

基于黄河水利系统网络安全场景下 YOLOv8 人耳识别方法改进的研究

杨子

河南省鼎信信息安全等级测评有限公司, 河南 郑州 450000

DOI: 10.61369/SSSD.2025090044

摘 要 : 随着网络信息技术的飞速发展, 网络空间安全威胁日益严峻, 身份认证与访问控制作为网络安全的核心环节, 其重要性不断凸显。黄河水利系统作为国家关键基础设施, 其网络信息安全尤为关键。本文在现有网络安全防控技术基础上, 提出一种基于 YOLOv8 的人耳识别方法, 以增强黄河水利系统的身份认证安全。人耳生物特征具有唯一性、稳定性和防伪性, 能有效提升认证的可靠性与安全性^[1-3]。本文针对人耳图像特点, 对 YOLOv8 模型进行了两方面的改进: 一是在中小尺度检测头前引入新型注意力模块 SaE, 通过多分支全连接结构增强对小目标的特征表征能力; 二是将骨干网络中的 C2f 模块替换为基于 Transformer 的 C2f_T 结构, 以增强模型对长程依赖的感知能力, 提升对耳部与背景混淆情况的区分能力。实验结果表明, 改进后的模型在遮挡、光照变化及姿态多样性等复杂场景下均表现出较高的鲁棒性, mAP50 达到 97.3%, 显著优于原模型及其他对比模型。本方法可应用于水利关键区域人员身份验证、视频监控中的异常人员检测及网络终端生物认证加固等多个场景^[4], 为提升黄河水利网络信息安全提供了有效的技术支撑。

关 键 词 : 人耳识别; YOLOv8; 黄河水利; 网络安全; 身份认证

Research on Improvement of YOLOv8 Ear Recognition Method Under Network Security Scenario of Yellow River Water Conservancy System

Yang Zi

Henan Dingxin Information Security Level Evaluation Co., Ltd., Zhengzhou, Henan 450000

Abstract : With the rapid development of network information technology, cyber space security threats have become increasingly severe. As the core links of network security, identity authentication and access control have become more and more prominent in their importance. The Yellow River Water Conservancy System, as a national key infrastructure, has particularly critical network information security. Based on the existing network security prevention and control technologies, this paper proposes an ear recognition method based on YOLOv8 to enhance the identity authentication security of the Yellow River Water Conservancy System. Ear biometrics possess uniqueness, stability and anti-counterfeiting properties, which can effectively improve the reliability and security of authentication [1-3]. Aiming at the characteristics of ear images, this paper makes two improvements to the YOLOv8 model: first, a new attention module (SaE) is introduced before the small and medium-scale detection heads, and the feature representation ability for small targets is enhanced through a multi-branch fully connected structure; second, the C2f module in the backbone network is replaced with a Transformer-based C2f_T structure to enhance the model's ability to perceive long-range dependencies and improve the ability to distinguish between ear and background confusion. Experimental results show that the improved model exhibits high robustness in complex scenarios such as occlusion, light changes and posture diversity, with mAP50 reaching 97.3%, which is significantly better than the original model and other comparison models. This method can be applied to multiple scenarios such as personnel identity verification in key water conservancy areas, abnormal personnel detection in video surveillance, and biological authentication reinforcement of network terminals [4], providing effective technical support for improving the network information security of the Yellow River Water Conservancy System.

Keywords : ear recognition; YOLOv8; yellow river water conservancy; network security; identity authentication

引言

黄河水利工程在国家水资源管理、防洪调度及生态保护中具有重要地位。随着数字化进程的加速，其网络信息系统规模不断扩大，数据量急剧增长，系统安全面临严峻挑战。根据《2023中国水利发展统计公报》，黄河流域已部署智能传感设备32.7万台，水利专用网络日均传输数据量达1.8PB。国家互联网应急中心监测显示，2022年水利工控系统遭受的高级持续性威胁攻击同比上升41%，其中身份伪造导致的未授权访问占比达37%。《网络安全法》与《关键信息基础设施安全保护条例》对水利行业提出了更严格的身份认证要求，传统“用户名+密码”方式已难以满足零信任安全架构的需求。

人耳作为一种生物特征，具有稳定性高、不易伪造、受表情和遮挡影响小等优势，非常适用于复杂环境下的身份认证。近年来，基于深度学习的人耳检测与识别方法取得了显著进展^[5]。例如，Bizjak 等人采用 Mask R-CNN 实现高精度人耳分割；Yuan 和 Lu 利用 YOLOv2-tiny 实现实时检测；Kamboj 等人提出上下文感知检测网络 CED-Net；Nguyen 则结合 YOLOv3 与 RetinaFace 实现联合检测^[6]。在识别方面，钱昱来等人基于改进 ICP 算法在三维人耳识别中取得98.55%的 Rank-one 识别率；杜保伟利用 SSD 与 MobileNet 结合实现98%的识别率^[7]；钱俊如基于 YOLOv3 在自建动态人耳库上达到90%以上识别精。

尽管已有成果显著，但现有方法在复杂场景下仍存在检测精度不足、鲁棒性差等问题。本文基于 YOLOv8 模型^[8]，结合注意力机制与 Transformer 结构，提出一种改进的人耳识别方法，旨在提升在黄河水利实际应用场景下的性能与可靠性。

一、改进的 YOLOv8 模型框架

YOLOv8 是一种高效的检测架构，支持分类、检测、分割及姿态估计等多种任务。其结构包括 Backbone、Neck 和 Head 三部分，具有多尺度检测能力，适用于不同尺寸目标。本文针对人耳目标的特性，重点在以下两方面进行改进：

（一）C2f_T 模块

为增强模型对长程依赖的建模能力，将原 C2f 模块替换为基于 Contextual Transformer (CoT) 的 C2f_T 模块。CoT 通过融合局部上下文与自注意力机制，在提取特征的同时保留丰富的语义信息。具体地，CoT 首先通过3×3卷积提取静态上下文特征，再通过多头注意力机制聚合动态上下文，最终将二者融合输出。该结构显著提升了模型对耳部轮廓与复杂背景的区分能力。

（二）SaE 模块

针对人耳多为中小尺寸目标的特点，在中小尺度检测头前引入 SaE (Squeeze Aggregated Excitation) 模块^[9]。该模块通过四个并行全连接分支进行多粒度特征提取，并利用前馈网络计算注意力权重，增强对细节特征的感知能力。其计算过程可表示为：

$$SaE=x+F(x \cdot Ex(\sum Sq(x)))$$

其中，Sq 为压缩操作，Ex 为激励操作，F 为特征融合函数。

改进后的 YOLOv8 模型结构如图1所示，在保持原有多尺度检测能力的基础上，显著提升了对小目标的检测精度与鲁棒性。

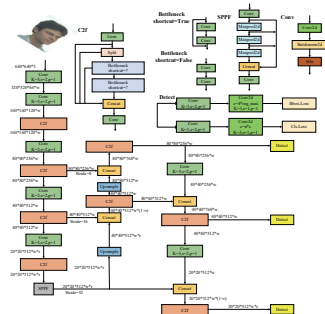


图1 YOLOv8网络框架图

二、实验结果与分析

（一）实验环境与数据

实验硬件采用 Intel Xeon Gold 8280 CPU 与4×NVIDIA RTX3090 GPU，软件环境为 PyTorch 框架与 Linux 系统。数据集包含7254张人耳图像，涵盖不同种族、年龄、性别及遮挡条件，数据示例如图2所示。

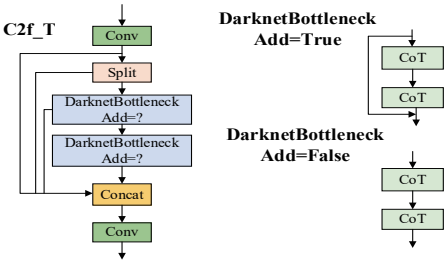


图2 C2f_T 结构图

为评估模型抗干扰能力，对测试集添加随机噪声、模糊、遮挡、光照和对比度变化，构建干扰测试集（图3）。

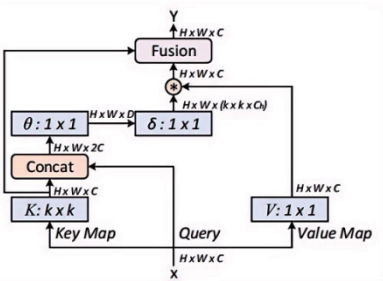


图3 CoT 模块

（二）评价指标

采用精确率 (Precision, P)、召回率 (Recall, R)、平均精度 (AP) 和平均精度均值 (mAP) 作为评价指标，计算公式如下：

$$PRAPmAP=TP+FTP=TP+FNTP=\int_0^1 P(R) dR=N\ln=1\sum NAPn$$

(三) 消融实验

在干扰数据上对 YOLOv8、YOLOv8+SaE、YOLOv8+C2f_T 及 YOLOv8+SaE+C2f_T 进行对比实验，结果如表 1 所示。

表 1 消融实验结果

方法	Precision	Recall	mAP50	mAP50-95
YOLOv8 (原始数据)	0.993	0.995	0.995	0.858
YOLOv8 (干扰数据)	0.973	0.804	0.889	0.748
+ SaE	0.980	0.921	0.967	0.765
+ C2f_T	0.968	0.910	0.954	0.766
+ SaE + C2f_T	0.983	0.946	0.973	0.789

实验表明，SaE 模块显著提升了小目标检测的召回率，C2f_T 增强了对复杂背景的区分能力，二者结合进一步提升了综合性能。

(四) 对比实验

将改进模型与 RT-DETR、YOLOv10、YOLOv11 在相同干扰数据下对比，结果如表 2 所示。

表 2 对比实验结果

方法	Precision	Recall	mAP50	mAP50-95
RT-DETR	0.965	0.844	0.884	0.676
YOLOv10	0.992	0.781	0.880	0.733
YOLOv11	0.982	0.797	0.883	0.732
YOLOv8+SaE+C2f_T	0.983	0.946	0.973	0.789

改进模型在召回率与 mAP 指标上均优于对比模型，尤其在复杂环境下保持较高鲁棒性。

三、结论

本文针对黄河水利网络信息安全需求，提出了一种基于 YOLOv8 的改进人耳识别方法。通过引入 C2f_T 模块与 SaE 注意力机制，显著提升了模型在复杂场景下的检测精度与鲁棒性^[10]。实验表明，该方法在干扰环境下仍能保持 97.3% 的 mAP50 与 94.6% 的召回率，优于现有主流模型。本方法可应用于水利设施人员准入控制、视频监控异常检测与网络终端生物认证等场景，为提升黄河水利系统的网络安全防护能力提供了可靠的技术手段。

参考文献

[1] Bizjak M, et al. Mask R-CNN for ear detection. IEEE, 2019.

[2] Yuan L, Lu F. Real-time ear detection based on embedded systems. IEEE, 2018.

[3] Kamboj A, et al. CED-Net: context-aware ear detection network. Pattern Analysis and Applications, 2021.

[4] Nguyen H Q, Truong H V. Real-Time Human Ear Detection Based on YOLO and RetinaFace. Complexity, 2021.

[5] 钱显来, 盖绍彦, 郑东亮, 等. 基于局部和全局信息的快速三维人耳识别 [J]. 仪器仪表学报, 2019, 40 (11): 99-106. DOI:10.19650/j.cnki.cjsi.J1904807.

[6] 杜保伟. 基于深度学习的人耳识别的研究和实现 [D]. 长春大学, 2021.

[7] 钱俊如. 基于深度学习的动态人耳识别方法研究 [D]. 长春大学, 2021.

[8] Terven J, et al. A comprehensive review of YOLO architectures. Machine Learning and Knowledge Extraction, 2023.

[9] Li Y, et al. Contextual transformer networks for visual recognition. IEEE TPAMI, 2022.

[10] Mahendran N. SENetV2: Aggregated dense layer for channelwise and global representations. arXiv:2311.10807, 2023.