

# 基于改进 SegFormer 网络的线激光分割和中心提取方法<sup>\*</sup>

韩佳鑫<sup>1</sup> 王生怀<sup>1</sup> 钟明<sup>2</sup> 陈哲<sup>2</sup> 张伟<sup>1</sup>

(1. 湖北汽车工业学院机械工程学院 十堰 442000; 2. 武汉筑梦科技有限公司 武汉 430000)

**摘要:** 线结构光条纹中心线提取是三维测量技术的关键,测量工件表面易反光等环境因素的干扰使中心线提取的精度低、稳定性差。本文提出了一种改进的激光中心线提取方法,首先在 SegFormer 网络编码层 Transformer backbone 分支提取线激光图像全局特征的基础上,引入 Vgg16 backbone 分支提取线激光图像中的浅层轮廓信息,再添加 MASPP 模块来提高模型对线形目标的分割效果,提高了激光条纹区域的分割精度,通过改进 SegFormer 网络模型为后续中心线提取提供高质量的图像源,再利用 Steger 法实现对激光中心线的精确提取。实验结果表明,该方法计算速度与 Steger 算法相比提升了 42%,其提取的精度提升约 0.3 个像素,并适用于多种复杂环境,在工业检测上满足精度和稳定性的要求。

**关键词:** 深度学习;线结构光;光条中心提取

中图分类号: TN209 文献标识码: A 国家标准学科分类代码: 510.2

## Improved SegFormer network-based line laser segmentation and center extraction method

Han Jiaxin<sup>1</sup> Wang Shenghuai<sup>1</sup> Zhong Ming<sup>2</sup> Chen Zhe<sup>2</sup> Zhang Wei<sup>1</sup>

(1. School of Mechanical Engineering, Hubei University of Automotive Technology, Shiyan 442000, China;

2. Wuhan Dream Technology Co., Ltd., Wuhan 430000, China)

**Abstract:** The extraction of the centerline from multi-line structured light is a critical technique in three-dimensional measurement technologies. Reflectivity and other environmental factors on the surface of the object being measured commonly result in low accuracy and instability in extracting the centerline. This thesis proposes an enhanced laser centerline extraction method. It begins by harnessing global features from line laser images, extracted through the Transformer backbone branch within the encoding layer of the SegFormer network. Additionally, the method integrates the Vgg16 backbone branch to capture shallow contour details from the line laser images. The incorporation of the MSASPP module significantly refines the model's ability to segment linear targets, thus elevating the segmentation accuracy within the laser stripe area. This refined SegFormer network model supplies a superior image source for subsequent centerline extractions, utilizing the Steger method to achieve precise detections. Experimental evidence indicates a 42% enhancement in computational speed over the Steger algorithm, with a notable increase in extraction accuracy by approximately 0.3 pixel. This method proves effective in diverse and complex environments, satisfying industrial demands for precision and stability in inspections.

**Keywords:** deep learning; line structured light; light strip center extraction

## 0 引言

中线激光扫描技术作为一种非接触式的三维测量技术,利用激光发射器通过投射激光到物体表面,并通过相机捕捉光线在物体表面上的变形来获取三维信息。线激光检测作

收稿日期:2024-08-06

\* 基金项目:国家自然科学基金(51675167)、湖北省重点研发计划项目(2021BAA056)、湖北省高等学校优秀中青年科技创新团队计划项目(T2020018)资助

为重要的视觉测量方法,具有精度高、抗干扰能力强、通用性强、结构简单等优点,被广泛应用于复杂环境下工件的三维测量、三维图像重建、质量检测等领域<sup>[1]</sup>。在测量过程中,使用线激光器投射激光束,相机拍摄光条图像,计算像素中心坐标,通过标定关系,再转换为三维坐标。线激光条纹中心提取的精度和速度是线激光三维测量技术的关键,其决定三维测量、重建及目标检测等的最终的精度和速度。

目前常见的中心线提取方法有极值法、灰度重心法、基于 Hessian 的 Steger 算法、骨架细化法等。极值法适合灰度分布满足理想高斯分布的光条,计算速度很快,但易受到图像背景噪声的影响,信噪比大时效果较差。陈云赛等<sup>[2]</sup>针对此问题提出新型条纹中心线提取算法,既保持多阈值法的快速性又融合高斯拟合法的准确性。灰度重心法是一种针对亮度分布不均匀的光条而提出的算法,计算效率高,但精度低,噪声影响比较大。万卓仁等<sup>[3]</sup>提出了一种自适应窗口灰度重心法,通过高斯曲线能量分布确定有效窗口,从而减少噪声和光照变化的影响。Steger 算法的核心特征是利用 Hessian 矩阵得到光条的法线方向,通过高斯拟合来得到光条纹中心的亚像素位置。该算法在局部分析的情况下精度高、鲁棒性好,但是计算复杂度高,提取速度慢。王志永等<sup>[4]</sup>提出了改进 Steger 算法流程的线激光中心提取方法,粗略估算光条宽度降低算法冗余度,再运用 Steger 算法对感兴趣区域内光条进行中心提取,提高了光条中心提取效率。骨架细化法将二值图像中的物体形状简化为细线骨架,同时保留物体的拓扑结构,但未考虑光条纹的横截面灰度分布特点,提取的光条纹中心线精度有限。周祥曼等<sup>[5]</sup>对图像进行滤波操作,基于改进的细化算法对提取中心进行冗余去除和补线处理,实现了线激光条纹中心快速和高精度的提取。

以上传统中心线提取算法在光条质量良好的前提下,可以保证检测精度。但在实际测量中,受到物体表面反光均匀度、平整程度、采集设备等因素的影响,导致光条产生变化,两侧产生噪点现象,对光条中心提取精度产生不同程度的干扰。

随着人工智能技术的不断发展,为了解决传统中心线提取算法的问题,学者们提出利用深度学习方法对原始光条图像进行预处理,消除了噪声点和反光等干扰因素的影响,以此提高后续处理算法的抗干扰能力。陈新禹等<sup>[6]</sup>提出一种在 U<sup>2</sup>-Net 网络中加入 TSA、TCA 模块的新方法,提高了模型的特征提取能力,有效去除图像中的噪声、毛刺,但存在分割后的中心线边缘不够清晰的问题;张勇停等<sup>[7]</sup>用端到端的深度学习方法提取线结构光条纹中心,处理速度提高,但是对环境要求比较严格;王胜春等<sup>[8]</sup>用 ENet 深度学习模型实现了激光条纹的多区段快速分割,提高了检测的时效性和准确性,但存在细节特征表达不充分的问题;王子涵等<sup>[9]</sup>用 YOLOv5 网络对于激光域进行预分割从而提取激光条纹中心线,模型复杂,处理速度不够;Ming

等<sup>[10]</sup>提出了一种基于卷积神经网络和多层感知的激光条纹中心线提取神经网络,克服了不同杂散光的干扰,适应各种照明条件,精度得到了提升,但未提及处理时间;张宇等<sup>[11]</sup>通过改进 UNet++ 模型将光条图像进行分割,再利用灰度重心法实现对激光中心线的准确提取,处理速度得到有效提高的同时,导致精度下降。以上深度学习的方法精度都有不同程度的提高,但是局部特征不能准确提取。在干扰较多的情况下,无法准确衡量光条连续性,无法得到有利于后续处理的像素值分布关键信息,对中心点定位的提升有限。

针对复杂物体三维测量中线激光条纹图像出现严重反光、毛刺、大量噪点的现象,中心线提取的精度和速度无法兼顾的问题。本文提出一种改进的 SegFormer 网络语义分割网络,在编码器部分添加两个编码路径,提取图像中目标局部特征和全局特征。此外,通过添加多尺度条纹池化的空洞空间金字塔池化模块 (atrous spatial pyramid module for multi scale stripe pooling, MASPP),提高全局信息融合的能力,减少分割后激光条纹断裂和错分反光条纹的情况。在此基础上,进一步结合 Steger 方法,实现光条中心线亚像素提取,改善提取精度和抗噪声的能力。

## 1 基于改进的 SegFormer 线结构光条分割

### 1.1 SegFormer 网络简介

SegFormer<sup>[12]</sup> 可高效准确处理图像分割任务,由 Transformer<sup>[13]</sup> 和轻量级多层感知机 (multi-layer perceptron, MLP) 解码器组成。

如图 1 所示,编码器首先采用了一定重叠的贴片嵌入 (overlap patch embeddings, OPE) 模块,用于将图像分割成多个固定大小且互相重叠的矩形图像块,每个块包括独立区域和周围区域的信息,这种重叠有助于捕捉图像中的细节和上下文信息,从而确保元素的连续性与完整性。将得到的特征输入 Transformer block 中,任一 Transformer block 均采用 N 个叠加的自注意力层 (efficient self-attention, ESA) 与局部和全局前馈神经网络 (mix feedforward neural network, MixFFN) 层,构建全局关系。通过将不同的图像块在多个尺度上进行自注意力计算,与在前馈层的两层线性变换层中间加入深度可分离卷积来减少参数量,可更快速地获取全局信息。SegFormer 的编码器在较低阶段自然产生类似卷积的局部注意力,有助于捕捉图像的局部特征。同时,编码器在 Stage-4 阶段能够输出高度非局部的注意力,有效地捕捉上下文信息,进一步增强了模型的性能,输出多尺度特征,改善了测试时分辨率与训练分辨率不同而导致性能下降的问题。

SegFormer 采用轻量级 All-MLP 解码器,通过聚合不同层级的信息,结合了局部和全局注意力,生成强大的表示。解码器首先将多级特征通过 MLP 层进行通道维度统一,依次进行特征的上采样和拼接,最后融合特征后,导人

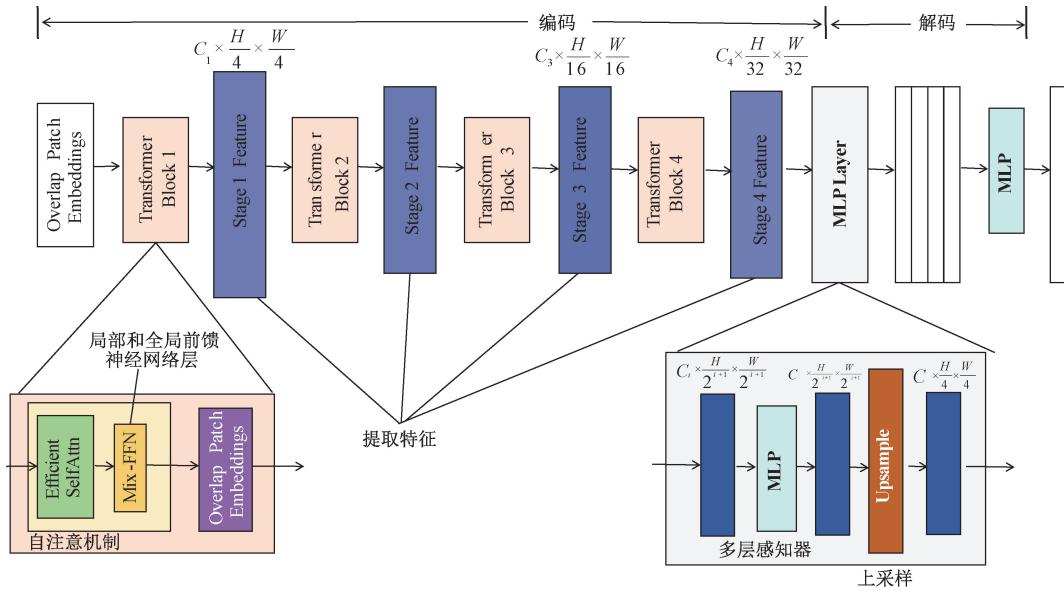


图 1 SegFormer 网络结构

Fig. 1 Network structure of SegFormer

另一个 MLP 模块来预测分割掩码。

虽然 SegFormer 的编码层 Transformer 在捕捉全局上下文信息方面的固有优势,但是 Transformer 模型的架构设计对于特征的提取和处理有一定的影响。Transformer 架构输入部分的编码器由多个子层组成,这种复杂的结构虽然能够有效地处理大量数据并提取出高级特征,但在转换分辨率过程中,引入位置编码进行插值时可能会因为信息处理和传递的复杂性而丢失大量边缘和轮廓等细节信息,在数据集不充分的情况下,局部浅层特征的提取也会受到影响。针对此问题,本文提出一种双分支特征提取框架,用来提升神经网络的提取能力。

## 1.2 改进的 SegFormer 网络结构

由于存在工件材料的反光特性不一,周围环境光的多次反射,相机和激光发射器中的光电噪声等多种因素的影响,导致图像中的光条亮度分布不均匀,个别图像出现严重反光现象,直接影响了光条中心点定位的精确度和稳定性。针对该难点,本文提出改进 SegFormer 网络对激光条纹图像进行预处理。

改进后的 SegFormer 的总体架构如图 2 所示,该网络基于编码器-解码器结构。相较于传统的 SegFormer 结构,改进的结构编码器部分由 Transformer backbone 分支和 Vgg16 backbone 分支<sup>[14]</sup>组成。此外,通过添加 MASPP 模块<sup>[15]</sup>,提高网络提取和融合特征,增强网络的分割性能。Vgg16 backbone 分支的负责提取图像的局部特征,填补 Transformer backbone 转化分辨率时丢失的浅层特征的空缺。MASPP 模块则通过多尺度空洞池化来提取局部条形特征,同时融合两分支提取到的全局和局部特征,从而提升图像分割的细致程度。将提取的特征输入到 MLP

Layer 中,进行特征融合,然后上采样进行恢复,进行解码后输出图片。

## 1.3 双分支特征提取编码部分

编码器部分结构为双分支编码结构,由 Transformer backbone 分支和 Vgg16 backbone 分支组成。Transformer backbone 分支进行全局特征提取,其中输入图像首先由嵌入图像块(patch embedding, PE)进行分割,最后输入到 4 个编码器块中,每个编码器块由 N 个叠加的 ESA 与 MixFFN 模块和 OPE 层组成。Vgg16 backbone 分支由 4 个 CNN 块组成,进行局部特征提取,将预训练的 Vgg16 作为骨干网络,其深度网络能产生高质量的分割掩码,每个编码块都利用最大池化操作来缩小输入特征并提取输入的代表性特征。前 3 个 Transformer backbone 和 Vgg16 backbone 编码模块的输出被馈送到 MLP Layer 中,为解码部分提供丰富的语义信息。每条分支的最后一个模块提取的特征进一步馈送到 MASPP 模块,用来聚合两个分支的最深层特征,并捕获不同尺度输入的上下文信息。

## 1.4 Transformer backbone 模块

Transformer backbone 模块如图 3 所示。

模块包括 ESA 和 MixFFN 层。在 ESA 模块中,网络的计算量主要体现在自注意力机制层上。为了降低网络整体的计算复杂度,在自注意力机制的基础上,添加缩放因子 R,来降低每一个自注意力机制模块的计算复杂度。原始的自注意力机制原理如式(1)所示。

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax} \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_{\text{head}}}} \mathbf{V} \quad (1)$$

式中: $\mathbf{Q}$  为查询矩阵,  $\mathbf{K}$  为键矩阵,  $\mathbf{V}$  为值矩阵, 都为  $N \times C$

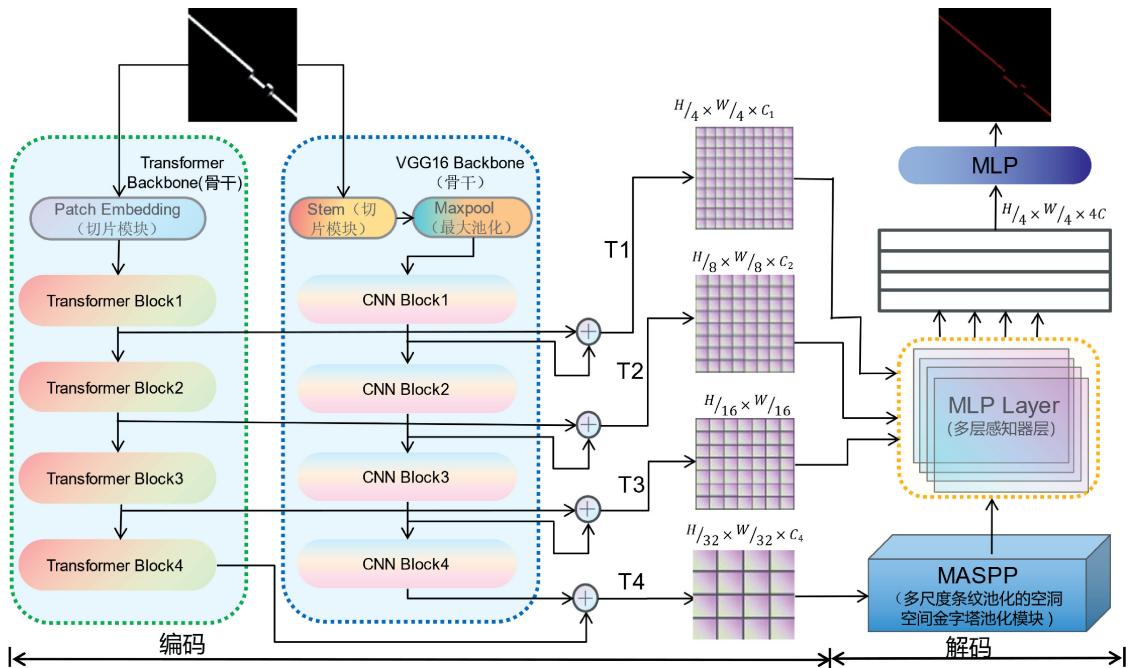


图 2 改进的 SegFormer 激光图像分割网络

Fig. 2 Improved SegFormer laser image segmentation network

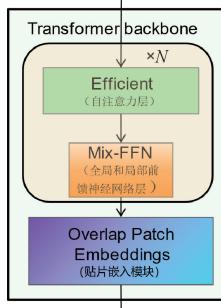


图 3 Transformer backbone 模块

Fig. 3 Transformer backbone module

的特征图,  $d_{head}$  是注意力的维度, 用于缩放查询和键之间的点积结果, 用来提高数值稳定性。

首先通过计算  $Q$  和  $K$  的点积  $QK^T$  得到相似性, 然后除以  $\sqrt{d_{head}}$  进行缩放。其次应用 Softmax, 将缩放后的点积结果通过 Softmax 函数转化为概率分布, 得到注意力权重。最后将得到的注意力权重应用于  $V$  上, 进行加权求和, 生成最终的注意力输出。添加缩放因子  $R$ 。具体操作如式(2)、(3)所示。

$$\hat{K} = \text{Reshape} \left( \frac{N}{R}, CR \right) (K) \quad (2)$$

$$K = \text{Linear}(CR, C)(\hat{K}) \quad (3)$$

式中:  $N$  为所有(片)Patch 的数量,  $C$  为每个 Patch 对应的维度。

通过 Reshape 操作把  $N \times C$  的特征图, 转化为  $(N/R \times CR)$  的特征图, 然后在经过一个全连接层, 把  $(N/R \times CR)$

转化为  $(N/R \times C)$ 。从 Transformer Block1 到 Transformer Block4 中, Transformer Block 的缩放因子分别为  $[64, 16, 4, 1]$ 。

为了缓解预测分辨率与训练分辨率不一致时位置编码插值导致的准确性下降问题, 引入了 MixFFN。MixFFN 考虑了零填充对位置信息泄漏的影响, 通过在前馈网络(feed forward network, FFN)中使用  $3 \times 3$  卷积来解决这一问题。每个 FFN 模块结合了一个  $3 \times 3$  卷积和一个 MLP, 用来提供位置信息, 通过使用深度可分离卷积来减少参数数量以提高效率。Mix-FFN 可以表示如式(4)所示。

$$\begin{cases} \mathbf{X}_{out} = \text{MLP}(\text{GELU}(Y)) + \mathbf{X}_{in} \\ Y = \text{Conv}_{3 \times 3}(\text{MLP}[\mathbf{x}_{in}]) \end{cases} \quad (4)$$

式中:  $\mathbf{X}_{in}$  是输入张量, 为上步自注意力模块产生的数据。

首先对输入  $\mathbf{x}_{in}$  使用 MLP 进行处理, 目的是对输入进行非线性变换。然后对经过 MLP 处理后的结果进行  $3 \times 3$  卷积操作, 提取局部特征, 接着用 GELU 激活函数进行激活, 将激活后的结果再通过 MLP 处理。最后将 MLP 处理后的结果与输入  $\mathbf{X}_{in}$  进行加法操作得到最终的输出结果。

## 1.5 MASPP 模块

线激光检测环境复杂, 工件形状各异, 多数检测图像中都是不同长度和角度的条形激光线。传统的条形池化呈方形或矩形, 只能按照特征图中固定的长度和宽度进行池化, 适应性有限。为了更好地提取线激光图像中不同长度和大小的条形特征, 条形池化(strip pooling, SP)模块<sup>[16]</sup>适合当下环境。SP 关注图像中的特定条形区域, 在保持图

像空间信息的同时,对局部区域进行汇聚,从而避免在标准池化中可能出现的信息丢失问题。为此,引用灵活可变的多尺度条纹池化(multi-scale stripe pooling, MSP)。MSP 的结构如图 4 所示。先将输入的特征图按横纵进行平均切分为 4 块,尺寸变为原来的  $1/4$ ,对每一块特征图进行横

向和纵向的条纹池化,特征图尺寸再次缩减为原来的  $1/3$ ,通过输入依赖深度卷积(ID Conv)对特征图进行扩展和融合,将池化后的特征图按初始位置拼接以恢复原始体积。模块通过设置特征图的划分数量来调整条纹池化的范围,以便捕捉不同尺寸目标的条状特征。

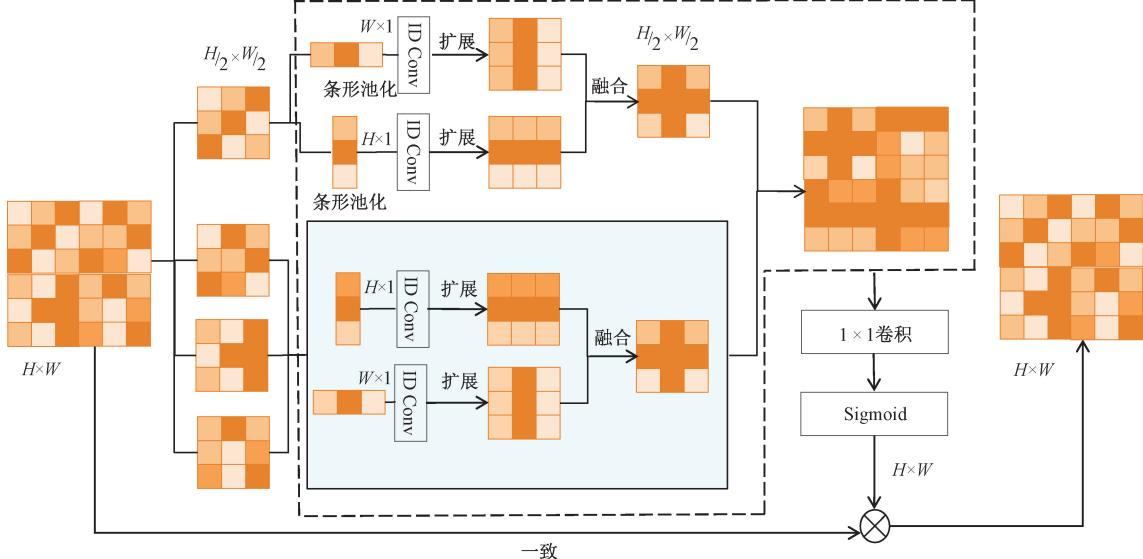


图 4 Strip Pooling 结构

Fig. 4 Strip pooling structure

空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP)模块是由不同尺度的空洞卷积形成的类金字塔结构,捕捉从局部到全局的多种特征信息。模块依次增加膨胀率,形成 3 种感受野的空洞卷积,卷积核间隔依次为 [6, 12, 18],特征单独输入到每个分支中进行处理,最后融合不同分支上的特征,以生成最终特征。因为线激光检测的数据集目标存在大量长短不同的线形激光目标,常规的全局平均池化提取线性激光目标效果佳,将原始 ASPP 中的全局平均池化替换为 3 个多尺度池化(MSP)模块,使得模块能够有效地提取和融合多种尺度的线性特征,增强模型对不同大小线性目标的理解能力。

MASPP 模块如图 5 所示。

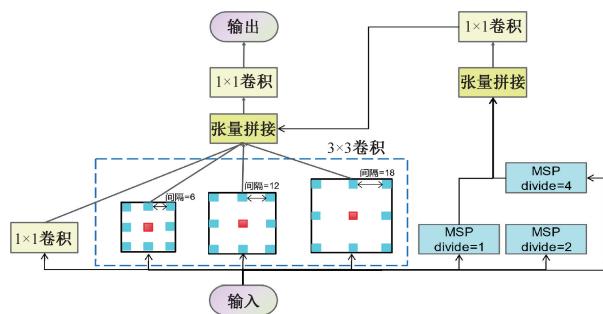


图 5 多尺度条纹池化的空洞空间金字塔池化模块

Fig. 5 Atrous spatial pyramid module for multi scale stripe pooling

## 1.6 特征解码部分

本文仍沿用 SegFormer 解码器,该解码器仅由 MLP<sup>[17]</sup>层组成,MLP 层如图 6 所示,避免了其他通用网络中使用计算量较大的组件。实现此简单有效的解码器的关键在于分层 Transformer 编码器具有比传统 CNN 编码器更大的有效感受野(effective receptive field,ERF)。

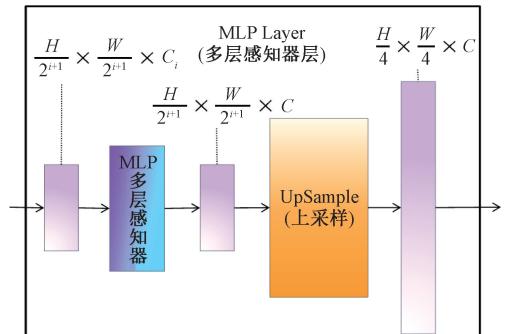


图 6 MLP 层结构

Fig. 6 MLP layer structure

全 MLP 解码器包括 4 个主要步骤:

1) 来自编码器的多级特征  $F_i$  通过一个 MLP 层进行通道维度统一。

$$\hat{F}_i = \text{Linear} = (C_i, C)(F_i), \forall i \quad (5)$$

2) 特征被上采样到  $1/4$  大小,进行拼接。

$$\widehat{\mathbf{F}}_i = \text{Upsample}\left(\frac{H}{4} \times \frac{W}{4}\right)(\widehat{\mathbf{F}}_i), \forall i \quad (6)$$

3) 采用 MLP 层来融合拼接后的特征  $\mathbf{F}$ 。

$$\mathbf{F} = \text{Linear}(4C, C)(\text{Concat}(\widehat{\mathbf{F}}_i)), \forall i \quad (7)$$

4) 另一个 MLP 层将融合后的特征输入, 预测分割掩码  $\mathbf{M}$ , 分辨率为  $H/4 \times W/4 \times N_{cls}$ 。

$$\mathbf{M} = \text{Linear}(C, N_{cls})(\mathbf{F}) \quad (8)$$

式中:  $C$  为维数,  $N_{cls}$  为类别的数量。

## 2 中心点亚像素提取

Steger 算法是目前使用最广泛的线结构光条纹中心提取算法之一, 该方法具有精度高, 稳定性好等优点。

Steger 算法是基于 Hessian 矩阵, 如式(9)所示。

$$\left\{ \begin{array}{l} \mathbf{H}(x, y) = \begin{bmatrix} \frac{\partial^2 \mathbf{G}(x, y)}{\partial x^2} & \frac{\partial^2 \mathbf{G}(x, y)}{\partial x \partial y} \\ \frac{\partial^2 \mathbf{G}(x, y)}{\partial x \partial y} & \frac{\partial^2 \mathbf{G}(x, y)}{\partial y^2} \end{bmatrix} \otimes \mathbf{I}(x, y) \\ \mathbf{I}(x, y) = \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{xy} & \mathbf{R}_{yy} \end{bmatrix} \end{array} \right. \quad (9)$$

式中:  $\mathbf{G}(x, y)$  是二维高斯卷积核, 用于突出光条的灰度分布特征。 $\mathbf{I}(x, y)$  是以图像上点  $(x, y)$  为中心, 与二维高斯核大小相等的图像矩阵。 $\mathbf{G}(x, y)$  如式(10)所示。

$$\mathbf{G}(x, y) = \frac{1}{\sigma \sqrt{2\pi}} \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \quad (10)$$

式中:  $\sigma$  为高斯函数的标准差, 一般取  $\sigma \geq \frac{w}{\sqrt{3}}$ ,  $w$  代表光条宽度<sup>[18]</sup>。

图像中点的法线方向由该点的 Hessian 矩阵最大特征值的绝对值对应的特征向量给出, 通过在法线方向上求极值点得到光条纹中心的亚像素位置。设某像素点  $P$  为  $(x_0, y_0)$ , 由像素  $P$  的灰度  $R(x_0, y_0)$  和二阶泰勒展开多项式(11)所示。

$$R(x_1, y_1) = R(x_0, y_0) + t\mathbf{n}[\mathbf{R}_x \mathbf{R}_y]^T + t\mathbf{n}\mathbf{H}(x, y)\mathbf{n}^T \quad (11)$$

式中:  $x_1 = x_0 + t\mathbf{n}_x$ ,  $y_1 = y_0 + t\mathbf{n}_y$ ,  $\mathbf{n}$  为单位向量,  $\mathbf{n} = (\mathbf{n}_x, \mathbf{n}_y)$ ,  $P$  点光条纹法相方向上的像素点为  $R = (x_0 + t\mathbf{n}_x, y_0 + t\mathbf{n}_y)$ 。

将  $t$  代入其中可求出光条纹中心的亚像素坐标。式(12)为  $t$  的表达式, 如下:

$$t = -\frac{\mathbf{n}_x \mathbf{R}_x + \mathbf{n}_y \mathbf{R}_y}{\mathbf{n}_x^2 \mathbf{R}_x + 2\mathbf{n}_x \mathbf{n}_y \mathbf{R}_{xy} + \mathbf{n}_y^2 \mathbf{R}_{yy}} \quad (12)$$

## 3 实验验证与分析

### 3.1 实验环境与数据集

为验证本文提出光条提取算法的可行性和精度, 自行搭建激光视觉测量实验平台, 如图 7 所示。

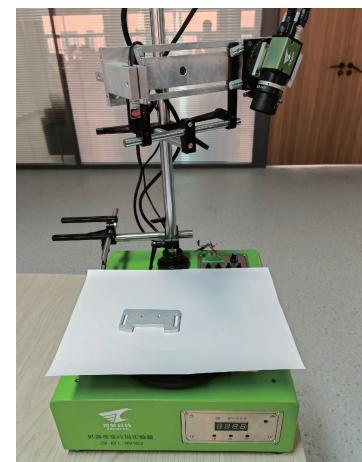
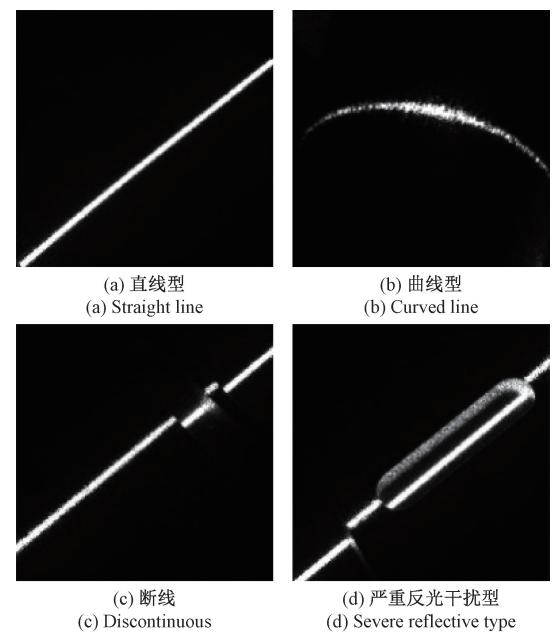


图 7 激光视觉测量平台

Fig. 7 Laser vision measurement platform

实验平台由筑梦科技有限公司 mv500 视觉实验平台、筑梦科技有限公司的 ZM-1300-30GM 相机、一字激光发射器、计算机等部分构成。其中计算平台配置为: Windows 操作系统、Pytorch 深度学习框架、Intel Core i5-13600F 型号 CPU (主频 3.6 GHz)、16 G 运行内存、NVIDIA GeForceRTX4060 显卡 (8 G 显存)。使用图 7 所示的线结构光三维视觉测量系统采集大量激光条纹图像, 并利用 LabelMe 软件对图像中的激光条纹用二分类法进行标记, 数据集图片如图 8 所示, 标注图像如图 8(f)所示, 本文采用自制数据集来评估提出的改进网络。如图 8 所示, 数据集包含了直线型激光 90 张、曲线型激光 66 张、长短不一断线激光 324 张、带有严重反光干扰的激光 236 张、边缘模糊激光 84 张, 每张检测图像具有  $1574 \times 1317$  像素分辨率。因为高分辨率图像较大, 会降低网络处理效率, 所以在对其进行语义分割之前, 要对图像进行剪裁。原始数据集中



(a) 直线型

(a) Straight line

(b) 曲线型

(b) Curved line

(c) 断线

(c) Discontinuous

(d) 严重反光干扰型

(d) Severe reflective type

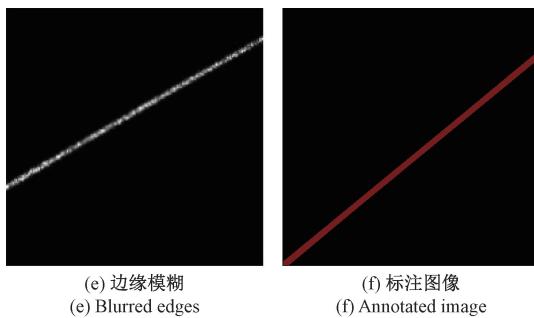


图 8 数据集图片

Fig. 8 Dataset images

的图片均被剪裁为 512 像素, 共 680 张图像用于训练, 80 张用于验证, 40 张用于测试。数据集的详细信息如表 1 所示。

表 1 数据集

Table 1 Datasets

项	激光
原始尺寸/px	1 574×1 317
分割后尺寸/px	512×512
训练集/张	680
验证集/张	80
测试集/张	40

### 3.2 激光分割实验评价指标

实验设置了 3 个针对分割性能的评价指标: 平均交并比 (mIoU)、平均像素准确率 (mPA) 和平均精确率 (mPrecision)。mIoU 衡量的是预测区域与真实区域的重叠程度; mPA 衡量的是全局的像素级别的分类准确性; mPrecision 衡量的是预测为某一类别的像素中实际属于该类别的比例。它反映了模型在局部区域的精确度, 可以反映局部特征的准确性。mIoU、mPA 和 mPrecision 的计算方式用式(13)~(15)表示<sup>[19]</sup>。

$$mIoU = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i + FN_i} \quad (13)$$

$$mPA = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i + TN_i + FN_i} \quad (14)$$

$$mPrecision = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i} \quad (15)$$

式中:  $N$ 、 $TP$ 、 $FP$ 、 $TN$ 、 $FN$  分别表示类别总数、真正例数、假正例数、真负例数和假负例数。

### 3.3 分割算法对比实验

本文将改进 SegFormer 同 SegFormer、Unet、DeepLab v3 和 Pspnet 网络在相同的环境下进行对比。线激光数据集分割评价指标如表 2 所示。

表 2 不同网络的评价指标

Table 2 Evaluation metrics of different networks

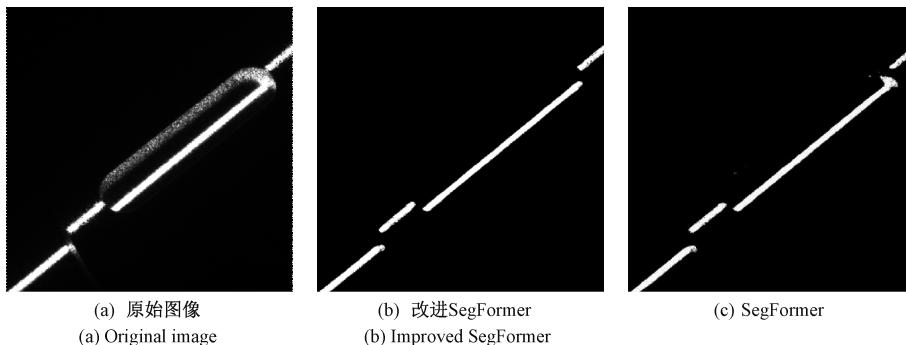
方法	mIoU	mPA	mPrecision
改进 segformer	0.907 2	0.953 2	0.953 3
Segformer	0.856 9	0.942 4	0.891 1
Unet	0.848 8	0.948 2	0.886 1
DeepLabv3	0.853 2	0.944 3	0.872 2
Pspnet	0.812 4	0.876 6	0.834 2

结果表明, 改进后的 SegFormer 网络分割精度最高。与其他的网络相比, 改进模型在所有评估指标上均表现最佳, 其中 mIoU、mPA 和 mPrecision 指标分别达到了 90.72%、95.32% 和 95.33%, mIoU 平均提高了 5%。

分割效果对比如图 9 所示。如图 9(b)所示, 改进后模型的分割图效果最好, 最接近标注图, 对激光线分割更加精确; 如图 9(c)所示, SegFormer 网络去除了明显的反光痕迹, 但是局部强反光没有去除, 局部轮廓也不够清晰; 如图 9(d)所示, Unet 网络存在一些反光干扰未被清除; 如图 9(e)所示, DeepLabv3 网络对图像中的局部强反光过处理, 激光条纹边缘呈现阶梯状; 如图 9(f)所示, Pspnet 网络处理得到边缘不平整激光条纹。

### 3.4 消融实验

本文选用原始 SegFormer 网络作为本次实验的基准网络。消融实验评价指标结果如表 3 所示。

(a) 原始图像  
(a) Original image(b) 改进 SegFormer  
(b) Improved SegFormer

(c) SegFormer

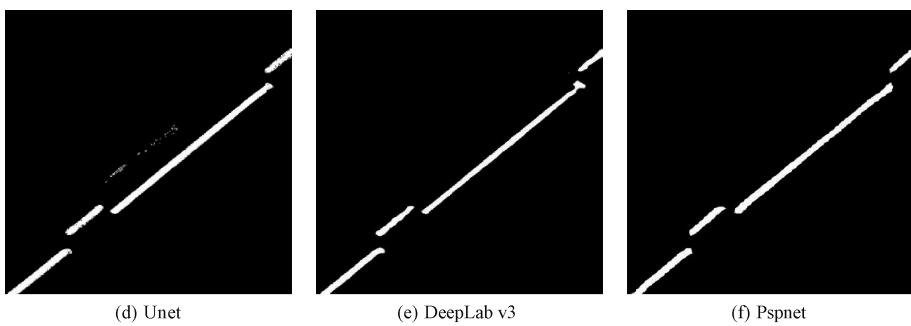


图 9 分割效果对比图

Fig. 9 Comparison of segmentation results

表 3 分割性能对比

Table 3 Comparison of segmentation performance

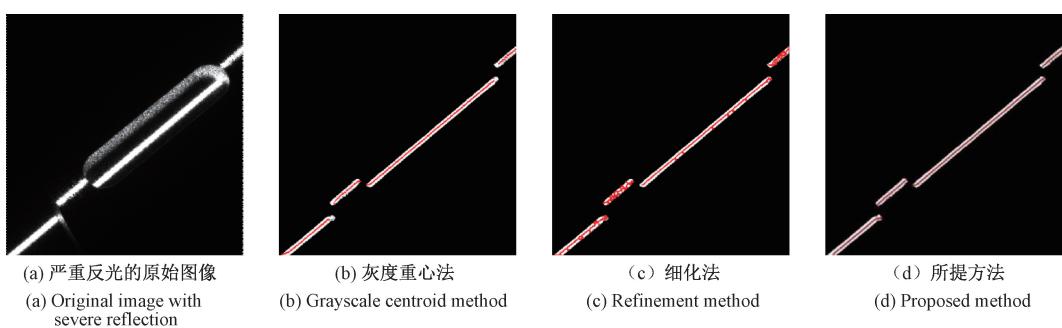
Base	Vgg16	Maspp	mIoU	mPA	mPrecision
✓	✗	✗	0.851	0.942	0.871
✓	✓	✗	0.897	0.951	0.914
✓	✓	✓	0.907	0.953	0.933

Base 表示基准网络。在编码层添加 Vgg16 分支后, 网络的 mIoU、mPA 指标平均提升 4%, mPrecision 指标上升 6%, 结果表明, 在添加 Vgg16 backbone 分支后模型能够有效提取目标的局部特征。进一步加入 MASPP 模块后, 网络的 mIoU、mPA 和 mPrecision 指标继续平均增长了 1.1%, 结果表明, MASPP 模块能较好的提取多尺度线形目标特征。随着这些关键模块的逐步加入, 集成所有模块的网络达到了最佳分割效果。

### 3.5 线激光条纹提取实验评价指标

采取光条中心点到拟合直线的标准差来评价提取算法的准确度, 如式(16)所示, 其中  $\bar{d}$  表示所有光条中心点到拟合直线的平均距离, 表示第  $d_i$  个光条中心点到拟合直线的距离,  $n$  为提取到的中心点个数。

$$\delta = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n}} \quad (16)$$



### 3.6 线激光提取实验

将本文方法同 steger 法、细化法和灰度重心法进行对比。中心点提取算法处理时间结果如表 4 所示。

表 4 各算法时间对比

Table 4 Comparison of algorithm times s

类型	本文	Steger	细化法	灰度重心法
1	0.253	0.361	0.582	0.029
2	0.257	0.364	0.598	0.031
3	0.266	0.391	0.626	0.032
4	0.285	0.424	0.693	0.034
5	0.334	0.453	0.773	0.037
平均	0.279	0.398	0.654	0.033

线激光提取效果对比如图 10 所示。由图 10(d)与(h)可知, 所提方法对激光线中心提取更加精确; 灰度重心法虽然处理时间较快, 但计算得到的重心位置不够准确, 结合图 10(b)与(f)可知, 中心线提取存在明显的断裂现象; 细化法涉及复杂的像素级别计算和迭代过程, 处理需要较长的时间, 在较多的噪声或者干扰情况下, 导致线中心提取的精度下降, 出现错误的条纹中心位置, 如图 10(c)所示; steger 法提取激光中心线比较精确, 但在受到干扰时会出现多个中心, 如图 10(g)所示, 并且处理时间较长。基于上述分析, 改进提取方法相较于列举的通用提取方法有更好的提取性能。

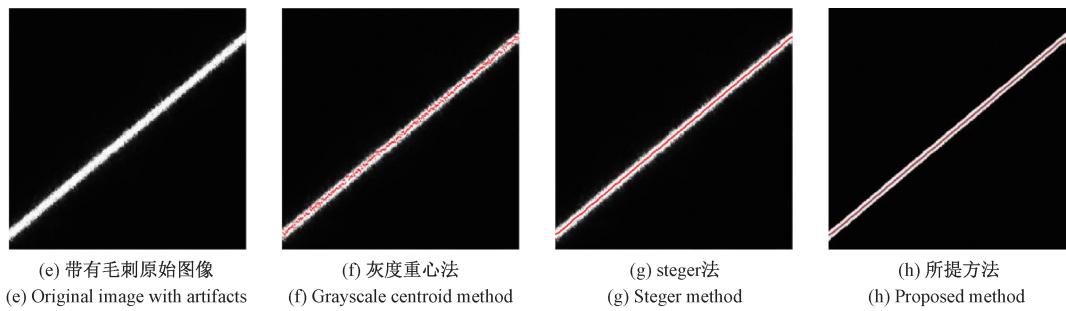


图 10 激光条纹中心提取效果对比图

Fig. 10 Comparison of laser stripe center extraction results

表 5 为线激光中心点提取算法标准差结果。

表 5 不同提取算法的标准差

Table 5 Std of different extraction algorithms px

类型	本文	Steger	细化法	灰度重心法
1	1.28	1.56	5.98	2.97
2	1.29	1.58	6.84	3.11
3	1.21	1.53	5.83	2.93
4	1.30	1.61	6.91	3.21
5	1.27	1.58	6.13	3.06
平均	1.27	1.57	6.34	3.06

由表 5 可知, 所提方法对激光线中心提取的标准差小, Steger 法次之, 相较于 Steger 法平均提升 0.3 个像素, 有较好的抗干扰效果。

## 4 结 论

本文针对表面易反光工件三维测量中线结构光图像出现强反光、噪声干扰, 造成光条中心线提取精度低、速度慢问题, 提出了一种改进的线激光中心提取方法。在 SegFormer 网络中引入 Vgg16 分支提取线激光图像的浅层轮廓信息, 加入 MASPP 模块来提高模型对线形目标的分割效果, mIoU 指标可达 0.907 2, 显著提高了激光条纹的分割精度。通过改进 SegFormer 网络模型将光条图像分割后, 再对分割后的条纹图像利用 Steger 法实现对激光中心线的精确提取。实验结果表明, 该算法计算速度与 Steger 算法相比提升了 42%, 提取的精度提升约 0.3 个像素, 可以在复杂情况下精确实现激光条纹中心线提取。

## 参考文献

- [1] 冀振燕, 宋晓军, 付文杰, 等. 激光光条中心线提取研究综述[J]. 测控技术, 2021, 40(6): 1-8.
- [2] JI ZH Y, SONG X J, FU W J, et al. A review on the extraction of laser stripe centerline[J]. Measurement & Control Technology, 2021, 40(6): 1-8.
- [3] 陈云赛, 杨磊, 宫垠, 等. 基于线结构光的水下目标扫描定位方法[J]. 机器人, 2018, 40(6): 786-793, 802.

CHEN Y S, YANG L, GONG Y, et al. Underwater target scanning and positioning method based on line structured light [J]. Robot, 2018, 40 (6): 786-793, 802.

- [4] 万卓仁, 赖磊捷. 基于自适应窗口的线结构光条中心提取方法[J]. 自动化仪表, 2020, 41(10): 43-46, 50.
- [5] WAN ZH R, LAI L J. A method for extracting the center of line structured light strips based on adaptive window [J]. Automation Instrumentation, 2020, 41(10): 43-46, 50.
- [6] 王志永, 于宇, 王武越, 等. 基于改进 Steger 算法流程的线激光中心提取[J]. 电子测量技术, 2023, 46(1): 84-89.
- [7] WANG ZH Y, YU Y, WANG W Y, et al. Line laser center extraction based on improved Steger algorithm process [J]. Electronic Measurement Technology, 2023, 46(1): 84-89.
- [8] 周祥曼, 王华, 李立军, 等. 基于改进细化法的线激光中心提取方法[J]. 激光杂志, 2023, 44(9): 70-74.
- [9] ZHOU X M, WANG H, LI L J, et al. Line laser center extraction method based on improved refinement algorithm[J]. Laser Journal, 2023, 44(9): 70-74.
- [10] 陈新禹, 孙晓雨, 孙延鹏. 改进 U<sup>2</sup>-Net 的激光条纹中心线高精度提取[J/OL]. 光电子·激光, 2024: 1-8 [2024-08-02]. <http://kns.cnki.net/kcms/detail/12.1182.04.20240321.1609.036.html>.
- [11] CHEN X Y, SUN X Y, SUN Y P. High precision extraction of laser stripe centerline based on improved U<sup>2</sup>-Net[J/OL]. Optoelectronics & Laser, 2024: 1-8 [2024-08-02]. <http://kns.cnki.net/kcms/detail/12.1182.04.20240321.1609.036.html>.
- [12] 张勇停, 李峰. 基于改进 UNet 网络的线结构光条中心提取[J]. 计算机与数字工程, 2023, 51 (9): 2092-2097.
- [13] ZHANG Y T, LI F. Line structured light strip center extraction based on improved UNet network [J]. Computer and Digital Engineering, 2023, 51 (9):

- 2092-2097.
- [8] 王胜春, 韩强, 王昊, 等. 行车环境下钢轨轮廓激光条纹中心的提取方法[J]. 光学学报, 2019, 39(2):175-184. WANG SH CH, HAN Q, WANG H, et al. Extraction method for rail profile laser stripe center under driving environment[J]. Acta Optica Sinica, 2019, 39 (2): 175-184.
- [9] 王子涵, 杨国田, 蓝天翔, 等. 基于线结构光和 YOLOv5 的管道保温层破损检测[J]. 激光与光电子学进展, 2023, 60(6):232-239. WANG Z H, YANG G T, LAN T X, et al. Pipe insulation layer damage detection based on line structured light and YOLOv5[J]. Progress in Laser & Optoelectronics, 2023, 60(6):232-239.
- [10] MING H, XIAO X. A method of laser stripe centerline extraction based on deep learning for structured light 3D reconstruction [J]. Journal of Physics: Conference Series, 2023, 2522(1): 012015.
- [11] 张宇, 黄丹平, 田颖, 等. 复杂物体线结构光中心线提取方法研究[J]. 电子测量技术, 2023, 46(1):167-172. ZHANG Y, HUANG D P, TIAN Y, et al. Research on line structured light centerline extraction method for complex objects [J]. Electronic Measurement Technology, 2023, 46(1): 167-172.
- [12] XIE EN Z, WANG W H, YU ZH D, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.
- [13] 黄星华, 吴天舒, 杨龙玉, 等. 一种面向旋转机械的基于 Transformer 特征提取的域自适应故障诊断[J]. 仪器仪表学报, 2022, 43(11):210-218. HUANG X H, WU T SH, YANG L Y, et al. Domain adaptive fault diagnosis based on Transformer feature extraction for rotating machinery[J]. Journal of Instrumentation and Control, 2022, 43 ( 11 ): 210-218.
- [14] LIU Q H, LIN Y J, HAN X X, et al. Dual-branch feature extraction network combined with Transformer and CNN for polyp segmentation [J]. International Journal of Imaging Systems and Technology, 2023, 34(1):1-16.
- [15] 张昊, 何灵敏, 潘晨. 改进的 SegFormer 遥感图像语义分割网络 [J]. 计算机工程与应用, 2023, 59 (24): 248-258. ZHANG H, HE L M, PAN CH, et al. Improved SegFormer network for remote sensing image semantic segmentation [J]. Computer Engineering and Applications, 2023, 59(24): 248-258.
- [16] HE K M, ZHANG X Y, REN SH 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.
- [17] 杨靖怡, 李芳, 康晓东, 等. 基于 SegFormer 的超声影像图像分割[J]. 计算机科学, 2023, 50(S1):414-419. YANG J Y, LI F, KANG X D, et al. Ultrasound image segmentation based on SegFormer [J]. Computer Science, 2023, 50(S1):414-419.
- [18] 吴玉波, 陈迪来, 杨超, 等. 基于 Steger 算法的多线结构光中心提取[J]. 应用激光, 2023, 43(10):188-195. WU Y B, CHEN D L, YANG CH, et al. Multiline structured light center extraction based on Steger algorithm[J]. Applied Laser, 2023, 43(10):188-195.
- [19] 于营, 王春平, 付强, 等. 语义分割评价指标和评价方法综述[J]. 计算机工程与应用, 2023, 59(6):57-69. YU Y, WANG CH P, FU Q, et al. A review of semantic segmentation evaluation metrics and methods[J]. Computer Engineering and Applications, 2023, 59(6):57-69.

## 作者简介

韩佳鑫, 硕士研究生, 主要研究方向为机器视觉与非接触式三维检测。

E-mail: 592547119@qq.com

王生怀(通信作者), 教授, 博士生导师, 主要研究方向为精密测量技术、表面形貌测量与评定。

E-mail: shwangkb@163.com