·医学工程·

基于多专家系统的长尾内窥镜视频伪影检测算法研究

金浩龙　冯庆敏　刘胜林*

（华中科技大学同济医学院附属协和医院生物医学工程研究所，武汉 430000）

【摘要】 **目的**　通过长尾学习的视角探讨内窥镜视频伪影检测的方法。**方法**　通过分析公开数据集EAD 2020中内窥镜视频伪影分布特性，采用解耦框架，提出一种基于多专家系统的长尾内窥镜伪影检测算法，将头部类别与尾部类别分为两组，不同专家处理不同组别，并且设计了专家结果聚合的方式，最后通过引入视觉提示（visual prompt）辅助专家，提高了伪影检测的全类平均精度（mean Average Precision, mAP）。**结果**　得益于更强大的主干网络与更先进的分类器设计，相比于现有流行的集成方法，内窥镜连续视频帧检测的mAP从0.440提高至0.491，独立视频帧检测的mAP从0.339提升至0.384，并且本方法在所有类别上全面领先于集成方法。**结论**　基于多专家系统的长尾内窥镜视频伪影检测框架通过分组与多专家有效平衡了不同伪影检测的表现，有效提高了尾部类别的检测性能，同时没有损害头部类别的性能，缓解了类别不平衡带来的问题，为数据不平衡场景下的医学图像检测任务提供了一个新的解决思路。

【关键词】 深度学习；长尾学习；内窥镜；伪影检测；多专家系统；视觉提示

【中图分类号】 TH77; TP18　　【文献标识码】 A　【DOI】 10.11876/mimt202503002

Long-tailed endoscopic artifact detection based on multi-expert system

Jin Haolong, *Feng Qingmin*, *Liu Shenglin**

(*Institute of Biomedical Engineering*, *Union Hospital*, *Tongji Medical College*, *Huazhong University of Science and Technology*, *Wuhan 430000*, *China*)

【Abstract】 **Objective**　Exploring the challenges of artifact detection in endoscopic videos through the perspective of long-tailed learning. **Methods**　Through analysis of artifact distribution characteristics in endoscopic videos from the public EAD2020 dataset, we develop a long-tailed endoscopic artifact detection algorithm via a decoupled framework. The proposed multi-expert system strategically partitions head and tail classes into two distinct subsets, with domain-specific experts processing each group separately. We innovatively design a hierarchical expert aggregation mechanism and further integrate visual prompts to enhance cross-group discrimination capabilities, ultimately achieving a significant improvement in artifact detection performance as measured by the comprehensive mean Average Precision (mAP) metric. **Results**　Benefiting from a more powerful backbone network and advanced classifier design, the proposed method achieves significant improvements in endoscopic video analysis: mAP for sequential video frame detection increases from 0.440 to 0.491, while mAP for individual video frame detection rises from 0.339 to 0.384. Notably, our approach comprehensively outperforms existing ensemble methods across all artifact categories. **Conclusion**　The proposed multi-expert framework for long-tailed endoscopic video artifact detection achieves balanced performance across artifact categories through grouping and multi-expert methods. This architecture significantly enhances detection performance for tail classes while maintaining head class accuracy, effectively mitigating the inherent class imbalance problem. This study provides a novel direction for medical image detection tasks, particularly in data-imbalanced scenarios.

【Key words】 Deep learning; Long-tailed learning; Endoscope; Artifact detection; Multi-expert system; Visual prompt

---

基金项目：湖北省重点研发计划项目（2022BCA025）

第一作者：金浩龙，男，在读硕士研究生，研究方向：医学图像处理、长尾学习

通信作者：刘胜林，男，博士研究生，高级工程师，研究方向：人工智能医疗器械、医学图像处理、裸眼三维可视化技术

内窥镜检查是一种广泛应用的临床检查手段，用于多种癌症的早期检测与治疗性手术。然而内窥镜影像视频帧存在缺陷，它们普遍受到多重伪影的严重干扰，如像素过饱和、运动模糊、失焦、镜面反射、气泡等[1]。这些伪影不仅会妨碍诊断过程中对目标组织的观察，还会影响后续分析所涉及的各项处理流程[2]，如报告生成所需的视频帧检索。精确识别这些伪影对于实现高质量的内窥镜帧修复至关重要，是开发可靠计算机辅助内镜工具的关键[3]。

在实际应用中，内窥镜影像伪影通常呈现长尾分布特性，现有研究已经贡献了相当多的学习方法[4]，大致可以分为3类[5]：类重新平衡、信息增强和模块改进。类重新平衡是长尾学习中的主流范式，旨在调整训练样本的分布或是损失权重，以便在所有类别上获得平衡的效果，如根据样本频率决定采样比率[6]、根据样本频率调整不同类别损失权重[7]、动态调整正负梯度权重[8]等。这类技术有2个缺陷，即尾部类别面临过拟合风险，以及尾部类别性能的提升往往以头部类别性能下降为代价。信息增强方法旨在将其他信息引入模型培训中，以改善模型在长尾学习中的表现，如从头到尾知识转移[9-10]、知识蒸馏[11-12]或数据增强[13-14]，这些方法避免了牺牲头部类别表现的缺陷，同时改善了尾部类别的表现。然而，信息增强依赖更多的算力资源以及对数据集专业领域的了解程度。

最近，改进网络模块的方法成为长尾任务的又一主流解决方案。部分方法通过改进模块结构来提高性能，如改进特征提取器的表示学习[15-16]，或是改进分类器设计以增强性能[17]，还有改进模块训练方式的，如解耦方法和集成学习。解耦方法[18-19]将长尾训练过程分为表示学习与分类器训练两个阶段；集成学习[20-21]则在策略上采用多个专家网络来解决长尾问题。现有的流行方法主要是集成学习，但同样有算力需求大、假阳性率高的问题。解耦方法在此类问题的解决上仍有潜力未被挖掘。本文在上述研究基础上，克服现有长尾学习方法的弱点，提出一种新的解耦学习框架，结合均衡分组的思想与多专家理念改进了网络，使得模型检测尾部类别伪影的性能显著提升，在长尾分布的内窥镜影像上表现良好。

1 算法原理与模型设计

1.1 模型设计

由于内窥镜视频伪影检测是后处理，对于速度与效率的要求并不高，更看重准确度。因此，选择两阶段的目标检测网络而非单阶段检测网络。为了提高伪影检测的准确度，选用Cascade R-CNN[22]作为网络主要结构，如图1所示。模型由3个部分构成：主干网络（backbone）用于提取输入图片的特征，本文采用先进的主干网络ConvNeXt-T[23]与Swin-T[24]作对比；区域提案网络（region proposal network, RPN）用于初步提取特征图中可能存在的感兴趣区（region of interest, ROI），仅区分前景与背景，要选出尽可能多的前景区域，初步调整边界框位置，然后会执行1次非极大值抑制（non-maximum supression, NMS）操作，消除多余、重复的候选框；最后Cascade R-CNN通过级联结构，将上一步获得的候选区域进行进一步的多次的分类、回归，细化结果。

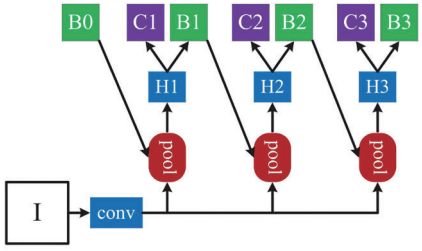


图1 Cascade R-CNN网络结构[22]

每个分类头的结构如图2所示，将每个分类头都改造成多专家分支形式，每个头部类别上的多专家分支推理完成后都需要进行专家聚合的操作并且映射回原数据标签。在该算法框架中额外引入了prompt方法，直接作用于特征图，每个专家分支单独维护1个prompt，直接提高了每个组分支的表现，同时只增加了少量须学习的参数。

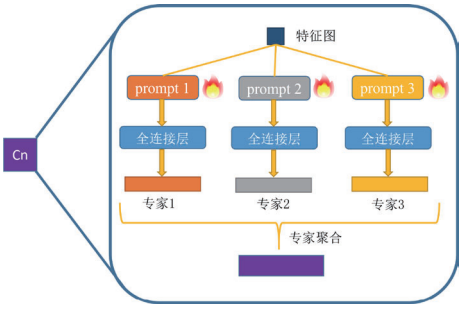


图2 多专家分支的分类头结构

1.2 均衡分组

本研究在数据集EAD 2020[1]上进行。将原始数据类别分为K组，不同组的类别由不同专家负责。将类别c分配给组G_k，则有：

$$n_k^l < N(c) < n_k^u,\ 0 < k < K \tag{1}$$

其中，$N(c)$是类别c的实例总数，n_k^l和n_k^u是组G_k所包含类别实例数量的下界和上界，是超参数。为了确保每个组之间没有重叠，每个类别只对应1个组，故令$n_{k+1}^l = n_k^u$。需要注意的是，原始数据是长尾数据集，分组是基于一个朴素的方法论，即遇到复杂问题就先拆解问题再解决。长尾问题的根源就在于头部类别信息过于丰富以至于淹没了尾部类别的信息，或者说是在梯度反向传播过程中，头部类别对尾部类别的负样本梯度惩罚远大于尾部类别的正样本梯度，分组将头部类别与尾部类别分隔开，大大减轻尾部类别的压倒性负样本梯度。

不平衡比[4]和Gini系数[5]都是代表数据集不平衡程度的指标。本文通过调整这些长尾指标来确定分组方式，目标是尽可能降低不平衡比与Gini系数这2个长尾指标，从图3数据集的样本统计情况来看，从气泡到光照不足导致的对比度过低，样本数量下跌最多，可以此为界限将数据集类别分为2组，并计算不平衡比与Gini系数加以验证。此外，为了辨别检测对象是否为背景，额外增加前景—背景的组别，记为G_0，同时能够避免背景类与其他所有类别间的不平衡。如表1所示，原始数据集被分为更加平衡的2组，而前景—背景组G_0用来辅助分类。

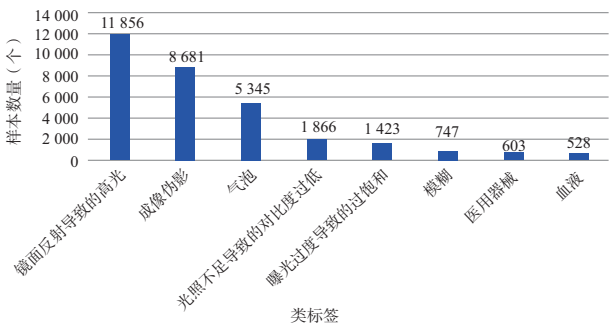


图3 训练数据的类别分布情况

表1 数据集分组及其对应的长尾指标

组别	包含的类别	不平衡比	Gini系数
G_0	背景、前景	—	—
G_1	血液、医用器械、模糊、曝光过度导致的过饱和、光照不足导致的对比度过低	3.53	0.27
G_2	气泡、成像伪影、镜面反射导致的高光	2.22	0.17
原始数据集	所有类别	22.45	0.54

1.3 专家聚合

分类头被分为多个分支，理想情况下可以视为多专家系统。但经过实践探索，传统的专家系统并没有很好地发挥功效，根源在于专家系统的门控网络。直接用1个硬门控分支对实例进行组别分类效果不佳，而本文采用了软门控（soft gating）机制。软门控提供额外的前景—背景概率，每位专家的结果将与前景概率相乘，再剪裁拼接，获得最终结果。

经过均衡分组后，不再对所有类别进行softmax操作，而是在组内进行softmax操作，有效抑制了专家分支遇到非本组类别由于softmax归一化产生的误报。组内softmax计算公式如下：

$$p_j = \frac{e^{z_j}}{\sum_{i \in Gn} e^{z_i}},\ \{n \mid j \in G_n\} \tag{2}$$

其中，p_j是输出的第j类预测概率值。G_n代表分组后的第n组类别集合；z_i是检测结果为第i类的预测logits值。

对于训练阶段，为求损失，需要将原始标签重新映射为分组后的标签，每个原始标签映射到每个组都各自有1个标签；接着每个组经组内softmax后各自与映射后的标签求损失。损失函数L的计算公式为：

$$L = -\sum_{n=0}^{N}\sum_{i \in G_n} y_i^n \log\left(p_i^n\right) \tag{3}$$

其中，y_i是映射后的标签；p_i是在测试或推理阶段得到模型输出logits值后进行组内softmax得到的概率预测值。除了前景—背景组G_0外，所有“其他组”均被舍弃，因为“其他组”的作用在组内softmax操作之后就已经结束了，它与组内前景类是此消彼长的关系。具体表达式如下：

$$\begin{aligned} P^0 &= [p_0^0, p_1^0]^T \\ P^1 &= [p_0^1, p_1^1, \ldots, p_{K_1}^1]^T \\ &\vdots \\ P^n &= [p_0^n, p_1^n, \ldots, p_{K_n}^n]^T \end{aligned} \tag{4}$$

舍去1至n组的“其他组”类预测：

$$\begin{aligned} P^{0\prime} &= [p_1^0] \\ P^{1\prime} &= [p_1^1, \ldots, p_{K_1}^1]^T \\ &\vdots \\ P^{n\prime} &= [p_1^n, \ldots, p_{K_n}^n]^T \end{aligned} \tag{5}$$

p_0^0是前景类概率，将其与其他组中所有类别相乘，再进行拼接得到：

$$P = \begin{pmatrix} P^{0\prime} \\ P_0^0 \cdot P^{1\prime} \\ \vdots \\ P_0^0 \cdot P^{n\prime} \end{pmatrix} \tag{6}$$

最后还需要将分组的类别标签映射回原始标签域，输出正常的全类别概率。

1.4 分支增强

为进一步增强各分支的表现，引入prompt方法。视觉提示（visual prompt），是可学习的像素块参数，不像语言模型中的prompt主要靠人工设计，该视觉提示只要经过合理的初始化，它们的作用相同，即引导模型关注关键信息或执行特定任务，提升输出结果的针对性、准确性和可控性。不论是计算机视觉还是自然语言处理领域，prompt都是作为参数高效微调模型的手段，而在本文中，创新性地将这种手段应用到长尾学习当中，利用可学习的prompt增强各专家分支的表现，辅助组间区分的同时优化特征表示。计算原理如下：

$$y = head[backbone(x) \times channel_prompt] \tag{7}$$

其中，x为输入数据，y为模型输出的概率结果，*head*指模型结构中的检测头，*backbone*指模型的主干网络，*channel_prompt*即为可学习的prompt模块。

2 实验结果与分析

2.1 实验数据

EAD 2020[1]数据集专为推进计算机视觉技术在医疗内镜智能化诊断中的转化应用而构建。该数据集整合全球7家顶级医疗中心的内窥镜影像数据，共计2 532个内镜视频帧，包含2 299个单帧样本和5组共计233个连续序列数据，标注总数高达31 049个。该数据集系统采集了涵盖不同解剖部位、病理阶段和操作环境的视频序列与关键帧图像，专注于内窥镜视频伪影检测领域。数据集中不仅提供精确的边界框标注，更包含像素级实例分割掩码的高质量标注信息，总共有8类典型的内窥镜伪影：模糊（blur）、气泡（bubbles）、镜面反射导致的高光（specularity）、成像伪影（imaging artefacts）、曝光过度导致的过饱和（saturation）、光照不足导致的对比度过低（contrast）、血液（blood）、医用器械（instrument），为实现高质量的内窥镜视频修复与重建提供了坚实的基础。图3提供了该数据集各类别的统计情况，该数据集具有长尾分布特征。

2.2 实验方法

本文对比了多种主流的内窥镜伪影检测算法，包括RetinaNet、Cascade R–CNN与多模型集成方法。对于已有方法采用原有设置，而本方法实验设置细节如下：首先，在非极大值抑制配置中，对于Cascade R–CNN的3个级联检测头的输出，NMS的阈值分别设置为0.4、0.5、0.6，这是为了去除重复的检测对象；其次，每个检测头输出的检测框置信度阈值同样设置为0.4、0.5、0.6，分类得到的最高概率低于该阈值则舍去对应边界框。所有输入图像尺寸缩放为512像素 × 512像素。本文注意到待检测样本中有大量的小目标，尤其是镜面反射导致的高光与成像伪影，故在检测网络中RPN的anchor默认大小需要作出调整，这里采用5种尺度的anchor：16^2，32^2，96^2，224^2，320^2，3种比例：1∶1，1∶2，

2∶1。

模型共训练（24+8）个epochs。实验采用混合优化器。解耦学习训练的第1阶段共24个epochs，采用随机梯度下降（stochastic gradient descent, SGD）优化器，初始学习率为0.02，动量为0.9，权重衰减系数为0.000 1。第2阶段共8个epochs，需要训练两部分：一是分类头，还是用初始学习率为0.02、动量为0.9、权重衰减系数为0.000 1的SGD优化器；二是特征提示部分，用初始学习率为0.001、权重衰减系数为0.01、其他参数默认的AdamW优化器。关于学习率调整策略，用LinearLR进行学习率预热，起始因子为0.001，迭代轮数为500；用MultiStepLR策略控制SGD优化器学习率，表征学习阶段第16个和第22个epoch时学习率乘以0.1，分类器训练阶段第5个和第7个epoch时学习率乘以0.1。实验指标统一采用全类平均精度（mean Average Precision, mAP）[25]。mAP是目标检测中衡量模型准确性的核心指标，通过计算多个类别在不同置信度阈值下的平均精度均值，综合反映模型对目标定位和分类的整体性能。

2.3 实验结果分析

2.3.1 不同检测方法的对比实验

如表2所示，本方法取得了最佳成绩，且计算量远低于集成方法。RetinaNet是单阶段检测方法，效果相较于其他两阶段的方法并不理想。集成方法是同时用3种网络进行检测，最终将3种网络的检测结果进行聚合与非极大值抑制，这种方法虽然有效，但要消耗更长的推理时间与更多的计算资源。其中，mAP_d是指独立视频帧的检测平均精度，mAP_{seq}是指连续的视频帧序列的检测平均精度。数据对比证明本研究提出的方法在内窥镜伪影检测上效果显著。

表2 不同方法实验结果对比

方法	mAP_d	mAP_{seq}
RetinaNet	0.286	0.391
Cascade R–CNN + DCN + ResNeXt101	0.331	0.437
集成方法(Cascade R-CNN, Faster R-CNN, RetinaNet)	0.339	0.440
本文方法	0.384	0.491

注：mAP_d为独立视频帧的检测平均精度，mAP_{seq}为连续的视频帧序列的检测平均精度。

如表3所示，对比了两种不同方法在各个类别上的平均精度（average precision, AP），最后一列为平均值，即mAP。其中类别标签是由头部类别向尾部类别依次排列，按提升幅度来看，尾部类别的血液、医用器械、曝光过度导致的过饱和提升最大，分别提升了0.066、0.058、0.061。比较本方法和集成方法在头部类别与尾部类别上的性能差异发现，本方法提升每个头部类别的mAP为0.014，提升每个尾部类别的mAP为0.061。本结果证明本文提出的方法能够有效提高尾部类别的性能，同时不损害头部类别的性能。

表3 不同方法在不同类别伪影检测上的平均精度对比

方法	镜面反射导致的高光	成像伪影	气泡	光照不足导致的对比度过低	曝光过度导致的过饱和	模糊	医用器械	血液	平均
集成方法	0.341	0.304	0.284	0.493	0.276	0.263	0.541	0.210	0.339
本文方法	0.382	0.321	0.301	0.537	0.337	0.318	0.599	0.276	0.384

2.3.2 不同主干网络的对比实验

本文采用的主干网络ConvNeXt–T是卷积结构的特征提取器，进一步探究本文提出的方法在不同主干网络下的表现，分别用ResNet–50、Swin–T、ConvNeXt–T 3个参数量相当的主干网络进行实验，对比结果如表4所示。通过实验数据可知，ConvNeXt–T要优于ResNet–50，对于独立视频帧，mAP达到了0.384，提升了0.01；对于序列视频帧，mAP达到了0.491，提升了0.02。ConvNeXt–T与Swin–T相比，mAP相差不大，前者有微小的优势。为了进一步分析，表5列出了3种主干网络在不同类别上的AP对比。比较了Swin–T和ConvNeXt–T主干后，发现Swin–T在光照不足导致的对比度过低、模糊和医用器械上更优，ConvNeXt–T在其他类别的伪影检测上更优，原因可能是恰好对比度过低、失焦模糊和器械遮挡，这些往往是大目标，Swin–T由于ViT的结构更容易捕获大范围的信息，而ConvNeXt–T由于卷积架构易捕获细粒度特征，在高光、气泡这种细小目标上表现良好。

表 4 不同主干网络实验结果对比

主干网络	mAP_d	mAP_{seq}
ResNet–50	0.374	0.471
Swin–T	0.383	0.485
ConvNeXt–T	0.384	0.491

注：mAP_d为独立视频帧的检测平均精度，mAP_{seq}为连续的视频帧序列的检测平均精度。

表 5 不同主干网络在不同类别伪影检测上的平均精度对比

主干网络	镜面反射导致的高光	成像伪影	气泡	光照不足导致的对比度过低	曝光过度导致的过饱和	模糊	医用器械	血液
ResNet–50	0.374	0.306	0.288	0.526	0.319	0.317	0.591	0.272
Swin–T	0.366	0.317	0.289	0.546	0.326	0.331	0.610	0.276
ConvNeXt–T	0.382	0.321	0.301	0.537	0.337	0.318	0.599	0.276

2.3.3 不同方法的检测结果定性分析

在该内窥镜视频伪影检测任务中，许多研究者采用了集成多个模型结果来提高性能的方法，如同时使用 Cascaded R–CNN、Faster R–CNN 和 RetinaNet 3个模型聚合处理它们的结果得到最终结果，这在运算复杂度上远高于本研究多专家分支的方法。在具体的复杂场景下，本框架好于集成多模型的方法，图 4A 为基于集成多模型框架的检测结果，图 4B 为本研究方法的检测结果，其中所有褐色检测框为镜面反射导致的高光，由于数量多且区域小故未注释名称。从图中可以看出，集成模型的检测结果出现密集且冗余的检测框，非极大值抑制并未很好地解决多个检测模型间的冲突，同时出现了单个检测目标被检测为多个子目标的情况，如图 4A 的右上区域所示，而本文检测方法避免了这些问题。图 5 展示了本方法的检测结果与真实标签值的对比，本方法的检测效果在一般场景下基本满足伪影检测的需求，能准确定位与识别伪影。

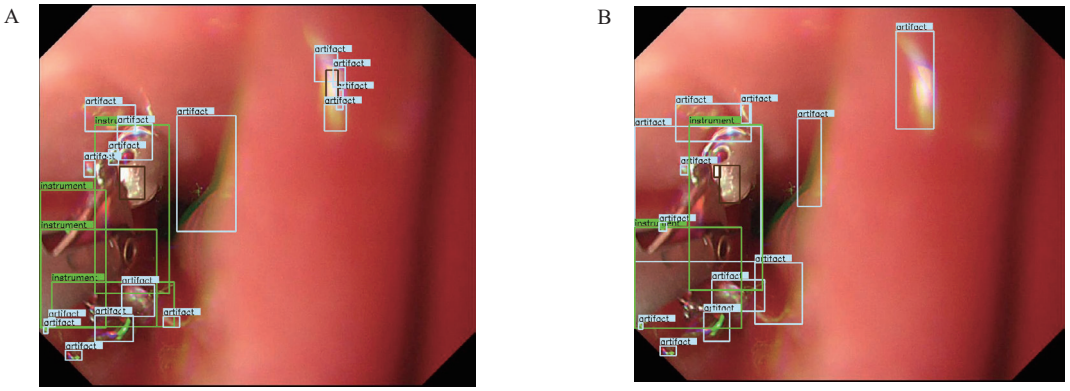


图 4 复杂场景检测结果示例 A 为集成方法检测结果；B 为本文方法检测结果

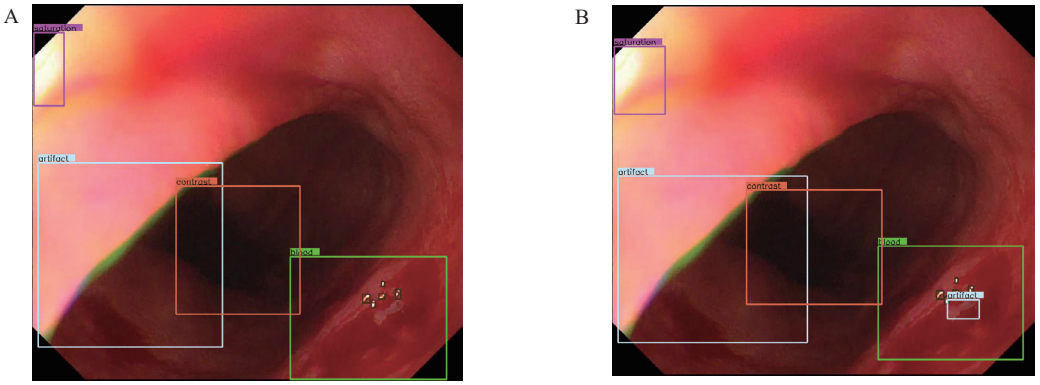


图 5 一般场景检测结果示例 A 为真实标签值；B 为本文方法检测结果

3 讨论

通过实验结果的详细分析，可以看出本文提出的多专家系统的解耦学习框架能有效提高长尾分布下内窥镜伪影检测的精度，应用更先进的主干网络能进一步改善模型的表现。相比于流行的集成方法，消耗更少的算力，获得更佳的性能，能够显著增强尾部类别上的表现。

本研究设计的多专家分组的模式，源于分而治之的朴素思想。对于长尾问题来说，本质上是样本数量特别大的类别得到了过度的关注，样本数量特别稀少的类别关注不足，因此将两者分开，各成1组，每组都实现了组内平衡，这是本方法能够有效的根本原因。其中，引入的prompt方法主要用于辅助专家分组，每个组维护1个可学习的prompt，相当于专家的“标识”。本文的方法可以说是把长尾问题转换成了普通的平衡分布问题，使头部类别与尾部类别得到了相同的关注，既克服了传统长尾学习方法提升尾部类别性能会以牺牲头部类别性能为代价的缺点，又避免了集成方法带来的计算成本高的问题。此外，实验还对比了不同主干网络下模型的表现，先进的主干网络利于提取更优的特征表示，故性能更优；而由于网络结构的差异即CNN结构和ViT结构，在不同类别上表现也有差异，前者更擅长提取细节，发现微小目标，后者着眼全局，对大目标更敏感。

本文实验均在Cascade R-CNN的网络结构上进行，而本方法是模型无关的，下一步应当采用更新的网络进行实验，或是加入更多的优秀结构，如FPN可以换为更先进的结构。本方案与数据增强方法并不冲突，可以采用一些数据增强方法来加强特征表示，如Mixup、RandAugment等。此外，长尾学习还有另外一种视角，即基于类别区分的难易程度进行加权，观察本研究的实验数据发现，医用器械虽然样本数量少但很容易区分，气泡样本数量多却相对难以精准检测，因此，在均衡分组基础上通过难易区分样本的视角进一步优化模型也是未来的一个研究方向。

本研究聚焦于内窥镜视频伪影检测的长尾分布情况，提出了1个基于多专家系统的解耦学习框架。该框架中，首先通过均衡分组的方式来决定各个专家负责的类别；其次为了各组专家间的配合，提出了有效的专家聚合方式；最后引出用于模型微调的prompt方法增强专家分支的表现。该框架有效提高了对尾部类别的检测表现，缓解了类别不平衡带来的问题，为长尾场景下医学图像检测任务提供了一个行之有效的解决方案。

[参 考 文 献]

[1] Ali S, Zhou F, Braden B, et al. An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy [J]. Sci Rep, 2020, 10(1): 2748.

[2] Ali S. Where do we stand in AI for endoscopic image analysis? Deciphering gaps and future directions [J]. NPJ Digit Med, 2022, 5(1): 184.

[3] Ali S, Zhou F, Daul C, et al. Endoscopy artifact detection (EAD 2019) challenge dataset [EB/OL]. (2019-05-08) [2025-04-14]. https://arxiv.org/abs/1905.03209.

[4] Yang L, Jiang H, Song Q, et al. A survey on long-tailed visual recognition [J]. International Journal of Computer Vision, 2022, 130: 1837-1872.

[5] Zhang Y, Kang B, Hooi B, et al. Deep long-tailed learning: a survey [J]. IEEE Trans Pattern Anal Mach Intell, 2023, 45(9): 10795-10816.

[6] Yaman B, Mahmud T, Liu C H. Instance-aware repeat factor sampling for long-tailed object detection [EB/OL]. (2023-05-14) [2025-04-14]. https://arxiv.org/abs/2305.08069.

[7] Tan J, Lu X, Zhang G, et al. Equalization loss v2: a new gradient balance approach for long-tailed object detection [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, Nashville, TN, USA. IEEE, 2021: 1685-1694.

[8] Wang J, Zhang W, Zang Y, et al. Seesaw loss for long-tailed instance segmentation [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, Nashville, TN, USA. IEEE, 2021: 9690-9699.

[9] Liu J, Sun Y, Han C, et al. Deep representation learning on long-tailed data: a learnable embedding augmentation perspective [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, Seattle, WA, USA. IEEE, 2020: 2967-2976.

[10] Liu B, Li H, Kang H, et al. GistNet: a geometric structure transfer network for long-tailed recognition [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, Montreal, QC, Canada. IEEE, 2021: 8189-8198.

[11] He Y Y, Wu J, Wei X S. Distilling virtual examples for long-tailed recognition [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, Montreal, QC, Canada. IEEE, 2021: 235-244.

[12] Zhang S, Chen C, Hu X, et al. Balanced knowledge distillation for long-tailed learning [J]. Neurocomputing, 2023, 527: 36-46.

[13] Chu P, Bian X, Liu S, et al. Feature space augmentation for long-tailed data [C]// ECCV 2020: 16th European Conference, 2020, Glasgow, UK. Springer, Cham, 2020: 694-710.

[14] Zang Y, Huang C, Loy C C. FASA: feature augmentation and sampling adaptation for long-tailed instance segmentation [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, Montreal, QC, Canada. IEEE, 2021: 3437-3446.

（下转第23页）

乏筛查方法：WS/T 677—2020［S/OL］.［2025-02-26］. https://www.nhc.gov.cn/wjw/yingyang/202005/4edc890001cd4041bec5108bee468d66/files/1739783560802_56160.pdf.
［8］ Lazzarino G, Longo S, Amorini A M, et al. Single-step preparation of selected biological fluids for the high performance liquid chromatographic analysis of fat-soluble vitamins and antioxidants［J］. Journal of Chromatography A, 2017, 1527: 43-52.
［9］ Zhou J, Wang F, Ma Y, et al. Vitamin D_3 contributes to enhanced osteogenic differentiation of MSCs under oxidative stress condition via activating the endogenous antioxidant system［J］. Osteoporosis International, 2018, 29(8): 1917-1926.
［10］ Maghsoumi-Norouzabad L, Javid A Z, Mansoori A, et al. Vitamin D_3 supplementation effects on spermatogram and oxidative stress biomarkers in asthenozoospermia infertile men: a randomized, triple-blind, placebo-controlled clinical trial［J］. Reproductive Sciences, 2022, 29(3): 823-835.
［11］ Tijerina A, Garza A, López A, et al. An HPLC-UV method to assess human plasma 25(OH)D_3［J］. Nutrients, 2024, 16(14): 2304.
［12］ Park J, Kim M, Kim W, et al. Ultrasensitive detection of 25-hydroxy vitamin D_3 in real saliva using sandwich-type electrochemical aptasensor［J］. Sensors and Actuators B: Chemical, 2022, 355: 131239.
［13］ 李普庆. 分子印迹膜的制备及应用进展［J］. 合成技术及应用, 2015(3): 21-25.
［14］ Sun Y, Luo Y, He Y, et al. CNF@MIP-based electrochemical molecularly imprinted sensor for detection of chloramphenicol in milk［J/OL］. Journal of Solid State Electrochemistry, 2025［2025-02-26］. https://link.springer.com/article/10.1007/s10008-025-06235-y.
［15］ Haupt K, Mosbach K. Molecularly imprinted polymers and their use in biomimetic sensors［J］. Chemical Reviews, 2000, 100(7): 2495-2504.
［16］ Garg M, Pamme N. Strategies to remove templates from molecularly imprinted polymer (MIP) for biosensors［J］. Trends in Analytical Chemistry, 2024, 170: 117437.
［17］ Pilvenyte G, Ratautaite V, Boguzaite R, et al. Molecularly imprinted polymer-based electrochemical sensors for the diagnosis of infectious diseases［J］. Biosensors, 2023, 13(6): 620.
［18］ Liu R, Zhao M, Zhang X, et al. Advances and challenges in molecularly imprinted electrochemical sensors for application in environmental, biomedicine, and food safety［J/OL］. Critical Reviews in Analytical Chemistry, 2025［2025-02-26］. https://www.tandfonline.com/doi/10.1080/10408347.2025.2460751.
［19］ Zhang Y, Yang Z, Qiao C, et al. Synergistic enhancement of wearable biosensor through Pt single-atom catalyst for sweat analysis［J］. Biosensors and Bioelectronics, 2024, 258: 116354.
［20］ 李梦飞, 李桂新. N-(4-氨基丁基)-N-(乙基异鲁米诺)/血红素/金纳米三功能化氧化石墨烯材料的合成及性质研究［J］. 广东化工, 2018, 45(10): 26-29.
［21］ Ahmed S A, Roushani M, Karazan Z M. Development of a molecularly imprinted electrochemical sensor for erlotinib detection using a modified screen-printed carbon electrode［J］. Microchimica Acta, 2025, 192: 138.
［22］ Liu X, Wang Q, Zhang Y, et al. Detecting miRNA-21 using dual-mode biosensor via click chemistry-mediated variable temperature amplification reaction［J］. Chemical Engineering Journal, 2024, 479: 147231.
［23］ Fang C, Chen Z, Megharaj M, et al. Potentiometric detection of AFFFs based on MIP［J］. Environmental Technology & Innovation, 2016, 5: 52-59.

（上接第16页）

［15］ Cui J, Zhong Z, Liu S, et al. Parametric contrastive learning［C］//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, Montreal, QC, Canada. IEEE, 2021: 695-704.
［16］ Zhu J, Wang Z, Chen J, et al. Balanced contrastive learning for long-tailed visual recognition［C］//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, New Orleans, LA, USA. IEEE, 2022: 6898-6907.
［17］ Tang K, Huang J, Zhang H. Long-tailed classification by keeping the good and removing the bad momentum causal effect［EB/OL］. (2020-09-28)［2025-04-14］. https://arxiv.org/abs/2009.12991.
［18］ Kang B, Xie S, Rohrbach M, et al. Decoupling representation and classifier for long-tailed recognition［EB/OL］. (2019-10-21)［2025-04-14］. https://arxiv.org/abs/1910.09217.
［19］ Li Y, Wang T, Kang B, et al. Overcoming classifier imbalance for long-tail object detection with balanced group softmax［C］//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, Seattle, WA, USA. IEEE, 2020: 10988-10997.
［20］ Zhou B, Cui Q, Wei X S, et al. BBN: bilateral-branch network with cumulative learning for long-tailed visual recognition［C］//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, Seattle, WA, USA. IEEE, 2020: 9716-9725.
［21］ Zhang Y, Hooi B, Hong L, et al. Self-supervised aggregation of diverse experts for test-agnostic long-tailed recognition［EB/OL］. (2021-07-20)［2025-04-14］. https://arxiv.org/abs/2107.09249.
［22］ Cai Z, Vasconcelos N. Cascade R-CNN: high quality object detection and instance segmentation［J］. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(5): 1483-1498.
［23］ Liu Z, Mao H, Wu C Y, et al. A ConvNet for the 2020s［C］//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, New Orleans, LA, USA. IEEE, 2022: 11966-11976.
［24］ Liu Z, Lin Y, Cao Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows［C］//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, Montreal, QC, Canada. IEEE, 2021: 9992-10002.
［25］ Zheng L, Shen L, Tian L, et al. Scalable person re-identification: a benchmark［C］//2015 IEEE International Conference on Computer Vision (ICCV), 2015, Santiago, Chile. IEEE, 2015: 1116-1124.