

# 基于语义信息增强的化纤丝线网络度检测方法

郑广智<sup>1</sup> 彭添强<sup>1</sup> 肖计春<sup>2</sup> 吴高昌<sup>1</sup> 李智<sup>1</sup> 柴天佑<sup>1,3</sup>

**摘要** 网络度是衡量化纤丝线及化纤织物性能的重要指标之一,在生产车间中通常采用人工方式进行检测.为解决人工检测误检率较高的问题,提出一种基于语义信息增强的化纤丝线网络度并行检测方法.首先,为提升单根化纤丝线网络结点识别的准确度,使用基于 MobileNetV2 优化的主干网络结构提取语义信息,以提高模型的运算速度.在所提主干网络的基础上,设计语义信息增强模块和多级特征扩张模块处理主干网络的特征信息,同时,设计像素级注意力掩膜对特征信息进行加权和融合,以提高网络度检测的准确性.然后,为实现多根化纤丝线网络度的批量计算,基于所提语义信息增强算法,设计网络度并行检测方法.使用算法检测丝线网络结点,同时使用连通域分析及掩膜提取的方法并行检测,提取视野内每条丝线的独立区域.随后,将并行检测结果融合,以准确获取每根丝线的网络度检测结果.为验证所提方法的有效性,使用自主研发的网络度检测设备建立了化纤丝线数据集,并进行了实验验证.结果表明,所提出的方法能够有效地提高检测的准确性.

**关键词** 网络度, 图像语义分割, 语义特征信息, 并行检测策略, 注意力机制

**引用格式** 郑广智, 彭添强, 肖计春, 吴高昌, 李智, 柴天佑. 基于语义信息增强的化纤丝线网络度检测方法. 自动化学报, 2024, 50(10): 1963–1976

**DOI** 10.16383/j.aas.c230649 **CSTR** 32138.14.j.aas.c230649

## A Detection Method for the Interlacing Degree of Filament Yarn Based on Semantic Information Enhancement

ZHENG Guang-Zhi<sup>1</sup> PENG Tian-Qiang<sup>1</sup> XIAO Ji-Chun<sup>2</sup> WU Gao-Chang<sup>1</sup> LI Zhi<sup>1</sup> CHAI Tian-You<sup>1,3</sup>

**Abstract** The interlacing degree serves as an important indicator for evaluating the performance of filament yarns and fabrics, typically detected manually in production workshop. To address the issues of high false detection rates in manual inspection, a parallel detection method for filament yarn interlacing degree based on semantic information enhancement is proposed. Firstly, to improve the recognition accuracy of interlacing nodes in a filament yarn, an improved backbone architecture based on MobileNetV2 is used for semantic information extraction to improve the computational speed of model. Building upon the proposed backbone architecture, semantic information enhancement module and multilevel feature dilated module are designed to process the feature information of the backbone architecture. Meanwhile, a pixel-level attention mask is designed to weight and fuse the feature, in order to improve the accuracy of interlacing degree detection. Then, based on the proposed enhancement algorithm for semantic information, a parallel detection method of interlacing degree is designed to achieve batch calculation for interlacing degree of multiple filament yarns. The algorithm is used to detect interlacing node, while connected domain analysis and mask extraction are used for parallel detection to extract independent regions of each filament yarn within the field. The parallel detection results are then fused to accurately obtain the interlacing degree detection results for each filament yarn. To validate the effectiveness of the proposed method, a synthetic filament yarn dataset is established using a self-developed interlacing degree detection device, and experimental verification is conducted. The results demonstrate that the proposed method can effectively improve the accuracy of detection.

**Key words** Interlacing degree, image semantic segmentation, semantic feature information, parallel detection strategy, attention mechanism

**Citation** Zheng Guang-Zhi, Peng Tian-Qiang, Xiao Ji-Chun, Wu Gao-Chang, Li Zhi, Chai Tian-You. A detection method for the interlacing degree of filament yarn based on semantic information enhancement. *Acta Automatica Sinica*, 2024, 50(10): 1963–1976

收稿日期 2023-10-24 录用日期 2024-04-03

Manuscript received October 24, 2023; accepted April 3, 2024

国家自然科学基金 (62173077, 61991404, 62103092) 资助

Supported by National Natural Science Foundation of China (62173077, 61991404, 62103092)

本文责任编辑 陈胜勇

Recommended by Associate Editor CHEN Sheng-Yong

1. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819 2. 东

北大学机械工程与自动化学院 沈阳 110819 3. 国家冶金自动化工程技术研究中心 沈阳 110819

1. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819 2. School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819 3. National Engineering Research Center of Metallurgy Automation, Shenyang 110819

我国化纤产业作为国民经济发展的关键支柱产业之一,近年来逐渐明确了高质量发展的需求,即推动化纤制造向中高端迈进,实现成为化纤强国的目标<sup>[1]</sup>.其中,化纤丝线作为一种重要的合成纤维产品,在纺织品制作等领域因其耐磨、高强度和高弹性等优点,得到广泛应用<sup>[2]</sup>.在化纤丝线的质量检测过程中,网络度作为质量评估的关键指标之一,因其具有直观且易于观察与统计的特点,在许多企业中广泛应用.网络度是指单位长度的化纤丝线中,具有一定牢度的未散开的网络结数量.以锦纶6型全拉伸丝为例,丝线在水面上散开后呈现的形态如图1所示.对于正常的化纤丝线,其网络结数需要在规定的范围内.当网络结数偏多时,使丝线在松弛加工过程中无法充分解捻,同时在染色时上染率不同,导致面料表面出现斑点.相反,当网络结数偏少时,使丝线在织造过程中,网络内部存在张力,导致丝线网络结点松散,进而产生毛羽.总之,化纤丝线的网络度作为重要的质量指标,对于化纤产品的制造工艺和织物的最终品质都有着不可忽视的影响,同时对网络度的准确评估具有重要意义.

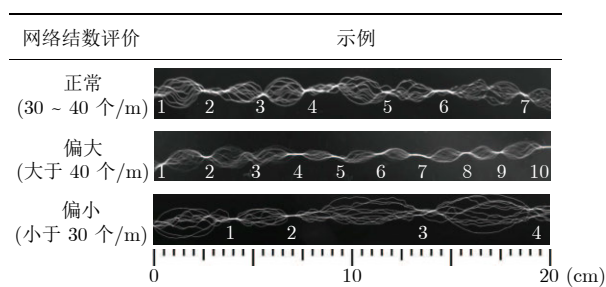


图1 化纤丝线网络度示例

Fig.1 Interlacing degree diagram of filament yarns

在工业领域,常见的网络度检测方法包括手工移针法、仪器移针法、手工重锤法和水浴法等<sup>[3]</sup>.目前常用的网络度检测设备主要基于仪器移针法研发,但是使用基于仪器移针法的检测设备时,每次只能检测单根丝线,难以适用于大规模检测.因此,许多化纤生产企业选择水浴法作为网络度检测的方法,将丝线放入盛有纯净水的水槽中,待丝线接触水面散开成网络状后,统计丝线网络结的数量.然而,此方法也存在一些局限性.首先,人眼目测依赖于主观评价,易受光照等外界因素的影响,使网络度检测结果不稳定;其次,在产量较大的情况下,人工检测的能力有限,难以同时对多根丝线进行检测.因此需要设计更为准确与高效的化纤丝线网络度检测方法.

与传统工业技术相比,基于计算机视觉与深度

学习的检测技术具有准确率高和抗干扰能力强等诸多优势.通过研发深度学习算法与自动化检测设备,实现对化纤丝线网络度的准确检测和量化分析,从而减少人为主观判断带来的不确定性.此外,深度学习算法能够自动学习并提取复杂特征,高效处理大规模数据,进而提升质量分析的效率和精确性.这为化纤行业提供了一种先进的技术手段,提升了化纤产品的质量控制效果和后道质检结果的准确性.目前,基于目标检测的深度学习算法已广泛应用于丝线断裂检测<sup>[4]</sup>、毛羽检测<sup>[5-6]</sup>和丝线直径均匀度<sup>[7]</sup>等方面.例如,基于注意力机制改进的CenterNet模型<sup>[8]</sup>在丝线毛羽检测方面表现出良好的性能,通过全局阈值化方法处理样本图像,可有效分割出丝线主干位置.然而,化纤丝线网络结点尺寸较小,使用上述基于目标检测的方法容易导致丝线结点位置信息关注不足,从而影响检测结果的准确性.

为提高化纤丝线网络度检测结果的准确性,本文将网络度检测问题视为像素级的二分类问题,即将丝线的网络结点位置视为感兴趣区域,并将样本图像中其他像素区域归属于背景区域,最后按照像素进行分类.目前针对此类问题,主要采用编码器-解码器结构进行特征提取.其中,编码器用于提取图像的抽象特征,解码器用于恢复特征至原始图像尺寸并细化特征.双线性插值<sup>[9-10]</sup>和跳跃连接<sup>[11]</sup>有助于在编码和解码过程中保留细节和边缘信息,多层空洞卷积<sup>[12-14]</sup>有助于在解码过程中增加图像的感受野.然而,对于丝线网络结点这种小目标而言,使用大量的卷积层和较大的空洞卷积率容易导致语义信息丢失.为增强对语义信息的提取能力及检测性能,注意力机制<sup>[15-16]</sup>和Transformer<sup>[17-19]</sup>等方法相继提出.其中,通过多尺度融合策略<sup>[20]</sup>和残差特征金字塔<sup>[21]</sup>等方式使得模型专注于特定的感兴趣区域,金字塔特征融合<sup>[22-24]</sup>和损失注意力机制<sup>[25]</sup>等方法在特征图融合和损失函数优化方面取得了显著的成果,提高了对语义信息的提取能力.然而,考虑其对计算资源的较高要求,对小样本数据集的适用性有限,同时对局部特征信息处理不足,训练难度较大,故难以适用于本文的网络度检测问题.

基于以上分析,本文提出一种基于语义信息增强的化纤丝线网络度检测方法,提高从化纤丝线图像中提取语义信息的能力,实现化纤丝线网络度的批量计算,同时解决人工检测误检率高的问题.本文的主要贡献如下:

1) 针对化纤丝线网络度结点目标较小,导致检测准确性低的问题,提出语义信息增强模块,使得模型更加关注丝线网络结点区域的信息.同时,设

计多级特征扩张模块,使用不同大小的扩张卷积核提取深层特征的语义信息,能够增强高质量的空间细节信息的提取能力,提高检测结果的准确性.

2) 针对丝线样本数量大且检测耗时长的问題,提出网络度并行检测方法,将化纤丝线图像中的丝线主干提取与网络结点提取两部分并行处理,实现网络度的批量计算.

3) 针对人工检测误检率高的问题,研制网络度检测设备,结合网络度检测算法与掩膜提取方法,实现网络度自动化视觉检测.通过使用所研制的设备进行丝线样本采集并建立数据集,对所提检测模型进行实验验证.

1 基于语义信息增强的网络度检测算法

本文提出一种基于语义信息增强的“编码-解码”两阶段语义分割检测算法,模型结构如图 2 所示.在编码阶段,首先基于 MobileNetV2<sup>[26]</sup> 进行改进,用于提取输入图像的特征信息,并获得包含局部细节信息的浅层特征图和全局语义信息的深层特征图.然后,使用交叉注意力模块和全局注意力模块增强浅层特征图中语义信息的提取能力.同时使用扩

张卷积方法提取深层特征图中的多级语义信息.最后,将得到的浅层特征信息与深层特征信息通过解码阶段的阶段性特征融合模块进行加权融合,并将特征图像尺寸恢复至输入图像大小,得到检测结果.

1.1 主干网络

为保证所提网络度检测算法的可靠性,本文基于 MobileNetV2 模型设计主干网络结构.在这一模型中,使用深度可分离卷积<sup>[27]</sup> 的方法进行卷积操作.对于传统卷积操作而言,这种方法有效地减少了模型的参数数量和计算量,从而在保持高质量特征提取的同时,显著地降低了模型的复杂度.深度可分离卷积将普通的卷积过程分解为深度卷积和逐点卷积两部分,其中深度卷积用于获得输入特征图中包含的空间信息,逐点卷积用于将提取的特征进行整合与变换.使用这种卷积方式能够减少模型的运算量,提高模型的运算效率.经下采样操作后,减小了特征图的尺寸,同时增加了特征图的通道数.这一操作通过扩大网络的感受野,使得模型能够提取更为抽象的特征.在本文中,下采样的层数位置设定为 2 层、4 层、8 层和 10 层,其网络架构如表 1 所示.

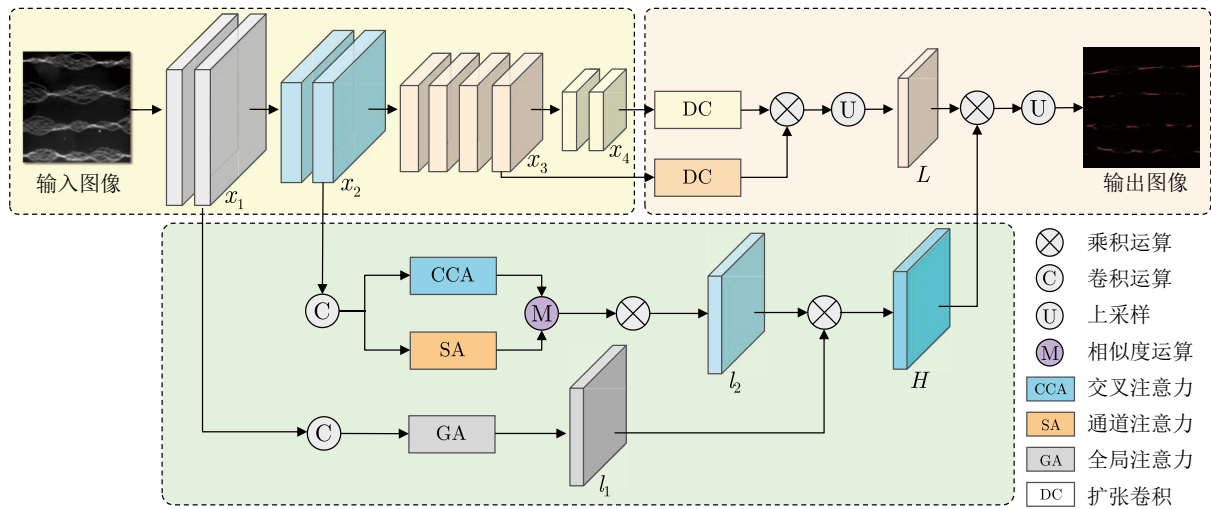


图 2 网络度检测算法结构

Fig.2 Architecture of interlacing degree detection algorithm

表 1 主干网络架构

Table 1 Architecture of the backbone network

特征尺寸 (像素)	扩展因子	循环次数	输出通道数	步长
512 × 512 × 3	—	1	32	2
256 × 256 × 32	1	2	32	1
128 × 128 × 32	6	4	64	2
64 × 64 × 64	6	2	96	2
32 × 32 × 96	—	—	—	—

## 1.2 基于语义信息增强模块

在主干网络中, 浅层特征包含原始输入图像中丰富的局部细节和位置信息. 为更加有效地从浅层特征中提取到原始输入图像的语义信息, 即提取像素点及其周围像素之间的关系信息, 本文提出一种基于语义信息增强模块, 对主干网络中的浅层特征图  $x_1$  和  $x_2$  进行处理, 模型结构如图 3 所示.

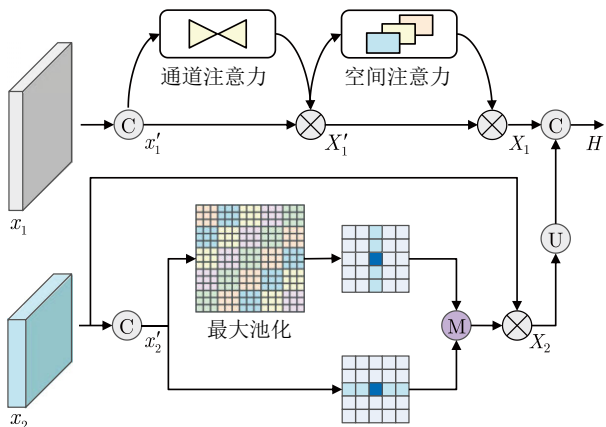


图 3 基于语义信息增强模块

Fig. 3 Enhancement module based on semantic information

对于特征图  $x_1$ , 由于其像素相对较大, 若多次进行特征提取, 将导致模型计算量增加, 影响模型的处理速度. 为确保输出特征图的维度与输入特征图一致, 同时避免增加模型的运算量. 本文首先对特征图  $x_1$  使用  $1 \times 1$  卷积核进行卷积操作, 将维度降至原始特征图的一半, 得到特征图  $x_1'$ , 然后采用全局注意力机制的方法进行特征提取, 计算方法为

$$X_1 = \sigma(f_{\text{con}}(f_{\text{avg}}(X_1'), f_{\text{glo}}(X_1'))) \times X_1' \quad (1)$$

$$X_1' = \sigma(f_{\text{avg}}(x_1') + f_{\text{glo}}(x_1')) \times x_1' \quad (2)$$

其中,  $\sigma(\cdot)$  表示 sigmoid 归一化函数,  $f_{\text{con}}(\cdot)$  表示特征图融合卷积操作,  $f_{\text{avg}}(\cdot)$  表示全局平均池化操作,  $f_{\text{glo}}(\cdot)$  表示全局最大池化操作,  $X_1'$  表示经通道注意力机制处理后的特征图. 对于特征图  $x_2$ , 首先使用  $1 \times 1$  卷积核进行卷积操作, 将维度降至原始特征图的一半, 得到特征图  $x_2'$ . 然后使用全局最大池化方法提取空间注意力信息, 将所得特征图与原始特征图  $x_2$  进行相似性判断, 通过归一化的方式获取像素注意力权重. 最后将权重加权赋值到原始图像上, 从像素和通道两种角度增强特征的语义信息, 得到增强后的特征图  $X_2$ . 为获得两个特征图中共同关注的感兴趣区域, 本文通过对比两个特征图中对应像素位置的相似度, 使模型更好地理解不同尺度下

的语义信息. 此外, 通过对不同尺度的特征图进行加权融合, 有助于提高语义分割的精度和鲁棒性. 在像素相似性判断中, 常用方法包括点积、余弦相似度和欧氏距离等<sup>[28-29]</sup>. 综合考虑计算复杂度与数据量, 本文使用自注意力机制中的像素点积方式进行处理, 得到特征图  $X_2$ , 并将其进行上采样至特征图  $X_1$  的尺寸后进行特征融合, 得到特征图  $H$ . 计算方法为

$$X_2 = x_2 \times \sigma((f_{\text{glo}}(x_2'))^T \times x_2') \quad (3)$$

$$H = f_{\text{con}}(X_1, f_{\text{up}}(X_2)) \quad (4)$$

其中,  $f_{\text{up}}(\cdot)$  表示上采样操作,  $X_1$  表示特征图  $x_1$  经全局注意力机制模块处理后得到的特征图,  $X_2$  表示特征图  $x_2$  经本文提出的自注意力机制模块处理后得到的特征图.

## 1.3 多级特征扩张模块

在主干网络中, 深层特征中包含更为抽象和丰富的语义信息, 对于图像分析和理解至关重要, 但是使用过于频繁的特征提取可能导致特征信息的丢失.

在 DeepLab 系列的网络中, 通过采用空洞卷积模块提取深层语义信息, 能够得到多尺度信息. 然而使用单一的空洞卷积结构进行特征提取时, 可能导致不同特征之间的空间相关性不足, 使提取到的有效信息减少, 从而增加模型的参数量等<sup>[30]</sup>. 因此, 本文提出一种多级特征扩张模块, 旨在增强深层特征图中的语义信息提取能力, 如图 4 所示.

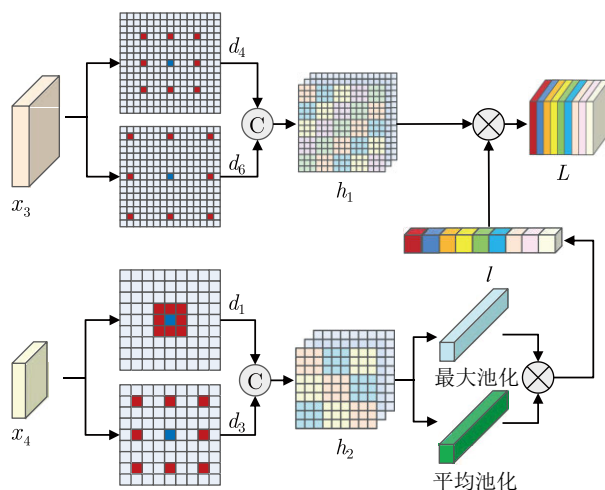


图 4 多级特征扩张模块

Fig. 4 Dilated module of multilevel features

对于从主干网络中提取的特征图  $x_4$  与上一级特征图  $x_3$ , 按照上述分析进行扩张卷积操作, 即对卷积核的相邻两个权值之间进行扩张, 计算方法为



$$d_k(m, n) = \sum_{p=1}^M \sum_{q=1}^N f(m+k \times p, n+k \times q) \cdot w(p, q) \quad (5)$$

其中,  $d_k(m, n)$  表示像素点  $(m, n)$  经扩张率为  $k$  的卷积核扩张后的特征图,  $w(p, q)$  表示卷积核,  $M$  和  $N$  分别表示卷积核的长度与宽度。

采用式 (5) 所示的方法, 使用扩张率为 4 和 6 的卷积核分别对特征图  $x_3$  进行上述操作, 经特征融合后得到特征图  $h_1$ 。同时, 使用扩张率为 1 和 3 的卷积核分别对特征图  $x_4$  进行卷积, 经特征融合后得到特征图  $h_2$ 。由于使用不同级别的特征图进行扩张卷积后, 生成的特征图的尺寸不同, 无法直接进行融合, 因此需要对特征图  $h_2$  进行上采样操作。目前传统的上采样方法主要包括最近邻插值、双线性插值和上池化等, 但是这些方法只是简单地复制或插值原始像素值, 忽略了对图像语义信息的保留以及对图像语义信息一致性的考虑, 容易导致特征信息的丢失。为了克服这一问题, 本文使用卷积注意力机制模块对特征图  $h_2$  进行处理。首先对特征图每个通道进行全局平均池化和全局最大池化操作, 得到缩放后的全局特征向量, 然后采用反转置卷积的方式将特征图放大至与特征图  $x_3$  相同的尺寸, 最后将上采样后的特征图与缩放后的全局特征向量逐元素相乘, 得到加权后的特征图  $L$ , 方法如下:

$$L = f_{\text{con}}(C^T \times l, h_1) \quad (6)$$

$$l = f_{\text{con}}(f_{\text{max}}(h_2), f_{\text{avg}}(h_2)) \quad (7)$$

其中,  $f_{\text{max}}(\cdot)$  表示全局最大池化操作,  $C^T$  表示转置卷积核矩阵,  $h_1$  和  $h_2$  表示特征图  $x_3$  和  $x_4$  经扩张卷积模块处理并融合后得到的特征图。由于全局特征向量涵盖整体信息, 因此使用所提方法有助于保持全局特征信息的一致性。通过与上采样后的特征图进行相似性判断, 使得模型更加专注于提取全局特征信息, 从而增强了模型的特征信息提取能力。

#### 1.4 阶段性特征融合模块

通过使用本文提出的基于语义信息增强模块和多级特征扩张卷积模块分别对主干网络的浅层特征图和深层特征图进行处理, 然后将得到的特征图进行融合, 构成模型的解码器部分。然而, 直接将语义特征信息通过上采样至输入图像的分辨率大小时, 容易丢失原特征图的语义信息。因此, 本文设计了阶段性特征融合模块, 如图 5 所示。该模块的输入分为两大部分, 包括使用多级特征扩张模块后得到的特征图  $L$  和使用基于语义信息增强模块处理后

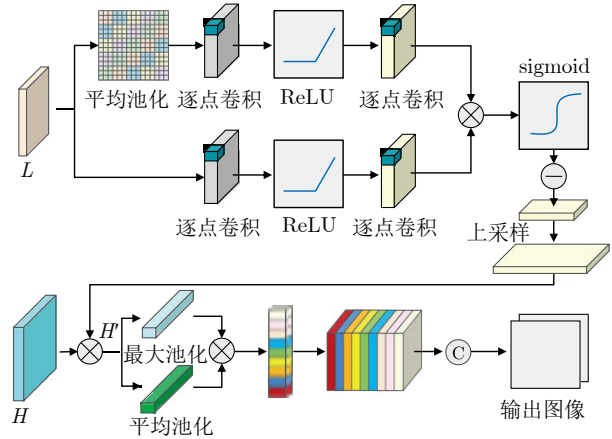


图 5 阶段性特征融合模块

Fig.5 Fusion module of staged feature

得到的特征图  $H$ 。首先, 采用两个不同尺度的分支提取特征  $L$  的通道注意力权重, 旨在从多个尺度上提取通道注意力信息。其中, 第 1 个分支采用全局平均池化的方法提取全局特征的注意力信息; 第 2 个分支则采用逐点卷积的方式提取局部特征的通道注意力信息。随后, 将这两个分支提取到的特征进行融合, 经过归一化操作后, 得到最终的特征权重值。这一方法旨在有效地结合多尺度的语义特征信息, 以实现更精细的特征融合和更准确的解码结果。为调节浅层特征图与深层特征图融合的权重, 使用像素级注意力掩膜进行权重分配。经过相应的注意力掩膜相乘后, 将这两种不同尺度的语义特征图对应的像素进行融合, 得到输出结果。计算过程为

$$H' = \alpha \times H + f_{\text{up}}((1 - \alpha) \times L) \quad (8)$$

其中,  $H'$  表示输出特征图,  $\alpha$  表示特征权重。将两特征图融合以后, 采用上采样方法将特征图恢复到原始图像的大小, 再经过  $1 \times 1$  卷积核进行卷积处理, 最后对特征图按照类别数量进行分类, 得到输出结果。

## 2 网络度并行检测方法

使用基于语义信息增强的网络度检测算法检测单根丝线的网络度时, 能够有效提取丝线网络结点的位置区域。然而, 如果图像中丝线样本数量较大, 则难以区分视野内每根丝线的网络度是否符合标准。因此, 针对多根化纤丝线的网络度的批量计算, 以及独立提取每根丝线的网络度检测结果的需求, 本文提出一种网络度并行检测方法。此方法在原有检测算法基础上, 增加一条并行检测支路。其中, 主支路用于检测视野内所有丝线网络结点的位置, 而并行检测支路用于提取每根丝线的位置。此并行检测支

路主要通过提取样本图像中丝线主干的边缘轮廓, 获得丝线的掩膜图像, 最后得到每根丝线主干的位置信息. 当图像同时经过两条支路处理后, 将所得结果进行融合, 从而得到每根丝线上的网络结点图像.

## 2.1 丝线掩膜提取算法

通过提取丝线掩膜, 定位化纤丝线图像中每一条丝线主干区域, 能够从含有多根丝线的图像中分离出每根丝线所在的位置, 检测过程如图 6 所示. 具体方法如下:

1) 获取丝线样本图像. 从化纤丝线生产车间获取待检测的丝线样本后, 取一段单位长度的丝线放置于盛有纯净水的水槽中, 使丝线在水槽中自然散开并形成较好的形态, 然后使用工业相机拍摄包含多根并行的待测丝线样本的图像. 为配合检测算法模型中对输入图像的分辨率需求, 设置图像分辨率为  $2\,048 \times 512$  像素.

2) 提取丝线主干. 使用工业相机拍摄图像的分辨率较高, 导致图像处理时间较长. 为减小模型处理时间, 同时确保获取到完整的化纤丝线轮廓区域, 在图像处理前将样本图像缩小至原尺寸的一半. 在实际检测过程中, 容易有小颗粒灰尘落入水槽中, 导致采样图像出现毛刺和杂点. 因此, 本文使用均值滤波算法处理图像. 使用最大类间方差算法对图像进行阈值分割, 通过遍历所有可能的阈值寻找最佳的分割结果. 最后对丝线所在区域与图像背景区域进行处理, 得到包含多根丝线掩膜区域的丝线主干图像.

3) 提取单根丝线掩膜. 由于丝线主干区域与背

景区域的对比度较为显著, 因此本文选择 Canny 边缘检测算子对上一步处理得到的丝线主干图像进行处理, 从而得到视野中每根丝线的边缘拟合图像. 然而, 由于边缘检测方法对图像中的噪音和复杂纹理较为敏感, 若目标边缘不清晰可能导致丝线主干边缘不完整, 进而影响丝线区域的完整性. 因此, 本文借助边缘提取方法, 以图像中每根化纤丝线轮廓区域的中心点为基准, 以丝线轮廓长度值作为掩膜图像的长度, 以轮廓最大宽度为掩膜的图像宽度, 得到轮廓尺寸信息. 然后根据所得尺寸信息, 将丝线拟合轮廓绘制在与输入图像相同尺寸的空白图像上, 得到最终的丝线掩膜图像.

4) 提取单根丝线图像. 由于丝线掩膜图像为二值图像, 即每个像素点由 0 或 1 组成. 因此, 将单根丝线掩膜与丝线样本图像的每个像素进行乘积处理, 得到单根丝线的图像.

## 2.2 网络度并行检测过程

在实际生产环境中, 使用工业相机采集到的含多根丝线图像进行网络度检测, 方法如图 7 所示. 首先使用基于语义信息增强的网络度检测算法检测经预处理后的丝线样本图像的结点区域, 同时使用丝线掩膜提取算法检测单根丝线的位置区域. 其次, 将得到的两个特征图进行对应像素乘积, 得到多张含单根丝线且已标注网络结点位置的图像. 然后根据前述轮廓提取算法得到的数据, 绘制结点特征的外接矩形, 并根据矩形尺寸与标准值比较, 将长度小于 2 mm 的结点视为可忽略的结点, 筛选之后的结点即为该丝线所有的网络结点. 最后统计丝线上

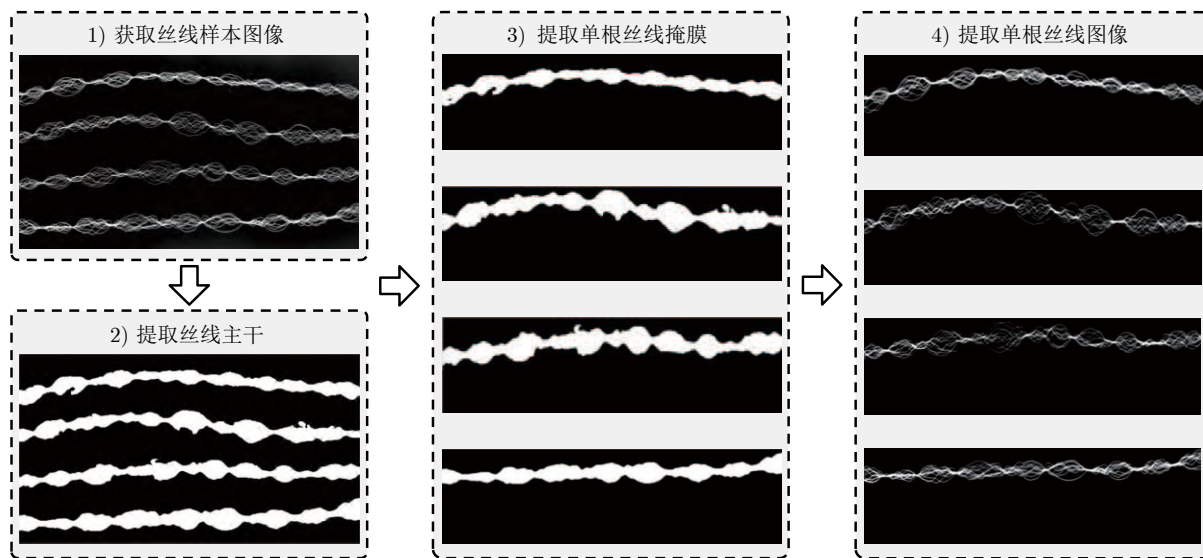


图 6 单根丝线图像提取过程

Fig. 6 Process of extracting single filament yarn image

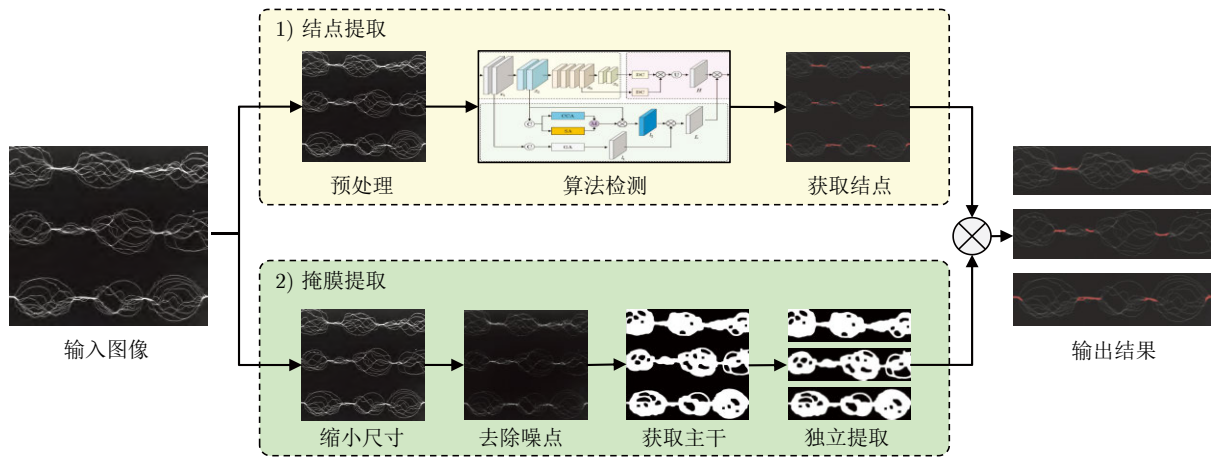


图7 多根丝线网络度并行检测方法

Fig.7 Parallel detection method for interlacing degree of the multiple filament yarns

所有网络结点数,并转换为网络度指标,计算方法为

$$k_i = \frac{m}{d} \quad (9)$$

其中,  $k_i$  表示单根丝线在单位长度内的网络度指标,单位为(个/m);  $i$  表示图像中丝线的位置序号( $i = 1, 2, \dots$ );  $m$  表示丝线中含有的网络结点数;  $d$  表示视野范围内待测丝线对应的长度,单位为 m.

### 3 实验结果与分析

#### 3.1 网络度检测设备

为实现化纤丝线网络度自动视觉检测,解决人工检测导致结果误检率高的问题,本文研制了网络度检测设备.该设备主要包括 CME 标准型闭环步进电机、M3ST502M-H 型工业相机、3030 型铝型材支撑架、亚克力水槽、丝线夹持装置、LED 灯源和滑轨,如图 8 所示.使用网络度检测设备检测化纤丝线网络度的流程如图 9 所示.在检测开始之前,需要对设备进行初始化操作,包括控制参数设置,以及使用上位机控制两个丝线夹持装置移动至同一侧.在丝线网络度检测过程中,首先将丝线挂载至夹取装置的固定孔中.待丝线夹紧后,由上位机控制滑轨转动,拖动水平滑轨上的夹爪至水槽的另一侧.然后控制垂直滑轨,将丝线浸入水中.最后使用工业相机采集丝线在水中的图像,建立丝线样本图像数据集.

与人工检测方法相比,使用本文提出的检测设备能够实现精准统计视野范围内每根丝线的网络度,降低因人工主观因素导致的检测结果偏差.同时,人工检测最多只能同时抓取 4 根丝线进行检测,而使用本文所提装置进行检测时,通过合理设计丝线夹持装置的固定孔数量及滑轨长度,可实现多达 10 根

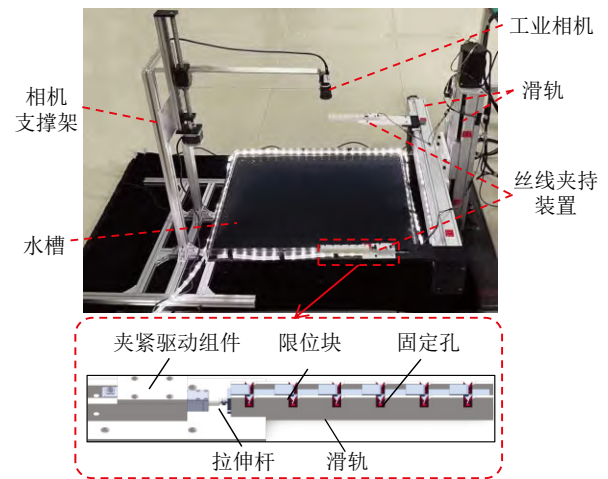


图8 多根丝线并行网络度检测设备

Fig.8 Parallel detection equipment for interlacing degree of multiple filament yarns

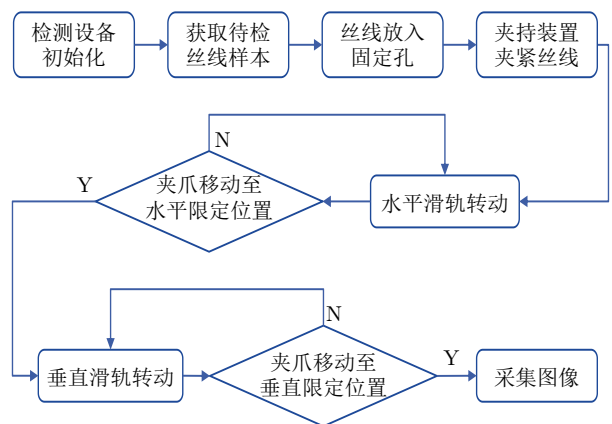


图9 网络度检测设备操作流程

Fig.9 Operation process of interlacing degree detection equipment



以上丝线的同时检测,从而提高丝线的检测效率。

### 3.2 数据集构建与实验配置

本文采用化纤丝线生产工厂提供的锦纶 6 型丝线进行样本采集,并制作数据集。经图像增强和裁剪等预处理操作后,得到 1 200 张  $512 \times 512$  像素大小的样本图像。训练集与测试集的比例设置为 9 : 1, 然后进行模型训练与验证。实验中使用的配置信息如表 2 和表 3 所示。

表 2 模型训练环境配置

Table 2 Configuration of model training environment

项目	版本参数
操作系统	Ubuntu 18.04.6 LTS
CUDA	cuda 11.3
GPU	NVIDIA RTX 3 090
训练框架	PyTorch 1.10.2
内存	128 GB
编程语言	Python 3.8

表 3 模型训练超参数配置

Table 3 Configuration of model training hyperparameter

参数	配置信息
输入图像尺寸	$512 \times 512$ 像素
下采样倍数	16
初始学习率	$5 \times 10^{-3}$
最小学习率	$5 \times 10^{-5}$
优化器	Adam
权值衰减	$5 \times 10^{-4}$
批量大小	12

### 3.3 评价指标

本文采用平均交并比、 $F_1$  分数和每秒传输帧数三个指标对训练模型进行评价,具体介绍如下:

1) 平均交并比: 用于计算真实值和预测值两个集合的交集和并集之比,计算方法为

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \sum_{j=0}^k \frac{p_{ii}}{p_{ij} + p_{ii}} \quad (10)$$

其中,  $MIoU$  表示平均交并比值,  $k$  表示类别数,  $i$  表示真实值,  $j$  表示预测值,  $p_{ij}$  表示将  $i$  预测为  $j$  的像素点数量。

2)  $F_1$  分数: 用于平衡模型的精确率和召回率,衡量模型在正负样本上的分割结果,计算方法为

$$F_1 = 2 \times \frac{pre \times rec}{pre + rec} \quad (11)$$

其中,  $pre$  表示模型的精确率,  $rec$  表示模型的召回率。

3) 每秒传输帧数: 用于评估模型在每秒内可以处理的图片数量或者处理一张图片所需时间,进而评估模型的检测速度。计算方法为

$$FPS = \frac{1}{t} \quad (12)$$

其中,  $FPS$  表示每秒传输帧数,  $t$  表示模型处理一张图片时消耗的时间。

### 3.4 损失函数

本文使用图像语义分割损失函数中评估样本相似性的度量函数 Dice 损失和交叉熵损失函数作为损失函数,具体介绍如下:

1) Dice 损失: 用于评估真实值  $T$  与预测值  $P$  的相似性,计算方法为

$$L_{dice} = 1 - \frac{2|T \cap P|}{|T| + |P|} \quad (13)$$

其中,  $L_{dice}$  表示样本相似值,  $|T \cap P|$  表示真实标签  $T$  和预测标签  $P$  之间的交集元素个数,  $|T|$  和  $|P|$  分别表示真实标签和预测标签中元素的个数。

2) 交叉熵损失函数: 用于衡量每个像素点的预测类别与真实类别之间的差异,计算方法为

$$L_{ce} = -\frac{1}{N} \sum_{i=1}^s [t_i \lg p_i + (1 - t_i) \lg(1 - p_i)] \quad (14)$$

其中,  $L_{ce}$  表示交叉熵损失值,  $N$  表示一个批次中含有的特征图数量,  $s$  表示图像中像素点的数量,  $t_i$  表示第  $i$  个像素点的真实标签,  $p_i$  表示第  $i$  个像素点的预测概率。

### 3.5 性能比较

为验证所提方法在化纤丝线网络度检测方面的有效性,本文使用语义分割领域常用模型进行对比实验,包括 PSPNet<sup>[9]</sup>, UNet<sup>[10]</sup>, DeepLabV3+<sup>[11]</sup>, BiSeNet<sup>[31]</sup>, CGNet<sup>[32]</sup>, HRNet<sup>[33]</sup> 和 SegFormer<sup>[34]</sup>, 并计算其平均交并比、 $F_1$  分数和每秒传输帧数三个评定指标,并比较训练模型权重的参数,比较结果如表 4 所示。由表 4 可以看出,本文提出的方法在平均交并比、 $F_1$  分数和每秒传输帧数三个指标均优于其他几种方法,特别是在模型的每秒传输帧数方面,本文方法的提升更加显著。此外,在模型参数量方面,所提方法相对于对比实验中的大部分模型具有优势。其中, BiSeNet 和 CGNet 使用相对简单的轻量级结构,模型运算时的处理速度较高。DeepLabV3+、PSPNet 和 SegFormer 中使用到运算较为复杂的多尺度特征融合模块与特征金字塔模块,使得模型权重参数量较大。本文提出的网络度



表 4 不同方法的评价指标比较  
Table 4 Comparison of evaluation indicators for different methods

方法	平均交并比 (%)	$F_1$ 分数 (%)	每秒传输帧数 (帧/s)	参数量 (MB)
BiSeNet	78.95	86.76	63.84	48.93
CGNet	79.00	86.79	33.17	<b>2.08</b>
DeepLabV3+	79.50	87.35	43.11	209.70
HRNet	78.74	86.69	12.43	37.53
PSPNet	73.58	82.46	49.80	178.51
SegFormer	79.04	86.84	40.87	14.34
UNet	79.83	87.63	22.54	94.07
本文方法	<b>81.52</b>	<b>88.12</b>	<b>76.16</b>	7.98

注: 加粗字体表示各列最优结果。

检测算法使用语义信息增强的方式对特征图进行处理, 同时使用交叉注意力的方法对特征图进行加权赋值, 增强了对特征信息的提取能力。为直观展示本文所提模型的识别效果, 根据丝线在水中呈现的样态频率, 选择了四种不同类型的化纤丝线的样本图像, 如图 10 所示。使用不同的模型检测这四种化纤丝线图像, 实验结果如图 11 ~ 14 所示。

图 11 所示为丝线散开大于 20 mm 时的图像检测结果, 此类图像中丝线与背景区域具有明显的对比度, 使得网络结点的位置容易区分。由实验结果可以看出, 多数模型能够识别出网络结点的位置。然而, 经过 BiSeNet、CGNet、HRNet 和 SegFormer 四个模型检测后得到的结果图中, 矩形框标记位置处出现了像素预测区域不连续的情况。此外, 在 PSPNet 模型的检测结果图中, 圆形框标记位置处出现了漏检。

图 12 所示为丝线散开小于 10 mm 时的图像检测结果, 其中误检结点区域均已在图中标注。在此类图像中, 网络结点区域和丝线主干区域的对比度较低, 而且存在许多易混淆的结点, 容易导致漏检。SegFormer 模型中使用到了自注意力机制, 从

结果可以看出识别效果相对较好。PSPNet 模型中使用金字塔池化提取语义信息, 然而在检测结果中, 容易将网络结点附近的背景区域误判为感兴趣区域。

图 13 所示为含倾斜丝线的图像检测结果。实验结果表明, 当图像存在一定倾斜角度时, 能够准确识别明显的网络结点像素区域。针对图中标注矩形框的位置区域, 在 BiSeNet 和 CGNet 的识别结果中, 部分结点像素未完全识别。当识别图中框选出来的不明显区域时, 采用本文提出的语义信息增强方法能够很好地识别出结点区域图像。而 HRNet、SegFormer 和 DeepLabV3+ 模型虽能识别出结点信息, 但是识别区域不完整, 且有断续现象。

图 14 所示为图像中含不明显结点的图像检测结果, 如图中矩形框标注位置所示。实验结果表明, 视野中存在不同尺寸大小的网络结点时, 容易在模型识别上存在混淆, 但网络结点位置的整体识别情况较好。针对图中第三根丝线, 使用 BiSeNet、CGNet、DeepLabV3+、HRNet 和 SegFormer 模型进行检测时均忽略了小网络结点, 使用 PSPNet 识别像素不完整, 体现出本文模型在细节信息提取上效果较好。

综上所述, 当图像中存在不同网络散开程度的丝线, 以及丝线倾斜与不明显结点等异常情况时, 与常用的语义分割模型相比, 使用本文提出的模型能够较好地获得网络结点区域, 且准确率较高。

### 3.6 消融实验

#### 3.6.1 模块有效性验证

本实验针对提出的模块进行有效性验证, 设计了 8 种实验方案, 分别是: 1) 仅使用主干网络; 2) 使用特征图  $x_4$  和所提语义信息增强模块直接融合; 3) 使用特征图  $x_2$  和所提多级特征扩张模块直接融合; 4) 使用特征图  $x_2$ ,  $x_4$  和所提阶段性特征融合模块; 5) 使用特征图  $x_2$ 、所提多级特征扩张和阶段性

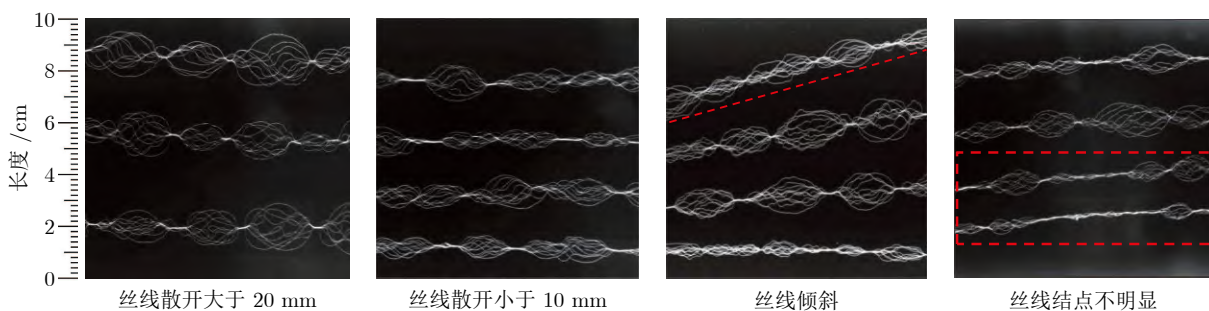


图 10 四种丝线样本图像

Fig.10 Sample images for four types of filament yarns

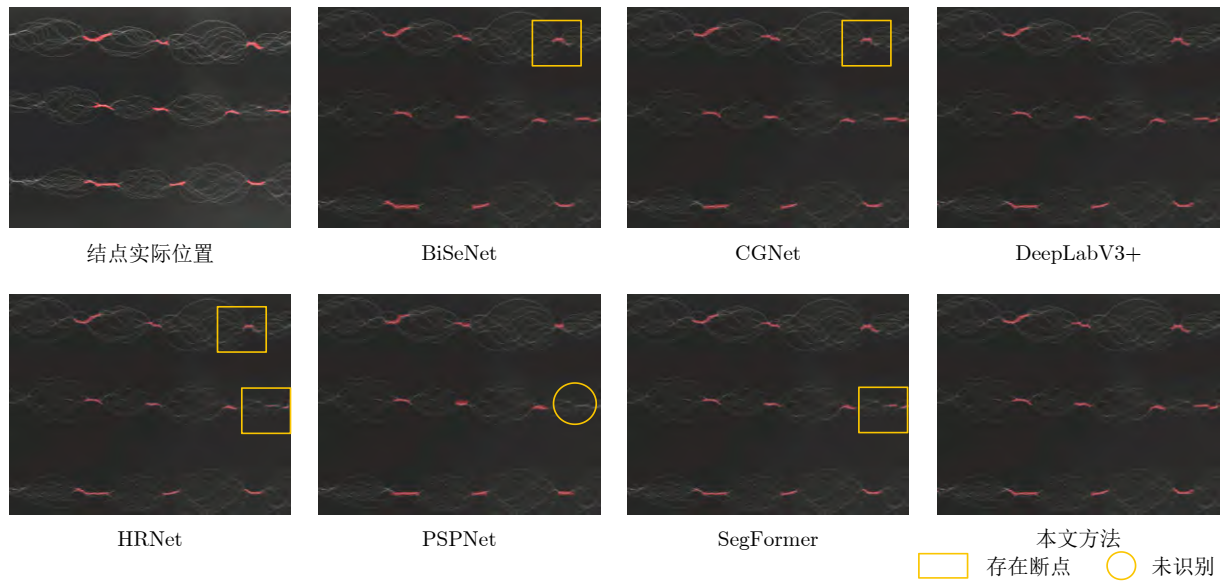


图 11 丝线散开大于 20 mm 时的图像检测结果

Fig. 11 Detection results of images with filament yarn spreading greater than 20 mm

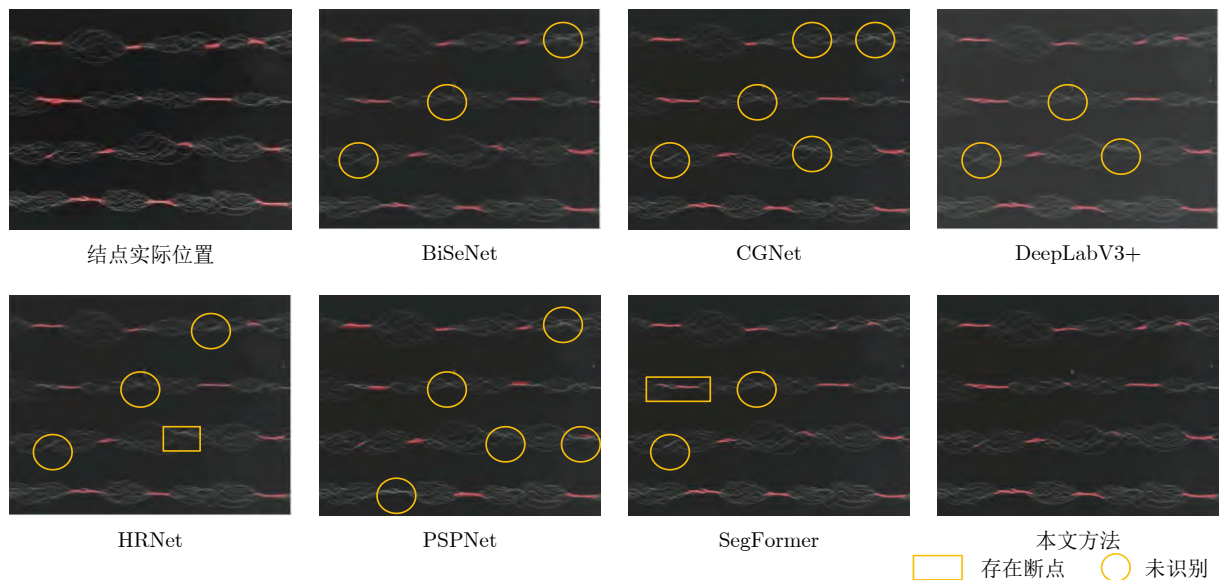


图 12 丝线散开小于 10 mm 时的图像检测结果

Fig. 12 Detection results of images with filament yarn spreading less than 10 mm

特征融合模块; 6) 使用特征图  $x_4$ 、所提语义信息增强和阶段性特征融合模块; 7) 使用所提语义信息增强和多级特征扩张模块直接融合; 8) 使用本文所提方法. 实验结果如表 5 所示.

由表 5 第 5 行和第 8 行可知, 当主干网络的浅层特征缺少语义信息增强模块时,  $MIoU$  值下降了 0.81%,  $FPS$  值下降了 20.86 帧/s. 当主干网络的深层特征缺少语义信息增强模块时,  $MIoU$  值下降了 0.37%,  $FPS$  值下降了 14.68 帧/s. 由此可见, 本文提出的全局注意力机制与扩张卷积方法能够增强模

型对语义信息的提取能力, 同时改善了模型的运算速度. 由表 5 的第 7 行和第 8 行可知, 使用本文提出的阶段性特征融合模块使得  $MIoU$  值增加 1.27%, 说明本文提出的阶段性特征提取方式能够提高检测结果的准确性.

### 3.6.2 主干网络提取效率实验比较

本文采用 6 种主干网络与本文所提主干网络进行对比实验, 包括 FCN, MobileNetV2, Xception, VGGNet, ResNet18 和 ResNet50. 实验结果如表 6 所示. 结果表明, 使用本文提出的主干网络在检测

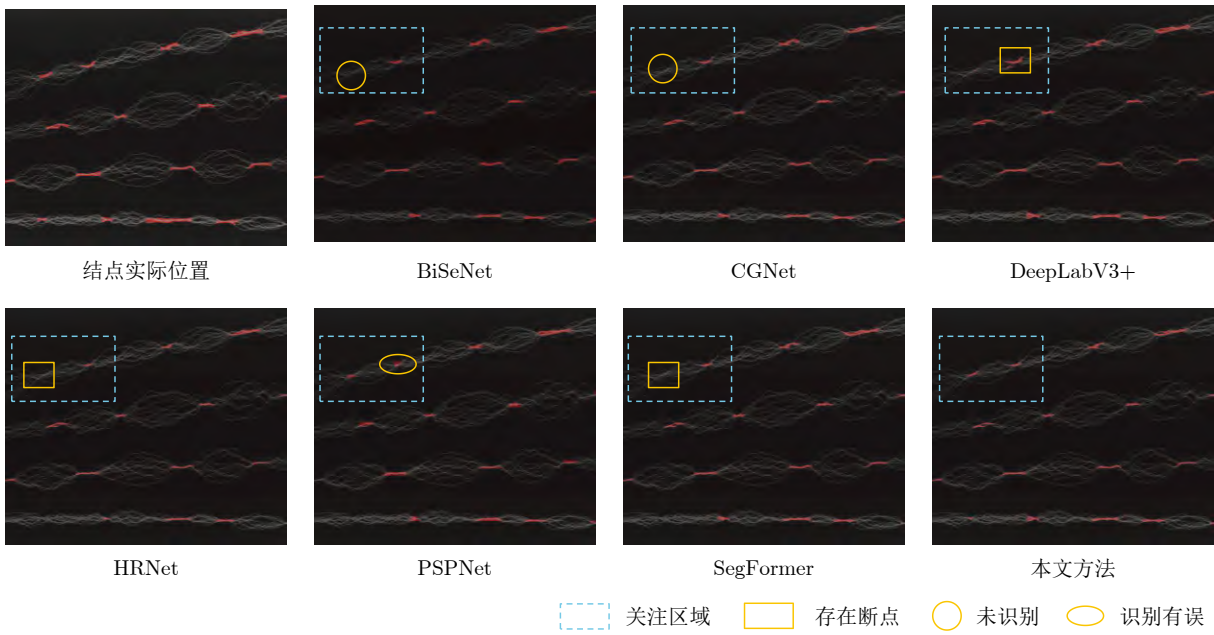


Fig.13 Detection results of images with inclined filament yarn

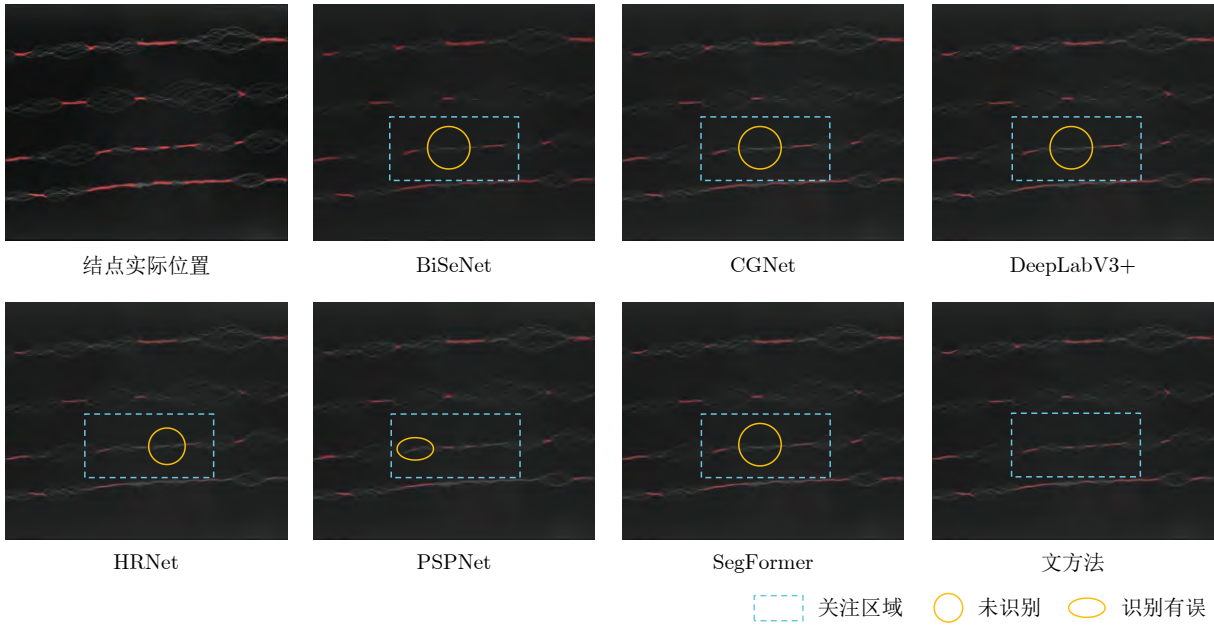


Fig.14 Detection results of images with indistinct interlacing nodes

结果准确性方面均有提升.

3.6.3 语义信息提取方法实验

提取主干网络浅层特征图的语义信息时, 本实验设计 4 种实验方案与所提方法进行对比, 分别是: 1) 只使用空间注意力机制 SA (Spatial attention); 2) 使用注意力机制 SE (Squeeze-and-excitation); 3) 使用空间注意力机制 SA 与通道注意力机制 CA

(Channel attention) 串行的注意力机制 CBAM (Convolutional block attention module); 4) 使用通道注意力机制 ECA (Efficient channel attention). 实验结果如表 7 所示. 结果表明, 只使用一种通道注意力机制的方法, 提取特征的效果相对较弱. 使用通道注意力机制 ECA 进行特征提取后, *MIoU* 值低于原通道注意力机制处理后的结果, 说

表 5 模块有效性验证实验结果  
Table 5 Results of module validity verification experimental

方案序号	语义信息增强模块	多级特征扩张模块	阶段性特征融合模块	<i>MIoU</i> (%)	<i>FPS</i> (帧/s)
1	×	×	×	77.18	72.75
2	✓	×	×	79.91	72.15
3	×	✓	×	79.85	66.32
4	×	×	✓	79.33	68.31
5	×	✓	✓	80.71	55.30
6	✓	×	✓	81.15	61.48
7	✓	✓	×	80.25	78.16
8	✓	✓	✓	81.52	76.16

注: ✓ 指使用此模块, × 指不使用此模块.

表 6 不同主干网络提取效率比较  
Table 6 Comparison of extraction efficiency of different backbone networks

方案序号	主干网络	<i>MIoU</i> (%)	<i>FPS</i> (帧/s)
1	FCN	79.65	33.45
2	MobileNetV2	80.25	43.11
3	Xception	79.61	27.45
4	VGGNet	77.45	30.12
5	ResNet18	77.52	45.21
6	ResNet50	78.01	47.06
7	本文方法	81.52	76.16

表 7 不同语义信息提取方法结果比较  
Table 7 Comparison results of extraction method for different context information

方案序号	注意力选择	<i>MIoU</i> (%)
1	SA	80.36
2	SE	80.44
3	CBAM	80.83
4	ECA	79.89
5	本文方法	81.52

明在特征提取时只关注局部注意力容易减少网络结点特征的提取效果. 将空间注意力机制 SA 与通道注意力机制 CA 串行处理后, 检测结果得到提升, 证明了多注意力机制的有效性.

3.6.4 扩张卷积模块实验比较

提取主干网络深层特征中的语义信息时, 针对主干网络处理后输出的最后两级特征图  $x_3$  和  $x_4$ , 设计 4 种扩张卷积提取方案, 分别是: 1) 不使用扩张卷积模块; 2) 只使用特征图  $x_3$  进行扩张卷积; 3) 只使用特征图  $x_4$  进行扩张卷积; 4) 使用本文所提方法. 实验结果如表 8 所示. 结果表明, 使用扩张卷积后, *MIoU* 指标提升 0.96%, 使用特征图  $x_3$  进行辅助特征提取能够有效增强扩张卷积的效果.

表 8 不同扩张卷积提取方式结果比较  
Table 8 Comparison results of different dilated convolution extraction methods

方案序号	$x_3$	$x_4$	<i>MIoU</i> (%)
1	×	×	80.56
2	✓	×	81.02
3	×	✓	81.13
4	✓	✓	81.52

注: ✓ 指使用此模块, × 指不使用此模块.

3.6.5 阶段性特征融合方法实验

对语义信息增强模块与多级语义信息扩张卷积模块使用通道注意力权重融合时, 设计 4 种方案, 分别是: 1) 只使用全局平均池化方法; 2) 只使用逐点卷积方法; 3) 使用全局平均池化方法和逐点卷积方法串行处理; 4) 使用全局平均池化方法和逐点卷积方法并行处理. 实验结果如表 9 所示. 实验结果表明, 将全局平均池化与逐点卷积方法进行串行连接时, 导致图像原有特征信息丢失严重. 而使用并行连接时, 能够更有效地保留图像的特征信息.

表 9 阶段性特征融合方法实验比较  
Table 9 Comparison results of staged feature fusion module

方案序号	全局平均池化	逐点卷积	组合方法	<i>MIoU</i> (%)
1	✓	×	无	80.91
2	×	✓	无	80.65
3	✓	✓	串行	78.91
4	✓	✓	并行	81.52

注: ✓ 指使用此模块, × 指不使用此模块.

4 结束语

针对人工检测化纤丝线网络度过程中存在误检率高的问题, 本文提出一种基于语义信息增强的化



丝线网络度检测方法, 同时研制网络度检测设备, 实现网络度的自动化视觉精准检测. 其中, 为解决丝线网络结点这一小目标检测问题, 设计语义信息增强模块和多级特征扩张模块, 增强结点信息的提取能力, 并设计阶段性特征融合模块减少特征融合时造成的语义信息丢失. 为实现多根丝线网络度的批量快速检测, 基于掩膜提取与检测算法并行处理机制设计了一种网络度并行检测方法. 利用自主研发的网络度检测装备对所提方法进行了实验验证. 结果表明, 本文所提方法能够有效地增强化纤丝线网络结点的提取能力, 同时提高目标识别的准确性. 下一步工作主要针对化纤工业现场环境进行实验验证, 从检测效率方面进一步对本文提出的检测方法和设备进行迭代更新.

## References

- Chen Xiang-Ling, Wang Hua-Ping, Ji Peng. Flexibility and multi-objective production of intelligent manufacturing of China's chemical fiber industry. *China Textile Leader*, 2020, **916**(3): 13-14, 16-18, 21-22, 24-25  
(陈向玲, 王华平, 吉鹏. 我国化纤智能制造的柔性性与多目标生产. 纺织导报, 2020, **916**(3): 13-14, 16-18, 21-22, 24-25)
- Wang M, Zhan Y L, Yao L. A new method to evaluate interlacing yarns. *Textile Research Journal*, 2020, **90**(7-8): 838-846
- Zhang Ye-Xing, Chen Hao-Qiu, Gao Guo-Hong, Huang Meng-Fu, Xiao Yong-Xin, Liu Li-Na. Comparative analysis about testing method of synthetic filament network degree. *China Fiber Inspection*, 2010, **359**(15): 54-57  
(张叶兴, 陈浩秋, 高国洪, 黄猛富, 肖永新, 刘丽娜. 合成纤维长丝网络度测试方法比较分析. 中国纤检, 2010, **359**(15): 54-57)
- Yan N, Zhu L L, Yang H M, Li N N, Zhang X D. Online yarn breakage detection: A reflection-based anomaly detection method. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: 1-13
- Guo M R, Gao W D, Wang J A. Online measurement of sizing yarn hairiness based on computer vision. *Fibers and Polymers*, 2023, **24**(4): 1539-1552
- Wang L, Lu Y C, Pan R R, Gao W D. Evaluation of yarn appearance on a blackboard based on image processing. *Textile Research Journal*, 2021, **91**(19-20): 2263-2271
- Khaddam H S, Ahmad G G. A method to evaluate the diameter of carded cotton yarn using image processing and artificial neural networks. *The Journal of the Textile Institute*, 2022, **113**(8): 1648-1657
- Tang Jia-Lu, Yang Zhong-Liang, Zhang Song, Mao Xin-Hua, Dong Qing-Qi. Detection of yarn hairiness combining microscopic vision and attention mechanism. *CAAI Transactions on Intelligent Systems*, 2022, **17**(6): 1209-1219  
(唐嘉璐, 杨钟亮, 张崧, 毛新华, 董庆奇. 结合显微视觉和注意力机制的毛羽检测方法. 智能系统学报, 2022, **17**(6): 1209-1219)
- Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 640-651
- Zhao H S, Shi J P, Qi X J, Wang X G, Jia J Y. Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 6230-6239
- Zhou Z W, Siddiquee M M R, Tajbakhsh N, Liang J M. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 2019, **39**(6): 1856-1867
- Chen L C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision. Munich, Germany: Springer, 2018. 801-818
- Chen L C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv: 1706.05587, 2017.
- Chen L C, Papandreou G, Kokkinos I, Murphy K, Yuille A L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **40**(4): 834-848
- Rush A M, Chopra S, Weston J. A neural attention model for abstractive sentence summarization. arXiv preprint arXiv: 1509.00685, 2015.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. arXiv preprint arXiv: 1706.03762, 2023.
- Cheng H X, Han X F, Xiao G Q. TransRVNet: Lidar semantic segmentation with transformer. *IEEE Transactions on Intelligent Transportation Systems*, 2023, **24**(6): 5895-5907
- Liu Z, Hu H, Lin Y T, Yao Z L, Xie Z D, Wei Y X, et al. Swin transformer v2: Scaling up capacity and resolution. In: Proceedings of the International Conference on Computer Vision. New Orleans, USA: IEEE, 2022. 12009-12019
- Huang T Y, Chen K X, Jiang L F. DS-UNet: Depthwise separable convolution network with large convolutional kernel for medical image segmentation. *Signal, Image and Video Processing*, 2023, **17**(5): 1775-1783
- Sinha A, Dolz J. Multi-scale self-guided attention for medical image segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2020, **25**(1): 121-130
- Peng Xiu-Ping, Tong Qi-Sheng, Lin Hong-Bin, Feng Chao, Zheng Wu. A deep residual-feature pyramid network framework for scattered point cloud semantic segmentation. *Acta Automatica Sinica*, 2021, **47**(12): 2831-2840  
(彭秀平, 仝其胜, 林洪彬, 冯超, 郑武. 一种面向散乱点云语义分割的深度残差-特征金字塔网络框架. 自动化学报, 2021, **47**(12): 2831-2840)
- Zhai W Z, Gao M L, Li Q L, Jeon G, Anisetti M. FPANet: Feature pyramid attention network for crowd counting. *Applied Intelligence*, 2023, **53**(16): 19199-19216
- Mei Y Q, Fan Y C, Zhang Y L, Yu J H, Zhou Y Q, Liu D, et al. Pyramid attention network for image restoration. *International Journal of Computer Vision*, 2023, **131**(12): 3207-3225
- Fan Bing-Bing, Ge Li-Yue, Zhang Cong-Xuan, Li Bing, Feng Cheng, Chen Zhen. A feature pyramid optical flow estimation method based on multi-scale deformable convolution. *Acta Automatica Sinica*, 2023, **49**(1): 197-209  
(范兵兵, 葛利跃, 张聪炫, 李兵, 冯诚, 陈震. 基于多尺度变形卷积的特征金字塔光流计算方法. 自动化学报, 2023, **49**(1): 197-209)
- Jin Xia-Ting, Wang Yao-Nan, Zhang Hui, Liu Li, Zhong Hang, He Zhen-Dong. DeepRail: Automatic visual detection system for railway surface defect using Bayesian CNN and attention network. *Acta Automatica Sinica*, 2019, **45**(12): 2312-2327  
(金侠挺, 王耀南, 张辉, 刘理, 钟杭, 贺振东. 基于贝叶斯 CNN 和注意力网络的钢轨表面缺陷检测系统. 自动化学报, 2019, **45**(12): 2312-2327)
- Sandler M, Howard A, Zhu M L, Zhmoginov A, Chen L C. MobilenetV2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 4510-4520
- Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 1251-1258
- Kim S, Min D B, Ham B, Jeon S, Lin S, Sohn K. FCSS: Fully convolutional self-similarity for dense semantic correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, **41**(3): 581-595
- Kirisci M. New cosine similarity and distance measures for fermatean fuzzy sets and topsis approach. *Knowledge and Information Systems*, 2023, **65**(2): 855-868
- Yu C Q, Wang J B, Peng C, Gao C X, Yu G, Sang N. BiSeNet:

Bilateral segmentation network for real-time semantic segmentation. In: Proceedings of the European Conference on Computer Vision. Munich, Germany: IEEE, 2018. 325–341

- 31 Meng Jun-Xi, Zhang Li, Cao Yang, Zhang Le-Tian, Song Qian. Optimization of image semantic segmentation algorithms based on Deeplab v3+. *Laser and Optoelectronics Progress*, 2022, **59**(16): 161–170  
(孟俊熙, 张莉, 曹洋, 张乐天, 宋倩. 基于 Deeplab v3+ 的图像语义分割算法优化研究. *激光与光电子学进展*, 2022, **59**(16): 161–170)
- 32 Wu T Y, Tang S, Zhang R, Zhang Y D. CGNet: A light-weight context guided network for semantic segmentation. *IEEE Transactions on Image Processing*, 2020, **30**: 1169–1179
- 33 Sun K, Xiao B, Liu D, Wang J D. Deep high-resolution representation learning for human pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019. 5693–5703
- 34 Xie E, Wang W H, Yu Z D, Anandkumar A, Alvarez J M, Luo P. SegFormer: Simple and efficient design for semantic segmentation with transformers. arXiv preprint arXiv: 2105.15203, 2021.



**郑广智** 东北大学流程工业综合自动化国家重点实验室硕士研究生。2021 年获得大连海洋大学学士学位。主要研究方向为机器学习, 深度学习和计算机视觉。E-mail: zgz4923@163.com  
(**ZHENG Guang-Zhi** Master student at the State Key Laboratory of

Synthetical Automation for Process Industries, Northeastern University. He received his bachelor degree from Dalian Ocean University in 2021. His research interest covers machine learning, deep learning, and computer vision.)



**彭添强** 东北大学流程工业综合自动化国家重点实验室硕士研究生。2020 年获得青岛理工大学学士学位。主要研究方向为机器学习和深度学习。E-mail: ptq15236155308@163.com

(**PENG Tian-Qiang** Master student at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. He received his bachelor degree from Qingdao University of Technology in 2020. His research interest covers machine learning and deep learning.)



**肖计春** 东北大学机械工程与自动化学院博士研究生。2020 年获得重庆理工大学硕士学位。主要研究方向为运动规划, 攀爬机器人和自主系统。E-mail: xc390297815@163.com  
(**XIAO Ji-Chun** Ph.D. candidate at the School of Mechanical Engin-

earing and Automation, Northeastern University. He received his master degree from Chongqing University of Technology in 2020. His research interest covers motion planning, climbing robots, and autonomous systems.)



**吴高昌** 东北大学流程工业综合自动化国家重点实验室副教授。主要研究方向为智能计算成像, 深度学习和异常工况智能感知与预测。

E-mail: wugc@mail.neu.edu.cn

(**WU Gao-Chang** Associate professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. His research interest covers intelligent computational imaging, deep learning, and intelligent sensing and prediction of abnormal working conditions.)



**李 智** 东北大学流程工业综合自动化国家重点实验室教授。主要研究方向为数据驱动的建模与控制方法, 精密运动控制和智能机器人。本文通信作者。E-mail: lizhi1@mail.neu.edu.cn

(**LI Zhi** Professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. His research interest covers data-driven modeling and control methods, precision motion control, and intelligent robots. Corresponding author of this paper.)



**柴天佑** 中国工程院院士, 东北大学教授, IEEE Life Fellow, IFAC Fellow, 欧亚科学院院士。主要研究方向为自适应控制, 智能解耦控制, 流程工业综合自动化与智能化系统理论、方法与技术。

E-mail: tychai@mail.neu.edu.cn

(**CHAI Tian-You** Academician of Chinese Academy of Engineering, professor at Northeastern University, IEEE Life Fellow, IFAC Fellow, and academician of the International Eurasian Academy of Sciences. His research interest covers adaptive control, intelligent decoupling control, and theories, methods and technology of synthetical automation and intelligent system for process industries.)