

DOI:10.16136/j.joel.2022.07.0842

基于改进 CenterNet 的椎间盘图像检测算法研究

周华平, 汪 婷*, 孙克雷

(安徽理工大学 计算机科学与工程学院, 安徽 淮南 232001)

摘要:针对椎间盘图像灰度值较高及成像图具有不均匀性导致的空间信息难捕捉、特征缺乏语义信息等问题,以磁共振腰椎 T2 矢状位的椎间盘识别为目标,本文提出一种基于改进 CenterNet 模型的椎间盘检测算法 TCA_CenterNet (top coordinate attention CenterNet),首先在主干特征提取网络顶层(Top)加入坐标注意力机制(coordinate attention,CA),加强网络对椎间盘的关注度,增强模型对目标位置的敏感性;其次采用深浅层特征融合,增强 CenterNet 提取有效特征的能力,并通过数据增强提高模型的泛化性能。实验结果表明,模型最终的平均精度均值(mean average precision, mAP)达到 81.15%,平均帧率为 14.2 frame/s,与其他对比算法相比,该改进算法具有更好的准确性与鲁棒性。

关键词:椎间盘; CenterNet; 坐标注意力机制; 无锚框检测网络

中图分类号: TP399 **文献标识码:** A **文章编号:** 1005-0086(2022)07-0760-10

Research on intervertebral disc image detection algorithm based on improved CenterNet

ZHOU Huaping, WANG Ting*, SUN Kelei

(School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan, Anhui 232001, China)

Abstract: Aiming at the problems of high gray value of intervertebral disc image and uneven imaging image, it is difficult to capture spatial information and lack of semantic information. Taking the recognition of lumbar intervertebral disc in T2 sagittal position by magnetic resonance as the object, this paper proposes an intervertebral disc detection algorithm, TCA_CenterNet (top coordinate attention CenterNet), based on improved CenterNet model, firstly, the coordinate attention (CA) mechanism is added to the top of the backbone feature extraction network to strengthen the network's attention to the intervertebral disc and enhance the sensitivity of the model to the target position; Secondly, deep and shallow feature fusion is used to enhance the ability of CenterNet to extract effective features, and the generalization performance of the model is improved through data enhancement. The experimental results show that the final mean average precision (mAP) of the model is 81.15% and the average frame rate is 14.2 frame/s. Compared with other comparison algorithms, the improved algorithm has better accuracy and robustness.

Key words: intervertebral disc; CenterNet; coordinate attention (CA) mechanism; no anchor frame detection network

1 引 言

椎间盘退变严重影响患者的生活质量,并且随着年龄的增长其发病率也会随之增加^[1]。磁共

振成像(magnetic resonance imaging, MRI)是目前公认的最有效果的一种无创检测方法, MRI 矢状位 T2 加权像 T2 weighted image, T2WI)对于观察病变特征更具有权威性,但是 MRI 矢状位 T2 图

* E-mail: 2675628905@qq.com

收稿日期: 2021-12-14 修订日期: 2022-01-20

基金项目: 国家自然科学基金(61703005)和安徽省重点研发计划国际科技合作专项(202004b11020029)资助项目

的成像具有不均匀性,且腰椎结构强度、椎间盘形状及位置都有较大的差异性,导致 MRI 成像图的灰度跳跃比较大,诊断的准确性完全依赖于放射科医师的诊断水平,且由于腰椎间盘图像的灰度值较近易产生误诊。

卷积神经网络(convolutional neural network, CNN)的发展带来了两种目标检测网络类型:两阶段检测网络和一级检测网络。两阶段检测网络的第一步是生成图像的感兴趣区域, Fast R-CNN (fast region-based convolutional neural network)^[2]、Faster R-CNN(faster region-based convolutional neural network)^[3]和 R-FCN(region-based fully convolutional networks)^[4]是两阶段检测网络的代表。在一级检测网络中,使用预定义的锚点进行对象分类和对象定位, SSD(single shot multi-box detector)^[5,6]和 YOLO(you only look once)^[7,8]系列是一级检测网络的代表。所有基于 CNN 的对象检测的一个共同特征是特征图在共享网络中编码。

目前,国内已有针对医学图像检测的算法研究。刘森^[9]提出了一种基于自适应提升算法(adaptive boosting, AdaBoost)的 MRI 图像检测算法, AdaBoost 算法首先通过分割脊髓图像来获取椎间盘的边缘及纹理信息,训练出分类器,其次按照椎间盘的灰度值及位置数据检测椎间盘。RUHAN 等^[10]使用 Faster-R-CNN 通过微调更深度深度学习网络来检测椎间盘图像,改善在医学图像处理领域由于缺乏大型标注训练数据集带来的问题。朱亚洲^[11]提出了一种称为深度回声状态网络(deep echo state network, DESN)的 MRI 图像去噪方法,该方法使用随机参数进行初始化,利用

CNN 直接从随机数值张量生成高质量的去噪 MRI 图像,从而将深度神经网络收敛到一个理想点。

针对目前目标检测算法检测密集的椎间盘图像时,产生边界模糊、梯度复杂,缺乏足够的高分辨率信息,特征图提取时缺乏浅层语义信息,导致检测效果不佳,本文提出一种基于改进 CenterNet 的椎间盘图像检测方法,采用中心点回归算法检测椎间盘的位置和类别,融合深层特征语义和浅层特征语义,设计了以深度残差网络(residual network, ResNet)为基础的主干特征提取网络。在深层语义特征和浅层语义特征融合过程中使用坐标注意力模块(coordinate attention, CA)实现特征的加权融合,从而突出有效语义信息、抑制无效语义信息,提高对椎间盘类别的检测精度与速度。

2 基本原理

2.1 CenterNet 网络框架

CenterNet^[12]是单阶段的目标检测网络,在 CornerNet^[13]的基础上进行改进。CornerNet 采取检测物体的左上角点和右下角点来确定目标,但无法感知到内部语义信息,易产生误检。CenterNet 是利用关键点三元组确定一个目标,通过感知内部信息可以有效抑制误检。

如图 1(a)所示, CenterNet 模型使用关键点估计来寻找物体的中心点并返回大小、位置等目标属性。CenterNet 算法无需计算量大且耗时的非极大值抑制(non-maximum suppression, NMS),提高了检测速度。如图 1(b)所示, CenterNet 的中心点为左上角右下角的点融合,并采取高斯函数分布,由中心点向外置信度从 1 下降到 0。CenterNet 相对于两阶段检

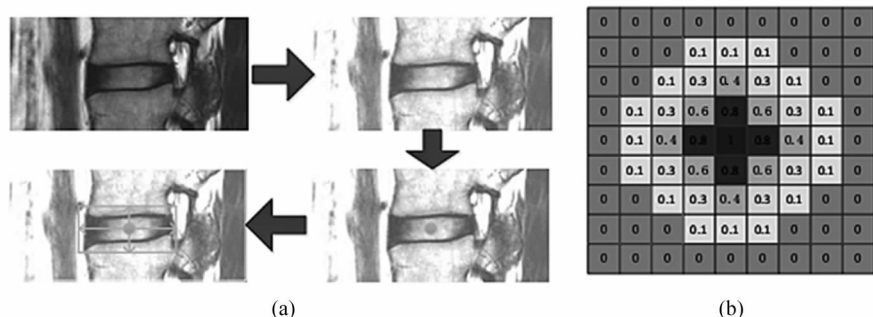


图 1 检测椎间盘目标 CenterNet 热力图:(a) 椎间盘热力图; (b) 高斯标签

Fig. 1 CentrNet thermograph of intervertebral disc target detection: (a) Heatmap of intervertebral disc; (b) Gaussian label

测器更加注重独立分类、回归分支量。

2.2 注意力机制

在计算机视觉领域,将注意力机制应用在目标检测上的模型都取得了良好的效果。SENet

(Squeeze-and-excitation networks)^[14]是经典的通道注意力机制, SENet 包括两个转换:挤压和激励,由于神经网络架构较浅, SENet 的感受野太小,导致 SENet 在大规模图像上无法取得良好的性能。

CBAM(convolutional block attention module)^[15]是基于卷积块的注意力机制,能够有效提高计算效率的一个轻量级的应用模块,但 CBAM 只能捕获局部信息,无法建立远距离的依赖关系。

CA 机制是一种轻量级的注意力机制,它可以同时在通道间和空间方向上获取特征信息与精确位置信息,为了避免 2D 全局池化导致位置信息损失,将通道注意力分解为两个并行的一维特征编码过程,以有效地整合空间坐标。

3 椎间盘检测算法设计

3.1 主干提取网络

考虑到对椎间盘检测的速度和精度要求,改进算法使用的主干网络是 Resnet-50,检测器可以找到中心点并返回大小、方向等属性,使用 Resnet-50 可

以实现速度和精度之间的最佳平衡。

ResNet 的诞生是因为 CNN 的飞速发展,业界普遍认为模型效果随卷积层的增加而增强。然而,多次实验表明,网络深度的增加会导致梯度爆炸甚至消失。后来又提出了随机梯度下降(stochastic gradient descent,SGD)方法,但 SGD 仅能应用在模型深度在十层以内。为了解决此问题,提出了一个新的解决方案:一个名为 ResNet 的深度残差网络。ResNet 的部分输入数据不经过 CNN 直接输出,可有效防止方向传播中的梯度下降问题。

为了结合它们的优势,本文使用 Pytorch 实现了 Resnet-50。基于 Pytorch 的 Resnet-50 的基本步骤如图 2 所示,输入(Input)经过 Resnet-50 主干网络后得到3个子分支,矩阵表示Resnet中每个卷积层的大小输出和训练次数。

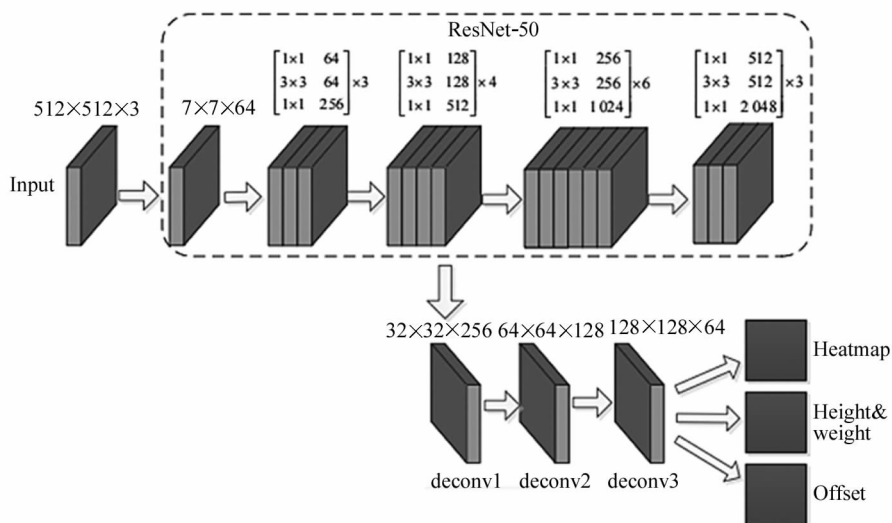


图 2 主干网络结构图

Fig. 2 Diagram of backbone network structure

3.2 融合坐标信息的 CenterNet 网络模型构建

CenterNet 在进行特征提取时,由于医学图像边界模糊、梯度复杂,需要较多的高分辨率信息,因此,将 CA 添加到 ResNet 网络顶层以保留椎间盘图像位置信息,融合坐标信息的 CenterNet 网络框架如图 3 所示,关键点为待检测目标中心,其次分别回归待检测物体的长度、宽度及中心点的偏移量。CA 机制嵌入主干网络顶层,增大网络特征图中包含椎间盘的区域的权重,使得网络更多关注此椎间盘区域,抑制背景区域,该检测结构通过关键点估计找到中心点,回归到其他目标属性进行检测,检测结构包含 Heatmap 层、Offset 层以及 Height&weight 3 个子

分支。

若是输入图片为 $I \in R^{W \times H \times 3}$, 想要预测的目标中心特征图为 $\tilde{Y} \in [0, 1]^{\frac{W}{R} \times \frac{H}{R} \times 3}$, 其中 R 为网络下采样的比例, C 为本次训练数据集中目标的类别数。当 $\tilde{Y} = 1$ 时,则代表此点为目标中心点, $\tilde{Y} = 0$ 时,则为背景。

TCA_CenterNet(top coordinate attention CenterNet)算法将 CA 机制加入主干网络顶层,以加强网络对椎间盘的关注度。CA 将通道注意分解为两个 ID 特征编码过程,两个 ID 特征编码过程分别沿着 x 和 y 两个空间方向聚合特征。通过聚合特征可以捕获沿一个空间方向的远程依赖关系,同时保留可以沿另一个空间方向的精确位置信息。

“x avg pool”指一维水平平均池化,“y avg pool”指一维垂直平均池化。CoordAttention 模块具体操作分为坐标信息嵌入和 CA 生成两个步骤,坐标信息嵌入,通道注意力中常用全局池化对空间进行编码,但池化之后难以保存位置信息,而椎间盘的位置信息对于捕获空间结构至关重要。为了使注意力块在空间上捕捉精确的位置信息,CA 机制将全局池化操作分解为沿着输入特征图的水平方向(kernel(H , 1))和垂直方向(kernel(1, W))方向分别进行池化操作,从而获得输入特征图的 x 、 y 轴相关的位置信息,所以水平方向得到的一维特征可以表示为:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < w} x_c(h, i), \quad (1)$$

式中, $z_c^h(h)$ 表示的是第 c 通道高度为 h 的输出, W 为特征图的宽度, $x_c(h, i)$ 表示的是第 c 通道 h 高度的 i 宽度的特征图的数值。垂直方向得到的一维特征可以表示为:

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < h} x_c(j, w), \quad (2)$$

式中, $z_c^w(w)$ 表示的是第 c 通道宽度为 w 的输出, H 为特征图的高度, $x_c(j, w)$ 表示的是第 c 通道 j 宽度的 w 高度的特征图的数值。

CA 模块的生成步骤包含把两个特征图进行拼接融合,使用两个 1×1 卷积和激活函数。CA 模块的输出表示如式(3)所示:

$$y_c(i, j) \times x_c \times g_c^h(i) \times g_c^w(j') \quad (3)$$

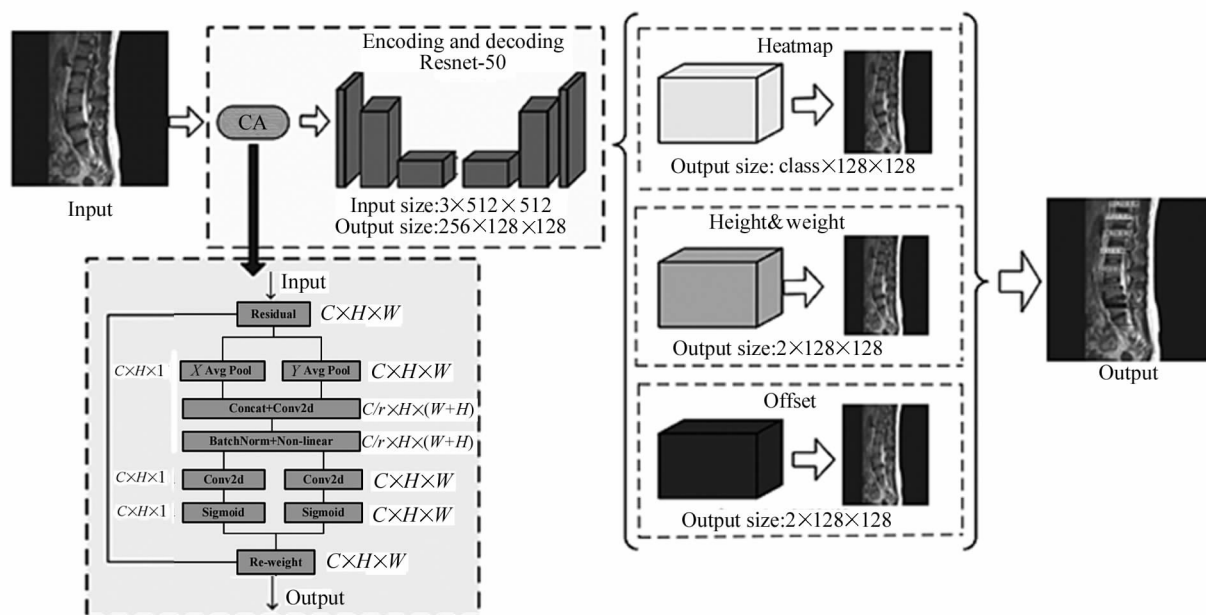


图3 融合坐标信息的 CenterNet 网络模型

Fig. 3 CenterNet network model integrating coordinate information

3.3 特征融合

核磁共振成像的椎间盘图像灰度值较高,且倾斜角在 0° — 45° 之间,导致在编码-解码的过程中浅层特征图提取的特征缺乏语义信息,因此引入深层特征与浅层特征融合机制以提高椎间盘的定位精度,如图 4 所示,在编码-解码过程中,网络编码层特征图由跳连接旁路引出,与解码层特征图进行融合,残差结构中的 1×1 卷积用于跨通道融合,使用坐标空间注意力机制重新调整特征在 X 坐标及 Y 坐标维度上的权重,然后再与解码层特征进行加权融合。

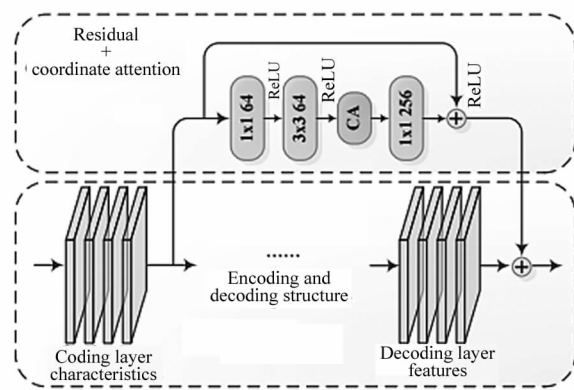


图4 深浅层特征融合

Fig. 4 Deep and shallow feature fusion

3.4 损失函数

在深度学习神经网络模型中,损失函数是网络模型的性能度量。Focal Loss 是一种交叉熵损失函数,通过降低容易分类的样本的权重,将模型的焦点更多地转移到训练中难以分类的样本上,从而解决训练过程中前景和背景极度不平衡的问题。其数学表达式如下所示:

$$FL(p_i) = -(1 - p_i)^\gamma \log(p_i), \quad (4)$$

式中, FL 表示图像上每个像素点的 Focal Loss 损失函数值, $(1 - p_i)^\gamma$ 是调制因子, p_i 表示模型在该像素点上对正类的输出值。从公式可以看出,当 p_i 的值较小,即样本难以分类时, $1 - p_i$ 的值接近于 1, 损失不受影响。相反,若 p_i 的值越接近 1, 如式中所示,当样本易于分类时, $1 - p_i$ 值接近 0。参数 $(1 - p_i)^\gamma$ 使总损耗值接近零。图 5 展示了不同 γ 取值下 $Loss$ 曲线。

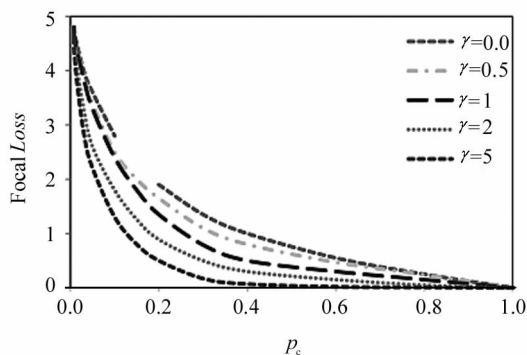


图 5 不同 γ 取值下的 $Loss$ 曲线

Fig. 5 $Loss$ curve under different γ

4 实验结果及分析

4.1 数据集及数据增强

4.1.1 数据集

本文实验采用 spineVocRect 数据集, spineVocRect 数据集包含了从百度飞桨 AI studio 和阿里云天池得到两个数据集, 腰椎间盘突出包含 9 类不同的椎间盘类别, 其中包括了 L5/S、L4/L5、L3/L4、L2/L3、L1/L2、T12/L1、T11/T12、T10/T11、T9/T10, 针对的应用场景是在磁共振成像中。将上述数据集使用 La-belImag 标注工具进行标注, 按照 PASCALVOC 数据集格式存储。图像中亮度和对比度多样, 如图 6 所示, 在对提取出的椎间盘图像进行模板匹配之前, 根据检测出的椎间盘倾斜角进行选取与变形, 按照图像规律, 椎间盘的倾斜角在 0° — 45° 之间。

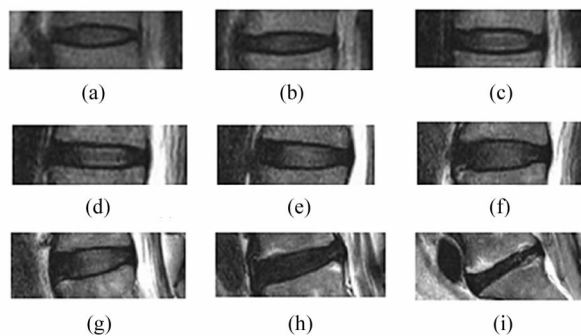


图 6 椎间盘分类图: (a) T9/T10; (b) T10/T11; (c) T11/T12; (d) T12/L1; (e) L1/L2; (f) L2/L3; (g) L3/L4; (h) L4/L5; (i) L5/S

Fig. 6 Classification diagram of intervertebral disc:

(a) T9/T10; (b) T10/T11; (c) T11/T12; (d) T12/L1; (e) L1/L2; (f) L2/L3; (g) L3/L4; (h) L4/L5; (i) L5/S

4.1.2 数据增强

模型的训练需要数以千计的图像信息数据, 但在现实中往往缺乏足够的图像数据。而数据增强对于模型的泛化能力有明显提升, 从而有效提高神经网络模型的识别率。

数据增强是有效地提高神经网络模型的识别率的一种方法。一方面, 数据增强意味着训练的数据量的增加, 模型泛化能力增强, 另一方面提高了模型的鲁棒性。

本文基于 CenterNet 融合了数据增强技术, 采用了仿射变换 warpAffine, 仿射变换 (affine transformation) 的原理是对原图进行裁剪, 然后缩放至 512×512 的大小 (长边缩放, 短边补 0)。实施过程是先确定一个中心点和一个裁剪的长宽, 然后进行仿射变换。

仿射变换是一种矩阵变换, 有多种变换可用于对与图像对相关的几何变换进行建模, 仿射变换通常用于多视图图像配准。仿射变换可以写为如下的形式:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_2 & a_1 & a_0 \\ b_2 & b_1 & b_0 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (5)$$

仿射变换由原图像转变到变换图像的过程如图 7 所示, 图 7 中仿射变换对应着 5 种变换, 平移、缩放、旋转、翻转、错切, 在训练模型时从 3 843 个图像数据中抽取进行仿射变换。

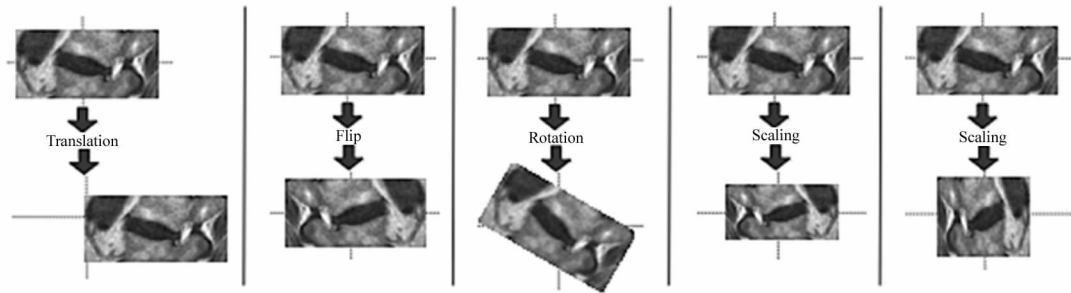


图 7 仿射变换图

Fig. 7 Graph of affine transformation

4.2 实验设置

本实验使用 VOC2007 数据集首先进行 300 个 epoch 的预训练,使用得到的预训练参数对 spineVocRect 数据集进行 400 个 epoch 训练和评估,实验环境配置如表 1 所示。

表 1 实验环境配置

Tab. 1 Experimental environment configuration

Parameter	Disposition
Operating system	Ubuntu 18.04
CPU	Inter Xeon E5-2678 v3
GPU	NVIDIA RTX 2080 Ti
Programing language	Python
Programming platform	Pycharm
Graphics environment	CUDA 10.1, cuDNN 7.6
ML framework	Pytorch 1.5.0

模型首先在 PASCALVOC 数据集上训练,得到可靠的权重参数,再使用 TCA_CenterNet 对 spineVocRect 数据集进行迁移学习,对数据集随机抽取 10% 的图片作为测试样本,剩余 90% 的图片中抽取 10% 作为验证集,其余作为训练集,两个子集覆盖了 9 类需要检测的椎间盘图像,且图像数据之间互不包含,采用余弦学习方案,初始学习率为 0.05。

4.3 评价指标

为了验证 TCA_CenterNet 算法的准确性,实验使用深度学习的评价指标,其中包含召回率 (Recall)、精度 (Precision)、平均精度均值 (mean average precision, mAP)。

具体的表达式如下:

$$Precision = \frac{TP}{TP + FP}, \quad (6)$$

$$Recall = \frac{TP}{TP + FN}, \quad (7)$$

$$mAP = \int_0^1 P(R) dR, \quad (8)$$

式中, TP (true positives) 为正确椎间盘检测数, TN (true negatives) 为漏检的椎间盘检测数, FP (false positives) 为虚检的椎间盘检测数, $P(R)$ 为精度和召回率曲线。

4.4 实验结果与分析

为更加直观地验证 TCA_CenterNet 算法的鲁棒性,选取含有大量不清晰、密集、灰度值较高等干扰因素的 spineVocRect 数据集图像进行训练和测试。通过训练样本数据集,获取改进 TCA_CenterNet 网络模型的总损失函数值变化,如图 8 所示,改进的 TCA_CenterNet 模型迭代到 80 代时,模型的 total loss 值达到稳定状态。

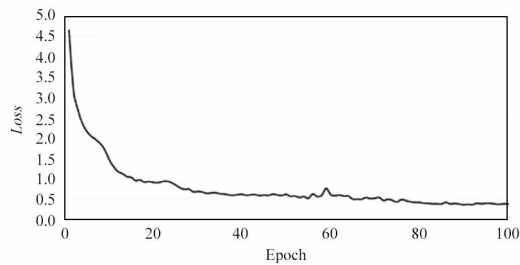


图 8 TCA_CenterNet 网络模型 Loss 函数值的变化

Fig. 8 The change of Loss function value of TCA_CenterNet network model

使用改进后的 CenterNet 模型在数据集上进行实验,实验得到的 $P-R$ (Precision-Recall) 曲线如图 9 所示,其横轴是召回率,纵轴是精度。曲线越靠近坐标 (1,1) 位置,算法的鲁棒性越好。实验得到的算法性能结果如表 2 所示。由表 2 可知,改进 TCA_CenterNet 算法的 Precision、Recall 和 mAP 分别达到 80.47%, 62.24%, 81.15%。实验结果证明,本文改进的算法对 MRI 椎间盘图像的识别精度更高。

表2 改进算法 TCA_CenterNet 的 Precision, Recall 和 mAP 实验结果

Tab. 2 Detection results of Precision, Recall and mAP of improved TCA_CenterNet

Disc type	Precision/%	Recall/%	mAP/%
L1/L2	90.00	69.23	83.89
L2/L3	91.18	72.09	85.97
L3/L4	90.00	83.72	96.22
L4/L5	97.44	88.37	93.72
L5/S	97.44	95.00	99.37
T9/T10	48.31	6.51	53.47
T10/T11	50.00	6.67	54.91
T11/T12	70.59	72.73	76.44%
T12/L1	89.29	65.79	78.96
All	80.47	62.24	81.15

根据设置参数对椎间盘数据集进行训练,图10分别为训练过程中 CenterNet 算法和本文改进算法 TCA_CenterNet 的训练损失曲线,横坐标代表训练次数,纵坐标代表 Loss 值。

由图10(a)训练损失曲线图可知,训练迭代次数 (Iter) 在 0—10 批次之间时,Resnet-50 网络与改进的 Resnet-50+CA 网络之间训练损失与训练精度都比较相近,而在之后的训练过程中,改进的 Resnet-50+CA 网络的平均损失值小于 Resnet-50 网络的平均损失值,由图10(b)训练精度曲线图可知,改进的 Resnet-50+CA 网络较 Resnet-50 网络精度提升较快,随迭代次数的增加,网络训练损失与精度都逐渐趋于稳定,由图说明,改进的 Resnet-50+CA 网络远

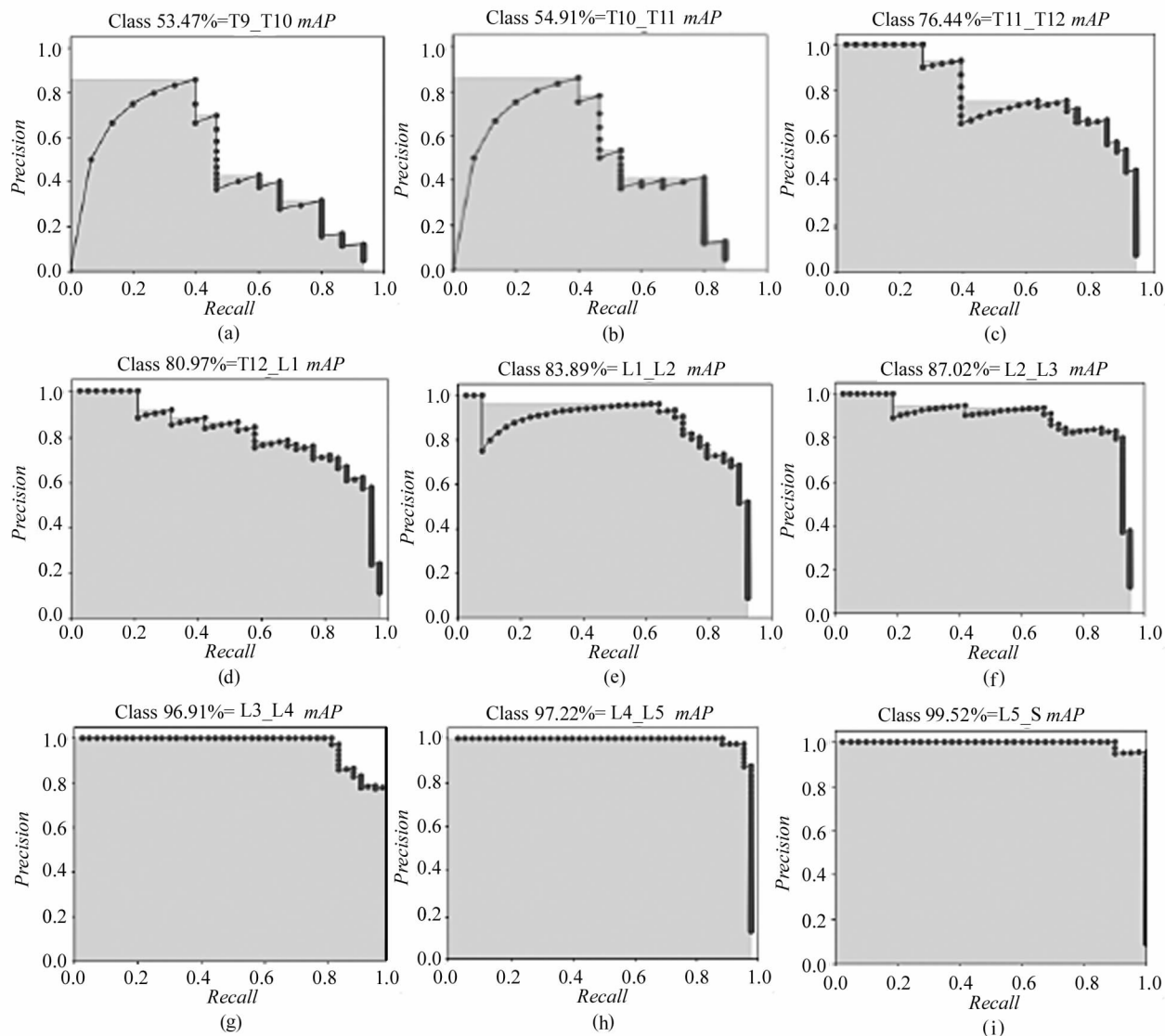


图9 TCA_CenterNet 的 P-R 曲线:(a) T9/T10; (b) T10/T11; (c) T11/T12; (d) T12/L1; (e) L1/L2; (f) L2/L3; (g) L3/L4; (h) L4/L5; (i) L5/S

Fig. 9 P-R curve of TCA_CenterNet:(a) T9/T10; (b) T10/T11; (c) T11/T12; (d) T12/L1; (e) L1/L2; (f) L2/L3; (g) L3/L4; (h) L4/L5; (i) L5/S

优于 Resnet-50 网络,因为模型不仅通过融入 CA 模块增强保留空间信息的能力,且语义特征信息由深浅层特征融合进行增强。

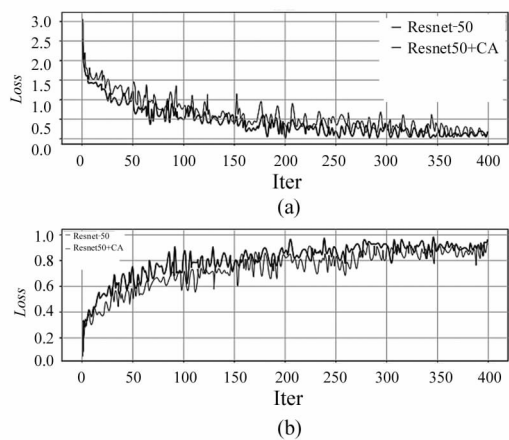


图 10 TCA_CenterNet 训练对比曲线:

(a) 训练损失; (b) 训练精度

Fig. 10 Training comparison curve of TCA_CenterNet;

(a) Train loss; (b) Train accuracy

4.5 对比实验结果分析

为了验证本文改进版 TCA_CenterNet 作为 backbone 网络对椎间盘图像特征提取更具优势,本文选用采用了目标数据集对 SSD、FasterR-CNN、YOLOv3、CornerNet、CenterNet 和改进的 TCA_CenterNet 网络进行识别,表 3 为基于 spineVocRect 数据集下的检测性能对比,从实验结果不难看出,相对于超分辨率测试序列视觉几何群 (visual geometry group, VGG)、Darknet 和 ResNet 系列网络, TCA_CenterNet 网络以少量的参数达到更高的椎间盘图像识别率,同时,本文设计的网络较原版的 ResNet 更好地适应椎间盘检测识别任务,改进的 TCA_CenterNet 算法引入 CA 模块,相对于原始 CenterNet 网络,提升了改进后网络在保持一定检测速度的情况下的检测精度,在实际应用中满足实时性要求。

同时本文进行了一系列的消融实验,相应的结果都列在表 4 中。改进算法一在原算法的基础上采

表 3 spineVocRect 数据集下的检测结果对比

Tab. 3 Comparison of detection results under spineVocRect dataset

Algorithm	Backbone	Processing speed /(frame · s ⁻¹)	mAP/%
SSD	VGG16	12.74	75.2
Faster R-CNN	VGG16	7.8	75.55
YOLOv3	Darknet53	18	69.36
CornerNet	Hourglass-104	4.2	75.8
CenterNet	Resnet-50	14.7	78.09
TCA_CenterNet	Resnet-50+CA	14.2	81.15

取深层特征与浅层特征融合,改进算法二在原算法的基础上嵌入 CA 模块,验证 CA 机制与特征融合的重要性。如表 4 所示。当特征融合和 CA 机制结合在一起时,可得到表 4 中所强调的最佳结果。实验

结果表明,坐标信息嵌入与特征融合具有可比较的学习参数和计算代价,更有利于椎间盘图像检测与识别。

不同注意力机制对 MRI 椎间盘图像的检测算

表 4 消融实验

Tab. 4 Ablation experiment

Algorithm	Feature fusion	CA	Precision/%	Recall/%	mAP/%
CenterNet	×	×	75.10	61.51	78.09
Improved algorithm I	√	×	75.9	61.82	78.99
Improved algorithm II	×	√	78.00	61.29	80.06
TCA_CenterNet	√	√	80.47	62.24	81.15

法的效果如表 5 所示,以原 ResNet50 网络为基础框架,分别加入 SENet 与 CA 机制,对比结果如表 5 所示,原 ResNet 的框架参数量为 25.56 M,与 SE、CBAM、CA 注意力机制的参数量相当,从结果可知,

加入 SE 注意力机制提升了 1.6%,加入 CBAM 注意力机制提升了 1.6%,而加入 CA 机制的网络相对原 ResNet 框架提升了 3.06%,CA 机制可通过精确的位置信息捕捉长距离的依赖关系,对于具有密集椎

间盘图像预测任务更有利。

表 5 不同注意力机制下的检测结果对比

Tab. 5 Comparison of detection results under different attention mechanisms

Algorithm	Parameter/M	mAP/%
ResNet	25.56	78.09
+SE	28.09	79.69
+CBAM	28.09	79.69
+CA	28.30	81.15

为了展示改进的 TCA_CenterNet 算法对于椎间盘图像的检测结果,图 11 所示为本文算法在磁共振成像场景下椎间盘的检测结果,可以看出本文算法对密集椎间盘能够完成精确识别。

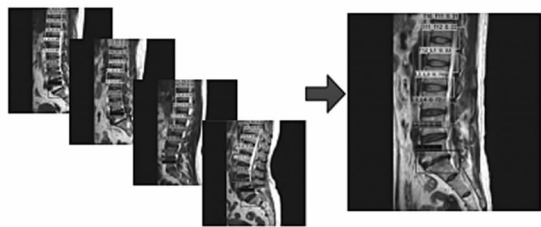


图 11 磁共振成像场景下的椎间盘检测

Fig. 11 Detection of intervertebral disc in magnetic resonance imaging scene

5 结 论

本文基于 CenterNet 框架提出一种改进的椎间盘检测算法,首先通过 CA 机制将所得到的特征图分别编码成一对方向感知和位置敏感特征图以增强感兴趣区域。再选择深浅层特征融合结合深层语义信息强、浅层特征纹理信息强的特点;与原算法对比,改进后的 TCA_CenterNet 在检测精度上提升了 3.06%,参数量减少了 10.7%,检测速度未有明显下降,实验表明,加入空间注意力机制与深浅层特征融合可有效改进目前椎间盘检测因灰度值较高出现的空间信息捕获难的问题,且模型对目标位置的敏感性进一步提升,为椎间盘图像的检测提供了一种较好的选择。

参考文献:

- [1] LI D, WANG L X, WU W J, et al. Quantitative evaluation of MRI T1rho on lumbar disc degeneration in patients with low back pain [J]. Magnetic Resonance Imaging, 2019, 10(12): 904-907.
李丹,王莉霞,吴文骏,等. MRI T1rho 对下腰痛者腰椎间

盘退变的定量评估[J]. 磁共振成像, 2019, 10(12): 904-907.

- [2] KIZRAK M A, BOLAT B. Crowd density estimation by using attention based capsule network and multi-column CNN[J]. IEEE Access, 2021, 9: 75435-75445.
- [3] MENG R, RICE S G, WANG J, et al. A fusion steganographic algorithm based on faster R-CNN[J]. Computers, Materials & Continua, 2018, 55(1): 1-16.
- [4] CHEN F, XU X, TAO Y, et al. An improved recognition method based on YOLOv3-45k for refrigerator item images[C]//2021 6th International Conference on Image, Vision and Computing (ICIVC), July 23-25, 2021, Qingdao, China. New York: IEEE, 2021: 85-89.
- [5] ZHAI S, SHANG D, WANG S, et al. DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion[J]. IEEE Access, 2020, 8: 24344-24357.
- [6] XU G Y, YIN M Y. Improved SSD target detection algorithm based on space channel attention[J]. Journal of Optoelectronics · Laser, 2021, 32(9): 970-978.
许光宇,尹孟园. 基于空间-通道注意力的改进 SSD 目标检测算法[J]. 光电子·激光, 2021, 32(9): 970-978.
- [7] TIAN Y, YANG G, WANG Z, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417-426.
- [8] LAROCA R, SEVERO E, ZANLORENSI L A, et al. A robust real-time automatic license plate recognition based on the YOLO detector[C]//2018 International Joint Conference on Neural Networks (IJCNN), July 8-13, 2018, Rio de Janeiro, Brazil. New York: IEEE, 2018: 1-10.
- [9] LIU M. Research and implementation of whole spine localization and labeling algorithm in MRI image[D]. Wuhan, Huazhong University of Science and Technology, 2017.
刘森. 核磁图像的全脊椎的定位和标记算法的研究与实现[D]. 武汉, 华中科技大学, 2017.
- [10] SA R, OWENS W, WIEGAND R, et al. Intervertebral disc detection in X-ray images using faster R-CNN[C]//2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 11-15, 2017, Jeju Island, Korea. New York: IEEE, 2017: 564-567.
- [11] ZHU Y, PAN X, LV T, et al. DESN: An unsupervised MR image denoising network with deep image prior[J]. The-

oretical Computer Science, 2021, 830: 97-110.

- [12] DUAN K, BAI S, XIE L, et al. Centernet: Keypoint triplets for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 02, 2019, Seoul, Korea (South). New York: IEEE, 2019: 6569-6577.
- [13] LAW H, DENG J. CornerNet: Detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2020, 128(3): 642-656.
- [14] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE Conference on Computer

Vision and Pattern Recognition, June 19-21, 2018, Salt Lake City, Utah, USA. New York: IEEE, 2018: 7132-7141.

- [15] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. Singapore: Springer, 2018: 3-19.

作者简介:

汪 婷 (1998—), 女, 硕士研究生, 主要从事目标检测方面的研究。