



# 先验知识辅助的金属涂层损伤分割方法

谢洲洋<sup>1</sup>, 舒畅<sup>2</sup>, 傅彦<sup>1,3</sup>, 周俊临<sup>1,3</sup>, 蒋家玮<sup>1</sup>, 陈端兵<sup>1,3\*</sup>

(1. 成都数之联科技有限公司, 成都 610094; 2. 西南技术工程研究所, 重庆 400039; 3. 电子科技大学 大数据研究中心, 成都 611731)

**摘要** 金属涂层损伤的自动识别是一项具有重大实际应用价值的研究。随着深度学习在各类材质的表面损伤图像分割任务中取得突破性进展, 大部分研究工作使用端到端的深度卷积神经网络分割模型分割损伤区域。然而, 端到端深度学习方法很难利用金属涂层损伤相关的先验知识来识别损伤尺度差异大、训练数据不足等问题。因此, 设计了一种基于先验知识的金属涂层损伤区域分割算法, 结合深度学习分割模型 U-Net 实现金属涂层损伤的自动识别。该算法基于 Hue 通道分布和边缘响应对金属涂层图像中的异常区域进行分割, 深度学习分割模型利用先验知识排除异常样本, 可有效避免过拟合。在含有开裂、起泡、生锈和脱落这 4 类损伤的金属涂层数据集中, 该算法取得了 81.24% mIoU 的分割效果, 优于端到端的深度学习方法。实验结果表明, 先验知识辅助能够有效提升深度学习分割模型对金属涂层损伤的分割效果。

**关键词** 深度学习; 损伤识别; 图像分割; 先验知识

中图分类号 TP391

文献标志码 A

DOI 10.12178/1001-0548.2022373

## Knowledge-Driven Metal Coating Defect Segmentation

XIE Zhouyang<sup>1</sup>, SHU Chang<sup>2</sup>, FU Yan<sup>1,3</sup>, ZHOU Junlin<sup>1,3</sup>, JIANG Jiawei<sup>1</sup>, and CHEN Duanbing<sup>1,3\*</sup>

(1. Chengdu Union Big Data Technology Incorporated, Chengdu 610094, China;

2. Southwest Technology and Engineering Research Institute, Chongqing 400039, China;

3. Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, China)

**Abstract** Automatic recognition of metal coating defects has significant value in realistic applications. As deep learning makes breakthrough in surface defect segmentation for a variety of materials, most of deep convolutional neural network segmentation models are trained in an end-to-end manner. However, it is difficult to exploit prior knowledge about metal coating defects in end-to-end deep learning and adapt to the variable scale of the defects and the limited training data. This paper proposes a defect segmentation algorithm based on prior knowledge about metal coating defects to unify U-Net, a deep learning segmentation model for automatic metal coating defect recognition. This anomaly segmentation is based on Hue channel distribution and edge response. Being trained in a knowledge driven manner, the model can exclude outliers from training data and effectively avoid over-fitting. On a metal coating defect image dataset with four defect types, including crack, blister, rusting and flaking, the proposed method achieves 81.24% mIoU, which is advantageous over end-to-end deep learning. The experiment shows that knowledge-driven model can boost the performance of deep learning models in metal coating defect segmentation.

**Key words** deep learning; defect recognition; image segmentation; prior knowledge

金属涂层是一种覆盖在金属材料上的防护层, 通过相对隔绝金属材料与外界环境因素 (如光照、氧气、腐蚀性物质等), 达到防锈、防腐、防破损等目的。金属涂层被广泛应用在电力、航空航天、装备制造、船舶制造等重要领域, 对铁基金属材料和有色金属材料 (如铝、铜、锡、镍等) 提供保

护。然而, 涂层本身也会因为各种因素产生损伤, 进而将其保护的金属材料暴露在环境中。因此, 金属涂层损伤的自动识别对避免安全隐患乃至人员财产损失具有重要意义。

目前, 针对各种材料的表面损伤识别问题已有较多研究。文献 [1] 在绝缘子缺陷检测任务中使用

收稿日期: 2022-10-28; 修回日期: 2022-12-14

基金项目: 国家自然科学基金 (61673085); 西南技术工程研究所合作基金 (HDHDW5902010301)

作者简介: 谢洲洋, 主要从事计算机视觉方面的研究。

\*通信作者 E-mail: dbchen@uestc.edu.cn

去噪自编码器来寻找异常区域并识别。文献 [2] 针对表面损伤识别问题中广泛存在的样本量小的问题, 提出了使用图像生成模型扩充样本的方法。文献 [3] 总结了基于统计分布、谱特征和机器学习 3 类织物表面损伤识别方法。文献 [4-5] 提出了以损伤分割为先验信息, 辅助损伤类别分类器学习的模型结构。文献 [6] 提出的基于残差结构、空间金字塔池化改进 U-Net, 以及文献 [7] 提出的基于“压缩-激活”模块改进 U-Net, 在路面损伤分割问题上取得了良好效果。文献 [8] 提出了一种混合监督信息的弱监督学习框架, 以适应表面损伤识别任务中标注难度大、标注样本少的问题。

目前针对金属涂层损伤识别的研究鲜有报道。涂层损伤相比其他类型的表面缺陷有一定特殊性。首先, 涂层损伤的起泡、开裂、脱落等损伤类型的视觉特征与常见的腐蚀性金属损伤有很大差异; 其次, 涂层损伤的目标尺度范围非常大, 既有小于 1 mm 的开裂, 也有几毫米宽的锈点和几厘米宽的脱落和起泡; 最后, 涂层表面除了上述待识别的损伤外, 一般不包含其他杂质和痕迹, 即背景几乎没有混杂, 这是涂层损伤识别任务中值得利用的先验知识。直接使用端到端 (End-to-End) 的深度卷积神经网络 (Deep Convolutional Neural Network) 分割模型, 既不能利用关于涂层损伤的先验知识, 也较难处理训练样本量小、模型设计难以处理不同尺度损伤、容易过拟合等问题。

因此, 本文提出了一种图像异常区域分割算法辅助深度卷积神经网络分割模型进行训练和预测, 实现金属涂层损伤的自动识别。本文利用涂层损伤的先验知识, 设计了一种基于计算机视觉技术的图像异常分割算法, 实现损伤区域的类别无关的图像分割, 并指导分割模型进行预测; 此外, 借鉴 YOLO 检测模型<sup>[9]</sup>, 本文设计了一种分割模型损失函数, 并利用先验知识指导分割模型的训练。在一个包含开裂、起泡、脱落、生锈 4 类损伤的涂层损伤数据集上验证了本文算法的有效性。

## 1 方法

### 1.1 算法设计思路

在常见的图像分割数据集 (如 PASCAL VOC<sup>[10]</sup> 和 COCO<sup>[11]</sup>) 中, 已验证深度卷积神经网络比传统的图像分割算法具有更强的拟合和泛化能力。相比传统的计算机视觉方法, 深度学习能自适应地学习有效的视觉特征, 并拥有更强的分类判别能

力。但是, 大多数深度图像分割模型致力于对自然图像中较大尺度的物体进行分割; 虽然现有的一些工作致力于在特殊场景或不同尺度物体上训练分割模型, 但都需要大规模数据的训练, 且难以利用先验知识。相对地, 传统的计算机视觉方法能较好地利用先验知识, 人为设计出针对性较强的特征提取或分割算法。因此, 本文利用金属涂层损伤的先验知识, 设计有针对性的分割算法, 并以此辅助深度学习模型进行训练和预测。本文提出的金属涂层损伤识别算法包括异常区域先验分割和深度分割模型两部分。一方面, 异常区域先验分割算法利用金属涂层损伤的先验知识, 分割出图像中所有疑似损伤的区域, 但不能对其损伤类型进行分类, 也不能基于标注数据进行自适应优化。另一方面, 本文使用 U-Net<sup>[12]</sup> 模型, 在标注数据上进行端到端的分割任务训练和预测。训练时, 深度分割模型被损伤标注和先验分割的结果同时监督; 预测时, 深度分割模型的预测被先验分割的结果加强。图 1 为本文所提方法的流程图, 展示了异常区域分割结果对模型输出的训练监督和辅助预测过程。

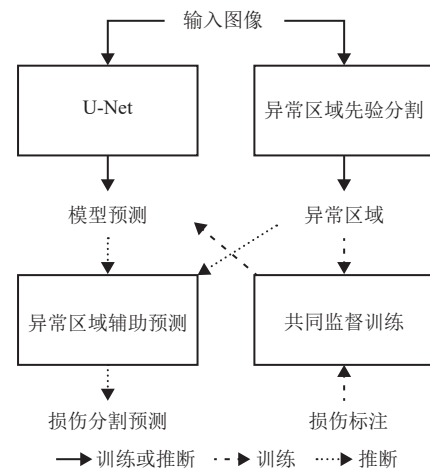


图 1 算法流程图

### 1.2 基于 Hue 分布的异常区域先验分布

对脱落和生锈这两类损伤, 颜色特征是区分损伤区域的重要依据。未损伤的金属涂层表面往往只包含单一的颜色, 很少包含杂质, 而脱落或生锈的区域会呈现出其他颜色。因光照不均匀或阴影, 有时损伤或未损伤区域的表面会产生亮度、饱和度不一致, 但像素的颜色仍具有一致性。图 2 中的示例图展示了这种特点。HSL 颜色空间是一种能有效分离像素中色彩 (Hue)、饱和度 (Saturation) 和亮度 (Lightness) 等成分的一种颜色表征方式。其中, Hue 维度是一个弧度值, 表示不同的颜色成

分,能很大程度避免因光照强度、光照角度、阴影而导致的像素值变化。因此,本文使用图像中像素值在 Hue 维度上的分布来分割异常区域。

本文基于多阈值分割算法 (Multi-level Otsu's Method)<sup>[13]</sup> 在 Hue 通道弧度值空间进行阈值搜索。首先,对图像进行高斯平滑过滤以排除噪声干扰,然后统计所有像素的 Hue 通道直方图,最后,对直方图执行多阈值分割算法。该算法在区间  $[0, 2\pi)$  内搜索  $N$  个递增的阈值  $\{t_1, t_2, \dots, t_N\}$ , 从而将  $[0, 2\pi)$  划分为  $N$  个区间。由于 Hue 通道像素是弧度值,当  $i=N$  时,区间为  $[t_N, 2\pi) \cup [0, t_1)$ 。算法的目标是类内方差最小,即式 (1)<sup>[13]</sup> 中的优化函数:

$$\min_{t_i} \sum_{i=1}^N w(t_i) \sigma^2(t_i) \quad (1)$$

式中,  $w(t_i)$  是 Hue 通道直方图在区间  $[t_i, t_{i+1})$  内分布的和;  $\sigma^2(t_i)$  为此区间内分布的方差。找到阈值后,取其中分布占比最大的前  $m$  个区间的像素点作为无异常区域,其他区域被分割为异常区域。

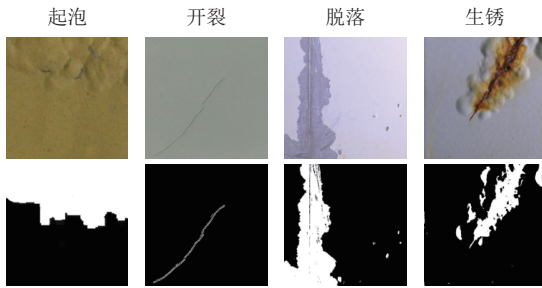


图 2 异常区域分割示例

### 1.3 基于边缘响应的异常区域先验分割

起泡和开裂两类是比较细微的金属涂层损伤,它们通常不会带来大面积的颜色变化或可辨别的纹理,但却会产生显著的边缘特征响应。开裂损伤以细微线条的形式呈现,线条的边缘会产生空间连续的、细小的边缘特征;起泡损伤的膨胀处与未损伤的平整处的光照反射角度不同,因此会在起泡的边界产生边缘特征。

本文使用经典的 Canny 边缘检测<sup>[14]</sup> 来寻找连续的、显著边缘响应。Canny 边缘检测算法的输出是一个边缘响应图  $E_i, i \in [0, 1]$ , 代表位于坐标点  $(i, j)$  的像素的边缘特征响应。本文通过设定一个边缘响应阈值  $\tau \in (0, 1)$ , 使边缘响应最高的像素被识别为异常区域。最后,由于起泡损伤的实际损伤区域是边缘轮廓所包围的区域,因此对起泡损伤的分割结果增加一步后处理,对边缘响应图进行最大值

滤波来“填充”其轮廓内部的区域。

### 1.4 先验辅助的分割模型训练与预测

本文提出的分割模型训练算法的核心思想是借助先验知识来减弱训练样本的标签噪声。由于标注人员知识和操作的不确定性,涂层损伤图像的像素级标注 (布尔值标签  $D_{\text{label}}$ ) 会存在一定的噪声,有:

$$D_{\text{label}} = \begin{cases} \widehat{D}_{\text{label}} & \varepsilon_{\text{label}} = 0 \\ !\widehat{D}_{\text{label}} & \varepsilon_{\text{label}} = 1 \end{cases} \quad (2)$$

$$\varepsilon_{\text{label}} \sim \text{Bernoulli}(p_{\text{label}})$$

如式 (2) 所示,真实的布尔值标签常量  $\widehat{D}_{\text{label}}$  服从伯努利分布的噪声  $\varepsilon_{\text{label}}$  干扰;噪声产生错误标签的概率为  $p_{\text{label}}$ 。而异常区域先验分割结果  $D_{\text{prior}}$  也具有相似的不确定性,有:

$$D_{\text{prior}} = \begin{cases} \widehat{D}_{\text{prior}} & \varepsilon_{\text{prior}} = 0 \\ !\widehat{D}_{\text{prior}} & \varepsilon_{\text{prior}} = 1 \end{cases} \quad (3)$$

$$\varepsilon_{\text{prior}} \sim \text{Bernoulli}(p_{\text{prior}})$$

但是,由于先验分割与人工标注过程相互独立,本文假设两者的标签噪声  $\varepsilon_{\text{label}}$  和  $\varepsilon_{\text{prior}}$  相互独立。因此,如果从训练集中排除先验分割与人工标注标签不同 (即  $D_{\text{label}}$  与  $D_{\text{prior}}$  不等) 的样本,则训练集中标签噪声  $\varepsilon$  为:

$$\varepsilon \sim \text{Bernoulli}\left(\frac{p_{\text{label}} p_{\text{prior}}}{p_{\text{label}} p_{\text{prior}} + (1 - p_{\text{label}})(1 - p_{\text{prior}})}\right) \quad (4)$$

易证明,此时标签噪声  $\varepsilon$  为 1 的概率小于  $p_{\text{label}}$ 。用先验分割的结果过滤不确定性较高的人工标签,能减小模型训练中标签噪声的干扰。

本文使用基于 Hue 分布和基于边缘响应的异常区域先验分割辅助深度学习分割模型进行训练和预测。假设图像中某点  $i$  的  $C$  类型损伤 (起泡、开裂、脱落、生锈等) 的深度学习模型分割结果概率值为  $y \in [0, 1]$ , 本文提出的训练和预测方式如下。

训练时,当  $D_{\text{label}}$  和  $D_{\text{prior}}$  同时取值为真,则设可信标签  $y$  为 1; 当  $D_{\text{label}}$  和  $D_{\text{prior}}$  同时取值为假,则设可信标签  $y$  为 0。本文使用交叉熵损失函数训练分割模型。此时点  $i$  处模型预测值  $y_i$  的损失函数为:

$$L(i) = -[y_i \log(y_i) + (1 - y_i) \log(1 - y_i)] \quad (5)$$

而当  $D_{\text{label}}$  和  $D_{\text{prior}}$  的真值不同时,点  $i$  的损失函数值为 0,回传的梯度值也将为 0。算法 1 给出了计算点  $i$  处损失函数值的伪代码。

算法 1 计算损失函数值

输入: 点  $i$  的损伤标注  $D_{\text{label}} \in \{\text{True}, \text{False}\}$ , 分割模型在点  $i$  的概率输出  $y_i \in [0, 1]$

输出: 损失函数值

计算点  $i$  的异常区域先验分割结果  $D_{\text{prior}} \in \{\text{True}, \text{False}\}$

if  $D_{\text{label}} \&\& D_{\text{prior}}$

$\widehat{y}_i \leftarrow 1$

else if  $\neg D_{\text{label}} \&\& \neg D_{\text{prior}}$

$\widehat{y}_i \leftarrow 0$

else

return 0

loss  $\leftarrow -[\widehat{y}_i \log(y_i) + (1 - \widehat{y}_i) \log(1 - y_i)]$

return loss

模型预测时, 仅当  $C$  类型的概率输出  $y_i$  为各类型的最大值, 且  $C$  类型的  $D_{\text{prior}}$  为真时, 才将点  $i$  预测为  $C$  类型损伤; 否则, 在  $D_{\text{prior}}$  与模型预测发生冲突的情况下, 将此点预测为无损伤。

本文选用 U-Net 来训练金属涂层损伤分割模型。U-Net 是一种在医学影像病变识别<sup>[15]</sup> 和材料损伤识别<sup>[16]</sup> 等领域广泛应用的图像分割模型。与其他常见的用于识别自然图像中宏观物体的分割模型 (如 FCN<sup>[17]</sup>、SegNet<sup>[18]</sup>、PSPNet<sup>[19]</sup>、DeepLab<sup>[20]</sup> 等) 相比, U-Net 具有独特的优势。如图 3 所示, U-Net 中编码器的各层输出直接与解码器连接, 使得模型有机会直接使用高分辨率的底层视觉信息来辅助分割预测。这一优势在医学图像分割、材料损伤分割和本文的金属涂层损伤分割任务中都非常重要, 因为它们都包含了比较细微的分割目标, 且往往依赖于底层的视觉信息来判别。另外, U-Net 使用卷积层代替池化层来扩大感受野, 能有效避免局部细节信息的丢失。

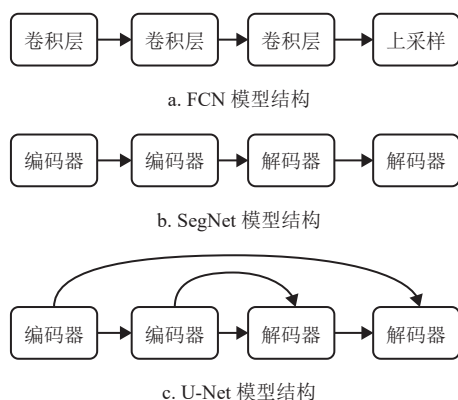


图 3 分割模型结构对比

## 1.5 方法对比

本文提出的训练和预测方式借鉴了目标检测模型 YOLO 的训练和预测方式。它们都使用一种可

靠度较高的方式来过滤大量的预测结果 (YOLO 过滤大量的密集检测框, 而本文过滤大量的像素级损伤预测结果)。不同的是, YOLO 使用一个后验预测的、类别无关的、可训练的似物性分数 (Objectness) 对预测结果进行过滤, 而本文采用基于先验知识的、特定类别的分割结果来达到相同目的。本文提出的训练方式能够找到先验知识和人工标注之间的冲突, 排除异常的损伤区域, 从而增强模型的泛化性能。采用此种先验知识辅助的训练方式时, 人工标注为异常区域先验分割排除了噪声干扰; 另一方面, 先验知识将人工标注中异常的、难以学习的样本剔除。两种信息共同监督模型训练, 使得模型专注于拟合稳定的、可泛化的样本, 而忽略异常样本。本文提出的预测方式延续了这一思路, 排除先验知识和模型预测的冲突, 只保留置信度最高的像素预测结果。

本文提出的模型训练方式对应的损失函数是以交叉熵损失函数的形式。它主要基于先验知识对训练样本进行选择, 减小标签噪声对训练的影响。与常用的加权交叉熵 (Weighted Cross Entropy, WCE) 损失函数相比, WCE 主要用于对正负样本加权, 不涉及对标签噪声的处理。与常用的 Focal Loss 相比, Focal Loss 对难样本动态地增加权重, 可能反而会增强标签噪声的影响, 导致难以收敛或过拟合。

在深度学习语义分割模型的研究中, 一些方法利用了空间先验信息改造模型结构, 引导模型学习空间上下文联系, 如 Context Prior<sup>[21]</sup>、EncNet<sup>[22]</sup> 等方法。这类方法并不显式地将领域知识编码进模型, 而是引导模型挖掘空间关系, 因此与本文提出的方法有较大区别。文献 [23-24] 提出利用全局信息 (如场景、物体类型、文本描述等) 对模型训练进行监督或对预测结果进行后处理。而本文提出的方法与它们的区别在于, 利用局部的图像特征信息为语义分割模型提供了更细粒度的监督信息, 并同时作用于训练和预测阶段。

## 2 实验与结果分析

### 2.1 实验数据

本文使用西南技术工程研究所采集的金属涂层损伤图像数据测试本文提出的金属涂层损伤分割算法的有效性。此数据集包含了 715 张金属涂层图像, 涵盖了起泡、开裂、生锈和脱落 4 类损伤。每张图像包含一个金属涂层材料样本, 每个样本可能存在多种损伤、单一类型损伤或无损伤。其中

302 张包含起泡损伤、175 张包含开裂损伤、280 张包含生锈损伤、211 张包含脱落损伤、210 张不包含任何损伤。本文随机选取数据集中的 100 张金属涂层图像作为测试集，其中 80 张包含损伤（大多数图像含两种，总体损伤类型分布均衡），另有 20 张不含损伤。剩余所有图像为训练集。

由于图像分辨率较高（大多为 100 万左右像素）且金属涂层损伤具有局部性（从局部观察也可以进行损伤分类、分割），所以本文在实际的模型训练和测试中将图像作如下处理：先使用最邻近插值将图像变形，使其长宽均为 256 的倍数；再将变形后的图像切分成若干个边长为 256 像素的正方形图像；对损伤标注做相同的变形和划分；将每个划分出的小图像和标注作为一个实例。这样，本文实际上在原有数据基础上划分出了约 10 000 个实例。

此外，为了广泛对比本文方法与其他深度学习模型在公开数据集的语义分割性能，本文使用了 MagneticTile 数据集<sup>[25]</sup>进行实验。MagneticTile 数据集包含了孔隙、开裂、磨损、不平整等损伤。虽然与涂层损伤的名称不同，但其表现损伤特征有相似之处。如 MagneticTile 的开裂与涂层开裂损伤都表现为裂纹；MagneticTile 的空洞、磨损等与涂层的生锈、脱落相似，都表现为与背景颜色差异较大的连续块状区域；MagneticTile 的不均匀损伤与起泡相似。在 MagneticTile 数据集的实验将验证本文为金属涂层损伤设计的异常区域先验分割方法能适用于其他具有相似图像特征的表面缺陷损伤，且与其他表面缺陷分割模型相比具有良好的识别效果。数据集划分比例为 50% 训练集，20% 验证集和 30% 测试集。

## 2.2 实验设置

对金属涂层损伤语义分割任务，本文主要使用 Mean Intersection-over-Union (mIoU) 作为损伤分割的评价指标。在  $N$  个实例上，其计算方式如下：

$$\text{mIoU} = \frac{1}{N} \sum_{i=1}^N \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i + \text{FN}_i} \quad (6)$$

另外，本文也使用准确率和召回率具体分析异常区域先验分割算法的特性：

$$\text{Precision} = \frac{1}{N} \sum_{i=1}^N \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \quad (7)$$

$$\text{Recall} = \frac{1}{N} \sum_{i=1}^N \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \quad (8)$$

式 (6)~式 (8) 中，TP、FP、FN 分别代表同一实例中真阳性、假阳性、假阴性预测的数量。

本文中各个深度学习分割模型的训练使用随机梯度下降算法 (Stochastic Gradient Descent, SGD) 和 Adam 优化算法。训练的初始学习率为 0.001，衰减系数为 10，每 50 个世代衰减一次，世代数为 200，mini-batch 大小为 64。

本文的实验硬件环境包括 2 个 16 GB 显存的 GPU 和一个 2.2 GHz 的 CPU，软件环境包括基于 PyTorch 实现的深度卷积神经网络和基于 OpenCV 实现的异常区域先验分割算法。另外，本文使用随机旋转和随机翻转作为数据增强手段，从而增强模型的泛化性能。这两种数据增强手段都不会影响基于 Hue 分布和基于边缘响应两种异常区域先验分割算法。

## 2.3 异常区域先验分割结果

对基于 Hue 分布的生锈、脱落损伤的异常分割，当选取的非异常区间数量  $m$  提高时，更多的区域会被识别为异常。一般而言，这会导致对损伤的召回率提高、准确率下降。而对基于边缘响应的开裂、起泡损伤的异常分割，当边缘响应阈值  $\tau$  降低，也会导致召回率提高、准确率下降。由于异常区域分割结果将与损伤标注共同监督模型训练，选择参数的依据是保证异常区域召回率较高，且此区间内准确率变化较小，以保证结果较为稳定。

表 1 的实验结果展示了不同参数设置下异常区域先验分割对损伤区域的识别准确率和召回率。其中， $N$  和  $m$  为基于 Hue 通道的异常区域先验分割的阈值数量和异常区间数量， $\tau$  为边缘响应的阈值。表 1 实验中，先设置了区间数量  $m/N$ ，再通过实验结果选取了最优的响应阈值  $\tau$ ，并报告其组合下的平均准确率和召回率。

表 1 不同参数设置下异常区域先验分割的分割效果

| 区间数量 ( $m/N$ ) | 响应阈值 ( $\tau$ ) | 平均准确率 | 平均召回率 |
|----------------|-----------------|-------|-------|
| 1/6            | 0.5             | 0.535 | 0.924 |
| 2/6            | 0.55            | 0.476 | 0.948 |
| 1/12           | 0.45            | 0.603 | 0.872 |
| 2/12           | 0.6             | 0.643 | 0.867 |
| 3/12           | 0.5             | 0.494 | 0.904 |

实验发现，基于先验的异常区域分割算法对损伤区域有较好的召回率，但因为噪声和多种损伤的干扰，准确率不佳。其中起泡和开裂这两种较为细微的损伤更难被准确识别，容易被其他带边缘响应

的噪声或无关物体干扰。

最终, 本文选取异常区域先验分割的参数分别为: 基于 Hue 通道的异常区域先验分割中使用  $N=6$  个阈值, 取占比最高的  $m=2$  个区间; 基于边缘响应的异常区域先验分割  $\tau=0.55$ 。选定参数后异常区域先验分割的结果如表 2 所示。

表 2 参数选定后异常区域先验分割的分割效果

| 评价指标 | 起泡    | 生锈    | 开裂    | 脱落    |
|------|-------|-------|-------|-------|
| 准确率  | 0.474 | 0.503 | 0.432 | 0.493 |
| 召回率  | 0.944 | 0.949 | 0.939 | 0.958 |
| mIoU | 0.492 | 0.531 | 0.460 | 0.513 |

## 2.4 金属涂层损伤分割结果

在 U-Net、FCN、SegNet、DeepLab 和 PSPNet 等模型上, 本文验证了基于异常区域先验分割的训练方式的有效性。如表 3 所示, 含有“+prior”标记的模型使用了本文提出的基于异常区域先验分割的训练方式, 其他模型使用常用交叉熵损失函数训练。本文提出的先验知识辅助 U-Net 分割训练与预测的方法达到了 81.24% mIoU 的分割效果, 实现了对金属涂层损伤的准确分割。本文提出的先验知识辅助方法能够在多种分割模型上实现分割效果提升, 在 FCN 和 SegNet 上, mIoU 提升到了 65% 以上。另一方面, 在自然图像语义分割问题上展现出较高性能的 DeepLab 和 PSPNet 并没有在金属涂层损伤图像上取得较好的效果, 但基于先验分割的训练对 mIoU 仍有近 10% 的提升。

表 3 金属涂层损伤分割结果

| 方法             | mIoU  |       |       |       |       | 时延/s |
|----------------|-------|-------|-------|-------|-------|------|
|                | 起泡    | 生锈    | 开裂    | 脱落    | 平均    |      |
| U-Net + prior  | 76.64 | 87.31 | 75.32 | 85.71 | 81.24 | 0.74 |
| U-Net          | 67.54 | 74.92 | 61.31 | 73.03 | 69.20 | 0.53 |
| FCN + prior    | 70.91 | 61.77 | 63.59 | 69.10 | 66.34 | 0.51 |
| FCN            | 46.84 | 57.04 | 43.83 | 64.78 | 53.12 | 0.33 |
| SegNet + prior | 67.50 | 63.00 | 64.10 | 65.89 | 65.12 | 0.43 |
| SegNet         | 44.12 | 59.20 | 51.21 | 61.33 | 53.97 | 0.18 |
| DeepLab +prior | 60.98 | 58.38 | 70.00 | 65.56 | 63.73 | 1.48 |
| DeepLab        | 40.98 | 55.82 | 59.64 | 62.02 | 54.62 | 1.23 |
| PSPNet +prior  | 59.26 | 55.35 | 62.32 | 67.88 | 61.20 | 1.33 |
| PSPNet         | 43.60 | 51.88 | 57.70 | 63.31 | 54.12 | 0.93 |

结合 1.4 节对模型结构的分析, U-Net 取得更好结果的原因, 来自于底层编码器与解码器的连接, 使模型能更好地捕获底层视觉特征。而其他四类模型都缺少这一特性。这一特性对涂层损伤分割

问题非常重要, 因为损伤的准确分割依赖于对颜色、边缘、纹理等底层特征的学习, 而较少利用高层语义信息。另一方面, DeepLab 和 PSPNet 的空洞卷积和空间金字塔池化主要是为了增强高层语义特征的空间分布, 但又加深了底层特征与模型输出层间的距离。对涂层损伤分割而言, 这一特性有可能反而损害了性能。

表 3 同时也对比了加入和不加入先验信息辅助条件下推理速度的差异。可以发现, 先验信息辅助本身涉及的边缘检测、阈值分割、预测结果后处理等计算引入了约 0.2 s 的时延, 且与模型本身无关, 不会构成显著的性能瓶颈。

图 4 对比了训练 U-Net 时使用和不使用先验知识辅助对模型收敛的影响。可以发现, 先验知识辅助下的训练收敛更快、更平稳。不使用先验信息时, U-Net 收敛时的损失函数值仍然高于 0.2, 且波动较大; 使用先验信息时, U-Net 收敛损失函数值小于 0.1, 且更为稳定。此实验结果表明, 先验知识除了能提高模型泛化性能, 也能降低数据拟合难度、提高训练效率。

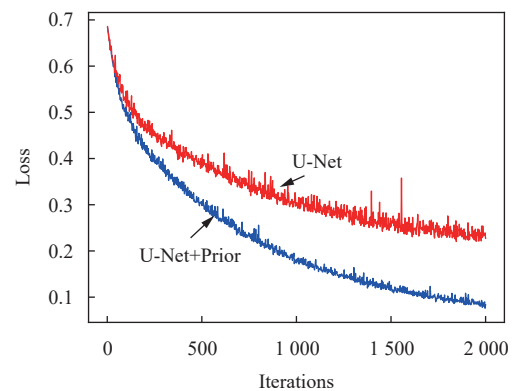


图 4 训练时先验知识辅助对损失函数收敛的影响

## 2.5 MagneticTile 数据集实验结果

表 4 对比了本文提出的方法 (U-Net 结合先验知识辅助) 与 U-Net、DeepLabV3<sup>[26]</sup>、HRNet<sup>[27]</sup>、FastFCN<sup>[28]</sup> 等语义分割模型在 MagneticTile 数据集上的损伤分割结果。由于对比方法的原论文并没有在 MagneticTile 数据集上进行实验, 因此本文使用作者公布的模型、代码和超参数设置在 MagneticTile 数据集上进行训练, 训练方式与 2.2 节中介绍的本文方法的训练方式相同。

实验显示, 异常区域先验分割虽然主要针对金属涂层的损伤而设计, 但对其他具有相似先验知识的表面缺陷仍有一定的效果, 比原始的 U-Net 具

有 8% 的 mIoU 提升。对比其他通常用于自然图像语义分割的模型, 本文方法具有一定的优势, 特别是对孔隙和磨损等样本量不丰富但可以借助先验知识的损伤类型。

表 4 MagneticTile 损伤分割准确率对比

| 方法            | 孔隙   | 开裂   | 磨损   | 不平整  | 平均   |
|---------------|------|------|------|------|------|
| U-Net + Prior | 86.2 | 75.3 | 69.4 | 62.5 | 73.4 |
| U-Net         | 72.4 | 68.0 | 62.3 | 59.1 | 65.4 |
| DeepLabV3     | 62.7 | 54.3 | 45.2 | 50.2 | 53.1 |
| HRNet         | 59.1 | 76.8 | 65.0 | 58.7 | 64.9 |
| FastFCN       | 62.9 | 64.7 | 54.6 | 64.0 | 61.6 |

### 3 结束语

本文基于计算机视觉和先验知识设计的金属涂层异常区域先验分割算法, 辅助深度学习分割模型, 能有效提高各类深度学习分割模型对金属涂层损伤的分割效果。本文选取的深度学习分割模型 U-Net 在先验知识的辅助下, mIoU 提高了 12.04%。本文提出的方法也可以应用到医学图像病变区域分割、其他材料的损伤区域分割端到端深度学习模型中。

### 参考文献

- [1] KANG G, GAO S, YU L, et al. Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 68(8): 2679-2690.
- [2] WEI T, CAO D, ZHENG C, et al. A simulation-based few samples learning method for surface defect segmentation[J]. Neurocomputing, 2020, 412: 461-476.
- [3] KUMAR A. Computer-Vision-Based fabric defect detection: A survey[J]. IEEE Transactions on Industrial Electronics, 2008, 55(1): 348-363.
- [4] HUANG Y, JING J, WANG Z. Fabric defect segmentation method based on deep learning[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-15.
- [5] TABERNIK D, SELA S, SKVARC J, et al. Segmentation-Based deep-learning approach for surface-defect detection[J]. Journal of Intelligent Manufacturing, 2020, 31: 759-776.
- [6] AUGUSTAUSKAS R. Improved pixel-level pavement-defect segmentation using a deep autoencoder[J]. Sensors, 2020, 20(9): 2557-2578.
- [7] MIAO X, WANG J, WANG Z, et al. Automatic recognition of highway tunnel defects based on an improved U-Net model[J]. IEEE Sensors Journal, 2019, 99: 11413-11423.
- [8] BOZIC J, TABERNIK D, SKOCAJ D. Mixed supervision for surface-defect detection: From weakly to fully supervised learning[J]. Computers in Industry, 2021, 129: 103459-103472.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 779-788.
- [10] EVERINGHAM M, ZISSERMAN A, WILLIAMS C, et al. The 2005 PASCAL visual object classes challenge [C]//Proceedings of Machine Learning Challenges Workshop. [S.l.]: Springer, 2005: 117-176.
- [11] LIN T, MARIE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer International Publishing, 2014: 740-755.
- [12] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer International Publishing, 2015: 234-241.
- [13] LIAO P, CHEN T, CHUNG P. A fast algorithm for multilevel thresholding[J]. Journal of Information Science and Engineering, 2001, 17: 713-727.
- [14] ELDER J H, ZUCKER S W. Local scale control for edge detection and blur estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(7): 699-716.
- [15] HENRY T, CARRE A, LEROUSSIEU M, et al. Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-net neural networks: A BraTS 2020 challenge solution[C]//International MICCAI Brainlesion Workshop. [S.l.]: Springer, 2020: 327-339.
- [16] SOUKUP D, HUBER-MRK R. Convolutional neural networks for steel surface defect detection from photometric stereo images[C]//International Symposium on Visual Computing. Berlin: Springer, 2014: 668-677.
- [17] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(12): 2481-2495.
- [18] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [19] ZHAO H, SHI J, QI, X, et al. Pyramid scene parsing network[C]//IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 6230-6239.
- [20] CHEN L C, PAPANDREOU G, KOKKINOS I, et al.

- DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [21] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2019: 3146-3154.
- [22] ZHANG H, DANA K, SHI J, et al. Context encoding for semantic segmentation[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2018: 7151-7160.
- [23] 刘昕, 袁家斌, 王天星. 基于场景先验知识的室内人体行为识别方法[J]. *计算机科学*, 2022, 49(1): 225-232.
- LIU X, YUAN J B, WANG T X. Interior human action recognition method based on prior knowledge of scene[J]. *Computer Science*, 2022, 49(1): 225-232.
- [24] ZHENG Z, YAN H, SETZER F C, et al. Anatomically constrained deep learning for automating dental CBCT segmentation and lesion detection[J]. *IEEE Transactions on Automation Science and Engineering*, 2020, 18(2): 603-614.
- [25] HUANG Y, QIU C, YUAN K. Surface defect saliency of magnetic tile[J]. *The Visual Computer*, 2020, 36(1): 85-96.
- [26] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-Decoder with atrous separable convolution for semantic image segmentation[C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. [S.l.]: ECCV, 2018: 801-818.
- [27] WANG J, SUN K, CHENG T, et al. Deep high-resolution representation learning for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 43(10): 3349-3364.
- [28] WU H, ZHANG J, HUANG K, et al. Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation[EB/OL]. [2022-03-12]. <https://arxiv.org/abs/1903.11816>.

编辑 税红