多模态生物特征信息安全防护体系研究

胡先智[1],陈 浩[2],梁 艳[3]
(1. 西安理工大学 信息化管理处,陕西 西安 710048;
2. 西安理工大学 计算机科学与工程学院,陕西 西安 710048;
3. 西安思源学院 理工学院,陕西 西安 710038)

摘 要:在新的计算能力和深度学习技术推动下,人工智能、大数据发展进入了繁荣期,导致多模态生物特征信息迅猛增加与应用。由于多模态生物识别具有自然性和多场景应用性,特征信息的采集、识别、分析不仅涉及个人隐私和人格尊严,还主动或被动暴露在现实环境中,高校面临着巨大的信息安全保护需求和风险挑战。通过对高校多模态生物特征信息安全问题及现状进行分析,提出从基础设施安全防护、网络安全防护、数据安全防护、应用安全防护四个层面构建多模态生物特征信息安全防护框架,并将管理数据防护、业务数据防护、用户鉴别信息防护、分类分级全生命周期信息防护策略与技术相结合,以着力解决不同维度的安全风险和隐患问题。结合国内高校多模态生物特征信息实施应用情况进行对比分析,验证了所提出的信息防护策略与技术有良好的应用效果,为加强高校多模态生物特征信息防护能力提供借鉴和参考。

关键词:人工智能;多模态生物特征;分类分级;安全防护;全面生命周期
中图分类号:TP391 **文献标识码**:A **文章编号**:1673-629X(2022)04-0086-06
doi:10.3969/j.issn.1673-629X.2022.04.015

Research on Information Security Protection System of Multimodal Biometrics Identification

HU Xian-zhi[1],CHEN Hao[2],LIANG Yan[3]
(1. Division of Information Management,Xi'an University of Technology,Xi'an 710048,China;
2. Faculty of Computer Science and Engineering,Xi'an University of Technology,Xi'an 710048,China;
3. School of Technology,Xi'an Siyuan University,Xi'an 710038,China)

Abstract:Driven by the new computing capabilities and deep learning technology,the development of artificial intelligence and big data has entered a boom period,resulting in the rapid increase and application of multimodal biometric information. Due to the naturalness and multi-scenario application of multimodal biometrics,the collection,identification and analysis of feature information not only involves personal privacy and dignity,but also actively or passively exposed to the real environment. Colleges and universities are facing huge information security protection needs and risk challenges. By analyzing the security problems and current situation of multimodal biometric information in colleges and universities,we propose to build an information security protection framework of multimodal biometrics from four levels:infrastructure security protection,network security protection,data security protection and application security protection. Then strategy and technology are combined by management data protection,business data protection,user identification information protection,classification and all life cycle information protection,which solve the security risks and hidden dangers in different dimensions. A comparative analysis of the implementation and application of multimodal biometric information in domestic colleges and universities verifies the good application effects of the proposed information protection strategy and technology. It provides reference for strengthening the information protection capabilities of multimodal biometrics.

Key words:artificial intelligence;multimodal biometrics;classification and gradation;security protection;all life cycle

0 引 言

人工智能(AI)[1]是新一代信息通信技术,作为“十四五”时期国家新型基础设施建设(简称新基建)七大核心领域之一[2],“AI+应用”已成为各行业数字

收稿日期:2021-05-10 **修回日期**:2021-09-12
基金项目:国家自然科学基金项目(61976175);陕西省教育厅专项科学研究计划项目(19JK0686)
作者简介:胡先智(1978-),男,硕士,工程师,研究方向为教育信息化、信息安全。

化的重要驱动。在这一新模式新业态下,高校积极开展人脸、虹膜、指纹、掌纹等人体多模态生物特征融合技术的研发与应用,围绕教学、消费、安防、图书、医疗、会议、考勤控等场景进行人机交互和身份认证,将大数据与多模态生物特征识别技术相结合,形成海量多模态生物特征信息[3-4]。然而,生物特征识别技术的推广应用也产生一定的负面影响,尤其对师生隐私信息泄露带来前所未有的风险挑战。新冠疫情发生以来,多模态生物特征信息泄露和技术滥用等造成的信息安全问题突出,引起了国家重视。其中,将“建设安全便捷的智能社会和更高水平的平安中国”列为其重要任务之一。因此,针对高校智慧校园多模态生物特征信息安全防护问题,为加强其安全风险防范和管理能力,提出构建高校多模态生物特征信息安全防护体系,为高校、社会等提供信息安全防护参考。

1　安全现状分析

1.1　安全问题

随着大数据、物联网、人工智能技术的发展,师生可方便地利用多模态生物特征信息实现身份识别验证与移动互联支付等服务,但此类识别信息的收集与频繁共享应用[5-6],却面临着较大的安全问题。

(1)网络安全防护技术手段不够。网络安全防护大多从多模态生物特征识别系统防护角度进行信息保护,缺乏边界、业务、运维等综合整体安全防护[7],信息隐私泄露面临巨大挑战。

(2)应用安全管理不到位。作为高敏感性的生物特征识别信息,由于应用管理机制不健全,产生信息泄露和被人盗取的风险。

(3)数据安全管理不足。部分高校对其安全治理需求不够明确、尚未制定数据管理和安全相关战略规划,信息主库面临被外界攻击及自身重视不力导致信息泄露的风险。

(4)基础设施不完善,存在物理安全漏洞,导致信息篡改、窃取等风险。

1.2　信息泄露风险分析

多模态生物特征信息泄露,导致身份信息被盗用、篡改、窃取,行踪被记录、财产损失等安全风险,造成比较严重的影响,其安全风险表现在四个方面:

(1)多模态生物特征识别平台的共享、融合引发信息泄露风险。随着多模态生物特征信息采集系统应用,原本分散存储的个人相关信息逐渐向平台统一汇合,形成数据的“黑匣子”[8],易成为黑客、不法分子的攻击对象。

(2)多模态生物特征识别平台的漏洞造成信息泄露风险。多模态生物特征识别平台存在漏洞未修复、端口开放等问题,会导致信息数据泄露、篡改、窃取等风险。

(3)多模态生物特征识别平台的互联互通带来信息泄露风险。

(4)多模态生物特征识别平台利用大数据、人工智能、与计算、物联网等新技术带来了信息泄露风险。

2　防御体系架构

2.1　设计思路

在人工智能背景下,高校多模态生物特征信息种类和形态呈现多样化,传统信息安全标准等难以维持多模态生物特征信息安全工作的开展[9]。因此,结合当前多模态生物特征信息实际应用要求和安全风险分析,以及安全防护的问题与难点,以信息系统安全等级保护 2.0 为指导标准,从基础设施安全、网络安全、数据安全、应用安全四类防护层面,提出多模态生物特征信息防护架构,如图 1 所示。结合信息安全防护内容,从统一的安全管理中心、信息安全运行体系等方面实现多模态生物特征信息安全防护措施。

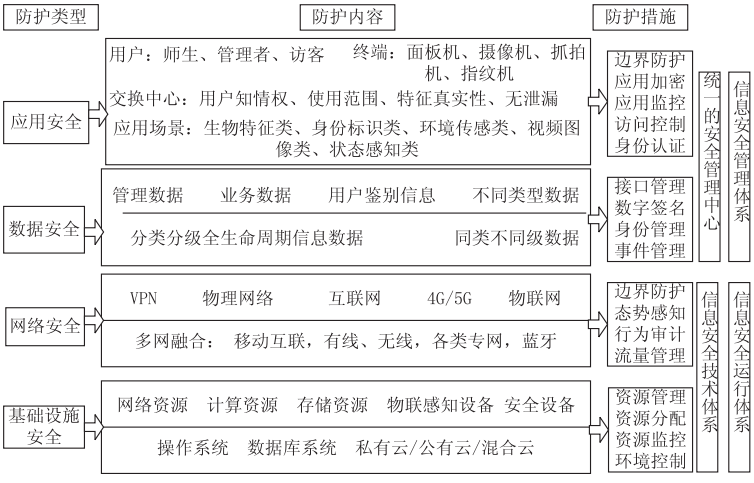

图 1　多模态生物特征信息防护架构

2.2 架构设计

本架构结合多模态生物特征信息安全管理和技术要求,从基础设施、网络、数据及应用等方面提出安全防护措施。重点根据不同类别,有针对性地提出多模态生物特征信息分类防护措施,解决各类信息差异化的防护问题。本架构利用“人防物防技防协防融合、因势利导、分类治理、分级管理”的综合防护思路,旨在解决不同类别、级别的多模态生物特征信息防护问题。

(1)人防、物防、技防、协防融合,实现四位一体安全防护。

(2)因势利导。按照“两地三中心、无感应用”容灾备份模式,既关注本地信息保护,也重视异地信息采集、传输、迁移、交换共享与公开披露等全流程安全防护。

(3)分类治理。根据多模态生物特征信息的需求及应用场景的不同,重点对敏感数据、重要数据、一般数据进行分类的信息安全防护[10],有针对性地提出多模态生物特征信息分类治理、多重保障,确保师生信息安全、稳定。

(4)分级管理。围绕多模态生物特征信息全生命周期,明确差异化的多模态生物特征信息分级安全防护要求。

3 建立安全防护体系

3.1 基础设施安全防护

从物理安全、主机安全两方面实现多模态生物特征信息的基础设施安全防护,明确多模态生物特征信息物理安全要求,并从系统安全防护的角度实施主机安全防护。

3.1.1 物理安全

(1)机房及基础设施管理。妥善选择机房地址,合理划分机房物理区域,并从使用空间、容量等方面科学布置基础设施,满足基础设施扩容需求,防范物理和环境潜在风险和非授权访问[11]。制定安全管理制度流程与规范,设立安全管理部门,明确关键管理人员,在运维中严格推动访问控制、监控审计、应急响应等措施落地,及时整改消除重要数据泄露、滥用等安全隐患,以确保多模态生物特征信息的物理环境安全。

(2)网络设备安全管理。结合多模态生物识别业务和安全状态需求,合理选择网络设备资源,同时对网络设备访问、协议和服务、日志、配置、软件版本升级等进行安全管理,保障设备安全。

(3)安全设备管理。由于多模态生物特征信息基础设施接入前后需要进行安全审核,因此,在进行成本/效益分析的基础上,通过部署堡垒机、下一代安全网关、虚拟化防火墙、安全态势感知设备、数据库审计系统、日志审计系统等安全设备,合理划分网络设备安全防护边界和安全区域,实现纵深防御、区域访问控制和有效安全隔离,消除多模态生物特征信息面临的各种风险。

3.1.2 主机安全

多模态生物特征识别利用机器学习为用户、主机带来大量的数据,需要严格地控制对关键主机、端口和服务的访问,替换服务器标识,隐藏内部网络结构,阻止 DoS 攻击和网络扫描,实现主机在数据存储和处理的保密性、完整性、可用性,帮助高校构建主机在资产管理、文件查杀、入侵检测、漏洞扫描、安全基线、攻击监测等方面安全防护体系。

3.2 网络安全防护

多模态生物特征信息网络安全防护包括区域安全、网络层安全。按安全等级保护和国密要求,安全区域内的节点采用相同的安全等级和信任关系。安全网络层承载安全隔离、接入控制和边界防护,严格执行相应的管控措施,进行攻防测试,确保多模态生物特征信息网络安全防护。

(1)架构安全。采取合理架构,利用信息网络安全传输策略,明确划分多模态生物特征信息网络边界和区域。

(2)访问管控。采用下一代安全网关、虚拟化防火墙、安全态势感知设备,实现横向纵向防护与安全隔离,确保多模态生物特征信息网络流量访问安全。同时,设定管理员和普通用户的权限,对信息访问形成日志记录。

(3)边界防护。对物理网络和虚拟网络、物理主机与虚拟机进行边界区域安全隔离、入侵监测。同时,建立多模态生物特征信息边界立法及惩罚机制。在非授权的情况下,对违规使用生物特征识别信息的用户进行惩戒。

(4)传输加密。利用信息加密、隐藏、水印等技术实现多模态生物特征信息主机、虚拟机间的保密传输[12],防止传输信息被盗取、修改等。另外,针对第三方开放应用,通过注册认证、引擎调用口令密钥的方式,确保多模态生物特征信息授权传输。

(5)网络监控。采取网络态势感知、日志审计等技术对多模态生物特征信息网络的运行状态、流量等进行监控,消除多模态生物特征信息面临的网络问题。

3.3 数据安全防护

多模态生物特征信息数据安全防护围绕管理数据、业务数据、用户鉴别信息[13-14],重点针对不同类别的多模态生物特征信息数据特征和安全防护需求,保证租户对多模态生物特征信息数据的隐私权、所有权

和控制权不受侵犯,为用户提供最切实有效的数据保护。同时,多模态生物特征信息数据安全保护在数据分类的基础上,围绕同一类数据全生命周期,标记数据,建立数据资产分级清单,按照数据级别确定安全管理策略和防护措施,明确差异化的数据分级安全防护要求,如图 2 所示。

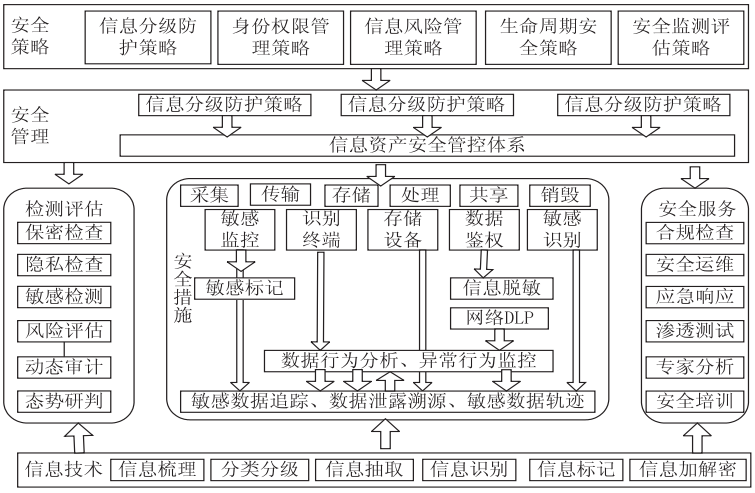

图 2　多模态生物特征信息分类分级治理防护

(1)管理数据防护。管理数据包括人员组织架构、多模态生物信息、运维、供应链等非结构化文档数据。其中,管理数据在传输、交换共享等方面需求大且保密性、完整性要求相对较高。因此,采用数据脱敏、数据水印、数据二维码、信息隐藏、身份识别、访问控制等措施实现管理数据安全防护。

(2)业务数据防护。业务数据包括消费、门禁、考勤、图书、医疗、会议等场景数据,主要是多模态生物特征信息、文档、时序数据、关系表数据等结构化与半结构化数据,涉及多模态生物特征信息服务平台核心内容,其保密性和可靠性要求相对较高。因此,为了保护多模态生物特征识别业务数据在采集、传输、比对等环节不被泄露、篡改和非授权分析,可通过操作权限管理、接口认证、不可逆转译成图像、数据加密、数字水印、虚拟专用网络、安全审计、冗余备份等措施,保障业务数据的有效性、可用性、安全性。

(3)用户鉴别信息防护。用户鉴别信息是通过用户/系统身份建立的一项信任,使用鉴别令牌、基于证书的鉴别、生物方式鉴别、认证加密或数位签章实现信息鉴定。结合多模态生物特征用户鉴别信息完整性的安全防护要求,采用签名等技术,防止对方破解信息;采用基于共享密钥、生物学特征、公开密钥等身份验证,防止伪用户窃取信息;通过对消息进行加密,实现信息鉴别码的保密认证,确保数据可用性,提高用户隐私保护和信息传输安全能力。

为此设计了用户鉴别信息隐私保护,如图 3 所示。针对各种生物特征识别技术,部署算法引擎池,将各商家、算法引擎进行集中注册和统一管理;采集各类生物特征信息,并建立多个特征库的管理器,与前端应用、智能终端建立起多对多映射,按需推送相关的生物特征供前端使用,确保原始信息不被传递到前端,防止用户鉴别信息隐私泄露。

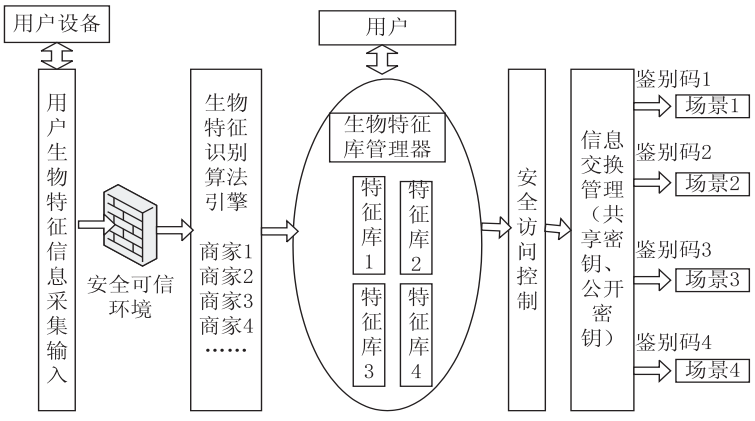

图 3　用户鉴别信息隐私防护

(4)全面生命周期防护。多模态生物特征信息安全治理包含管理、运维、风险管控、技术支撑、标准化等一系列内容,是高校教育信息化发展到一定程度,利用已有数据推动各项业务进入一种新的数据治理形态。

但在现有体系下全生命周期的多模态生物特征信息采集、传输、存储、处理、交换共享、销毁等阶段存在安全风险,需要进行安全防护,如表 1 所示。

表 1 多模态生物特征信息全生命周期防护

生命周期	安全风险点	防护措施
信息采集	数据库管理系统、数据标签、采集设备、数据源	利用统一身份认证平台
信息传输	传统加密、明文传输;多路径、跨区域传输	采用规范的信息传输接口,使用加密算法、信息隐藏等
信息存储	容灾备份、访问控制及身份认证	提供“两地三中心”信息容灾备份与恢复机制
信息处理	敏感信息处理、信息处理平台间认证机制	采用信息隐藏、加密、鉴定、脱敏、信息回退机制
信息交换共享	访问控制权、交换共享、信息公开披露及认证	采用身份认证、信息加密、共享监控、信息水印措施
信息归档销毁	信息删除、存储介质	建立信息归档机制,格式化或物理销毁存储介质,日志记录

3.4 应用安全防护

针对多模态生物特征信息鉴别、控制、审计、保护等应用安全,可部署防护墙、网闸、定制边界安全防护设备,对多模态生物特征信息的访问权限进行精细化划分,设定不同权限等级,实现安全访问控制,从而防止越权访问和各功能区间的病毒感染;可设立多模态生物特征信息身份认证管理机制,对关键平台、设备,根据信息重要程度,采用单因子或多因子认证方式;还可为每条数据赋予可信时间戳,利用区块链技术防伪溯源,避免对信息篡改、窃取,保障身份认证合规性,通过管控越权来降低信息违规访问风险[15-18]。

另外,加强关键网络节点、攻击行为、异常流量、信息有效性等方面检测,加强操作系统安全防护,同时部署安全审计系统,对内网访问互联网行为内容进行审计和监控,可有效抵御外部攻击、内部人员私自调整账号权限等行为。

4 实施部署

结合该文构建的多模态生物特征信息安全防护框架,利用管理数据防护、业务数据防护、用户鉴别信息防护、分类分级全生命周期信息防护策略与技术,实施部署多模态生物特征信息平台如下:

硬件方面:GPU 服务器(2 颗英特尔至强银牌 5218 处理器,256 GB 内存,8 块 1200 GB-SAS 12 Gb/s -10K rpm 硬盘;4 块 480G SSD 硬盘,1 块 NVIDIA-Tesla T4-16 GB GPU 卡)。

软件方面:Oracle 11g。

安全方面:利用 SDK 和特征值进行数据隐私保护;整个平台部署云环境,通过“IP + MAC + 端口”的方式,限定可访问后台的终端设备。借助账户密码、生物特征、二维码、移动设备进行多因素认证;调用密钥、数据隔离、多级授权等进行应用控制。考虑到对第三方开放共享应用,通过引擎调用口令密钥的方式,确保合规授权。

传输方面:通过 https 进行传输加密。

数据备份与容灾:通过 Lustre 和 Oracle 分别进行特征数据存储,并实现主备存储。

5 应用验证与分析

根据多模态生物特征信息平台部署情况以及多模态生物特征信息安全防护体系架构,对师生提供的信息进行验证和比对,同时,在不了解师生信息的情况下利用生物特征在数据库中进行搜索其身份,从而验证信息防护策略的应用效果。

由于区域教育发展水平不同和经济水平的限制,各高校多模态生物特征信息应用侧重点有所不同。截止 2021 年 3 月,对高校多模态生物特征信息应用及安全防护情况进行统计和验证其应用,选取了中国东部、西部、南部、北部、中部 60 余所高校进行了统计,其中世界一流大学 15 所,世界一流学科高校 15 所,非一流本科高校 15 所,高职院校 15 所,统计见表 2。

表 2 多模态生物特征信息应用及安全防护情况统计

应用	世界一流大学(15 所)	世界一流学科高校(15 所)	非一流本科高校(15 所)	高职院校(15 所)	四类及全生命周期防护
人脸	14	11	9	5	35 所
虹膜	4	0	0	0	4 所
指纹	4	2	0	0	6 所
掌纹	0	0	0	0	0 所

从表 2 中可以看出,国内高校多模态生物特征信息应用及安全防护情况:世界一流大学应用技术及安全防护走在前列,世界一流学科高校应用技术次之,高职院校应用技术相对较弱。下面从三方面进行效果

分析。

(1)应用验证。

表 2 中,大部分高校利用人脸信息进行身份识别,但各高校采用文中的信息防护策略和技术也不尽相同,其中 14 所世界一流大学、11 所世界一流学科高校比较重视信息系统安全等级防护,人脸特征信息没有泄露和篡改发生,但有 4 所非一流本科高校、高职院校存在局部信息泄露隐患。

(2)共享与开放。

表 2 所统计高校中已建多模态生物特征信息平台不到 50%,其开放不足限制了业务场景应用。另外,由于多模态生物特征信息平台缺乏规范性和统筹性,导致信息资源冗余和分散,难以形成合力应用环境。另外,部分学校在信息共享方面,通过 APP 接口,直接传输人脸照片,而不是特征值或者加密水印信息,导致信息泄露或者篡改,无法进行追踪分析。

(3)安全防护。

结合该文所设计的多模态生物特征信息安全防护体系,针对表 2 中四类高校,对比分析多模态生物特征信息应用的安全防护问题,12 所世界一流大学在基础设施安全、网络安全、数据安全、应用安全四类及全生命周期防护层面认识程度高,建设投入力度大,所面临的网络病毒纵向横向攻击、黑客攻击、信息篡改、信息泄密等安全问题基本没有,而 6 所世界一流学科高校、非一流高校所面临的安全问题,2 所高职院校由于资金投入有限,重视程度不够,所面临的安全问题比较严峻,甚至影响多模态生物特征信息的应用与推广。

6　结束语

对多模态生物特征信息安全问题与风险进行了剖析,总结分析了当前多模态生物特征信息安全防护存在的问题和安全风险。针对多模态生物特征信息安全问题与风险,设计了从基础设施安全、网络安全、数据安全、应用安全四类防护层面构建多模态生物特征信息安全防护框架,将管理数据、业务数据、用户鉴别信息、分类分级全生命周期信息防护策略与技术相结合,解决不同层次的安全问题,为高校智慧教育信息化建设提供参考和实践指导。下一步,将进行多模态生物特征融合防假体攻击技术方面的研究。

参考文献:

[1] 杨晓哲,任友群. 教育人工智能的下一步—应用场景与推进策略[J]. 中国电化教育,2021(1):89-95.

[2] 国家标准化管理委员会(2020). 国家新一代人工智能标准体系建设指南[EB/OL]. 2020-07-27. http://www.gov.cn/zhengce/zhengceku/2020-08/09/content_5533454.htm.

[3] 王会勇,唐士杰,丁　勇,等. 生物特征识别模板保护综述[J]. 计算机研究与发展,2020,57(5):1003-1021.

[4] YANG H X,SUN E,CHENG C,et al. Multi-modal biometrics based on data fusion[J]. Journal of Physics: Conference Series,2020,1684(1):012023.

[5] 张雪莹,杨帅锋,王冲华,等. 工业互联网数据安全分类分级防护框架研究[J]. 信息技术与网络安全,2021,40(1):2-9.

[6] VHADURI S, POELLABAUER C. Summary:multi-modal biometric-based implicit authentication of wearable device users[J]. IEEE Transactions on Information Forensics and Security,2019,14(12):3116-3225.

[7] 张　宁,藏亚丽,田　捷. 生物特征与密码技术的融合——一种新的安全身份认证方案[J]. 密码学报,2015,2(2):159-176.

[8] 毋立芳,马玉琨,周　鹏,等. 生物特征模板保护综述[J]. 仪器仪表学报,2016,37(11):2407-2420.

[9] TIONG L C O, SONG T K , YONG M R. Multimodal facial biometrics recognition:dual-stream convolutional neural networks with multi-feature fusion layers[J]. Image and Vision Computing,2020,102:103977.

[10] 杨雪鹤,刘欢喜,肖建力. 多模态生物特征提取及相关性评价综述[J]. 中国图象图形学报,2020,25(8):1529-1538.

[11] GUPTA S,BURIRO A,CRISPO B. Driver auth:a risk-based multi-modal biometric-based driver authentication scheme for ride-sharing platforms[J]. Computers & Security,2019,83:122-139.

[12] 李唐薇,叶学义,夏经文,等. 面向生物特征的信息隐藏算法[J]. 杭州电子科技大学学报:自然科学版,2017,37(2):39-44.

[13] KANT C,CHAUDHARY S. A watermarking based approach for protection of templates in multimodal biometric system[J]. Procedia Computer Science,2020,167:932-941.

[14] 潘林青. 面部特征信息法律保护的技术诱因、理论基础及其规范构造[J]. 西北民族大学学报:哲学社会科学版,2020(6):75-85.

[15] SADHYA D,SINGH S K. Privacy preservation for soft biometrics based multimodal recognition system[J]. Computers & Security,2016,58:160-179.

[16] 肖建力,张　静. 联合人脸与指纹的多模态生物特征识别方法综述[J]. 上海理工大学学报,2017,39(1):54-60.

[17] 荆继武,刘丽敏,回春野,等. GB / T 36651-2018 信息安全技术基于可信环境的生物特征识别身份鉴别协议框架[S]. 北京:中国标准出版社,2018.

[18] 钟　陈,秦日臻,张　璋,等. 生物特征识别系统安全性分析与对策[J]. 信息技术与标准化,2018(7):31-35.