

# 基于多任务学习的焊缝图像激光条纹分割与特征点定位方法研究

黄义庚<sup>1,2</sup>, 王大庆<sup>1\*</sup>, 江曼<sup>1</sup>, 殷浩宇<sup>1</sup>, 高理富<sup>1,2</sup>

<sup>1</sup>中国科学院合肥物质科学研究院智能机械研究所, 安徽 合肥 230031;

<sup>2</sup>中国科学技术大学, 安徽 合肥 230026

**摘要** 焊缝信息的快速准确获取是实现自动化焊接的首要问题。然而,在实际焊接过程中,电弧、飞溅、强反射光等噪声会严重污染采集的图像,导致焊缝定位偏移,最终导致跟踪失败。为了提高跟踪过程中的焊缝定位精度与图像处理速度,本文提出了一种将激光条纹分割与焊缝特征点定位相结合的轻量级多任务深度学习模型。该模型由编码器和解码器组成,激光条纹分割子任务与焊缝特征点定位子任务共用编码器主干网络,解码器包含激光条纹分割分支和基于可微空间到数值转换(DSNT)的焊缝特征点定位分支,整个模型遵从轻量化设计思想,同时利用多个子任务之间的相关信息,进一步提升各子任务的性能。实验结果表明,所提模型能够有效克服各类焊接噪声,完成焊缝特征的提取,单幅图像的处理时间约为 11.45 ms,特征点定位精度可达 0.1872 pixel,在自动化焊接方面具有广阔的应用前景。

**关键词** 激光技术; 焊缝跟踪; 轻量级; 多任务; 激光条纹分割; 特征点定位

中图分类号 O436 文献标志码 A

DOI: 10.3788/CJL221057

## 1 引言

随着科技的进步,焊接技术已由人工焊接逐步向自动化焊接、智能化焊接方式转变<sup>[1-2]</sup>,如何准确识别并提取焊缝信息是实现自动化焊接的首要问题。基于激光视觉的焊缝识别方法以高精度、非接触的的优点在工程中被广泛应用。该方法借助主动视觉三维测量技术,将激光束投射到待焊接工件表面,再利用工业相机拍摄含有激光条纹的图像进行处理,得到焊缝的三维信息。

在借助激光主动视觉得到的焊缝图像中,激光条纹的拐点被称为焊缝的特征点。它反映了焊缝的结构,可以表征焊缝的坡口形状、深度、宽度等信息。从焊缝图像中提取出焊缝特征点通常需要经过激光条纹分割、条纹细化、直线检测或曲线拟合等步骤。但在实际焊接过程中识别焊缝时,采集到的原始焊缝图像往往会受到强反射光、焊接飞溅、电弧噪声等的干扰,影响焊缝的识别精度。因此,研究如何从含有大量噪声的图像中准确、快速地提取激光条纹进而获得焊缝信息,对于提高焊接质量具有重要意义<sup>[3]</sup>。

国内外学者针对这一问题提出了许多方法,如: Sun等<sup>[4]</sup>提出了利用灰度值的瞬间守恒定律提取激光条纹中心的方法,该方法能从含有噪声的图像中准确地提取激光条纹的中心;杜健准等<sup>[5]</sup>针对焊接过程易受噪声干扰而导致焊缝位置识别精度不高的问题,提出了一种基于核相关/卡尔曼滤波算法的焊缝路径识别方法;Jia等<sup>[6]</sup>提出了一种带有时空上下文算法的卡尔曼滤波器,该滤波器可以精确地提取焊接过程中激光条纹被电弧和飞溅噪声遮挡时的焊缝特征点。传统方法是针对特定场景设计的,可解释性强,但抗干扰能力较差,对不同工况条件的适应性较差。

随着基于深度学习的特征提取方法的发展,人工智能技术应运而生。Wu等<sup>[7]</sup>设计了一种激光条纹边缘引导网络,在VGG-Net作为主网络的框架下,首先对各阶段特征图引入一对一制导模块,实现含有重噪声焊缝图像的激光条纹的提取;然后利用灰度质心法细化得到激光条纹的中心线;之后针对单道焊缝和多层多道焊缝分别设计了最小二乘法直线拟合和二阶导数非均匀有理B样条函数拟合,并通过拟合分别获得了焊缝特征点。Zou等<sup>[8]</sup>提出一种基于条件对抗网络

收稿日期: 2022-07-15; 修回日期: 2022-08-19; 录用日期: 2022-09-13; 网络首发日期: 2022-09-25

基金项目: 国家自然科学基金重大研究计划重点支持项目(92067205)、中国科学院战略性先导科技专项(A类)(XDA22040303)、安徽省重点研发计划(2022a05020035)、安徽省科技重大专项(202103a05020022)、中国科学院合肥物质科学研究院院长基金青年项目(YZJJ2021QN25)

通信作者: \*dqwang@iim.ac.cn

(CGAN)的焊缝图像修复方法,并运用条件对抗网络对含有噪声的焊缝图像进行修复得到了相应的只含激光条纹的无噪声图像;同时,为了提高焊缝的跟踪精度,他们将优化后的修复网络集成到跟踪器中进行训练样本的恢复。该方法提高了焊缝跟踪系统的抗噪声干扰能力。之后,Zou等<sup>[9]</sup>又提出结合卷积滤波和深度强化学习的两阶段焊缝特征点定位方法,并采用该方法稳健而准确地对每个焊接图像中的焊缝特征点实现了定位。Xu等<sup>[10]</sup>采用一种特征监督跟踪器驱动的生成对抗网络(FT-GAN)来修复受噪声干扰的激光条纹图像,在编码器的特征提取过程中设计了特征监督模块和特征选择模块,并在损失函数中加入了判别相关滤波器网络(DCFnet)来跟踪响应损失,指导面向跟踪的特征恢复。Yang等<sup>[11-12]</sup>设计了一种轻量级多层卷积神经网络,通过生成先验框区域将焊缝坡口边缘的全局检测和定位转化为局部先验框区域的置信度预测和位置坐标值回归,保证了网络的精度和鲁棒性。张世宽等<sup>[13]</sup>基于参数软共享机制设计了一种语义分割与目标检测相结合的深度学习模型,并采用该模型在复杂噪声环境下从焊缝图像中精确地提取出了结构光条纹。

上述基于深度学习的特征提取方法大多是将深度学习算法作为预分割模型分割出激光条纹,进而根据焊缝的类型来设计方法,以提取焊缝特征点。虽然这些特征提取方法的精度相对于传统算法有了很大提升,但存在以下不足:1)设计的网络模型过深并且需要通过后处理提取焊缝特征点,导致图像处理速度慢,难以满足实时性跟踪的要求;2)基于目标检测的架构检测焊缝特征点,没有考虑焊缝特征点位于激光条纹上的信息。针对这些不足,本文提出了一种基于子任务相关性多任务学习模型的激光条纹分割与焊缝特征点提取方法,所构建的模型遵从轻量化设计思想,同时利用多个相关学习任务之间的有用信息提升各子任务的性能,实现了激光条纹的分割和焊缝特征点的定位。

具体来说,本文的主要内容包括:1)对结构光视觉测量技术进行原理介绍,引入特征点像素坐标到空间三维坐标的转化。2)设计了一种轻量级多任务的焊缝特征提取网络模型,通过多级监督使编码器网络学习到激光条纹的特征,同时使解码器网络得到激光条纹的分割结果和焊缝特征点的定位。3)针对焊缝特征点定位方法中高斯热图回归法(heatmap)不可微和全连接层回归法(FC)精度低的缺点,引入可微空间到数值转换(DSNT)模块,实现高斯热图法和全连接层回归法的融合,进而实现基于DSNT的焊缝特征点的高精度定位。4)针对强干扰噪声图像,开展所提方法的激光条纹分割和特征点定位实验,并对实验结果进行分析与讨论。

## 2 结构光视觉测量技术

结构光视觉测量系统由线结构激光发射器、工业

相机、光学滤光片组成。测量的基本原理如图1所示,线激光发射器发射出的线结构激光照射到工件表面,在工件表面形成激光条纹,激光条纹随着焊缝轮廓的变化而变化。工业相机前端安装有特定波长的滤光片,使得采集的图像表现为亮条纹暗背景,以减小图像处理难度。

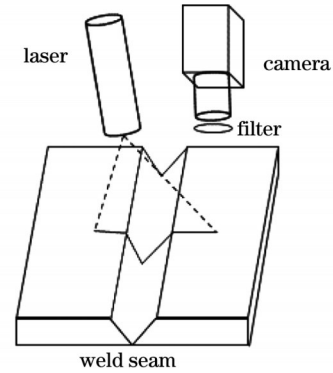


图1 焊缝信息测量原理图

Fig. 1 Schematic of weld information measurement

激光条纹的成像过程如图2所示,其中平面 $XOY$ 和 $O_w-X_wY_wZ_w$ 分别为成像平面和测量平面。将相机的成像模型简化为小孔成像模型,根据相机的成像模型可以得到图像坐标系 $uv$ 下特征点坐标 $P'$ 与相机坐标系下三维坐标 $P$ 的关系为

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \frac{1}{z_c} \mathbf{M}_c \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}, \quad (1)$$

式中: $\mathbf{M}_c$ 为相机的内参矩阵; $(u, v)$ 为特征点的像素坐标; $k_x$ 和 $k_y$ 为 $x$ 轴和 $y$ 轴的耦合放大系数; $u_0$ 和 $v_0$ 为光轴中心的图像坐标; $(x_c, y_c, z_c)$ 是特征点在相机坐标系

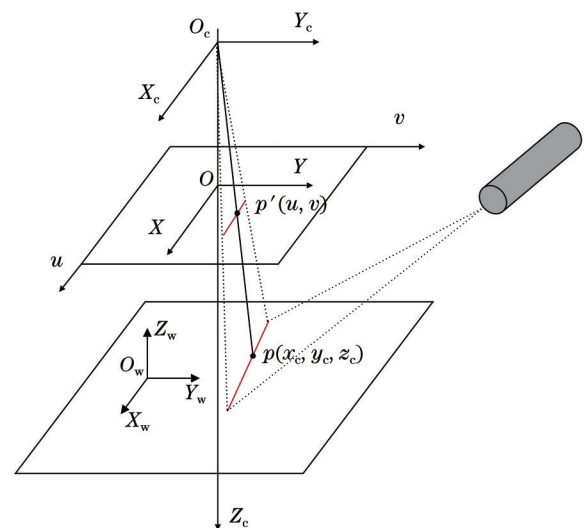


图2 测量投影模型( $O_c-X_cY_cZ_c$ 和 $uv$ 分别为相机坐标系和图像坐标系)

Fig. 2 Projection model measurement ( $O_c-X_cY_cZ_c$  and  $uv$  are camera coordinate system and image coordinate system, respectively)

下的三维坐标。

设结构光平面在相机坐标系下的平面方程为

$$ax + by + cz + 1 = 0. \quad (2)$$

其中内参矩阵  $M_c$  和平面方程的参数可以通过标定实验获取。联立式(1)和式(2)可以得到

$$\begin{cases} z_c = \frac{-k_x k_y}{ak_y(u - u_0) + bk_x(v - v_0) + ck_x k_y} \\ x_c = \frac{z_c(u - u_0)}{k_x} \\ y_c = \frac{z_c(v - v_0)}{k_y} \end{cases}, \quad (3)$$

$$\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = T_6 T_m \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}, \quad (4)$$

式中:  $T_6$  和  $T_m$  分别为机器人末端坐标到基坐标系的旋转矩阵和机器人手眼标定矩阵。

将获得的特征点像素坐标  $(u, v)$  代入式(3)即可得到特征点在相机坐标系下的三维坐标,再代入到式(4)中就可得到机器人基坐标系下的三维坐标。

### 3 基于多任务学习的焊缝特征提取方法

多任务学习利用子任务之间的相关性来提升各子任务的性能<sup>[14]</sup>。深度学习中的多任务学习通常分为参数软共享机制和参数硬共享机制<sup>[15]</sup>。软共享机制下的每个任务都有各自的模型和参数,而硬共享机制共享编码器主干网络的参数,各任务输出层有不同的参数。因此,参数硬共享机制具有更少的参数量。本文网络采用的是参数硬共享机制。本文提出的模型由编码器结构和解码器结构组成,如图 3 所示。子任务共享编码器结构参数,解码器包含激光条纹分割子任务和特征点定位子任务,子任务拥有各自的参数。本文利用编码器网络完成焊缝位置信息和边缘信息的特征提取,利用解码器网络完成激光条纹和焊缝特征点位置的输出。同时,本文也对编码器网络结构进行了多级监督,包括 Stage 3 的边缘细节信息监督,以及空间路径(context info)和上下文路径(spatial info)中的激光条纹输出监督,以便使编码器学习到激光条纹的特征。

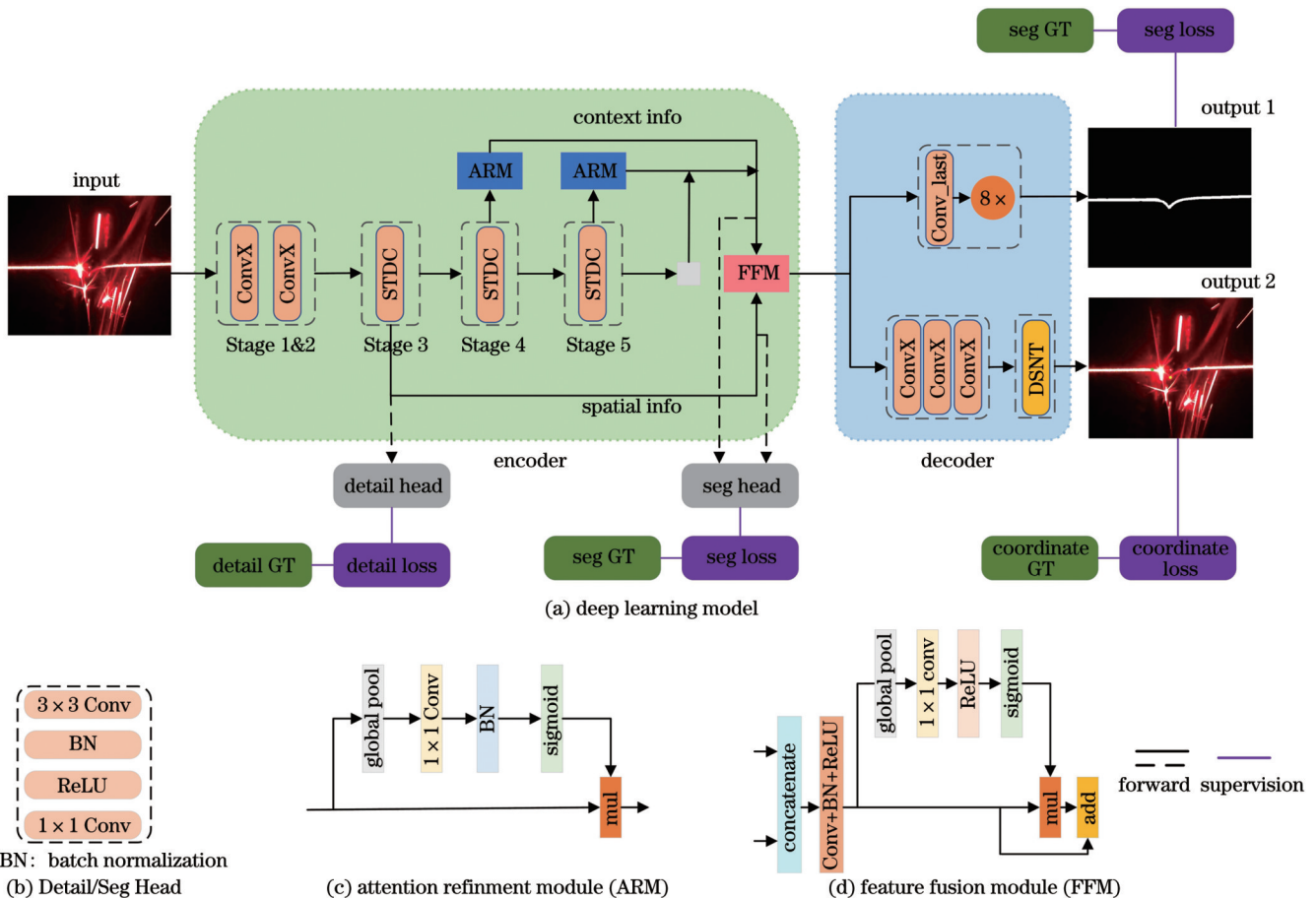


图 3 网络结构。(a)本文提出的网络;(b)~(d)模型中用到的Detail/Seg Head、ARM、FFM模块

Fig. 3 Network structure. (a) Our network; (b)~(d) Detail/Seg Head, ARM, and FFM modules used in the model

#### 3.1 编码器网络

对于基于短期密集连接模块(STDC)<sup>[16]</sup>搭建的双边

分割网络(BiSeNet)<sup>[17-18]</sup>进行改进,提出了新的编码器结构。新提出的编码器在上下文路径中减少了3个



STDC 模块,使得网络更加轻量。

STDC 模块如图 4 所示。在激光条纹分割任务中主要关注的是可伸缩的感受野和多尺度信息,需要低层提供足够的焊缝边缘细节信息和高层语义信息。为了丰富特征信息和减少网络的深度,STDC 模块通过跳跃机构将 Block  $x$  中的特征串联起来作为 STDC 模块的输出。

编码器网络包含上下文路径(context info)和空间路径(spatial info)。上下文路径采用的是 STDC 模块搭建的主干网络,在 Stage 3、Stage 4、Stage 5 阶段分别

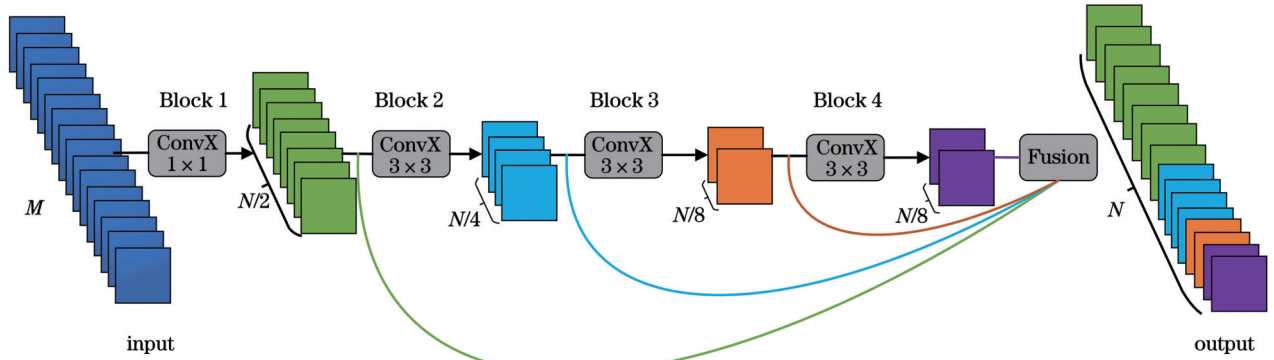


图 4 STDC 模块

Fig. 4 STDC module

### 3.2 解码器网络

解码器网络包含激光条纹分割子任务和特征点定位子任务。

#### 3.2.1 激光条纹分割

激光条纹的分割为二分类问题,最终的输出为 2 个通道,编码器网络的结构输出为 256 个通道的 1/8 尺寸特征图。如图 3(a)所示,编码器经过 Conv\_last 之后再行 8 倍上采样得到最终输出,其中 Conv\_last 的结构与 Seg Head 的结构相同,如图 3(b)所示。

在训练过程中,优化器也对编码器网络的上下文路径和空间路径进行指导,将上下文路径和空间路径提取的特征进行 Conv\_last 卷积后再进行 8 倍上采样,同时利用得到的输出与标签图像计算损失。

#### 3.2.2 特征点定位

目前,关键点的定位方法主要有全连接层回归法<sup>[19]</sup>和高斯热图法<sup>[20]</sup>。其中:全连接层回归法采用全连接层输出关键点坐标,这种做法是可微分的并且推理简单,但是网络会出现过拟合问题,从而丢失了空间泛化能力;高斯热图法对关键点生成高斯热图,将高斯热图作为网络的输出。相比全连接层回归法,高斯热图法的空间泛化能力更强,但它并不是一种端到端可微的模型,无法直接获取坐标值,最终通过在热图上搜索像素值最大的点并返回其位置(Argmax)来获取关键点坐标。如果热图尺寸小于原图尺寸,则会存在理论上误差。DSNT 模块是文献[21]提出的一种可微空间到数值转换的方法,该方法是高斯热图法和全连

下采样生成原图 1/8、1/16、1/32 尺寸的特征图,然后采用全局平均池化来获得具有最大感受野的全局上下文信息。编码器网络中的上下文路径采用 U 形结构对全局特征进行上采样(采样到原图 1/8 尺寸),并将全局特征与编码器的最后两个阶段结合,结合使用的是注意力细化模块(ARM)<sup>[18]</sup>。ARM 的结构如图 3(c)所示。空间路径是编码器网络第三阶段 1/8 尺寸的特征图。最后用图 3(d)所示的特征融合模块(FFM)<sup>[18]</sup>将上下文路径的特征与空间路径的特征融合。

接层回归法的融合,是一种全微分的模型,具有高斯热图法的空间泛化能力。

特征点定位部分的整体结构如表 1 所示。

表 1 特征点定位分支结构

Table 1 Feature point positioning branch structure

Stage	Output	Ksize	Stride	Padding	Channel quantity
ConvX L1	60×70	3	1	1	64
ConvX L2	60×70	3	1	1	32
Conv2d HM	60×70	1	1	0	3
DSNT	3×2				

图 5 为 DSNT 模块的结构示意图,展示了高斯热图向特征点坐标的转换。

Conv2d HM 之后获得的 3 张高斯热图分别对应 3 个焊缝特征点的坐标热图,再经过 Softmax 将坐标值归一化至 0~1 内(整张高斯热图中的所有元素之和为 1),即将每个元素的值转化为所在位置的概率值。该部分操作为

$$H_{ij} = \text{Softmax}(h_{i,j}) = \frac{\exp(h_{i,j})}{\sum_{i=1}^n \sum_{j=1}^m \exp(h_{i,j})}, \quad (5)$$

式中: $m$ 和 $n$ 为热图尺寸。接着根据输出的高斯热图的尺寸构造两个矩阵,该矩阵为常数矩阵,分别记为 $\mathbf{X}$ 和 $\mathbf{Y}$ ,矩阵中元素的计算公式为

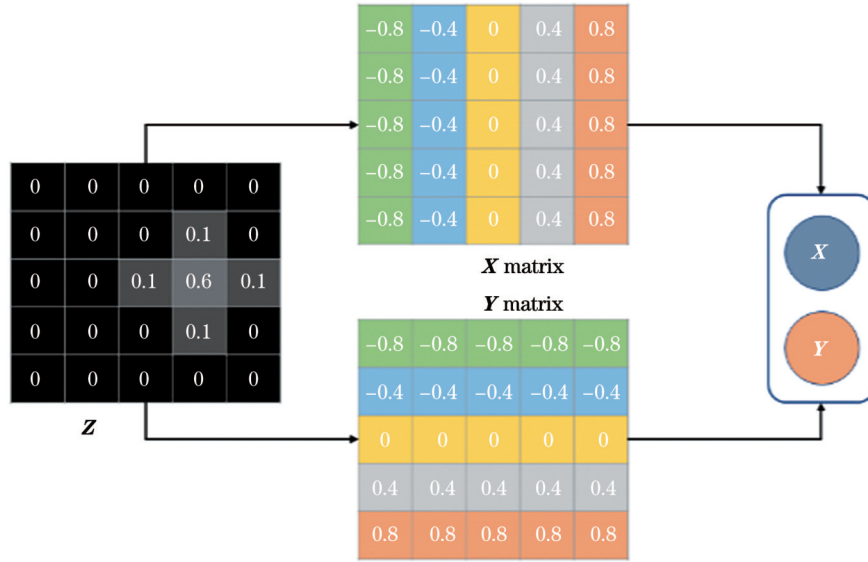


图 5 DSNT 模块  
Fig. 5 DSNT module

$$\begin{cases} X_{ij} = \frac{2j - (n + 1)}{n} \\ Y_{ij} = \frac{2i - (m + 1)}{m} \end{cases} \quad (6)$$

最后根据式(7)进行 Frobenius 内积求和,计算焊缝特征点坐标。

$$(x, y) = \text{DSNT}(H) = [\langle H, X \rangle_F, \langle H, Y \rangle_F] \quad (7)$$

### 3.3 损失函数

在激光条纹的分割任务中,图像中只包含条纹和背景,并且条纹在图像中的占比很小。因此,激光条纹分割属于类不平衡分类问题。在分割任务中,损失函数选用 Ohem 函数,该损失函数能够提高正样本的召回率,缓解类不平衡问题。同时,对空间路径和上下文路径的输出进行监督,使得 FFM 模块的输出是只含有激光条纹的图像。激光条纹分割的损失为

$$L_{\text{seg}} = L_{\text{ohem}}(g, g_{\text{gt}}) + L_{\text{ohem}}(g_{\text{sp}}, g_{\text{gt}}) + L_{\text{ohem}}(g_{\text{co}}, g_{\text{gt}}), \quad (8)$$

式中:  $g_{\text{gt}}$ 、 $g$ 、 $g_{\text{sp}}$ 、 $g_{\text{co}}$  分别为分割图像的真实标签、激光条纹分割预测输出、空间路径输出、上下文路径输出。

图像中含有大量噪声,部分激光条纹会被飞溅噪声掩盖,失去激光条纹的轮廓。为了保留激光条纹的轮廓信息,本文采用细节损失对 Stage 3 的细节轮廓信息进行监督,同时使用二元交叉熵 Bce 和 Dice 损失联合优化细节信息,即

$$L_{\text{detail}} = L_{\text{bce}} + L_{\text{dice}} \quad (9)$$

其中 Bce 损失函数和 Dice 损失函数的表达式分别为

$$L_{\text{bce}}(p_d, g_d) = -(1 - g_d) \lg(1 - p_d) - g_d \lg(p_d), \quad (10)$$

$$L_{\text{dice}}(p_d, g_d) = 1 - \frac{2 \sum_{i=1}^{H \times W} p_d^i g_d^i + 1}{\sum_{i=1}^{H \times W} (p_d^i)^2 + \sum_{i=1}^{H \times W} (g_d^i)^2 + 1}, \quad (11)$$

式中:  $p_d$ 、 $g_d$  分别为细节信息的预测值和真实值;  $H$  和  $W$  分别输出图像的高和宽。

定位部分的损失函数  $L_p$  包含焊缝特征点真实值和模型预测值之间的损失  $L_{\text{euc}}$  以及高斯热图规范化损失  $L_D$ 。特征点定位损失  $L_{\text{euc}}$  为

$$L_{\text{euc}}(c, c_{\text{gt}}) = \|c - c_{\text{gt}}\|_2, \quad (12)$$

式中:  $c$ 、 $c_{\text{gt}}$  分别为特征点的预测坐标和标签坐标,将真实值和预测值的二维欧氏距离作为损失。

高斯热图规范化损失  $L_D$  的表达式为

$$L_D(h, h_{\text{gt}}) = D_{\text{JS}}(h \| h_{\text{gt}}) = \frac{1}{2} D_{\text{KL}}\left(h \left\| \frac{h + h_{\text{gt}}}{2}\right.\right) + \frac{1}{2} D_{\text{KL}}\left(h_{\text{gt}} \left\| \frac{h + h_{\text{gt}}}{2}\right.\right), \quad (13)$$

式中:  $h$ 、 $h_{\text{gt}}$  分别为预测的高斯热图和真实的高斯热图标签;  $D_{\text{JS}}(\cdot \| \cdot)$  和  $D_{\text{KL}}(\cdot \| \cdot)$  分别为 JS 散度和 KL 散度。其中 KL 散度的表达式为

$$D_{\text{KL}}(p \| q) = \sum_{i=1}^N p(x_i) [\lg p(x_i) - \lg q(x_i)]. \quad (14)$$

同时,使用超参数  $\lambda$  对  $L_D$  进行约束,得到特征点定位损失为

$$L_p = L_{\text{euc}} + \lambda L_D. \quad (15)$$

综上,网络的总损失  $L$  为

$$L = L_{\text{seg}} + L_{\text{detail}} + L_p. \quad (16)$$

## 4 实验与结果

### 4.1 实验数据与环境

本文将采集的 U 形钢管焊接过程中的 700 张图像作为图像数据。首先对图像数据进行预处理,设置感兴趣区域,得到分辨率为 480 pixel × 560 pixel 的图像,并通过翻转、平移、旋转等操作将图像由 600 张扩充到 2000 张;然后按照 7:3 的比例选取训练集、验证集,其

中训练集和验证集包含原始图像、激光条纹标签图像、坐标标签。同时,对剩余的 100 张测试集图像也进行标注,用于验证模型的精度。

实验使用的计算机和系统配置为:Windows 11 操作系统,16 GB 内存,NVIDIA GeForce RTX2060 显卡、AMD Ryzen 7 4800H CPU,Pytorch 1.6.0,CUDA10.1,CUDNN7.6.5。

#### 4.2 算法的训练与评估

使用动量为 0.9、权重衰减为  $5e^{-4}$  的小批量随机梯度下降(SGD)方法进行优化训练,batch\_size 设置为 4,使用 ploy 学习速率策略,初始学习速率为 0.01,速率设置为 0.9。在训练集上进行 20000 次迭代。图 6 为模型训练过程中激光条纹分割损失、Stage 3 细节损失、特征点定位损失的变化过程。

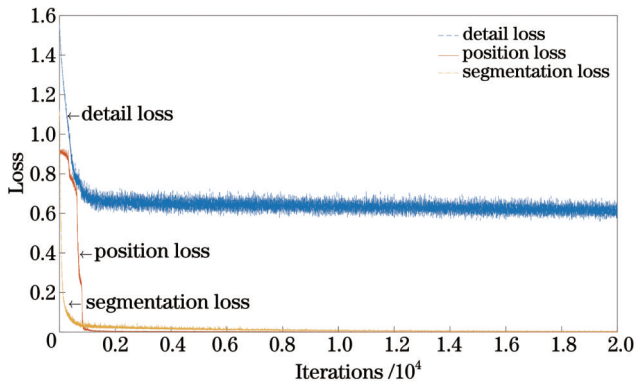


图 6 模型的训练损失

Fig. 6 Training loss of proposed model

本文主要从激光条纹的分割精度和特征点的定位精度出发对算法进行评估,用平均交并比(MIOU)对分割进行评估,用预测坐标与真实坐标之间的位移误

差对定位进行评估(单位是 pixel)。

#### 4.3 激光条纹分割结果与讨论

本节主要对有无使用 3.3 节所述细节信息损失函数的分割结果进行对比。网络在训练过程对编码器的低层(即 Stage 3)进行了细节信息监督。图 7 为细节标签的生成过程,细节标签是使用拉普拉斯卷积核对分割标签图像分别进行步长为 1、2、4 的卷积(目的是生成不同尺寸的特征图),再对步长为 2、4 的特征图进行 2、4 倍上采样(将特征图恢复到原图尺寸),最后通过  $1 \times 1$  的卷积核将三张特征图融合得到的。

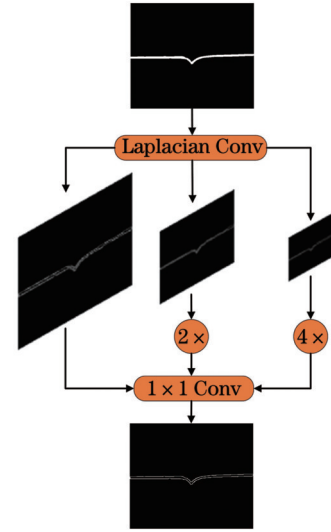


图 7 细节标签生成

Fig. 7 Detail label generation

同时将本文所提算法与文献[22]中使用的全卷积神经网络(FCN)的改进版 FCN-8s 提取的激光条纹进行分割性能对比,对比结果如图 8 所示。

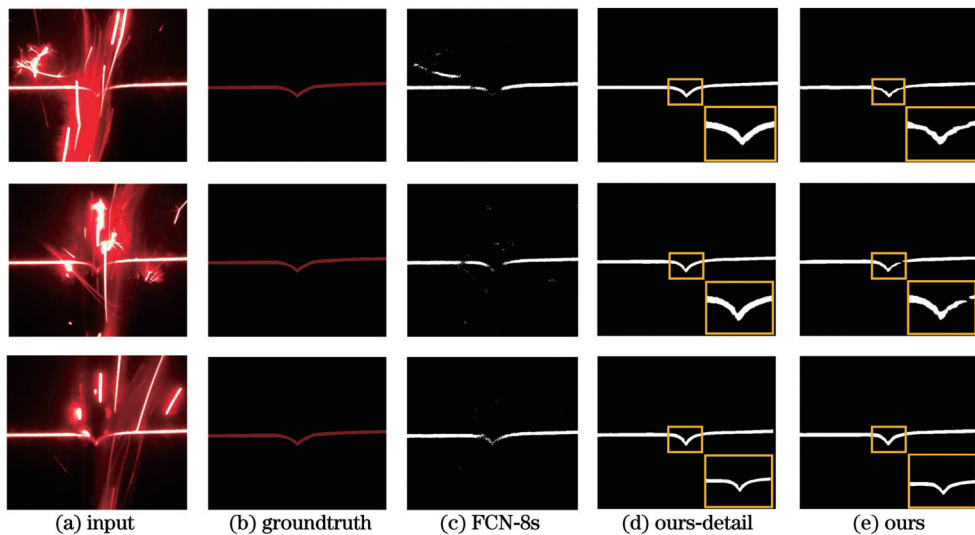


图 8 激光条纹分割结果。(a)原图;(b)激光条纹标签图像;(c)FCN-8s 提取的焊缝特征;(d)本文方法在有细节信息监督下提取的特征;(e)本文方法在没有细节信息监督下提取的特征

Fig. 8 Laser stripe segmentation results. (a) Original images; (b) laser stripe label images; (c) weld seam features extracted by FCN-8s; (d) features extracted by our method with detailed information supervision; (e) features extracted by our method without detailed information supervision



从图 8(d)、(e)可以看出使用细节轮廓信息对低层特征(即 Stage 3)进行监督可以提高分割效果,尤其是在激光条纹被飞溅噪声掩盖的图像中可以很好地恢复出条纹的形状。测试了各网络的分割精度 MIOU 和速度指标 FPS,结果如表 2 所示。由表 2 可以看出,所提网络相对于大部分网络在精度和速度上取得了平衡。

表 2 激光条纹分割精度对比  
Table 2 Comparison of laser stripe segmentation accuracy

Model	Resolution / (pixel×pixel)	MIOU / %	FPS / (frame·s <sup>-1</sup> )
FCN-8s	480×560	89.07	23
BiSeNetV1	480×560	96.67	42
BiSeNetV2	480×560	93.95	103
STDC1-seg	480×560	99.12	72
U-Net	480×560	93.85	16
Ours	480×560	95.97	87
Ours-detail	480×560	98.82	87

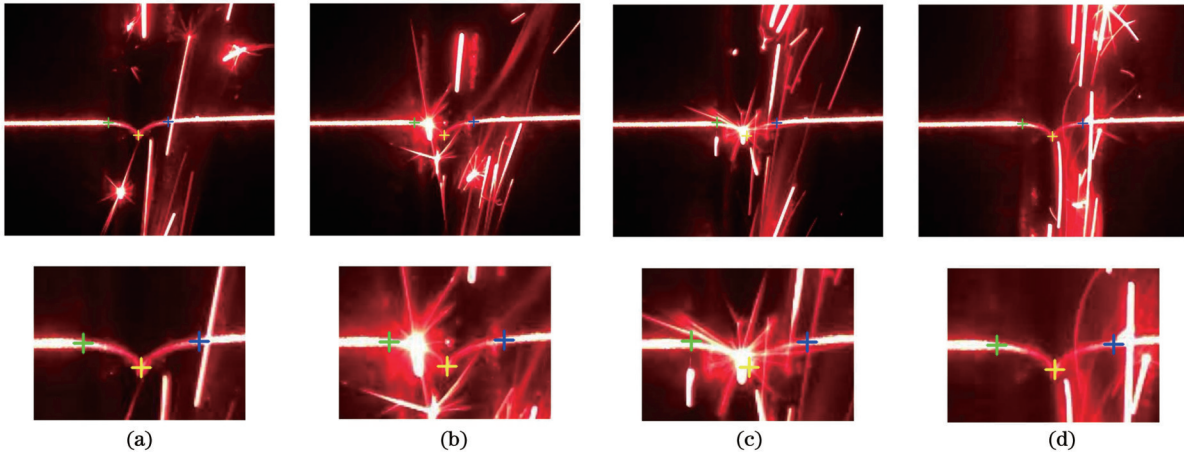


图 9 不同噪声干扰下 DSNT 方法的焊缝特征点定位结果,其中绿色和蓝色“+”为左右特征点,黄色“+”为中间特征点  
Fig. 9 Location results of weld feature points by DSNT method under different noise interferences, where the green and blue “+” are left and right feature points and yellow “+” is intermediate feature point

将中间特征点坐标作为焊缝的特征点坐标。本文使用中间特征点的预测坐标与真实坐标之间的平均位移误差(MDE,在公式中记为 $f_{\text{MDE}}$ )进行评估,MDE 越小表明特征点的定位精度越高。MDE 的计算公式为

$$f_{\text{MDE}} = \sum_{i=1}^N \frac{|u_i^{\text{gt}} - u_i^{\text{p}}| + |v_i^{\text{gt}} - v_i^{\text{p}}|}{2}, \quad (17)$$

式中: $(u_i^{\text{gt}}, v_i^{\text{gt}})$ 和 $(u_i^{\text{p}}, v_i^{\text{p}})$ 分别是特征点的真实坐标和预测坐标; $N$ 为测试图像的数量。使用 100 张测试图像得到 DSNT 方法的 MDE 为 0.1872 pixel。

#### 4.5 子任务相关性

为了研究分割子任务对特征点定位子任务的影响,本文对网络结构进行更改(解码器的输出只含特征点定位,不对编码器的上下文路径和空间路径输出进行激光条纹分割监督,也不对 Stages 3 阶段的激光条

#### 4.4 特征点定位结果与讨论

使用 DSNT 方法定位焊缝特征点,部分定位结果如图 9 所示。本节主要通过改变特征点来定位部分网络层,并将全连接层回归法(FC)、高斯热图回归法(heatmap)、DSNT 法的定位精度进行了对比,图 10 展示了 3 种方法对 3 个特征点定位误差的对比。

高斯热图回归法的输出结构采用的是热图回归法定位中常用的结构,输出特征图尺寸为原图的 1/4,主要是为了避免上采样到原图尺寸时增加网络参数。从图 10 中可以看出高斯热图回归法的定位结果。由于坐标是在高斯热图上通过 Argmax 获得的,无法做到亚像素级,并且热图尺寸为原图的 1/4,最后的像素坐标乘以 4 之后存在理论上限误差问题,因此像素误差相对较大。全连接层回归法网络结构直接用全连接层替代 DSNT 层,全连接层输出特征点坐标。从图 10 所示的特征点定位误差可以看出,全连接层回归法的误差最大。综合比较可知 DSNT 不仅可以实现亚像素级定位,而且定位误差变化最小,从而验证了 DSNT 法在焊缝特征点定位中的可行性。

纹细节信息进行监督),使得整个网络为特征点定位单任务网络。

对焊缝中间特征点定位进行误差分析,结果如图 11 所示,其中图 11(a)为  $u$  轴方向特征点定位误差,图 11(b)为  $v$  轴方向特征点定位误差。由图 11(a)可以看出,对于部分噪声污染严重的图像,不使用激光条纹分割子任务的特征点提取误差比使用多任务模型的误差大。这一结果验证了激光条纹分割子任务对特征点定位任务具有提升作用。

#### 4.6 网络实时性评估

图像的处理时间也是一个重要指标。因此,本节将对 100 张图像进行算法耗时分析。对 100 张图像中的每一张运行 300 次求平均推理时间,得到 100 张图像的前向推理时间。100 张图像的平均处理时间为

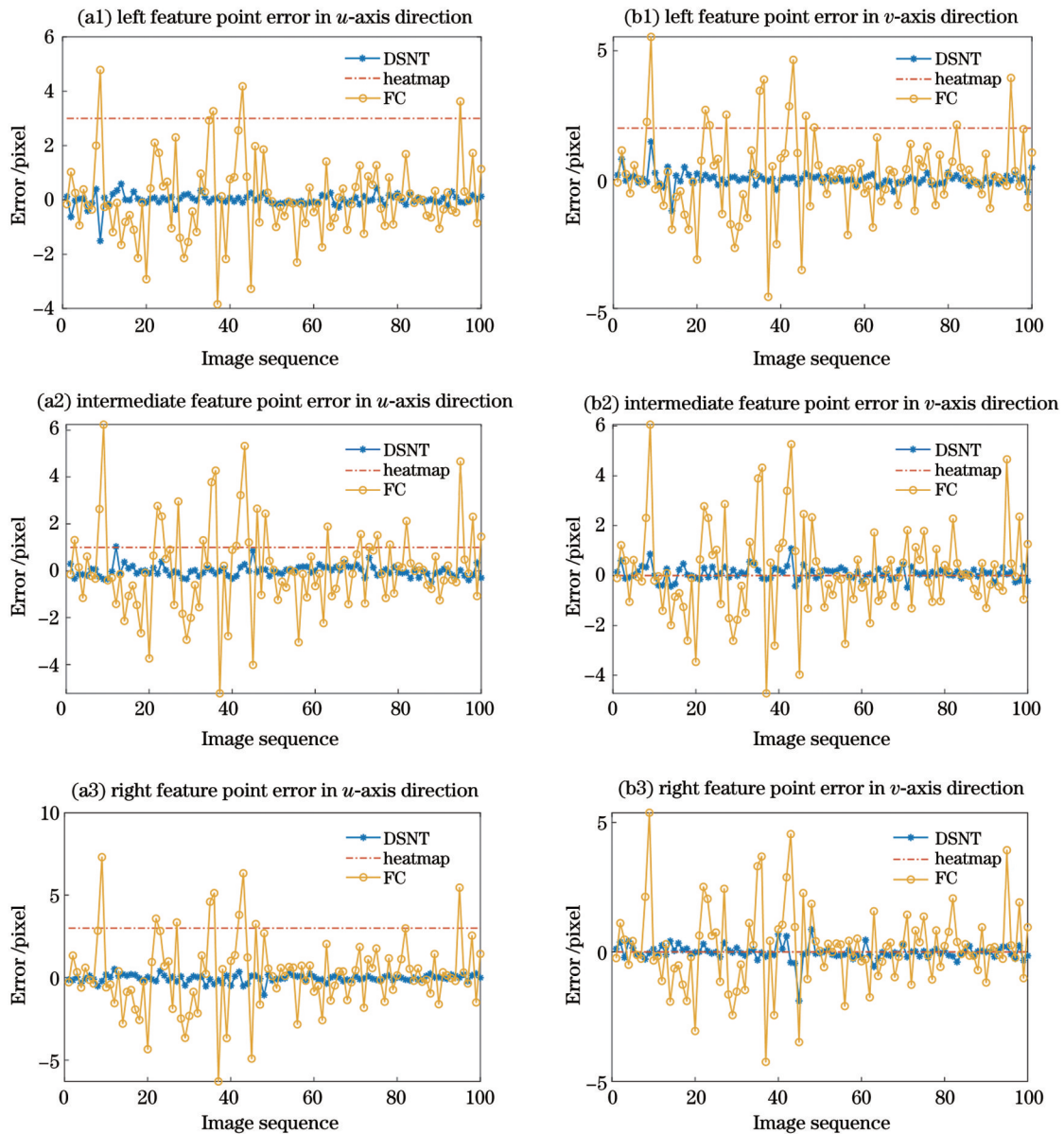


图 10 特征点定位结果对比。(a1)~(a3)焊缝图像中  $u$  轴方向左特征点、中间特征点、右特征点的提取误差；(b1)~(b3)焊缝图像中  $v$  轴方向左特征点、中间特征点、右特征点的提取误差

Fig. 10 Comparison of feature point location results. (a1)–(a3) Extraction errors of left feature point, intermediate feature point, and right feature point in  $u$ -axis direction in weld image, respectively; (b1)–(b3) left feature point, intermediate feature point, and right feature point in  $v$ -axis direction in weld image, respectively

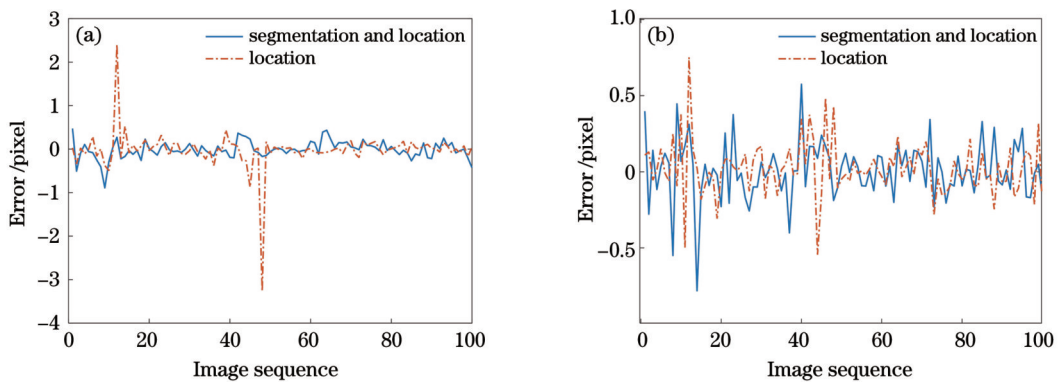


图 11 子任务相关性特征点定位误差对比。(a)  $u$  轴方向；(b)  $v$  轴方向

Fig. 11 Comparison of location errors of subtask correlation feature points. (a) In  $u$ -axis direction; (b) in  $v$ -axis direction



11.4478 ms, 最长推理时间为 11.8872 ms, 最短推理时间为 11.0841 ms, 如图 12 所示。可见, 网络的推理速度可以满足 20 frame/s 的实时性要求。最后, 基于本文的焊缝图像数据集, 将所提模型与其他模型的推理时间进行对比, 对比结果如表 3 所示。表 3 显示, 本文设计的网络模型在推理时间保持优势的情况下可以同时完成激光条纹分割和特征点定位。

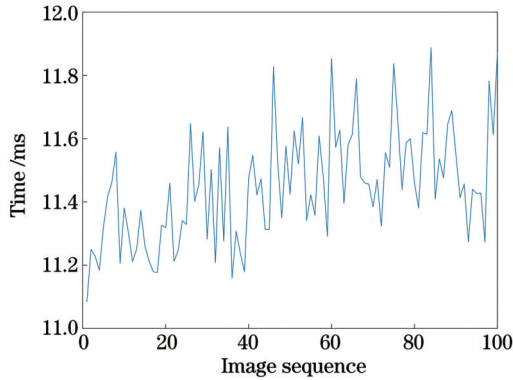


图 12 图像序列的处理时间

Fig. 12 Time consumption in processing for image sequence

表 3 不同网络的推理时间对比

Table 3 Comparison of inference time of different networks

Model	Task		Inference time /ms
	Laser fringe segmentation	Feature point location	
FCN-8s	Yes	No	43.4783
BiSeNetV1	Yes	No	23.6123
BiSeNetV2	Yes	No	9.6583
ICNet	Yes	No	48.2594
Ours	Yes	Yes	11.4478

## 5 结 论

本文针对含有多重噪声的焊缝图像提出了一种激光条纹分割和焊缝特征点定位的多任务学习模型, 在基于实时性语义分割网络 BiSeNet 的基础上搭建了由 STDC 模块组成的编码器, 并在焊缝特征点定位子任务的解码器中使用了 DSNT 模块。DSNT 模块具有全连接层的全微分特性以及高斯热图的空间泛化能力, 并且对低分辨率特征图具有很高的精度, 符合轻量化设计思路。模型训练过程中使用了多级监督。实验结果表明: 1) 使用细节轮廓信息对低层特征进行监督可以提高激光条纹的分割性能; 2) 通过改变特征点定位部分的网络层, DSNT 模块相比高斯热图法和全连接层回归法具有更高的焊缝特征点定位精度; 3) 多任务学习的方式有助于提升焊缝特征点的定位精度; 4) 网络的推理时间可以满足焊缝跟踪对图像处理的实时性要求。

## 参 考 文 献

- [1] Li L Y, Fu L J, Zhou X, et al. Image processing of seam tracking system using laser vision[M]//Tam T J, Chen S B, Zhou C J. Robotic welding, intelligence and automation. Lecture notes in control and information sciences. Heidelberg: Springer, 2007, 362: 319-324.
- [2] Liu S Y, Liu L T, Zhang H, et al. Study of robot seam tracking system with laser vision[C]//2009 International Conference on Mechatronics and Automation, August 9-12, 2009, Changchun, China. New York: IEEE Press, 2009: 1296-1301.
- [3] Ogbemhe J, Mpofu K. Towards achieving a fully intelligent robotic arc welding: a review[J]. Industrial Robot: an International Journal, 2015, 42(5): 475-484.
- [4] Sun Q C, Chen J, Li C J. A robust method to extract a laser stripe centre based on grey level moment[J]. Optics and Lasers in Engineering, 2015, 67: 122-127.
- [5] 杜健准, 张艳喜, 王静静, 等. 核相关/卡尔曼滤波焊缝路径识别新方法研究[J]. 中国激光, 2022, 49(2): 0202009.
- [6] Du J Z, Zhang Y X, Wang J J, et al. Weld path identification based on kernel correlation/Kalman filters[J]. Chinese Journal of Lasers, 2022, 49(2): 0202009.
- [7] Jia Z W, Wang T Q, He J J, et al. Real-time spatial intersecting seam tracking based on laser vision stereo sensor[J]. Measurement, 2020, 149: 106987.
- [8] Wu K X, Wang T Q, He J J, et al. Autonomous seam recognition and feature extraction for multi-pass welding based on laser stripe edge guidance network[J]. The International Journal of Advanced Manufacturing Technology, 2020, 111(9): 2719-2731.
- [9] Zou Y B, Wei X Z, Chen J X. Conditional generative adversarial network-based training image inpainting for laser vision seam tracking[J]. Optics and Lasers in Engineering, 2020, 134: 106140.
- [10] Zou Y B, Chen T, Chen X Z, et al. Robotic seam tracking system combining convolution filter and deep reinforcement learning[J]. Mechanical Systems and Signal Processing, 2022, 165: 108372.
- [11] Xu F J, Zhang H J, Xiao R Q, et al. Autonomous weld seam tracking under strong noise based on feature-supervised tracker-driven generative adversarial network[J]. Journal of Manufacturing Processes, 2022, 74: 151-167.
- [12] Yang G W, Wang Y Z, Zhou N. Detection of weld groove edge based on multilayer convolution neural network[J]. Measurement, 2021, 186: 110129.
- [13] 杨国威, 周楠, 杨敏, 等. 融合卷积神经网络和相关滤波的焊缝自动跟踪[J]. 中国激光, 2021, 48(22): 2202011.
- [14] Yang G W, Zhou N, Yang M, et al. Automatic weld tracking based on convolution neural network and correlation filter[J]. Chinese Journal of Lasers, 2021, 48(22): 2202011.
- [15] 张世宽, 吴清潇, 林智远. 焊缝图像中结构光条纹的检测与分割[J]. 光学学报, 2021, 41(5): 0515002.
- [16] Zhang S K, Wu Q X, Lin Z Y. Detection and segmentation of structured light stripe in weld image[J]. Acta Optica Sinica, 2021, 41(5): 0515002.
- [17] Chennupati S, Sistu G, Yogamani S, et al. MultiNet++: multi-stream feature aggregation and geometric loss strategy for multi-task learning[EB/OL]. (2019-04-15) [2022-03-06]. <https://arxiv.org/abs/1904.08492>.
- [18] Ruder S. An overview of multi-task learning in deep neural networks[EB/OL]. (2017-06-15) [2022-06-03]. <https://arxiv.org/abs/1706.05098>.
- [19] Fan M Y, Lai S Q, Huang J S, et al. Rethinking BiSeNet for real-time semantic segmentation[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 9711-9720.
- [20] Yu C Q, Gao C X, Wang J B, et al. BiSeNet V2: bilateral

- network with guided aggregation for real-time semantic segmentation[EB/OL]. (2020-04-05)[2022-03-05]. <https://arxiv.org/abs/2004.02147>.
- [18] Yu C Q, Wang J B, Peng C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[EB/OL]. (2018-08-02)[2022-03-06]. <https://arxiv.org/abs/1808.00897>.
- [19] Wei S E, Ramakrishna V, Kanade T, et al. Convolutional pose machines[C]//29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4724-4732.
- [20] Toshev A, Szegedy C. DeepPose: human pose estimation via deep neural networks[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, Columbus, OH, USA. New York: IEEE Press, 2014: 1653-1660.
- [21] Nibali A, He Z, Morgan S, et al. Numerical coordinate regression with convolutional neural networks[EB/OL]. (2018-01-23)[2022-03-06]. <https://arxiv.org/abs/1801.07372v1>.
- [22] 张永帅, 杨国威, 王琦琦, 等. 基于全卷积神经网络的焊缝特征提取[J]. 中国激光, 2019, 46(3): 0302002.
- Zhang Y S, Yang G W, Wang Q Q, et al. Weld feature extraction based on fully convolutional networks[J]. Chinese Journal of Lasers, 2019, 46(3): 0302002.

## Laser Fringe Segmentation and Feature Points Location Method of Weld Image Based on Multi-Task Learning

Huang Yigeng<sup>1,2</sup>, Wang Daqing<sup>1\*</sup>, Jiang Man<sup>1</sup>, Yin Haoyu<sup>1</sup>, Gao Lifu<sup>1,2</sup>

<sup>1</sup>*Institute of Intelligent Machines, Hefei Institute of Physical Science, Chinese Academy of Science, Hefei 230031, Anhui, China;*

<sup>2</sup>*University of Science and Technology of China, Hefei 230026, Anhui, China*

### Abstract

**Objective** With advancements in science and technology, welding technology has progressed from manual to automated and intelligent welding. The widely used weld tracking technology based on laser vision can improve the ability of welding robots to perceive their environments, with the added advantages of non-contact and high precision. However, in real-time weld tracking, the collected weld images are often severely affected by strongly reflected light, splash, and arc noise. Therefore, laser stripes are accurately and quickly extracted from images containing a large amount of noise, and then obtaining weld information from them is a prerequisite for high-quality welding. To improve the weld location accuracy and the image processing speed in the weld tracking process, this paper proposes a lightweight multi-task deep learning model that combines laser strip segmentation and weld feature point location. The model consists of an encoder and a decoder. The laser fringe segmentation subtask and the weld feature point location subtask share the encoder backbone network. The decoder includes a laser fringe segmentation branch and a weld feature point location branch based on differentiable space-to-numerical transformation (DSNT). The entire model is designed in a lightweight manner, and it simultaneously adopts relevant information between multiple subtasks to further improve the performance of each subtask. In summary, we expect that the designed deep learning model can achieve accurate and rapid acquisition of weld features during the welding process.

**Methods** In order to improve the weld location accuracy and image processing speed in the weld tracking process, a lightweight multi-task deep learning model combining laser strip segmentation and weld feature point location is proposed. The proposed model adopts the parameter hard sharing mechanism in multi-task learning such that the model uses fewer parameters. Specifically, the model consists of an encoder and a decoder. The encoder completes the feature extraction of weld position and edge information, while the decoder implements the output of the laser stripe segmentation and feature point location subtasks. The encoder network adopts the concept of a more efficient bilateral segmentation network, including context and spatial paths. The context path realizes the extraction of high-level semantic features of the image, and the spatial path provides edge detail information. In addition, to make up for the loss of detailed information, the spatial path is supervised with detailed information. To utilize the information that the weld feature point is located on the laser fringe, multi-stage supervision is adopted to make the encoder structure learn the characteristics of laser fringes. Therefore, the structure of the laser fringe segmentation subtask in the decoder only contains a simple convolutional layer and an upper sampling layer, which can realize the output of laser fringes. The DSNT module is introduced into the feature point location subtask to realize the fusion of the Gaussian thermal diagram method and the fully connected layer method so that the model is completely differentiated and has the spatial generalization ability of the Gaussian heat map method.

**Results and Discussions** The results of laser fringe segmentation on images disturbed by noise demonstrate that our model exhibits good segmentation accuracy, and the detail information supervision of low layers can further improve the segmentation accuracy (Fig. 8); in addition, our model achieves a good balance between accuracy and speed (Table 2). The location results of the weld feature points show that DSNT can accurately locate the feature points of the weld under different noise interference conditions (Fig. 9). Through an experiment where the output layer structure of the network was changed, we verified that compared with the Gaussian thermal diagram method and fully connected layer regression method, the DSNT method can achieve subpixel-level location with minimal errors (Fig. 10). By changing the output structure of the decoder, it is experimentally verified that the laser stripe segmentation subtask can improve the performance of the weld feature point location subtask (Fig. 11). Finally, the experimental

results verified that, compared with various deep learning models, the proposed network model can complete the segmentation of laser stripes and the localization of feature points while maintaining the inference time (Table 3).

**Conclusions** In this study, a multi-task learning model for laser fringe segmentation and weld feature point location is proposed for weld images with multiple noises. Using detailed contour information to supervise the characteristics of the lower layer can improve the segmentation performance of laser fringes. By changing the network layer of the feature point location part, the DSNT module exhibits a higher weld feature point location accuracy than the Gaussian heat map method and the fully connected layer regression method. The multi-task learning method improves the accuracy of the location of weld feature points. In addition, the inference time of the proposed network can meet the real-time requirements of image processing in weld tracking. In summary, our model can effectively handle all types of welding noise and complete the feature extraction of welds, demonstrating good application prospects in automated welding.

**Key words** laser technology; weld tracking; lightweight; multi-task; laser fringe segmentation; feature point location