PR U-Net: 工业缺陷分割中的特征融合和 渐进监督策略

王新宇

南京邮电大学理学院, 江苏 南京

收稿日期: 2024年6月26日; 录用日期: 2024年7月23日; 发布日期: 2024年7月30日

摘要

在工业缺陷检测任务中,由于缺陷与周围环境的相似性,准确识别图像中的缺陷区域是一项重要的挑战。 本文提出了一种名为渐进细化(PR) U-Net的新型语义分割模型,该模型基于Swin U-Net架构,通过集成 PR解码块和焦点模块(FM)来进行改进。PR解码块使模型能够逐步细化特征,并在每个解码阶段引导监 督,而FM通过融合分割预测和特征融合机制,有助于更好地区分缺陷和非缺陷区域。大量实验结果表明, 本文的PR U-Net在区分背景区域中的细微缺陷方面具有更高的准确性和鲁棒性,同时在推理时间方面也 具有高效性,使其适用于各种场景下的工业缺陷检测任务。

关键词

渐进监督,工业缺陷分割,特征焦点,Swin U-Net

PR U-Net: Feature Fusion and Progressive Supervision Strategy in Industrial Defect Segmentation

Xinyu Wang

School of Science, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

Received: Jun. 26th, 2024; accepted: Jul. 23rd, 2024; published: Jul. 30th, 2024

Abstract

The paper presents a novel semantic segmentation model named Progressive Refining (PR) U-Net, designed to enhance industrial defect detection. The model builds upon the Swin U-Net architec-

ture, incorporating PR Decoder Blocks and Focus Modules (FMs) to improve feature refinement and segmentation accuracy. The PR Decoder Blocks allow for progressive feature refinement, while FMs aid in distinguishing defects from non-defects by integrating segmentation predictions and feature fusion mechanisms. Extensive experiments demonstrate that PR U-Net outperforms existing models in accuracy, robustness, and efficiency in industrial defect detection tasks.

Keywords

Progressive Supervision, Industrial Defect Segmentation, Feature Focus, Swin U-Net

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

CC O Open Access

1. 引言

图像分割技术在工业领域中起着至关重要的作用,它有助于将图像划分为语义上有意义的区域。其 应用使得工业企业能够实现自动化、提高产品质量、降低成本,并增强竞争力。

图像分割技术广泛应用于工业环境中[1]-[4]。例如,它广泛应用于工业机器人视觉系统中,提高了机器人准确识别和定位目标的能力[5]。此外,图像分割技术对工业自动化和监控也有重大贡献[6]。通过分析工厂内部环境和设备状态的分割图像,企业能够实现生产过程的实时监控和优化,从而及时发现和解决潜在问题,提高生产效率和资源利用率。值得注意的是,其最有价值的应用之一在于装配线生产中的缺陷检测[3][7]。利用图像分割技术,工业企业能够快速、准确地识别产品表面的缺陷、瑕疵或其他不合格之处,从而实现产品质量的有效监控和控制。这对于提高产品质量、降低售后服务成本以及维护品牌声誉至关重要。

然而,工业领域中的图像分割技术面临着诸多挑战和问题。Liu [8]等人指出,矿物分割模型在部署、 分割精度和区域重叠方面面临的挑战。同样,Hesamian [9]等人指出,深度学习分割模型在工业领域中面 临的数据标注不足、类别不平衡以及过拟合导致的分割精度下降等问题。总体而言,分割结果存在大量 的假阳性和假阴性,以及普遍的边界错误。

为应对这些挑战,本文提出了渐进式精细化 U-Net (PR U-Net)。在 Swin U-Net [10]架构的基础上,本 文重新设计了解码器模块,加入了三个渐进式精细化解码块(PR 解码块)。通过传播和监督分层损失,并 逐步精细化分割结果,本文旨在提高分割结果的质量,同时减少边界错误。同时,在每个 PR 解码块中, 本文引入了聚焦模块结构,以帮助模型更有效地区分缺陷区域和正常区域,从而减少假阳性和假阴性的 发生。

2. 相关工作

2.1. 基于 CNN 的方法

最初,学者们通过将卷积神经网络(CNN)中的最后一层全连接层替换为卷积层,使模型输出特征图 作为分割结果。考虑到不同层次之间尺寸和语义信息的差异,FCN [11]引入了上采样和跳跃连接,使高 层特征能够直接添加到低层特征中,初步形成了特征金字塔的雏形。然而,FCN 没有考虑全局特征,忽 略了潜在有用的场景级语义上下文[12]。

2.2. 基于编码器 - 解码器的方法

在图像分割领域,编码器 - 解码器模型也是一类流行的深度学习模型。通用分割的编码器 - 解码器 模型包括 SegNet [13]、HRNet [14]等。其中,Noh [15]等人提出了一种基于反卷积的语义分割方法,使用 VGG 16 层网络的卷积层作为编码器,并通过反卷积网络生成像素级类别概率图。SegNet 利用编码器 -解码器架构,通过在解码器中使用池化索引进行非线性上采样实现图像分割。在特定应用领域,如医学 图像处理[16]-[18]和工业图像[19]-[21]处理,U-Net [22]和 V-Net [23]受到了广泛关注。U-Net 由 Ron 博士 提出,具有一个收缩路径和对称扩展路径,能够有效地从少量标注图像中学习。U-Net 的许多变体在医 学和工业领域发挥了重要作用。V-Net 通过引入基于 Dice 系数的新目标函数,解决了前景和背景体素数 量不平衡的问题。

2.3. 基于注意力机制的方法

注意力机制在计算机视觉领域得到了持续探索,使模型在处理过程中能够更专注于输入数据的相关 部分。因此,在语义分割中应用注意力机制是一种自然趋势。Chen [24]等人提出了一种注意力机制,学 习每个像素位置上多尺度特征的软权重。Huang [25]等人提出了一种相反于传统方法的方法,他们设计了 一个反向注意力网络(RAN),捕捉与目标类别无关的特征。为了增强特征,Fu [26]等人引入了双重注意 力网络,利用自注意机制捕捉图像中的丰富上下文依赖关系。通过在扩张卷积网络中附加两种类型的注 意力模块,该模型能够模拟空间和通道维度的语义相互依赖关系,以理解图像的语义结构并实现更好的 分割结果。

除了这些方法外,还有一些工作探索了注意力机制在语义分割中的应用[27]-[29]。SwinU-Net 结合了 Swin Transformer 和 U-Net,是一个重要的里程碑。一方面,Swin Transformer [30]是一种强大的注意力机 制网络,能够有效处理不同尺度的特征。另一方面,U-Net 结构是一种经典的编码器 - 解码器结构,用 于保留和融合不同分辨率的特征。Swin U-Net 结合了两者的优势,在语义分割任务中表现出色,通过跨 区域信息交换和多尺度特征融合实现更准确和稳定的图像分割。

2.4. 计算机视觉在工业中的应用

视觉技术在现代工业中发挥着至关重要的作用。其基本原理是使用摄像机捕捉目标物体的图像,然 后使用计算机算法对这些图像进行处理和分析,实现目标物体的检测、识别和定位等功能。在制造业中, 视觉技术广泛应用于产品工艺监控、缺陷检测等阶段。

以图像分割为例,它不仅是计算机视觉领域的重要研究方向,而且在工业领域的各种任务场景中得到了广泛应用。Song [31]等人提出了一种结合空间注意力和前馈神经网络的混合残差模块(SAFM),替代编码器中的下采样层以提取缺陷特征,代表了工业缺陷检测领域的前沿方法。

总的来说,视觉技术在工业中的应用前景广阔且意义重大。随着计算机视觉和人工智能技术的发展, 视觉技术为工业生产的智能化和自动化提供了强有力的支持,推动了工业领域的进步与发展。

3. 本文模型

在工业缺陷检测任务中,模型需要准确识别图像中的缺陷区域,但由于缺陷(正样本)与周围环境(负 样本)的相似性,这一任务具有一定的难度。Swin U-Net 作为一种结合了 Swin Transformer 和 U-Net 的语 义分割模型,在通用语义分割任务中表现出色。然而,其特征提取和信息融合能力不足以有效地区分背 景中的细小缺陷区域,从而降低了在工业缺陷检测任务中的效率。

为了解决这些问题,本文提出了渐进细化 U-Net (PR U-Net)。见图1所示, PR U-Net 保留了 Swin U-Net

的编码器部分,以保留其特征提取和侧向连接能力。为了进一步增强模型区分前景和背景的能力,本文 重新设计了整个解码器部分,包含三个渐进细化解码块(PR 解码块)。每个 PR 解码块包含两个焦点模块 (FM),用于在当前分辨率下细化特征图,并更好地区分目标和非目标区域。通过传递编码器中获得的特 征并经过三个 PR 解码块,模型逐步提高其语义分割精度。此设计使 PR U-Net 能够更有效地利用特征细 化和信息融合机制,从而提高其在工业缺陷检测任务中的性能。



3.1. 焦点模块

为了更好地区分缺陷和非缺陷区域,减少假阳性和假阴性的数量,本文设计了一个焦点模块。见图 2 所示,焦点模块包含三个输入:深层语义特征 *F_H*、浅层语义特征 *F_L*与和分割预测 *F_{seg}*焦点模块的目标是 通过特征融合,提高最终分割结果的准确性。

步骤 1: 对深层语义特征 F_H 、和分割预测 F_{seg} 进行两倍上采样,得到上采样后的特征 F'_H 和 F'_{seg} 此步骤的目的是将低分辨率特征图上采样到与高分辨率特征图一致,以确保后续操作的一致性和有效性。

步骤 2: 在特征融合过程中,首先通过交叉熵操作,将浅层语义特征 *F_L*与分割预测 *F_{seg}* 逐元素相乘, 强调分割预测中的缺陷区域。然后,获得的特征与权重参数α相乘,并从深层语义特征 *F'_H* 中减去,得到 特征 *F*₁。此步骤的目的是强调模型应关注的缺陷区域,并减少模型对背景区域的注意,从而减少假阳性。

如公式(1)所示:

$$F_1 = \alpha \times \text{CE Block} \left(F'_{seg} \odot F_L \right) - F'_H \tag{1}$$





在第二个分支中,将浅层语义特征 *F_L*与反向分割预测1-*F'_{seg}* 逐元素相乘,以排除分割预测中的缺陷 区域。然后,同样通过交叉熵操作并乘以权重参数 β,特征 *F*₂。此步骤的目的是强调模型应忽略的区域, 并减少模型对非缺陷区域的注意,从而减少假阴性。

如公式(2)所示:

$$F_2 = \beta \times \text{CE Block}\left(\left(1 - F'_{seg}\right) \odot F_L\right)$$
(2)

步骤 3:从两条分支获得的特征进行逐元素异或操作,得到焦点模块的输出 out。如公式(3)所示:

$$out = F_1 \oplus F_2 \tag{3}$$

通过这种设计,焦点模块有效地融合了不同层次的特征信息,并关注模型需要关注的区域,从而提 高了模型在缺陷检测任务中的性能。

3.2. 渐进细化解码块

渐进细化解码块设计用于语义分割任务,通过逐步细化特征并融合上层解码器的分割结果,生成最终分 割输出。见图 3 所示,渐进细化解码块包含两个分支,每个分支集成了一个焦点模块和一系列卷积块操作。



步骤 1: 对输入特征进行两倍上采样,以保持分辨率一致。然后,渐进细化解码块由两个分支组成, 每个分支包含一个焦点模块和一系列卷积块操作。

步骤 2: 在第一个分支中,输入分割结果 F_{seg} 、解码器特征 e_D 和卷积块操作后的编码器特征 e_E 进入 焦点模块 FM_1 ,得到特征 z_1 。然后,将 z_1 与卷积块操作后的编码器特征 e'_E 拼接,生成融合特征表示。最 后,通过两次卷积操作,进一步细化融合特征,得到特征表示 z'_1 。

如公式(4) (5) (6)所示:

$$z_1 = FM_1(F_{seg}, \text{CBlock}(e_E), e_D)$$
(4)

$$z_{1} \leftarrow \left[z_{1}; \operatorname{CBlock}\left(e_{E}^{\prime}\right) \right]$$
(5)

$$z_1' \leftarrow \text{CBlock}(\text{CBlock}(z_1)) \tag{6}$$

在该分支中,本文强调了分割结果信息的及时融合,突出模型感兴趣的区域。通过将分割结果与解 码器和编码器特征联合处理,模型能够更好地学习目标区域的特征表示。该分支增强了分割结果中目标 区域的特征学习,提升了模型对缺陷等目标的表示能力和检测能力。

在第二个分支中,输入分割结果 F_{seg} 、上采样的编码器嵌入 e'_{E} 和解码器嵌入 e'_{D} 进入焦点模块 FM_{2} ,得到特征 z_{2} ,然后,将第一个分支生成的细化特征与第二个分支生成的特征进行级联操作,最后,通过三次卷积操作,进一步提取和融合级联后的特征,生成最终的输出 *out*。

如公式(7)(8)(9)所示:

$$z_2 = FM_2(F_{seg}, \text{CBlock}(e'_E), \text{CBlock}(e'_D))$$
(7)

$$z = \begin{bmatrix} z_1'; z_2 \end{bmatrix} \tag{8}$$

$$out = \text{CBlock}\left(\text{CBlock}\left(\text{CBlock}\left(z\right)\right)\right)$$
(9)

最后,为了逐层引导 F_{seg} 的生成过程,在 PR 解码器块的每一层计算每层的 F_{seg} 与实际分割结果 y_{gt} 之间的损失,具体公式(10)如下:

$$\mathcal{L}_{total} = \sum_{l=1}^{L} \mathcal{L} \left(F_{seg,l}, y_{gt} \right)$$
(10)

这里, *L* 表示 PR 解码块的层数, *F_{seg,l}* 表示第*l* 层的分割结果, *y_{gt}* 是实际分割结果。这个损失计算 机制确保每层的分割结果受到监督信号的引导,有助于提高模型在语义分割任务中的学习效果,从而使 模型能够更好地理解图像内容并生成准确的分割结果。

4. 实验

4.1. 数据集

所有实验均在 MVTec AD 异常检测(MVTec AD)数据集上进行,该数据集是检测工业制造中表面缺陷的广泛使用的数据集。MVTec AD 数据集由 MVTec Software GmbH 创建,包含各种类型的高分辨率图像,这些图像包含裂纹、凹痕、划痕和斑点缺陷等各种表面异常。这些图像覆盖了包括塑料、纸张、纺织品和金属在内的多种材料。

MVTec AD 数据集旨在促进工业表面缺陷检测算法的研究,作为评估算法在实际工业场景中性能的基准。每类图像分为训练集和测试集,训练集包含正常样本和异常样本,测试集仅包含异常样本。

该数据集的一个关键特性是其丰富的变异性,包括光照条件、视角、缺陷类型和表面材料的变化。 这使得 MVTec AD 数据集成为评估算法鲁棒性和泛化能力的理想选择。

4.2. 参数设置

所有实验均在一个综合的软件和硬件环境中进行。操作系统为 Ubuntu 20.04,主要库包括 Python 3.9、 CUDA 11.8 和 PyTorch 2.0。硬件配置为 Intel i9 12900KF CPU 和 NVIDIA RTX 3090TI GPU,系统配备了 32GB 3200MHz 的内存。本文的 PR U-Net 模型训练了 100 个 epoch,批处理大小为 32。优化器为 Adam, 学习率为 0.001。这些配置在整个实验过程中保持一致,确保了结果的可靠性和可重复性。

4.3. 消融实验

为了评估本文提出的创新的有效性和贡献,本文以 Swin U-Net 作为基准模型进行了消融研究。

见表 1 所示,消融实验结果表明,本文提出的 PR U-Net 在所有指标上均优于基准模型 Swin U-Net。 具体来说,PR U-Net 在 Dice 系数、召回率、精确度、F1 分数和 IoU 等指标上都表现出更高的数值。例 如,在 Dice 系数方面,PR U-Net 在负类上的表现为 99.8593,显著高于 Swin U-Net 的 99.6877,而在正 类上的表现为 96.5790,明显高于 Swin U-Net 的 92.4085。此外,PR U-Net 在平均指标上也有显著提升, 其表现为 98.2192,而 Swin U-Net 为 96.0481。

见图 4 所示,在正常和异常区域交界处的分割准确性方面, PR U-Net 模型显著优于 Swin U-Net 模型, 这意味着本文实现了精细分割的研究目标。总体而言, PR U-Net 在工业缺陷检测任务中表现出优越的性能,进一步验证了其在特征提取和信息融合方面的有效性和优越性。



Figure 4. Example of comparison of segmentation results in ablation experiments 图 4. 消融实验分割结果对比示例

4.4. 与经典模型的比较

本文使用其他经典的图像分割模型在 MVTec AD 数据集上进行了实验。为了确保比较的严格性和公 平性,这些模型都在相同的硬件配置和相同的方式下进行了训练。

见表 2 所示, PR U-Net 在工业数据集上显著优于 ViT [23]和 TransUnet [27]这两个强大的分割模型。 这证明了逐步细化分割在工业异常分割任务中的关键作用。

4.5. 与最新模型比较

在本节中,本文将 PR U-Net 模型与当前工业缺陷分割领域的最新模型 SAFM Res UNet [31]在准确性和推理时间方面进行比较。

见表 3 所示, PR U-Net 在这两个特定数据集上的图像分割能力均优于 SAFM Res UNet。这意味着 PR U-Net 更能准确地识别工业产品中的细微缺陷和异常。这样的性能优势表明, PR U-Net 在实际工业应 用中更为可靠,特别是在需要极高精度和准确度的制造和质量控制领域。它还展示了 PR U-Net 在不同工 业缺陷场景中的优秀泛化能力。本文还评估了前面提到的几种常见分割方法的推理时间。同样,所有测 试均在相同配置下使用 MVTec AD 数据集进行,每个模型在该数据集上独立运行 10 次,以确定平均推 理时间。为了更好地呈现实验结果,本文对推理时间进行了归一化处理。

见表 4 所示, PR U-Net 在推理速度方面略优于 SAFM Res UNet, 这使得 PR U-Net 适用于需要快速和精确处理的工业应用。

	类别	Dice	召回率	精确度	F1 分数	IoU
Swin U-Net (基准)	负类	99.6877	99.6805	99.6949	99.6877	99.3774
Swin U-Net (基准)	正类	92.4085	92.5712	92.2465	92.4086	85.8885
Swin U-Net (基准)	平均	96.0481	96.1258	95.9707	96.0481	92.6329
PR U-Net	负类	99.8593	99.8528	99.8658	99.8593	99.719
PR U-Net	正类	96.5790	96.7325	96.4260	96.5700	93.3843
PR U-Net	平均	98.2192	98.2926	98.1459	98.2192	96.5517

Table 1. Comparison of ablation experimental result 表 1. 消融实验结果对比

 Table 2. Comparison results of different classical models

表 2. 不同经典模型的比较结果

	ViT	TransUnet	Swin U-Net	PR U-Net
Dice	93.2581	95.3832	96.0481	98.2192
F1 分数	93.2546	95.3217	96.0481	98.2129
IoU	90.4587	91.2321	92.6329	96.5517

Table 3. PR U-Net and SAFM Res Unet based on Dice comparison 表 3. PR U-Net 与 SAFM Res Unet 基于 Dice 系数泛化性比较

	MVTec AD	KolektorSDD
SAFM Res UNet	97.8	95.1
Swin U-Net	96.0	94.5
PR U-Net	98.2	95.3

 Table 4. Comparison of normalized inference time between different models

 表 4. 不同模型推理时间经归一化后的比较

	Vit	TransUnet	SAFM Res UNet	PR U-Net
平均推理时间	1	0.6	0.09	0

5. 结论

在本文中,本文提出了一种新型的语义分割模型——渐进细化(PR) U-Net,以提高工业缺陷检测的性能。PR U-Net 基于 Swin U-Net 架构,通过集成 PR 解码块和焦点模块(FM)来进行改进。PR 解码块允许逐步细化特征,并在每个解码阶段进行监督,而焦点模块通过特征融合和分割预测,增强了模型区分缺陷本文和非缺陷区域的能力。

本文在 MVTec AD 和 KolektorSDD 两个广泛使用的工业缺陷检测数据集上进行了实验,结果表明, PR U-Net 在所有评估指标上均显著优于现有的基准模型 Swin U-Net 以及其他经典的分割模型,如 ViT 和 TransUnet。此外, PR U-Net 在准确性、召回率、精确度、F1 分数和 IoU 等方面均表现出色,特别是 在细小缺陷区域的分割能力上,取得了显著的提升。

通过消融研究,本文进一步验证了 PR 解码块和焦点模块在提高模型性能方面的有效性。这些模块 的设计不仅提高了特征提取和信息融合的能力,还有效地减少了假阳性和假阴性的发生,从而提高了分 割的精度和鲁棒性。

本文还对比了 PR U-Net 与最新的工业缺陷检测模型 SAFM Res UNet,结果显示 PR U-Net 在多个数 据集上的性能均优于 SAFM Res UNet,证明了本文模型在工业应用中的有效性和广泛适用性。此外, PR U-Net 在推理时间上也表现出色,适合需要快速处理的工业场景。

总的来说, **PR** U-Net 在工业缺陷检测任务中表现出优越的性能和高效性,证明了其在实际应用中的 潜力。未来的研究可以进一步优化模型结构,并探索其在其他图像分割任务中的应用,进一步提升模型 的通用性和适应性。

本文计划未来的工作将集中在以下几个方面:首先,进一步优化 PR U-Net 的结构,提高其在更大规 模数据集上的训练和推理效率;其次,探索更多种类的焦点模块,以进一步增强模型在复杂背景下的分 割性能;最后,将 PR U-Net 应用于其他领域的图像分割任务,如医学图像分割和卫星图像分割,以验证 其通用性和适用性。

综上所述, PR U-Net 是一种有效且高效的工业缺陷检测模型, 其创新性设计在多个实验中表现出色, 具有广泛的应用前景和研究价值。

参考文献

- Liu, Y., Wang, X., Zhang, Z. and Deng, F. (2023) Deep Learning in Image Segmentation for Mineral Production: A Review. *Computers & Geosciences*, 18, Article 105455. <u>https://doi.org/10.1016/j.cageo.2023.105455</u>
- [2] Hesamian, M.H., Jia, W., He, X. and Kennedy, P. (2019) Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *Journal of Digital Imaging*, **32**, 582-596. <u>https://doi.org/10.1007/s10278-019-00227-x</u>
- [3] Gruosso, M., Capece, N. and Erra, U. (2021) Human Segmentation in Surveillance Video with Deep Learning. *Multi-media Tools and Applications*, 80, 1175-1199. <u>https://doi.org/10.1007/s11042-020-09425-0</u>
- [4] Colleoni, E. and Stoyanov, D. (2021) Robotic Instrument Segmentation with Image-to-Image Translation. IEEE Robotics and Automation Letters, 6, 935-942. <u>https://doi.org/10.1109/LRA.2021.3056354</u>
- [5] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, 7-12 June 2015, 3431-3440. https://doi.org/10.1109/CVPR.2015.7298965
- [6] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 21-26 July 2017, 2117-2125. <u>https://doi.org/10.1109/CVPR.2017.106</u>
- Liu, W., Rabinovich, A. and Berg, A.C. (2015) ParseNet: Looking Wider to See Better. arXiv: 1506.04579. https://doi.org/10.48550/arXiv.1506.04579
- [8] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L. (2014) Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. arXiv: 1412.7062. <u>https://doi.org/10.48550/arXiv.1412.7062</u>

- Schwing, A.G. and Urtasun, R. (2015) Fully Connected Deep Structured Networks. arXiv: 1503.02351. https://doi.org/10.48550/arXiv.1503.02351
- [10] Lin, G., Shen, C., Van Den Hengel, A. and Reid, I. (2016) Efficient Piecewise Training of Deep Structured Models for Semantic Segmentation. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 3194-3203. <u>https://doi.org/10.1109/CVPR.2016.348</u>
- [11] Liu, Z., Li, X., Luo, P., Loy, C.C. and Tang, X. (2015) Semantic Image Segmentation via Deep Parsing Network. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 7-13 December 2015, 1377-1385. <u>https://doi.org/10.1109/ICCV.2015.162</u>
- [12] Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 2481-2495. https://doi.org/10.1109/TPAMI.2016.2644615
- [13] Li, X., Chen, H., Qi, X., Dou, Q., Fu, C. and Heng, P. (2018) H-DenseUNet: Hybrid Densely Connected Unet for Liver and Tumor Segmentation from CT Volumes. *IEEE Transactions on Medical Imaging*, 37, 2663-2674. https://doi.org/10.1109/tmi.2018.2845918
- [14] Weng, Y., Zhou, T., Li, Y. and Qiu, X. (2019) NAS-Unet: Neural Architecture Search for Medical Image Segmentation. *IEEE Access*, 7, 44247-44257. <u>https://doi.org/10.1109/access.2019.2908991</u>
- [15] Jiang, J., Zhu, J., Bilal, M., Cui, Y., Kumar, N., Dou, R., et al. (2023) Masked Swin Transformer Unet for Industrial Anomaly Detection. *IEEE Transactions on Industrial Informatics*, **19**, 2200-2209. https://doi.org/10.1109/tii.2022.3199228
- [16] McMillan, M., Haber, E., Peters, B. and Fohring, J. (2021) Mineral Prospectivity Mapping Using a VNet Convolutional Neural Network. *The Leading Edge*, 40, 99-105. <u>https://doi.org/10.1190/tle40020099.1</u>
- [17] Cheng, L., Yi, J., Chen, A. and Zhang, Y. (2023) Fabric Defect Detection Based on Separate Convolutional UNet. *Multimedia Tools and Applications*, 82, 3101-3122. <u>https://doi.org/10.1007/s11042-022-13568-7</u>
- [18] Chen, L.C., Yang, Y., Wang, J., Xu, W. and Yuille, A.L. (2016) Attention to Scale: Scale-Aware Semantic Image Segmentation. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 27-30 June 2016, 3640-3649. <u>https://doi.org/10.1109/CVPR.2016.396</u>
- [19] Huang, Q., Xia, C., Wu, C., Li, S., Wang, Y., Song, Y. and Kuo, C.C.J. (2017) Semantic Segmentation with Reverse Attention. arXiv: 1707.06426. <u>https://doi.org/10.48550/arXiv.1707.06426</u>
- [20] Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z. and Lu, H. (2019) Dual Attention Network for Scene Segmentation. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, 15-20 June 2019, 3146-3154. <u>https://doi.org/10.1109/CVPR.2019.00326</u>
- [21] Yuan, Y., Huang, L., Guo, J., Zhang, C., Chen, X. and Wang, J. (2018) OCNet: Object Context Network for Scene Parsing. arXiv:1809.00916. <u>https://doi.org/10.48550/arXiv.1809.00916</u>
- [22] Li, X., Zhong, Z., Wu, J., Yang, Y., Lin, Z. and Liu, H. (2019) Expectation-Maximization Attention Networks for Semantic Segmentation. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, 27 October-2 November 2019, 9167-9176. <u>https://doi.org/10.1109/ICCV.2019.00926</u>
- [23] Huang, Z., Wang, X., Huang, L., Huang, C., Wei, Y. and Liu, W. (2019) CCNet: Criss-Cross Attention for Semantic Segmentation. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, 27 October-2 November 2019, 603-612. <u>https://doi.org/10.1109/ICCV.2019.00069</u>
- [24] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S. and Guo, B. (2021) Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. Proceedings of the 2021 *IEEE/CVF International Conference on Computer Vision* (*ICCV*), Montreal, 10-17 October 2021, 9992-10002. <u>https://doi.org/10.1109/ICCV48922.2021.00986</u>
- [25] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., et al. (2023) Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. Computer Vision-ECCV 2022 Workshops, Tel Aviv, 23-27 October 2022, 205-218. https://doi.org/10.1007/978-3-031-25066-8_9
- [26] Muhammad, K., Hussain, T., Del Ser, J., Palade, V. and de Albuquerque, V.H.C. (2020) DeepReS: A Deep Learning-Based Video Summarization Strategy for Resource-Constrained Industrial Surveillance Scenarios. *IEEE Transactions* on Industrial Informatics, 16, 5938-5947. <u>https://doi.org/10.1109/TII.2019.2960536</u>
- [27] Tan, S.H., Chuah, J.H., Chow, C. and Kanesan, J. (2023) Coarse-to-Fine Context Aggregation Network for Vehicle Make and Model Recognition. *IEEE Access*, 11, 126733-126747. <u>https://doi.org/10.1109/access.2023.3330114</u>
- [28] Liu, Y., Zhang, Z., Liu, X., Wang, L. and Xia, X. (2021) Efficient Image Segmentation Based on Deep Learning for Mineral Image Classification. Advanced Powder Technology, 32, 3885-3903. <u>https://doi.org/10.1016/j.apt.2021.08.038</u>
- [29] Huang, Y., Jing, J. and Wang, Z. (2021) Fabric Defect Segmentation Method Based on Deep Learning. IEEE Transactions on Instrumentation and Measurement, 70, 1-15. <u>https://doi.org/10.1109/TIM.2020.3047190</u>

- [30] Shi, J., Li, Z., Zhu, T., Wang, D. and Ni, C. (2020) Defect Detection of Industry Wood Veneer Based on NAS and Multi-Channel Mask R-CNN. Sensors, 20, Article 4398. <u>https://doi.org/10.3390/s20164398</u>
- [31] Song, Y., Xia, W., Li, Y., Li, H., Yuan, M. and Zhang, Q. (2024) AnomalySeg: Deep Learning-Based Fast Anomaly Segmentation Approach for Surface Defect Detection. *Electronics*, 13, Article 284. <u>https://doi.org/10.3390/electronics13020284</u>