

DOI:10.19651/j.cnki.emt.2416707

## 轻量化锻件表面小目标裂纹分割算法\*

张上<sup>1,2,3</sup> 邹扬<sup>1,2,3</sup> 张岳<sup>1,2,3</sup>

(1. 水电工程智能视觉监测湖北省重点实验室 宜昌 443002; 2. 三峡大学计算机与信息学院 宜昌 443002;  
3. 三峡大学湖北省建筑质量检测装备工程技术研究中心 宜昌 443002)

**摘要:** 锻件在制造过程中容易出现裂纹等各种表面缺陷,影响产品质量。针对复杂可见光环境下小目标裂纹容易出现漏检问题,并考虑生产线上高效部署需求,提出 LSC-PoolFormer 算法。首先,采集湖北三环锻造有限公司汽车转向节生产线的磁粉探伤图像,标注后制成 FDMPI 数据集;然后,使用基于 PoolFormer 骨干网络的编码器,实现轻量级高效的特征提取;其次,引入渐近特征金字塔作为颈部网络,减少不同尺度特征之间的语义差距;最后,基于动态蛇形卷积提出 DS-Seghead 作为解码头,强化模型对条状裂纹的感知能力,并提出 DDS 训练策略,降低小目标裂纹丢失的概率。LSC-PoolFormer 在 FDMPI 上的实验结果表明,相较基准模型,该算法的参数量和计算量分别下降 9.2% 和 48.78%,F1 分数和 IoU 分别提升 1.1% 和 1.69%;同时在公开数据集 NEU-Seg 上的表现也证明了该算法的泛化能力,相较基准模型,大幅度降低参数量和计算量的情况下,mF1 分数和 mIoU 分别提升 0.66% 和 1.04%。实验证明,本文算法在保持检测精度的同时,显著降低算法复杂度,有利于实际部署。

**关键词:** 语义分割;缺陷检测;渐近特征融合;形态学

中图分类号: TP391.4;TN98 文献标识码: A 国家标准学科分类代码: 510.4050

## Lightweight segmentation algorithm for small target cracks on forging surface

Zhang Shang<sup>1,2,3</sup> Zou Yang<sup>1,2,3</sup> Zhang Yue<sup>1,2,3</sup>

(1. Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang 443002, China; 2. College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China;  
3. Hubei Province Engineering Technology Research Center for Construction Quality Testing Equipment, China Three Gorges University, Yichang 443002, China)

**Abstract:** Forgings are prone to various surface defects such as cracks during the manufacturing process, which affects product quality. Aiming at the problem that small target cracks are easily missed in complex visible light environments, and considering the requirement of efficient deployment in production line, the LSC-PoolFormer algorithm is proposed. First, the magnetic particle inspection images from the automobile steering knuckle production line of Hubei Sanhuan Forging Co., Ltd. were collected, annotated and made into a FDMPI data set; then, an encoder based on the PoolFormer backbone network was used to achieve lightweight and efficient feature extraction; secondly, the asymptotic feature pyramid is introduced as the neck network to reduce the semantic gap between features of different scales; finally, DS-Seghead is proposed as the decoding head based on dynamic snake convolution to enhance the model's perception of stripe cracks, and a DDS training strategy is proposed, reducing the probability of missing small target cracks. The experimental results of LSC-PoolFormer on FDMPI show that compared with the baseline model, the parameter amount and calculation amount of this algorithm decreased by 9.2% and 48.78% respectively, and the F1 score and IoU increased by 1.1% and 1.69% respectively. At the same time, the performance on the public data set NEU-Seg also proves the generalization ability of the algorithm. Compared with the baseline model, the mF1 score and mIoU increased by 0.66% and 1.04% respectively while greatly reducing the number of parameters and calculations. Experiment shows that the algorithm in this paper significantly reduces the complexity of the algorithm while maintaining detection accuracy, which is beneficial to the actual deployment.

**Keywords:** semantic segmentation; defect detection; asymptotic feature fusion; morphology

## 0 引言

锻件被广泛应用于船舶部件制造、汽车零件生产以及

石油化工的运输设备和存储设备。生产中,其表面无可避免会出现裂纹、孔隙、氧化层等缺陷,为保障产品质量,有必要针对锻件进行无损检测<sup>[1]</sup>,其中,磁粉检测技术通常用于

收稿日期:2024-08-21

\* 基金项目:国家级大学生创新创业训练计划(202011075013)项目资助

解决锻件的裂纹缺陷<sup>[2]</sup>。传统的物理检测主要通过人工视觉直接识别零件的裂纹,效率低下且容易造成误判,同时,重量较大的锻件易对操作人员产生较高的体能消耗,这进一步降低了检测效率,进而影响整个生产过程<sup>[3]</sup>。

近年来,伴随互联网的不断发展,深度学习因其出色的特征提取能力被广泛应用于产品缺陷检测<sup>[4]</sup>,利用其进行质量管理在整个工业制造中得到广泛认可,新的检测算法在不断被探索和创造<sup>[5]</sup>。目前,基于深度学习的表面缺陷检测方法包括图像分类<sup>[6-7]</sup>,目标检测<sup>[8-9]</sup>和语义分割<sup>[10-11]</sup>。针对磁粉探伤检测锻件表面缺陷的场景,文献[12]指出图像分类只能识别目标锻件图像中是否存在缺陷,没有实现缺陷的准确定位,这导致其应用受到限制。目标检测虽然可以实现缺陷的进一步定位,但无法显示目标缺陷的准确形状,在实际的流水线检测任务中仍然存在缺点。因此本文使用语义分割,实现像素级的定位和分类,不仅可以准确定位缺陷,同时能显示缺陷的具体形状。

实际生产中,条状裂纹作为锻件表面的一种常见缺陷,通常呈现细长且占据图片面积比例较小的特点,在检测过程中容易出现漏检,而生产线上高效部署的需求使得此类现象更加严重。文献[13]提出一种语义先验与高效空洞卷积网络(semantic prior and extremely efficient dilated convolution network, SPEED),在不增加计算复杂度的前提下,高效捕捉金属表面的纹理,提高网络对缺陷的感知能力,但未解决小目标容易丢失的难题。文献[14]提出一种基于对比学习的语义分割模型,引入缺陷记忆库来存储和检索未标记训练数据中的类别感知语义记忆,利用类别感知的语义对比驱动神经网络学习,增强困难样本的利用率,提高检测精度,但难以满足高效部署的需求。文献[15]针对磁砖表面缺陷在对比度低于背景噪声时难以识别,特别是微小缺陷容易被忽略的情况,提出高斯残差注意力卷积(gaussian residual attention convolution, GRAC),通过建模判别特征表示来区分相似特征,此外,还提出多视图循环卷积(multi-view cyclic convolution, MVCC),通过学习多视图特征图来分割背景噪声中的微小缺陷,但数据集是针对低对比度的工业环境。

针对上述问题,考虑实际生产环境的复杂光照条件和高效部署需求,本文提出了一种轻量化条状卷积模型(lightweight striped convolution-PoolFormer, LSC-PoolFormer),用于分割可见光状态下锻件表面的条状小目标裂纹缺陷,主要工作如下:

1)模型基于 Encoder-Decoder<sup>[16]</sup>结构,使用 PoolFormer<sup>[17]</sup>骨干网络作为编码器,在保持优异性能的同时,大幅度减少计算量和参数,实现高效、轻量级的特征提取过程。

2)在解码器中引入渐进特征金字塔网络(asymptotic feature pyramid network, AFPN)<sup>[18]</sup>,添加颈部网络,实现轻量化特征融合的同时,加强非相邻尺度的相互作用,减少不同尺度特征融合后产生的语义损失,确保模型在可见光

环境下仍能感知到详细的特征信息<sup>[19]</sup>。

3)引入动态蛇形卷积(dynamic snake convolution, DSConv)<sup>[20]</sup>设计 DS-Seghead 模块,进一步强化解码器对锻件表面条状裂纹缺陷的感知能力,并提出双重蛇形卷积(double dynamic snake, DDS)的训练策略,通过重构 AFPN 的上采样部分,降低多尺度特征融合过程中小目标裂纹丢失的概率。

4)收集湖北三环锻造有限公司戴姆勒奔驰重卡转向节生产线磁粉检测过程的图像,并制作了锻造缺陷磁粉检测数据集(forging defect magnetic particle inspection, FDMPI),用于模型训练和结果验证。

## 1 LSC-PoolFormer

文献[21]提出的视觉模型(vision transformer, ViT)将 TransFormer<sup>[22]</sup>引入了视觉任务,文献[23]提出的 SegFormer 算法沿用 ViT 模型的框架,着重考虑其 Decoder 结构的改进,提出全连接多层感知机,实现轻量级解码,解决了 ViT 模型计算代价高的问题。由文献[17]提出的 PoolFormer 算法则是着重对 ViT 中 Encoder 结构进行改进,提出一种没有参数的空间池化算子,替换 TransFormer 结构中的 token mixer 模块,实现轻量级编码。

本文提出轻量级条状卷积模型 LSC-PoolFormer 由两部分构成:基于 PoolFormer 的 Encoder 结构和重新设计的 Decoder 结构,其中 Decoder 由 AFPN 颈部网络和 DS-Seghead 共同组成。在 Encoder 结构实现轻量级编码的基础上,进一步降低解码计算代价,提高模型分割精度。网络结构如图 1 所示。

### 1.1 基于 PoolFormer 的轻量化 Encoder 结构

PoolFormer<sup>[17]</sup>旨在验证 TransFormer 系列算法架构的有效性,并将此类通用架构称为 MetaFormer(由 token mixer 和 channel MLP 构成)。作为 MetaFormer 模型的一种,PoolFormer 通过使用一个平均池化层 Pooling 取代复杂的 token mixer 模块,在多个视觉任务中取得了良好效果,展示了 MetaFormer 架构在特定任务中的高效性能,证明 token mixer 模块对模型整体的性能影响不大,若输入为  $I \in \mathbb{R}^{W \times H \times C}$ ,则池化操作如式(1)所示。

$$I'_{i,j,c} = \frac{1}{k \times k} \sum_{p,q=1}^k I_{i+p-\frac{k+1}{2},j+q-\frac{k+1}{2},c} - I_{i,j,c} \quad (1)$$

其中,  $k$  是池化核大小;  $p$  和  $q$  表示宽高坐标;  $C$  为通道数,由于池化后续包含残差连接,因此需减去池化前的输入。

若输入一个序列  $X \in \mathbb{R}^{N \times C}$ ,首先通过归一化层 Norm Layer 和池化操作平均聚合序列中不同切片的信息,再与处理前的序列进行残差连接得到新的序列  $Y$ ,如式(2)所示。

$$Y = Pooling(Normlize(X)) + X \quad (2)$$

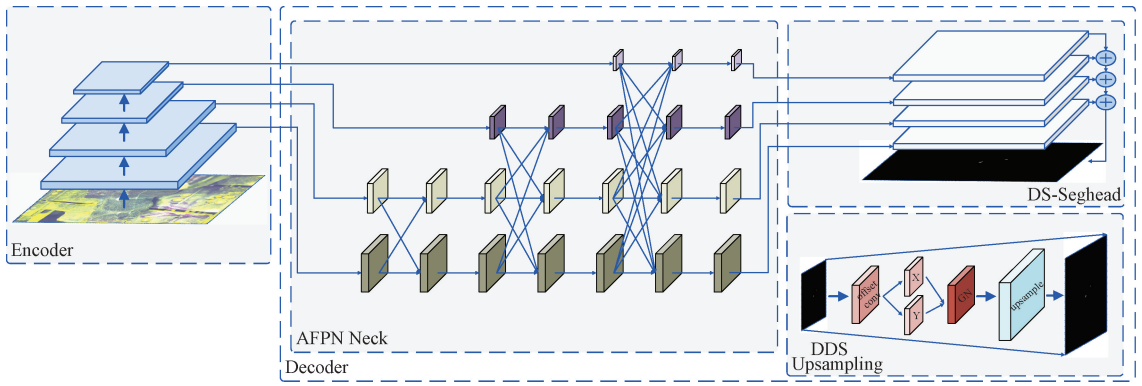


图 1 LSC-PoolFormer 算法结构

Fig.1 LSC-PoolFormer algorithm structure

其中,  $Pooling(\cdot)$  是池化操作;  $Normalize(\cdot)$  是归一化操作;  $N$  为序列长度。

最后将序列  $Y$  通过双层非线性激活层 MLP, 并与未处理的序列  $Y$  进行残差连接, 得到序列  $Z$ , 完成 PoolFormer 块的特征提取工作, 如式(3)所示。

$$Z = \sigma(Normlize(Y)W_1)W_2 + Y \quad (3)$$

其中,  $\sigma(\cdot)$  是非线性激活函数;  $W_1 \in \mathbb{R}^{C \times r}$  和  $W_2 \in \mathbb{R}^{r \times C}$  是 MLP 扩展率  $r$  的可学习参数。

PoolFormer 在计算机视觉的多个领域得到广泛应用<sup>[24-25]</sup>。在本研究中, 为了满足锻件表面缺陷检测场景中实时性和高性能的要求, 采用 PoolFormer 作为编码器, PoolFormer Encoder 结构如图 2 所示。其中图 2(b)为网络结构, 图 2(a)为 PoolFormer 块的详细架构。由于 PoolFormer 的池化层没有可学习参数, 可以有效降低网络的计算量, 此外, 得益于 MetaFormer 架构的出色设计, 可以实现有效的特征提取。

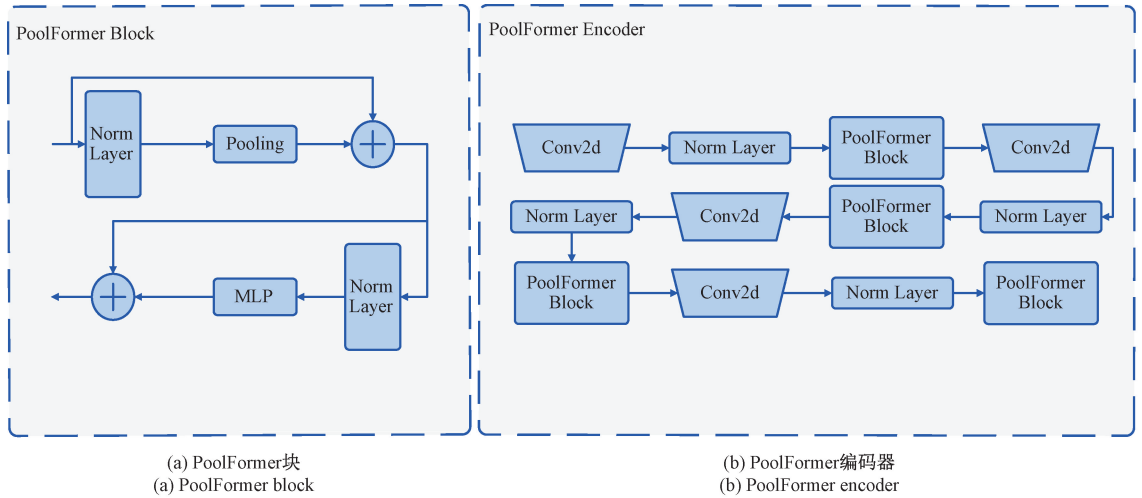


图 2 PoolFormer Encoder 结构

Fig.2 PoolFormer Encoder structure

### 1.2 轻量级颈部网络 AFPN

条状裂纹细长且占据图片面积比例较小, 这些特点使得此类缺陷很难从背景中分割出来。此类问题的典型解决方案是使用多尺度特征提取网络, 保留图像中不同尺度的语义特征信息以实现有效识别, 而由于其创新的架构和优越的性能, 文献[26]中提出的特征金字塔网络(feature pyramid networks, FPN)已被广泛应用<sup>[27]</sup>。尽管金字塔顶部的特征从上到下传播并与中间层进行特征融合, 但其中的语义信息仍可能丢失, 于是文献[28]提出了路径聚合特征金字塔网络(path aggregation feature pyramid

network, PAFPN), 增加了一条自下而上的路径, 但未能解决特征信息在传播和交互过程中可能发生的退化问题。文献[29]提出图特征金字塔网络(graph feature pyramid network, GraphFPN), 引入图神经网络(graph neural network, GNN)进行信息传播和交互, 但代价是参数量和计算量爆炸性增长。

文献[18]提出的 AFPN 通过自下而上的方法提取特征, 其网络架构如图 3 所示。对于从骨干网络提取到的不同层特征, AFPN 首先将低层特征输入到网络, 然后通过混合上下采样的方式逐级合并高层特征, 从而有效缓解非

相邻层之间的语义差距。此外 AFPN 还引入了文献[30]中提出的自适应空间融合(adaptive spatial feature fusion, ASFF)操作,当 AFPN 进行相邻层特征融合操作时,ASFF 会为不同级别的特征分配不同的权重,增强关键特征的重要性,缓解不同空间位置的多目标信息在融合过程中可能引起的冲突。AFPN 通过渐近特征融合和 ASFF 解决了

非相邻层之间的语义差距,保留各层的特征语义,缓解特征融合中的信息冲突,同时由于 AFPN 并没有引入新的神经网络,相较于其他的多尺度特征提取方法,在保持优异性能的同时,不会造成模型参数量和计算量的大幅度增长,而 ASFF 操作的存在,使得 AFPN 会自动过滤训练中的无用信息,在一些任务中可能会优化模型的网络结构。

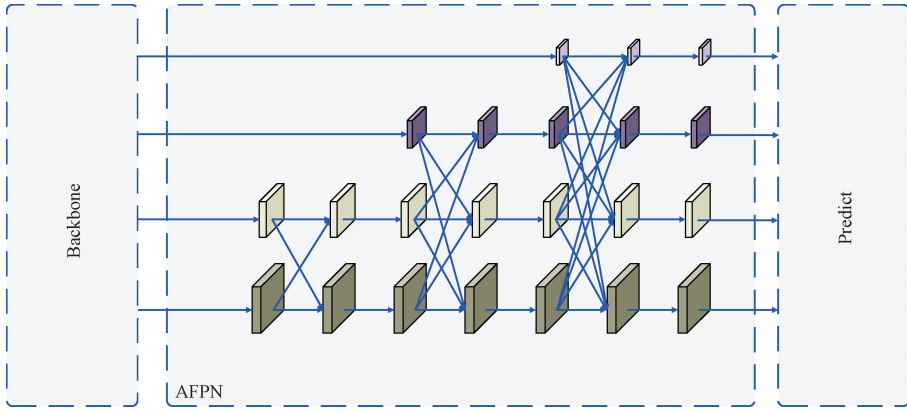


图 3 AFPN 结构

Fig. 3 AFPN structure

### 1.3 强感知分割头 DS-Seghead

解决了小目标缺陷在特征提取阶段容易出现信息丢失和退化的问题,还需强化网络对小目标缺陷的感知能力。文献[27]利用 FPN 提取的特征,在输出新增一条语义分割的附加分支,提出了一种新的语义分割网络,该网络通过双线性插值上采样将各级提取的特征调整为相同分辨率的特征图,然后将各个特征图按矩阵元素求和,通过逐点卷积、上采样和 Softmax 生成最终的语义分割分支。

由于网络中使用的标准卷积在小目标检测任务中性能有限,联想到条状缺陷呈现细长特征,本文引入动态蛇形卷积<sup>[20]</sup>,利用形态学预测目标的位移方向,网络结构如图 4 中 DSConv 所示。当网络学习到整个卷积核变化的所有偏执后,添加一个偏执约束,用以动态调整卷积核位置,每一个卷积的移动都以前一个位置作为基准,类似蛇移动时由头带动身体,一节一节摆动的姿态,实现对条状缺陷的特征提取。以一个  $1 \times 1$  的卷积核  $K$  为例,如果卷积中心点坐标表示为  $K_i = (x_i, y_i)$ ,偏执约束记为  $\Delta = \{\lambda \mid \lambda \in [-1, 1]\}$ ,则第  $c$  次调整后卷积核中心点坐标如式(4)所示。

$$K_{i+c} = \begin{cases} (x_{i+c}, y_{i+c}) = (x_i + c, y_i + \sum_{i+1}^{i+c} \Delta_i) \\ (x_{i-c}, y_{i-c}) = (x_i - c, y_i + \sum_{i-c}^{i-1} \Delta_i) \end{cases} \quad (4)$$

使用 DSConv 替换原始语义分割分支中的标准卷积,本文提出了 DS-Seghead,网络架构如图 4 所示。针对经过颈部网络融合后输出的多尺度特征,进行卷积操作,随后上采样处理,通过残差连接将每一层的输出进行融合,有

效结合各级特征信息,最终实现对缺陷的精准语义分割。

### 1.4 DDS 训练策略

鉴于颈部网络在整体模型中的重要作用,连接解码器和编码器,丰富两者间信息交流,减少特征在传播过程中的损失和退化,确保信息的完整性和准确性,本文提出了 DDS 训练策略,借助 DSCConv 的思想,重构 AFPN 中的上采样操作,使得模型对条状表面缺陷的特征感知能力进一步加强,提升小目标检测的性能,增强模型对复杂缺陷场景的适应能力,优化后的网络结构如图 5 所示。

## 2 实验与分析

### 2.1 实验环境

实验基于 Windows 11 操作系统;Python 版本 3.8;CUDA 版本 11.8;GPU 为 NVIDIA RTX A4000,16 GB 显存。在训练过程中,迭代次数设置为 80 000, batch size 为 4,验证时 batch size 为 1,损失优化器使用 AdamW,学习率采用 PolyLR 策略随着迭代不断更新,如式(5)所示。

$$lr = lr_{base} \times \left(1 - \frac{iter}{maxiter}\right)^{power} \quad (5)$$

其中,  $lr$  为新的学习率;  $lr_{base}$  为初始学习率,训练设置为 0.000 2;  $iter$  为迭代次数;  $maxiter$  为最大迭代次数;  $power$  用于控制学习率更新速度,训练设置为 0.9。

### 2.2 评价指标

为了评估算法的分割性能,使用预测精度( $Pre$ )、召回率( $Rec$ )以及交并比( $IoU$ )作为分割精度评估指标,由于预测精度体现模型对负样本的区分能力,而召回率体现模型对正样本的识别能力,为平衡精度与召回率,引入 F1 分数( $F1$ )评价

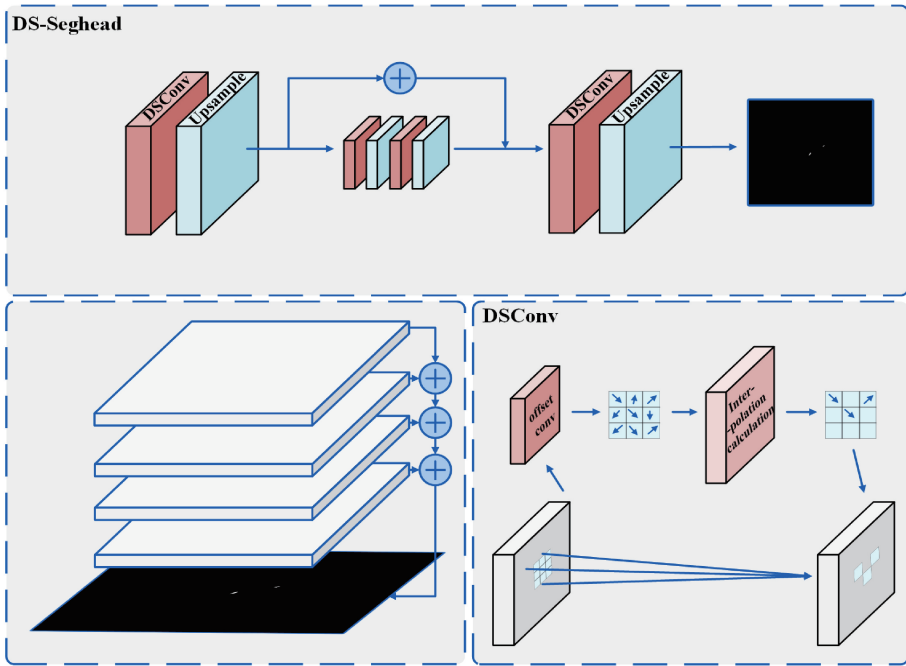


图 4 DS-Seghead 结构

Fig. 4 DS-Seghead structure

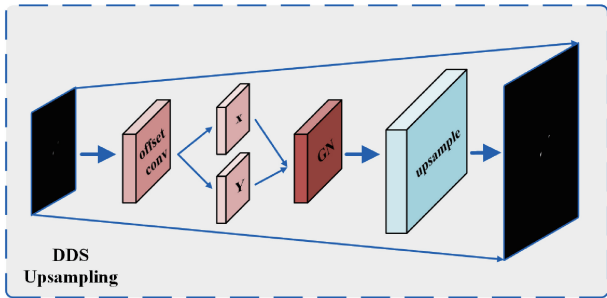


图 5 DDS Upsampling 结构

Fig. 5 DDS Upsampling structure

模型的鲁棒性,定义如式(6)~(9)所示。

$$Pre = \frac{TP}{TP + FP} \quad (6)$$

$$Rec = \frac{TP}{TP + FN} \quad (7)$$

$$F1 = \frac{2 \times Pre \times Rec}{Pre + Rec} \quad (8)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (9)$$

其中,  $TP$  表示将标注为正样本的目标预测为正例的数量,  $FP$  表示将标注为负样本的目标预测为正例的数量,  $FN$  表示将标注为正样本的目标检测为负例的数量。

此外,使用模型可学习参数量(Parameters)、浮点数量(FLOPs)作为算法的轻量化评估指标。

### 2.3 数据集分析与预处理

FDMPI 数据集中图像来自湖北三环锻造有限公司戴

姆勒奔驰重卡转向节生产线的磁粉检测过程。数据集中包含 362 张尺寸为  $1\ 024 \times 768$  的可见光状态下重卡转向节表面裂纹图像,按 4:1 的比例划分为训练集和验证集,测试集与验证集相同。数据集中部分图像如图 6 所示。由于算法针对图片进行像素级检测,通过在每次迭代中对图片执行随机裁剪和翻转等变换,可以有效实现数据集的动态扩充,增加训练样本的多样性,提高模型学习的随机性,从而增强算法的泛化能力。

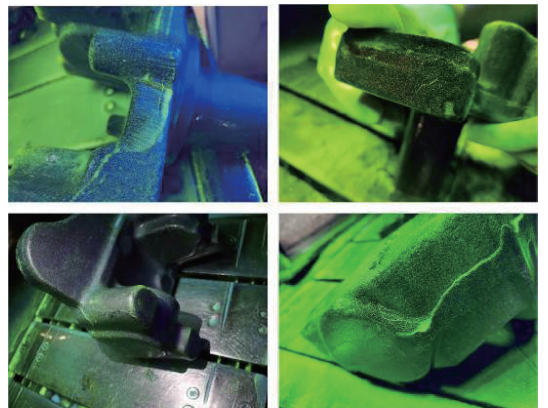


图 6 FDMPI 数据集

Fig. 6 FDMPI dataset

NEU-Seg 数据集由东北大学宋克臣团队提出<sup>[31]</sup>,包含 3 类热轧带钢表面缺陷,数据集中部分图片如图 7 所示。其中图 7(a)为夹杂缺陷,图 7(b)为斑点缺陷,图 7(c)为划伤缺陷。数据集包含尺寸为  $200 \times 200$  的灰度图像 4 470 张,其中训练集 3 630 张,验证集 840 张,测试集与验证集

相同。与 FDMPI 数据集相比, NEU-Seg 数据集中的缺陷更为宽大。

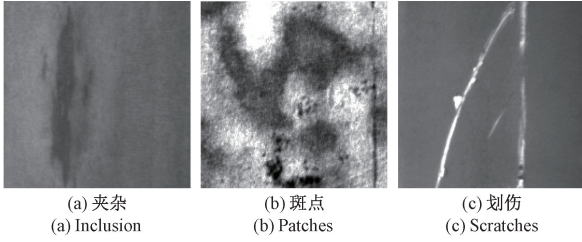


图 7 NEU-Seg 数据集

Fig. 7 NEU-Seg dataset

## 2.4 实验分析

### 1) 消融实验

首先,本文通过消融实验在 FDMPI 数据集上探索 AFPN、DS-Seghead 以及 DDS 训练策略对模型的影响,实验结果如表 1 所示。No. 1 是基准模型的实验结果,代表使用 PoolFormer 作为编码器,FPN<sup>[26]</sup>与语义分割分支<sup>[27]</sup>作为解码器组成的算法。基准模型的精度为 74.32%、召回率为 64.71%、F1 分数为 69.18%、IoU 为 52.88%、模型参数数量为 15.65 M、计算量为 246 G。No. 2 使用 AFPN 替换 FPN 网络,在 F1 分数提升约 1% 的情况下,精度和召回率

分别提升至 73.52%、66.85%,并且 IoU 提升至 53.88%,说明 AFPN 的渐进特征提取策略和空间自适应融合方法对锻件表面裂纹缺陷的检测具有显著效果,此外模型的参数量、计算量均大幅度下降,轻量化效果明显。No. 3 使用 DS-Seghead 替换语义分割分支,在 F1 分数有小幅度提升的情况下,召回率小幅度下降,但精度提升至 75.93%,同时 IoU 的提升效果小于 No. 2,说明 DS-Seghead 虽然能提升模型的检测性能,但如果缺少 AFPN 提供丰富的语义信息,提升效果有限。No. 4 同时使用 AFPN 和 DS-Seghead 对基准模型进行改进,与基准模型相比,精度、召回率、F1 分数以及 IoU 分别提高了 0.1%、2.52%、1.47% 和 1.73%,同时模型参数量、计算量分别下降了 9.33%、48.78%,说明 LSC-PoolFormer 算法相较基准算法鲁棒性更强、分割精度更高、模型更轻量化。No. 5 在 No. 4 的改进基础上引入 DDS 训练策略,此时 F1 分数、IoU 以及轻量化指标均与 No. 4 达到同一水平,相较基准模型,虽然精度下降了 3.71%,但召回率实现了 6.23% 的提升,说明模型牺牲了部分负样本区分的性能,以提高对正样本的识别能力。在锻件表面缺陷检测应用中,缺陷漏检比非缺陷误识别为缺陷带来的安全代价更大,因此 No. 5 提出的模型更适合实际生产。

表 1 消融实验

Table 1 Ablation experiment

No	Base	AFPN	DS-Seghead	DDS	Parameters/M	FLOPs/G	Pre/%	Rec/%	F1/%	IoU/%
1	✓	—	—	—	15.65	246	74.32	64.71	69.18	52.88
2	✓	✓	—	—	13.88	147	73.52	66.85	70.03	53.88
3	✓	—	✓	—	15.96	224	75.93	64.14	69.54	53.30
4	✓	✓	✓	—	14.19	126	74.42	67.23	70.65	54.61
5	✓	✓	✓	✓	14.21	126	70.61	70.94	70.28	54.57

FDMPI 上的消融实验可视化结果如图 8 所示,基于图中第 2 行和第 3 行, No. 5 模型对小目标裂纹有更好的识别效果,相较于其他 4 种模型,漏检率最低。基于图中第 1 行,相比 No. 2 和 No. 4 的分割结果, No. 3 更接近真实情况,这一结果表明 AFPN 虽然实现了轻量化的特征融合,丰富语义信息,但是可能导致模型出现过拟合,结合 No. 5 的分割结果,说明基于 DDS 能有效增强模型对小目标裂纹的感知能力,抑制过拟合,实现更精准的分割,进一步证明了 No. 5 提出的模型在实际生产中能更精准分割出锻件表面条状小目标裂纹缺陷。

### 2) FDMPI 上的对比实验

为进一步证明 LSC-PoolFormer 的性能,本文使用 FCN<sup>[32]</sup>、EncNet<sup>[33]</sup>、DeepLabv3 +<sup>[34]</sup>、UPerNet<sup>[35]</sup>、SegFormer<sup>[23]</sup>、PoolFormer<sup>[17]</sup> 在 FDMPI 数据集上进行对比实验,结果如表 2 所示。表中前 4 种算法均使用 Resnet-50 作为主干网络,后 4 种算法使用基于 Transformer 的改

进算法作为编码器。对比模型轻量化的两类评价指标,后 4 种算法均展现了显著的优势,这一结论凸显 Transformer 系列算法架构设计上的简洁性,可以有效降低模型复杂度。

与 LSC-PoolFormer 算法相比,FCN 虽然在预测精度上提高了 1.72%,但召回率下降了 10.64%;EncNet 作为四类使用 Resnet-50 作为主干网络中最轻量化的算法,精度、召回率以及 IoU 均达不到 LSC-PoolFormer 算法的同等水平;UPerNet 算法虽然在模型分割精度和鲁棒性的评价指标上呈现与 LSC-PoolFormer 算法相近的水平,但其参数量、计算量呈现翻倍增长的现象。与 SegFormer 相比,尽管 LSC-PoolFormer 由于引入 AFPN 颈部网络造成参数量上涨,但计算量涨幅较小,同时模型分割精度和鲁棒性的评价指标均处于领先地位,其中 F1 分数与 IoU 分别提升 1.36%、1.99%,在追求模型轻量化的同时,综合考虑了分割精度和鲁棒性;与 PoolFormer 相比,本文提出的

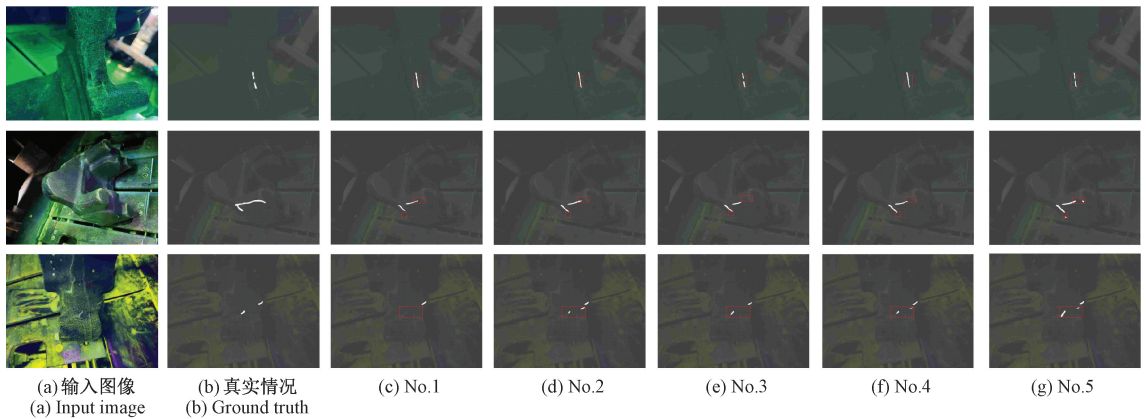


图 8 消融实验分割预测图

Fig. 8 Ablation experiment segmentation prediction map

表 2 FDMPI 上的对比实验

Table 2 Comparative experiments on FDMPI

算法	Parameters/M	FLOPs/G	Pre/%	Rec/%	F1/%	IoU/%
FCN	47.12	1583	76.14	56.59	64.92	48.06
EncNet	33.51	1127	70.41	60.35	64.99	48.14
DeepLabv3+	41.22	1412	68.84	54.26	60.68	43.56
UPerNet	64.04	1895	75.44	64.29	69.42	53.16
SegFormer	3.716	121	71.68	66.37	68.92	52.58
PoolFormer	15.65	246	74.32	64.71	69.18	52.88
LSC-PoolFormer(本文)	14.19	126	74.42	67.23	70.65	54.61
LSC-PoolFormer+DDS(本文)	14.21	126	70.61	70.94	70.28	54.57

改进算法在模型轻量化、分割精度和鲁棒性均取得先进水平。模型启用 DDS 训练策略之后,在维持相同预测精度的前提下,实现高召回率,在保持轻量化和强鲁棒性的同时,强化模型对小目标缺陷的感知能力,降低缺陷漏检概率,更适合实际生产需求。

图 9 展示了 FDMPI 数据集上的对比实验可视化结果,图中第 2 行和第 5 行显示,FCN、EncNet、DeepLabv3+ 和 UPerNet 算法对小目标均出现了不同程度的漏检,另外基于 Transformer 的 4 类算法效果较好,结合表 2 中数据分析,可能是前 4 类算法的低召回率导致。基于第 3 行、第 4 行和第 6 行,EncNet、SegFormer 和 PoolFormer 均出现了一定程度上的误检。综合,LSC-PoolFormer 算法启用 DDS 训练策略前后相较其他算法,在实现轻量化的同时,均有更好的分割表现,实验结果表明,针对复杂可见光环境下小目标缺陷,本文提出的算法可以提供更好的分割结果。

### 3) NEU-Seg 上的对比试验

为验证 LSC-PoolFormer 的泛化能力,本文在 NEU-Seg 数据集上进行对比试验,实验结果如表 3 所示(表中 m 表示对 3 类缺陷,相关评价指标的平均计算)。对比 LSC-PoolFormer 算法,前 4 类算法中,只有 DeepLabv3+ 和

UPerNet 在平均预测精度方面分别实现 1.42% 和 0.56% 的提升,其余指标均未达到同等水平。对比 SegFormer 算法,LSC-PoolFormer 在轻量化方面不及,但鲁棒性和分割精度方面仍略微领先;相比 PoolFormer,本文提出的算法在轻量化、分割精度以及鲁棒性方面均有提升。

图 10 展示了 NEU-Seg 数据集上的对比实验可视化结果,第 2 行和第 5 行显示,表中前 6 种算法均出现不同程度的误检,其中 FCN 最明显,结合表 3 中数据分析,可能是 FCN 高召回率、低精度导致。在第 3 行和第 4 行中,EncNet、DeepLabv3+、SegFormer 和 PoolFormer 针对小目标均出现较明显的漏检。对比启用 DDS 训练策略前后的结果,发现启用 DDS 的 LSC-PoolFormer 算法在第 2 行和第 4 行的检测效果强于未启用,这表明可能由于 NEU-Seg 数据集中小目标缺陷占比较少,最终导致未启用 DDS 的数据表现强于已启用,这进一步表明 DDS 策略有利于模型对小目标的检测。

综合实验数据,相较同样采用 Transformer 的改进算法作为编码器的 SegFormer 和 PoolFormer 算法,本文提出的模型在 NEU-Seg 数据集上仍实现了分割精度提升,不局限于单一特定的数据集,尽管提升幅度相较于 FDMPI 数据集上的表现,显得更为温和,但这表明,LSC-PoolFormer

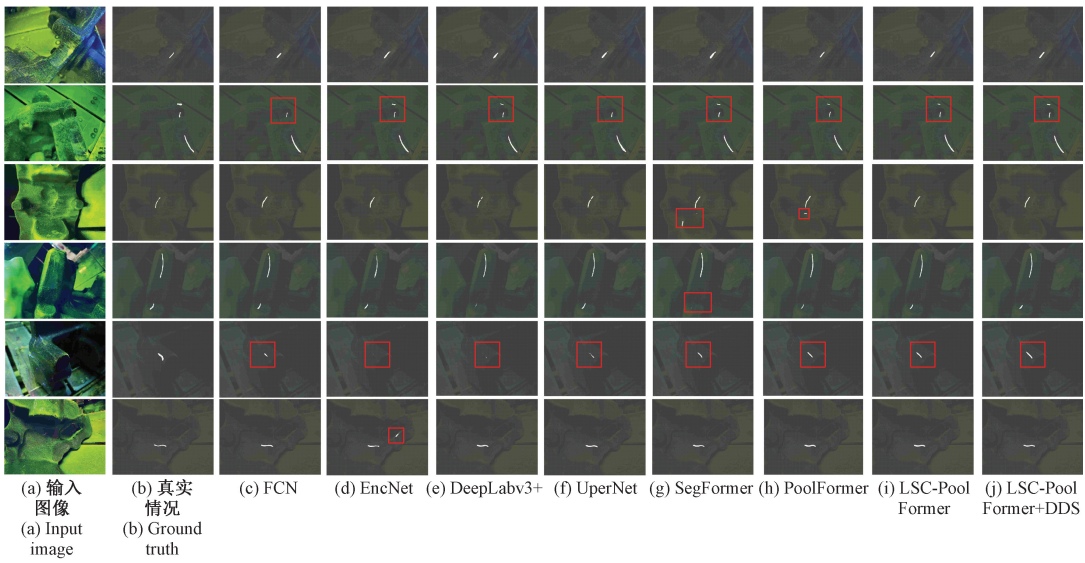


图 9 FDMPI 上的对比实验分割预测图

Fig. 9 Comparative experimental segmentation prediction map on FDMPI

表 3 NEU-Seg 上的对比实验

Table 3 Comparative experiments on NEU-Seg

算法	Parameters/M	FLOPs/G	MPre/%	MRec/%	mF1/%	MIoU/%
FCN	47.13	1583	85.84	87.99	86.89	77.09
EncNet	33.51	1127	89.21	85.99	87.55	78.11
DeepLabv3+	41.22	1412	91.15	82.62	86.65	76.74
UPerNet	64.04	1895	90.29	83.73	86.86	77.05
SegFormer	3.716	121	88.45	88.82	88.63	79.75
PoolFormer	15.65	246	90.14	86.40	88.23	79.13
LSC-PoolFormer(本文)	14.19	126	89.73	88.07	88.89	80.17
LSC-PoolFormer+DDS(本文)	14.21	126	89.76	87.77	88.75	79.97

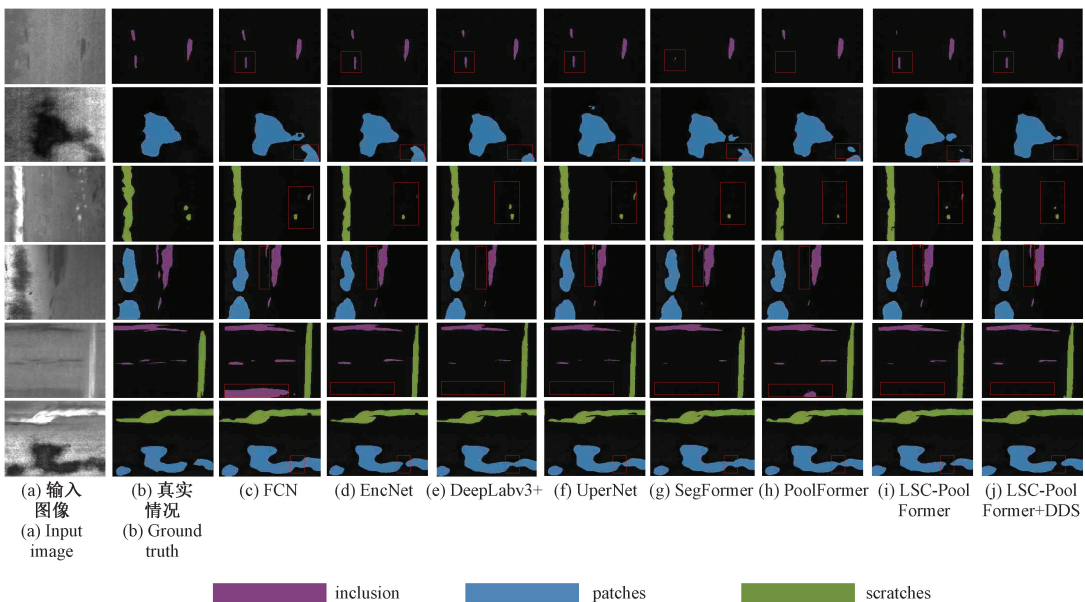


图 10 NEU-Seg 上的对比实验分割预测图

Fig. 10 Comparative experimental segmentation prediction map on NEU-Seg

具有较强泛化能力。模型在不同数据集上的性能表现会受到数据特性的影响,具体而言,FDMPI 主要包含小目标裂纹缺陷,而 NEU-Seg 中缺陷类型多样且大多占据较大的画面比例,小目标缺陷占比较低。这种数据分布差异导致 LSC-PoolFormer 在 NEU-Seg 上的性能提升幅度相对有限。结合可视化结果图,本文提出的算法针对锻件表面的小目标缺陷具有更好的分割效果。

### 3 结 论

本文提出了一种用于分割锻件表面缺陷的轻量级条状卷积算法 LSC-PoolFormer。针对复杂可见光环境下的条状小目标缺陷容易产生漏检的问题,首先采用 PoolFormer 作为编码器实现轻量级特征提取,其次引入 AFPN 作为颈部网络实现多尺度特征融合,减少层与层之间语义信息损失,并借助 DSConv 设计 DS-Schead 强化解码器对缺陷的感知能力,最后提出 DDS 训练策略进一步提升网络对小目标缺陷的检测能力。本文在提出的 FDMPI 数据集上进行实验,证明了 LSC-PoolFormer 算法的性能,并在公开数据集 NEU-Seg 上验证了其泛化能力,在未来工作中,将进一步扩充 FDMPI 数据集,将算法部署到真实工业场景中做进一步评估。

### 参考文献

- [1] 张辉,宋雅男,王耀南,等. 钢轨缺陷无损检测与评估技术综述[J]. 仪器仪表学报,2019,40(2):11-25.  
ZHANG H, SONG Y N, WANG Y N, et al. Review of rail defect non-destructive testing and evaluation [J]. Chinese Journal of Scientific Instrument, 2019, 40(2): 11-25.
- [2] 王宸,唐禹,张秀峰,等. 基于改进 EfficientNet 的锻件磁粉探伤智能检测方法研究[J]. 仪器仪表学报,2021,42(9):89-96.  
WANG CH, TANG Y, ZHANG X F, et al. An intelligent magnetic particle testing method for forgings based on the improved EfficientNet [J]. Chinese Journal of Scientific Instrument, 2021, 42(9): 89-96.
- [3] ZHAO ZH K. Review of non-destructive testing methods for defect detection of ceramics[J]. Ceramics International, 2021, 47(4): 4389-4397.
- [4] 朱栋,胡伟笑,赵腾. 太阳能电池片图像校正与表面缺陷检测[J]. 电子测量技术,2024,47(8):126-133.  
ZHU D, HU W X, ZHAO T. Image calibration and surface defect detection of solar photovoltaic cells[J]. Electronic Measurement Technology, 2024, 47(8): 126-133.
- [5] TANG B, CHEN L, SUN W. Review of surface defect detection of steel products based on machine vision[J]. IET Image Processing, 2022, 17(2): 303-322.
- [6] FU G ZH, SUN P Z, ZHU W B, et al. A deep-learning-based approach for fast and robust steel surface defects classification[J]. Optics and Lasers in Engineering, 2019, 121: 397-405.
- [7] CHEON S, LEE H, KIM C O, et al. Convolutional neural network for wafer surface defect classification and the detection of unknown defect class[J]. IEEE Transactions on Semiconductor Manufacturing, 2019, 32(2): 163-170.
- [8] ZHANG J B, SU H, ZOU W, et al. CADN: A weakly supervised learning-based category-aware object detection network for surface defect detection [J]. Pattern Recognition, 2021, 109, DOI: 10.1016/j.patcog.2020.107571.
- [9] 卢开喜,段先华,陶宇诚,等. KThin-YOLOV7:轻量级的焊接件表面缺陷检测[J]. 电子测量技术,2024,47(7):9-18.  
LU K X, DUAN X H, TAO Y CH, et al. KThin-YOLOV7: Lightweight inspection of surface defects on welded parts [J]. Electronic Measurement Technology, 2024, 47(7): 9-18.
- [10] GUO F, QIAN Y, YU H Y. Automatic rail surface defect inspection using the pixel-wise semantic segmentation model[J]. IEEE Sensors Journal, 2023, 23(13): 15010-15018.
- [11] DANG L M, WANG H X, LI Y F, et al. Lightweight pixel-level semantic segmentation and analysis for sewer defects using deep learning [J]. Construction and Building Materials, 2023, 371: 130792.
- [12] SILVA M I, MALITCKII E, SANTOS T G, et al. Review of conventional and advanced non-destructive testing techniques for detection and characterization of small-scale defects[J]. Progress in Materials Science, 2023, 138: 101155.
- [13] GUO B Y, WANG Y T, ZHEN SH, 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.
- [14] WANG K. Contrastive learning-based semantic segmentation for in-situ stratified defect detection in additive manufacturing[J]. Journal of Manufacturing Systems, 2023, 68:465-476.
- [15] LUO F, CUI Y, LIAO Y. MVRA-UNet: Multi-view residual attention U-Net for precise defect

- segmentation on magnetic tile surface [J]. IEEE Access, 2023, 11: 135212-135221.
- [16] 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.
- [17] YU W H, LUO M, ZHOU P, et al. Metaformer is actually what you need for vision [C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2022: 10809-10819.
- [18] YANG G Y, LEI J, ZHU ZH K, et al. AFPN: Asymptotic feature pyramid network for object detection[C]. 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2023: 2184-2189.
- [19] ZHANG Y J, YIN Y, SHAO Z Y. An enhanced target detection algorithm for maritime search and rescue based on aerial images[J]. Remote Sensing, 2023, 15(19): 4818.
- [20] QI Y L, HE Y T, QI X M, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation[C]. 2023 IEEE/CVF International Conference on Computer Vision(ICCV), 2023: 6047-6056.
- [21] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth  $16 \times 16$  words: Transformers for image recognition at scale [C]. International Conference on Learning Representations, 2021.
- [22] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]. 31st International Conference on Neural Information Processing Systems (NIPS), 2017: 6000-6010.
- [23] XIE EN Z, WANG W H, YU ZH D, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [C]. 35th International Conference on Neural Information Processing Systems(NIPS), 2024: 12077-12090.
- [24] WANG H Q, WANG K, YAN T, et al. Endoscopic image classification algorithm based on Poolformer[J]. Frontiers in Neuroscience, 2023, 17: 1273686.
- [25] LI B, FAN W K, YANG M X, et al. DUP-Net: Double U-poolformer networks for renal artery segmentation in CT urography[C]. 2023 International Joint Conference on Neural Networks(IJCNN), 2023: 1-8.
- [26] 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(CVPR), 2017: 936-944.
- [27] KIRILLOV A, GIRSHICK R, HE K M, et al. Panoptic feature pyramid networks[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2019: 6392-6401.
- [28] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2018: 8759-8768.
- [29] ZHAO G M, GE W F, YU Y ZH. GraphFPN: Graph feature pyramid network for object detection[C]. 2021 IEEE/CVF International Conference on Computer Vision(ICCV), 2021: 2743-2752.
- [30] LIU S T, HUANG D, WANG Y H. Learning spatial fusion for single-shot object detection [J]. ArXiv preprint arXiv.1911.09516, 2019.
- [31] BAO Y Q, SONG K CH, LIU J, et al. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [32] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2015: 3431-3440.
- [33] ZHANG H, DANA K, SHI J, et al. Context encoding for semantic segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2018: 7151-7160.
- [34] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. European Conference on Computer Vision(ECCV), 2018: 833-851.
- [35] XIAO T, LIU Y CH, ZHOU B L, et al. Unified perceptual parsing for scene understanding [C]. European Conference on Computer Vision (ECCV), 2018: 432-448.

## 作者简介

张上,副教授,主要研究方向为图像处理和人工智能。

E-mail: wetoo@sina.com

邹扬,硕士研究生,主要研究方向为计算机视觉和深度学习。

E-mail: zouyang@ctgu.edu.cn

张岳(通信作者),硕士研究生,主要研究方向为计算机视觉和深度学习。

E-mail: zhangyue980202@ctgu.edu.cn