

团 体 标 准

T/GZAA 02-2023

AI 行人重识别技术规范

Technical Specification for AI Pedestrian Re-identification

2023-11-13 发布

2023-11-20 实施

目 次

前言	II
1 范围	3
2 规范性引用文件	3
3 术语和定义	3
4 要求	3
5 实现方法	4
参考文献	5

前 言

本文件按照 GB/T 1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由广州市自动化学会提出并归口。

本文件主要起草单位：华南理工大学、广州隐盾智能科技有限公司、广东省科学院智能制造研究所、佳都新太科技股份有限公司、广州城市理工学院、广东电网有限责任公司东莞供电局、江门市交通建设投资集团有限公司、华南师范大学、广东技术师范大学、广东农工商职业技术学院。

本文件主要起草人：张浪文、余孝源、谢巍、王智东、钟震宇、王昱、冯瑞珏、刘飘、罗金满、李博、邓丰强、何伟、罗旭、黎阳、谢嘉乐。

本文件为首次发布。

AI 行人重识别技术规范

1 范围

本文件规定了 AI 行人重识别技术的要求与实现方法。
本文件适用于所有 AI 行人重识别技术。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 41867—2022 信息技术 人工智能 术语

3 术语和定义

下列术语和定义适用于本文件。

3.1

AI Artificial intelligence

针对人类定义的给定目标，产生诸如内容、预测、推荐或决策等输出的一类工程系统。

[来源：GB/T 41867-2022，3.1.8]

3.2

行人重识别技术 Re-identification technology

使用AI算法进行对同一行人实现多次识别与综合判断的技术。

4 要求

4.1 要求组成

AI行人重识别要求主要由四部分组成，分别为技术、数据、系统和组织，参见图1。

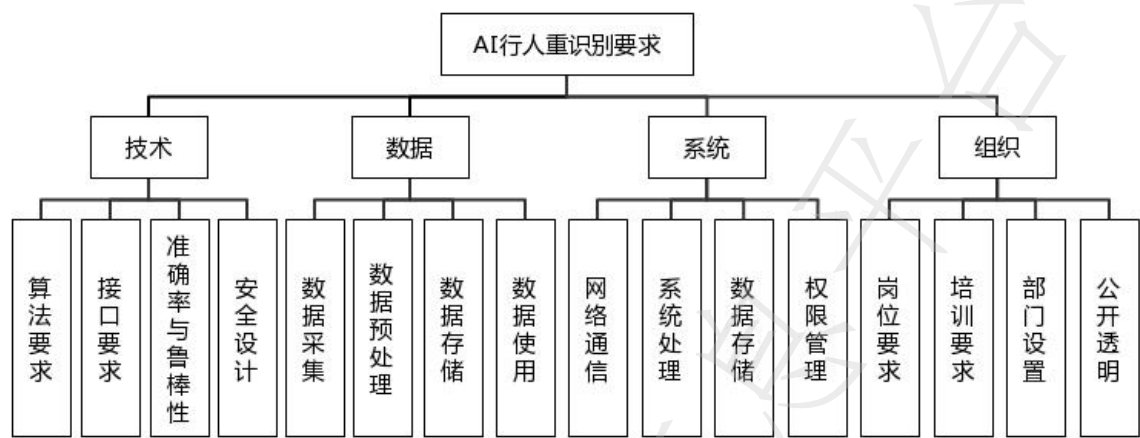


图 1 AI行人重识别要求组成结构图

4.2 技术

4.2.1 算法要求

- 4.2.1.1 根据应用场景需求和数据特点, 评估选择适宜的模型框架, 如深度学习模型、SVM 模型等。
- 4.2.1.2 算法训练和测试数据集应涵盖不同人群, 具有代表性, 避免数据偏差。
- 4.2.1.3 重识别算法应具备可解释性, 能够对识别结果进行解释, 如关键特征划重点等。

4.2.2 接口要求

- 4.2.2.1 提供符合标准规范的重识别服务接口, 便于与第三方系统对接。
- 4.2.2.2 设置反馈接口, 将重识别结果返回给请求方。

4.2.3 准确率与鲁棒性

- 4.2.3.1 重识别系统整体准确率不低于 95%。
- 4.2.3.2 在不同条件下测试重识别鲁棒性, 确保准确率稳定。
- 4.2.3.3 根据应用类型设置不同的置信度阈值。
- 4.2.3.4 评估误报率和漏报率指标, 降低错误识别风险。

4.2.4 安全设计

- 4.2.4.1 数据传输接口应采用加密等安全措施。
- 4.2.4.2 对算法模型加密, 控制访问权限。
- 4.2.4.3 采用访问控制、身份验证等方法, 保证系统安全。
- 4.2.4.4 对重识别模型实施差异化保护。
- 4.2.4.5 对行人图像和信息进行脱敏和拟人化处理, 防止隐私泄露。

4.3 数据

4.3.1 数据采集

- 4.3.1.1 图像采集范围限定为执行重识别任务所需的人脸和步态特征。
- 4.3.1.2 向被采集者明示重识别系统使用功能, 取得同意后方可采集其图像。
- 4.3.1.3 对采集对象的身份标识、同意确认进行记录, 用于证明采集行为合规性

4.3.2 数据预处理

- 4.3.2.1 使用人脸检测和关键点匹配算法,实现人脸图像的自动对齐。
- 4.3.2.2 对采集图像进行清洗和抽样,剔除无效样本,保证样本均衡。

4.3.3 数据存储

- 4.3.3.1 对图像样本进行分类标注,标明类别、特点等关键信息。
- 4.3.3.2 采用加密、访问控制等手段对样本库进行安全存储。

4.3.4 数据使用

- 4.3.4.1 图像数据仅可用于重识别系统功能的实现,不得用作他用。
- 4.3.4.2 重识别系统应持续更新图像样本库,优化识别效果。

4.4 系统

4.4.1 网络通信

- 4.4.1.1 图像及关键数据传输必须采用加密技术,防止泄露。
- 4.4.1.2 对外部系统访问设置过滤控制,防止非法访问重识别系统。
- 4.4.1.3 开放适当查询接口。

4.4.2 系统处理

- 4.4.2.1 设置冗余系统部署方案,实现容灾备份能力。
- 4.4.2.2 重识别结果应可追溯到对应的输入图像,以便复核。
- 4.4.2.3 提供行人数据删除和修改的功能接口,保障数据控制权。

4.4.3 数据存储

- 4.4.3.1 图像和模型数据必须采用加密技术进行存储,并设置访问权限。
- 4.4.3.2 制定冗余备份策略,防止重要数据丢失。

4.4.4 权限管理

- 4.4.4.1 根据角色设定不同的系统访问权限,遵循最小权限原则。
- 4.4.4.2 系统操作行为应全部记录在审计日志中,用于安全审查。
- 4.4.4.3 根据内部查询与外部查询设置不同权限控制流程,限制非授权访问。

4.5 组织

- 4.5.1 明确重识别系统相关岗位的职责,如系统管理员、识别操作员等。
- 4.5.2 重识别系统的操作人员应定期接受不少于 20 小时/年的培训,培训内容包括相关法律法规、系统操作流程、信息安全知识、职业道德等。
- 4.5.3 重识别系统的管理方应专门设置系统运营管理部门,配备 2 名以上专职人员负责系统的日常运行监督、管理和维护工作。
- 4.5.4 重识别系统管理方应面向公众公开系统的法律依据、运营方信息、投诉举报渠道等,并妥善响应公众的查询、建议和投诉。
- 4.5.5 建立第三方审计机制,接受社会组织监督。

5 实现方法

5.1 技术实现

5.1.1 算法实现

5.1.1.1 根据场景和数据特征评估算法模型,选择深度学习或机器学习框架,给出选型分析报告。

5.1.1.2 构建具有代表性的训练和测试数据集,包含不同人群,每类样本量不少于 500 张。

5.1.1.3 在模型结构中加入可解释组件,标注关键特征或设置解释参数等。

5.1.1.4 构建公平性评估指标,如不同群体的误识率,防止算法歧视。

4.1.3.4 评估误报率和漏报率指标,降低错误识别风险。

5.1.2 接口实现

5.1.2.1 设计标准化重识别服务接口,采用 Representational State Transfer 规范。

5.1.2.2 返回结果数据集成识别置信度、系统日志等,格式符合标准。

5.1.2.3 设置人工审查机制,避免获取不必要数据。

5.1.3 准确率与鲁棒性

5.1.3.1 在数据集上多次测试评估整体准确率水平,进行模型优化以达到指标要求。

5.1.3.2 构建包含不同情况的图像数据集,测试识别鲁棒性并优化。

5.1.3.3 根据使用场景设置匹配的置信度阈值分级标准。

5.1.4 安全实现

5.1.4.1 采用 HTTPS 等安全的网络传输协议,使用 AES 等加密算法。

5.1.4.2 对核心算法模型文件加密,设置访问权限控制。

5.1.4.3 完善系统登录验证、操作日志审计等访问控制机制。

5.1.4.4 采用模型差异化保护技术

5.2 数据开发

5.2.1 数据采集

5.2.1.1 在采集设备和存储中设置仅采集人脸和步态所需图像范围的逻辑限制。

5.2.1.2 在采集场所设置重识别系统说明,提供电子签名等获取使用同意的方式。

5.2.1.3 对采集对象设置身份编码,记录同意采集的确认信息。

5.2.2 数据预处理

5.2.2.1 使用 OpenCV 等开源工具包实现人脸检测和关键点提取、匹配。

5.2.2.2 设置数据筛选算法,自动清洗错误样本,保证样本库均衡性。

5.2.3 数据存储

5.2.3.1 为图像样本建立结构化数据库,包含类别、标签等信息字段。

5.2.3.2 在样本数据库设置访问权限控制,使用加密和冗余存储提高安全性。

5.2.4 数据使用

5.2.4.1 在系统设计中限定仅可将样本数据应用于重识别功能。

5.2.4.2 通过新增采集模块实时更新样本库,保证样本量和覆盖面。

5.3 系统建设

5.3.1 网络实现

5.3.1.1 采用 SSL/TLS 等加密传输协议,使用 AES 等算法加密数据。

5.3.1.2 设置访问控制列表,对外部访问 IP 进行过滤。

5.3.2 处理实现

5.3.2.1 构建双机热备架构或云容灾方案,实现系统容灾能力。

5.3.2.2 系统返回结果时包含对应输入数据的标识符,可追溯到源。

5.3.2.3 开发数据删除和修改接口。

5.3.3 存储实现

5.3.3.1 对存储系统加密,采用访问控制列表等限制非授权访问。

5.3.3.2 制定每日“全量+增量”数据备份策略,备份到异地。

5.3.4 权限实现

5.3.4.1 根据用户角色设定访问权限,操作权限等,启用最小权限机制。

5.3.4.2 核心操作行为均记录到统一审计日志中,每月审阅。

5.3.4.3 根据内部查询与外部查询设定不同审批流程,实施访问权限分级控制。

5.4 组织构建

5.4.1 制定岗位说明书时,应当参考相关法律法规,明确重识别系统运营相关岗位的主要职责、任务和权限等。

5.4.2 重识别系统操作员培训应采取理论培训与实操培训相结合的形式,确保操作员掌握法规知识和操作流程,并通过考核评估合格方可上岗。

5.4.3 重识别系统运营管理部门的设置应符合国家网络安全管理的相关规定,专门负责系统的日常监督、运维和安全管理,具体人员配备比例应与系统规模匹配。

5.4.4 面向公众公开的相关信息应符合国家网络安全和信息公开制度的规定,妥善响应和处理公众查询、建议和投诉,确保公众知情权。

5.4.5 聘请独立第三方机构,开展定期审计评估

参 考 文 献

- [1] GB/T 35295-2017信息技术 大数据 术语
 - [2] GB 5768.7-2018 道路交通标志和标线 第7部分：非机动车和行人
 - [3] 中华人民共和国数据安全法
-