

DOI:10.19651/j.cnki.emt.2209382

基于多尺度特征融合的皮肤烧伤创面分级算法*

韩旭晖 刘宇 何圭波 宋年秀

(青岛理工大学机械与汽车工程学院 青岛 266525)

摘要:为实现遭受重大火灾等灾害后,对伤员皮肤烧伤自动化分级,加快诊断效率,提出一种用于皮肤烧伤分类的轻量化模型 BI-YOLOv5 算法。替换 Swish 激活函数,提高模型收敛能力及检测效率;使用 K-means++ 算法对 anchors 聚类分析,增强对不同尺度目标的适应能力;修改特征提取网络,提取多个尺度的特征信息,建立多尺度特征融合网络,提高模型对深层特征信息的利用率,提高小面积烧伤的识别精度。实验结果表明,BI-YOLOv5 算法在检测并区分不同烧伤类别及环境干扰下烧伤检测有较高的精度和效率,mAP 达到 97.6,对比 YOLOv5 提升 8.4 个百分点。

关键词:深度学习;机器视觉;YOLO;火灾;皮肤烧伤

中图分类号: TP391.7 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Skin burn wound classification algorithm based on multi-scale feature fusion

Han Xuhui Liu Yu He Guibo Song Nianxiu

(College of Automotive and Mechanical Engineering, Qingdao University of Technology, Qingdao 266525, China)

Abstract: In order to realize the automatic classification of skin burns of the wounded after suffering a major fire and other disasters and speed up the diagnosis efficiency, a lightweight model BI-YOLOv5 algorithm for skin burn classification was proposed. Replace the Swish activation function to improve the convergence ability and detection efficiency of the model; use the K-means++ algorithm to perform cluster analysis on anchors to enhance the adaptability to targets of different scales; modify the feature extraction network to extract feature information of multiple scales and establish multi-scale features. The fusion network improves the utilization rate of the deep feature information by the model and improves the recognition accuracy of small-area burns. The experimental results show that the BI-YOLOv5 algorithm has high accuracy and efficiency in detecting and distinguishing different burn types and environmental disturbances, and the mAP reaches 97.6, which is 8.4 percentage points higher than that of YOLOv5.

Keywords: deep learning; machine vision; YOLO; fire disaster; skin burn

0 引言

烧伤是一种常见的组织损伤,在全球意外伤害死亡中排第六位^[1],我国因烧伤而住院的情况近年来呈不断上升趋势^[2],对创面的早期诊断可以有效地控制创口恶化并提高治愈率。在遭遇火灾、爆炸等突发事件时,会出现大量不同等级的烧伤患者,传统烧伤诊断依靠医生的经验判断^[3],且长时间的诊断会对医生造成视觉疲劳及心理负担,增加误诊漏诊的概率。现代医学利用各类光学、声学仪器辅助诊断,如前视红外技术、热成像技术、活组织染色、激光多普勒等^[4],激光多普勒技术^[5]是目前最有效的方法之一,其判断不同等级烧伤的准确率能达到 90% 以上,但此类医学仪器成本高,无法广泛使用。

随着深度学习的不断发展,神经网络在医学领域也得到应用,辅助医生进行诊断治疗。费洪磊等^[6]建立基于二维神经网络的癫痫脑电分类算法用于癫痫检测,有效提高了模型精度及效率;高文波等^[7]利用多尺度生成对抗网络对 CT 图像进行降噪处理,有效提高 CT 图像质量,但算法对浅层信息的利用率较低,且模型运算速度慢;Dezs 等^[8]提出基于 Faster R-CNN 的乳房肿块检测算法,用以检测人眼无法注意到的微小肿块图像,但对相似度较高的目标分辨能力较差,易得出假阳性结果。上述算法为深度学习在医疗领域的应用提供思路,但目前尚无对烧伤创面进行检测的研究,且由于烧伤创面存在相似性高、特征不明显等特点,常规目标检测算法易出现漏检、误检等问题。

收稿日期:2022-03-24

* 基金项目:山东省重点研发计划项目(2019GGX101038)资助

针对上述问题分析,提出一种基于深度学习的轻量化模型 BI-YOLOv5 算法,实现皮肤烧伤创面的大规模、自动化检测和分级。为了满足算法实时性及准确度的要求,在 YOLOv5 算法的基础上,由于创面尺度变化较大,因此首先进行 K-means++ 算法聚类^[9],调整 anchors 的尺寸,加快模型收敛速度;其次用更高效的 Swish 激活函数^[10]代原有 Leaky-ReLU 激活函数^[11],加快模型训练速度及收敛能力;最后重构主干特征提取网络,基于 EfficientDet^[12] 中双向特征融合的思想,建立跨尺度特征融合网络,将主干网络提取到多个尺度的特征图加权融合,使浅层特征信息得到充分利用,提高模型检测精度。实验证明,BI-YOLOv5 算法能够有效地区分不同类别的烧伤创面,且具有较高的精确度及效率。

1 实验数据集

本文所用数据集为 Kaggle 开源皮肤烧伤数据集,该数据集包含未烧伤(0类),I 级烧伤(1类)和 II 级烧伤(2类)的照片共计 1334 张,其中:0 类样本约 800 张,1 类样本约 1 000 张,2 类样本约 400 张,均以 jpg 格式存储。图像尺寸及单张图片包含目标的数量、尺度和种类具有很大的随机性,满足实际诊断情况。图 1 为部分数据集样本图像,其中图 1(a)为未烧伤样本,该类样本易被误检为轻度烧伤,图 1(b)为 I 级烧伤样本,图 1(c)为 II 级烧伤样本。

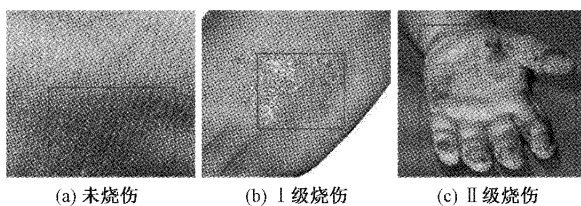


图 1 数据集样本

由于数据集样本容量较小,为了提高算法精度,对数据集进行增强处理,包括翻转、随机剪裁和马赛克等操作。图 2 为数据增强后部分图片。其中,样本信息被外接矩形框包围,矩形框上方所示数字为样本对应的类别所属序号。



图 2 数据增强

2 BI-YOLOv5 算法

其结构如图 3 所示。自深度学习被引入目标检测领域

后,为目标检测带来了许多突破性的进展,当前流行的目标检测算法主要分为两种类型:双阶段网络,代表模型为 R-CNN^[13]等;单阶段网络,代表模型为 SSD^[14]和 YOLO 系列^[15]等。本文将以 YOLOv5 模型为基础进行研究,其结构如图 3 所示。

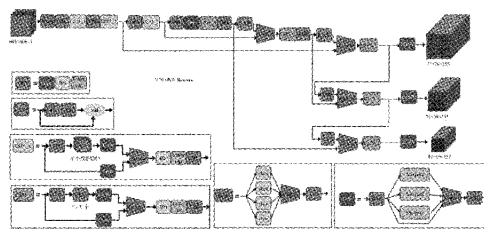


图 3 YOLOv5 网络结构

2.1 K-means++ 聚类

模型收敛速度与 anchors 选择有很大关系,根据不同的目标检测任务应选择与任务目标相对应的 anchors 尺寸。K-means++ 算法在聚类效果上优于传统 K-means 算法,能够显著提升模型收敛速度。算法在样本中随机选择一个作为聚类中心,计算任务中所有样本到这个中心最短距离,利用轮盘法选择下一个中心,重复上述步骤直至得到 K 个中心,随后按照传统 K-means 算法的流程直至聚类中心不再发生变化。

使用 K-means++ 进行聚类分析得到新的 anchors 尺寸如表 1 所示。

表 1 锚框参数

Anchors	Fixed (weight/pixel), high/pixel)
A1	[40,41, 63,53, 64,80]
A2	[99,77, 100,116, 147,118]
A3	[160,180, 242,176, 217,227]

2.2 优化激活函数

神经网络的每个神经元中,输出层输出的数据为矩阵的形式,激活函数的作用是将输出的线性方程转换为非线性方程,提高模型深层神经元的表征能力。

YOLOv5 使用的激活函数 Leaky-ReLU 为非饱和激活函数,对比传统的 ReLU 激活函数,该函数为输入为负的值设置了一个非零的斜率,当输入负值时仍保留一定的梯度,解决了饱和激活函数存在的梯度消失问题,其公式如下:

$$f(x) = \begin{cases} x^i, & x^i \geq 0 \\ \frac{x^i}{\alpha^i}, & x^i < 0 \end{cases} \quad (1)$$

Leaky-ReLU 解决了传统激活函数容易发生梯度消失的问题,在对输入激活函数中为负的值也可以进行计算,但其无法保证正负输入值使用相同的预测关系。本文通过实验对比不同激活函数对模型带来的影响,最终发现 Swish 激活函数能够提高模型检测速度,因此将 YOLO v5 的激

活函数替换为 Swish。因为 Swish 无上边界且有下边界，当输入值为负时，输出结果会逐渐趋近于零，产生很强的正则化效果，且 Swish 处处连续可导，在深层模型上的效果优于 Leaky-ReLU。

2.3 多尺度特征融合网络

YOLO v5 的特征融合网络结构简单，FPN 只建立了自上而下的信息传递通道，对中、大尺度的目标检测效果好，但对小目标的检测效果并不理想；PANet 增加了自下而上的通道，以解决对低层信息利用差的问题，提高对小目标的识别效果，但并未考虑其输出贡献不等的问题；NAS-FPN^[16]利用神经网络搜索技术，在一定空间内进行搜索，建立了一个不规则的拓扑结构提取网络，但计算量也随之增加，模型效率下降。三种网络如图 4 所示。

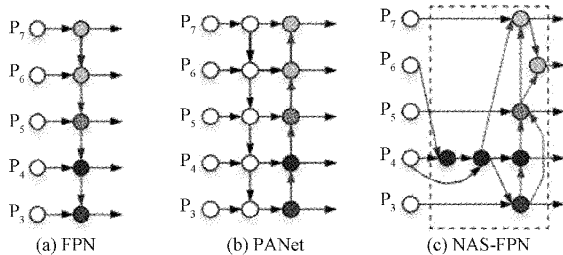


图 4 特征提取网络结构

为了提高 YOLO v5 对小尺度目标的检测效果，在 CPSPDarknet53 主干网络中添加分辨率为 768 * 768 的卷积层，并输出尺度为 64、32、16、8 的四个层级的特征，将不同尺度的特征层相连接，得到双向特征融合网络；为了融合更多的特征，将特征提取网络节点的输入端与输出端直接连接，建立自上而下和自下而上的双向通道设置为一个提取层，同级和跨级连接，得到多尺度特征融合网络。在不引入新数据的前提下，提高模型检测精度效率。网络结构如图 5 所示。

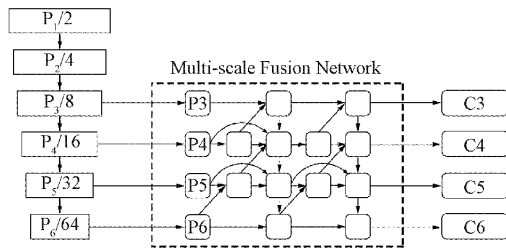


图 5 BI-YOLOv5 网络结构

传统目标检测算法对不同特征采用了相同的处理方式，但由于目标分辨率有差异，导致不同分辨率下提取到的目标特征对模型做出的贡献也不同，若对其进行相同的处理，会导致模型训练效率差。为保证复杂场景下检测精度，本文引入加权特征融合机制，将权重归一化处理，利用 ReLU 函数保证权重为非负值。归一化加权融合公式为：

$$0 = \sum_i \frac{e^{w_i}}{\epsilon + \sum_j e^{w_j}} \cdot I_i \quad (2)$$

其中，在每个 W_i 之后应用 Relu 以确保 $W_i \geq 0, \epsilon$ 以避免数值波动。

3 分析与讨论

3.1 模型训练

实验在 Ubuntu 18.04 操作系统下进行，GPU 为 Intel i7-7700K，显卡为 NVIDIA GeForce GTX 1070，16 GB 内存，环境为 Pytorch 1.9.1、CUDA10.2。

模型训练参数设置如下：输入图片的尺寸为 640 * 640，训练时使用在 VOC 2007 数据集上所得到的权重作为预训练权重，设置 200 个训练周期；超参数中：学习率为 0.0001，训练 100 个 Epoch 后将学习率降至 0.00005，batch_size 为 16，余弦退火参数为 0.12，动量为 0.937。

3.2 评价指标

本文采用目标检测模型常用评估指标，利用 mAP 和 FPS 进行模型评估，其公式如下：

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$R = \frac{TP}{TP + FN} \quad (4)$$

$$mAP = \int_0^1 P(R) dR \quad (5)$$

3.3 实验结果

利用可视化函数，可以很方便地观察训练过程中模型各项参数的变化，如图 6 所示。在训练的前 100 个 Epoch 中，模型各项数值波动幅度较大，因此降低学习率后继续训练。在训练 200 个 Epoch 后，模型各项数据趋于稳定，Precision(精确率)、Recall(召回率)和 mAP@0.5(平均精确度)均达到 0.95 左右，表示模型已达到收敛，训练结果较为理想。

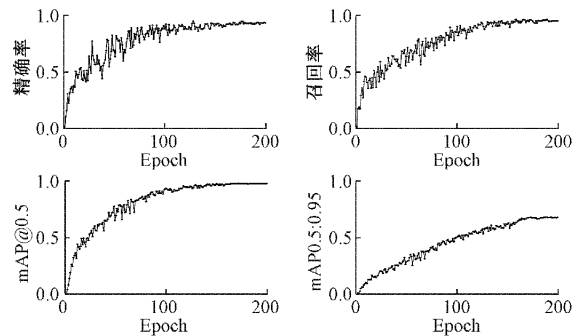


图 6 训练结果

将改进后算法与其他主流目标检测算法在数据集上进行对比测试，其结果如表 2 所示，加粗显示的数值为该指标下最优值，检测结果如图 7 所示。

通过表 2 可以看出，BI-YOLOv5 对比使用浅层主干网络的模型，如 SSD 和 Faster R-CNN 模型，mAP 提升超过 20 个百分点，且每秒帧速率提升也超过 10；对比使用相同主干网络的 YOLOv4 模型，mAP 提升 11.2 个百分点，FPS

表 2 模型对比

模型	Backbone	mAP@0.5/ %	FPS	Size/ MB
SSD	VGG16	72.2	61	93.6
Faster R-CNN	ResNet-50	75.7	39	108.3
YOLOv4	Darknet53	78.8	58	244.2
YOLOv5	Darknet53	89.2	85	56.9
YOLOv5-MS	Darknet53	97.3	83	58.8
BI-YOLOv5	Darknet53	97.6	82	65.5

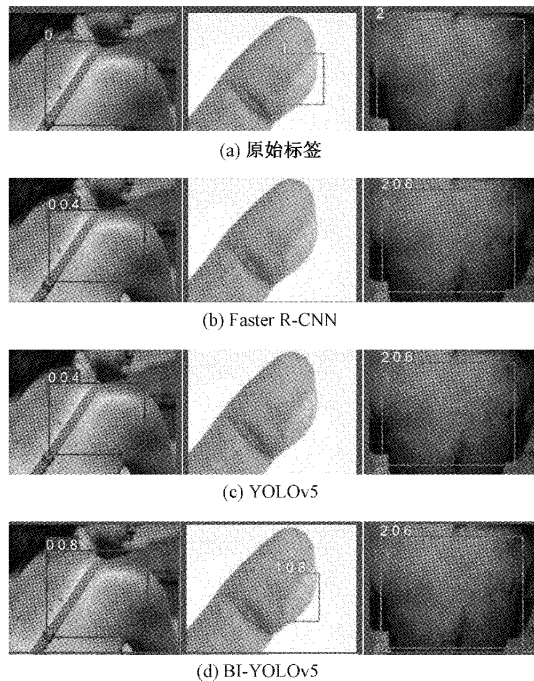


图 7 检测结果

提升 14; 与原始 YOLOv5 模型对比, 得益于更深层的主干网络及激活函数的优化, 在 mAP 提升 8.4 个百分点的基础上, FPS 仅下降 3; 与仅添加多尺度特征融合网络的模型相比, mAP 上升 0.3 个百分点, 在提高精确的同时, 保证了检测效率。

图 7 中分别为数据集原始标签、Faster R-CNN、原 YOLOv5 模型及改进后模型检测结果的对比, 可以看出, 三种模型均能检测出易混淆的未烧伤部位, 但改进后模型置信度更高; 在其他两类图片中, 改进后模型能够准确检测出烧伤部位, 较其他两种算法置信度更高。实验结果表明改进后算法效果更好, 准确度较高。

4 结 论

本文基于 YOLO v5 算法, 针对烧伤创面分类进行改进, 应用 K-means++ 聚类算法对 anchors 进行调整, 优化激活函数, 重新设计特征提取网络并引入双向特征融合网络, 使模型在保证效率的前提下, 提高烧伤创面分类精度。

通过与原始的 YOLOv5 算法进行对比, 实验结果表明, BI-YOLOv5 算法较原始算法有着更高的检测精度, 漏检率更低, 在烧伤初诊中可辅助医生进行初步分级。此外, 算法受数据集限制, 无法对烧伤类别进行检测, 将来可对数据集进行扩充, 提高模型的泛化能力。

参考文献

- [1] 叶翠萍, 李勤, 张晋蔚, 等. 2017—2020 年广州市某医院烧烫伤病例流行病学特征[J]. 职业卫生与应急救援, 2022, 40(1): 90-94, DOI: 10.16369/j. oher. issn. 1007-1326. 2022. 01. 019.
- [2] 张慧君, 赵筱卓, 王成, 等. 223 例成人严重烧伤住院患者的流行病学调查[J]. 中国医刊, 2022, 57(2): 184-190.
- [3] RESCH T R, OSLAND J S, JOST G D, et al. Estimation of burn depth at burn centers in the United States: A survey [J]. Journal of Burn Care & Research: Official Publication of the American Burn Association, 2014, 35(6): 465-533.
- [4] ALICJA R, MARIUSZ K, ANTONI N, et al. Active dynamic infrared thermal imaging in burn depth evaluation [J]. Journal of Burn Care & Research Official Publication of the American Burn Association, 2014(5): 294.
- [5] PAUL D W, GHASSEMI P, RAMELLA-ROMAN J C, et al. Noninvasive imaging technologies for cutaneous wound assessment: A review [J]. Wound Repair & Regeneration, 2015, 23(2): 149-162.
- [6] 费洪磊, 袁琦, 郑玉叶. 基于深度学习的癫痫脑电不平衡分类方法[J]. 仪器仪表学报, 2021, 42(3): 231-240, DOI: 10.19650/j. cnki. cjsi. J2007199.
- [7] 高文波, 孔慧华, 连祥媛. 基于多尺度生成对抗网络的低剂量 CT 去噪算法[J]. 国外电子测量技术, 2021, 40(8): 1-6, DOI: 10.19652/j. cnki. femt. 2102746.
- [8] DEZS, RIBLI, ANNA, et al. Detecting and classifying lesions in mammograms with Deep Learning [J]. Scientific Reports, 2018, DOI: 10.1038/s41598-018-22437-z.
- [9] ARTHUR D, VASSILVITSKII S. K-means ++: The Advantages of Careful Seeding [C]. Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7-9, 2007, ACM, 2007.
- [10] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions [J]. ArXiv Preprint, 2017, ArXiv: 1710.05941.
- [11] XU B, WANG N, CHEN T, et al. Empirical evaluation of rectified activations in convolutional network [J]. ArXiv Preprint, 2015, ArXiv: 1505.00853.
- [12] TAN M, PANG R, LE Q V. EfficientDet: Scalable

- and efficient object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), IEEE, 2020.
- [13] Feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [14] 徐晓光,李海. 多尺度特征在 YOLO 算法中的应用研究[J]. 电子测量与仪器学报, 2021, 35(6): 96-101, DOI:10.13382/j.jemi.B2003350.
- [15] XU B, WANG N, CHEN T, et al. Empirical evaluation of rectified activations in convolutional network[J]. ArXiv Preprint, 2015, ArXiv:1505.00853.
- [16] GHIASI G, LIN T Y, LE Q V. NAS-FPN: Learning scalable feature pyramid architecture for object detection [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2019.

作者简介

韩旭晖, 硕士研究生, 主要研究方向为图像处理、机器视觉等。

E-mail:734798155@qq.com