

基于红外注意力提升机制的热成像测温区域实例分割

易诗^{*①②③} 李俊杰^① 贾勇^①

^①(成都理工大学信息科学与技术学院(网络安全学院、牛津布鲁克斯学院) 成都 610059)

^②(工业物联网与网络化控制教育部重点实验室 重庆 400065)

^③(太赫兹科学技术四川省重点实验室 成都 610054)

摘要: AI+热成像人体温度监测系统被广泛用于人群密集的人体实时温度测量。此类系统检测人的头部区域进行温度测量,由于各类遮挡,温度测量区域可能太小而无法正确测量。为了解决这个问题,该文提出一种融合红外注意力提升机制的无锚点实例分割网络,用于实时红外热成像温度测量区域实例分割。该文所提出的实例分割网络在检测阶段和分割阶段融合红外空间注意力模块(ISAM),旨在准确分割红外图像中的头部裸露区域,以进行准确实时的温度测量。结合公共热成像面部数据集和采集的红外热成像数据集,制作了“热成像温度测量区域分割数据集”用于网络训练。实验结果表明:该方法对红外热成像图像中头部裸露测温区域的平均检测精度达到88.6%,平均分割精度达到86.5%,平均处理速度达到33.5 fps,在评价指标上优于大多数先进的实例分割方法。

关键词: 红外热成像; 人体体温监测系统; 红外注意力提升机制; 无锚点实例分割网络; 热成像温度测量区域分割数据集

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2021)12-3505-08

DOI: 10.11999/JEIT200862

Instance Segmentation of Thermal Imaging Temperature Measurement Region Based on Infrared Attention Enhancement Mechanism

YI Shi^{*①②③} LI Junjie^① JIA Yong^①

^①(College of Information Science and Technology(College of Cyber Security, College of Oxford Brookes),
Chengdu University of Technology, Chengdu 610059, China)

^②(Key Laboratory of Industrial Internet of Things & Networked Control,
Ministry of Education, Chongqing 400065, China)

^③(Terahertz Science and Technology Key Laboratory of Sichuan Province, Chengdu 610054, China)

Abstract: AI+thermal imaging human body temperature monitoring system is widely used for real-time temperature measurement of human body in dense crowds. The artificial intelligence method used in such systems detects the human head region for temperature measurement. The temperature measurement area may be too small to measure correctly due to occlusion. To tackle this problem, an anchor-free instance segmentation network incorporating infrared attention enhancement mechanism is proposed for real-time infrared thermal imaging temperature measurement area segmentation. The instance segmentation network proposed in this paper integrates the Infrared Spatial Attention Module (ISAM) in the detection stage and the segmentation stage, aiming to accurately segment the bare head area in the infrared image. Combined with the public thermal imaging facial dataset and the collected infrared thermal imaging dataset, the "thermal imaging temperature measurement area segmentation dataset" is produced. Experimental results demonstrate that this method reached an average detection precision of 88.6%, average mask precision of 86.5%, average

收稿日期: 2020-10-09; 改回日期: 2021-04-22; 网络出版: 2021-07-15

*通信作者: 易诗 549745481@qq.com

基金项目: 国家自然科学基金(61771096), 工业物联网与网络化控制教育部重点实验室开放基金(2020FF06), 太赫兹科学技术四川省重点实验室开放基金(THZSC202001)

Foundation Items: The National Natural Science Foundation of China (61771096), The Key Laboratory of Industrial Internet of Things & Networked Control Foundation, Ministry of Education (2020FF06), The Terahertz Science and Technology Key Laboratory Foundation of Sichuan Province (THZSC202001)

processing speed of 33.5 fps. This network is superior to most state of the art instance segmentation methods in objective evaluation metrics.

Key words: Infrared thermal imaging; Human body temperature monitoring system; Infrared attention enhancement mechanism; Anchor-free instance segmentation network; Thermal imaging temperature measurement area segmentation dataset

1 引言

热成像使用物体温度差成像,可以实时反映物体温度以进行非接触式快速温度测量^[1]。热成像在医学中广泛使用,是一种非侵入性和非接触式检测方法,已用于温度调节^[2]、乳腺癌检测^[3,4]、新生儿监测^[5]、泌尿系统检测^[6]和血管领域疾病的检测^[7,8]。

在2020年爆发的新冠肺炎疫情中, AI+热成像实时体温监控系统得到了广泛的应用。这种类型的系统的原理是使用深度学习目标检测方法检测红外图像中的行人头部并测量头部区域的温度。然而,使用红外热成像人体温度监测系统的有效温度测量区域是裸露的面部区域,仅在红外图像中检测人的头部区域会导致以下问题:(1)头发、帽子、口罩和眼镜等遮挡物会导致温度测量区域过小而无法正确测量。(2)目标检测精度难以分割到像素级温度测量区域。使用实例分割方法,可以对红外图像中的头部的温度测量区域进行像素级精确分割,并且同时可以分割出面部的各类遮挡物,从而划定红外图像中头部裸露区域进行更准确的温度测量,同时对于遮挡过多导致测温区域过小无法正确测量的情况进行警告提示。

可见光环境下的实例分割方法发展迅速。He等人^[9]提出了最具代表性的深度学习实例分割方法(Mask Region Convolutional Neural Networks, Mask R-CNN),该方法在可见光环境下对目标进行实例分割达到较高的分割精度但实时性不理想。文献^[10]提出了一种实时实例分割方法YOLACT(You Only Look At CoefficienTs),该方法针对实例分割的实时性做出了改进。2020年,实时无锚实例分割网络以其准确性和实时性而广受欢迎, Lee等人^[11]提出了一种具有准确性和实时性的实时无锚点实例分割方法(Center Mask, CenterMask)。Xie等人^[12]提出了一种以中心点向外极化方式进行分割的无锚点实例分割方法(Polarization Mask, PolarMask),进一步提升了无锚点实例分割网络的分割准确性。对于可见光环境下的面部分割, Benini等人^[13]提出了通过语义面部分割进行面部分析的方法,能够较精准地分割可见光环境下的面部。红外图像中检测面部方面研究近年来也有所发展,文献^[14]全面回顾了红外面部识别的当前方法和数据

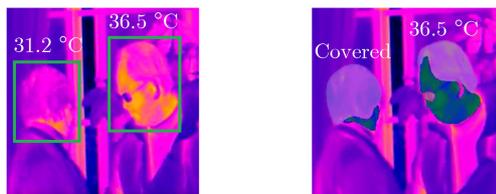
集, Yang等人^[15]提出了一种基于YWF-SIFT(Y-styled Window Filter Scale Invariant Feature Transform)和多尺度融合的红外图像人脸识别的良好匹配方法,文献^[16]提出了热成像图像中人脸和眼睛的定位算法,用于测量人眼温度。

上述方法涉及可见光环境下目标实例分割、面部分割、红外图像中的面部识别与定位,然而目前尚未出现应用于红外测温系统中的测温区域实时精确分割方法的研究。为此,本文提出了一种融合红外注意力提升机制的无锚点实例分割网络,用于红外热像测温系统中头部裸露测温区域的实时准确分割。该网络根据红外图像的特性对网络架构进行了改进,在特征提取主干网络与分割模块中融入红外空间注意力模块(Infrared Spatial Attention Module, ISAM)以优化提取到的红外图像特征,使像素分割注意力集中于红外图像中信息丰富的像素,同时抑制红外图像中的噪点。采集制作了针对红外热成像测温区域分割的“热成像温度测量区域分割数据集”用于模型训练。最后通过实验证明,本文提出的融合红外注意力提升机制的无锚点实例分割网络对红外热成像测温区域的分割对比其他实例分割方法更具优势。

2 算法

2.1 问题描述

热成像体温测量方法如图1所示,红外图像中左边测温人由于头发遮挡而无法正确测温。图1(a)所示目前普遍使用的红外热成像头部检测温度测量方法,头部测温框检测到了头部,对遮挡部分进行了测温,显示了31.2 °C的温度测量值,而图1(b)所示本文提出的红外热成像测温区域实例分割温度测量方法,分割结果显示头发遮挡了有效测温区域无法准确测量,给出了遮挡的警告提示。



(a) 头部检测温度测量 (b) 测温区域实例分割温度测量

图1 热成像体温测量方法

2.2 网络框架

实例分割可以准确地分割所需的目标轮廓, 从而达到像素级的精度。本文需要构架一个实时实例分割网络来分割红外热成像面部温度测量区域和各种遮挡物。

本文构建了一种实时无锚点实例分割网络, 该网络适合于根据红外热成像的特征分割红外热成像温度测量区域和各种遮挡物。网络架构如图2所示。

该网络分为红外目标无锚点检测与红外目标分割两个阶段。

红外目标无锚点检测阶段, 主干特征提取网络考虑到检测精度与实时性选用残差网络ResNet-50^[17], 阶段2—阶段5为主干特征提取网络中步长 $2^2 \sim 2^5$ 层所提取的特征图, 低层的特征图拥有较高的空间分辨率, 能提供更多的定位信息, 而高层的特征图则融合更多的语义信息, 因此可将低层和高层的特征图融合成一个特征图的金字塔网络 (Feature Pyramid Networks, FPN) 特征提取方式用于检测任务。出于减少算法复杂度的考虑, 这里采用一种简单的特征融合方式: 首先将每一特征提取阶段输出特征图通过红外空间注意力模块 (ISAM) 以强化特征图中红外图像特征, 再利用由512个 3×3 卷积核所构成的反卷积层1、由1024个 3×3 卷积核所构成的反卷积层2、由2048个 3×3 卷积核所构成的反卷积层3分别将第3、4和5阶段的特征图分辨率提升到和第2阶段的特征图分辨率保持一致, 之后对所有要融合的特征图进行L2归一化, 将这些特征图在通道维度上拼接起来, 得到增强型红外特征图。获取增强型红外特征图后进行目标检测, 由于无锚点检测方式提供了一种将目标检测中密集滑窗或锚点框铺设转换为目标关键点检测的思路, 提高了目标检测的精确性、实时性与通用性。因此, 无锚点检测方式更加适用于密集人群中实时检测红外图像测温区域, 本文对获取的增强型红外特征图采用文献^[18]提到的无锚点检测方法得到检测结果。

红外目标分割阶段, 分割部分第1层输入特征

图大小为 112×112 , 第2层输入特征图大小为 56×56 , 第3层输入特征图大小为 28×28 , 第4层输入特征图大小为 14×14 , 每个卷积层均由80个 3×3 卷积核构成。在分割部分第2层与第3层间插入红外空间注意力模块, 有助于将注意力集中在红外图像中信息丰富的像素上, 同时还能抑制红外图像中的噪声, 得到更为精确的测温区域实例分割结果。

2.3 红外空间注意力模块

深度神经网络的注意力提升机制近年来广泛应用于提升特征提取网络所提取的图像的特征, 其中具有代表性的为Woo等人^[19]提出的卷积块注意力模块 (Convolutional Block Attention Module, CBAM), Park等人^[20]提出的瓶颈注意力模块 (Bottleneck Attention Module, BAM)。普遍使用的注意力提升模块采用通道注意力分支 (channel attention branch) 与空间注意力分支 (spatial attention branch) 相融合的结构以在特征图通道关系与特征图空间位置关系上进行注意力提升, 增强特征表达能力。

而对于红外图像这类单通道图像, 反映的是物体温差, 成像原理为目标温度差形成的灰度图, 因此设计提升红外图像特征的注意力模块时, 应当考虑单独使用强化的空间注意力分支结构。本文在构建的实时无锚点实例分割网络中融入了红外空间注意力模块针对红外图像的特点进行了特征强化。本文设计的红外空间注意力模块结构如图3所示。

红外空间注意力模块工作原理为将输入红外特征向量在空间路径上进行强化, 经过模块中两条路径, 生成增强红外特征向量, 如式(1)所示

$$\mathbf{FE} = \mathbf{F} + \mathbf{M}(\mathbf{F}) \otimes \mathbf{F} \quad (1)$$

式中, \mathbf{F} 表示输入红外特征向量, $\mathbf{M}(\mathbf{F})$ 为红外空间注意力提升向量, \otimes 为逐元素相乘, \mathbf{FE} 为输出的增强红外特征向量。

红外空间注意力提升分支路径由包含1个 1×1 卷积核的卷积层1, 2个混合空洞卷积层包含扩张率 (Dilation Rate, DR) 分别取1, 2, 5的3个 3×3 空洞卷积核, 包含1个 1×1 卷积核的卷积层2, 批标准化层

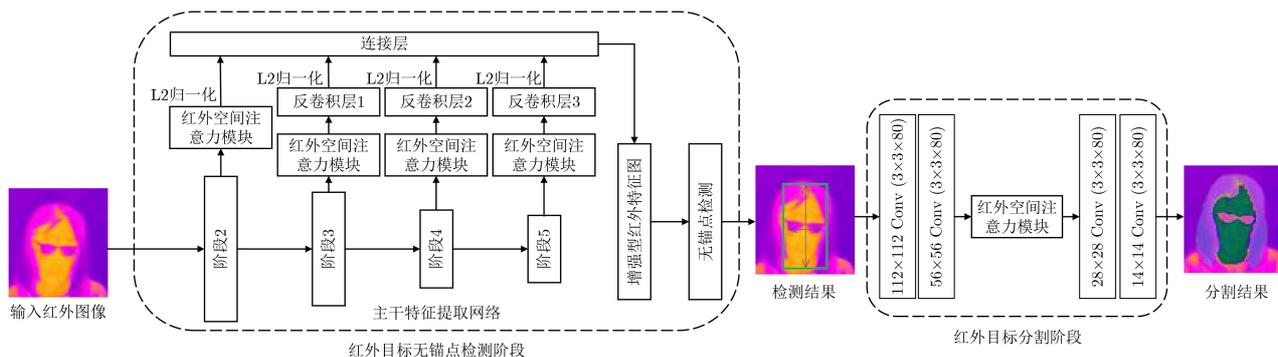


图2 用于红外热成像温度测量区域分割的无锚点实例分割网络结构

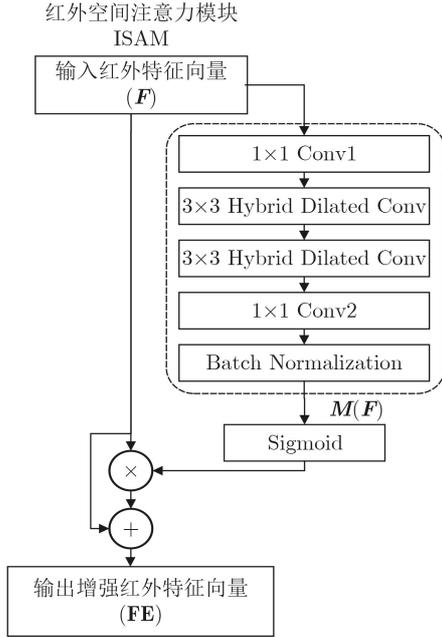


图3 红外空间注意力模块结构

和S型激活函数构成。该路径中，重要的改进在于使用了混合空洞卷积层(Hybrid Dilated Convolution, HDC)，使用混合空洞卷积层可以在提升感受视野的同时，避免使用普通空洞卷积层对红外特征产生的栅格效应，红外空间注意力提升向量 $M(F)$ 的计算如式(2)所示

$$M(F) = \text{BN}(f_1(f_{\text{hd}}(f_{\text{hd}}(f_1(F)))))) \quad (2)$$

式中，BN为批标准化操作， f_1 为 1×1 卷积操作， f_{hd} 为 3×3 混合空洞卷积操作。

2.4 损失函数

在本文提出网络训练中，所使用的损失函数为多任务损失函数，如式(3)所示

$$L = \lambda L_{\text{cls}} + \beta L_{\text{center}} + \alpha L_{\text{box}} + \phi L_{\text{mask}} \quad (3)$$

式中， L 为总损失函数， L_{cls} 为红外目标检测阶段的分类损失函数， L_{center} 为红外目标检测阶段的中心位置损失函数， L_{box} 为红外目标检测阶段的边界回归损失函数， L_{mask} 为红外目标分割阶段的平均掩模交叉熵损失函数。 $\lambda, \beta, \alpha, \phi$ 为平衡几个损失函数分量的参数，根据参考文献[19]分别取1, 0.01, 0.10, 0.25。

L_{cls} 分类损失函数在红外目标检测阶段训练中提供分类损失，使用Softmax Loss函数，如式(4)所示

$$L_{\text{cls}} = -\log_2 \frac{e^s}{\sum_{j=1}^C e^{s_j}} \quad (4)$$

式中， C 为类别个数， s 为正确类别对应的Softmax函数输入。

L_{center} 中心位置损失函数在红外目标检测阶段训练中提供目标中心点位置损失，如式(5)所示

$$L_{\text{center}} = \frac{1}{2N} \sum_{i=1}^N |x_i - c|^2 \quad (5)$$

式中， N 为特征层阶数， x_i 为第 i 阶特征向量， c 为该类别中心点。

L_{box} 边界回归损失函数在红外目标检测阶段训练中提供目标检测框的大小范围损失，如式(6)所示

$$L_{\text{box}} = -\ln \frac{\text{Intersection}(P, G)}{\text{Union}(P, G)} \quad (6)$$

式中，Intersection为预测框与标注框的交叉范围，Union为预测框与标注框的联合覆盖范围， P 为预测框， G 为标注框。

L_{mask} 平均掩模交叉熵损失函数在红外目标分割阶段训练中提供了分割损失，如式(7)所示

$$L_{\text{mask}} = -\log_2(S(y))/2 \quad (7)$$

式中， $y \in (0, 1)$ 为2维分割掩模标签， S 为Softmax函数输出。

3 实验结果与分析

3.1 实验环境

在实验中，软件平台使用TensorFlow平台和Pytorch平台，对应的版本分别为1.12.0和1.1，使用CUDA深层神经网络库10.0。硬件平台使用NVIDIA GTX 1080 GPU(8 GB)进行训练。

3.2 实验数据集

本研究所使用数据集，结合来自IRIS红外/可见面部数据库，Terravic Facial IR数据库的5000张图像以及根据温度测量环境收集的5000张热成像图像。为了有效训练实例分割模型，需要大量的热成像图像。数据集增强是一种用于创建状态与原始数据不同的新数据并进行数据集扩充的方法。本文使用的数据集增强方法包括亮度增强、对比度改变、翻转、反转和旋转，通过原始数据集与增强部分融合，制作了“热成像温度测量区域分割数据集”。训练集、验证集和测试集划分比例为6:3:1，其中验证集与测试集使用原始数据。训练集中训练数据包括：戴口罩的面部、裸露的面部和具有各种遮挡物的面部。训练集中各类数据分布的比例如表1所示。

数据集使用labelme工具进行实例分割的标注。标签类别包括头部裸露的测温区域、头发、帽子、眼镜和口罩。数据集中的所有图像均调整为 512×512 ，以减少计算复杂度。实例分割任务在标注分割类别时，要求对需要分割的每个类别赋予不同颜

色掩模标签(Mask), 每种类别掩模标签对应不同标注色的RGB值对应的标注色如表2所示。

3.3 训练细节

在模型训练期间, 选择随机下降梯度作为优化器, 并将学习率、学习率衰减和动量分别设置为0.001, e^{-6} 和0.9。训练批次大小设置为64, 每个批次包括15个数据。

训练过程中, 准确率不断提高, 损失不断降低。在第180个批次达到了最大的验证准确率, 此后大致保持不变。因此, 在第200个批次之后停止训练, 并保存第200个批次的训练模型。

3.4 评价指标

本文从3个方面评估实时实例分割的性能: 目标预测的准确性, 目标的分割的准确性以及平均处理速度^[20]。

IOU(Intersection-Over-Union)等同于将两个区域的重叠部分除以两个区域的合计部分的结果。正确预测结果(True Positives, TP)表示IOU>0.5的分割结果, 假阳性(False Positives, FP)表示模型预测的阴性样本为阳性, 假阴性(False Negatives, FN)表示模型预测的阳性样本为阴性。在这项研究中, 平均精度(Average Precision, AP)用于评估目标预测的准确性。平均精度定义如式(8)所示

$$AP = \frac{1}{|cl|} \sum_{cl} \left(\frac{1}{|th|} \sum_t \frac{TP(t)}{TP(t)+FP(t)} \right) \quad (8)$$

式中, cl表示检测类别, th表示每一类别的阈值, t为检测样本数。

平均分割精度(Average Mask Precision, AMP)用于在预测目标后评估分割掩模覆盖目标区域的平均精度。平均分割精度定义如式(9)所示

$$AMP = \frac{1}{c} \sum_t \frac{TP(t)}{TP(t) + FP(t)} \quad (9)$$

式中, c表示分割类别总数, t为检测样本数。

每秒帧数(frames per second, fps)用于评估平均处理速度。每秒的帧数定义如式(10)所示

$$fps = \frac{fr}{s} \quad (10)$$

式中, fr为图像帧数, s为秒。

3.5 实验仿真与结果分析

3.5.1 红外空间注意力模块有效性验证

为了验证融入红外空间注意力模块对红外图像特征增强有效性, 采用无注意力提升机制的无锚点实例分割网络, 融合CBAM模块^[19]的无锚点实例分割网络, 融合BAM模块^[20]的无锚点实例分割网络以及本文提出的融合ISAM模块的无锚点实例分割网络进行热成像图像中测温区域的实例分割实验。对比结构均以ResNet50网络为主干特征提取网络, 由于在ResNet50网络阶段2至阶段5提取特征, 检测部分使用无锚点检测方法, 因此训练过程中冻结ResNet50网络输入卷积层与最大池化层, 去除网络末端的平均池化层、全连接层以及softmax层, 在其他层上重新进行训练。用于对比的网络结构均采用前文提到的训练方法与训练参数, 使用本文制作的“热成像温度测量区域分割数据集”进行训练。红外空间注意力模块有效性验证分为红外目标无锚点检测阶段的验证与红外目标分割阶段的验证。

红外目标无锚点检测阶段验证目的在于验证红外空间注意力模块对目标检测效果的提升, 该测试中红外目标分割部分不引入红外空间注意力模块, 在目标检测阶段进行对比验证。测试结果如图4所示。

图4(a)为原始测试红外图像, 图4(b)为真实标签值, 图4(c)为无注意力提升机制的无锚点实例分割网络目标检测效果, 图4(d)为融合CBAM模块的无锚点实例分割网络目标检测效果, 图4(e)为融合BAM模块的无锚点实例分割网络目标检测效果, 图4(f)为本文提出的融合ISAM模块的无锚点实例

表 1 训练集中各类数据分布

数据类型	戴口罩的面部	裸露的面部	存在各类遮挡的面部
比例(%)	40	20	40

表 2 数据集中标签对应的标注色

分割类别	温度测量区域	头发	帽子	眼镜	口罩
R	0	139	0	255	255
G	100	0	255	105	255
B	0	139	127	180	210

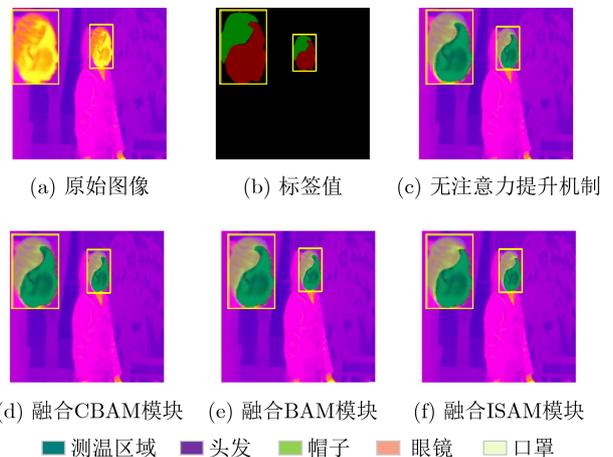


图 4 红外空间注意力模块目标检测阶段有效性测试

分割目标检测效果,从图4可以看出本文提出的融入红外空间注意力模块无锚点实例分割网络对热成像图像中测温区域的检测效果相比其他3种结构更为优良,因为检测效果很大程度决定后续分割效果,因此可以看出在检测阶段本文提出的融入红外空间注意力模块使分割的细节也更为细腻精确。

在“热成像温度测量区域分割数据集”测试集中进行ISAM模块在目标检测阶段有效性对比测试,测试结果统计如表3所示。精度指标方面,本文提出的融合ISAM模块的无锚点实例分割网络在目标检测阶段的检测效果最佳,在AP指标上相对其他3种对比结构达到首位。AP达88.6%,相比无注意力提升机制的无锚点实例分割网络提高4.9%,相比融合CBAM模块的无锚点实例分割网络提高3%,相比融合BAM模块的无锚点实例分割网络提高3.8%。由于目标检测的精度提升对将提升后续目标分割的精度,因此,目标检测阶段融入ISAM模块的网络AMP达84.5%,相比无注意力提升机制的无锚点实例分割网络提高4.5%,相比融合CBAM模块的无锚点实例分割网络提高3%,相比融合BAM模块的无锚点实例分割网络提高3.5%。在实时性指标方面,本文提出的实例分割网络由于融入了红外注意力空间注意力模块加深了网络结构,因此低于无注意力提升机制的无锚点实例分割网络2.5 fps,由于ISAM模块结构上相对其他两种注意力提升模块更为简单高效,因此,速度指标高于融合CBAM模块的无锚点实例分割网络2 fps,高于融合BAM模块的无锚点实例分割网络1 fps。实验结果表明,在目标检测阶段融入ISAM模块有助于增强网络对红外图像目标特征的提取能力,对热成像测温区域的检测与分割效果起到关键的提升作用。

红外目标分割阶段验证目的在于验证红外空间注意力模块对红外目标分割效果的提升,该测试在红外目标部分引入红外空间注意力模块的基础上,进一步在目标分割阶段进行对比验证。测试结果如图5所示。

图5(a)为原始测试红外图像,图5(b)为真实标签值,图5(c)为无注意力提升机制的无锚点实例分割网络分割效果,图5(d)为融合CBAM模块的无锚

点实例分割网络分割效果,图5(e)为融合BAM模块的无锚点实例分割网络分割效果,图5(f)为本文提出的融合ISAM模块的无锚点实例分割网络分割效果,从图5可以看出目标分割阶段融入红外空间注意力模块相对其他3种结构进一步提升了红外目标的分割效果。

在“热成像温度测量区域分割数据集”测试集中进行目标分割阶段的结构对比测试结果统计如表4所示。由于在该阶段的测试仅对分割精度与速度有影响,由表4可见,精度指标方面,AP值均一致,而本文提出的在目标分割阶段融合ISAM模块的无锚点实例分割网络分割效果最佳,AMP指标上相对其他3种结构达到首位。目标分割阶段融合ISAM模块结构AMP达86.5%,相比无注意力提升机制结构提高2%,相比融合CBAM模块结构提高2.2%,相比融合BAM模块结构提高2.4%。在实时性指标方面,本文提出的实例分割网络由于融入了红外空间注意力模块加深了网络结构,因此低于无注意力提升机制结构2 fps,但达到33.5 fps,高于注意力提升模块结构更为复杂的其他2种结构,高于融合CBAM模块结构2.3 fps,高于融合BAM模块的无锚点实例分割网络1 fps。实验结果表明,在目标分割阶段融入ISAM模块有助于进一步提升热成像测温区域分割的能力。

3.5.2 本算法与主流算法的比较

为进一步验证本文提出算法的有效性,在本

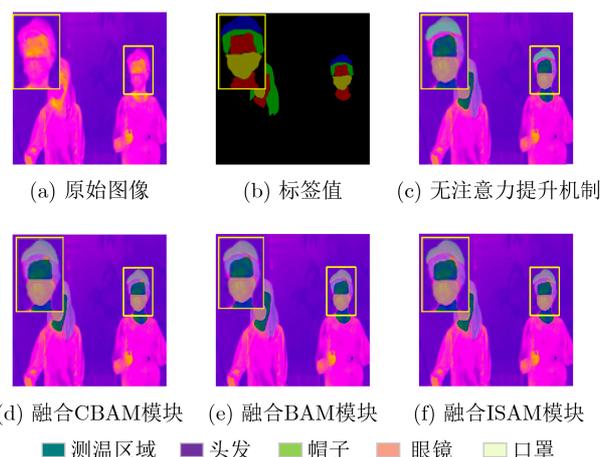


图5 红外空间注意力模块目标分割阶段有效性测试

表3 ISAM模块目标检测阶段有效性验证实验结果

结构	AP(%)	AMP(%)	fps
无注意力提升机制	83.7	80.0	38.0
融合CBAM模块	85.6	81.5	33.5
融合BAM模块	84.8	81.0	34.5
融合ISAM模块	88.6	84.5	35.5

表4 ISAM模块目标分割阶段有效性实验结果

结构	AP(%)	AMP(%)	fps
无注意力提升机制	88.6	84.5	35.5
融合CBAM模块	88.6	84.3	31.2
融合BAM模块	88.6	84.1	32.5
融合ISAM模块	88.6	86.5	33.5

文制作的“热成像温度测量区域分割数据集”上使用其他4种目前主流代表性可见光环境下实例分割方法采用相同训练参数训练得到热成像测温区域实例分割模型,对模型进行对比测试。本文选取用作对比的4种实例分割方法包括在引言部分提到的Mask R-CNN^[10], YOLACT^[11], CenterMask^[12], PolarMask^[13]。其中除Mask R-CNN根据参考文献^[10]以ResNet101网络为主干特征提取网络外,其余参与对比的网络均以ResNet50网络为主干特征提取网络。

在“热成像温度测量区域分割数据集”测试集中进行方法对比测试的结果统计如表5所示。其中由于Mask R-CNN实例分割网络使用了ResNet101网络为主干特征提取网络,目标检测与分割精度由于主干特征提取网络的深度优势而超过了本文未融合红外空间注意力模块网络,但实时性上将会大幅下降。由表5可见,精度指标方面,本文提出的引入红外注意力提升机制无锚点实例分割网络分割效果最佳,在AP与AMP指标上相对其他4种算法达到首位。AP达88.6%,相比Mask R-CNN实例分割网络提高4.8%,比YOLACT实例分割网络提高13%,比CenterMask实例分割网络提高5.8%,比PolarMask实例分割网络提高5%。AMP达86.5%,相比Mask R-CNN实例分割网络提高6.2%,比YOLACT实例分割网络提高16.4%,比CenterMask实例分割网络提高8.4%,比PolarMask实例分割网络提高7.5%。在实时性指标方面,本文提出的实例分割网络由于引入红外注意力提升机制加深了网络结构,因此低于针对分割实时性特别强化的YOLACT实例分割网络4.5 fps,但达到33.5 fps,满足分割实时性要求,同时速度指标高于其他3种算法,高于Mask R-CNN实例分割网络22 fps,高于CenterMask实例分割网络0.5 fps,高于PolarMask实例分割网络3 fps。实验结果表明,本文提出的引入红外注意力提升机制的无锚点实例分割网络对于热成像测温区域实例分割各项性能指标相对其他算法最为优良均衡,适用于在红外热成像图像中实时精确地分割有效测温区域。

表5 热成像测温区域实例分割方法对比实验结果

实例分割方法	AP(%)	AMP(%)	fps
Mask R-CNN	83.8	80.3	11.5
YOLACT	75.6	70.1	38.0
CenterMask	82.8	78.1	33.0
PolarMask	83.6	79.0	30.5
本文方法	88.6	86.5	33.5

4 结论

本文提出一种融入红外注意力提升机制的无锚点实例分割网络用于热成像图像中测温区域的实时实例分割。该实例分割网络在红外目标的检测阶段采用金字塔特征提取方式,在4个提取阶段融入本文设计的红外空间注意力模块以增强红外特征图,提高红外目标检测精度,采用无锚点检测方式提高红外目标检测实时性,在检测完成后的红外目标分割阶段,卷积层中插入红外空间注意力模块提升对红外目标分割的精度。在本文制作的“热成像温度测量区域分割数据集”进行训练与测试,对比多种最新主流实例分割算法,实验结果表明,对于热成像中测温区域的实时实例分割,本文方法在主观视觉和客观评价两个方面都能超越大多数可见光环境下的实例分割算法。

参考文献

- [1] USAMENTIAGA R, VENEGAS P, GUEREDIAGA J, *et al.* Infrared thermography for temperature measurement and non-destructive testing[J]. *Sensors*, 2014, 14(7): 12305–12348. doi: [10.3390/s140712305](https://doi.org/10.3390/s140712305).
- [2] KAMMERSGAARD T S, MALMKVIST J, and PEDERSEN L J. Infrared thermography—a non-invasive tool to evaluate thermal status of neonatal pigs based on surface temperature[J]. *Animal*, 2013, 7(12): 2026–2034. doi: [10.1017/S1751731113001778](https://doi.org/10.1017/S1751731113001778).
- [3] ARORA N, MARTINS D, RUGGERIO D, *et al.* Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer[J]. *The American Journal of Surgery*, 2008, 196(4): 523–526. doi: [10.1016/j.amjsurg.2008.06.015](https://doi.org/10.1016/j.amjsurg.2008.06.015).
- [4] KONTOS M, WILSON R, and FENTIMAN I. Digital Infrared Thermal Imaging (DITI) of breast lesions: Sensitivity and specificity of detection of primary breast cancers[J]. *Clinical Radiology*, 2011, 66(6): 536–539. doi: [10.1016/j.crad.2011.01.009](https://doi.org/10.1016/j.crad.2011.01.009).
- [5] CLARK R P and STOTHERS J K. Neonatal skin temperature distribution using infra-red colour thermography[J]. *The Journal of Physiology*, 1980, 302: 323–333. doi: [10.1113/jphysiol.1980.sp013245](https://doi.org/10.1113/jphysiol.1980.sp013245).
- [6] RASOR J S, ZLOTTA A R, EDWARDS S D, *et al.* TransUrethral Needle Ablation (TUNA): Thermal gradient mapping and comparison of lesion size in a tissue model and in patients with benign prostatic hyperplasia[J]. *European Urology*, 1993, 24(3): 411–414. doi: [10.1159/000474339](https://doi.org/10.1159/000474339).
- [7] BAGAVATHIAPPAN S, SARAVANAN T, PHILIP J, *et al.* Investigation of peripheral vascular disorders using thermal imaging[J]. *The British Journal of Diabetes &*

- Vascular Disease*, 2008, 8(2): 102–104.
- [8] BAGAVATHIAPPAN S, SARAVANAN T, PHILIP J, *et al.* Infrared thermal imaging for detection of peripheral vascular disorders[J]. *Journal of Medical Physics*, 2009, 34(1): 43–47. doi: [10.4103/0971-6203.48720](https://doi.org/10.4103/0971-6203.48720).
- [9] HE Kaiming, GKIOXARI G, DOLLÁR P, *et al.* Mask R-CNN[C]. IEEE International Conference on Computer Vision, Shenzhen, China, 2017: 2961–2969. doi: [10.1109/TPAMI.2018.2844175](https://doi.org/10.1109/TPAMI.2018.2844175).
- [10] BOLYA D, ZHOU Chong, XIAO Fanyi, *et al.* YOLACT++: Better real-time instance segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, To be published. doi: [10.1109/TPAMI.2020.3014297](https://doi.org/10.1109/TPAMI.2020.3014297).
- [11] LEE Y and PARK J. CenterMask: Real-time anchor-free instance segmentation[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 13903–13912.
- [12] XIE Enze, SUN Peize, SONG Xiaoge, *et al.* PolarMask: Single shot instance segmentation with polar representation[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 12190–12199.
- [13] BENINI S, KHAN K, LEONARDI R, *et al.* Face analysis through semantic face segmentation[J]. *Signal Processing: Image Communication*, 2019, 74: 21–31. doi: [10.1016/j.image.2019.01.005](https://doi.org/10.1016/j.image.2019.01.005).
- [14] GHIASS R S, ARANDJELOVIĆ O, BENDADA A, *et al.* Infrared face recognition: A comprehensive review of methodologies and databases[J]. *Pattern Recognition*, 2014, 47(9): 2807–2824. doi: [10.1016/j.patcog.2014.03.015](https://doi.org/10.1016/j.patcog.2014.03.015).
- [15] YANG Changcai, ZHOU Huabing, SUN Sheng, *et al.* Good match exploration for infrared face recognition[J]. *Infrared Physics & Technology*, 2014, 67: 111–115.
- [16] BUDZAN S and WYŻGOLIK R. Face and eyes localization algorithm in thermal images for temperature measurement of the inner canthus of the eyes[J]. *Infrared Physics & Technology*, 2013, 60: 225–234.
- [17] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770–778.
- [18] ZHOU Xingyi, WANG Dequan, and KRÄHENBÜHL P. Objects as points[C]. IEEE Conference on Computer Vision and Pattern Recognition, Los Angeles, 2019: 2678–2689.
- [19] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional block attention module[C]. 15th European Conference on Computer Vision, Munich, Germany, 2018: 3–19.
- [20] PARK J, WOO S, LEE J Y, *et al.* BAM: Bottleneck attention module[C]. European Conference on Computer Vision, Munich, Germany, 2018: 68–86.
- 易 诗: 男, 1983年生, 高级实验师, 研究方向为深度学习红外图像处理.
- 李俊杰: 男, 1997年生, 硕士生, 研究方向为深度学习图像处理.
- 贾 勇: 男, 1986年生, 副教授, 研究方向为穿墙雷达图像处理.

责任编辑: 马秀强