

融合类别注意力的铝硅合金显微图像分割方法

沈 韬^{1,2}, 金 凯^{1,2}, 司昌凯^{1,2}, 郑剑锋^{1,2}, 刘英莉^{1,2}

(1. 昆明理工大学信息工程与自动化学院, 昆明 650093; 2. 云南省计算机技术应用重点实验室, 昆明 650500)

摘 要: 为了提取铝硅合金显微图像的初晶硅区域, 提出一种结合类别注意力块(Class attention block, CAB)的改进模型类别注意力网络(Class attention network, CA-Net)。类别注意力块从特征图中计算各通道与每个类别的相关性信息, 并将不同类别的相关性信息融合产生注意力权重, 以使特征通道的权重与其对任务类别的贡献相关, 从而增强重要特征的表达, 并抑制无关特征的干扰。在铝硅合金显微图像数据集上进行实验, 本文提出的方法在Dice系数、Jaccard相似度、敏感度、特异度和分割准确率上的结果分别为94.82%、90.16%、94.54%、98.80%和97.97%。相比CCNet、SPNet和TA-Net等方法, CA-Net能够有效改进铝硅合金显微图像中初晶硅区域的分割效果。

关键词: 铝硅合金显微图像; 图像分割; 注意力机制; 卷积神经网络; 深度学习

中图分类号: TP391 文献标志码: A

Segmentation of Al-Si Alloy Microscopic Image by Fusing Class Attention

SHEN Tao^{1,2}, JIN Kai^{1,2}, SI Changkai^{1,2}, ZHENG Jianfeng^{1,2}, LIU Yingli^{1,2}

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650093, China;
2. Yunnan Key Laboratory of Computer Technologies Application, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: An improved model of class attention network (CA-Net) incorporating a class attention block (CAB) is proposed to extract the primary silicon regions of the microscopic images of Al-Si alloys in this paper. The correlation information of each channel to each class is calculated from the feature map by class attention block, and the correlation information of different classes is fused to generate attention weights for correlating the weights of feature channels with their contributions to the class in the task, thus the representation of important features is enhanced and the interference of irrelevant features is suppressed. Experiments are conducted on the Al-Si alloy microscopic image dataset, and the proposed method obtains results of 94.82%, 90.16%, 94.54%, 98.80%, and 97.97% for Dice coefficient, Jaccard similarity, sensitivity, specificity, and segmentation accuracy, respectively. The proposed CA-Net can effectively improve the segmentation effect of the primary silicon region in Al-Si alloy microscopic images compared with CCNet, SPNet, TA-Net, and other methods.

Key words: microscopic image of Al-Si alloy; image segmentation; attention mechanism; convolutional neural network; deep learning

引 言

高端制造业以及先进技术的发展依赖于新材料的研发。传统的材料研究采用以直觉和实验为主的“试错法”,其耗费的时间、人力成本相当高昂。在这种研究模式下,一种新材料从发现到应用之间往往需要10~20年的时间^[1],严重阻碍了新材料的研发进程。为变革传统的材料研究模式,加速新材料的发现和开发,美国在2011年宣布启动了“具有全球竞争力的材料基因组计划”(Materials genome initiative for global competitiveness, MGI)^[2],材料基因组计划有一个宏大的目标,就是将先进材料的发现、开发、制造以及使用的速度减半^[3]。在材料基因组计划中,数据库扮演着至关重要的角色,高质量的材料数据不仅能够用于模拟计算,而且可以为知识发现提供样本支撑^[4],是在加速新材料、新性能的探索与发现中不可或缺的基础。

金属材料是材料科学的研究重点之一,其中铝硅合金因具备耐磨性高、膨胀性低、密度低、耐腐蚀性良好以及力学性能优异等特点,在汽车、电子和航空航天等领域得到了广泛应用^[5]。在铝硅合金中,存在着针状的共晶硅和块状的初晶硅,其中初晶硅的尺寸、形状和分布是影响合金性能的关键因素。因此,从铝硅合金科研文献所包含的微观组织图像中提取出初晶硅的等效直径、形状因子和纵横比等特征参数,并与文献中的其他实验信息一起存入材料数据库中,可以记录材料的成分、组织、工艺和性能之间的关系。随着数据的积累,未来可以利用机器学习手段从中学习材料的成分-组织-工艺-性能间的关联关系^[6],从而实现材料性能的预测。通过这种方式,可以为材料研究人员提供相关的参考,从而避免许多不必要的实验,加速新材料的研发和应用。为提取微观组织图像中的特征参数,首要工作便是通过分割取得初晶硅颗粒的区域。自20世纪70年代起,图像分割方法便开始得到广泛研究,传统的图像分割方法主要基于图像的灰度、色彩、空间纹理和几何形状等特征来区分图像中的不同组成成分。传统图像分割方法主要有阈值分割法、边缘检测法以及基于数学形态学、图论等特定理论的方法,这些方法往往需要手动调整参数,且鲁棒性差,不适用于自动处理复杂图像。

近年来,深度学习的研究取得了巨大突破,卷积神经网络(Convolutional neural network, CNN)^[7]在图像分类^[8-11]、图像分割^[12-18]及目标检测等多个计算机视觉任务上取得了令人瞩目的进步。基于深度学习的图像分割算法能够自动地从数据中提取图像的隐含特征,并从中获取高级语义信息,具有较高的鲁棒性,因此更适合用在复杂图像的处理任务中。目前,主流的深度学习分割方法采用了全卷积网络(Fully convolutional networks, FCN)^[14]结构,这种结构能够实现端到端的图像分割。Ronneberger等^[18]通过增加模型的深度,并融合更多的高分辨率特征图信息,构建了U-Net,因其对小数据集效果较好的特点在医学图像分析领域被广泛研究和应用^[19]。Badrinarayanan等^[17]提出池化索引,并减少模型的整体参数量和计算内存需求,构建了SegNet。由于图像中含有不同大小的对象,且不同大小的感受野也会影响模型对上下文的理解,因而有很多工作致力于研究多尺度特征的提取和融合。Zhao等^[16]和Hou等^[20]利用改进的池化模块来聚合多尺度特征,以同时捕获图像的局部空间信息和全局上下文信息。Gu等^[21]在他们的工作中通过并行的方式组合不同膨胀率的空洞卷积,以提取不同尺度的上下文特征。Raza等^[22]提出了Micro-Net,通过将不同分辨率的原始图像输入到模型中,并融合不同分辨率的特征图来获取多尺度特征,使模型能够更好地识别微观组织图像中不同大小的细胞和腺体等对象。自SENet^[11]在ImageNet 2017挑战赛取得冠军后,注意力机制也得到了广泛的关注。SE模块^[11]通过全局平均池化和多层感知机来学习每个特征通道的权重,以增强重要的特征,弱化不重要的特征。CBAM^[23]在SE(Squeeze-and-excitation)模块的基础上添加了全局最大池化,并设计了

空间注意力对空间特征进行加权,改进了注意力模块的性能。Huang等^[24]提出了一种纵横交叉注意力模块,通过收集纵轴和横轴两个方向的像素信息来获取全局上下文。Fu等^[13]提出了DANet,以自注意力方式获取位置注意力和通道注意力权重,将通道与通道和位置与位置之间的关系融合到了特征的表达之中。Peng等^[25]使用平均池化聚合每个像素周围的局部空间特征生成注意力权重,将局部上下文嵌入到特征中,改进了模型在医学图像上的分割性能。Pang等^[26]提出了TA-Net,通过整合CBAM使模型更关注与肿瘤相关的特征,改进了肿瘤图像的分割效果。在材料显微图像分割任务中,深度学习方法也得到了一些研究。Azimi等^[27]将最大投票策略结合到FCN中,提出MVFCNN模型用于分割钢铁显微图像。Strohmann等^[28]提出一种加权损失函数,对铝硅合金显微图像中的硅以及硅和铝化物间的交接处赋予更高的权重,使模型能够更好地处理对应区域,并通过将三维序列中的多个切片输入网络,融入了三维信息。Ma等^[29]考虑到铝镧合金显微图像三维切片之间的关系,提出了一种对称矫正策略用于克服由噪声导致的欠分割问题。马博渊等^[30]提出一种区域感知损失函数,根据像素点到多晶纯铁晶粒(前景)中心及晶界(背景)的距离对不同通道的像素加权,加强了模型对晶粒形状信息的关注。

当前,基于深度学习方法分割材料显微图像的工作还比较少,且缺乏提取铝硅合金显微图像中初晶硅区域的研究。铝硅合金显微图像外观复杂,背景中大量的共晶硅在颜色上与初晶硅极为相似,容易造成混淆,对分割带来了很大的挑战。为增强模型对不同类别像素的识别能力,改进铝硅合金显微图像的分割质量,本文提出了一种新颖的类别注意力块(Class attention block, CAB),并结合类别注意力块构建了类别注意力网络(Class attention network, CA-Net)。在现有的通道注意力方法中,SE模块和CBAM基于全局池化获取的空间上下文来产生通道权重,DANet通过自注意力计算通道和通道间的关系来获取通道权重。与这些方法不同,本文提出的CAB从类别角度出发,通过计算特征通道和每个类别的相关性,并基于这些相关性信息来获取通道权重,以增强与任务类别相关的重要特征,抑制无关特征,提升模型的分割能力。与SE模块和CBAM相比,CAB显式地建模了特征和类别间的关系,有利于模型学习到更具判别性的特征。相比DANet的通道注意力模块,CAB占用更少的内存和计算资源,降低了模型对运行设备的要求。

1 类别注意力网络

CA-Net是一个全卷积网络,由特征编码器和特征解码器两部分构成。其中,特征编码器用于从输入图像中自动提取特征,特征解码器负责将包含丰富语义信息的低分辨率特征图恢复到原图分辨率,并预测出最终的分割结果,模型的整体结构如图1所示。

1.1 特征编码器

特征编码器如图1(a)所示,该模块由5个编码模块和3个CAB构成,编码模块上的数字指出了模块中卷积层的输出通道数,每个编码模块由两个 3×3 卷积层和残差连接^[8]组成,图1(c)展示了编码模块的具体结构。整个编码器结构从上至下可看成5个运算阶段,在每2个运算阶段之间,特征编码器会通过最大池化对特征图进行下采样,下采样有利于模型学习图像的上下文信息,并能有效减少模型的整体计算量和内存开销。在特征编码器后3个阶段的编码模块后分别放置了CAB,能够提取输入特征图中不同通道特征与各类别的相关性信息,并将这些信息结合到特征图中,从而增强与任务类别相关的特征,抑制无关特征,使特征具有更强的区分度。

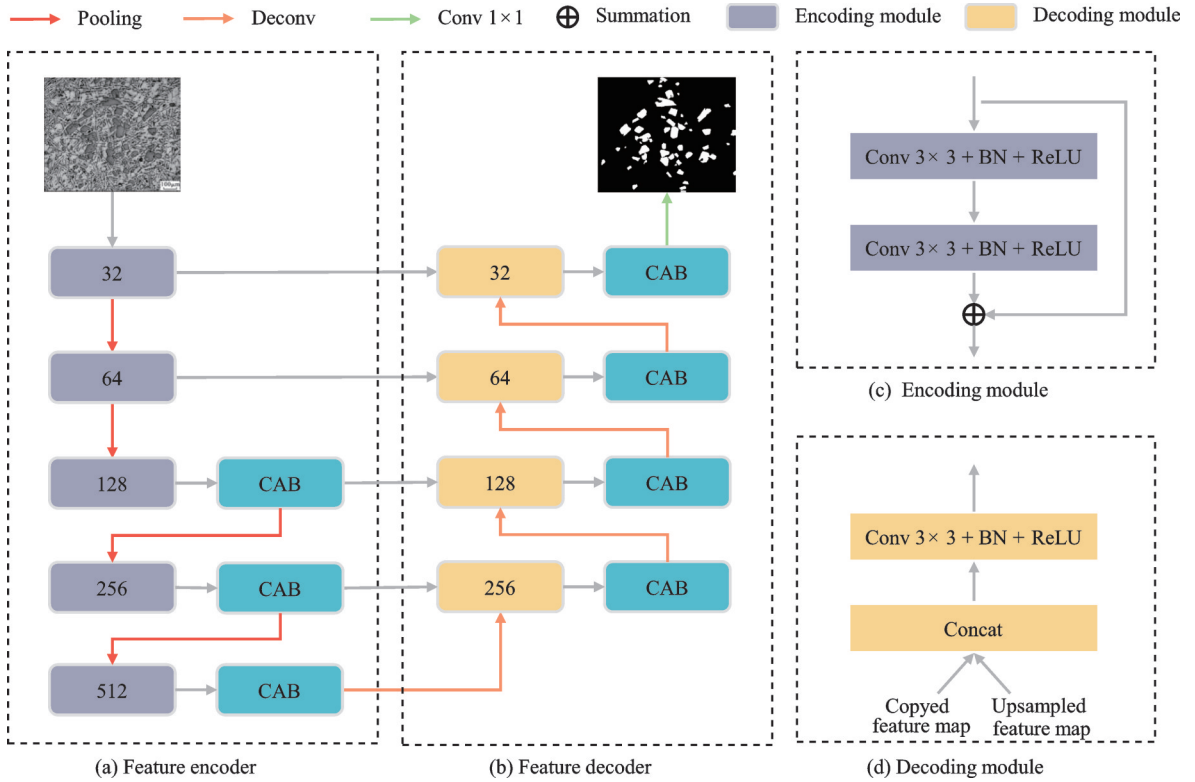


图1 CA-Net模型结构

Fig.1 CA-Net model structure

1.2 特征解码器

特征解码器如图1(b)所示,其中包含4个解码模块和4个类别注意力块,解码模块上的数字表示模块中卷积层的输出通道数,每个解码模块由1次特征图拼接操作和1个 3×3 卷积层构成,图1(d)中给出了解码模块的细节。对于每个解码模块,其接受来自编码器的高分辨率特征图,以及由低分辨率特征图上采样后的特征图,将它们拼接在一起并通过 3×3 卷积进行进一步处理,高分辨率的特征图能够弥补池化过程中丢失的特征信息,有利于模型产生更精细的结果。在第4个类别注意力块之后,特征解码器对输出特征图进行分类,得到最终的分割结果。

1.3 类别注意力块

为突出与任务类别相关的重要特征,减少无关特征的干扰,使模型能够更好地区分图像中不同类型的像素,本文提出了类别注意力块,其结构如图2所示。类别注意力块首先通过 1×1 卷积和softmax函数对输入特征 $F_{in} \in \mathbf{R}^{H \times W \times C}$ 进行分类,得到类别图 $S \in \mathbf{R}^{H \times W \times N}$,即

$$S = \text{softmax}(f_{1 \times 1}(F_{in})) \tag{1}$$

式中: $f_{1 \times 1}$ 表示 1×1 卷积;softmax表示softmax函数。然后通过将 S 在宽和高维度上展开,得到矩阵 $S_r \in \mathbf{R}^{(H \times W) \times N}$ 为

$$S_r = \text{reshape}(S) \tag{2}$$

式中reshape表示维度调整。调整维度后, S_r 的每一列即 S 对应通道展开后的向量。将输入特征图 F_{in}

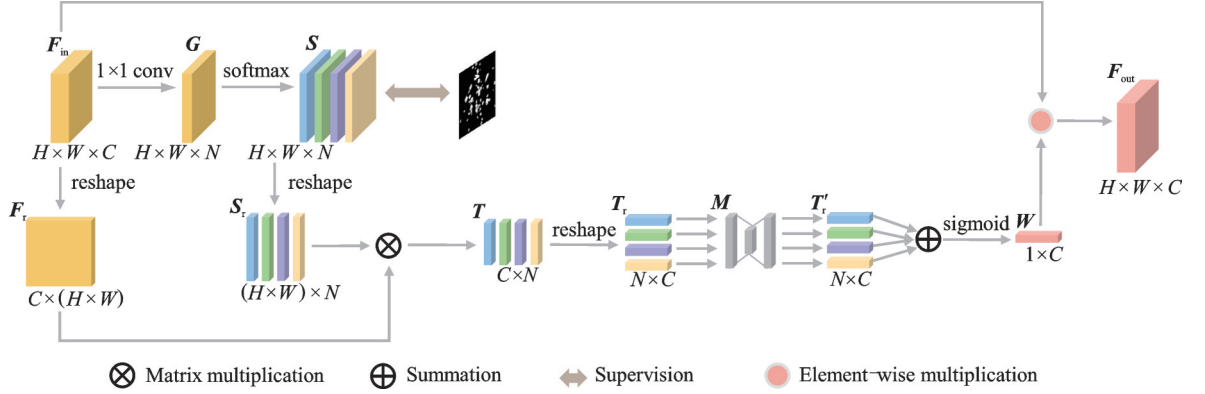


图2 类别注意力块

Fig.2 Class attention block

在宽和高维度上展开,并调整维度得到矩阵 $F_r \in \mathbf{R}^{C \times (H \times W)}$,即

$$F_r = \text{reshape}(F_{in}) \quad (3)$$

F_r 的每一行即 F_{in} 对应通道展开后的特征图。然后在 F_r 和 S_r 之间进行矩阵乘法运算,并调整维度得到特征通道和类别之间的相关性矩阵 $T_r \in \mathbf{R}^{N \times C}$,即

$$T_r = \text{reshape}(F_r \otimes S_r) \quad (4)$$

式中 \otimes 表示矩阵乘法。矩阵 T_r 的每一行反映了每一个类别与 F_{in} 的各通道之间的相关性,其中的值越大,则对应特征通道和相应类别间的关系越紧密。接下来,将 T_r 的每一行分别传入共享的多层感知机得到注意力图 $T_r' \in \mathbf{R}^{N \times C}$,再将 T_r' 中的每一行相加,通过 sigmoid 函数归一化得到最终的通道注意力权重 $W \in \mathbf{R}^{1 \times C}$,即

$$W = \sigma \left(\sum_{i=1}^N \text{MLP}(T_r^i) \right) \quad (5)$$

式中: $i \in \{1, 2, \dots, N\}$ 表示矩阵 T_r 中第 i 行;MLP为多层感知机; N 为类别数; σ 表示 sigmoid 函数。 W 综合考虑了每一个特征通道对不同类别的贡献,因此将特征通道的重要性嵌入到了注意力权重之中。最终,将输入特征图 F_{in} 和注意力权重 W 相乘得到输出特征图 $F_{out} \in \mathbf{R}^{H \times W \times C}$,即

$$F_{out} = F_{in} \odot W \quad (6)$$

式中 \odot 表示逐元素乘法。 F_{out} 中结合了特征通道与类别之间的相关性信息,从而能够加强与任务类别相关的重要特征,并抑制无关特征。此外,在类别注意力块中,为了获取类别图,包含了一个辅助的分类器,辅助分类器在训练时能够起到深度监督的作用,有利于加速模型的收敛。

2 实验结果及分析

2.1 数据集

为实现对文献中铝硅合金显微图像的分割,本文从各类中英文期刊网站上收集了与铝硅合金研究相关的 pdf 文献,并使用 Photoshop 工具从中提取出铝硅合金显微图像。通过与材料专家交流,本文去除了非常模糊以及初晶硅较少的图像,并最终选取出 71 张图像用于构建数据集,为降低模型的计算开销,所有的图像被保存为 8 bit 的灰度图。本文在材料专家的帮助下对图像进行了像素级标注,标签中

含有初晶硅和背景两个类别。由于数据集较小,为克服过拟合问题,本文使用了大量的数据增强方法,包括水平翻转、垂直翻转、转置、随机90°旋转、0.5~1.5倍随机尺度缩放、随机亮度对比度调整和随机裁剪等。其中,图像的剪裁尺寸被统一为257像素×257像素,以方便在训练时并行处理多张图像,加速训练过程。此外,本文对数据集进行了5次不同的划分,并分别在这些划分上进行实验,以更充分地评估模型的泛化性能。具体而言,在每次划分中,训练集、验证集和测试集分别包含70%、10%和20%的图像,最终的结果为5次实验的平均结果。

2.2 评估指标

为了客观地评价模型性能,本文选择了Dice系数 D 、Jaccard相似度 J 、敏感度 E 、特异度 S 和准确率 A 作为评估指标。Dice系数和Jaccard相似度从集合的角度反映了分割结果和标签之间的相似程度,敏感度表示正确分割的初晶硅像素在所有初晶硅像素中的比例,特异度表示正确分割的背景像素在所有背景像素中的占比,准确率即所有像素中被正确分类的比例。各项指标的计算公式为

$$D = \frac{2TP}{2TP + FP + FN} \quad (7)$$

$$J = \frac{TP}{TP + FP + FN} \quad (8)$$

$$E = \frac{TP}{TP + FN} \quad (9)$$

$$S = \frac{TN}{FP + TN} \quad (10)$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

式中:TP表示正确预测的初晶硅像素数量;FP表示错误预测的初晶硅像素数量;TN表示正确预测的背景像素数量;FN表示错误预测的背景像素数量。

2.3 实验环境及实现细节

本文实验所用环境为Ubuntu 20.04操作系统,NVIDIA Tesla V100 GPU,显存16 GB,Intel Xeon Gold 5120 CPU,运行内存256 GB,实验采用深度学习框架PyTorch进行。

在训练时,使用随机梯度下降优化器,动量设置为0.99,初始学习率设置为0.003,batch size设置为16。为了克服过拟合,设置了0.0005的权重衰减,训练总共迭代了30000次,采用poly学习率衰减策略,计算方法为 $lr = base_lr \times \left(1 - \frac{iter}{max_iter}\right)^{power}$,其中 lr 表示当前学习率, $base_lr$ 表示初始学习率, $iter$ 表示当前迭代次数, max_iter 表示总迭代次数, $power$ 用于控制衰减速度,并设置 $power$ 为0.9。

2.4 消融实验

2.4.1 类别注意力块个数对模型性能的影响

本文在验证集上对类别注意力块的个数进行了消融实验,由于类别注意力块需要通过分类获取类别图,而卷积神经网络模型在浅层的分类能力较弱,深层的分类能力较强,因此本文从模型的最后一层开始逐渐往前增加类别注意力块。类别注意力块被放置在编码模块或解码模块之后。实验结果如表1所示,其中:未使用类别注意力块的模型被称为BaseNet;使用类别注意力块的模型为BaseNet + CAB(NUM);NUM表示类别注意力块的个数。所有的评估指标中,每一列的最佳结果用粗体进行了标注。

表1 使用不同类别注意力块个数的模型的分割性能

Table 1 Segmentation performance of models using different numbers of class attention blocks

模型配置	Dice系数/%	Jaccard相似度/%	敏感度/%	特异度/%	准确率/%
BaseNet	94.74	90.01	94.30	99.00	98.15
BaseNet + CAB(1)	94.72	89.97	94.43	98.96	98.15
BaseNet + CAB(2)	94.93	90.35	94.82	98.95	98.22
BaseNet + CAB(3)	95.25	90.94	94.97	99.07	98.33
BaseNet + CAB(4)	95.31	91.06	95.19	99.05	98.36
BaseNet + CAB(5)	95.26	90.95	95.03	99.05	98.34
BaseNet + CAB(6)	95.32	91.07	95.08	99.07	98.36
BaseNet + CAB(7)	95.47	91.33	95.37	99.06	98.42
BaseNet + CAB(8)	95.46	91.32	95.22	99.11	98.42
BaseNet + CAB(9)	95.48	91.37	95.29	99.10	98.43

从表1中可以看出,所有使用了CAB的模型配置在敏感度指标上的结果均高于BaseNet,这表明CAB有效地改进了模型对初晶硅区域的分割能力。在CAB的个数从1增至4以及从5增至7的过程中,模型的Dice系数、Jaccard相似度、敏感度和准确率逐渐上升,这反映了CAB的数量越多,模型的性能改进越大。但是,当CAB的个数大于7时,模型的Dice系数、Jaccard相似度和准确率只产生了非常微小的改变,且敏感度有所下降,这表明模型浅层的分类能力较弱,难以有效获取类别信息,从而限制了类别注意力块的性能。因此,本文最终选择在模型中放置7个类别注意力块。

2.4.2 深度监督 and 不同注意力方法对模型性能的影响

为探究类别注意力块对模型性能的影响,本文在验证集上对深度监督和不同的注意力方法进行了消融实验,所有的实验配置如表2所示。实验结果如表3所示,每列的最佳结果用粗体标示。由表3可见,所有使用了深度监督的版本相较于没使用的版本均对模型性能带来了一些改进,这说明深度监督能够有效地使模型收敛到更好的结果。对于BaseNet + SE配置,Dice系数、Jaccard相似度、敏感度、特异度和准确率相比BaseNet分别提升0.18%、0.34%、0.22%、0.03%和0.08%。BaseNet + CBAM配置相比于BaseNet,指标

表2 不同实验配置及说明

Table 2 Different experimental configurations and descriptions

模型配置	说明
BaseNet	不使用注意力方法和深度监督
BaseNet + DS	仅使用深度监督
BaseNet + CAB	使用类别注意力块
BaseNet + SE	使用SE模块
BaseNet + SE + DS	使用SE模块和深度监督
BaseNet + CBAM	使用CBAM
BaseNet + CBAM + DS	使用CBAM和深度监督

分别上升0.41%、0.75%、0.36%、0.11%和0.16%。上述结果说明SE模块和CBAM均能够有效改进模型性能,且CBAM的效果优于SE模块。对于BaseNet + CAB配置,相比BaseNet + DS配置,Dice系数、Jaccard相似度、敏感度、特异度和准确率分别提升了0.72%、1.30%、1.02%、0.07%和0.27%。上述结果说明本文提出的类别注意力块不仅拥有深度监督的优点,而且能够在这个基础上进一步结合特征与类别间的相关性信息,从而有效提升模型的分割性能。其中,敏感度的提升反映了模型能够更好地分割初晶硅区域,特异度的提升说明背景像素的识别也得到了改进。这说明类别注意力块有

效地增强了与任务类别相关的重要特征的表达,并改进了模型对不同类别像素的区分能力。从表3同时可以看到,BaseNet + CAB配置在Dice系数、Jaccard相似度、敏感度和准确率4个指标上取得了最佳的结果,相较于BaseNet + SE + DS配置,这些指标分别改进了0.12%、0.20%、0.17%和0.05%,相较于BaseNet + CBAM + DS配置,这些指标分别提升了0.11%、0.18%、0.37%和0.04%,这表明相比SE模块和CBAM,类别注意力块能够获取到更具判别性的特征,从而使模型得到更大的改进。

表3 不同模型配置的分割性能

Table 3 Segmentation performance of different model configurations

模型配置	Dice系数/%	Jaccard相似度/%	敏感度/%	特异度/%	准确率/%
BaseNet	94.74	90.01	94.30	99.00	98.15
BaseNet + DS	94.75	90.03	94.35	98.99	98.15
BaseNet + CAB	95.47	91.33	95.37	99.06	98.42
BaseNet + SE	94.92	90.35	94.52	99.03	98.23
BaseNet + SE + DS	95.35	91.13	95.20	99.07	98.37
BaseNet + CBAM	95.15	90.76	94.66	99.11	98.31
BaseNet + CBAM + DS	95.36	91.15	95.00	99.11	98.38

2.5 类别注意力块可视化

为了观察类别注意力块是否能够正确地获取到特征通道与类别之间的相关性信息,本文对网络中的第2个类别注意力块进行了可视化。本文首先从类别注意力块中取得相关性矩阵 T_r ,每一行最大值和最小值对应的通道编号,然后将输入特征图在这些编号处的通道取出,并以热力图的形式展示。可视化结果如图3所示,其中Class 0 min和Class 0 max分别表示 T_r 第0行最小值和最大值对应通道的特征,Class 1 min和Class 1 max分别表示 T_r 第1行最小值和最大值对应通道的特征, T_r 的第0行和第1行分别代表背景类别和初晶硅类别。

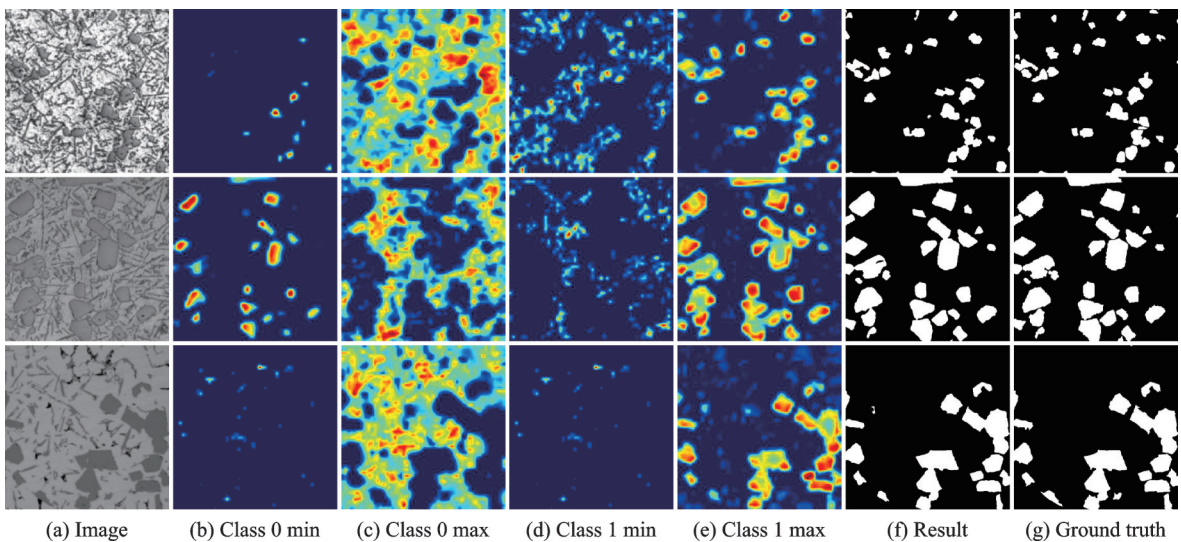


图3 网络中第2个类别注意力块的可视化结果

Fig.3 Visualization results for the 2nd class attention block in the network

从图3结果中可以看出,对于这3张图:Class 0 min在背景处响应非常少;Class 0 max在背景的大部分区域中获得了较高的响应;Class 1 min的响应较低,且基本分布在背景上;Class 1 max的响应较高,主要聚焦于初晶硅区域。这说明矩阵 T_r 获取到了每一个特征通道和每一个类别之间的相关性信息,并且对于每一个类别,与其相关性高的特征通道在属于该类别的区域获得了较高的响应,而与其相关性低的特征通道在这些区域获得了较低的响应。此外,从图3中还可以发现,前两张图的Class 0 min虽然也在识别初晶硅区域,但他们所聚焦的区域要比Class 1 max少,这表明相比Class 0 min,Class 1 max所对应的通道特征与初晶硅类别有着更高的相关性,而Class 1 max为 T_r 第1行中的最大值对应的通道特征,Class 0 min为其他通道的特征,这足以说明不同特征通道与类别之间相关性的强弱关系也反映到了 T_r 中。

2.6 结果分析

本文在测试集上评估了CA-Net与其他的8种方法,实验结果如表4所示,其中用粗体标出了每一列的最好结果。结果显示,CA-Net在Dice系数、Jaccard相似度、敏感度和准确率4个指标上都取得了最佳性能:相比SPNet,这些指标分别提升了0.81%、1.43%、1.73%和0.29%;相比TA-Net,这些指标分别改进了0.86%、1.51%、1.46%和0.32%。这表明类别注意力块能够有效地增强特征表达,从而更好地利用学习到的特征。此外,表4中也列出了各模型的参数量和每秒浮点运算次数FLOPs,结果显示,CA-Net的参数量为 7.59×10^3 ,仅比参数最少的MultiResUNet多 0.35×10^3 ,而CA-Net的FLOPs为 22.89×10^6 ,在所有模型中位列第3,与第1名CE-Net只相差 3.44×10^6 ,这说明本文提出的方法比较轻量,且计算效率相对较高。

表4 不同方法的分割性能

Table 4 Segmentation performance of different methods

模型	Dice系数/%	Jaccard相似度/%	敏感度/%	特异度/%	准确率/%	参数量/ 10^3	FLOPs/ 10^6
SegNet ^[17]	93.74	88.26	92.58	98.81	97.58	29.44	83.68
PSPNet ^[16]	93.92	88.57	93.16	98.74	97.63	49.07	97.94
CCNet ^[24]	93.58	87.96	92.07	98.85	97.51	52.28	109.53
SPNet ^[20]	94.01	88.73	92.81	98.89	97.68	47.72	92.63
U-Net ^[18]	93.83	88.43	92.79	98.80	97.60	7.76	27.44
MultiResUNet ^[31]	93.92	88.60	92.62	98.91	97.65	7.24	37.21
CE-Net ^[21]	93.82	88.42	92.60	98.86	97.61	29.00	19.45
TA-Net ^[26]	93.96	88.65	93.08	98.79	97.65	29.56	20.34
CA-Net(Proposed)	94.82	90.16	94.54	98.80	97.97	7.59	22.89

图4展示了CCNet、SPNet、TA-Net和CA-Net在4张不同图像上的分割结果,其中用红圈标出了CA-Net取得明显改进的部分。可以看到,CA-Net识别出了更多的初晶硅像素,得到了更完整的分割对象,且分割对象的形状更接近于真值标签。

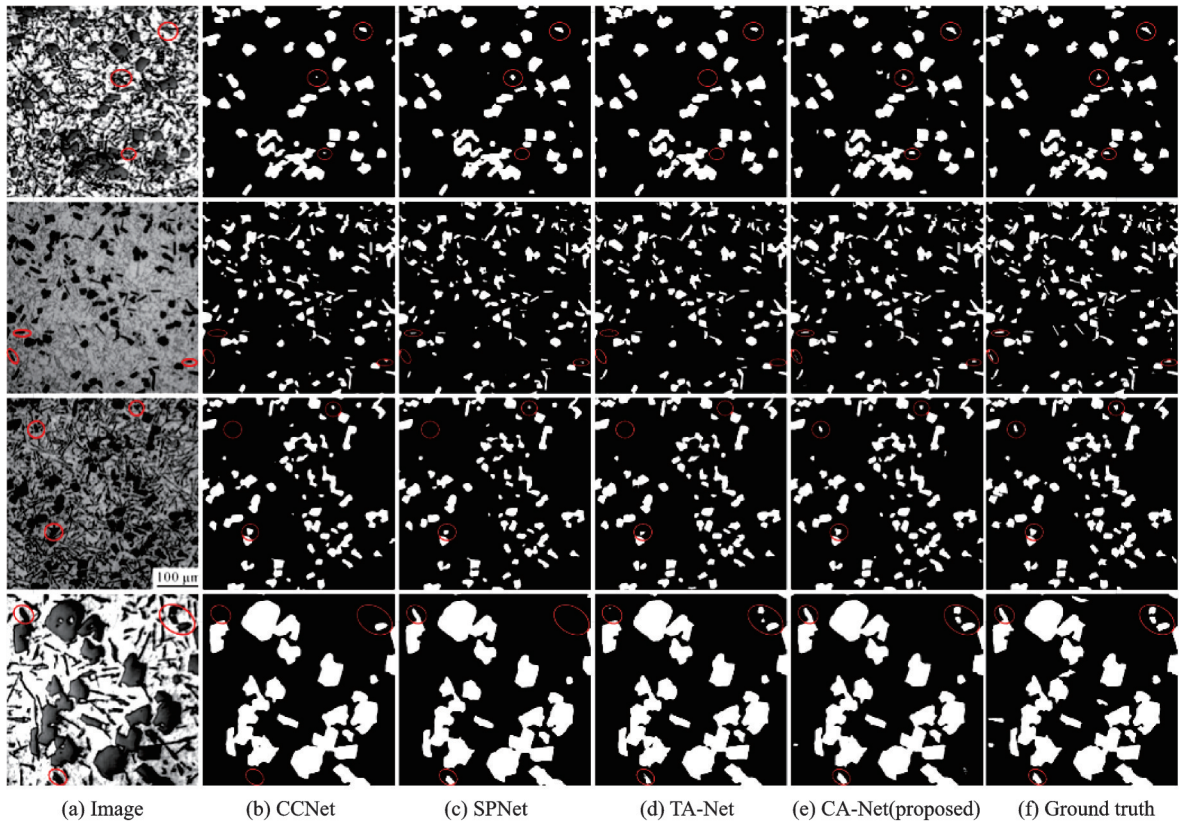


图4 CA-Net与其他方法的分割结果

Fig.4 Segmentation results of CA-Net and other methods

3 结束语

为了提升铝硅合金显微图像的分割质量,本文提出了一种结合类别注意力块的分割方法CA-Net。类别注意力块对输入特征图进行分类,然后计算特征图各通道与每个类别之间的相关性信息,并将不同类别的相关性信息融合产生类别注意力权重。通过将这些权重信息结合到特征图中,能够增强与任务类别相关的重要特征,并抑制无关特征。本文对深度监督和类别注意力块进行了消融实验,结果表明深度监督能够使模型收敛到更好的结果,类别注意力块除了具备深度监督的优点,还能对模型带来较大的改进。最后,本文对8种不同的分割方法进行了实验,并与所提方法进行对比,实验结果表明,本文提出的方法轻量、高效,且能够有效改进铝硅合金显微图像的分割效果。

参考文献:

- [1] 黄河, 陈宏生. 人工智能将推动材料基因组技术加速发展[J]. 全球科技经济瞭望, 2019, 34(Z1): 38-47.
HUANG He, CHEN Hongsheng. Artificial intelligence will accelerate the development of material genome technology[J]. Global Science, Technology and Economy Outlook, 2019, 34(Z1): 38-47.
- [2] White House Office of Science and Technology Policy. Materials genome initiative for global competitiveness[EB/OL].(2011-06-24) [2021-11-20]. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf.
- [3] 赵继成. 材料基因组计划简介[J]. 自然杂志, 2014, 36(2): 89-104.

- ZHAO Jicheng. A perspective on the materials genome initiative[J]. *Chinese Journal of Nature*, 2014, 36(2): 89-104.
- [4] 杨丽, 苏航, 柴锋, 等. 材料数据库和数据挖掘技术的应用现状[J]. *中国材料进展*, 2019, 38(7): 672-681, 650.
- YANG Li, SU Hang, CHAI Feng, et al. Material database and application status of data mining technology[J]. *Materials China*, 2019, 38(7): 672-681, 650.
- [5] ANAND S, SRIVATSAN T S, WU Y, et al. Processing, microstructure and fracture behaviour of a spray atomized and deposited aluminium-silicon alloy[J]. *Journal of Materials Science*, 1997, 32(11): 2835-2848.
- [6] 汪洪, 项晓东, 张澜庭. 数据驱动的材料创新基础设施[J]. *Engineering*, 2020, 6(6): 56-61.
- WANG Hong, XIANG Xiaodong, ZHANG Lanting. On the data-driven materials innovation infrastructure[J]. *Engineering*, 2020, 6(6): 56-61.
- [7] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [9] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA: IEEE, 2015: 1-9.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60: 84-90.
- [11] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [12] 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.
- [13] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]// *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, 2019: 3141-3149.
- [14] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA: IEEE, 2015: 3431-3440.
- [15] NOH H, HONG S, HAN B. Learning deconvolution network for semantic segmentation[C]// *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile: IEEE, 2015: 1520-1528.
- [16] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 6230-6239.
- [17] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]// *Proceedings of Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer International Publishing, 2015: 234-241.
- [19] FALK T, MAI D, BENSCH R, et al. U-Net: Deep learning for cell counting, detection, and morphometry[J]. *Nature Methods*, 2019, 16(1): 67-70.
- [20] HOU Q, ZHANG L, CHENG M M, et al. Strip pooling: Rethinking spatial pooling for scene parsing[C]// *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE, 2020: 4002-4011.
- [21] GU Z, CHENG J, FU H, et al. CE-Net: Context encoder network for 2D medical image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2281-2292.
- [22] RAZA S E A, CHEUNG L, SHABAN M, et al. Micro-Net: A unified model for segmentation of various objects in microscopy images[J]. *Medical Image Analysis*, 2019, 52: 160-173.
- [23] WOO S, PARK J, LEE J-Y, et al. CBAM: Convolutional block attention module[C]// *Proceedings of Conference on Computer Vision*. Cham: Springer International Publishing, 2018: 3-19.

- [24] HUANG Z, WANG X, WEI Y, et al. CCNet: criss-cross attention for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020: 1.
- [25] PENG D, XIONG S, PENG W, et al. LCP-Net: A local context-perception deep neural network for medical image segmentation[J]. Expert Systems with Applications, 2021, 168: 114234.
- [26] PANG S, DU A, ORGUN M A, et al. Tumor attention networks: Better feature selection, better tumor segmentation[J]. Neural Networks, 2021, 140: 203-222.
- [27] AZIMI S M, BRITZ D, ENGSTLER M, et al. Advanced steel microstructural classification by deep learning methods[J]. Scientific Reports, 2018, 8(1): 2128.
- [28] STROHMANN T, BUGELNIG K, BREITBARTH E, et al. Semantic segmentation of synchrotron tomography of multiphase Al-Si alloys using a convolutional neural network with a pixel-wise weighted loss function[J]. Scientific Reports, 2019, 9(1): 19611.
- [29] MA B, BAN X, HUANG H, et al. Deep learning-based image segmentation for Al-La alloy microscopic images[J]. Symmetry, 2018, 10(4): 107.
- [30] 马博渊, 刘楚妮, 高鸣飞, 等. 基于深度学习和区域感知的多晶体显微组织图像分割方法[J]. 中国体视学与图像分析, 2020, 25(2): 120-127.
- MA Boyuan, LIU Chuni, GAO Mingfei, et al. Region aware image segmentation for polycrystalline micrographic image using deep learning[J]. Chinese Journal of Stereology and Image Analysis, 2020, 25(2): 120-127.
- [31] IBTEHAZ N, RAHMAN M S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. Neural Networks, 2020, 121: 74-87.

作者简介:



沈韬(1984-),男,博士,教授,研究方向:区块链技术、人工智能、能源互联网, E-mail: shentao@kust.edu.cn。



金凯(1994-),男,硕士研究生,研究方向:图像处理、深度学习, E-mail: jinkai0916@outlook.com。



司昌凯(1993-),男,硕士研究生,研究方向:图像处理、深度学习, E-mail: sck_2020@163.com。



郑剑锋(1997-),男,硕士研究生,研究方向:图像处理、深度学习, E-mail: 1157645879@qq.com。



刘英莉(1978-),通信作者,女,博士,副教授,研究方向:机器学习、自然语言处理、材料基因组计划, E-mail: ly12002@126.com。

(编辑:张黄群)