数字中医中药

DOI:10.13193/j.issn.1673-7717.2026.08.003

基于耳穴分区的YOLOv12n－BiGS中医耳穴检测算法研究

夏书剑[1],周伟杰[1],姚伟涛[1],荆秦[1],周子艺[1],庞立健[2],秦钰莹[1],吕晓东[1],王琳琳[2]
(1.辽宁中医药大学,辽宁 沈阳 110847;2.辽宁中医药大学附属医院,辽宁 沈阳 110032)

摘要:中医耳穴作为人体脏腑经络的体表映射区域,其精准识别与定位是耳针诊疗、中医健康监测等场景的关键技术支撑。传统耳穴检测高度依赖专业医师的临床经验,不仅主观性强、效率低下,且难以满足标准化应用需求;现有目标检测算法在耳穴检测中普遍存在精度与速度难以兼顾的痛点。为此,提出一种基于YOLOv12的改进型中医耳穴检测算法。首先,结合耳穴国标与9种中医病症,构建包含31个耳穴分区的专用数据集,收录2140张涵盖不同年龄段、耳型特征及拍摄环境的耳穴图像。在模型优化方面,引入C3k2－GV3S轻量残差特征聚合模块(Lightweight residual feature aggregation module)替换原始主干网络的C3k2模块,借助"基础特征生成+廉价幻影特征补全"的双阶段策略,在保留残差特征复用优势的前提下,将参数量与计算量降低约50%,同时强化耳穴小目标的关键特征表征;设计双线性自适应边界增强上采样模块BABEU(Bilinear adaptive boundary－enhanced upsampling module)替代传统插值操作,通过"双线性插值平滑特征+可学习卷积细化边界"的组合策略,自适应学习耳穴边界的高频细节特征,避免像素块效应与棋盘格伪影,显著提升耳穴小目标的边界特征恢复精度。最终构建基于YOLOv12n的双线性自适应边界增强上采样与轻量残差特征聚合网络YOLOv12n－BiGS(Bilinear adaptive boundary－enhanced upsampling and lightweight residual feature aggregation network)。实验结果表明,该模型在专用数据集上的平均Dice系数达0.582 1,推理时间仅3.3 ms,综合性能优于YOLOv8n、YOLOv11n、DeepEar等主流轻量化算法。本研究构建的专用数据集贴合临床需求,改进模型实现了耳穴检测精度与实时性的协同优化,为中医耳穴诊疗的智能化、标准化推广提供技术支撑。

关键词:中医耳穴分区;深度学习;目标检测;YOLOv12n－BiGS;双线性自适应边界增强上采样模块BABEU;C3k2－GV3S

中图分类号:R245;TP391.41 **文献标志码**:A **文章编号**:1673-7717(2026)08-0012-10

Research on YOLOv12n－BiGS Traditional Chinese Medicine Auricular Point Detection Algorithm Based on Auricular Point Partition

XIA Shujian[1],ZHOU Weijie[1],YAO Weitao[1],JING Qin[1],ZHOU Ziyi[1],
PANG Lijian[2],QIN Yuying[1],LYU Xiaodong[1],WANG Linlin[2]
(1. Liaoning University of Traditional Chinese Medicine,Shenyang 110847,Liaoning,China;
2. Affiliated Hospital of Liaoning University of Traditional Chinese Medicine,Shenyang 110032,Liaoning,China)

Abstract:As the body surface mapping area of human organs and meridians,the accurate identification and location of auricular points in traditional Chinese medicine is the key technical support for auricular acupuncture diagnosis and treatment,traditional Chinese medicine health monitoring and other scenes. Traditional ear acupoint detection is highly dependent on the clinical experience of professional doctors,which is not only subjective and inefficient,but also difficult to meet the needs of standardized application. The existing target detection algorithms generally have pain points that are difficult to balance accuracy and speed in auricular point detection. To this end,this study proposes an improved traditional Chinese medicine auricular point detection algorithm based on YOLOv12. First of all,combined with the national standard of auricular points and 9 kinds of TCM diseases,a special data set containing 31 auricular points partitions was constructed,including 2140 auricular points images covering different age groups,ear shape characteristics and shooting environment. In terms of model optimization,the C3k2－GV3S Lightweight Residual Feature Aggregation Module is introduced to replace the C3k2 module of the original backbone network. With the help of the two－stage strategy of "basic feature generation+cheap phantom feature completion",the parameter quantity and calculation amount are reduced by about 50% on the premise of retaining the advantage of residual feature multiplexing,and the key feature

基金项目:国家自然科学基金面上项目(82274440);国家中医药管理局中医络病重点学科建设项目(T〔2023〕85);辽宁省科技计划联合计划项目(2023JH2/101700240);辽宁省中医药创新团队项目(LNZYYCXTD－CCCX－001);辽宁中医药大学人文社科类项目(2023LNZYQM004)

作者简介:夏书剑(1984－),女,辽宁沈阳人,讲师,硕士,研究方向:中医药人工智能与大数据分析。

通讯作者:王琳琳(1982－),女,辽宁盘锦人,主任医师,硕士研究生导师,硕士,研究方向:中医药治疗肺脾系病证。E-mail:wanglinlin820324@163.com。

庞立健(1982－),男,辽宁葫芦岛人,主任中医师,博士研究生导师,博士,研究方向:中医肺系疾病基础和临床、中医药人工智能。E-mail:hahaya123@126.com。

representation of small auricular targets is strengthened. The bilinear adaptive boundary – enhanced upsampling module BABEU (Bilinear Adaptive Boundary – Enhanced Upsampling Module) is designed to replace the traditional interpolation operation. Through the combination strategy of "bilinear interpolation smoothing feature + learnable convolution thinning boundary", the high – frequency detail features of the auricular point boundary are adaptively learned to avoid pixel block effect and checkerboard artifacts, and the boundary feature recovery accuracy of the auricular point small target is significantly improved. Finally, a bilinear adaptive boundary enhanced upsampling and lightweight residual feature aggregation network YOLOv12n – BiGS (Bilinear Adaptive Boundary – Enhanced Upsampling and Lightweight Residual Feature Aggregation Network) based on YOLOv12n has been constructed. The experimental results show that the average Dice coefficient of the model on the dedicated data set is 0.5821, and the reasoning time is only 3.3 ms. The comprehensive performance is better than the mainstream lightweight algorithms such as YOLOv8 n, YOLOv11 n and DeepEar. The special data set constructed in this study meets the clinical needs, and the improved model realizes the collaborative optimization of auricular point detection accuracy and real – time performance, which provides technical support for the intelligent and standardized promotion of auricular point diagnosis and treatment in traditional Chinese medicine.

Keywords: auricular point partition of traditional chinese medicine; deep learning; object detection; YOLOv12n – BiGS; bilinear adaptive boundary enhanced upsampling module BABEU; C3k2 – GV3S

中医耳穴疗法利用耳穴反射区进行诊疗，在临床多个领域都发挥了积极的辅助作用，例如利用耳穴按摩对急症高热患者的中医证候改善，利用耳穴贴压治疗原发性痛经[1]，耳穴压豆辅助治疗肾虚性腰肌劳损[2]，耳穴刮痧联合耳穴压豆治疗心脾两虚型失眠[3]及带状疱疹后遗神经痛[4]等，同时在保护视力[5]、调整人体生理功能、缓解与改善病情等方面也有显著疗效。

由于耳廓面积小、解剖结构复杂以及耳穴数量多、穴位分布密集等因素，耳穴的精确定位对医师的经验要求极高。传统耳穴定位与识别主要依赖医师的临床经验，通过肉眼观察耳部形态、触摸感知穴位反应点等方式完成，不仅耗费大量时间和精力，还存在较强的主观依赖性，定位结果易受医师经验水平、判断标准差异等因素影响，难以实现标准化、客观化与精准化定位。

为解决上述问题，图像识别、计算机视觉[6-7]等现代技术已逐步应用于中医耳穴检测领域，为耳穴的客观、高效识别提供了技术支撑。鉴于中医耳穴诊疗常采用耳穴贴压、压豆等临床方法，其疗效发挥均依赖对耳穴区域的精准定位，因此本研究依据中医脏腑关联理论，结合深度学习算法与计算机视觉技术，攻克耳穴区域自动检测的技术难点。通过智能化辅助诊疗提升耳穴区域检测的精度与效率，为临床耳穴疗法的规范化应用提供技术支持，从而精准、有效地改善人体健康状况。

1 中医脏腑经络理论

耳穴作为中医经络脏腑理论在耳部的体表映射，耳廓分区对应人体不同脏腑功能区域，其分区遵循明确的中医标准化规范。

中医遵循辨证取穴，目前已开展的相关研究都只是针对某一种疾病，这对耳穴疗法的应用意义体现得不够全面[8-10]，所以本研究基于耳穴与脏腑关联性，对肺系病症、脾胃系病症和肝胆系病症进行深入挖掘，从中选取常见的9种病症，包括：感冒、哮喘、咳嗽、呕吐、胃痛、泄泻、便秘、呃逆、胁痛。组织中医领域相关专家进行决策分析，结合辨证取穴，通过分型分证，使疾病在应用主穴的基础上可结合配穴进行诊疗，如："感冒"应用主穴肺、内鼻、气管、咽喉，其中发热可配耳尖、屏尖放血；前头痛可配额；偏头痛可配颞；后头痛可配枕。如表1所示，是9种病症采用耳穴诊疗时可关联的主穴和配穴。

按表1，结合9种病症，最终共确定了49个穴位作为研究范围。参照中医"耳穴名称与定位(GB/T13734 – 2008)"国家标准，将49个耳穴所对应的分区进行确认。例如：在"感冒"中如果是后头痛，在应用主穴治疗的基础上，可以搭配配穴"枕"进行辅助治疗。"枕"在"耳穴名称与定位"国家标准中对应的区域为"对耳屏三区"，再通过视觉检测技术对中医耳穴区域进行智能化识别和定位，从而提升中医耳穴辅助诊疗的精准性和规范化。

插页Ⅴ图1展示了研究范围为31个耳穴分区的耳朵标注示意图。插页Ⅴ图2展示了以"咳嗽"病症实例关联的耳穴分区。表2简列了5种穴位对应的分区以及从中医辨证取穴的角度所关联的中医病症。例如"肺"在"耳穴名称与定位"国家标准中对应的区域为"耳甲14区"，而肺在咳嗽、泄泻、便秘等中医病症中都对中医耳诊起到了辅助治疗的作用，比如肺可作为主穴应用于咳嗽中的外感咳嗽和内伤咳嗽。肺还可以作为配穴应用于泄泻中的寒湿者来辅助改善病症。

2 研究方法

2.1 *数据集的构建* 本研究所采用的耳朵图像样本采集于辽宁中医药大学及附属医院，通过高分辨率手机拍摄而成，确保了图像质量的清晰与一致性，为实验分析提供了可靠的数据基础。考虑到人耳形态的多样性，确保数据集的广泛性和代表性，图像数据是在不同环境、不同光线、不同角度的情况下，采集了从20～70岁的真实人耳图像，共获取2911张，去除无效图，构建了一个包含2140张图像的数据集。数据集中对每幅原始耳朵图像标注了31个耳穴分区。本研究联合辽宁中医药大学及附属医院开展，所有数据采集流程严格遵循世界医学会《赫尔辛基宣言》的人体受试者伦理准则，研究方案已通过辽宁中医药大学附属医院伦理委员会审查[批准文号：2024050XZ(KT)－016－03]。研究开展前已向所有志愿者告知研究目的、图像采集方式及数据使用范围，在充分获取志愿者知情、确认其自愿参与研究的前提下，采用手机拍摄志愿者双侧耳部图像，完成无创耳部图像数据采集。所有采集的耳部图像均进行脱敏处理，去除全部个人身份识别信息，严格保护受试者隐私安全，不用于其他研究及商业用途。随后由具备中医专业知识的技术人员借助Labelme工具开展耳穴分区的手工标注工作，同时通过人工微调的方式对模型标注数据进行修正，以保障数据标注质量；最终将所得数据集按照7∶1.5∶1.5的比例划分为训练集、验证集与测试集，整个研究过程严格遵循医学伦理规范。

2.2 *语义分割算法选择* 传统的耳穴检测算法的精度和速度

表1 病症-穴位关联表

病症		主穴	配穴
感冒		肺、内鼻、气管、咽喉	发热:耳尖、屏尖放血 前头痛:额 偏头痛:颞 后头痛:枕
哮喘	发作期	肺、气管、肾上腺、皮质下、内分泌、交感	
	缓解期	肺、脾、肾、交感	
咳嗽	外感咳嗽	气管、咽喉、肺、内鼻、肾上腺	发热时:耳尖、屏尖放血
	内伤咳嗽	脾、肺、气管、三焦、神门	痰多而黏(痰湿阻肺):脾、内分泌 干咳少痰(肺阴亏虚):肝、肾 急躁易怒(肝火犯肺):肝、内分泌、皮质下
胃痛	实证	脾、胃、神门、胰胆、食道	寒者(寒邪客胃):肾上腺、屏尖 胃脘胀痛(肝气犯胃):肝
	虚证	脾、胃、神门	胃脘隐隐灼痛(胃阴亏耗):肝、肾 胃脘隐痛,喜暖喜按(脾胃虚寒):肾上腺、屏尖
呕吐	实证	贲门、脾、胃、交感、肝	呕吐痰涎清水(痰饮内停):肾上腺、屏尖 呕吐酸腐(饮食停滞):食道、肝
	虚证	脾、胃、神门	饮食稍多即吐(脾胃虚寒):肾上腺、屏尖 反复发作,或时作干呕(胃阴亏虚):肝、肾
泄泻		脾、胃、小肠、结肠	寒湿者:肺、屏尖 湿热者:耳尖 食滞:三焦 肝脾不和:肝、交感 脾肾两虚:肾、肾上腺
便秘	实秘	胃、肺、脾、大肠	热秘:耳尖、三焦 冷秘:肾、肾上腺 气秘:肝
	虚秘	脾、大肠、肺	血虚:心
呃逆	实证	胃、神门、交感	寒者(胃寒积滞):屏尖 气滞者:肝 食滞者:脾、大肠
	虚证	脾、胃、神门、肾	
胁痛		胰胆、肝、三焦、内分泌	肝郁气滞:肝、交感 肝胆湿热:三焦、耳尖 瘀血阻络:内分泌 肝络失养:肾、肾上腺

表2 耳穴分区-穴位-病症关联表

穴位分区	穴位	中医病症
耳甲14区	肺	咳嗽、泄泻、便秘
耳甲16区	气管	咳嗽
耳甲4区	胃	呕吐、胃痛、泄泻、便秘、呃逆
耳甲7区	大肠	便秘、呃逆
三角窝4区	神门	咳嗽、呕吐、胃痛、呃逆

都很难达到实际应用需求,直到基于深度学习的目标检测与语义分割算法的出现,从语义分割领域的开创性模型开始,全卷积网络FCN(Fully convolutional network)首次实现端到端像素级分割,为耳穴分区这类细粒度分割任务提供了基础思路,但它对小目标边界的分割精度不足,且缺乏目标定位能力[11-12];DeepLabv3+通过空洞卷积与多尺度特征融合,大幅提升了语义分割的边界精度,理论上适配耳穴密集分区的分割需求,却因计算复杂度高导致实时性差,难以满足临床快速检测场景[13-14];掩膜区域卷积神经网络Mask R-CNN(Mask region-based convolutional neural network)则在快速区域卷积神经网络Faster R-CNN(Faster region-based convolutional neural network)基础上新增掩码分支,实现了“目标检测+实例分割”一体化,完美契合耳穴“定位+分区”的双重需求,但作为两阶段算法,其参数量大、推理效率低,无法适配实时检测需求和轻量化部署[15]。目标检测领域的演进同样围绕“精度与速度平衡”展开,从R-CNN[15]到Faster R-CNN[16],逐步提升了目标定位的效率,但仍属于多阶段算法,检测速度受限;而YOLO系列的出现,将多阶段目标检测算法转变为单阶段的目标检测算法,检测效率也大幅提升。

但目前针对耳穴的深度学习检测算法大多集中于耳穴特征点识别,例如王甘红等2025年提出的YOLOv8n关键点检测模型,在外部测试集上耳部识别准确率达0.991,50%交并比阈值0.5下的平均精度均值(Mean average precision at IoU threshold 0.5,mAP50)为0.990,具备良好的点定位性能,但未涉及耳穴分区任务[17]。耳穴分区检测方面的研究仍较少,高枕岳等人提出的K-YOLACT方法,通过“耳部轮廓检测-解剖部位分割-关键点估计”三阶段实现8个耳穴解剖部位分割,平均精度均值(Mean average precision,mAP)达83.2%,但未结合中医经络脏腑关联性,且多阶段流程导致推理效率受限[18]。何强等[19]发明的多层次卷积分割装置,可获得耳部结构与功能分区,但仅基于结构特征划分,未融合中医辨证理论,

也未实现精度与速度的协同优化。

YOLO系列算法[20]的演进逻辑始终围绕“精度与速度协同优化”展开,从v1的端到端框架突破,到v3～v5的多尺度检测与轻量化探索,再到v9～v12的注意力机制核心化、特征融合深度增强与动态推理策略升级,精度与速度呈持续上升趋势。尤其是2024—2025年发布的v9～v12,通过区域注意力模块(A2)、残差高效层聚合网络(Residual efficient layer aggregation network,R-ELAN)等创新,YOLOv12n更是在保持轻量化特性(参数量仅数M级别)的同时,mAP较YOLOv10n提升2.1%以上[20],依托高端GPU在定制场景下的推理速度可达450 FPS,且原生支持实例分割任务,完美适配耳穴“小目标、密集分区、实时检测”的核心需求。目前该算法已在医学影像小目标检测领域逐步显现技术潜力[21-24],为耳穴检测的高精度、实时性需求提供了坚实基础。因此,结合耳穴小目标、需实时检测的核心需求,本研究采用当前轻量化与精度平衡性能较优的YOLOv12算法开展中医耳穴区域检测任务。

2.3 YOLOv12目标检测算法 YOLOv12作为YOLO系列目标检测算法的最新升级版本[20],构建了以注意力机制为核心的全新网络框架,提供了从n到x的多尺度模型选型,能够满足高精度与实时性并存的目标检测需求。其中,YOLOv12n属于纳米级轻量化模型变体,核心优势体现在参数量精简、计算成本低廉及推理效率优异,同时延续了该系列算法高效的特征提取性能,尤其适配小目标检测及边缘端部署场景。YOLOv12系列(含YOLOv12n)的核心网络结构统一包含四大模块:数据输入层、Backbone主干特征提取网络、Neck多尺度特征融合层以及解耦式检测头,具体网络架构如插页Ⅵ图3所示。

数据输入层集成了多种训练阶段专属的数据增强手段,具体涵盖多尺度缩放、随机翻转、平移变换、马赛克拼接以及光照自适应调节等策略。通过丰富训练样本的分布多样性,可有效增强模型的泛化能力,提升其在复杂场景下的适配性能;需说明的是,推理阶段将禁用马赛克等强增强操作,以确保检测结果的稳定性。

Backbone主干网络摒弃了传统以CNN为核心的架构设计,采用残差高效层聚合网络(R-ELAN)作为核心特征提取单元,并嵌入区域注意力模块(A2)形成协同架构。R-ELAN借助块级残差连接与特征聚合机制,有效缓解了深层注意力网络中易出现的梯度阻塞问题,增强了模型训练的稳定性与特征表达的完整性;A2模块则通过将特征图划分为垂直或水平区域的轻量化处理方式,替代传统复杂的全局注意力计算,在保留较大感受野的前提下,将注意力计算复杂度由n2hd缩减至(1/2)n2hd,显著提升了特征提取的效率。

Neck多尺度特征融合结构:采用A2C2f模块堆叠与跨尺度特征拼接相结合的架构,通过上采样插值实现不同尺度特征图的对齐,配合FlashAttention技术优化特征长距离依赖匹配效率,实现浅层细节特征与深层语义特征的深度融合,有效强化小目标与密集目标的特征表征,提升检测精度。其中,A2C2f模块为A2注意力模块与C2f特征融合块的集成结构,兼顾注意力聚焦与特征聚合能力。

解耦式检测头:将分类任务与回归任务分离独立处理,通过3个不同尺度的检测层分别适配小、中、大尺寸目标,配合自适应锚框匹配策略提升特征响应的针对性。模型通过FlashAttention技术优化训练阶段的内存访问效率,减少冗余计算开销;前向推理阶段的高效性则主要依赖A2模块的轻量化设计与多层感知机(Multilayer perceptron,MLP)比例调整,最终实现高精度与高速度的协同提升。

损失函数采用解耦化设计思路,分别对分类损失、回归损失及置信度损失进行针对性优化:分类损失选用多标签二元交叉熵损失函数(Multi-label binary cross-entropy,ML-BCE);回归损失采用高效交并比损失函数(Efficient intersection over union loss,EIOU_Loss),该损失函数通过拆分纵横比影响因子与中心距离误差,可更精准地量化目标框的位置偏差;置信度损失采用二元交叉熵(Binary cross-entropy,BCE)损失函数,以提升目标存在性判断的可靠性。

YOLOv12系列嵌入了多项核心优化策略,能够很好地适配YOLOv12n的轻量化需求:A2模块不仅可降低计算开销,还能提升小目标检测的平均精确率(Average precision,AP)值;R-ELAN可改善深层注意力网络的梯度传播路径,解决大规模模型训练过程中的不稳定性问题;FlashAttention技术可优化GPU高速内存的利用效率,提升模型训练速度;将MLP通道扩展比例从4调整为1.2,实现了注意力机制与前馈网络计算量的平衡;通过移除传统位置编码,并引入7×7深度可分离卷积作为位置感知单元,进一步简化了网络结构,提升了整体计算效率。

3 YOLOv12n-BiGS目标检测算法

针对耳穴尺度小、分布密集且需平衡效率与精度的核心需求,原始YOLOv12n存在两大关键问题:一是主干网络C3k2模块参数量与计算量偏高,推理效率受限;二是传统最近邻插值上采样缺乏特征学习能力,耳穴边界特征恢复模糊。为此,本研究提出YOLOv12n-BiGS改进模型(BiGS为双线性自适应边界增强上采样模块BABEU与C3k2-GV3S轻量残差特征聚合模块的缩写组合),通过“轻量化特征提取+可学习上采样融合”双策略优化:引入C3k2-GV3S轻量残差特征聚合模块替换原始C3k2模块,实现轻量化特征提取;设计双线性自适应边界增强上采样模块(BABEU)替代传统插值操作,提升耳穴小目标边界特征恢复精度。改进模型在保持轻量化优势的同时,强化了耳穴小目标检测与分割性能。

3.1 基于C3k2-GV3S的轻量残差特征聚合模块设计 原始YOLOv12n主干网络的C3k2模块采用标准卷积构建特征提取单元[24],存在特征冗余计算问题,导致模型参数量偏大、推理速度难以满足耳诊的实时检测需求。C3k2模块作为C3模块的增强版本,通过改进核结构(K2)扩展感受野以提升小目标检测能力,但标准卷积的固有冗余仍限制其轻量化潜力。结合耳穴检测“轻量化+高精度”的双重需求,本研究引入GhostNetV3的核心轻量化思想[25],提出C3k2-GV3S轻量残差特征聚合模块,替换原始YOLOv12n中的C3k2模块,在保留跨阶段局部(Cross Stage Partial,CSP)结构特征复用优势的基础上,通过“基础特征生成+廉价幻影特征补全”的双阶段策略,大幅降低计算开销,同时强化耳穴小目标特征表征能力。

插页Ⅵ图4为C3k2-GV3S轻量残差特征聚合模块的网络结构图,其中该模块的核心是GhostBottleneckV3单元,其通过1×1卷积完成通道维度压缩与基础特征生成,再利用深度可分离卷积(Depthwise separable convolution,DWConv)生成廉价幻影特征,最终拼接基础特征与幻影特征得到完整输出,避免了标准卷积的冗余特征计算。为量化其轻量化效果,明确C3k2-GV3S与原始C3k2模块的参数量、计算量差异,基于标准卷积与GhostConvV3的核心计算逻辑[26],核心公式如公式(1)～(4)。

原始C3k2模块核心卷积(标准卷积)参数量公式:

$$Params_{std} = c_{in} \times c_{mid} \times k_1 \times k_1 + c_{mid} \times c_{out} \times k_2 \times k_2 \quad 式(1)$$

原始 C3k2 模块核心卷积计算量公式：

$$FLOPs_{std} = c_{in} \times c_{mid} \times k_1 \times k_1 \times h \times w + c_{mid} \times c_{out} \times k_2 \times k_2 \times h \times w \quad 式(2)$$

C3k2 – GV3S 模块核心卷积（GhostConvV3）参数量公式：

$$Params_{ghost} = c_{in} \times \frac{c_{mid}}{r} \times 1 \times 1 + \frac{c_{mid}(r-1)}{r} \times d \times d + \frac{c_{mid}}{r} \times c_{out} \times 1 \times 1 \quad 式(3)$$

C3k2 – GV3S 模块核心卷积计算量公式：

$$FLOPs_{ghost} = c_{in} \times \frac{c_{mid}}{r} \times 1 \times 1 \times h \times w + \frac{c_{mid}(r-1)}{r} \times d \times d \times h \times w + \frac{c_{mid}}{r} \times c_{out} \times 1 \times 1 \times h \times w \quad 式(4)$$

C_{in}代表耳穴特征图的通道维度即输入通道数，C_{mid}为中间特征图通道数（卷积层维度转换的过渡通道），C_{out}为输出特征图通道数，k_1、k_2为标准卷积核尺寸（原始 C3k2 模块中均取1 × 1 和 3 × 3），h、w 代表输入/输出特征图的高度与宽度，即耳穴特征图的空间维度，因卷积步长为 1，输入输出尺寸一致。r 为 GhostV3 的特征压缩比（本研究取 r = 4，采用激进压缩策略实现轻量化），d 为深度可分离卷积核尺寸（取 3 × 3，用于生成廉价幻影特征）。

C3k2 – GV3S 轻量残差特征聚合模块的整体结构延续 C3k2 的 CSP 架构，将输入特征图分为两部分：主分支通过 2 个串联的 GhostBottleneckV3 单元提取深层特征，残差分支直接传递浅层特征，最终通过 Concat 操作融合两分支特征，实现“轻量化提取 + 特征复用”的协同优化。结合公式（1）与（3）、（2）与（4）对比可知，当 r = 4 时，C3k2 – GV3S 模块的参数量与计算量可降低至原始 C3k2 模块的约 50%，在大幅减少冗余计算的同时，通过 GhostBottleneckV3 的特征生成策略保留耳穴小目标的关键纹理与边界特征，为实时检测奠定基础。

3.2 基于 BABEU 的双线性自适应边界增强上采样模块设计

原始 YOLOv12n 的 Neck 结构采用最近邻插值（nn. Upsample）完成上采样操作，该方法属于固定插值策略，仅通过像素复制提升特征图分辨率，缺乏特征学习能力，导致耳穴这类小目标的边界特征恢复模糊，分割边界精度偏低。双线性插值虽能通过像素距离加权获得更平滑的特征过渡，但同样无参数可学习，无法适配耳穴这类细粒度目标的特征分布。针对这一问题，本研究引入双线性自适应边界增强上采样模块（BABEU），替换原始上采样操作，通过“双线性插值 + 卷积细化”的组合策略[27]，实现耳穴特征图分辨率提升与边界特征自适应学习的协同，强化耳穴区域的分割边界质量。

插页Ⅶ图 5 为双线性自适应边界增强上采样模块（BABEU）的网络结构图，该模块的核心逻辑是：先通过双线性插值初步提升特征图分辨率，利用双线性插值的平滑特性保留耳穴特征的整体结构；再通过 3 × 3 卷积层对插值后的特征图进行细化，自适应学习耳穴边界的高频细节特征，同时引入批量归一化（Batch Normalization，BatchNorm）与 Sigmoid 线性单元（Sigmoid Linear Unit，SiLU）激活函数优化梯度传播，避免特征过拟合。其核心特征转换公式基于双线性插值原理与可学习卷积细化逻辑，具体如公式（5）～（7）。

双线性插值上采样特征映射公式：

$$F_{up}(x,y) = \sum_{i=0}^{1} \sum_{j=0}^{1} w_{i,j} \times F_{in}(x_i, y_i) \quad 式(5)$$

卷积细化特征公式：

$$F_{refine} = SiLu(BatchNorm(Conv3 \times 3(F_{up}))) \quad 式(6)$$

双线性自适应边界增强上采样模块（BABEU）最终输出公式：

$$F_{out} = F_{refine} \quad 式(7)$$

$F_{up}(x,y)$为双线性插值后（x，y）位置的特征值。$w_{i,j}$为双线性插值的权重系数，由（x，y）与原始特征图像素点的距离决定，确保距离越近的原始像素权重越大。$F_{in}(xi,yj)$原始输入特征图（上采样前）中邻近像素点的特征。Conv3 × 3（）为 3 × 3 标准卷积操作，用于细化插值后的特征，增强耳穴边界特征表达。BatchNorm（）为批量归一化操作，用于标准化特征分布，提升模型训练稳定性。SiLu（）为 Sigmoid Linear Unit 激活函数，用于引入非线性特征转换，强化耳穴特征的区分度。F_{refine}为卷积细化后的特征图。F_{out}为 BABEU 的最终输出特征图。

相较于原始最近邻插值，BABEU 具有三大优势：一是通过双线性插值的权重自适应特性，避免了最近邻插值的像素块效应，使耳穴特征过渡更平滑；二是通过可学习的 3 × 3 卷积层，能够自适应学习耳穴边界的高频细节，解决了传统插值“只升分辨率、不增特征”的缺陷[28]；三是完全避免了转置卷积常见的棋盘格伪影，确保耳穴分割边界更精准。

综上，YOLOv12n – BiGS 模型通过 C3k2 – GV3S 轻量残差特征聚合模块实现主干网络轻量化，通过双线性自适应边界增强上采样模块（BABEU）提升耳穴边界特征恢复精度，在保持检测头结构不变的前提下，实现了“轻量化部署”与“高精度检测”的协同优化，适配耳穴检测的实际应用需求。

4 实验分析

4.1 实验设置 本实验基于 Ultralytics YOLOv12 框架对 YOLOv12n 模型进行改进并开展耳穴检测任务，软件环境为 Ubuntu 18.04 操作系统，Python 3.11，PyTorch 2.5.1；并行计算平台统一计算设备架构（Compute unified device architecture，CUDA）12.8，CUDA 深度神经网络库（CUDA deep neural network library，cuDNN）8.9.7；标注工具使用 LabelMe5.5.0。硬件环境配置为 Intel（R）Xeon（R）Platinum 8352V 中央处理器（CPU，2.10 GHz），搭载 32 GB 显存的图形处理器（GPU），内存 90 GB。

模型核心配置将检测类别设置为 31 类，以匹配耳穴分类需求，训练超参数设置：优化器选用 AdamW 自适应权重衰减优化器（Adam weight decay，AdamW），初始学习率 lr = 0.002，动量因子 momentum = 0.9，权重衰减系数 weight_decay = 0.000 468 75；最大训练迭代轮次设为 1000 轮（epoch），早停机制容忍度（patience）设置为 100。AutoBatch 自适应批次自动调整批量大小为 20；训练前期迭代轮次采用线性热身（warmup）策略，后续采用余弦退火策略完成学习率衰减；数据加载采用 8 线程数据加载器（dataloader）提升训练效率。

4.2 评价指标 为综合评估模型在实例分割任务中的性能，本文采用的评估指标包括：参数量、计算量、平均 Dice 系数、平均 95% 豪斯多夫距离 meanHD95，以及推理时间 Inference Time。各项指标的计算方法如公式（8）～（10）所示。

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad 式(8)$$

Dice 系数（Dice Coefficient），亦常被称为 F1 – Score 或重叠指数，是医学图像分割等领域中衡量预测分割掩码与真实标签（Ground Truth）之间相似度的核心指标。其中 X 与 Y 分别代表预测区域与真实区域的像素集合。该指标的本质是计算两个集合交集大小的两倍与其总像素数之和的比值，其取值范围在［0，1］之间，值越高表明两者的空间重叠度与一致性

越好。

$$Mean\ Dice = \frac{1}{C}\sum_{c=1}^{c}\left(\frac{1}{N_c}\sum_{i=1}^{N_c}\frac{2\times TP_{c,i}}{2\times TP_{c_i}+FP_{c_i}+FN_{c_i}}\right) \quad 式(9)$$

平均Dice系数则为各类别单独计算出来的Dice系数的算术平均值。参数量是指神经网络中所有可学习参数的总数量，反映了模型的记忆容量和复杂度。计算量通常指浮点运算次数(Floating Point Operations，FLOPs)，衡量模型的计算复杂度。

$$meanHD95 = \frac{1}{C}\sum_{c=1}^{c}\frac{1}{N_c}\sum_{i=1}^{N_c}HD95_{c,i} \quad 式(10)$$

平均95%豪斯多夫距离(meanHD95)衡量的是预测分割边界与真实分割边界之间的平均最大偏差，但通过取95%分位数的方式，降低了极端异常值的影响，使其更加稳健。HD95值越小，说明预测的轮廓与实际标注的真实轮廓越贴合，分割的边界越准确。推理时间(Inference time)是指训练好的机器学习/深度学习模型在部署阶段处理输入数据并生成预测结果所需的时间。这是衡量模型计算效率和实时性的关键指标。

4.3 消融实验分析 从消融实验对比表3中可以看出，基础模型中引入C3k2－GV3S轻量化卷积后组成的算法(YOLOv12＋C3k2－GV3S)相较于YOLOv12n，算法参数量降低14 324、计算量减少0.2 GFLOPS，推理时间缩短4.2 ms，但Mean Dice略有下降，表明该模块可有效实现轻量化与提速，但对分割性能无增益。基础模型中引入双线性自适应边界增强上采样模块(BABEU)组成的算法(YOLOv12＋BABEU)相较于YOLOv12n，参数量提升738 048，Mean Dice提升1.36%，但推理时间延长3.2 ms，表明BABEU的引入虽使参数量增加、速度下降，但能显著提升模型对耳穴区域的分割能力。引入两种模块叠加组成的YOLOv12n－BiGS算法，相较于YOLOv12＋C3k2－GV3S，Mean Dice提升3.93%，meanHD95降至70.344 6 px(边界分割精度提升)；相较于YOLOv12＋BABEU，推理时间下降6.2 ms，参数量降低15 324；相较于YOLOv12n，参数量增加723 724，Mean Dice提升3.86%，推理时间缩短3.0 ms，meanHD95从93.624 7 px降至70.344 6 px(边界分割精度显著提升)。综上分析可知，本研究提出的YOLOv12n－BiGS模型虽参数量略有增加，但实现了分割能力(Mean Dice提升、边界精度优化)与推理效率的协同提升，验证了双模块叠加优化的有效性。

表3 消融实验算法对比

算法名称	参数量	GFLOPS	Mean Dice	meanHD95/(px)	inference time/(ms)
YOLOv12n	2 815 189	10.2	0.543 5	93.624 7	6.3
YOLOv12＋C3k2－GV3S	2 800 865	10.0	0.542 8	103.158 2	2.1
YOLOv12＋BABEU	3 553 237	14.0	0.557 1	106.439 9	9.5
YOLOv12n－BiGS	3 538 913	13.7	0.582 1	70.344 6	3.3

4.4 对比实验分析 表4为YOLOv12n－BiGS算法与其他主流轻量化目标检测算法的对比实验数据。本研究选取的对比算法包括YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n、YOLOv11n及基础模型YOLOv12n，从检测精度、分割性能、模型效率3个维度开展对比验证[23]。从表3数据可以看出，YOLOv12n－BiGS算法在分割任务核心指标Mean Dice上达到0.582 1，显著高于其他对比算法(YOLOv5n 0.519 3、YOLOv8n 0.570 1、YOLOv12n 0.543 5)；同时其meanHD95低至70.344 6 px，相较于基础模型YOLOv12n(93.624 7 px)及性能较优的YOLOv8n(89.455 7 px)均大幅降低，表明模型对耳穴区域的边界分割精度显著提升。在模型效率方面，YOLOv12n－BiGS的参数量约为3.54 M，计算复杂度(GFLOPS)为13.7 G，相较于YOLOv6n(4.41 M、15.2 G)实现了轻量化优化；推理时间为3.3 ms，相较于基础模型YOLOv12n(6.3 ms)缩短3.0 ms，虽长于YOLOv5n(1.4 ms)、YOLOv8n(1.1 ms)等超轻量化模型，但仍满足实时检测需求。综上，YOLOv12n－BiGS算法在分割精度(Mean Dice最高、meanHD95最低)上显著领先于主流轻量化算法，同时在模型效率上具备竞争力，验证了本研究提出的双模块改进策略的有效性，适配耳穴小目标精准检测的需求。

为验证本研究提出的YOLOv12n－BiGS算法在中医耳穴分区检测任务中的性能优势，选取耳穴分割领域经典算法U－Net、SegNet、AttentionUnet、专用耳穴分割算法DeepEar[29]及主流医学分割算法DeepLabv3＋进行对比分析，结果如表5所示。

表4 轻量化目标检测算法对比实验

算法名称	参数量	GFLOPS	Mean Dice	meanHD95/(px)	inference time/(ms)
YOLOv5n	2 761 405	11.0	0.519 3	134.083 4	1.4
YOLOv6n	4 406 285	15.2	0.548 9	172.491 1	1.8
YOLOv8n	3 264 109	12.0	0.570 1	89.455 7	1.1
YOLOv10n	2 844 909	11.7	0.568 0	95.290 5	3.9
YOLOv11n	2 840 613	10.2	0.540 9	113.473 4	1.5
YOLOv12n	2 815 189	10.2	0.543 5	93.624 7	6.3
YOLOv12n－BiGS	3 538 913	13.7	0.582 1	70.344 6	3.3

表5 相关研究检测算法对比实验

算法名称	参数量	GFLOPS	Mean Dice	inference time/(ms)
U－Net	31 039 583	302.63	0.438 3	45.07
SegNet	29 460 957	257.90	0.410 9	51.85
AttentionUnet	96 930 532	1 060.22	0.536 5	1 536.00
DeepEar	5 228 287	26.58	0.559 6	24.05
deeplabv3＋	5 228 544	26.59	0.570 6	19.47
YOLOv12n－BiGS	3 538 913	13.70	0.582 1	3.30

从模型轻量化与计算效率来看，传统分割算法U－Net、SegNet、AttentionUnet参数量与计算量庞大，AttentionUnet参数量达96.93 M、计算量1 060.22 GFLOPS，推理时间长达1536 ms；U－Net、SegNet计算量也分别达302.63 GFLOPS、257.90 GFLOPS，推理时间均超45 ms，难以满足临床实时性需求。DeepEar与DeepLabv3＋虽实现轻量化，计算量仍为26.58 GFLOPS、26.59 GFLOPS，推理时间分别为24.05 ms、19.47 ms。本算法参数量仅3.54 M，计算量13.7 GFLOPS，推理时间压缩至3.3 ms，在轻量化与实时性上均显著优于对比方法。从分割精度来看，传统算法未针对耳穴小目标、边界模糊的特性优化，Mean Dice系数最高仅0.536 5；DeepEar、DeepLabv3＋虽将Mean Dice分别提升至0.559 6、0.570 6，但未针对中医31个标准耳穴分区做精细化设计。本算法取得0.582 1的Mean Dice系数，为所有对比方法中最优，充分验证了其在耳穴精细化分区检测的性能优势。而且现有耳穴检测研究多聚焦于耳廓整体分割或解剖结构划分，未结合中医耳穴国标规范，无法支撑临床辨证取穴需求。本算法基于中医耳穴分区标准，检测结果可直接对接耳针诊疗、耳穴贴压等临床操作，具备独特的临床落地优势。

综上，本算法在轻量化、实时性、分割精度及临床适配性上均优于现有耳穴检测相关研究，实现了精度与效率的协同

优化。

4.5 中医耳穴分区检测效果分析 本研究针对31个耳穴分区的检测效果如插页Ⅶ图6所示，为两组原始标注图像与yolo12n算法以及YOLOv12n－BiGS算法检测的效果对比图，可见从耳穴分区的区域形态、区域边缘精度以及平滑度上看，本文提出的YOLOv12n－BiGS算法更具优势。所以在YOLOv12n－BiGS比yolov12n的推理时间显著缩短的前提下，YOLOv12n－BiGS又能保持高精度的分割检测效果，实现了精度与效率的平衡优化，适配中医耳穴临床应用的大部分场景。

插页Ⅶ图6中的Class 0～Class 30分别为：耳甲14区、耳甲13区、耳甲15区、耳甲16区、耳甲17区、耳甲18区、三角窝2区、三角窝3区、三角窝4区、耳轮、耳舟、对耳轮11区、对耳轮9区、耳甲1区、耳甲2区、耳甲3区、耳甲4区、耳甲5区、耳甲6区、耳甲7区、耳甲9区、耳甲10区、耳甲11区、耳甲12区、耳轮1区、耳轮2区、对耳轮8区、对耳屏1区、对耳屏2区、对耳屏3区、耳垂。

5 结语

本研究针对传统耳穴定位主观性强、效率低的痛点，围绕耳穴分区自动化精准检测需求，开展了数据集构建与模型优化研究。研究结合中医国标规范构建含2140张图像的专用耳穴数据集，提出YOLOv12n－BiGS改进算法：通过C3k2－GV3S轻量残差特征聚合模块实现主干网络轻量化（参数量与计算量降低约50%），采用双线性自适应边界增强上采样模块（BABEU）提升耳穴边界特征恢复精度。实验验证表明，改进算法Mean Dice系数达0.582 1，推理时间3.3 ms，优于主流轻量化算法，实现精度与效率平衡。本研究创新点在于构建贴合临床的耳穴数据集，提出轻量化模型，突破小目标检测瓶颈，推动中医诊疗智能化。未来将扩大数据集规模、优化网络结构，开发便携式检测终端，促进成果临床转化。

参考文献

[1] 林莺，柳伟婷，张亮平，等. 耳穴周期贴压治疗原发性痛经41例临床观察[J]. 甘肃中医药大学学报，2025，42(3)：49－53.
[2] 袁铭，潘露. 金匮肾气丸与耳穴压豆疗法治疗肾虚性腰肌劳损的临床研究[J]. 内蒙古中医药，2025，44(7)：40－42.
[3] 靖瓒，吕若琳，魏晋晶. 基于经络检测指导耳部刮痧联合耳穴压豆治疗心脾两虚型失眠患者的临床研究[J]. 中国医药科学，2024，14(1)：103－107.
[4] 刘晓博，马迎存，杨彬. 芍药甘草汤加味联合耳穴压豆治疗带状疱疹后遗神经痛的效果[J]. 中国老年学杂志，2025，45(5)：1117－1121.
[5] 付晋，吕婧，姚靖，等. 耳穴压丸防控儿童青少年近视的临床观察[J]. 中国中医眼科杂志，2025，35(9)：840－845.
[6] 董文轩，梁宏涛，刘国柱，等. 深度卷积应用于目标检测算法综述[J]. 计算机科学与探索，2022，16(5)：1025－1042.
[7] 郭庆梅，刘宁波，王中训，等. 基于深度学习的目标检测算法综述[J]. 探测与控制学报，2023，45(6)：10－20，26.
[8] 王诗怡，于淼，王梓，等. 基于数据挖掘的耳穴贴压治疗肺系疾病的辨证选穴规律研究[J]. 湖南中医杂志，2025，41(8)：49－54.
[9] 常丹阳，方桂珍. 基于数据挖掘的耳穴贴压治疗变应性鼻炎的选穴规律分析[J]. 循证护理，2024，10(17)：3212－3216.
[10] 倪斐琳，陈改平，黄馨睿，等. 基于数据挖掘探析耳穴压豆治疗脑卒中后吞咽障碍的选穴规律[J]. 护理研究，2025，39(9)：1470－1475.
[11] SHELHAMER E，LONG J，DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell，2017，39(4)：640－651.
[12] BERTASIUS G，SHI J B，TORRESANI L. Semantic segmentation with boundary neural fields[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27－30，2016. Las Vegas，NV，USA. IEEE，2016：3602－3610.
[13] 柯寅，陈丹. 基于改进DeepLabv3＋网络的图像语义分割方法[J]. 智能计算机与应用，2025，15(4)：17－24.
[14] 米吾尔依提·海拉提，热娜古丽·艾合麦提尼亚孜，李莉，等. 基于改进DeepLabV^{3+}的囊型肝包虫病超声图像分割算法[J]. 中国医学物理学杂志，2024，41(6)：702－709.
[15] HE K M，GKIOXARI G，DOLLAR P，et al. Mask R－CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22－29，2017. Venice. IEEE，2017：2980－2988.
[16] REN S，HE K，GIRSHICK R，et al. Faster R－CNN：Towards Real－Time Object Detection with Region Proposal Networks[J]. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015：91－99.
[17] 王甘红，张子豪，夏开建，等. 基于YOLO神经网络构建耳穴特征点辅助检测的人工智能辅助系统[J]. 中国针灸，2025，45(4)：413－420.
[18] 高枕岳，贾世瑾，李青峰，等. 基于深度学习的耳部穴区自动分割方法[J]. 生物医学工程学杂志，2024，41(1)：114－120.
[19] 何强，郑毅，赵子豪，等. 耳部穴位反射区识别方法、装置、针灸机器人及存储介质：CN119762483B[P]. 2025－05－09.
[20] SAPKOTA R，FLORES－CALERO M，QURESHI R，et al. YOLO advances to its genesis：a decadal and comprehensive review of the You Only Look Once (YOLO) series[J]. Artif Intell Rev，2025，58(9)：274.
[21] QIU H F，HE Y. X－YOLO：a method for detecting wrist fractures in children based on dynamic feature enhancement and lightweight design[J]. Biomed Signal Process Control，2026，112：108874.
[22] YOU Z P，SUN K X，ZHANG L X，et al. SKF－YOLO：an enhanced method for real－time blood cell detection[J]. Biomed Signal Process Control，2026，112：108646.
[23] MEHTA P，VAGHELA R，PANSURIYA N，et al. Benchmarking YOLO variants for enhanced blood cell detection[J]. Int J Imaging Syst Tech，2025，35：e70037.
[24] ZOPH B，VASUDEVAN V，SHLENS J，et al. Learning transferable architectures for scalable image recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18－23，2018. Salt Lake City，UT. IEEE，2018：8697－8710.
[25] REDMON J，FARHADI A. YOLOv3：an incremental improvement[EB/OL]. 2018：arXiv：1804.02767. https://arxiv.org/abs/1804.02767
[26] LIU Z H，HAO Z W，HAN K，et al. GhostNetV3：exploring the training strategies for compact models[EB/OL]. 2024：arXiv：2404.11202. https://arxiv.org/abs/2404.11202
[27] SHI W Z，CABALLERO J，HUSZÁR F，et al. Real－time single image and video super－resolution using an efficient sub－pixel convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27－30，2016，Las Vegas，NV，USA. IEEE，2016：1874－1883.
[28] REDMON J，FARHADI A. YOLOv3：an incremental improvement[EB/OL]. 2018：arXiv：1804.02767. https://arxiv.org/abs/1804.02767
[29] CHEN Y H，KE W D，LI Q F，et al. DeepEar：a deep convolutional network without deformation for ear segmentation[J]. J Image Graph，2023，11(3)：242－247.

基于耳穴分区的 YOLOv12n – BiGS 中医耳穴检测算法研究

（正文见 12 – 18 页）

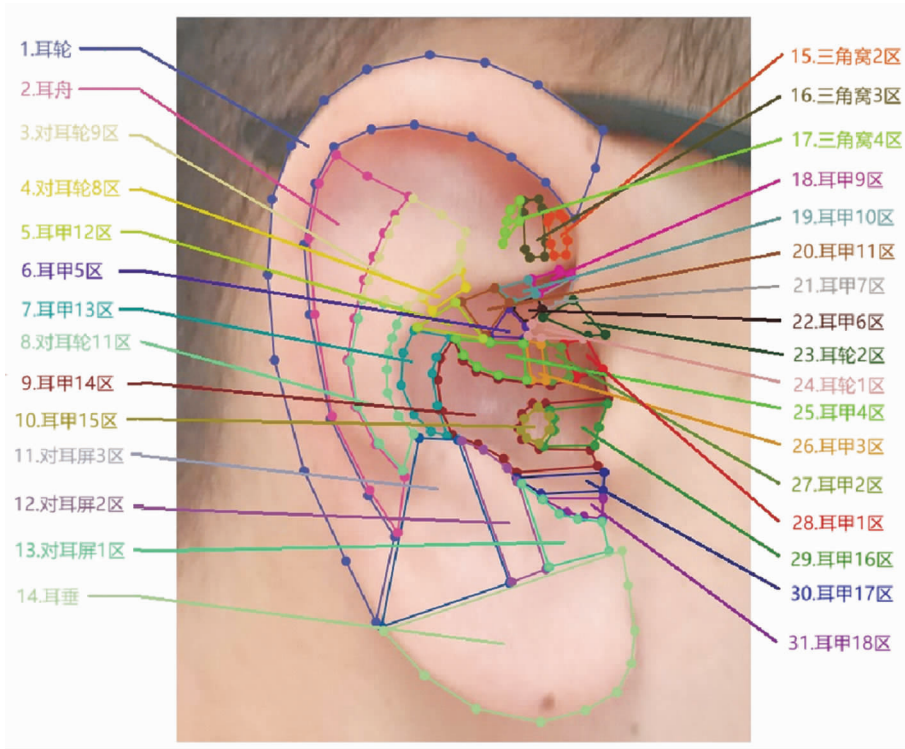


图 1　中医耳穴分区标注示意图

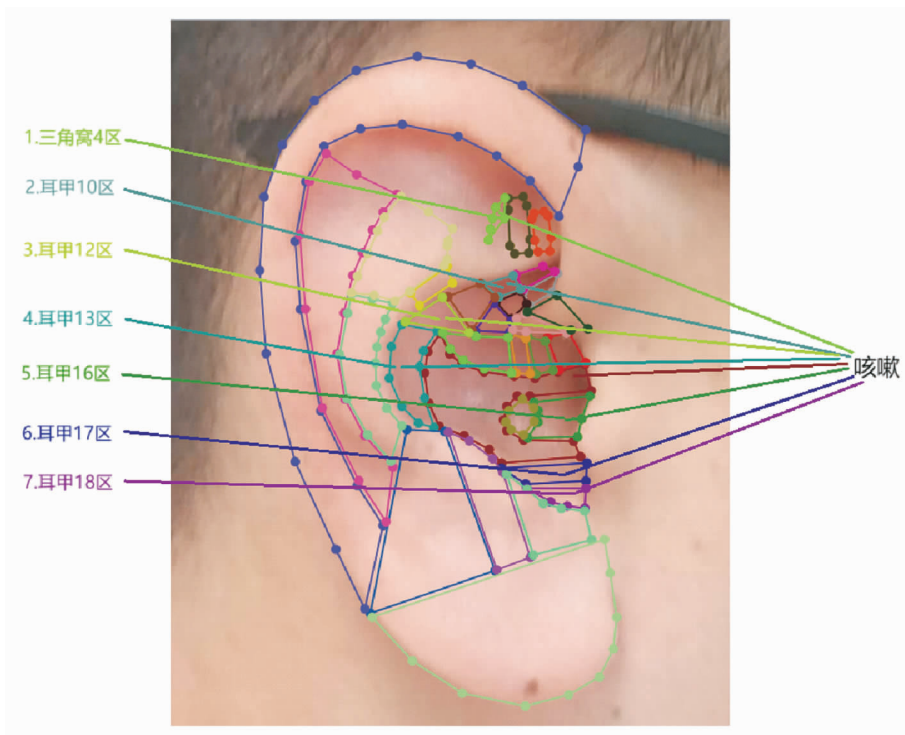


图 2　咳嗽病症关联耳穴分区的映射关系图

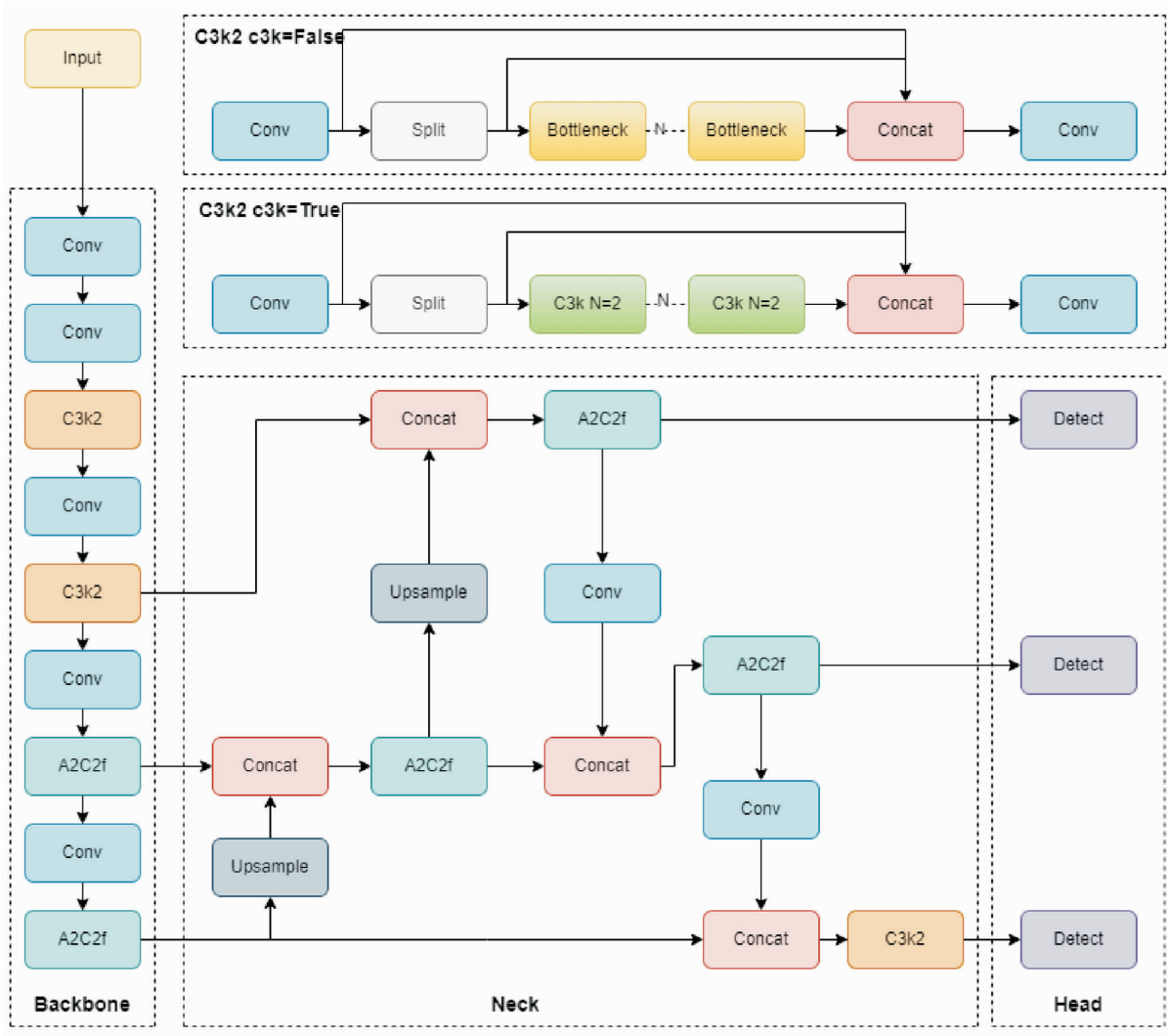


图 3 YOLOv12 网络结构图

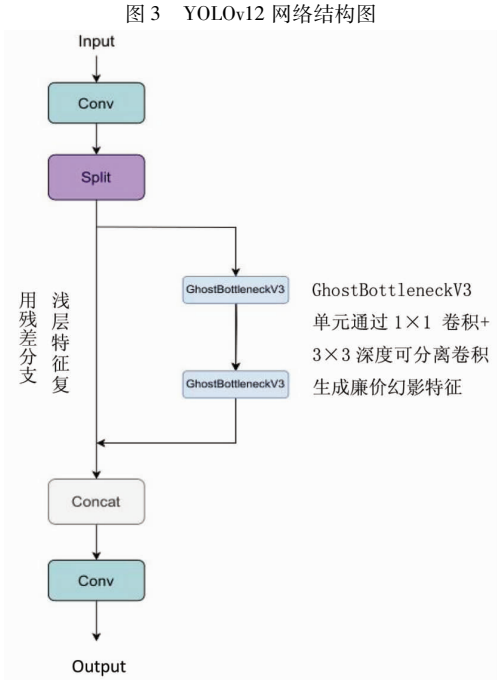


图 4 C3k2 – GV3S 轻量残差特征聚合模块结构图

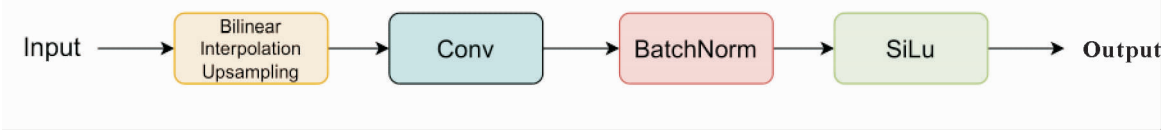


图 5　双线性自适应边界增强上采样模块(BABEU)结构图

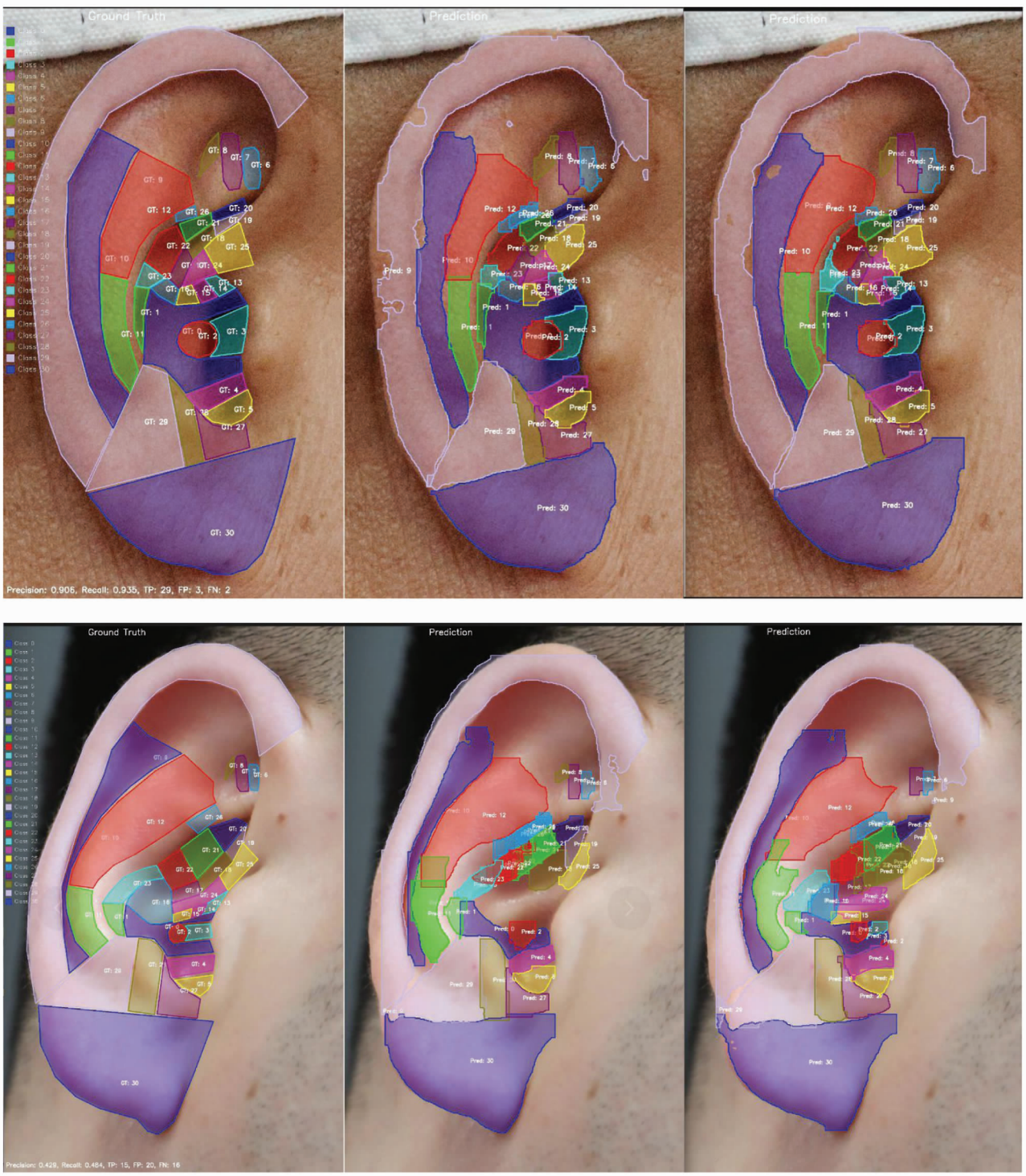


图 6　原始标注图像与 yolo12n 算法以及 YOLOv12n – BiGS 算法检测效果对比(两组)