# 介绍

结直肠癌(CRC)是第三位最常见的恶性肿瘤，也是第二位最常见的恶性肿瘤死亡原因。中低收入国家发病率持续上升、大部分发达国家的发病率居高不下，结直肠癌对健康的威胁将进一步加重[19]。结直肠息肉是结直肠癌的早期特征，是结直肠癌发病率和死亡率的驱动因素。相关研究表明，结肠镜息肉筛查是预防结直肠癌的一种有效技术，它可以提供结肠直肠息肉的位置和外观信息，使医生能够在其发展为结直肠癌之前将其切除，将CRC发病率降低30%[1]。但是结肠镜检查的精度依赖于内镜医生的经验，临床研究表明，22% - 28%的息肉在接受结肠镜检查的患者中被遗漏[20]，熟练医师遗漏率8.65%，不熟练医师遗漏率57.14%[23]。漏诊息肉可导致结肠癌的晚期诊断，生存率低至10%[21]。因此，在临床诊断中，准确全面的息肉检测是非常重要的。

然而，由于以下原因，对医生而言这是一项极具挑战的任务。首先，随着结直肠癌患者的不断增多，结肠镜检查人数持续上升,而具有经验的内镜医生医生却严重不足，这使得很多医生疲劳工作，导致较高的息肉漏检。此外，结肠壁有粘膜，粘膜会形成大量褶皱，其中一些褶皱与结肠息肉相似，使息肉难以识别；有些结肠息肉形状很小，甚至处于初期凸起阶段，极其不易察觉。这也导致息肉检测的不准确和漏检。因此，一种能够在早期发现所有可能息肉的计算机辅助技术（CAD）对息肉检测具有重要意义。

在各种息肉检测方法中，早期的检测方法依赖分类器区分息肉与环境，利用颜色小波、局部二值模式(local binary pattern, LBP)、纹理、Haar、定向梯度直方图(histogram of oriented gradients, HoG)[2][3][4][5][6][7]等基于颜色、纹理和形状的特征或特征的组合来区分息肉和正常粘膜。然而，息肉与周围环境的这些特征经常相似，手工特征的表示能力非常有限 ，导致这些模型存在较高的漏检率[15]。

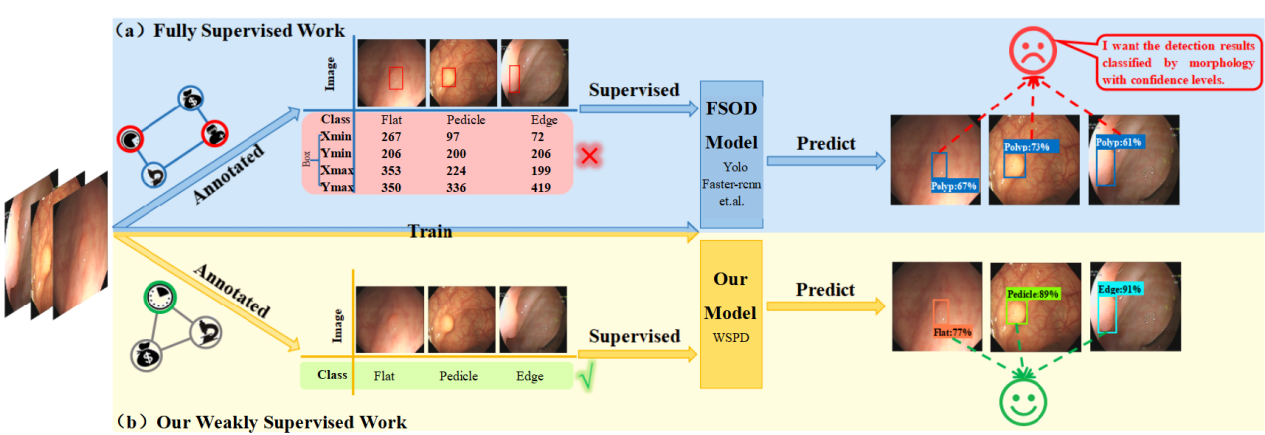
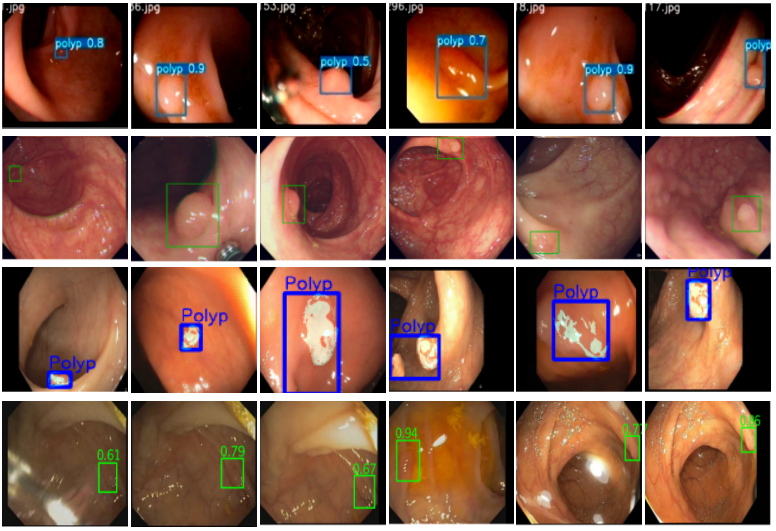


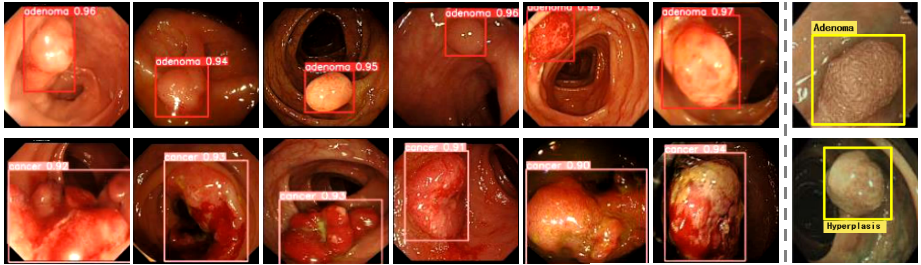
Fig. 1. (a) Traditional fully supervised polyp detection (FSPD) uses the instance-level annotations as supervision. (b) Our weakly supervised polyp detection (WSPD)uses the image-level annotations as supervision.

深度学习彻底改变了手工特征学习的范式，在各种医学任务中取得了显著更好的性能[8]。包括分析医学图像用于疾病诊断和治疗，如肺结节检测[9]、肺癌检测[10][11]、皮肤病分类与检测[18]、乳腺癌诊断[13]。最近，以Faster-Rcnn[16]和YOLO[17]为代表的全监督检测算法取得了不错的成果。这些方法先对数据进行边界框标注，给出目标物体的xmin、xmax、ymin、ymax，如图1(a)所示，通过以CNN为特征提取器、边界框作为标签的全监督网络训练检测器，最终在图像中框出待检测结肠息肉、给出结肠息肉的置信度。

(a)Only detection without classification



(b)Bicategory detection



(c)Challenges in weakly supervised polyp detection

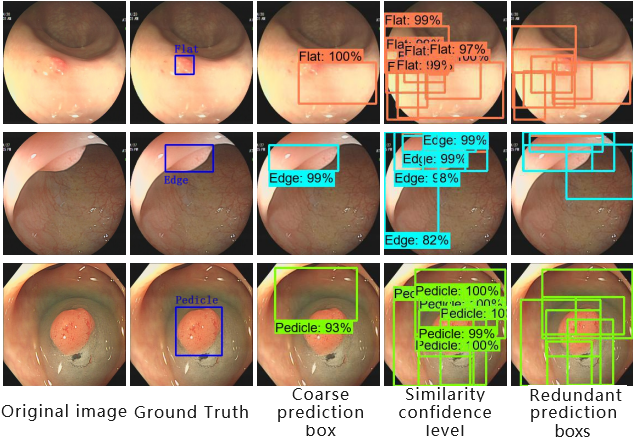


Fig. 2. Detection results from established methods and challenges within our framework.From (a) to (c), they are only detection without classification ,bicategory detection and challenges in weakly supervised polyp detection

全监督方法虽然取得了可喜的成果，但是由于以下原因，许多完全监督的深度学习方法很难扩展到结肠息肉诊断。全监督方法训练精确的模型需要大规模对象级注释的数据集，精确的对象级标注耗费了大量的人力资源。结肠息肉对象级注释对注释者有较高的要求，因此注释者多为经验丰富的临床医生，而结肠息肉图像不会在临床诊断期间被注释[14]，这需要医生花费额外的时间对数据进行注释，增加医生的负担。已有的检测网络根据检测结果可以分为两类：一类仅检查是否出现息肉不区分类别，如图2(a)所示，另一类检测息肉并进行腺瘤/增生二分类，如图2(b)所示。我们与合作医院交流后发现，在实际结肠镜检查过程中，这些网络只有当医生发现息肉并将结肠镜对准其所在位置时，网络才能检测出息肉。因此评价这些网络的检测性能对于实际临床而言是没有意义的。值得注意的是，在临床诊断中医生的漏诊率与息肉形态相关，调查显示扁平息肉漏诊率高达23.52%，而隆起息肉漏诊率8.95%[23]，因此息肉检测网络的训练与评价应聚焦于息肉形态。

考虑到我们在前文中所提到的问题，在本文中，我们设计了图像级标签的弱监督结肠息肉检测方法,如图1（b）所示。我们的方法对图像进行图像级息肉形态注释，将图像送入提出的弱监督息肉检测网络（WSPD）进行训练，在图像中输出检测框和息肉类别置信度。

我们的方法不是第一个解决弱监督目标检测问题(WSOD)的工作,例如Bilen et al.[22]使用预训练的CNN来描述图像区域，结合MIL形成端到端的弱监督检测网络。网络从区域建议中选择对象预测，例如通过选择性搜索窗口(SSW)[24]方法生成，使用预训练的CNN来描述图像区域、学习对象类别，通过汇总分类流和检测流的得分来选择阳性样本。受其启发，我们设计了用于结肠息肉的弱监督检测网络。然而我们发现，网络的检测结构倾向于选择那些覆盖大量背景区域却仅包含息肉局部的候选框，如图2（c）coarse prediction box所示。我们进一步观察到，网络的分类结构在训练开始时并没有学到不同类别最具判别性的部分，而是随着训练次数的增加逐渐过拟合每个候选框，如图2（c）similarity confidence level所示。此外，观察同一张图像的候选框我们还发现，SSW产生的不同候选框存在大量空间冗余，冗余的候选框类别得分极其相近，网络的后处理无法准确筛选，存在大量假阳性误报,如图2（c）redundant prediction boxs所示。

考虑到以上观察结果，我们提出了跨领域对照模块，通过公开的结肠息肉数据集预训练息肉分割网络；去中心化产生私有数据集的伪标签；将伪标签作为数据预处理阶段候选框筛选参照。通过这种方式，网络能够在训练开始之前就过滤候选框，降低不准确候选框对网络检测结构的干扰，同时减少检测时的计算量。

为了进一步增强所提出网络的辨别能力，我们提出在主干网中使用空间类别模块（Spatial Category Module，SCM）。据我们所知，通过全局池化的方法分类会使得网络参数骤减，因此网络会将学习的压力前导至特征提取阶段。此时，特征提取阶段不仅需要学习高级的分类特征，同时也要学习通用的低级特征，例如位置、形状、大小。因此，我们增加了通道引导的全局注意模块，充分考虑对象的全局特征，提高了网络的判别能力和检测精度。

为了降低网络的假阳性误报，我们提出了双阈值后处理策略。我们策略的动机是这样一个假设:紧围物体的box区域置信度要比非紧包物体的box置信度高。我们对最终结果双重筛选，分别设置了类别阈值与区域阈值。先由类别阈值对候选框进行第一次过滤，除去类别得分较低的候选框;再由区域阈值进行二次过滤，除去区域得分较低的候选框。通过这种后处理机制，能够在第二次过滤时筛除空间冗余候选框，降低检测假阳率。

值得注意的是，虽然分类和盒回归的联合学习已经被证明对完全监督的息肉检测是有益的，但对于弱监督的息肉检测，它仍然是不平凡的，需要有更多创新的想法和洞察力。虽然我们的方法在概念上形式简单，但它显著减轻了检测器需要大量数据注释、难以区分不同息肉的问题，与以前的检测方法大为不同。我们的贡献可以概括如下。

* 我们设计了跨领域对比筛选模块，减少了网络定位干扰因素，提高了网络定位准确性。
* 我们设计了一个空间类别模块，增强了学习特征的能力，定位与分类能力取得了显著的进步。
* 我们设计了双阈值后处理策略，细化了筛选条件，类别平均精度显著提升。
* 我们是第一个以结肠形态类别为指导的弱监督检测工作，在公开/私有数据集上取得了显著的成果。

# **相关工作**

## 息肉分割网络

基于编码-解码结构的端到端结构，Ronneberger et al.提出了医学分割网络Unet[27]，由于其良好的精度、高度的任务泛化性以及易于实现性，后续工作都以此标准范例。例如Unet++[28]运用剪枝策略、加快推理速度、解决深度未知问题，重新设计跳跃连接、聚合解码器不同语义尺度的特征、形成高度灵活的特征融合方案；ResUnet++[29]融合残差策略解决梯度消失问题，加入SE块（squeeze and excitation block）增强相关特征敏感性、抑制干扰，ASPP 模块融合不同尺寸上下文信息。

然而，这些方法不能很好地处理息肉的边界。随后，PraNet[26]、PsiNet[30]、MSNet[32]和SFANet[33]、LODNet[31]强制模型学习息肉边界差异，极大地增强了模型对息肉边界的辨别能力，取得了令人满意的效果。此外，ACSNet[34]、HRENet[35]和CCBANet[36]通过关注上下文信息，聚合多尺度上下文，减少局部特征模糊性，获得了高可信度的预测。ICGNet[37]利用反轮廓制导模块聚合底层边缘信息、约束反向区域，利用自适应上下文模块提取当前层的局部全局信息和上一层的互补信息，扩大密集特征以解决尺寸差异、形状不规则、边界模糊的问题。

## 弱监督目标检测

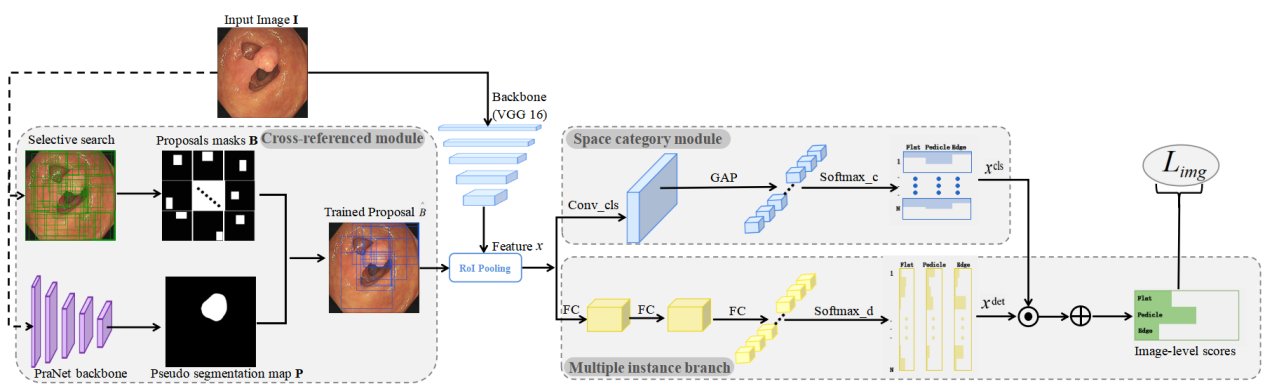
弱监督目标检测所需解决的问题与弱监督学习中的多示例学习（MIL）[38]研究目标相吻合，因此通常将弱监督目标检测视为多示例学习问题来处理。Bilen[22]首次提出基于多示例学习的弱监督目标检测框架。解决了将实例级别的候选框得分映射到图像级别的类别标记这一重要难题。完整框架包含三个主要部分：1) 候选框生成器。采用Selective search windows[24]或Edge boxes[39]算法在输入图像上生成大量的目标候选框；2) 特征提取。采用CNN对输入图像进行特征提取，再通过空间金字塔池化[40]或感兴趣区域池化[41]生成固定尺寸的候选框特征矩阵。3) 检测器。将候选框特征映射到图像类别标记，计算多示例学习损失函数，完成对图像中目标物体的定位和分类。

不幸的是，由于弱监督方法没有标准的实例级标签，网络学习经常出现过拟合。WACV2019提出........ECCV2022进一步改进..........。受启发于自然图像领域弱监督检测方法，近年来，适用于医学领域的弱监督方法也不断涌现。例如MICCAI2020...............MICCAI2022...................MICCAI2023。

* *Anatomy-Driven Pathology Detection on Chest X-rays MICCAI2023*
* *Object Discovery via Contrastive Learning for Weakly Supervised Object Detection ECCV 2022*
* *Diffusion Models for Medical Anomaly Detection MICCAI 2022*
* *Region Proposals for Saliency Map Refinement for Weakly-supervised Disease Localisation and Classification MICCAI2020*
* *Combinational Class Activation Maps for Weakly Supervised Object WACV2019*

# **方法**

在本节中，我们将介绍弱监督息肉检测网络的结构以及双阈值后处理策略。该网络的结构由三个主要部分组成:跨领域对照模块、空间类别模块(SCM)、多实例检测模块。所建议的网络的总体架构如图3所示。给定图像数据，图像特征由ImageNet[42]上预训练的ConvNet提取，候选框由跨领域对照模块生成。将候选框和图像特征送入ROI池化层进行区域特征映射，结果分别发送到空间类别模块和多实例检测模块。多实例检测模块给出位置得分，空间类别模块给出类别得分。点乘得分、夹紧列和作为预测结果。网络的回归训练由分类标签y = [y1, y2，···，yC]∈{0,1}C进行引导，其中yc = 1或0表示存在或不存在对象类c。本节的其余部分将详细讨论这三个模块以及双阈值后处理策略。



## 跨领域对照模块

我们采用SSW算法生成息肉候选框来定位息肉。但产生的候选框存在大量错误和空间冗余的候选框,给检测到来了噪声。为了解决上述问题，我们提出跨领域对照模块来产生噪声干扰较小的候选框。

具体来说，我们预先在公开的结肠息肉数据集训练了一个与我们数据异中心的分割网络，之后将我们的数据输入网络得出分割结果（P）。同时，使用SSW算法产生我们数据的初始候选框并将其转化为二进制掩码(B)。直观地看，如果如果B与P相差较大，则B可能是硬样本或错标样本，因为预先训练的分割模型必定学到了息肉的某些通用特征。在这种情况下，该边界框将会被过滤掉，不参与模型训练。因此，图1（d）所示的问题将得到缓解。

在实践中,我们按照8:2的比例划分数据集CVC-ClinicDB[25]，选取PraNet[26]作为分割网络（S）进行训练。其中80%的数据训练网络、20%的数据测试网络性能，保存网络性能最好时的权重参数。给定我们的私人数据集图像I，我们将其作为输入分割网络S，加载保存的权重参数，得到预测结果P，公式如下:



where I denotes the input image and w is the weight of the pre-trained segmentation network S.P is a binary segmentation map for two class with a size of h×w.

同时，对图像应用SSW算法并转化二进制掩码得到粗候选框B。选择Jaccard Coefficient jc= 来衡量B和P之间的差异，只选择 的候选框,最终输出为细候选框={b1,b2,...,bN}(where N is the number of proposals)，公式如下:





where denote is filtering threshold and SSW is the Selective search windows, Details are referred to[24].For each image,is a shortlist of candidate object proposals .Each element B(i,:) indicates the coordinate value of the proposals and the form is (xmin，ymin)、(xmax，ymax).

根据(3)优化后，潜在的噪声干扰可以最大限度地减少，精细的候选框被送入网络进行后续训练。

After optimization according to Eq.(3), potential noise interference can be minimized, and fine candidate boxes are fed into the network for subsequent training.

## 空间类别模块

一般来说,训练多分类网络的范式为卷积网络提取特征、全连接层拉平特征至一维向量、softmax对不同类别给分。然而，直接拉平做全连接的方式会保留大量的空间信息，因此具有更多变化的较大区域更有可能在整个训练集中具有高类别一致性。

Generally, the pattern of training multiclassification network is to extract features from convolutional network, flatten the features to one-dimensional vectors by fully connected layer, and give classification to different categories by softmax. However, the ways of directly flattening to make fully connected retains a lot of spatial information,since larger region with more variation are more likely to have high category consistency across the whole training set.

为了保证分类的准确性，我们设计了空间类别模块。形式上,空间类别模块包含一个卷积层、一个全局平均池化层和一个Flatten层as follows:



where  is a flatten operation and  is a Global Average Pooling (GAP) layer to get the classification score map. w and b are the parameters of the spatial category module, which is a 3 × 3 convolutional layer.然后将其传递给softmax运算符.The classification scores can be calculated by:



The classification scores is a matrice with the same shape as the detection scores,which is described in detail in the next section. In this case,we expect it to learn more foreground/background information to help categorize objects.

## 多实例模块

我们只有图像级别的标签来指示对象类别是否出现。为了用回归训练一个标准的对象检测器，有必要挖掘实例级监督。多实例检测模块通过对彼此相对的区域进行评分来执行检测。给定区域特征映射图extracted from a ConvNet,the multiple instance detection module takes it as input and outputs a detection scores matrice 。具体来说，特征映射图被送入三个串联的全连接层FC1~3，得到候选框拉伸向量。然后将其传递给softmax运算符，这次定义如下：





Here, has  channels,which differents from the input feature map channels, and the former is the result of the fully connected layer processing. Each elementindicates the probability of the ith proposal bi belonging to the jth category.

事实上，在这种情况下，softmax运算符独立地比较每个类别的不同区域的分数。通过这种方法，模块可以选择哪些区域更好的包含信息图像片段。

In the case, in fact, the softmax operator compares, for each region independently, class scores。In this method, In this method, the module can evaluate which region holds a better information image fragment.

After that, the scores of all proposals are generated by element-wise product⊙.During the training stage, the loss function can be formulated as follows:



## 双阈值后处理策略

为了获得更准确的预测，既不会太大而覆盖过多背景，也不会太小而退化为对象部分。我们提出了双阈值后处理策略，通过二个不同性质的阈值被设计对检测结果进行过滤。

**类别过滤策略。**虽然分类可能无法在类别上区分空间高度覆盖的候选框，但它们可以有效地将目标与背景区分开来。为了滤除不可信的候选框，我们设定类别阈值,只有类别置信度大于阈值的候选框被保留。After that，non-maxima suppression (with the threshold) is applied to the proposals.Finally, proposals with scores higher thanand overlap lighter than are held for the further filtering process.

**区域过滤策略。**As mentioned,仅对类别过滤会导致大量假阳（FP）预测。这是由于分类无法选出最具价值的区域，它更倾向于将包含大量背景的候选框作为预测结果，因为其有更多的分类信息。与其不同区域置信度更关注于候选框是否紧围物体。So naturally, 我们进一步设置了区域阈值，与固定的类别阈值不同，区域阈值是动态变化的。考虑到最糟糕的情况，输入到网络的候选框空间重叠度很高并且紧框物体的不同部分。这种情况下，候选框的区域得分会被平均分给各个候选框。因此，我们将区域阈值设置为最差情况下的平均归一，公式如下：



where N denotes the number of proposals in an image, and the result is between 0 and 1.

# **实验**

## 在本节中，我们将介绍评估数据集，介绍我们提出的方法的实施细节，然后进行全面实验，分析我们的 WSPD 及其用于弱监督息肉检测的组件。

In this section, we introduce the evaluation datasets, present the implementation details of our proposed approach.Then we perform thorough experiments to analyse our WSPD and its components for weakly-supervised polyp detection.

## **基准数据**

We evaluate our method on the popular PASCAL VOC 2007 and 2012 datasets [6] which have 9963 and 22531 image

As we focus on weakly supervised detection, only image-level labels are utilized during training. Average Precision (AP) and the mean of AP (mAP) are taken as the evaluation metrics to test our model on the testing set. Correct localization (CorLoc) is also used to evaluate our model on the trainval set to measure the localization accuracy.Both metrics are evaluated on the PASCAL criteria, i.e., IoU> 0.5 between ground truths boxes and predicted boxes.

## **实现细节Implementation details**

## **消融实验Ablation experiments**

**基干Baseline**

**跨域对照模块Cross-domain reference module(CRM)**

**空间类别模块Spatial category modules(SCM)**

**双阈值后处理Dual-threshold post-processing(DPS)**

## **与其他方法比较Comparison with other methods**

# 结论

1. Haggar, F.A., Boushey, R.P.: Colorectal cancer epidemiology: incidence, mortality,survival, and risk factors. Clinics in colon and rectal surgery 22(04), 191–197 (2009)
2. S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, and M. Tzivras, “Computer-aided tumor detection in endoscopic video using color wavelet features,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 3, pp. 141–152, Sep. 2003.
3. S. Hwang, J. Oh, W. Tavanapong, J. Wong, and P. C. De Groen, “Polyp detection in colonoscopy video using elliptical shape feature,” in *Proc. IEEE Int. Conf. Image Process.*, 2007, vol. 2, pp. II-465–II-468.
4. L. A. Alexandre, N. Nobre, and J. Casteleiro, “Color and position versus texture features for endoscopic polyp detection,” in *Proc. Int. Conf. BioMed. Eng. Informat.*, 2008, vol. 2, pp. 38–42.
5. S. Ameling, S. Wirth, D. Paulus, G. Lacey, and F. Vilarino, “Texturebased polyp detection in colonoscopy,” in *Proc. Bildverarbeitung fur die Medizin*, 2009, pp. 346–350.
6. J. Bernal, J. Sanchez, and F. Vilarino, “Towards automatic polyp detection with a polyp appearance model,” *Pattern Recognit.*, vol. 45, no. 9, pp. 3166–3182, 2012.
7. S. Park, D. Sargent, I. Spofford, K. G. Vosburgh, and Y. A-Rahim, “A colon video analysis framework for polyp detection,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1408–1418, May 2012.
8. Qadir H A , Balasingham I , Solhusvik J ,et al.Improving Automatic Polyp Detection Using CNN by Exploiting Temporal Dependency in Colonoscopy Video[J].IEEE Journal of Biomedical and Health Informatics, 2020, 24(1):180-193.DOI:10.1109/JBHI.2019.2907434.
9. Jiang B, Li N, Shi X, Zhang S, Li J, de Bock GH, Vliegenthart R, Xie X. Deep Learning Reconstruction Shows Better Lung Nodule Detection for Ultra-Low-Dose Chest CT. Radiology. 2022 Apr;303(1):202-212. doi: 10.1148/radiol.210551. Epub 2022 Jan 18. PMID: 35040674.
10. Wang L. Deep Learning Techniques to Diagnose Lung Cancer. Cancers (Basel). 2022 Nov 13;14(22):5569. doi: 10.3390/cancers14225569. PMID: 36428662; PMCID: PMC9688236.
11. Feng J, Jiang J. Deep Learning-Based Chest CT Image Features in Diagnosis of Lung Cancer. Comput Math Methods Med. 2022 Jan 19;2022:4153211. doi: 10.1155/2022/4153211. PMID: 35096129; PMCID: PMC8791752.
12. Nguyen VD, Bui ND, Do HK. Skin Lesion Classification on Imbalanced Data Using Deep Learning with Soft Attention. Sensors (Basel). 2022 Oct 4;22(19):7530. doi: 10.3390/s22197530. PMID: 36236628; PMCID: PMC9572097.
13. Balkenende L, Teuwen J, Mann RM. Application of Deep Learning in Breast Cancer Imaging. Semin Nucl Med. 2022 Sep;52(5):584-596. doi: 10.1053/j.semnuclmed.2022.02.003. Epub 2022 Mar 24. PMID: 35339259.
14. A N R , B M W S , A J G ,et al.Automated Detection of Acute Respiratory Distress Syndrome from Chest X-Rays Using Directionality Measure and Deep Learning Features[J].Computers in Biology and Medicine, 2021.DOI:10.1016/j.compbiomed.2021.104463.
15. Yu, L., Chen, H., Dou, Q., Qin, J., Heng, P.A.: Integrating online and offline threedimensional deep learning for automated polyp detection in colonoscopy videos.IEEE JBHI 21(1), 65–75 (2016)
16. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun.Faster R-CNN: Towards real-time object detection with region 候选框 networks. In NIPS, 2015.
17. Redmon J , Divvala S , Girshick R ,et al.You Only Look Once: Unified, Real-Time Object Detection[C]//Computer Vision & Pattern Recognition.IEEE, 2016.DOI:10.1109/CVPR.2016.91.
18. Naqvi M, Gilani SQ, Syed T, Marques O, Kim HC. Skin Cancer Detection Using Deep Learning-A Review. Diagnostics (Basel). 2023 May 30;13(11):1911. doi: 10.3390/diagnostics13111911. PMID: 37296763; PMCID: PMC10252190.
19. Zheng Ying, Wang Zezhou. Interpretation of global colorectal cancer statistics. Chinese Journal of Epidemiology, 2021, 42(1): 149-152.
20. A. M. Leufkens, M. G. H. van Oijen, F. P. Vleggaar, and P. D. Siersema, “Factors influencing the miss rate of polyps in a back-to-back colonoscopy study,” Endoscopy, vol. 44, no. 5, pp. 470–475, 2012.
21. L. Rabeneck, J. Souchek, and H. B. El-Serag, “Survival of colorectal cancer patients hospitalized in the veterans affairs health care system,” Am. J. Gastroenterology, vol. 98, no. 5, pp. 1186–1192, 2003.
22. H. Bilen and A. Vedaldi. Weakly supervised deep detection networks. In CVPR, pages 2846–2854, 2016. 1, 2, 3, 4, 7, 8
23. Chen Y L. Analysis of missed diagnosis rate and related factors of colorectal polyps in colonoscopy [J]. 2020. (in Chinese)
24. Jasper RR Uijlings, Koen EA Van De Sande, Theo Gevers, and Arnold WM Smeulders. Selective search for object recognition. IJCV, 2013.
25. Bernal, J., S´anchez, F.J., Fern´andez-Esparrach, G., Gil, D., Rodr´ıguez, C., Vilari˜no, F.: Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. CMIG 43, 99–111 (2015)
26. Deng-Ping Fan, Ge-Peng Ji, Tao Zhou,Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao.Pranet: Parallel reverse attention network for polyp segmentation. MICCAI, pages 263–273, 2020.
27. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015).
28. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested U-Net architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 3–11. Springer, Cham (2018).
29. ha, D., et al.: Resunet++: an advanced architecture for medical image segmentation. In: 2019 IEEE International Symposium on Multimedia (ISM), pp. 225–2255. IEEE (2019)
30. Murugesan, B., Sarveswaran, K., Shankaranarayana, S.M., Ram, K., Joseph, J., Sivaprakasam, M.: Psi-Net: shape and boundary aware joint multi-task deep network for medical image segmentation. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 7223–7226 (2019)
31. Cheng, M., Kong, Z., Song, G., Tian, Y., Liang, Y., Chen, J.: Learnable orientedderivative network for polyp segmentation. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 720–730. Springer, Cham (2021).
32. Zhao, X., Zhang, L., Lu, H.: Automatic polyp segmentation via multi-scale subtraction network. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 120–130. Springer, Cham (2021)
33. Fang, Y., Chen, C., Yuan, Y., Tong, K.: Selective feature aggregation network with area-boundary constraints for polyp segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 302–310. Springer, Cham (2019).
34. Zhang, R., Li, G., Li, Z., Cui, S., Qian, D., Yu, Y.: Adaptive context selection for polyp segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12266, pp. 253–262. Springer, Cham (2020).
35. Shen, Y., Jia, X., Meng, M.Q.-H.: HRENet: a hard region enhancement network for polyp segmentation. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 559–568. Springer, Cham (2021).
36. Nguyen, T.-C., Nguyen, T.-P., Diep, G.-H., Tran-Dinh, A.-H., Nguyen, T.V., Tran, M.-T.: CCBANet: cascading context and balancing attention for polyp segmentation. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 633–643. Springer, Cham (2021)
37. Xiuquan Du,Xuebin Xu,and Kunpeng Ma.:ICGNet: Integration Context-Based Reverse-Contour Guidance Network for Polyp Segmentation. IJCAI 2022.
38. Dietterich T G, Lathrop R H and Lozano-pérez T.1997.Solving the multiple instance problem with axis-parallel rectangles. Artificial Intelligence, 89(1-2): 31 - 71. [DOI: 10.1016/S0004-3702(96)00034-3]
39. C. L. Zitnick and P. Dollar. Edge boxes: Locating object 候选框 from edges. In Proc. ECCV, pages 391–405, 2014.
40. K. He, X. Zhang, S. Ren, and J. Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. In Proc. ECCV, pages 346–361, 2014.
41. Girshick R.2015.Fast r-cnn// Proceeding of the International Conference on Computer Vision[C]. Santiago, Chile: IEEE, 1440 - 1448.[DOI: 10.1109/ICCV.2015.169]
42. Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy,Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. IJCV, 2015