

文献综述：于改进的 YOLOv3 和 FaceNet 的无人机影像人脸识别

1. 文献信息

- **作者：** 高锦风，陈玉，魏永明，李剑南
- **发表年份：** 2023
- **文献标题：** 于改进的 YOLOv3 和 FaceNet 的无人机影像人脸识别
- **期刊/会议名称：** 中国科学院大学学报
- **DOI/网址：** 10.7523/j.ucas.2021.0019

2. 研究背景与目的 (Why)

本研究基于无人机影响的高精度识别在应急救援、嫌疑人员跟踪等场景，针对传统人脸识别主要识别静态正面人脸，而无人机的拍摄角度很难聚焦正脸的问题，提出改进后的yolov3+Facenet结构，在动态复杂环境中增强了对人脸的区分能力，提高了精度以及遮挡和模糊的鲁棒性。

相关文献提及：

高刘雅等 [11] 在 YOLOv3 的基础上提出轻量级注意力机制的网络实现了人脸的准确检测，准确率达到 94.08%

潘浩然 [12] 通过改进 YOLOv3 的损失函数改善 YOLOv3 的错检漏检问题，实现了人脸的高精度检测

刘长伟 [15] 利用 MT CNN 结合 FaceNet 进行人脸识别，在 LFW (labeled faces in the wild) 数据集上实现了 99% 的准确率

李林峰等 [16] 结合 MT CNN 和 FaceNet 设计实现了人脸识别考勤系统

3. 研究方法与技术路线 (How)

提出了一种基于改进 YOLOv3 的人脸检测与 FaceNet 人脸匹配相结合的识别方案：

改进yolov3：针对召回率低：采用 SE-ResNeXt代替Darknet53 中的ResNet，加入SEnet (squeeze-and-excitation networks) 引入注意力机制，提升特征提取能力；并使用CloU (Complete IoU) 优化损失函数，提升目标框定位精度

FaceNet模型：采用 Inception-ResNet-v1 网络做为主体网络，在 Triplet Loss 指导直接对比图像特征向量之间的欧式距离，从而判断人脸图像的相似性。

数据集：WIDERFACE、自制数据集、CASIA-FaceV5

4. 核心内容与创新点 (What)

优化yolov3: 引入注意力机制 (SE-ResNeXt) 以提升检测精度, 损失函数替换为 CIoU, 增强预测框的几何一致性

提出改进后的**yolov3+Facenet结构**适用于高动态复杂场景的人脸检测: 对人脸的区分能力更强, 检测及识别精度更高, 同时对于无人机影像存在的遮挡、模糊等情况的鲁棒性也更高

5. 研究结论与实际意义

改进后的 YOLOv3+FaceNet 对人脸的区分能力更强, 检测及识别精度更高, 同时对于无人机影像存在的遮挡、模糊等情况的鲁棒性也更高, 在 CASIA-FaceV5 数据集上识别准确率达到 74.52%, 优于传统 YOLOv3+FaceNet (71.16%) 及 MTCNN+FaceNet (66.20%)。同时, 检测速度达到 MTCNN 的 3 倍以上, 能够有效运用于无人机应用场景下拍摄高度造成的影像非正脸的情况。

6. 本文的不足与未来展望

自制数据集容量不够大

7. 个人思考与选题关联

本文对我深入理解“复杂动态场景下的目标识别”提供了启发。YOLOv3 在速度上的优势与 FaceNet 的判别能力结合, 为实际部署场景提供了实用框架。未来我可能会考虑从**Transformer-based轻量人脸识别**角度切入, 在不牺牲性能的前提下, 进一步提升模型在边缘计算设备(如无人机、车载终端)中的识别效果与部署效率
