

## 基于改进 YOLOv4 模型的脊柱磁共振成像影像检测

代宁<sup>1\*</sup>, 谷玉海<sup>1</sup>, 张志成<sup>2</sup>, 张阳<sup>2</sup>, 徐湛<sup>1</sup><sup>1</sup>北京信息科技大学现代测控教育部重点实验室, 北京 100192;<sup>2</sup>解放军总医院骨科医学部派驻七中心骨科, 北京 100700

**摘要** 针对脊柱解剖结构较为复杂的问题,提出一种用于脊柱磁共振成像影像检测的 YOLOv4-disc 算法。首先,针对真实病例样本数量较少的问题,使用限制对比度的自适应直方图均衡(CLAHE)数据增强方法提高模型的泛化能力。然后,使用 K-means 算法对数据集中真实框的尺寸进行聚类,得到合适的锚框尺寸并确定锚框数量。其次,在 CSPDarknet-53 骨干特征提取网络中使用深度可分离卷积替代普通卷积,减少网络参数并降低运算量。最后,基于 Focal 损失改进原生网络的损失函数,解决 one-stage 目标检测中正负样本比例严重失衡的问题。实验结果表明,所提 YOLOv4-disc 算法的平均精度均值(mAP)达到了 90.80%,相比原生 YOLOv4 提高了 3.51 个百分点。

**关键词** 医用光学;深度学习;YOLOv4;脊柱;目标检测;计算机视觉

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP213059

## Improved YOLOv4 Model-Based Spinal Magnetic Resonance Imaging Image Detection

Dai Ning<sup>1\*</sup>, Gu Yuhai<sup>1</sup>, Zhang Zhicheng<sup>2</sup>, Zhang Yang<sup>2</sup>, Xu Zhan<sup>1</sup><sup>1</sup>Key Laboratory of Modern Measurement and Control Ministry of Education, Beijing Information Science and Technology University, Beijing 100192, China;<sup>2</sup>Department of Orthopedics, PLA General Hospital, Beijing 100700, China

**Abstract** Aiming at the complex anatomical structure of the spine, a YOLOv4-disc algorithm for spinal magnetic resonance imaging image detection is proposed. First, aiming at the problem of small number of real case samples, the adaptive histogram equalization (CLAHE) data enhancement method with limited contrast is used to improve the generalization ability of the model. Second, K-means algorithm is used to cluster the size of real frames in the dataset to obtain the appropriate anchor frame size and determine the number of anchor frames. After that, depth separable convolution is used in CSPDarknet-53 backbone feature extraction network instead of ordinary convolution to reduce network parameters and reduce computation. Finally, the loss function of the native network is improved based on Focal loss to solve the problem that the proportion of positive and negative samples is seriously unbalanced in one-stage target detection. The experimental results show that the mean average precision (mAP) of the proposed YOLOv4-disc algorithm reaches 90.80%, which is 3.51 percentage points higher than that of the native YOLOv4 algorithm.

**Key words** medical optics; deep learning; YOLOv4; spine; object detection; computer vision

## 1 引言

腰椎间盘突出在临床中是一种非常常见的脊柱退行性疾病。据有关统计数据显示,40岁以下的人群中患有脊柱相关疾病的比例超过40%<sup>[1]</sup>。许多诊断方法被应用于脊柱疾病的诊断中。磁共振成像(MRI)不会对人体造成电离辐射损伤,能够获得三维断面成像,其

余多方位的图像也无需重建就可获得,所以关于脊柱疾病的诊断非常依赖于MRI影像<sup>[2]</sup>。如何在MRI及电子计算机断层扫描(CT)图片中定位到所要观察的层面并判断是否患有腰椎间盘突出症和患病的部位是对影像学检查方法及临床医师技术水平、体力、精力的一大挑战。由于我国医疗资源分配还不够均衡,在医疗资源发达的地区对腰椎间盘突出的诊断准确率尚

收稿日期: 2021-11-25; 修回日期: 2021-12-25; 录用日期: 2022-01-11; 网络首发日期: 2022-01-21

基金项目: 北京市科委、京津冀基础研究合作专项(J200012)

通信作者: \*18501301242@163.com

可,但在不发达地区容易出现误诊、漏诊等情况。因此,如何帮助下级医院医护人员充分发掘 MRI 和 CT 影像中的有效病理信息,来提高腰椎间盘突出诊断效率、准确率成为一个亟待解决的问题。

与此同时,基于深度学习的图像识别等人工智能技术飞速发展,在许多算法上都取得重大突破。罗浚铠等<sup>[3]</sup>提出一种基于改进无锚框的目标检测算法,该算法通过构建基于上下文结合的自校准双重注意力模块,有效地提高了在大型多类目标数据集 MSCOCO 上的检测效果。程叶群等<sup>[4]</sup>提出一种适用于嵌入式平台的轻量化目标检测网络(BENet),该网络在轻量化网络 MobileNetv2 的基础上加入通道特征交织模块,在保持较高目标检测精度和速度的同时有效降低计算复杂度。许多医疗领域的专家也与人工智能方向的专家展开深度合作,探讨如何将人工智能技术于医疗领域:侯高峰等<sup>[5]</sup>使用一种基于深度学习级联架构参数

优化的检测模型,有效解决了眼底病变的多尺度和小目标问题;马树志<sup>[6]</sup>使用全卷积神经网络提取影像深层次特征,进而实现了自动分割 CT 影像的功能。由于当前深度学习在脊柱诊断方面的应用依然局限于脊柱分割方面,在脊柱 MRI 影像方面的应用极少。因此,针对以上问题,本文提出一种用于脊柱磁共振成像影像检测的 YOLOv4-disc 算法。所提算法以当下最优秀的目标检测算法之一——YOLOv4 作为基准,并从数据增强、网络结构、损失函数等方面对其进行改进,使其达到更高的检测精度。

## 2 YOLOv4 网络简介

YOLOv4 目标检测模型是由 Bochkovskiy 等<sup>[7]</sup>于 2020 年提出的,是在 YOLOv3<sup>[8]</sup>网络模型的基础上,组合深度学习领域多项研究成果的模型。YOLOv4 网络结构如图 1 所示。

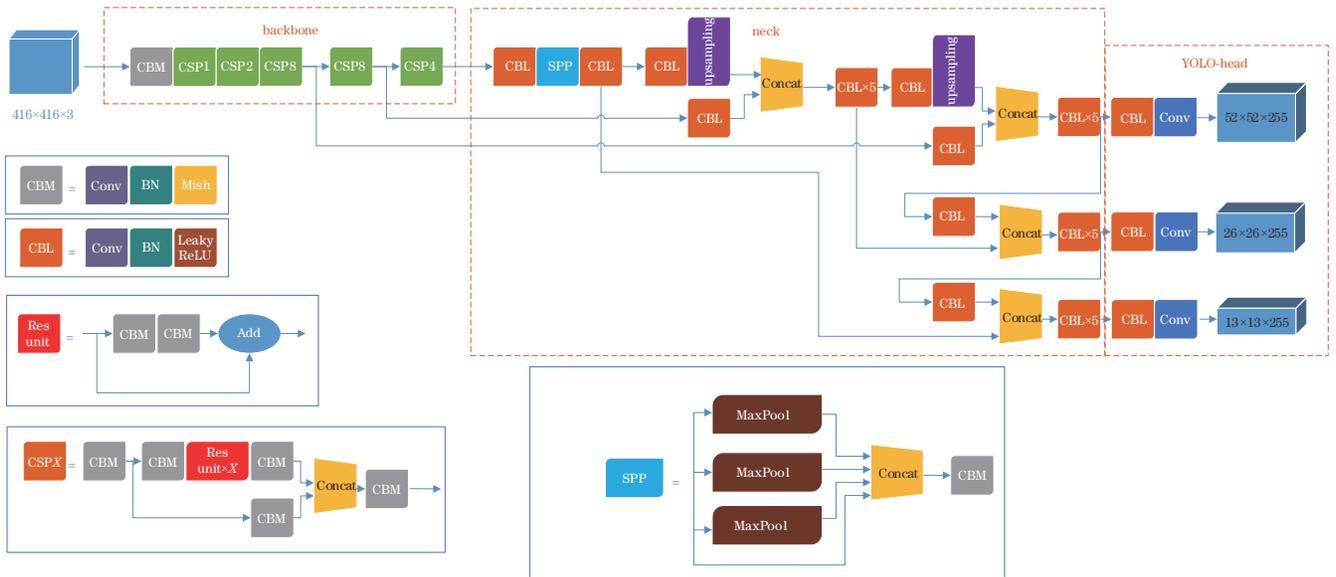


图 1 YOLOv4 网络结构图

Fig. 1 YOLOv4 network structure

### 2.1 网络结构

YOLOv4 的 CSPDarkent-53 主干特征提取网络在 Darkent-53 网络的基础上引入跨阶段局部(CSP)网络块思想,并将其中的 RES 网络块替换成 CSPNet。CSPNet 结构主要包括两个分支,上面的分支由两个 CBM 卷积层和 X 个残差组件构成,下面的分支有一个 CBM 卷积层,最后通过 Concat 操作将两个分支的输出连接起来。下方的分支是一条可以绕过很多残差结构的大残差边,这种拼接方式可以在减少计算量和降低内存成本的同时保证准确率从而使得卷积神经网络具有更强的学习能力<sup>[9]</sup>。颈部(neck)网络与 YOLOv3 的颈部网络相比,首先增加了一个空间金字塔池化(SPP)模块,同时也对特征金字塔网络(FPN)结构进行了改进,增加了两个 path-aggregation neck(PAN)结构。在提取 3 个不同尺度的图像特征后,将其传入

SPP 和 PAN 结构进行特征融合。最后将输出的 3 个尺度特征图输入 YOLO-head 部分进行预测,并输出结果。

### 2.2 激活函数

由图 1 可以看出,YOLOv4 将 YOLOv3 中大部分的 CBL 结构替换成 CBM 结构。CBM 和 CBL 结构的区别就在于激活函数是 Leaky ReLU 还是 Mish。Mish 激活函数的表达式为

$$f(x) = x \cdot \tanh[\zeta(x)], \quad (1)$$

$$\zeta(x) = \ln[1 + \exp(x)]. \quad (2)$$

Mish 激活函数的图像如图 2 所示。Mish 激活函数的无下界属性有助于实现强正则化效果。同时作为非单调函数,可以保持小的负值,从而稳定网络梯度流。

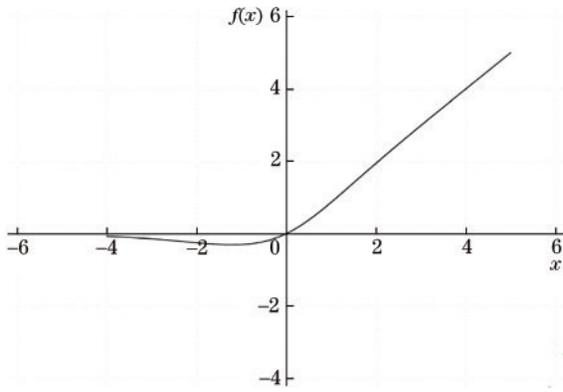


图 2 Mish 激活函数

Fig. 2 Mish activation function

### 2.3 损失函数

在网络模型的训练初期,由于预测准确率较低,需要设置损失函数以衡量预测值和实际值之间的差值,来达到反向更新网络参数的目的。目标检测网络通常会设置 3 种损失函数,分别是用于目标框定位的回归损失函数 ( $L_{CIoU\_loss}$ )、分类损失函数 ( $L_{Conf\_loss}$ ) 和置信度损失函数 ( $L_{Class\_loss}$ )。损失函数的表达式为

$$L = L_{CIoU\_loss} + L_{Conf\_loss} + L_{Class\_loss}, \quad (3)$$

$$L_{CIoU\_loss} = 1 - \left[ R_{IOU} - \frac{\rho^2(b, b^g)}{c^2} - av \right], \quad (4)$$

$$a = \frac{v}{(1 - R_{IOU}) + v}, \quad (5)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^p}{h^p} \right)^2, \quad (6)$$

式中:  $R_{IOU}$  为预测框与真实框的交并比;  $b^{gt}$  为真实框的中心点位置;  $b$  为预测框的中心点位置;  $\rho$  为真实框和预测框中心点之间的欧氏距离;  $c$  是同时包含预测框和真实框的最小封闭区域的对角线距离;  $av$  是对长宽比的惩罚项,  $v$  表示度量长宽比的相似性系数<sup>[10]</sup>;  $w^{gt}$ 、 $h^{gt}$  和  $w^p$ 、 $h^p$  分别表示目标框的宽、高和预测框的宽、高。

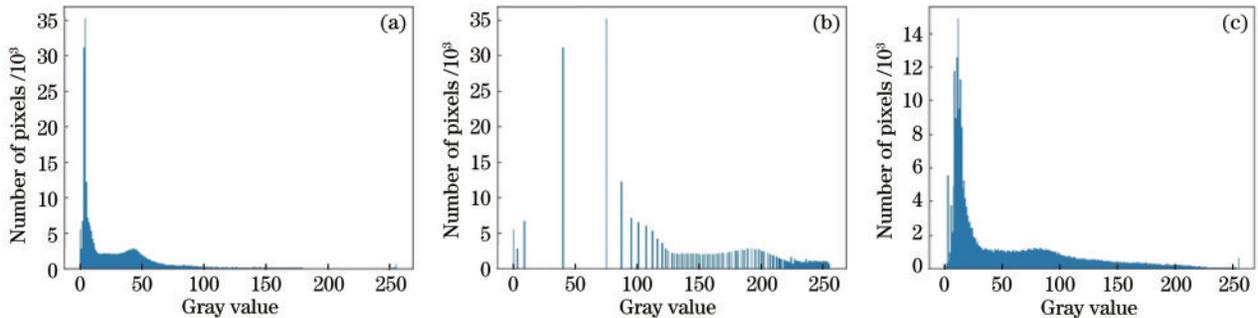


图 4 变换前后灰度值直方图。(a)原图; (b)直方图均衡化; (c) CLAHE

Fig. 4 Histogram of gray value before and after transformation. (a) Original drawing; (b) histogram equalization; (c) CLAHE

针对直方图均衡化造成的图像局部纹理消失的问题,本研究使用限制对比度的自适应直方图均衡化 (CLAHE) 算法,效果如图 5(c) 所示。从图中可以看

## 3 改进 YOLOv4 的 MRI 影像检测分析

### 3.1 数据预处理

脊柱侧面 MRI 影像主要包括腰椎间盘、脊神经、椎体等部分,如图 3 所示。要诊断患者是否患有腰椎间盘突出症,不仅需要检测 MRI 影像中腰椎间盘的形态,还要根据脊神经的纹理、形态等等因素综合判断。但由于实际医院影像拍摄环境中的各种因素会影响到影像的成像质量,例如当影像的对比度不明显时,会极大增加检测的难度,可能会出现误诊、漏诊的情况。因此,本研究首先采用图像数据预处理的方式使脊柱 MRI 影像特征更加明显。

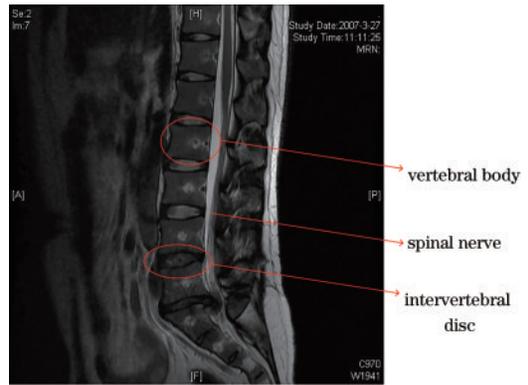


图 3 脊柱 MRI 影像

Fig. 3 MRI image of spine

常见的图像数据预处理方式有直方图均衡化、膨胀、腐蚀等。直方图均衡化首先使用直方图对图像灰度值的数据空间进行量化,直方图可以反映出图像中所有像素的灰度值分布情况,之后再对灰度直方图进行非线性拉伸,重新分配各个灰度单位中的像素点数量,使一定灰度范围内像素点的数量大致相等,如图 4(b) 所示。经过直方图均衡化的图像,对比度会得到增强,但会丢失图像局部的纹理信息,如图 5(b) 所示,图中右侧脊神经的纹理形态已经模糊。

出,CLAHE 在提高椎间盘和脊神经对比度的同时,也很大程度上保留了脊神经的纹理特征。



图 5 均衡化前后影像。(a)原图；(b)直方图均衡化；(c) CLAHE

Fig. 5 Images before and after equalization. (a) Original drawing; (b) histogram equalization; (c) CLAHE

### 3.2 新设定锚框的尺寸与数量

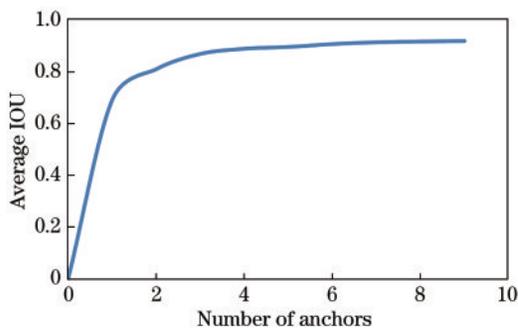
由于 YOLOv4 网络的初始锚框尺寸是基于公开数据集的聚类结果得到的,而公开数据集的真实框种类繁多,尺寸差别较大,因此不宜在训练中直接使用初始锚框。为了使锚框和本次实验真实框的尺寸尽可能匹配,需要对训练集中真实框的尺寸使用 K-means 算法进行聚类操作。算法中对距离公式的定义为

$$d(\text{box}, \text{centroid}) = 1 - R_{\text{IOU}}(\text{box}, \text{centroid}), \quad (7)$$

$$R_{\text{IOU}} = \frac{\text{Area}(A \cap B)}{\text{Area}(A \cup B)}. \quad (8)$$

由于每个检测尺度下的锚框数量会影响网络模型的检测效果,锚框数量较多时可以使预测框的回归运算匹配度更高,提高模型预测的召回率。但过多的锚框会导致计算量增大,从而影响模型的检测速度。因此,合适的锚框数量需要同时兼顾网络模型的检测性能和检测速度。

将 K-means 算法中的锚框数量分别设置为 1~9,得到不同的锚框尺寸,同时计算不同锚框数量下的平均  $R_{\text{IOU}}$  值,结果如图 6 所示。从图 6 可以看出:当锚框数量在 [0, 3] 范围内增加时,平均  $R_{\text{IOU}}$  效果提升显著;而在 [4, 9] 范围内增加时,平均  $R_{\text{IOU}}$  值变化不明显。为了确保模型的检测准确率,同时提高模型的检测速度,满足在实际检测环境中对实时性的要求,将锚框的数量设置为 3。输入图像的尺寸为 (512×512),作为锚框的聚类中心的尺寸为 (72, 55)、(77, 38)、(398, 508)。

图 6 不同锚框数量的平均  $R_{\text{IOU}}$ Fig. 6 Average  $R_{\text{IOU}}$  of different anchor frames

### 3.3 YOLOv4 网络结构优化

原生 YOLOv4 网络的骨干特征提取网络为

CSPDarknet-53, 该网络足以提取脊柱 MRI 影像中的椎间盘特征信息,但由于网络参数较多、结构复杂,不能很好地满足实际诊断环境对实时性的要求。因此,需要在保证图像特征提取效果的同时,简化特征提取网络,减小网络参数数量和运算量。本研究采用更为轻量级的 MobileNetv3<sup>[41]</sup> 替换 CSPDarknet-53。MobileNet 内部的卷积层为深度可分离卷积 (DSC),能在对精度的影响可以忽略的情况下,大幅降低卷积层的参数数量和网络的运算量。在通用卷积中,每个卷积核 (filter) 的通道数和输入图片的通道数保持相同,而 DSC 则将卷积过程分为逐通道卷积和逐点卷积两步,如图 7 所示。



图 7 深度可分离卷积

Fig. 7 Depthwise separable convolution

对于一个尺寸为  $W \times H \times C$  的输入图像来说,使用  $K$  个尺寸为  $Z \times Z \times C$  大小的 filter 进行卷积运算,设置 stride 为 1, padding 为 0,则输出特征图的尺寸为  $N \times M \times K$ :

$$N = W - Z + 1, \quad (9)$$

$$M = H - Z + 1. \quad (10)$$

则通用卷积在此次运算过程中,总运算次数为

$$C_{\text{conv}} = K \times Z \times Z \times C \times N \times M, \quad (11)$$

逐通道卷积的运算次数为

$$C_{\text{depth\_conv}} = C \times Z \times Z \times 1 \times N \times M, \quad (12)$$

逐点卷积的运算次数为

$$C_{\text{pnt\_conv}} = K \times 1 \times 1 \times C \times N \times M, \quad (13)$$

因此深度可分离卷积的总运算次数为

$$C_{\text{DSC}} = (Z \times Z + K) \times C \times N \times M, \quad (14)$$

则 DSC 与通用卷积的计算量之比为

$$\frac{C_{\text{DSC}}}{C_{\text{conv}}} = \frac{1}{K} + \frac{1}{Z^2}. \quad (15)$$

通常来说,为了提取输入图像的多种特征,会将  $K$  取一个较大的值,则  $1/K$  的值可忽略。由于一般将每个卷积核的尺寸设置为  $3 \times 3$ ,即  $Z = 3$ ,因此深度可分

离卷积的计算量约为通用卷积的 1/9,可大幅降低原生 YOLOv4 网络的运算量<sup>[12]</sup>。

同时由于 neck 网络部分中 PANet 的两个下采样

和特征融合操作容易导致小目标物体的特征丢失,将其去除,并在  $52 \times 52$  的尺度下进行特征融合,改进后的网络结构如图 8 所示。

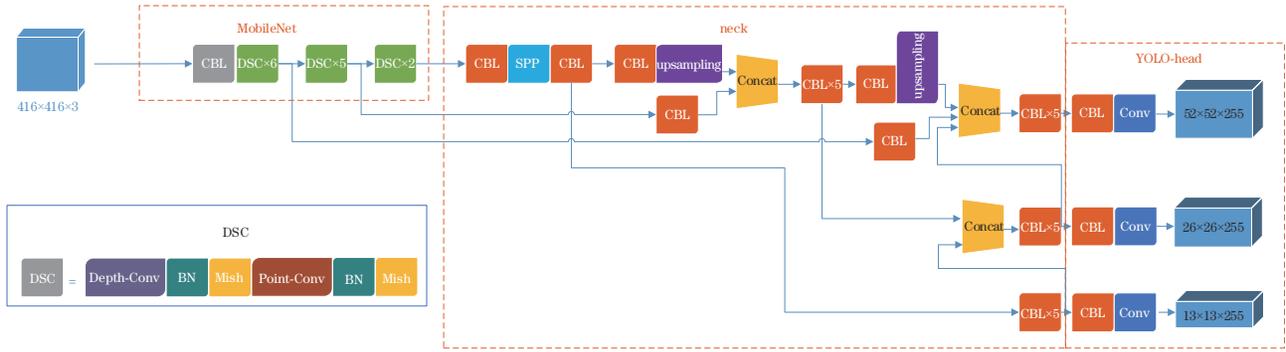


图 8 优化后的 YOLOv4 网络结构图

Fig. 8 Optimized YOLOv4 network structure

### 3.4 损失函数改进

YOLOv4 的损失函数主要由 3 个部分组成,分别是置信度损失、目标框定位损失、分类损失。在脊柱 MRI 影像中,作为诊断依据的椎间盘末端部分仅占整张影像的一小部分,大部分区域都不能作为诊断依据。将包含椎间盘末端的部分作为图像中的正样本,包含其他部分的部分作为图像的负样本。对于正样本,需要计算 3 个部分的损失并相加,而对于负样本,由于其不包含目标,只需要计算置信度损失。

原生 YOLOv4 网络使用的置信度损失的表达式<sup>[13]</sup>为

$$L_{ce} = -y \log y' - (1-y) \log(1-y') = \begin{cases} -\log y', & y=1 \\ -\log(1-y'), & y=0 \end{cases} \quad (16)$$

式中: $y$ 为真实值; $y'$ 为预测值;1表示正样本;0表示负样本。由于在脊柱的 MRI 影像中,负样本数量会远远大于正样本的数量,网络更倾向于给出单元格中不包含目标的预测,出现漏检的情况。

焦点损失 (FL) 函数<sup>[14]</sup>本质上是对交叉熵函数的一种改进,通过在原生 YOLOv4 交叉熵损失函数上设置调制系数和样本平衡系数来解决正负样本失衡的问题。首先为了减少易分类的样本的损失,引入一个大于零的因子  $\gamma$ ,并增加正样本的调制系数  $(1-y)^\gamma$ 、负样本的调制系数  $y^\gamma$ ,另外,为了解决正负样本失衡的问题再引入另一个平衡因子  $\alpha$ 。则改进后的 FL 函数的表达式为

$$L_{FL} = -\alpha y (1-y')^\gamma \log y' - (1-\alpha)(1-y) y'^\gamma \log(1-y') = \begin{cases} -(1-y')^\gamma \log y', & y=1 \\ -(1-\alpha) y'^\gamma \log(1-y'), & y=0 \end{cases} \quad (17)$$

则改进后的 YOLOv4 置信度损失函数为

$$L_{CONF} = -\sum_{i=0}^{S \times S} \sum_{j=0}^M I_{ij}^{obj} [\hat{C}_i \log C_i + (1-\hat{C}_i) \log(1-C_i)] - \lambda_{noobj} \sum_{i=0}^{S \times S} \sum_{j=0}^M I_{ij}^{noobj} [\hat{C}_i \log C_i + (1-\hat{C}_i) \log(1-C_i)], \quad (18)$$

式中: $S \times S$ 为当前检测尺度下图像被划分的网格数量; $M$ 为当前尺度下锚框的数量; $I_{ij}^{obj}$ 表示编号为  $i$  的网格的第  $j$  个锚框处有目标, $I_{ij}^{noobj}$ 表示没有目标; $C_i$ 和  $\hat{C}_i$ 分别表示第  $i$  个网格存在目标的置信度的预测值和真实值; $\lambda_{noobj}$ 为权重系数; $\gamma$ 一般取 2。

## 4 实验结果及分析

### 4.1 实验数据

YOLO 系列目标检测网络属于监督学习,需要在数据集中手动标注出真实框位置及种类。本实验所标记的标签特征种类和数量如表 1 所示。由于腰椎间盘末端根据突出的严重程度不同可以分为健康、膨出、突出等 3 种类型,严重程度依次递增。数据集资源为真实患者脊柱 MRI 影像,共 842 张,包含 3 种类型的椎间盘。考虑到影像资源较少,使用 Mosaic 数据增强和 CLAHE 图像增强对原始数据集进行扩充,扩充后的数据集数量为 2564 张。对训练集、验证集、测试集采用 6:2:2 的比例进行划分,训练集、验证集、测试集分别包含 1539 张、512 张、513 张真实患者腰椎间盘 MRI 影像资料。数据增强后的数据集较为复杂,可以提高网络模型的鲁棒性。

表 1 特征标签

Table 1 Feature label

Feature label	Number
Ptrotusion	6410
Swollen	3846
Healthy	2564

## 4.2 实验配置与训练

实验环境配置如下:标注工具为 LabelImg 开源工具;CPU 使用英特尔酷睿 i7-10700 3.00 GHz;显卡使用英伟达 RTX2080;运算平台使用英伟达 CUDA 8.0、CuDNN 6.0;操作系统为 ubuntu 20.04.05(64位);网络基于 PyTorch 框架实现。

网络训练参数配置如下:迭代次数设置为 20000;将输入图片的 batch、subdivisions 分别设置为 64、16;采用分步策略,初始值为 0.001,比率为 0.12。

## 4.3 模型评价指标

为了验证所提模型的性能,选取 5 种常用的评价指标来衡量网络模型的性能,分别是精确率( $P$ )、召回率( $R$ )、平均精度(AP)、每秒帧数(FPS)及 F1 值<sup>[15]</sup>:

$$\begin{cases} P = \frac{T_P}{T_P + F_P} \times 100\% \\ R = \frac{T_P}{T_P + F_N} \times 100\% \\ F_1 = \frac{2 \times P \times R}{P + R} \\ R_{AP} = \int_0^1 P(R) dR \end{cases}, \quad (19)$$

式中: $T_P$ 、 $F_P$ 、 $F_N$  分别表示真阳性样本、假阳性样本、假阴性样本的数量。

## 4.4 实验结果分析

使用原生 YOLOv4 损失函数的网络训练损失变化曲线如图 9(a)所示,使用焦点损失函数改进过的 YOLOv4-disc 网络损失变化曲线如图 9(b)所示。在迁移学习训练过程中,首先将特征提取网络前端的 50 个 epoch 冻结,之后再将其解冻。从图中可以看出,在解冻后 YOLOv4-disc 的损失函数大幅下降,收敛速度相比于原生 YOLOv4 更快,并低于原生 YOLOv4 模型的损失,最终稳定在 0.8 左右。

表 2 为不同网络模型的验证结果。从表 2 可以看出,原生 YOLOv4 网络相比于 YOLOv3,在 3 种类型椎间盘上的检测精确率和召回率均有不同程度的提升。但由于网络参数的增多,模型大小增加了 117 MB,降低了网络的轻量化程度。而在 YOLOv4-disc 中,由于深度可分离卷积等改进方法,在提高精确率、召回率和 FPS 的同时也降低网络模型大小,提升了轻量化程度。YOLOv4-disc 网络对 3 种椎间盘类型的精确率-召回率( $P-R$ )曲线图如图 10 所示。

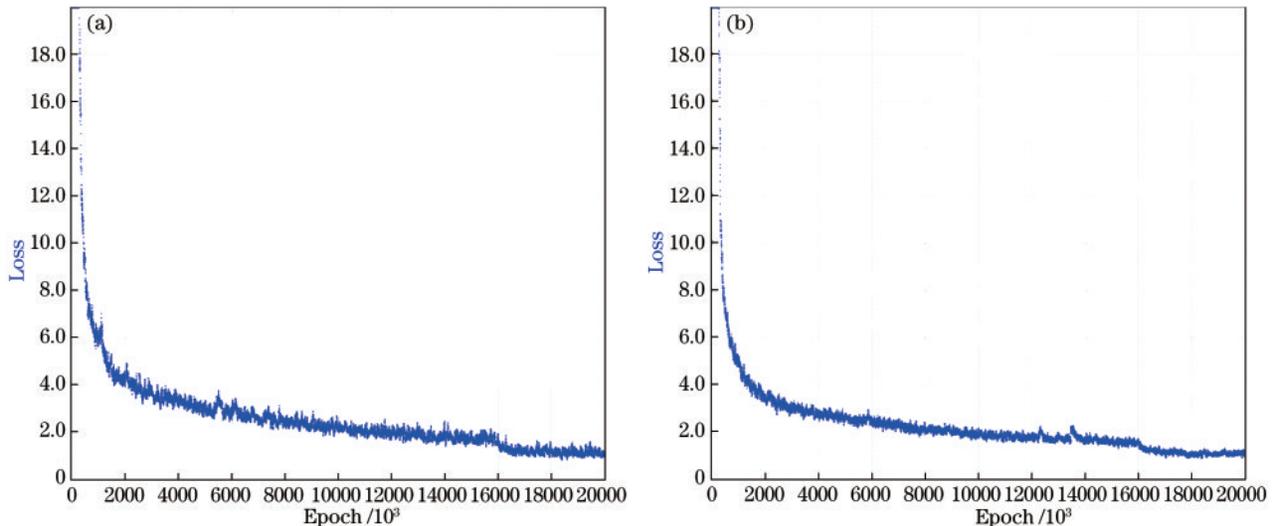


图 9 损失曲线。(a) YOLOv4; (b) YOLOv4-disc

Fig. 9 Loss curves. (a) YOLOv4; (b) YOLOv4-disc

表 2 不同网络模型的验证结果

Table 2 Verification results of different network models

Network	Healthy		Protrusion		Swollen		Size / MB	FPS
	Precision / %	Recall / %	Precision / %	Recall / %	Precision / %	Recall / %		
YOLOv4-disc	95.18	93.54	97.34	96.71	82.46	81.82	281	34.7
YOLOv4	92.61	91.16	92.35	91.84	82.42	81.61	354	28.5
YOLOv3	83.42	78.65	82.77	82.13	74.36	70.25	237	30.9
Faster-RCNN	81.35	80.22	80.67	79.45	70.75	68.92	146	27.6

由式(19)计算出的不同网络模型的 F1 值如表 3 所示。从表中可以看出, YOLOv4-disc 在 3 种病变类

型上的 F1 值均优于其他网络,优化后的 YOLOv4 更加适合对脊柱 MRI 影像中的腰椎间盘进行识别。

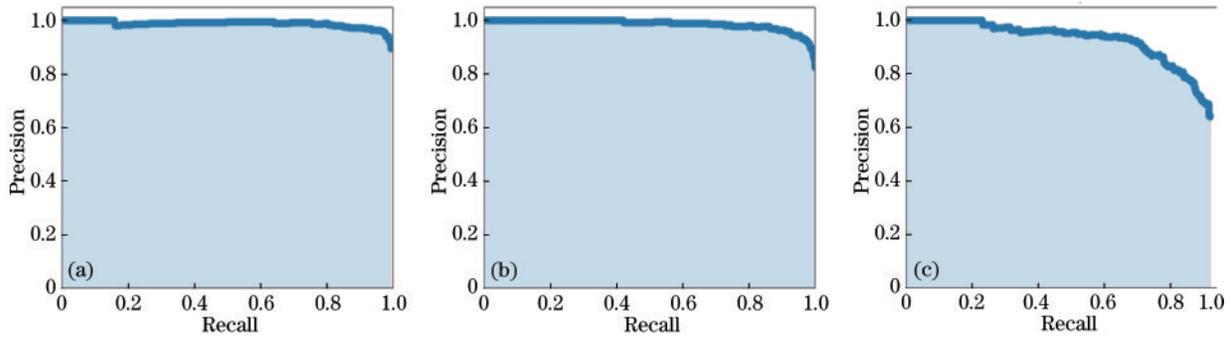


图 10 YOLOv4-disc  $P-R$  曲线图。(a) Healthy; (b) protrusion; (c) swollen  
Fig. 10 YOLOv4-disc  $P-R$  curve. (a) Healthy; (b) protrusion; (c) swollen

表 4 为不同网络对于 3 种病变类型检测的 AP 值, mAP 为平均精度均值。从图 11 可以更直观地看出, YOLOv4-disc 对 3 种类型的 AP 值均不同程度优于其余网络。

表 3 不同网络的 F1 值

Table 3 F1 values for different networks

Network	Healthy-F1	Protrusion-F1	Swollen-F1
YOLOv4-disc	0.94	0.97	0.82
YOLOv4	0.91	0.92	0.81
YOLOv3	0.81	0.82	0.72
Faster-RCNN	0.80	0.80	0.69

表 4 不同网络的 AP 值

Table 4 AP values of different networks unit: %

Network	Healthy	Protrusion	Swollen	mAP
YOLOv4-disc	94.48	96.21	81.73	90.80
YOLOv4	91.17	92.28	81.42	87.29
YOLOv3	82.01	83.34	72.77	79.37
Faster-RCNN	79.58	79.72	68.48	75.92

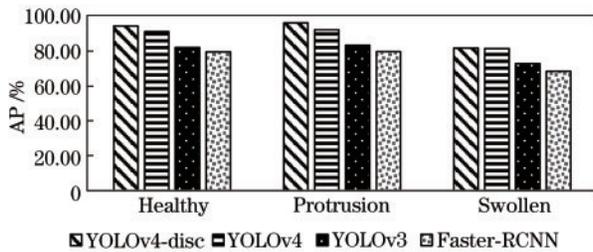


图 11 不同网络的 AP 值

Fig. 11 AP values of different networks

图 12(a) 和图 12(b) 是分别使用 YOLOv4 和 YOLOv4-disc 网络在椎间盘末端和周围软组织影像对比度不明显的情况下的检测结果, YOLOv4 出现了漏检的情况, 而 YOLOv4-disc 检测出了最后一节椎间盘的突出, 并且置信度高达 0.90。图 13(a) 和图 13(b) 是在椎间盘末端轻微膨出的情况下的检测结果, YOLOv4 出现了一盘多检的情况, 而 YOLOv4-disc 正确地诊断出了椎间盘的膨出。图 14(a) 和图 14(b) 是在

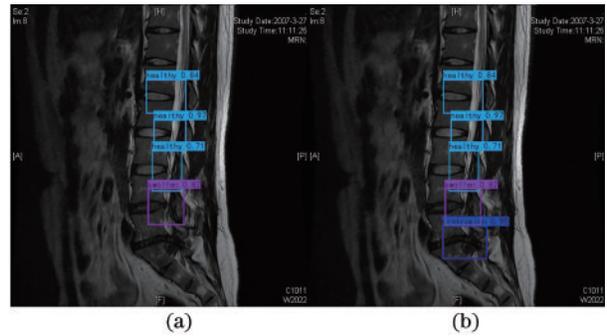


图 12 漏检情况的对比。(a) YOLOv4; (b) YOLOv4-disc  
Fig. 12 Comparison of missed inspection. (a) YOLOv4; (b) YOLOv4-disc

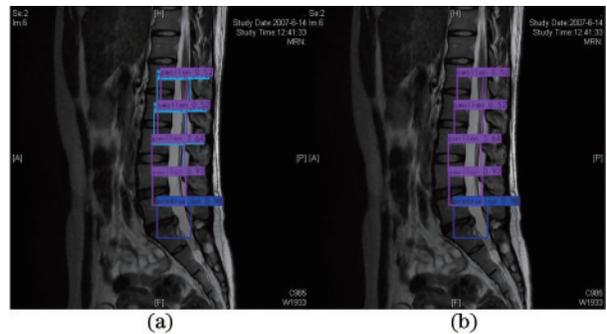


图 13 重复检测情况的对比。(a) YOLOv4; (b) YOLOv4-disc  
Fig. 13 Comparison of repeated detection. (a) YOLOv4; (b) YOLOv4-disc

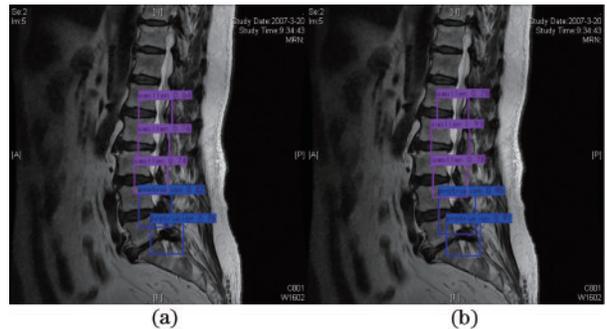


图 14 非最佳拍摄位置的对比。(a) YOLOv4; (b) YOLOv4-disc  
Fig. 14 Comparison of nonoptimal shooting positions. (a) YOLOv4; (b) YOLOv4-disc

影像拍摄位置未在最佳横切面时的检测结果, YOLOv4 对 5 个椎间盘的健康状况诊断的置信度较低, 而 YOLOv4-disc 诊断正确且整体置信度较 YOLOv4 有较大的提升。

## 5 结 论

为了提高对腰椎间盘突出患者脊柱 MRI 影像检测的实时性与准确性, 并对各椎间盘以及各椎间盘末端健康状况等体积较小目标进行识别诊断, 提出一种优化 YOLOv4 的 YOLOv4-disc 的 MRI 影像检测算法。该算法通过 CLAHE 与 Mosaic 数据增强提升网络模型的鲁棒性, 改进特征提取网络的结构、根据数据集预设锚框来提高网络的精度、实时性。实验结果表明, YOLOv4-disc 模型在健康、突出两种病变类型上相比 YOLOv4 有较大的提升, 并将平均精确率提高了 3.78 个百分点, 且 mAP 提升了 3.51 个百分点、FPS 提升了 6.2。但在膨出病变类型的检测上, 精确率和召回率仍然有较大的提升空间, 未来将尝试在网络中添加注意力机制使网络模型更加关注椎间盘的边缘信息, 进一步提高网络的性能。

## 参 考 文 献

- [1] 张晓静, 李伟峰. 腰椎间盘突出症的 CT 影像与 MRI 影像的比较研究[J]. 数理医药学杂志, 2020, 33(10): 1471-1472.  
Zhang X J, Li W F. Comparative study of CT and MRI images of lumbar disc herniation[J]. Journal of Mathematical Medicine, 2020, 33(10): 1471-1472.
- [2] 曹锐. 独活寄生汤治疗腰间盘突出所致的腰腿痛的临床疗效[J]. 内蒙古中医药, 2020, 39(8): 52-53.  
Cao R. Clinical effect of Huojisheng Decoction on lumbago and leg pain caused by lumbar disc herniation[J]. Nei Mongol Journal of Traditional Chinese Medicine, 2020, 39(8): 52-53.
- [3] 罗浚铠, 张宝华, 张艳月, 等. 基于上下文自校准双重注意力的目标检测算法[J]. 激光与光电子学进展, 2021, 58(12): 1210013.  
Luo J K, Zhang B H, Zhang Y Y, et al. An object detection algorithm based on contextual self-calibration and dual-attention mechanism[J]. Laser & Optoelectronics Progress, 2021, 58(12): 1210013.
- [4] 程叶群, 王艳, 范裕莹, 等. 基于卷积神经网络的轻量化目标检测网络[J]. 激光与光电子学进展, 2021, 58(16): 1610023.  
Cheng Y Q, Wang Y, Fan Y Y, et al. Lightweight object detection network based on convolutional neural network[J]. Laser & Optoelectronics Progress, 2021, 58(16): 1610023.
- [5] 侯高峰, 房丰洲. 基于深度学习的糖尿病眼底病变检测的研究[J]. 激光与光电子学进展, 2023, 60(2): 0217001.  
Hou G F, Fang F Z. Research on diabetic retinopathy detection based on deep learning[J]. Laser & Optoelectronics Progress, 2023, 60(2): 0217001.
- [6] 马树志. 基于深度学习的肝脏 CT 影像分割方法的研究与应用[D]. 长春: 吉林大学, 2017.  
Ma S Z. Research of liver segmentation in CT image based on deep learning[D]. Changchun: Jilin University, 2017.
- [7] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2021-02-05]. <https://arxiv.org/abs/2004.10934>.
- [8] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2021-02-04]. <https://arxiv.org/abs/1804.02767>.
- [9] 李妮妮, 王夏黎, 付阳阳, 等. 一种优化 YOLO 模型的交通警察目标检测方法[J]. 图学学报, 2022, 43(2): 296-305.  
Li N N, Wang X L, Fu Y Y, et al. A traffic police object detection method based on optimized YOLO model[J]. Journal of Graphics, 2022, 43(2): 296-305.
- [10] 刘正波, 鲍义东, 孟庆伟. 基于改进 YOLOv4 的棉花检测算法[J]. 计算机系统应用, 2021, 30(8): 164-170.  
Liu Z B, Bao Y D, Meng Q W. Cotton detection algorithm based on improved YOLOv4[J]. Computer Systems & Applications, 2021, 30(8): 164-170.
- [11] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2021-06-04]. <https://arxiv.org/abs/1704.04861>.
- [12] 严开忠, 马国梁, 许立松, 等. 基于改进 YOLOv3 的机载平台目标检测算法[J]. 光电与控制, 2021, 28(5): 70-74.  
Yan K Z, Ma G L, Xu L S, et al. Improved YOLOv3 based target detection algorithm for airborne platform[J]. Electronics Optics & Control, 2021, 28(5): 70-74.
- [13] 周宇杰, 徐善永, 黄友锐, 等. 基于改进 YOLOv4 的输送带损伤检测方法[J]. 工矿自动化, 2021, 47(11): 61-65.  
Zhou Y J, Xu S Y, Huang Y R, et al. Conveyor belt damage detection method based on improved YOLOv4[J]. Industry and Mine Automation, 2021, 47(11): 61-65.
- [14] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [15] 程松, 戴金涛, 杨洪刚, 等. 基于改进型 YOLOv4 的焊缝图像检测与识别[J]. 激光与光电子学进展, 2022, 59(16): 1610002.  
Cheng S, Dai J T, Yang H G, et al. Weld image detection and recognition based on improved YOLOv4[J]. Laser & Optoelectronics Progress, 2022, 59(16): 1610002.