文章编号: 2096-1472(2023)-03-30-06

DOI:10.19644/j.cnki.issn2096-1472.2023.003.007

# 基于改进SSD算法的胃部息肉图像检测

范明凯,肖满生,胡一凡,吴宇杰

(湖南工业大学计算机学院,湖南 株洲 412007) ⊠gangxinjiaoyongg@163.com; 349407041@qq.com; 373502340@qq.com; 2275128586@qq.com



摘 要:针对胃部内窥镜图像中的小尺度息肉检测精度不高的问题,提出了一种改进的SSD(Single Shot MultiBox Detector)算法。首先在SSD网络的池化层设计最大池化Dropout,使神经元稀疏化,然后在反向传播求损失函数梯度时,引入基于权重衰减的正则化技术以降低模型的复杂度,最后通过反卷积进行特征融合,使图像特征得以充分利用,解决内窥镜图像中小尺度物体检测不足的问题。实验表明,所提出的方法对胃部息肉具有良好的检测效果,相对于传统SSD算法,提高了5%的胃部息肉检测的平均精度(*mAP*)。

关键词:SSD;权重衰减;卷积神经网络;过拟合;Dropout 中图分类号:TP391 文献标识码:A



Image Detection of Gastric Polyp based on Improved SSD Algorithm

FAN Mingkai, XIAO Mansheng, HU Yifan, WU Yujie

(School of Computer Science, Hunan University of Technology, Zhuzhou 412007, China) Sangxinjiaoyongg@163.com; 349407041@qq.com; 373502340@qq.com; 2275128586@qq.com

Abstract: This paper proposes an improved SSD (Single Shot MultiBox Detector) algorithm to address the problem of poor detection accuracy of small polyps in gastric endoscope images. First of all, the max pooling Dropout is designed in the pooling layer of SSD network to sparse out the neurons. Then, a regularization technique based on weight decay is introduced to reduce the complexity of the model when calculating the gradient of the loss function by back-propagating. Finally, feature fusion via deconvolution helps make full use of image features and solve the problem of insufficient detection of small and medium-sized objects in endoscope images. The experimental results show that the proposed solution has a good detection effect on gastric polyps. Compared with the traditional SSD algorithm, it improves the mean average precision (mAP) of gastric polyp detection by 5%.

Keywords: SSD; weight attenuation; convolution neural network; over-fitting; Dropout

# 1 引言(Introduction)

胃癌是最常见的消化道肿瘤之一,是一种严重威胁人的 生命健康的重大疾病。胃中的大多数息肉具有恶性潜能,并 且具有与胃癌相同的一些危险因素和机理,因此可能发生癌 变<sup>[1-2]</sup>。胃息肉如果能在早期发现和治疗,对预防胃癌至关重 要。然而由于胃结构复杂,息肉影像数据量大,小目标息肉 不易识别,仅通过医生判别可疑病变息肉,效率低下,而且 精准度不高<sup>[3-4]</sup>。人工智能系统可以自动勾画病灶,提高影像 医生的诊断效率。作为人工智能发展的一个方向,深度学习 (Deep Learning, DL)为自动病变检测的医学图像分析提供了 新的思路<sup>[5]</sup>。DL模型对胃肠病学的主要贡献之一是其快速、 可靠的息肉检测能力,提高了诊断质量<sup>[6]</sup>。

近年来,智能化图像识别技术在医疗诊断领域得到了极 大发展,如DL方法在胃息肉检测中的应用。HIRASAWA

基金项目:湖南省自然科学基金项目(2021JJ50049),湖南省教育厅重点项目(21A0607),湖南省自然科学基金项目(2022JJ50077).

等<sup>[7]</sup>利用13,584 张内窥镜图像数据集,通过训练One-Stage 的目标检测模型<sup>[8]</sup>(SSD)检测胃内镜照片癌变区域,同时用 一套独立的测试图像集进行测试,共包括连续69 例患者的 2,296 张胃部内窥镜图像, 检测结果中含有77 个胃癌病灶, 最终正确诊断为71个,整体敏感性为92.2%,但没有改变SSD 的任何细节。WANG等<sup>[9]</sup>使用改进的Faster RCNN<sup>[10]</sup>算法 (Faster Region-based Convolutional Neural Network)检测 息肉,在Faster RCNN中用ROI(Region Of Interest)对齐操 作代替ROI池操作,用GIoU(Generalized Intersection Over Union)损失代替原有平滑L1损失,用Soft-NMS(Soft Non-Maximum Suppression)代替传统的NMS,在胃息肉图像检 测中取得了良好的效果,但网络较为复杂,ZHANG等<sup>[11]</sup>提 出了一种基于增强的SSD架构(SSD-gpnet)用于息肉检测的卷 积神经网络方法(Convolutional Neural Network, CNN), 共收集了215 名胃息肉患者的404 张内窥镜图像。经过增强 处理后,随机选取708幅图像和50幅图像分别进行训练和测 试步骤,该研究的平均精度(mAP)为90.4%; LADDHA等<sup>[12]</sup> 提出了检测胃息肉的YOLOv3模型,数据集的来源为文献[10] 采用的数据集,测试结果mAP为82%;XIA等<sup>[13]</sup>开发出基于 Faster-RCNN的胃病变自动检测系统,用于检测糜烂、息 肉、溃疡、黏膜下肿瘤、黄瘤、正常黏膜和无效图像。系统 接收器工作特性(ROC)分析给出了84%的结果,使用787名患 者的1,023,955 张内窥镜图像数据集,得到准确率、召回率和 调和平均指标(F1-Score)分别为78.96%、76.07%和77.49% 但准确率不够高。CAO等<sup>[14]</sup>使用1,941 张图像组成的私人数1 集,使用YOLOv3进行胃息肉检测,该方法的计算准确率为 91.6%, 召回率为86.2%, F1-Score为88.8%。

以上方法各有优势,但都存在不足,特别是精确度可以 继续提升。检测中的主干网络CNN模型在训练过程中,数据 集过小、含有大量噪声等,或者模型包干复杂、训练参数过 多、训练过度等,都极有可能使模型陷入过拟合。基于此, 本文在充分分析CNN特点及目前国内外学者对CNN模型过拟 合问题的研究基础上,提出了一种面向最大值池化Dropout与 权重衰减过拟合问题算法,并将该算法与SSD<sup>[8]</sup>网络相结合, 并通过理论推导与实验对比,证明了该方法能有效避免检测 中主干CNN模型训练过程的过拟合问题,提高了模型的泛化 能力,进一步提高了目标的检测能力。

# 6统的SSD检测算法及其局限性(Traditional SSD detection algorithm and its limitations)

SSD算法是一种目标检测算法,由LIU等<sup>[8]</sup>在欧洲计算机 视觉国际会议(ECCV2016)上提出。SSD借鉴Faster-RCNN<sup>[10]</sup> 中的锚框机制,在特征图上生成具有不同长宽比例的默认框 进行预测,并使用金字塔结构,在低层检测小尺寸目标,在 高层检测大尺寸目标。

#### 2.1 SSD基本原理

SSD网络结构如图1所示,该结构由两个部分组成,一 部分是VGG16<sup>[15]</sup>基础网络,另一部分是附加特征层。其中, 基础网络用来提取低尺度的特征映射图,附加特征层用来 提取高尺度的特征映射图。SSD将VGG16网络中的全连接 层FC6和FC7替换成卷积层,去掉DropOut层和FC8层,将 池化层pool5从2×2的大小及步长为2更改为3×3的大小和步 长为1。VGG16中的Conv4\_3作为第一个特征图,经过多次 卷积得到大小不同的特征图Conv7、Conv8\_2、Conv9\_2、 Conv10\_2、Conv11\_2,通过对多个特征图同时进行分类 预测和位置回归,最后经过非极大值抑制(Non-Maximum Suppression, NMS)策略过滤掉重叠度较大的预测框,完成对 不同大小目标的检测。



Fig.1 The SSD model diagram

SSD借鉴了Faster-RCNN中的预定义边框(Anchor Box)的理念,为每个单元设置了宽高比不同的先验框。每个单元 有k个先验框,每个先验框会预测c个分类的分数和4个位置 偏移量,大小为 $m \times n$ 的特征图将会产生(4+c)× $k \times m \times n$ 个预 测参数。

先验框架的比例随着特征图大小的减小,其线性增加, 设 S<sub>k</sub>表示第 k 个特征图先验框和原图大小的比例, *m* 是特征 图的个数,则默认框大小计算公式(1)如下:

$$S_k = S_{\min} + \frac{S_{\max} - S_{\min}}{m - 1} (k - 1), \quad k = 1, 2, \cdots, m$$
(1)

根据文献[8], 公式(1)中的 $S_{\min}$ 和 $S_{\max}$ 的取值分别为0.2和0.9, 此处不再赘述。

设长宽比为 $a_r$ , 一般选取 $a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}\}$ , 此外每个特征图会有一个 $a_r = 1$ 且尺度分别为 $s_k \ \pi s'_k = \sqrt{s_k s_{k+1}}$ 的默认框。每个默认框的中心为 $\left(\frac{i+0.5}{|f_k|}, \frac{j+0.5}{|f_k|}\right)$ ,  $i, j \in [0, |f_k|)$ , 其中  $|f_k|$ 是第k 个特征图的大小。

在模型训练过程中,需要确定真实目标匹配的默认框。第一,每个默认框找到与其交并比(Intersection over Union, IoU)最大的先验框,作为正样本,如果一个先验框没有与任何默认框匹配,则作为负样本,第二,在未匹配的先验框中,如果存在与某个默认框的IoU大于一个阈值,那么也作为正样本。设真实目标的默认框大小为S<sub>GT</sub>、先验框的目标大小为S<sub>PB</sub>,则IoU计算如公式(2)所示:

$$IoU = \frac{S_{GT} \cap S_{PB}}{S_{GT} \cup S_{PB}}$$
(2)

#### 2.2 SSD的局限性

SSD不同层提取的特征图感受野如图2所示,图2(a)为低 层特征图的感受野,图2(b)为中层特征图的感受野,图2(c) 为高层特征图的感受野,模型的低层特征由于经过的卷积运 算较少,所以特征图包含更多的纹理和细节信息,但语义信 息不足,难以区分目标和背景,高层特征图包含较多的语义 信息,但是在层层卷积下采样过程中丢失了大量细节信息。 SSD采用金字塔结构的特征层进行目标检测,利用大感受野的 高级特征层预测大目标,利用小感受野的低级特征层预测小 目标。然而,不同类别的特征层是独立的,低层特征图没有 利用高层特征图的语义信息,导致真实目标与先验框难以匹 配,无法满足检测小目标更加精确的要求。同时,每个特征 层是独立的,信息不能共享。综上,SSD算法对于小目标的检 测并不理想。





# 3 改进的SSD(The improved SSD)

上文介绍了传统的SSD模型,本部分在上述基础上设计 一种最大值池Dropout与权重衰减CNN模型,与传统的SSD模 型相结合,以求在模型训练过程中最大限度地避免过拟合情 况的发生,并且将高层语义信息与底层细节信息进行融合, 通过反卷积将SSD网络结构中富含语义信息的高层特征映射到 低层网络,以求模型能更好地检测小目标。

## 3.1 最大值池化Dropout

第二部分已说明,最大值池化是CNN池化层常用的方法,为了在模型训练过程中避免过拟合,本文在CNN池化层引入最大值池化Dropout。

设待池化的特征图(Feature Map)中每个池化区域的 保留概率为 P, P的大小可根据实际情况手动调节, — 般设 P=0.5, 则池化区域中每个单元均有的抑制概率为 q=1-p。同时, 假设第 l 层的每个池化区域m中的单元值  $(s_{m,1}^{l}, s_{m,2}^{l}, ..., s_{m,n}^{l})$ 按从小到大顺序重新排列, 排列后的单元值为  $0 < d_{m,1}^{l} < d_{m,2}^{l} < ... < d_{m,n}^{l}$ (注:此处采用半线性激活函数ReLU使 得所有单元的激活值为非负, 故最小的 $d_{m,1}^{l} > 0$ ), 那么 $d_{m,j}^{l}$ 被选 择作为整个池化区域最大值池化后的输出情况为所有大于 $d_{m,j}^{l}$ 的单元值 $(d_{m,j+1}^{l}, d_{m,j+2}^{l}, ..., d_{m,n}^{l})$ 被抑制, 只有小于等于 $d_{m,j}^{l}$ 的值 被保留下来, 在这些保留值中,由于采用最大值池化(取最大 值), 故池化最后的输出值为 $d_{m,i}^{l}$ , 其发生的概率为 $P_{i}$ , 即

 $p_{j} = poss(s_{m}^{l+1} = d_{m,j}^{1}) = pq^{n-j}, j = 1, 2, ..., n$  (3) 分析公式(3)可知, 在池化区域执行最大值池化Dropout

时,通过多项式排列选择池化区域中经过递增排列的第*j*个激活值*d<sup>l</sup><sub>m,j</sub>*,作为该池化区域的输出值,即

$$s_m^{l+1} = d_{m,j}^l, \ p_j \in (p_1, p_2, p_3, \cdots, p_n)$$
(4)

假设第1 层有 r 个Feature Map,每一个的大小为s,池 化区域大小为t×t,又设池化步长为t,即不考虑重叠池化, 则有 rs/t 个池化区域,则第1 层要训练的模型参数为(t+1)<sup>rs/t</sup> (其中加一个偏置),即最大值池化要训练的模型数量与输入 池化层的池化区域单元数量呈指数关系。引入最大值池化 Dropout后,池化单元被随机抑制,即*t*减少,要训练的模型 参数量则呈指数减少,这样有效降低了模型复杂度,因而能 更有效地抑制过拟合。

## 3.2 权重衰减计算

在训练大型CNN时,除了使用上述最大值池化Dropout 抑制单元避免过拟合,如果模型在某些区域里函数值变化剧 烈,则意味着函数的参数值(权重)偏大,使得该区域里的导数 值绝对值大、模型也变得复杂,权重衰减通过约束参数的范 数使其不能过大,以此降低模型的复杂度,减小噪声输入的 影响,从而在一定程度上减少过拟合发生,该方法也叫正则 化方法。

设模型的损失函数 L<sub>0</sub> 如公式(5)所示,进行权重衰减计算时,在原损失函数 L<sub>0</sub> 中加一个惩罚项,即

$$L_0 = -\frac{1}{N} \sum_{i=1}^{N} (y_i \ln(o_i))$$
(5)

$$L = L_{w} + \frac{\lambda}{2} \sum_{w} \left\| w \right\|^{2} \tag{6}$$

公式(6)中,  $L_0$ 是原损失函数, w是网络权重,即神经元 的连接系数,  $\lambda(\lambda > 0)$ 是惩罚项系数,用来衡量惩罚项与 $L_0$ 的 比例关系, 1/2是为了求导方便而设计的。上述惩罚项为网络 权重w的平方和,对公式(6)进行求导:

$$\begin{cases} \frac{\partial L}{\partial w} = \frac{\partial L_0}{\partial w} + \lambda w \\ \frac{\partial L}{\partial b} = \frac{\partial L_0}{\partial b} \end{cases}$$
(7)

公式(7)中, b为网络神经单元(如卷积核)的偏置,其包含 在 $L_0$ 的 $o_n$ 中,即 $o_n = w \cdot x_{n-1} + b$ ,可以发现加上惩罚项后,对偏置 b的更新没有影响,设n为学习率,对于权重值w有公式如下:

$$w' = w - \eta \left( \frac{\partial L_0}{\partial w} + \lambda w \right) = w - \eta \frac{\partial L_0}{\partial w} - \eta \lambda w = (1 - \eta \lambda) w - \eta \frac{\partial L_0}{\partial w}$$
(8)

分析公式(8)得出,在没加惩罚项前,权重值w前的系数为 1, 即w'= w- $\eta \frac{\partial L_0}{\partial w}$ , 加上惩罚项后, w前的系数变为1- $\eta\lambda$ , 由于 $\eta$ 、 $\lambda$ 都为小于1的正数,所以1- $\eta\lambda$  <1,即公式(8)的作用 是为了减小w值,这就是权重衰减的理论意义。注意公式(8)中  $\eta \frac{\partial L_0}{\partial w}$ 项是反向传播的权重变化梯度,无论加不加惩罚项,其表 达式都一样。因此,这里所说的权重衰减项不包括此项。进一 步对公式(8)进行分析,当w为正时,更新后的w'变小,当w为 负值时,更新后的w'变大,由于|w|<1(网络归一化后的权重), 因此公式(8)的效果就是让w向0靠近,即 $|w| \rightarrow 0$ ,使w的值尽可 能变小,也相当于减小了网络的权重,降低了网络复杂度, 从而避免过拟合。需要指出的是,公式(8)中参数λ值的大小 设置很重要, λ太大, 权重 w 减小过快, 可能会出现欠拟合, 甚至无法训练, 而 λ 太小, 又会出现过拟合。 λ 大小设置可采 用基于Bayes决策规则调整法,该方法假定网络的权重与偏置 是具有特定分布的随机变量,用统计方法进行自动计算,其 详细内容可参考文献[16],限于篇幅,此处不做详述。其中, 权重衰减的计算流程如下。

```
Step1: 输入学习率\eta \leftarrow 0, 惩罚项系数\lambda \in R。
Step2: 初始化时间步长t \leftarrow 0, 网络权重w \in R。
Step3: repeat。
Step4: t \leftarrow t+1。
Step5: \nabla L(w_i) \leftarrow SelectBatch(w_{i-1}) / / 返回相对应的梯度。
Step6: g_t \leftarrow \nabla L(w_t) + \lambda w_{t-1}。
Step7: w_t \leftarrow w_{t-1} - \eta g_t。
Step8: until满足停止判据。
Step9: return优化参数w_{to}
```

#### 3.3 反卷积

低层特征层包含的细节信息较多,丰富其特征细节信息 可以更准确地检测小目标。本文采用反卷积<sup>[17-18]</sup>操作,通过参 数调整将高层低分辨率特征图映射为高分辨率特征图,然后 将高分辨率特征图和低层对应的高分辨率特征图进行拼接, 使用于检测小目标的低层特征图的特征更加丰富,提高特征 表达能力。反卷积操作过程与卷积操作相反,具体操作过程 如图3所示。

图3(a)表示卷积过程,其将3×3卷积核作用于4×4特征图 中,卷积得到2×2的特征图。图3(b)表示反卷积过程,其操作 与卷积相反,输入2×2的特征图,通过填充补零,并将3×3 卷积核作用于特征图,得到4×4特征图。



为了表示方便,本文提出的最大值池化Dropout与权重

家城方法简称为MDWS(Maxpooling Dropout and Weight Scaled),并将此方法与传统SSD方法相结合,将传统SSD 的池化操作替换为MDWS,如图4所示,网络的基本结构为VGG16,遵循特征金字塔进行目标检测,输入的图片经过改进的SSD网络(SSD+MDWS网络)主要提取出Conv4 第三次卷积的特征,Fc7卷积的特征,Conv6、Conv7、Conv8、Conv9第二次卷积的特征,其中Conv4\_3和Fc7提取到38×38、19×19的特征图是用来检测小目标的,为了让小目标检测效果更佳,将Fc7提取到的特征进行融合,以增加特征,融合的过程如图5所示,并且在Conv4和Fc7的池化处理中设置 MDWS模块,对于这两种特征图进行分类和回归预测,在训练的时候减少网络的复杂度,抑制过拟合,提高特征图内物 体种类识别的精确度,有利于先验框参数的调整。



图4 SSD+MDWS网络模型图

Fig.4 Diagram of SSD+MDWS network model



图5 特征融合模块

Fig.5 Feature fusion module

# 4 实验结果分析(Analysis of experimental results) 4.1 实验环境

操作系统为Ubuntu 18.04,处理器为Intel Core i7 CPU@3.00,RAM 16 GB,GPU为NVIDIA GeForce GTX 1080Ti(11 GB),深度学习框架为Tensorflow,利用CUDA 0.1和cuDNN 9.1加速训练,使用Python作为主要编程语言。

#### 4.2 数据集和训练设置

实验采用两个数据集,分别是Hyper-Kvasir<sup>[19]</sup>和CVC-ClincDB<sup>[20]</sup>。Hyper-Kvasir是一个大型的多类公共胃肠道数 据集,数据来自挪威Baerum医院,所有的标签都是由经验丰 富的医生制作的,该数据集包含异常(不健康)和正常(健康)患 者的110,079 张图像和374 个视频,共产生约100万张图像和 视频帧。本文采用Hyper-Kvasir-Segmented-Images数据 集,该数据集提供了来自息肉类的1,000 张原始图像,并提供 了分割掩码和边界框,这个数据集是由制作Hyper-Kvasir数 据集的作者提供的。CVC-ClincDB包含23 位病人的31 个序 列的612 个标准清晰图片,每张图片的分辨率为384×288。部 分样本如图6所示。



图6 部分数据集内容 Fig.6 Contents of partial dataset

训练时将数据集按照8:2划分为训练集和测试集,输入图 像分辨率设置为300×300,初始学习率为1e-4,如果损失连 续在10个轮次之后不降低,则自动调节下降学习率,本实验 采用与传统SSD相同的图像增强方法。

## 4.3 评价指标

对于检测任务,使用精确率(Precision)、召回率(Recall)、 F1分数(F1-Score)及平均精确率均值(mAP)作为评估指标。设TP(True Positive)表示被正确判定为正类的样本; FP(False Positive)表示被错误判定为正类的负样本;FN(False Negative)表示被错误判定为负类的正样本;TN(True Negative)表示被正确判定为负类的负样本。

精确率(*Precision*)表示预测的正样本的样本数占所有预 测为正样本的样本数的比例。精确度越高,分类器的性能越 好,其定义见公式(9):

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

召回率(*Recall*)表示正确预测为正样本的样本数占实际正 样本数的比率,其定义见公式(10):

$$Recall = \frac{TP}{TP + FN}$$
(10)

(11)

F1-Score表示调和平均指标,其定义见公式(11):

$$F1-Score = 2 \times \frac{Precision \times R \circ call}{Precision + R \circ call}$$

得到模型的精确率和召回率,绘制PR曲线,PR曲线下的 面积代表平均精确率(AP)。计算方法是先确定目标的最分置信 度阈值,然后计算出符合条件的各个置信度下,测试集各图 像中每类目标的精确率和召回率,接着绘制PR曲线图,最后 计算PR曲线图的曲线下的面积,得到该类目标在该图像上的 AP值。将所有的图像样本求平均,即可得到该类目标的最终 AP标准值,然后对所有类别的AP求平均,得到mAP,见定义 公式(12)(其中N是类别数目):

$$mAP = \frac{\sum_{k=1}^{N} AP(k)}{N}$$
(12)

### 4.4 实验结果

#### 4.4.1 Hyper-Kvasir数据集实验

首先,将改进后的算法SSD+MDWS与其他四种流行的 模型在Hyper-Kvasir的100 张图片的测试集中做对比,模 型为传统SSD<sup>[7]</sup>、YoLoV3<sup>[12-13]</sup>、Faster R-CNN<sup>[9]</sup>和SSDgpnet<sup>[11]</sup>。相应的实验结果如表1所示。分析表1中的数据发 现,SSD+MDWS的整体性能是最好的,并且取得了最高的 *mAP*,为92.23%,虽然在召回率的表现上,Faster R-CNN 比SSD+MDWS多了约2%,但是准确率和F1-*Score*,都是本 文提出的SSD+MDWS的效果更好,也在一定程度反映出网络 训练时能够更好地减少过拟合。

表1 Hyper-Kvasir实验不同算法测试结界	ミ对비
---------------------------	-----

Tab.1 Test results comparison of different algorithms in Hyper-Kyasir experiment

方法	准确率/%	召回率/%	F1-Score/%	mAP/%	
YOLOv3	87.44	81.90	85.00	87.44	
Conventional SSD	82.06	85.81	84.00	88.47	
Faster R-CNN	66.32	89.25	76.00	88.19	
SSD-gpnet	84.30	87.85	86.00	90.12	
SSD+MDWS	88.68	87.85	88.00	92.23	

#### 4.4.2 CVC-ClincDB数据集实验

同"4.2"部分介绍的实验,采取上述四种模型(SSD、 YoLoV3、Faster R-CNN和SSD-gpnet),用划分好的测试 集在各自训练好的网络中测试,得到的结果如表2所示。分 析表2中的数据发现,SSD-gpnet的mAP、F1-Score、召回 率分别为96.26%、94.00%、95.35%,比其他四个模型表现 更好,但YOLOV3、Faster R-CNN、SSD-gpnet的准确率 比SSD+MDWS表现更好,最高的多出4.66%。综上所述, SSD+MDWS算法的综合表现最好。图7展示了本算法的检测 效果。

表2 CVC-ClincDB实验不同算法测试结果对比 10.2 Test results comparison of different algorithms in CVC-ClincDB

方法	准确率/%	召回率/%	F1-Score/%	mAP/%	
YOLOv3	95.56	66.67	83.60	93.34	
Conventional SSD	91.51	75.19	83.00	91.89	
Faster R-CNN	97.14	79.07	87.00	94.71	
SSD-gpnet	95.61	84.50	90.00	95.13	
SSD+MDWS	92.48	95.35	94.00	96.27	



图7 检测效果图 Fig.7 Detection rendering

#### 5 结论(Conclusion)

本文采用改进的SSD算法检测息肉,在模型设计中,通 过反卷积将低层和高层的语义信息进行融合,提高对小息肉 的检测能力,并且通过在池化层设计一个最大池化Dropout方 法及在反向传播中引入权重衰减机制,减少模型的复杂度, 进而避免了模型训练中的过拟合现象。本文的创新点主要有 两个:一是最大值池化Dropout的设计,采用池化区域单元值 排序设计单元(神经元)丢弃方式,二是在反向传播中引入惩罚 项设计权重衰减的执行过程。通过理论分析与实验对比验证 了本文提出的方法能有效避免过拟合现象,提高网络泛化性 能。通过实验评估,本文设计的网络显示出更大的息肉检测 潜力。在每个评价指标中,都取得了比以前其他方法更好的 结果。并且,良好的召回率和*F1-Score*表明本文提出的方法 可以降低胃镜检查中误检和漏检的风险,可以为医生和患者 提供更多的帮助。下一步,将该方法与更好的分类网络相结 合,以期准确分析息肉的类别,通过改进CNN架构和功能模 块实现更好的性能。

# 参考文献(References)

- WALDUM H, FOSSMARK R. Gastritis, Gastric Polyps and Gastric Cancer[J]. International Journal of Molecular Sciences, 2021, 22(12):6548.
- [2] LADABAUM U, FIORITTOA, MIANI A, et al. Realtime optical biopsy of colon polyps with narrow band imaging in community practice does not yet meet key thresholds for clinical decisions[J]. Gastroenterology, 2013, 144(1):81–91.
- [3] YANASE J, TRIANTAPHYLLOU E. A systematic survey of computer-aided diagnosis in medicine: Past and present developments[J]. Expert Systems with Applications, 2019, 138:112821.
- [4] HOOGEBBOOM S A, BAGCI U, WALLACE M B. Artificial intelligence in gastroenterology. The current state of play and the potential. How will it affect our practice and when?[1]. Techniques and Innovations in Gastrointestinal Endoscopy, 2020, 22(2):42–47.
- [5] LUNDEROLD A S, LUNDERVOLD A. An overview of deep learning in medical imaging focusing on MRI[J]. Zeitschrift f ü r Medizinische Physik, 2019, 23(2):102–127.
- [6] ABADIR A P, ALI M F, KARNES W, et al. Artificial intelligence in gastrointestinal endoscopy[J]. Clinical endoscopy, 2020, 53(2):132.
- [7] HIRASAWA T, AOYAMA K, TANIMOTO T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images[J]. Gastric Cancer, 2018, 21(4):653–660.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. Computer Vision, 2016, 9905(12):21–37.
- [9] WANG R, ZHANG W, NIE W, et al. Gastric polyps detection by improved Faster R-CNN[C]// Association for Computing Machinery. Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition. New York, USA: Association for Computing Machinery, 2019: 128-133.

- [10] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Patter Analysis and Machine Intelligence, 2017,39(6):1137-1149.
- [11] ZHANG X, CHEN F, YU T, et al. Real-time gastric polyp detection using convolutional neural networks[J]. PLOS One, 2019, 14(3):e0214133.
- [12] LADDHA M, JINDAL S, WOJCIECHOWSKI J. Gastric polyp detection using deep convolutional neural network[C]// Association for Computing Machinery. Proceedings of the 2019 4th International Conference on Biomedical Imaging, Signal Processing. New York, USA: Association for Computing Machinery, 2019:55–59.
- [13] XIA J, XIA T, PAN J, et al. Use of artificial intelligence for detection of gastric lesions by magnetically controlled capsule endoscopy[J]. Gastrointestinal endoscopy, 2021, 93(1): 133–139.
- [14] CAO C, WANC R, YU Y, et al. Gastric polyp detection in gastroscopic images using deep neural network[J]. PLOS One, 2021, 16(4):e0250632.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[DB/OL]. (2015– 04–10)[2023–01–18]. https://arxiv.org/abs/1409.1556.
- 16] MACKAY D J C. BAYESIAN interpolation[J]. Neural computation, 1992, 4(3):415-447.
- [17] 陈幻杰,王琦琦,杨国威,等.多尺度卷积特征融合的SSD目标 检测算法[]].计算机科学与探索,2019,13(6):1049-1061.
- [18] PENG L Y, ZHANG L, ZHANG Y, et al. Deep deconvolution neural network for imagae super-resolution[J]. Journal of Software, 2018, 29(4):927–934.
- [19] BORGLI H, THAMBAWITA V, SMEDSRUD P H, et al. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy[J]. Scientific data, 2020, 7(1):283.
- [20] BERNAI J, TAJKBAKSH N, SANCHEZ F J, et al. Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge[J]. IEEE transactions on medical imaging, 2017, 36(6):1231–1249.

# 作者简介:

范明凯(1996-),男,硕士生.研究领域: 图像处理.

- 肖满生(1968-),男,硕士,教授.研究领域:智能信息处理.本文通信作者.
- 胡一凡(1997-),男,硕士生.研究领域: 图像处理.
- 吴宇杰(1997-), 男, 硕士生.研究领域: 图像处理.