有效抑制误检和漏分现象. 该优异表现得益于双层函数嵌套式融合结构对全局语义与局部细节的高效协同建模,显著提升了模型在非结构化场景中的精分割能力.

综合实验结果可见, DynKANet 在多个语义分割任务中均展现出较优性能,尤其在边界保持、小目标识别以及复杂背景适应性方面表现突出,进一步验证其在跨领域任务中的良好泛化能力. 该优异表现主要归功于 DynKANet 在特征提取阶段引入双分支结构,高效提

取全局与局部特征;并在特征融合阶段采用基于 KA 定理的动态融合机制,有效缓解双路融合平衡难题,显著提升分割性能.

4.8 消融实验定量分析

为深入评估 DynKANet 各模块对性能的影响,本文在统一实验设置下,选取 6 个具有代表性的数据集开展消融实验. 通过在基线模型上逐步引入 4 个核心模块,系统分析各模块的独立贡献及其协同效果. 实验结果如表 9 所示.

表 9 各模块在 6 个数据集上的参数量和 mIoU 指标对比

RSUD	SSFE	DynKAF	CE/TD	Params/M	FSSD-12	ACDC	KITTI	LuSNAR	PlantSegV2	Synapse
					mIoU	mIoU	mIoU	mIoU	mIoU	mIoU
×	×	×	×	1.50	0.781 3	0.787 6	0.861 5	0.927 6	0.559 2	0.727 6
√	×	×	×	1.58	0.803 4	0.833 6	0.871 2	0.941 9	0.583 3	0.774 4
×	√	×	×	1.50	0.789 2	0.798 3	0.863 8	0.931 6	0.559 5	0.749 0
×	×	√	×	2.13	0.808 4	0.807 8	0.872 9	0.952 6	0.578 6	0.755 4
√	√	×	×	1.58	0.804 5	0.835 4	0.872 5	0.944 0	0.585 1	0.779 9
×	√	√	×	2.15	0.801 7	0.815 2	0.879 1	0.948 9	0.569 4	0.769 1
√	×	√	×	2.21	0.816 8	0.835 2	0.879 6	0.949 5	0.593 4	0.782 1
√	√	√	×	2.26	0.824 0	0.853 8	0.885 2	0.949 8	0.599 4	0.803 4
√	√	√	√	2.26	0.832 6 ↑	0.875 5 ↑	0.889 4 ↑	0.959 2 ↑	0.617 8 ↑	0.842 0 ↑

注: RSUD 为 RSUDownsampling 模块, DynKAF 为 DynKAFusion 模块, CE/TD 为 CE/TopK+Dice($r=0.8, K=0.2$)损失函数模块.

由表 9 可得,单独加入 RSUDownsampling、SSFE、DynKAFusion 和 CE/TopK+Dice 损失函数模块均能提升基线模型性能. 其中, RSUD 模块带来的增益最为明显, mIoU 在各数据集上平均提升 2.72%, 在 PlantSegV2、ACDC 和 Synapse 数据集上分别提升 2.41%、4.60% 和

4.68%,表明 RSUD 模块通过引入递归结构,增强下采样路径的跨层语义聚合能力,具备较强的全局建模能力和空间语义连续性保持能力. 但在 KITTI 数据集上提升有限,仅提升 0.97%,说明其在尺度统一且语义边界清晰的场景下仍存在泛化瓶颈. SSFE 模块同样带来较

为明显的增益,其引入差异引导的特征增强机制,有效提升关键细粒度区域的语义感知能力,mIoU 平均提升 0.78%。在 LuSNAR、ACDC 和 Synapse 数据集上,mIoU 分别提高 0.40%、1.07% 和 2.14%,进一步验证 SSFE 模块在增强局部细节特征表达、提升复杂场景下精细分割能力方面的有效性。

DynKAF 模块同样带来较大性能提升,mIoU 平均提升 2.18%,在 FSSD-12 和 Synapse 数据集分别提升 2.71% 和 2.78%,充分验证其在钢材缺陷检测与医疗影像任务中的特征融合优势。该性能优势得益于,DynKAF 模块采用 KA 表示定理启发的双层嵌套函数结构,并结合可学习核函数与动态融合策略,有效增强全局与局部特征间的非线性交互与融合能力。进一步引入 CE/TopK+Dice 损失函数模块,mIoU 平均提升 1.68%,在 ACDC 和 Synapse 数据集分别提升 2.17% 和 3.86%。这一结果充分验证,通过两阶段优化策略,CE/TD 模块能够显著增强模型对困难样本的学习与表征能力,进而提升其在医疗影像等复杂场景中的判别性能与鲁棒性。

进一步组合各模块进行消融实验,表 9 结果表明,将 RSUD 和 SSFE 模块进行组合,相较单独使用,mIoU 平均提升 0.23% 和 2.16%,说明二者在特征层级上形成有效互补:RSUD 侧重构建宏观语义结构,SSFE 强化局部细节响应,两者协同显著增强整体表征能力。联合使用 SSFE 和 DynKAF 模块,相较单独使用,mIoU 平均提升 1.53% 和 0.13%,说明细粒度特征增强能够为动态融合机制提供更优基础特征,促进跨层信息的高效集成。将 RSUD 和 DynKAF 模块进行组合,相较单独使用,mIoU 分别平均提升 0.81%、1.34%,表明上下文增强可为多尺度融合过程为多尺度融合提供语义连续的特征支撑,进而提升融合过程中的表达一致性。当三模块联合使用时,相较单独使用,分别提升 1.79%、3.73% 和 2.20%,进一步说明 RSUD、SSFE 和 DynKAF 在全局建模、局部增强与跨层融合方面具备显著协同效应。当同时引入 4 个模块,模型性能达到最优,相较于基线模型,mIoU 平均提升 6.19%,说明 CE/TopK+Dice 损失函数能够显著提升模型对困难样本的特征学习能力,实现性能全面提升。综合上述实验结果可见,RSUD+SSFE+DynKAF+CE/TD 组合在各数据集上均取得最佳性能,说明其能够稳定提升模型性能,适合作为最终方案。

4.9 消融实验定性分析

为更直观地展示本文方法中各模块的贡献,依次移除 DynKANet 中不同模块,在所选 6 个数据集上进行可视化对比,实验结果如图 11 所示。

由图 11 可见,移除 DynKANet 中不同模块后,模型在各数据集上分割性能普遍下降,尤其在小目标识别、

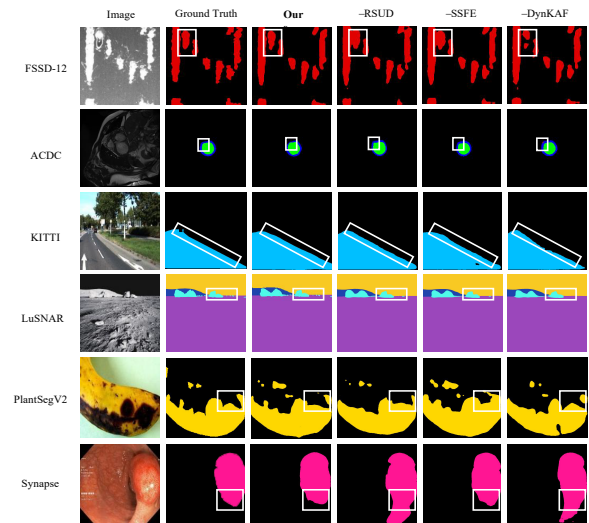


图 11 各模块对比的可视化分割结果

边界保持和局部错误处理方面退化显著。具体而言,去除 RSUD 模块后,模型在 FSSD-12 与 LuSNAR 数据集上分割精度和边界感知能力明显下降,具体表现为远距离目标轮廓模糊、复杂背景处理能力减弱;而在 Synapse 等数据集中,则表现为多尺度建模能力受限,导致低对比度和物体重叠区域中的误分与漏分现象增多。该实验现象表明,RSUD 模块通过多层次下采样捕获多尺度特征,能够显著增强模型的全局语义与局部细节表达。

进一步移除 SSFE 模块,模型在 FSSD-12 和 LuSNAR 数据集上目标与背景边界识别能力下降,细节感知减弱;而在其他数据集中,则表现为小目标分割性能下降、误分与漏分现象增多。该实验现象表明,SSFE 模块通过差异图驱动的特征增强策略,有效提升局部细节表征能力,进而显著增强目标与背景之间的判别性。去除 DynKAF 模块后,模型在 FSSD-12 与 Synapse 数据集上出现局部误分与漏分,其余数据集则表现为目标边缘模糊、轮廓退化,凸显该模块在全局与局部特征融合中的关键作用。该现象充分验证,DynKAF 模块受 KA 表示定理启发,构建双层函数动态融合机制,能够显著增强全局与局部特征的融合效果。

4.10 可解释性分析

为进一步验证 DynKANet 各模块有效性,以 LuSNAR 数据集为例,在消融实验的基础上引入热力图与类激活图,直观展示各模块在训练过程所学习的关键特征信息及其分布情况。

特征热力图和类激活图依赖于中间特征与分类输出之间的梯度或权重关联,本质上反映模型在特征空间的注意力分布,因而中间特征表达的优化将直接影响激活图的聚焦性与语义连贯性。基于此,图 12 直观呈现引入 RSUD、SSFE 与 DynKAF 模块前后模型响应图

的可视化对比结果. 由图可见, 引入 3 个模块后, 特征热力图的激活区域更为集中显著, 表明模型对判别性特征的关注能力得到显著提升. 相比之下, 基线模型的激活区域较为分散, 侧重捕捉整体轮廓特征, 缺乏对关键区域的精细建模. 类激活图进一步验证该趋势: 基线模型中激活区域分散且存在误关注现象, 而引入 3 个模块后, 模型激活区域与真实目标高度重叠, 模型识别能力显著增强.

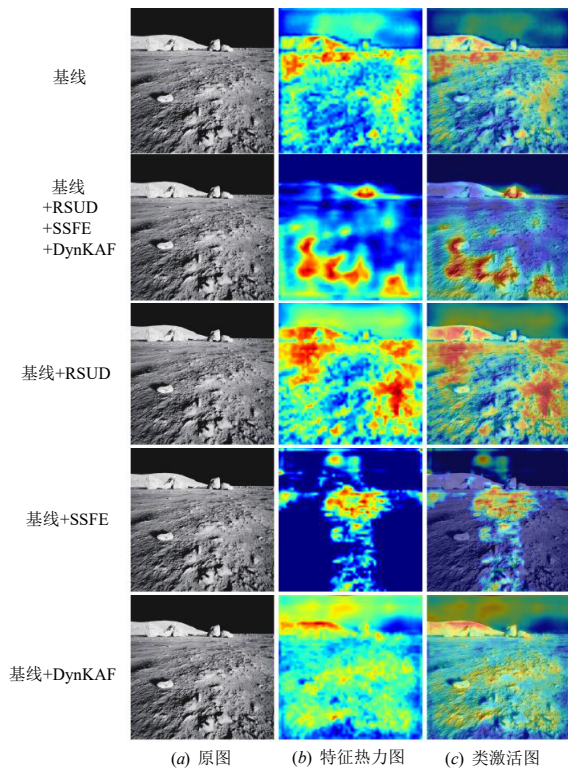


图 12 LuSNAR 数据集下各模块原图、特征热力图和类激活图

进一步从模型机制层面加以剖析, RSUD 模块通过残差连接构建 U 型上下文路径, 有效提升模型对长距离依赖的建模能力, 从而在特征热力图中呈现出覆盖范围更广、结构更清晰的响应模式; SSFE 模块借助语义差异图显式刻画多尺度语义间的区域差异, 引导模型强化对边缘及局部关键区域的响应, 使得类激活图中的边界识别与目标细节更为清晰; DynKAF 模块则在前序模块基础上, 引入受 KA 表示定理启发的双层嵌套函数融合机制, 实现多粒度特征的动态融合. 具体而言, 其通过标准卷积和可学习样条插值核构建双层嵌套函数结构, 并结合动态加权融合策略, 有效建模全局与局部特征间的非线性关联, 显著提升特征融合的代表性与适应性. 得益于上述设计, DynKAF 模块在可视化结果中呈现出更精确且语义对齐性更强的激活响应图, 进一步验证 KA 表示定理在多粒度特征融合中的适

用性与有效性. 整体来看, 当 3 个模块协同作用时, 模型在目标分割与边界识别方面取得显著提升. 该互效应进一步验证, DynKANet 通过精准提取与动态融合不同粒度特征, 能够精准捕捉并整合图像的全局与局部信息, 显著提升语义分割性能.

5 结束语

针对现有语义分割方法在复杂场景下全局语义与局部细节特征融合时难以平衡的问题, 本文提出 KA 表示定理启发的图像语义分割动态融合网络 DynKANet, 通过设计多级特征提取与动态特征融合机制, 有效平衡全局语义与局部细节的协同表达, 实现模型分割精度的有效提升. 此外, 设计 CE/Top-K+Dice 动态联合损失函数, 并引入条件触发策略, 增强模型对复杂场景的适应能力. DynKAFusion 模块实验表明, DynKAFusion 相较其他特征融合策略表现出更优性能, 具备良好的参数-性能平衡; 对比实验表明, DynKANet 在涉及 5 个领域的 10 个基准数据集上分割性能优于主流方法; 消融实验进一步验证各模块对分割性能贡献; 可解释性分析通过可视化揭示 KA 表示定理驱动的动态特征融合策略的合理性. 总体而言, DynKANet 在特征提取与融合方面取得较大进展, 能够实现全局语义与局部细节的平衡互补, 在分割精度和泛化能力上均表现出优异性能. 然而, 在复杂动态场景下, 高分辨率任务仍面临较高计算开销, 模型效率仍有提升空间. 未来将聚焦于 DynKANet 轻量化设计, 并探索自适应学习策略, 提升模型在动态环境中的适应性与鲁棒性.

参考文献

- [1] 秦飞巍, 沈希乐, 彭勇, 等. 无人驾驶中的场景实时语义分割方法[J]. 计算机辅助设计与图形学学报, 2021, 33(7): 1026-1037.
- [2] QIN F W, SHEN X Y, PENG Y, et al. A real-time semantic segmentation approach for autonomous driving scenes[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(7): 1026-1037. (in Chinese)
- [3] ASGARI TAGHANAKI S, ABHISHEK K, COHEN J P, et al. Deep semantic segmentation of natural and medical images: A review[J]. Artificial Intelligence Review, 2021, 54(1): 137-178.
- [4] 董思俊, 孟小亮. 基于文本语义驱动的遥感影像要素提取[J]. 航天返回与遥感, 2024, 45(3): 82-91.
- [5] DONG S J, MENG X L. Text-semantics-driven feature extraction from remote sensing imagery[J]. Spacecraft Recovery & Remote Sensing, 2024, 45(3): 82-91. (in Chinese)
- [6] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes

- dataset for semantic urban scene understanding[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 3213-3223.
- [5] 郝铁文. 基于DeepLabv3+的轻量化铝带表面缺陷分割方法[J]. 信息技术与信息化, 2024(12): 83-87.
HAO T W. Surface defect segmentation method of lightweight aluminum strip based on DeepLabv3+[J]. Information Technology and Informatization, 2024(12): 83-87. (in Chinese)
- [6] 高常鑫, 徐正泽, 吴东岳, 等. 深度学习实时语义分割综述[J]. 中国图象图形学报, 2024, 29(5): 1119-1145.
GAO C X, XU Z Z, WU D Y, et al. Deep learning-based real-time semantic segmentation: A survey[J]. Journal of Image and Graphics, 2024, 29(5): 1119-1145. (in Chinese)
- [7] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Cham: Springer, 2015: 234-241.
- [8] XU J C, XIONG Z X, BHATTACHARYYA S P. PIDNet: A real-time semantic segmentation network inspired by PID controllers[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 19529-19539.
- [9] IBTEHAZ N, KIHARA D. ACC-UNet: A completely convolutional UNet model for the 2020s[M]//Medical Image Computing and Computer Assisted Intervention - MICCAI 2023. Cham: Springer Nature Switzerland, 2023: 692-702.
- [10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [11] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 6230-6239.
- [12] YU C Q, GAO C X, WANG J B, et al. BiSeNet V2: Bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129(11): 3051-3068.
- [13] PENG Y P, CHEN D Z, SONKA M. U-Net V2: Rethinking the skip connections of U-Net for medical image segmentation[C]//2025 IEEE 22nd International Symposium on Biomedical Imaging. Piscataway: IEEE, 2025. DOI: 10.1109/ISBI60581.2025.10980742.
- [14] PAN H H, HONG Y D, SUN W C, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(3): 3448-3460.
- [15] PENG J C, LIU Y, TANG S Y, et al. PP-LiteSeg: A superior real-time semantic segmentation model[EB/OL]. (2022-04-06)[2025-03-23]. <https://arXiv.org/abs/2204.02681>.
- [16] POUDEL R P K, LIWICKI S, CIPOLLA R. Fast-SCNN: Fast semantic segmentation network[C]//Proceedings of the British Machine Vision Conference, 2019: 60441195.
- [17] LIU Z M, WANG Y X, VAIDYA S, et al. KAN: Kolmogorov-Arnold networks[EB/OL]. (2025-02-09) [2025-03-23]. <https://arXiv.org/abs/2404.19756>.
- [18] LI C X, LIU X Y, LI W Y, et al. U-KAN makes strong backbone for medical image segmentation and generation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(5): 4652-4660.
- [19] YANG X Y, WANG X C. Kolmogorov-Arnold transformer[EB/OL]. (2024-09-16) [2025-03-23]. <https://arxiv.org/abs/2409.10594>.
- [20] ZHANG M Y, WANG L, CHEN Z H, et al. Path-SAM2: Transfer SAM2 for digital pathology semantic segmentation[EB/OL]. (2024-09-24)[2025-03-23]. <https://arXiv.org/abs/2408.03651>.
- [21] YU C Q, WANG J B, PENG C, et al. BiSeNet: Bilateral segmentation network for real-time semantic segmentation[C]//Computer Vision - ECCV 2018. Cham: Springer, 2018: 334-349.
- [22] LIU M S, DAN J, LU Z Q, et al. CM-UNet: Hybrid CNN-mamba UNet for remote sensing image semantic segmentation[EB/OL]. (2024-03-17)[2025-03-23]. <https://arXiv.org/abs/2405.10530>.
- [23] WANG Z Y, MA C. Weak-mamba-UNet: Visual mamba makes CNN and ViT work better for scribble-based medical image segmentation[EB/OL]. (2024-01-16) [2025-03-23]. <https://arXiv.org/abs/2402.10887>.
- [24] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 936-944.
- [25] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [26] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Confer-

- ence on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [27] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [28] CAO Y, XU J R, LIN S, et al. Global context networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(6): 6881-6895.
- [29] FU J, LIU J, TIAN H J, et al. Dual attention network for scene segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 3141-3149.
- [30] HU Y, CHEN Y, LI X, et al. Dynamic featurefusion for semantic edge detection[EB/OL]. (2019-02-25)[2025-03-23]. <https://arxiv.org/abs/1902.09104>.
- [31] WANG C Y, ZHONG C M. Adaptive feature pyramid networks for object detection[J]. IEEE Access, 2021, 9: 107024-107032.
- [32] QIAO D H, ZULKERNINE F. Adaptive feature fusion for cooperative perception using LiDAR point clouds[C]//2023 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2023: 1186-1195.
- [33] XIE X Y, ZHENG X B, YU Z T, et al. FusionMamba: Dynamic feature enhancement for multimodal image fusion with Mamba[J]. Visual Intelligence, 2024, 2(1): 37.
- [34] YANG J, QIU P J, ZHANG Y C, et al. D-net: Dynamic large kernel with dynamic feature fusion for volumetric medical image segmentation[EB/OL]. (2024-10-17)[2025-03-23]. <https://arXiv.org/abs/2403.10674>.
- [35] WU B Q, XIAO Q, LIU S W, et al. E2ENet: Dynamic sparse feature fusion for accurate and efficient 3D medical image segmentation[EB/OL]. (2025-02-19)[2025-03-23]. <https://arXiv.org/abs/2312.04727>.
- [36] TA H T, DUY QUY T, TRAN A N, et al. Prkan:Parameter-reduced Kolmogorov-Arnold networks[EB/OL]. (2025-02-11)[2025-03-23]. <https://arxiv.org/abs/2501.07032>.
- [37] MA X P, WANG Z Y, HU Y, et al. Kolmogorov-Arnold network for remote sensing image semantic segmentation[EB/OL]. (2025-01-13)[2025-03-23]. <https://arXiv.org/abs/2501.07390>.
- [38] YANG L H, ZHAO J X, WANG Z M, et al. M-KANUNet: Enhanced defect segmentation in X-ray images of copper pipe welds via multi-scale representation and Kolmogorov-Arnold Networks[J]. The Visual Computer, 2025, 41(10): 7201-7214.
- [39] ZHANG B H, HUANG H R, SHEN Y, et al. MM-UKAN: A novel Kolmogorov-Arnold network-based U-shaped network for ultrasound image segmentation[J]. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2025, 72(4): 498-514.
- [40] PANG S C, ZHAO X, ZHANG Y L, et al. FKAN-GMF-Net: Fourier Kolmogorov-Arnold-based group multi-scale fusion network for aneurysm image segmentation[C]//ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2025: 1-5.
- [41] LI Y H, LIU S, WU J, et al. Multi-scale Kolmogorov-Arnold network (KAN) -based linear attention network: Multi-scale feature fusion with KAN and deformable convolution for urban scene image semantic segmentation[J]. Remote Sensing, 2025, 17(5): 802.
- [42] LIN S Y, HU R, LI Z Y, et al. KAC-unet: A medical image segmentation with the adaptive group strategy and Kolmogorov-Arnold network[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 5015413.
- [43] LI C S, XU X. Semi-supervised learning with Kolmogorov-Arnold network for MRI cardiac segmentation[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 2515311.
- [44] YUAN J, ZHOU L J, HE M R, et al. A lightweight dual path Kolmogorov-Arnold convolution network for medical optical image segmentation[J]. Neurocomputing, 2026, 659: 131776.
- [45] 汤红忠, 王蔚, 王涛, 等. 一种基于课程学习的胚胎图像语义分割方法[J]. 电子学报, 2023, 51(11): 3365-3376.
- TANG H Z, WANG W, WANG T, et al. A semantic segmentation method of embryo image based on curriculum learning[J]. Acta Electronica Sinica, 2023, 51(11): 3365-3376. (in Chinese)
- [46] 梁燕, 易春霞, 王光宇, 等. 基于多尺度语义编解码网络的遥感图像语义分割[J]. 电子学报, 2023, 51(11): 3199-3214.
- LIANG Y, YI C X, WANG G Y, et al. Semantic segmentation of remote sensing image based on multi-scale semantic encoder-decoder network[J]. Acta Electronica Sinica, 2023, 51(11): 3199-3214. (in Chinese)
- [47] QIN X B, ZHANG Z C, HUANG C Y, et al. U²-Net: Going deeper with nested U-structure for salient object detection[J]. Pattern Recognition, 2020, 106: 107404.
- [48] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway:

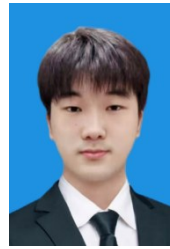
- IEEE, 2017: 1800-1807.
- [49] DING X H, GUO Y C, DING G G, et al. ACNet: Strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 1911-1920.
- [50] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2016-04-30)[2025-03-23]. <https://arXiv.org/abs/1511.07122>.
- [51] 周勇, 陈思霖, 赵佳琦, 等. 基于弱语义注意力的遥感图像可解释目标检测[J]. 电子学报, 2021, 49(4): 679-689.
- ZHOU Y, CHEN S L, ZHAO J Q, et al. Weakly semantic based attention network for interpretable object detection in remote sensing imagery[J]. Acta Electronica Sinica, 2021, 49(4): 679-689. (in Chinese)
- [52] 张志文, 刘天歌, 聂鹏举. 基于实景数据增强和双路径融合网络的实时街景语义分割算法[J]. 电子学报, 2022, 50(7): 1609-1620.
- ZHANG Z W, LIU T G, NIE P J. Real-time semantic segmentation for road scene based on data enhancement and dual-path fusion network[J]. Acta Electronica Sinica, 2022, 50(7): 1609-1620. (in Chinese)
- [53] 孟磊, 徐磊, 郭嘉阳. 一种基于改进的 MobileNetV2 网络语义分割算法[J]. 电子学报, 2020, 48(9): 1769-1776.
- MENG L, XU L, GUO J Y. Semantic segmentation algorithm based on improved MobileNetV2[J]. Acta Electronica Sinica, 2020, 48(9): 1769-1776. (in Chinese)
- [54] LI H C, XIONG P F, FAN H Q, et al. DFANet: Deep feature aggregation for real-time semantic segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 9514-9523.
- [55] BODNER A D, TEPSICH A S, SPOLSKI J N, et al. Convolutional Kolmogorov-Arnold networks[EB/OL]. (2024-06-19)[2025-03-23]. <https://arXiv.org/abs/2406.13155>.
- [56] DROKIN I. Kolmogorov-Arnold convolutions: Design principles and empirical studies[EB/OL]. (2024-06-19)[2025-03-23]. <https://arXiv.org/abs/2407.01092>.
- [57] FAN Y B, LYU S W, YING Y M, et al. Learning with average top-k loss[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 497-505.
- [58] GUO B Y, WANG Y T, ZHEN S, et al. SPEED: Semantic prior and extremely efficient dilated convolution network for real-time metal surface defects detection[J]. IEEE Transactions on Industrial Informatics, 2023, 19(12): 11380-11390.
- [59] FENG H, SONG K C, CUI W Q, et al. Cross position aggregation network for few-shot strip steel surface defect segmentation[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 5007410.
- [60] WANG H Y, XIE S, LIN L F, et al. Mixed transformer U-Net for medical image segmentation[C]//ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2022: 2390-2394.
- [61] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: The KITTI dataset[J]. International Journal of Robotics Research, 2013, 32(11): 1231-1237.
- [62] ZHANG X L, LIANG L, ZHAO S L, et al. GRFB-UNet: A new multi-scale attention network with group receptive field block for tactile paving segmentation[J]. Expert Systems with Applications, 2024, 238: 122109.
- [63] LIU J Y, ZHANG Q Y, WAN X, et al. LuSNAR: A lunar segmentation, navigation and reconstruction dataset based on Multi-sensor for autonomous exploration[EB/OL]. (2024-09-26)[2025-03-23]. <https://arXiv.org/abs/2407.06512>.
- [64] CHEBROLU N, LOTTES P, SCHAEFER A, et al. Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields[J]. International Journal of Robotics Research, 2017, 36(10): 1045-1052.
- [65] WEI T Q, CHEN Z, YU X, et al. PlantSeg: A large-scale in-the-wild dataset for plant disease segmentation[EB/OL]. (2024-09-06)[2025-03-23]. <https://arXiv.org/abs/2409.04038>.
- [66] 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.
- [67] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [68] RUAN J C, GAO J S, XIE M Y, et al. Learning multi-axis representation in frequency domain for medical image segmentation[J]. Machine Learning, 2025, 114(1): 10.

作者简介



梁 欣 女, 2005年5月出生于湖南省郴州市. 现为湖南工商大学人工智能与先进计算学院(湘江书院)本科生. 主要研究方向为语义分割.

E-mail: 3278028915@qq.com



刘志义 男, 2005年6月出生于湖南省衡阳市. 现为湖南工商大学人工智能与先进计算学院(湘江书院)本科生. 主要研究方向为语义分割.

E-mail: 3147941056@qq.com



姜 林 男, 1977年11月出生于湖南省常德市. 现为湖南工商大学人工智能与先进计算学院(湘江书院)教授. 主要研究方向为智能语音处理、机器视觉.

E-mail: jlcdf@163.com



黄依莎 女, 2004年2月出生于江西省萍乡市. 现为湖南工商大学人工智能与先进计算学院(湘江书院)本科生. 主要研究方向为语义分割.

E-mail: 2142827479@qq.com



彭 超 女, 2004年6月出生于湖南省浏阳市. 现为湖南工商大学人工智能与先进计算学院(湘江书院)本科生. 主要研究方向为语义分割.

E-mail: 166267867@qq.com



艾彦迪 男, 1982年9月出生于辽宁省沈阳市. 现为湖南工商大学前沿交叉学院副教授. 主要研究方向为计算机视觉、边缘智能.

E-mail: 444176993@qq.com