Smart Learning Environments

REVIEW

Open Access

Continuous user identification in distance learning: a recent technology perspective

David Portugal[1*], José N. Faria[1], Marios Belk[2], Pedro Martins[1], Argyris Constantinides[3], Anna Pietron[2], Andreas Pitsillides[3], Nikolaos Avouris[4] and Christos A. Fidas[4]

*Correspondence:
davidbsp@isr.uc.pt

[1] Institute of Systems and Robotics, University of Coimbra, 3030-290 Coimbra, Portugal
[2] Cognitive UX GmbH, 69253 Heidelberg, Baden-Württemberg, Germany
[3] Department of Computer Science, University of Cyprus, 1678 Nicosia, Cyprus
[4] Department of Electrical and Computer Engineering, University of Patras, 265 04 Patras, Greece

Abstract

The worldwide shift to distance learning at Higher Education Institutions (HEIs) during the COVID-19 global pandemic has raised several concerns about the credibility of online academic activities, especially regarding student identity management. Traditional online frameworks cannot guarantee the authenticity of the enrolled student, which requires instructors to manually verify their identities, a time-consuming task that compromises academic quality. This article presents a comprehensive review of existing efforts around continuous user identification, focusing on intelligent proctoring systems and automatic identification methods, as well as their applicability in this domain. We conclude that there is a clear need for continuous user identification technology by HEIs, but existing systems lack agile system integration models that combine many inputs, such as face, voice and behavioural data in a practical manner, and encounter numerous barriers related to data protection during implementation.

Keywords: Continuous user identification, Distance learning, Intelligent proctoring systems, Image-based identification, Voice-based identification, Biometrics, Data privacy-preservation

Introduction

Most Higher Education Institutions (HEIs) underwent a fast transition towards a completely remote academic, teaching and learning paradigm in a first phase of response to the COVID-19 pandemic. Worldwide, HEIs quickly enhanced digital learning opportunities for both students and teachers and encouraged new forms of teacher collaboration. According to a recent survey (Schleicher, 2021), online platforms have been widely utilized at all levels of education across countries. However, this transition embraced severe challenges related to deploying trustworthy and credible ICT user identity management solutions.

Continuous user identity management refers to an iterative process, throughout a users' session with a remote service, that confirms that the interaction between the user and the system is performed continuously by the same person who initially logged in, and is therefore eligible to use remote services and computational resources.

From an academic and pedagogical perspective, being able to verify the students' identity continuously is mandatory within several online learning scenarios, like remote laboratories, exams and lectures, to prevent fraudulent behavior where individuals intentionally impersonate others in order to unethically participate in academic activities. Such uncertainty puts at risk the whole mission of HEIs, which is to ensure that individuals have acquired knowledge and competencies in order to acquire a profession. From another perspective, continuous student identification not only improves the credibility and trustworthiness of remote learning systems, but can also be used as a tool for student presence awareness. Being able to accurately perceive who is attending in remote synchronous and asynchronous scenarios scaffolds and reproduces social situations that occur in the physical classroom, such as group attendance and awareness of classmates' presence.

The recent transition to online **distance learning** led to a major adoption of video conferencing tools by HEIs worldwide due to their feasibility, flexibility, and accessibility (Cai & King, 2020). As a result, an exponential growth on the use of video conferencing tools for higher education purposes was observed, increasing inevitably the popularity of software tools such as Zoom[1] and Microsoft Teams[2] (Massner, 2021). While providing interesting features, such solutions are not tailored to classroom use cases, as videoconferencing is only viable if teachers and instructors use the tools effectively (Martin, 2005), which requires training, careful planning and commitment to maintain student engagement in online courses. Realizing this, key players of the video conferencing market felt the need to retain the large number of recent users by developing exclusive features for online education (see (Omar & Abdul Razak, 2020), (Kim & Lee, 2020) and (Okmawati, 2020)).

In the response to the COVID-19 pandemic and global shutdowns, institutions have gradually improved their support for delivery of distance learning to students through e-learning infrastructures, and according to Huang et al. (2020) and Coman et al. (2020), this became an opportunity to optimize the process by pushing education institutions to: (i) improve internet infrastructure to avoid interruptions, e.g., during video-conferences; (ii) use more intuitive tools to help students assimilate information; (iii) provide compelling and interactive electronic resources; (iv) build online communities for students to counter isolation; (v) foster creativity by using techniques, such as debates, or learning based on discovery and experience; and (vi) provide services to disseminate trends and policies adopted by universities and governments. Given the aforementioned, some institutions developed new in-house Learning Management Systems (LMSs), such as the web-based UCTeacher solution at the University of Coimbra.[3]

Besides flexibility and accessibility (Bakia et al., 2012), online learning eliminates barriers of space and time, facilitating collaboration, and allowing students to learn in their own rhythm (Arkorful & Abaidoo, 2014), overall students tend to assimilate information in the same way than in traditional classrooms (Navarro & Shoemaker, 2000) and online learning seems to be particularly beneficial for shy and slow learning students, who are

[1] https://zoom.us/.

[2] https://www.microsoft.com/microsoft-teams.

[3] https://ucpages.uc.pt/ucframework/apps/ucteacher/.

able to express themselves more often in the online classroom (Stern, 2004). However, there are also downsides in distance learning, such as students missing the social aspects of learning on campus, leading to feelings of isolation. Other disadvantages include the susceptibility to distractions, dependency on the internet and computers, which may fail unexpectedly, decreased motivation, and physical health consequences when spending several hours working with computers (Coman et al., 2020).

Statement of contributions

Distance learning has been the object of a substantial amount of research for decades. Regardless, continuous user identification within distance learning has proven to be a complex problem that has not been robustly solved, primarily due to challenges in non-intrusive authentication during online tasks, precision and accuracy of intelligent biometric techniques, trade-off between computational requirements and usability, security and data privacy aspects, and challenges related to integrating such technologies in existing learning management systems of HEIs.

This article aims to shed light on a contemporary debate on higher education, following a technology perspective. To this end, we propose the following contributions:

- **Comprehensive review** of existing efforts in continuous user identification for distance learning, specifically focusing on intelligent proctoring systems and automatic identification methods. This review offers a valuable resource for researchers and practitioners in the field.
- Explore various **image-based identification, voice-based identification, and biometric trait combination methods**. By discussing their applicability in the context of distance learning, we provide insights into the identification technologies and their potential integration.
- Highlight the relevance of **data privacy-preservation** issues in distance learning. This emphasis contributes to the ongoing discussion about protecting sensitive student information while implementing continuous user identification systems.
- Identify **research gaps, open issues, and prospects** for the advancement of continuous student identification systems. By pinpointing areas that require further investigation, the study guides future research efforts and informs the development of innovative frameworks.
- Propose a **tentative roadmap** for the future, aiming to design an innovative framework for student identity management. This roadmap serves as a starting point for researchers and practitioners interested in implementing privacy-preserving techniques for face, voice, and interaction-based continuous user identification.

Review structure

We begin by introducing the concept of **distance learning** in "Distance learning" section. Next, we present a literature review on intelligent user interfaces for distance learning, with a specific focus on **online proctoring systems** in "Intelligent online proctoring systems" section. Afterwards, we turn our attention to technologies for continuous user identification in "Technologies for continuous user identification" section, providing a comprehensive review of the most relevant methods of **image-based identification**, **voice-based identification** and **combination of biometric traits**.

Following the literature review, we briefly discuss the relevance of **data-privacy preservation issues** within the context of distance learning in "Data privacy-preservation issues" section, and the current scientific and technological landscape including an analysis of open research questions is discussed in "Research gaps, open issues and opportunities" section . Finally, in "Conclusion" section the article concludes by summarizing key findings, drawing final conclusions and proposing a tentative roadmap for the future, that we will attempt to pursue in our own ensuing research efforts.

In this study, we have employed a mixed methodology for selecting papers in our literature review. Our approach consisted of three main strategies. Firstly, we conducted database searches to identify relevant studies on prevailing intelligent proctoring systems and automatic user identification methods. Secondly, we sought expert consultation to obtain valuable input and recommendations in the different subject areas, thereby suggesting relevant studies and pointing to emerging as well as seminal works. Lastly, we have applied snowball sampling by expanding our search through the examination of the reference lists of the initially identified relevant papers for the study. By employing this combination of methodologies, we aimed to ensure a comprehensive and diverse selection of literature for our review.

Distance learning

Recent developments in distance learning have been driven by the increasing availability of digital technologies and the growing demand for flexible and accessible education. One major trend is the rise of Massive Open Online Courses (MOOCs), which offer free or low-cost courses from top universities and institutions around the world. Another trend is the use of virtual and augmented reality technologies to enhance the learning experience, such as simulating real-world scenarios for medical students (Zafar et al., 2020).

Distance learning refers to a pedagogic model that utilizes technologies, such as the Internet, allowing students to learn remotely without the need to physically attend a traditional classroom.

In distance learning, students can access course materials, communicate with instructors, and interact with peers through virtual platforms such as video conferencing, discussion forums, and LMSs. This model of education has recently grown in popularity, mostly due to the COVID-19 pandemic, which accelerated the adoption of distance learning as a necessary alternative to in-person classes. In this model, content is generally presented and delivered online through two different methods:

- **Synchronous learning** involves real-time interaction between the teacher, students, and course content. This method is similar to traditional classroom instruction, as the class meets together virtually at the same time. Students can engage in discussions, ask questions, and share their thoughts with each other and the instructor.
- **Asynchronous learning**, on the other hand, relies on self-directed study and group collaboration via online platforms. Students can access course materials such as documents, videos, or journals, on their own schedule, and engage in discussions with their peers and instructor at their own pace.

In general, it is common to use two types of assessments for evaluating student progress: formative and summative assessments. **Formative assessments** are designed to evaluate the student's understanding of the material and learning during the course. These are typically low-stakes assessments that provide ongoing feedback to students and instructors to identify areas where additional instruction or practice may be needed, allowing them to adjust teaching strategies accordingly. Formative assessments can take many forms, such as quizzes, homework assignments, or group projects (Wang & Tahir, 2020). **Summative assessments** are often administered at the end of a course or unit to evaluate student learning outcomes. These assessments are typically high-stakes and may determine a student's final grade or certification. Examples of summative assessments include final exams, term papers, and presentations. Since these assessments have a higher impact in the final grade, students might feel more compelled to engage in academic fraud (Genereux & McLeod, 1995), which is particularly aggravating in the case of online learning, requiring the implementation of preventative measures during these evaluations, which for mid-term or final exams are usually in the form of proctoring systems to monitor students during the exams.

Intelligent online proctoring systems

Intelligent user interfaces have been studied within distance learning, which primarily focus on: (i) providing personalized features based on the knowledge of students on particular subjects, their emotions, mood, personality, etc.; and (ii) on building intelligent proctoring systems for online examinations.

Several works on intelligent user interfaces examined e-learning platforms in the context of different learning styles paired with users' expectations, motivation, habits, and needs. These factors result in building an adaptive learning system providing the users with a unique learning experience *"based on the learner's personality, interests and performance in order to achieve goals, such as learner academic improvement, learner satisfaction, effective learning process and so forth"* (El Bachari et al., 2010; Truong, 2016; Kulaglić et al., 2013; Alexandru et al., 2015; Klašnja-Milićević et al., 2016; Montebello, 2018). In this section, we primarily focus on **online proctoring systems** with Artificial intelligence (AI) technology enhancement.

Online proctoring system refers to software used for examination supervision running on a student's computers after his/her identity has been approved. During the examination, the proctor (either a real person or an AI agent) is granted access to the student's web camera, computer screen, microphone, and in some cases computer mouse and keyboard.

The new reality of the COVID-19 pandemic has proven more than ever the importance of online proctoring systems. Even though there are many controversies related to the application of this technology concerning the potential invasion of students' privacy, civil rights and leading to additional stress or anxiety just to name a few (Helms, 2021; Coghlan et al., 2021), still 54% of HEIs utilize them and the statistics foresee that the further growth will continue, reaching a market size value of US$ 1,187.57 Million by 2027 (Grajek, 2021; Partners, 2021).

Online proctoring systems have been presented as a supporting tool in remote education for over 20 years. Initially, they were implemented as a feature of computer-based examinations to bridge the gap between remote and 'on-campus' conditions. With time, several online proctoring systems have been developed to serve as an 'off-campus' examination practice, fostering increased ownership of laptops and tablet computers and supporting remote education (Selwyn et al., 2021). The extensive and persistent evolution of proctoring systems has not only inspired numerous research works investigating their application and ethical concerns (Henry & Oliver, 2021; González-González et al., 2020; Coghlan et al., 2021), but also closely relates to user identity verification and access management (Fidas et al., 2021; Gonzalez-Manzano et al., 2019).

With the development of information systems and online accessibility to e-learning, e-banking, e-gambling, or e-government platforms, the necessity of authentication and providing correct access only to the authorized individuals became an integrated part of every identity management system. This is especially critical for all educational organizations that offer MOOCs and whose online certification and accreditation relies on students' online verification and assurance that all academic achievements were earned honestly. When executed inadequately, the reliability of credentials and certification earned online are affected and harmed through questioning their authenticity (Fidas et al., 2021; Labayen et al., 2021).

A recent research work (Nigam et al., 2021) provides a thorough review of the existing AI-based proctoring systems (AIPS) and online proctoring systems (OPS). The evolution from OPS to AIPS can be traced through the adoption of the technological development, transitioning from human invigilators manually verifying individuals' identification (e.g., by checking their identity and asking a few proofing questions) and overseeing the test-taker, to the application of AI processes that analyze and continuously monitor biometrics (e.g., facial recognition to match the photographed identity with the student's face, eye tracking, voice recognition or facial detection to detect any signs of malpractice). Nigam et al. (2021) provide a comprehensive overview of OPS, distinguishing between three types: (i) live proctoring, characterized by the use of the proctoring system in real-time and involvement of the human proctor who can flag students engaged in malpractice; (ii) recorded proctoring, characterized by registering the video for later human proctor analysis of face and eye movements; (iii) automated proctoring, characterized by limited involvement of the human proctor and automated identification of malpractice behavior (Hussein et al., 2020). By incorporating the technological advancement of AI processes in the last type of OPS, this model represents the group of AIPS. However, all OPS focus on two critical requirements: granting access to the web camera (for recording purposes) and preventing access to other web browsers, as well as preventing the opening of new web browser tabs on the computer during the examination (Alessio et al., 2017).

However, these systems are not foolproof and are vulnerable to various attack vectors (Constantinides et al., 2023). One such vector is a student violating identification proofs, wherein the student may use fraudulent identification documents (e.g. using still photographs of someone else) to bypass the system. Another attack vector is a student switching seats after identification, wherein the student may swap places with another person after passing the identification step. A non-legitimate access to shared LMS credentials can be used to bypass standard password based authentication. Additionally, computer-mediated communication through voice or text-written chat and screen sharing and control applications can enable cheating during exams. Students may also access forbidden online resources or receive assistance from non-legitimate individuals on their computer or through a secondary input device. Furthermore, students may communicate and collaborate with others in the same physical context or receive answers on a whiteboard or computing device. All these attack vectors pose significant challenges to proctoring systems, highlighting the need for continuous improvement and adaptation to new forms of academic dishonesty.

To address these threats, the design of AIPS systems should consider a variety of parameters depending on the hardware that is available for students (Slusky, 2020; Atoum et al., 2017; O'Reilly & Creagh, 2016; Li et al., 2015). These parameters include: (i) video recording of the user and their surroundings using a camera, since it is integrated in the majority of today's laptops (Machuletz et al., 2018) and available through a simple web camera add-ons for desktop computers, which provides the proctor with live or recorded monitoring of the user's identity and activity, preventing impersonation and providing control over background movements (Harish et al., 2021); (ii) audio recording using a microphone, which is also commonly found in modern laptops, allowing the analysis of the audio for biometrics and background sounds (Sinha & Yadav, 2020; Prathish et al., 2016); (iii) involvement of a human proctor, which due to inaccuracy of existing solutions is still needed in the AIPS design model. With the supervision of a human proctor, who oversees the accuracy of the system by manually flagging suspicious behavior, intelligent mechanisms learn to recognize and mark more accurate future suspicious activities (Li et al., 2015); (iv) video recording of the desktop environment, revealing whether the user has any opened tabs on a web browser, ensuring that only the allowed materials are accessed during the examination (Slusky, 2020; Beust et al., 2018); (v) restrictions on applications running on the deskop environment, to ensure users only access applications and/or websites that are allowed during the examination (Slusky, 2020), flagging any forbidden attempts (Metzger and Maudoodi, 2020); (vi) biometric verification, which not only can help to detect possible impersonation threats (Chirumamilla et al., 2020) and improve the security of user authentication (Labayen et al., 2021), but also could automate and support attention tracking, mind wandering and facial behavior analysis (Villa et al., 2020; Blanchard et al., 2014; Baltrušaitis et al., 2016); (vii) eye tracking can prevent malpractice of using external sources of information, such as notes or textbooks (Maniar et al., 2021; Atoum et al., 2017; Li et al., 2015; Villa et al., 2020), albeit with a margin of error allowing users to keep their natural behavior movement; (viii) random question bank methods, in order to provide every individual with a unique paper or set of questions generated only for that user. This parameter can prevent students from trying to share the answers as the examination questions should not repeat among them (Chua et al., 2019; Norris, 2019).

Research-oriented proctoring systems

The importance of student identification and authentication in online activities has been broadly recognized in the academic field, with numerous studies aiming to design a solution supporting user identity verification and authentication. In Table 1, we present an overview of academic driven technologies, that aim to improve credibility of authenticated users and limit the possibilities of most common acts

Table 1 Research-oriented proctoring systems

System properties/identified threats	PS Zhang et al. (2016)	PS Khlifi and El-Sabagh (2017)	PS Musambo and Phiri (2018)	PS Monaco et al. (2013)	PS Atoum et al. (2017)	PS Fenu et al. (2018)
Security Measures	Face Verification, Real-time face tracking for behaviour analysis	Knowledge and behaviour based Verification	Face Verification, QR code authentication	Keystroke and stylometry behavioural analysis	User verification, text detection, voice detection, active window detection, gaze estimation and phone detection	Face, voice, touch, mouse and keystroke verification
Integration with LMSs	✗	✗	✗	✗	✗	✓
Type of Application	Not specified	MATLAB Script	Web app	Not specified	Native, Windows only	Web app and mobile
Student violating identification proofs	✓	✓	✓	✓	✓	✓
Student switching seats after identification	✓	✓	✗	✗	✓	✓
Non-legitimate person provides answers through shared LMS credentials	✗	✗	✗	✗	✗	✗
Computer mediated communication through voice or text-written chat	✗	✗	✗	✗	✓	✓
Computer mediated collaboration through screen sharing and control applications	✗	✗	✗	✗	✗	✗
Student access to forbidden online resources	✗	✗	✗	✗	✓	✗
Non-legitimate person providing answers on the student's computer through the student's main input device or a secondary input device	✗	✗	✗	✗	✗	✗
Student communicating/collaborating with another person within the same physical context	✗	✗	✗	✗	✓	✗
Non-legitimate person providing answers on a white board/computing device/hard-copy messages	✗	✗	✗	✗	✓	✗

PS stands for Proctoring System

of misconduct. A detailed comparison of research-oriented work can be found in Labayen et al. (2021); Guillén-Gámez et al. (2018). The incorporated solutions concentrate their applicability mainly on user authentication aspects, and therefore address only impersonation threats, leaving a space for misconduct related with the communication, collaboration and resource access threats. Among the presented solutions, the Proctoring System (PS) from Atoum et al. (2017) addresses several threat scenarios, by combining the continuous estimation components, and applying application of multimedia (audio and visual) for continuous user verification, gaze estimation, text on printed papers on the computer screen or keyboard text detection, speech detection, active window detection, phone presence and its usage, and cheating behavior detection. Initial results from a study with 24 test takers, which were evaluated during real-world behaviors scenarios in online examinations, show that the capabilities of the designed system demonstrated nearly 87% segment-based detection rate across different types of malicious behavior threats at a fixed “False Alarm Rate” of 2% (Atoum et al. 2017).

Commercial proctoring systems

Initial attempts to introduce commercial online proctoring solutions took place in the mid 2000’s by Kryterion,[4] which involved the engagement of human proctors who monitored online exams via web cameras. The company pioneered the market (Foster & Layman, 2013) giving a lead to the development and application of the online proctoring system. Since then, several other organizations have followed, commercializing proctoring technologies through the support of the authentication process, by monitoring the whole session, and/or recording the data for later session analysis. Table 2 presents a comparison of state-of-the-art commercial online proctoring systems, which were selected based on their infrastructure, user verification features, privacy policy of recorded data, integration with LMSs (OReilly & Creagh, 2016; Foster & Layman, 2013; Atoum et al., 2017; Labayen et al., 2021), and their features to help eliminate a number of commonly faced security threats during an online examination (Labayen et al., 2021; Ullah et al., 2016), such as impersonation, communication, collaboration and resource access threats. Accordingly, Smiley Owl (SMOWL) provides complete support in remote learning activities, potentially eliminating several threats. Moreover, this work has been inspired by a thorough analysis of the current market situation and an evident lack of proctoring systems that guarantee a comprehensive and reliable solution. The authors attempt to combine *“multi-biometric continuous authentication with continuous visual and audio monitoring, with device activity monitoring and lock-down options with human supervision (only when required)”* to fill the gaps in online authentication processes foreseen in remote learning (Labayen et al., 2021).

[4] https://www.kryteriononline.com/.

Table 2 Commercial proctoring systems

System properties/identified threats	ProctorU	Respondus	Proctorio	AIProctor	Kryterion	Examity	SMOWL
Security Measures	Face verification, live human proctor monitoring, unauthorized applications	Browser lockdown, IP tracking, flags suspicious behaviour for post-exam review	Gaze tracking, face detection, device monitoring	Live proctor monitoring, device monitoring	Keystroke and stylometry behavioural analysis, live human proctor monitoring	Face verification, live human proctor monitoring	Face verification, device monitoring, post-exam human proctor review, flags suspicious events in the users' environment
Integration with LMSs	✓	✓	✓	✓	✓	✓	✓
Type of Application	Browser plugin, Windows and MacOS only	Browser plugin	Browser plugin	Mobile and Native, Windows, MacOS and Linux	Native, Windows and MacOS only	Web app	Web app and Native, Windows and MacOS only
Student violating identification proofs	✗	✗	✗	✗	✗	✓	✓
Student switching seats after identification	✓	✓	✓	✓	✓	✓	✓
Non-legitimate person provides answers through shared LMS credentials	✗	✗	✗	✗	✗	✗	✓
Computer mediated communication through voice or text-written chat	✗	✓	✓	✓	✓	✗	✓
Computer mediated collaboration through screen sharing and control applications	✗	✓	✓	✓	✓	✗	✓
Student access to forbidden online resources	✗	✓	✓	✓	✓	✗	✓
Non-legitimate person providing answers on the student's computer through the student's main input device or a secondary input device	✗	✗	✗	✗	✗	✗	✓
Student communicating/collaborating with another person within the same physical context	✓	✓	✓	✓	✓	✓	✓

Table 2 (continued)

System properties/ identified threats	ProctorU	Respondus	Proctorio	AIProctor	Kryterion	Examity	SMOWL
Non-legitimate person providing answers on a white board/ computing device/hardcopy messages	✗	✗	✗	✗	✗	✗	✓

Technologies for continuous user identification

Biometric technologies are being considered lately for student identity management in HEIs, as they provide several advantages over the traditional knowledge-based and token-based authentication methods. While biometric technologies have many benefits from both a security and usability point of view, there is still a need for innovative continuous user identification to authenticate students during academic and teaching activities. User identity management is a critical aspect of any information system today aiming to assure that end-users have the appropriate access to sensitive data and services. Core components of user identity management relate to:

1. *User authentication* aiming to validate that the end-users are allowed to access the system by requiring them to provide various authentication factors, or a combination of them (e.g., textual and graphical passwords, push notifications on smartphones, Time-based One Time Passwords (TOTP), graphical Transaction Authentication Numbers (TAN), biometrics, etc.) (Mare et al., 2016; Ometov et al., 2018; Constantinides et al., 2021);
2. *Continuous user identification* aiming to verify the end-user's identity in real-time (after successfully authenticating), while carrying out tasks (Gonzalez-Manzano et al., 2019; Buschek et al., 2015);
3. *Access control* aiming to regulate user access to the system resources (Rouhani and Deters, 2019).

In this context, biometric-based authentication within user identity management represents a significant and evolving field of research and practice (Bhalla, 2020). Specifically, biometrics can create high entropies of the secret biometric data used for authentication, minimize administration expenses, offer convenience to end-users compared to traditional knowledge-based (e.g., passwords) and token-based (e.g., TOTP) solutions, and they provide a sense of technological modernity to the end-users (Leaton, 2017; Pagnin & Mitrokotsa, 2017). Common approaches for biometric-based authentication are based on the end-users' physical (e.g., fingerprint, iris, face, voice, etc.) and/or behavioral characteristics (e.g., typing patterns, interaction patterns, engagement patterns, etc.) (Jain et al., (2016; Rui & Yan, 2018). Such technologies have become an important means for enforcing strict security policies in a variety of domains, such as government, education, etc. (Bhalla 2020; Leaton, 2017; Labati et al., 2016).

HEIs already started considering the adoption of biometric technology for seamlessly identifying and/or authenticating students within teaching and learning activities, academic services, etc. This is a key aspect in ensuring the development of suitable procedures and contexts that prevent students from engaging in malicious activities, including prohibited communication and collaboration among students, as well as impersonation cases (i.e., intentionally pretending someone's identity in order to unethically participate in academic activities).

Through the LMS, students commonly authenticate using a single-point textual password system, which assumes integrity of the student's attendance within the whole academic activity. As for video conferencing tools, student identification and verification procedures are primarily conducted manually through human intervention (Fidas et al., 2023); for instance, instructors or invigilators visually compare the student's identity with an identity (ID) card or student card, which is presented by the student using a web camera. The approaches mentioned earlier fall short in detecting fraudulent student activities after the single entry-point of authentication has been performed successfully (Frank et al., 2012). In addition, manual and individual student identification is time-consuming, adds low value and presents a difficult endeavor for instructors throughout the whole academic activity. This also brings consequences at the level of assuring a satisfactory implementation of the HEI's curriculum and a fair students' evaluations process.

The literature is rich in solutions for image-based identification and voice-based identification. Below, we review relevant techniques for biometric identification, focusing firstly on single-modality identification through image and voice, and then addressing alternative interaction-based identification and the combination of biometric traits afterwards.

Image-based identification

We start by looking into image-based identification and specifically *face recognition*.

Face recognition is a contactless biometric technology that matches (authenticates) images or videos of human faces against a database of known individuals.

Such systems are widely used in a broad range of applications, such as access control, financial services (validating transactions), video surveillance, law enforcement, social media, smart advertising, automotive industry, etc.

In general, an image-based face recognition framework follows a pipeline that requires three main modules: (1) face detection, (2) image registration/normalization and (3) classification. The first stage, as the name implies, is used to localize the face in the image. Typically, it is a process that exhaustively scans the input image and returns the bounding box that covers the face region. The registration step deals with normalization tasks. Most approaches rely on non-rigid image aligning techniques to locate a set of facial landmark features, while others use simple image warping methods. Regardless, the main goal is to establish a common reference frame, sometimes known as the canonical reference frame, ensuring that a normalized face always has the same spatial dimensions. The last stage is responsible for the recognition/classification itself. It takes a normalized (warped) version of the face, from previous stages, and makes use of

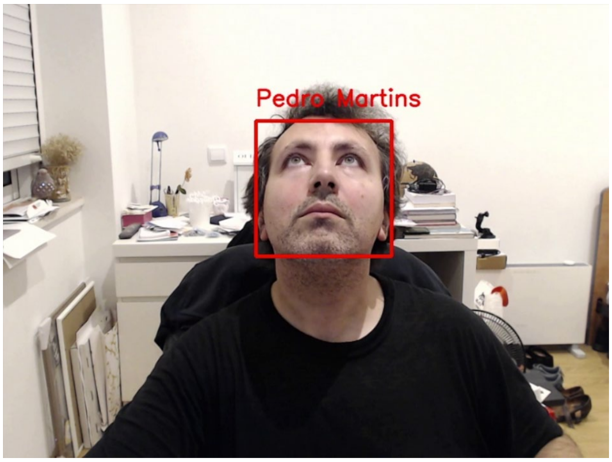

Fig. 1 Illustration of a face recognition system under ongoing development within the TRUSTID EU research project (https://trustid-project.eu/) Faria et al. (2023)

pattern recognition and machine learning techniques to infer the target individual, as illustrated in Fig. 1. In this section, we primarily focus on the last recognition module.

Facial recognition has been a widely studied research topic in the computer vision community. A substantial body of work has been reported in the past, with varying degrees of success. In fact, the consistent identification of human faces in unconstrained environments is still an open issue. Note that this is no easy task, as such systems need to account for variations in facial appearances, expression, ageing, motion and orientation (3D pose), as well as image acquisition issues, e.g., lighting, occlusion, resolution, focus and motion blur.

Early approaches, rely on low-dimensional representations of the (normalized) face. Linear models such as Turk and Pentland (1991a, 1991b) or Belhumeur et al. (1997) use Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), respectively. Later, manifold-based face embedding learning was proposed, and methods such as He et al. (2005) or Dornaika et al. (2015) employ locality preserving projection, i.e., a linear approximation of Laplacian Belkin and Niyogi (2003). Still under the low-dimensional design, sparse coding-based approaches (Wright et al., 2009; Zhang et al., 2011; Yang et al., 2011) have also been proposed.

Nonlinear learning based methods, such as kernel PCA Kim et al. (2002), Support Vector Machine (SVM) Heisele et al. (2001), Boosting Guo and Zhang (2001) or Random Forests Kremic and Subasi (2016); Mady and Hilles (2018), were also applied to face recognition tasks. Despite the relative success of the previous approaches, in general they fail under target images that differ from seen data—the so-called training set. Feature-based techniques (local texture descriptors) were later introduced to mitigate this issue. Popular approaches in facial appearances include Local Binary Patterns (LBP) Ahonen et al. (2006), Discrete Cosine Transform (DCT) Hafed and Levine (2001) or Gabor filters Liu and Wechsler (2002). These techniques are usually applied to extract local face characteristics and they can be combined with any of the previous linear or nonlinear learning strategies.

After the Krizhevsky et al. (2012) technique won (by a large margin) the ImageNet Deng et al. (2009) object detection challenge in 2012, Deep Convolutional Neural Networks (DCNNs) have become a reference to solve many computer vision problems, including face recognition.

In short, DCNNs are a cascade of multiple network layers, where convolution operations, activation functions and pooling are the basic building blocks. The configuration of blocks and layers defines the architecture of the CNN. Convolution layers aim to extract features (i.e., to learn a set of weights or filters). These layers are usually followed by an activation function that applies nonlinear transformations, constraining the outputs of the filter responses. Pooling is a nonlinear down-sampling strategy that expands the receptive field and reduces the number of parameters, and hence the overall computation cost. In these settings, the loss function needs to be specified. This function, defined during the learning/training stage, measures the error in the network's prediction compared to the ground truth.

The DCNNs are compelling techniques that allow learning their own feature representations (Yi et al., 2014), while simultaneously leveraging large amounts of data, thus reinforcing discriminability. It is worth mentioning that DCNNs are highly parallelizable and therefore can be accelerated by Graphics Processing Unit (GPU) computation. A notable approach in deep face recognition is the Taigman et al. (2014) network, which achieved human-level performance, for the first time, in the Labeled Faces in the Wild benchmark (Huang et al., 2007). DeepFace is a 9-layer CNN (120 M parameters) that utilizes several locally connected layers without weight sharing, whose input is a normalized warped image obtained through a 3D nonrigid alignment process.

Currently, the most common face recognition techniques rely on the He et al. (2016) architecture and its variants. Several reasons contribute to this. Firstly, ResNet has demonstrated exceptional performance in the 2015 Deng et al. (2009) and Microsoft Common Objects in Context Lin et al. (2014) challenges. Secondly, the authors of Cao et al. (2018) have used the ResNet-50 architecture to successfully validate face recognition performance in their proposed dataset. Finally, and perhaps one of the key aspects, is the performance/computation effort ratio. ResNet can deliver satisfactory results while being relatively lightweight. Its computational advantage relates to the single fully connected layer that is used at the end of the network, as opposed to the popular Simonyan and Zisserman (2015) network, which includes 3 large fully connected layers.

Later, other DCNNs research directions were exploited. Several networks architectures aimed to improve the performance by learning deeper features (Zheng et al., 2016; Simonyan & Zisserman, 2015; Gruber et al., 2017). Alternatively, other solutions aim to learn embeddings directly through metric learning instead of training a multi-class classifier. Schroff et al. (2015) introduced the triplet loss, which enforces faces within the same class to be closer to each other than to faces from different classes (with a soft margin). Parkhi et al. (2015) follows a similar approach and proposes a fine-tuned version of FaceNet.

Alternative solutions propose enhanced loss functions, such as the center loss (Wen et al., 2016), the Wang et al. (2017) (cosine similarity), the Liu et al. (2017) (angular softmax) or the Deng et al. (2019) (additive angular margin).

Voice-based identification

Research on digital audio processing, voice and speech understanding and computational linguistics has focused on two main fronts: *speech recognition* and *speaker recognition* (Peacocke & Graf, 1995).

Speech recognition involves techniques and methodologies aimed at achieving close to real-time recognition of speech by a computer (Vicens, 1969). It allows translation of spoken language into text (Dimauro et al., 2017), thus it is also known as automatic speech recognition (ASR) (Yu & Deng, 2016), computer speech recognition (Zhang & Liu, 2018) or speech-to-text (STT) (Dimauro et al., 2017).

Generically, the main goal of speech recognition is to develop techniques that enable computers to accept speech input (Reddy, 1976; Waibel & Lee, 1990). Research in the topic dates back to 1952 when a rudimentary recognition system called *Audrey*, developed at Bell Labs, was able to identify the first ten English digits (Meng et al., 2012). Since then, speech recognition has evolved into a significant research field.

The work of Huang and Lee (1993); Huang et al. (1993) is considered a major milestone in speech recognition research, marked by the release of the Sphinx-II speech recognizer at Carnegie Mellon University. This system was the first to achieve speaker-independent continuous speech recognition with support for a large vocabulary set of 1000+ words. It utilized dynamic and speaker-normalized features, and for modelling acoustic-phonetic phenomena, it employed semi-continuous Hidden Markov Models (HMMs), senone, and a tree-based allophonic model. It also led to reduced error rates in vocabulary- and speaker-independent speech recognition by supporting speaker adaptation, efficient search, and language modelling.

While there have been predecessor methods, such as Dynamic time warping (DTW)-based speech recognition (Wan & Carmichael, 2005; Amin & Mahmood, 2008), the vast majority of general-purpose speech recognition systems rely on HMMs, a statistical method for speech processing, that employ a Markov state diagram to capture the temporal properties of speech, and a Gaussian mixture model (GMM) to represent the spectral properties of speech (Rabiner, 1989). HMMs are used in a wide range of applications, from isolated word recognition systems to large vocabulary speech understanding systems. The keys to the success of HMMs include automatic training, simplicity and computational feasibility (Benesty et al., 2008). In recent years, with the increase of computational power, HMM recognition has also been combined with neural networks for pre-processing, feature transformation or dimensionality reduction (Hu & Zahorian, 2010). For instance, in Hadian et al. (2018), a simple HMM-based end-to-end method for large vocabulary continuous speech recognition is presented. The authors propose a one-stage training approach using a lattice-free maximum mutual information (LF-MMI) objective function in a flat-start manner, i.e., without running common HMM-GMM training and tree-building pipeline. This method outperforms other state-of-the-art approaches under similar conditions, reducing the word error rates (WER) to 10% to 25% on well-known speech databases.

More recently, benefiting from large training data and faster hardware, researchers have started effectively training deep neural networks for speech recognition, using a larger number of context-dependent output units to improve performance (Hinton et al., 2012; Deng et al., 2013). This has led to the proliferation of applications of deep feed-forward neural networks (DFFNNs) for speech recognition (Nassif et al., 2019), and widening the performance gap between acoustic models based on DFFNNs and those based on GMMs. Nowadays, modern end-to-end automatic speech recognition systems (e.g., from Apple, Google, Amazon and Microsoft) are deployed on the cloud to overcome the impracticality of deployment on personal devices.

In spite of the importance of speech recognition for modern applications such as vehicles, healthcare, military, telephony and others (Vajpai & Bora, 2016) as well as its extensive research footprint, continuous user identity management may particulary benefit from the topic of speaker recognition for voice-based biometrics.

Speaker recognition addresses the process of automatically identifying a speaker from voice samples (Poddar et al., 2018). For this reason, it is also known as voice recognition (Van Lancker et al., 1985).

Speaker recognition allows for verifying the identity of a speaker as part of a security process, using the acoustic features of speech which differ between individuals (Sambur, 1975). Speaker recognition research dates back to the 1960s (Hargreaves & Starkweather, 1963; Atal, 1969) and is traditionally divided into two main axes: Speaker verification (SV) and speaker identification (SI). On one hand, SV addresses the authentication issue of a claimed identity of a person from his/her own voice samples (Kinnunen & Li, 2010), enabling the confirmation of one's identity. On the other hand, SI aims to identify a speaker from a given set of speakers from the input speech signal (Chakroborty & Saha, 2009), allowing, for instance, the determination of an unknown speaker's identity.

The relationship between speech recognition and speaker recognition is clear. For instance, recognized words enable the use of text-dependent speaker modelling techniques. Also, the choice of words or pronunciation can be a useful indicator of speaker identity, as described in Stolcke et al. (2007). In that work, authors survey speaker recognition techniques that make use of speech recognition, e.g., text-dependent modelling and extraction of higher-level features such as speech prosody (supra-segmental pitch, duration, and energy patterns), showing the potential of combining both types of techniques.

Researchers from Google Deepmind teamed up with the French National Centre for Scientific Research (Seurin et al., 2020) to present a new paradigm: Interactive Speaker Recognition (ISR). For this, personalized utterances are requested from users in order to build a representation of the speakers, and the speaker recognition task is then translated to a sequential decision-making problem through a Markov Decision Process, which the authors propose to solve using Reinforcement Learning (RL). The method adapts a standard RL algorithm, which builds an iterative strategy to maximize the identification accuracy, while querying only a few words. Results show that the RL enquirer steadily improves upon training, and it consistently outperforms two non-interactive heuristic baselines, while using minimal speech signal data.

Current state-of-the-art techniques on speaker recognition follow the same trend as in speech recognition, i.e., they mostly rely on machine learning, and specifically

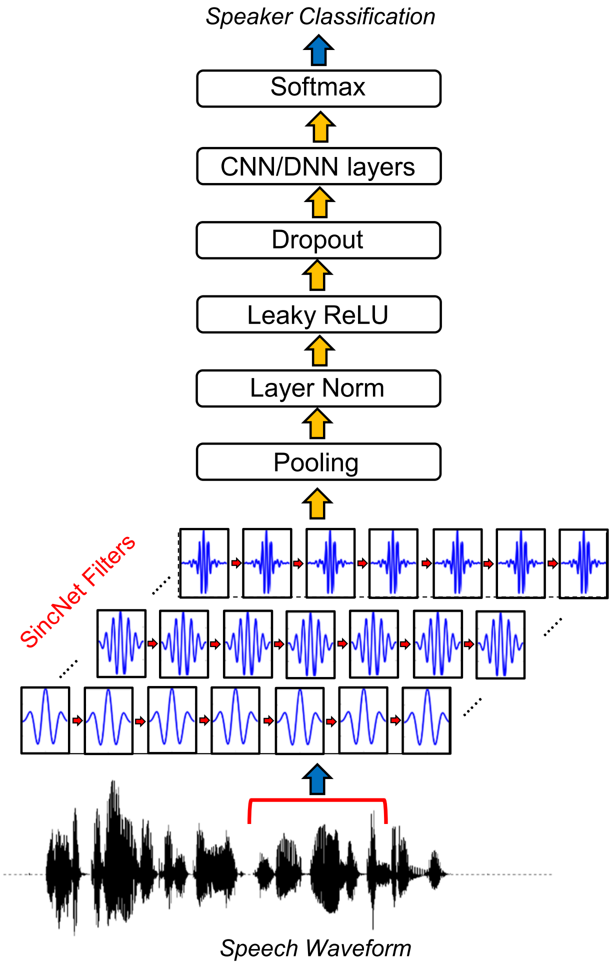

Fig. 2 The SincNet architecture Ravanelli and Bengio (2018). The speech waveform is convoluted with a set of parametrized sinc functions that implement band-pass filters. Then, a standard CNN pipeline (pooling, normalization, activations, dropout) is employed. Multiple standard convolutional, fully-connected or recurrent layers are then stacked to finally perform speaker classification with a softmax classifier

deep learning, which has shown the most promising results for SV and SI within the literature, due to the ability of using big data. In Boles and Rad (2017), a design of a text-independent voice identification system is presented. Audio features are extracted using Mel-Frequency Cepstral Coefficients (Hasan et al., 2004), and the lower 20 coefficients are fed into an SVM neural network algorithm for speaker identification, with an accuracy of around 97% for a full 40-person development set. Ravanelli and Bengio (2018) propose to drastically reduce the number of parameters in the first convolutional layer of a Convolutional Neural Network (CNN) to force the network to focus only on the filter parameters with a major impact on performance for SV and SI tasks. Thus, the architecture proposed, named SincNet (see Fig. 2), converges faster than standard CNNs by discovering more meaningful filters in the input layer, outperforming other speaker identification systems (sentence error rates $<$ 1%) and speaker verification systems (equal error rate $\simeq$ 0.5%) on a variety of datasets. Focusing on very challenging noisy and unconstrained conditions, citenagrani2020voxceleb present Voxceleb: a large-scale audio-visual dataset using a fully automated pipeline, which obtains videos from

YouTube and performs active speaker verification using a two-stream synchronization CNN and confirms the identity of the speaker using CNN-based facial recognition. The resulting dataset consists of over a million utterances from over 6000 gender-balanced speakers. The authors then compare different CNN architectures and new training strategies to identify voice under various conditions and conclude that the proposed trained relation module added into a Siamese network outperforms the CNN and non-CNN architectures tested.

Interaction-based identification and combination of biometric traits

Several works have focused on identifying users based on alternative methods to image and voice-based identification, via their behavioral traits and interactions (Lamiche et al., 2019; Bailey et al., 2014; Shen et al., 2015; Draffin et al., 2013). In particular, a large body of research has investigated a behavioral biometric-based technique known as keystroke analysis (Bergadano et al., 2002), which captures users' typing characteristics during keyboard interactions and uses them for authentication purposes. A feasibility study in Clarke and Furnell (2007) investigated two types of common interactions on mobile phones (i.e., entering telephone numbers and typing text messages) and revealed that neural network classifiers can be used for authenticating users based on their typing characteristics on mobile phone keypads. In Tse and Hung (2019), the authors presented an authentication scheme for touchscreen mobile devices that combines the user's password with features extracted from typing and swiping patterns, revealing that the combined behavioral biometric features enhance the performance of user identification on mobile devices compared to the use of a single set of features.

Keystroke dynamic-based analysis has also been used in graphical-based authentication systems (Chang et al., 2012), by capturing the time and pressure features when users enter their graphical password on a touch screen mobile device, revealing appropriate performance and suitability for even low-power mobile devices. Furthermore, the analysis of touch interactions in handheld devices, namely touch dynamics, has been used for user identification. A study conducted in Sandnes and Zhang (2012) proposed an approach for identifying users based on touch dynamics by considering touch features (e.g., left-hand *vs.* right-hand dominance, one-handed *vs.* bimanual operation, gesture size, gesture timing), revealing the effectiveness of using touch dynamics for successfully identifying users. Moreover, touch dynamics have been used for continuous user authentication by considering schematic and motor-skill touch features (Shen et al., 2015).

Other works focused on identifying users by analyzing mouse interaction behaviors, which can be broadly categorized based on the type of authentication (i.e., static authentication and active re-authentication (Shen et al., 2017)). Static authentication usually checks the authenticity of the user once during the login process and requires users to perform pre-defined mouse actions that will be compared with the legitimate user's profile (Sayed et al., 2013; Shen et al., 2012), while active re-authentication operates continuously by acquiring mouse data during the user's interactions and implicitly verifying the continued presence of the user (Zheng et al., 2016; Mondal and Bours, 2013; Shen et al., 2012).

Also, behavior-based methods that rely on mobile application usage have been used for authentication purposes (Ashibani & Mahmoud, 2019). Examples include identification of users and detection of anomalies based on users' interaction with their mobile applications, the use of text messages and calling behavior (Li et al., 2011), implicit or continuous authentication based on the user's habits and activities with respect to text messages, phone calls, browser history, and location (Shi et al., 2010), and behavioral profiling that authenticates users based on historical application usage (Li et al., 2014). Furthermore, recent works revealed that the generated traffic during accessing mobile applications and the time of accessing these applications can be used to effectively identify users (Ashibani et al., 2018).

Numerous works on continuous or implicit authentication methods have been proposed as an additional non-intrusive security countermeasure (Frank et al., 2012; Shahzad & Singh, 2017; Dzulkifly et al., 2020; Rathgeb et al., 2020). However, existing solutions that simply monitor face and/or body cues are not adequate to prevent fraudulent behavior in online examinations, since they lack students' interactions and are not able to capture scenarios in which the camera stream switches over to other video sources (Fenu et al., 2018). Other works have focused on the combination of multiple biometric traits. Kaur et al. (2016) discuss and address strong authentication mechanisms through biometrics and propose the idea of a framework model that combines voice recognition with typing pattern/keystroke mechanism recognition to develop an advanced and more secure way to authenticate users in e-learning systems. Focusing on continuous authentication, in Prakash et al. (2020) multimodal biometric traits considering finger and iris print images are extracted to enforce higher security and combined using an optimal feature level fusion (FLF) process. Results report a 92% accuracy for the proposed model when compared to other techniques. Moini and Madni (2009) examine the problem of remote authentication in online learning environments and analyze the challenges of using biometric technology to defend against user impersonation attacks by certifying the presence of the user in front of the computer at all times. They design a client–server architecture for continuous user authentication through combined continuous facial recognition with periodic fingerprint matching to verify the identities of its users. Combining fingerprint with mouse patterns for authentication is discussed in Asha and Chellappan (2008). The authors propose using a multimodal physiological (user fingerprint) and behavioral (mouse dynamics) biometric approach. For mouse dynamics, the authors aim to evaluate mouse movement speed, movement direction, action type, traveled distance, and elapsed time. However, no details about an actual implementation of the said system are given.

A common weakness of works that use multiple-biometric solutions is that they often operate in an intrusive fashion that interferes with students' activities and requires additional devices. However, these methods offer effective means to prevent and protect against impersonation attacks by unauthorized users. In contrast to the existing one-time authentication methods, continuous authentication goes a long way to ensure that the intended user is present in front of the workstation at all times. However, it cannot detect or prevent fraudulent behavior on the part of the authorized user, nor does it guarantee that only the authenticated and authorized user is present in the same room (Moini & Madni, 2009).

Existing proctoring tools (see "Intelligent online proctoring systems" section) are usually used only during examinations, tend to not consider the rest of the course

Table 3 Main features of the most relevant user identification works surveyed

Work	ID type	Core method	Offline	Computation cost	ID precision	Needed hardware
Belhumeur et al. (1997)	Image-based	LDA	✓	⋆	⋆	Webcam
He et al. (2005)	Image-based	LPP + Nearest-neighbor classifier	✓	⋆⋆	⋆⋆	Webcam
Ahonen et al. (2006)	Image-based	LBP + Bayesian classifier	✓	⋆⋆	⋆⋆⋆	Webcam
Heisele et al. (2001)	Image-based	SVM	✓	⋆⋆	⋆⋆⋆	Webcam
Cao et al. (2018)	Image-based	ResNet50 CNN (DL)	✓	⋆⋆⋆	⋆⋆⋆⋆	Webcam+GPU
Taigman et al. (DeepFace) (2014)	Image-based	DeepFace CNN (DL)	✓	⋆⋆⋆⋆⋆	⋆⋆⋆⋆⋆	Webcam+GPU
Taigman et al. (FaceNet) (2014)	Image-based	Inception CNN (DL) + Distance threshold	✓	⋆⋆⋆⋆	⋆⋆⋆⋆⋆	Webcam+GPU
VGGFace Parkhi et al. (2015)	Image-based	VGGFace CNN (DL)	✓	⋆⋆⋆⋆	⋆⋆⋆⋆⋆	Webcam+GPU
Stolcke et al. (2007)	Voice-based	MLLR transforms + SVM	✓	⋆⋆⋆	⋆⋆⋆	Microphone
Seurin et al. (2020)	Voice-based	MDP + RL	✓	⋆⋆⋆	⋆⋆⋆⋆	Microphone
Boles and Rad (2017)	Voice-based	MFCCs + SVM NN (DL)	✗	⋆⋆⋆⋆⋆	⋆⋆⋆⋆⋆	Microphone
Ravanelli and Bengio (2018)	Voice-based	Band-Pass Filters + CNN (DL)	✓	⋆⋆⋆⋆	⋆⋆⋆⋆	Microphone
Nagrani et al. (2020)	Voice-based	Relation Module + Two-stream synchronization CNN (DL)	✓	⋆⋆⋆⋆⋆	⋆⋆⋆⋆⋆	Microphone
Bergadano et al. (2002)	Interaction-based	Keystroke tri-graph duration	✓	⋆	⋆⋆⋆	Keyboard
Clarke and Furnell (2007)	Interaction-based	FF MLP Neural Network	✓	⋆⋆⋆⋆	⋆⋆⋆	Mobile Phone-Handset
Shen et al. (2015)	Interaction-based	Feature-distance vectors + SVM	✓	⋆⋆	⋆⋆⋆	Touch-basedSmartphone
Zheng et al. (2016)	Interaction-based	Mouse angle-based metrics + SVM	✓✗	⋆	⋆⋆⋆	Mouse
Chang et al. (2012)	Password and interaction-based	Graphical password + Touch time and pressure statistical classifier	✓	⋆	⋆⋆⋆⋆	Touch-based Mobile Device

Computation cost and Identification (ID) precision rated as Very Low (⋆), Low (⋆⋆), Medium (⋆⋆⋆), High (⋆⋆⋆⋆) or Very High (⋆⋆⋆⋆⋆)

CNN: Convolutional neural network

DL: Deep learning

FF MLP: Feed forward multi-layered perceptron

LDA: Linear discriminant analysis

LBP: Local binary patterns

LPP: Locality preserving projection

MDP: Markov decision process

MFCC: Mel-frequency cepstral coefficient

MLLR: Maximum likelihood linear regression

RL: Reinforcement learning

SVM: Support vector machine

participation and other important threats related to collaboration and communication with other people, and access to resources. They are not scalable, require technological infrastructure and setup [20], and fall short in adequately addressing privacy concerns related to the recorded videos [20–22]. Specifically, automatic proctoring solutions also exhibit variability in the accuracy of the algorithms and the limited scenarios that they support (Fenu et al., 2018).

Recent advances in biometrics for authentication include touchless fingerprint scanners, in-display fingerprint readers, fingerprint on card, 3D facial recognition, long-range iris recognition, spoof and liveness detection software, and improved machine learning systems that surpass the accuracy of traditional pattern recognition biometric systems (Bhalla, 2020). Despite their potential, these methods often require high-end computing devices and/or additional specific hardware, which are not commonly available in the HEIs domain.

Table 3 summarizes the main features of the most relevant user identification implementations reviewed in this section.

Data privacy-preservation issues

Student identity management in HEIs presents several challenges in preserving the privacy of stored biometric data concerning end-users. The key threats and challenges associated with designing secure and privacy-preserving biometric technologies, as discussed in previous studies (Jain et al., 2016; Rui & Yan, 2018; Pagnin & Mitrokotsa, 2017; Tran et al., 2021), include the following: (i) security of biometric data: in many cases cases, biometric data is not kept secret (e.g., fingerprints can be obtained from surfaces touched by the user, faces can be easily acquired from public online sources, voices can be recorded, etc.) (Rui & Yan, 2018); (ii) privacy of biometric data: stored biometric data has the potential to reveal sensitive information about end-users, such as ethnic origin and health details (Pagnin & Mitrokotsa, 2017); and (iii) revocability of biometric data: if biometric data is compromised, revoking the data of end-users becomes extremely difficult (Rui & Yan, 2018).

Therefore, there is a pressing need to implement and deploy innovative solutions that ensure the secure processing and storage of biometric data while maintaining high levels of security and privacy. State-of-the-art approaches (refer to (Jain et al., 2016; Rui & Yan, 2018; Pagnin & Mitrokotsa, 2017; Sarier, 2018; Tran et al., 2021)) for privacy preservation in biometric-driven data include the use of *biometric templates*. These are digital representations of specific features extracted from a biometric sample, such as the shape of a user's hand, without storing the exact raw biometric data. This approach avoids potential privacy issues if the data set is compromised. A widely used approach for preserving the privacy of biometric templates involves transforming the biometric template into a new domain through a non-invertible integration of biometric data with externally generated randomness, which provides protection similar to a cryptographic cipher (Teoh et al., 2006). Various non-invertible transformation methods have been proposed, including Cartesian, polar, and surface folding (Ratha et al., 2007). Moreover, the use of biometric cryptosystems, which associate a key with the biometric template (Uludag et al., 2004; Cavoukian et al., 2008), can be combined with neural networks-based transformation approaches (Kumar Jindal et al., 2018; Pandey et al., 2016; Jindal et al., 2019). However, these methods often store the biometric templates in an unprotected manner and are thus vulnerable to attacks.

Furthermore, *biometric encryption techniques* have been employed to address privacy concerns in biometrics. Traditional cryptographic hashing approaches may not be suitable for biometric data due to its high variability (Pagnin & Mitrokotsa, 2017). Hence, different cryptographic such as homomorphic encryption have been applied to protect biometric templates. In this approach, the encrypted biometric template is stored in the database and during verification, the matching module calculates the similarity score between the encrypted stored template and the encrypted query template (Jindal et al., 2020; Boddeti, 2018). Nevertheless, there is a tradeoff between matching performance, privacy, and computational cost, as the feature extraction methods and template protection methods have been developed independently.

Protocol-based approaches have also been proposed to safeguard the privacy of biometric data, e.g., secure multiparty computation (SMC) protocol, zero-knowledge proof (ZKP) protocol, etc. (Tran et al., 2021). SMC protocols are cryptographic protocols that preserve the privacy of each participant and can be utilized in privacy-preserving biometric systems (Bringer et al., 2013; Chun et al., 2014; Tian et al., 2018). An overview of SMC's application in privacy-preserving biometric systems, focusing on secure face identification and secure distance computation for fingerprints and iris is presented in Bringer et al. (2013). ZKP protocols, on the other hand, enable a user to prove certain knowledge to the verifier without revealing any additional information and can also be employed in privacy-preserving biometric systems (Bhargav-Spantzel et al., 2010), (Gunasinghe & Bertino, 2017).

Finally, *distributed ledger technologies* (e.g., private blockchain technologies) possess specific features that can address several challenges in privacy-preserving biometrics. Their distributed nature helps overcome single points of failure, eliminate the need for third parties and mitigate potential privacy breaches. They also facilitate monitoring and access to trustworthy and unmodifiable history logs (Rouhani & Deters, 2019; Sarier, 2018; Zhang et al., 2019; Tran et al., 2021). A biometric recognition architecture, which utilizes a private blockchain for feature extraction and performs decentralized matching is presented in Goel et al. (2019). Recent blockchain-based works in the literature for privacy preservation of biometrics include a protocol for decentralized storage of biometric credentials using decentralized identifiers and W3C Verifiable Claims (Othman & Callahan, 2018), as well as methods for protecting fingerprint templates using blockchain technology, which involves extracting fingerprint features, encrypting them with a AES block cipher, and uploading them to a symmetric distributed storage system (Acquah et al., 2020).

Research gaps, open issues and opportunities

Despite significant progress in identity management for distance learning, the current scientific and technological landscape still allows for substantial future work in deploying systems in practical application scenarios for HEIs. In the remainder of this section, we present a number of specific issues where we believe further improvements can be made to enhance the state of the art.

Clear need for adoption of continuous user identification technology for HEIs

Continuous user identification is a critical technology for HEIs as it aims to verify the identity of the end-users in real-time (after successfully authenticating), while they are performing tasks. This technology plays a crucial role in ensuring that HEIs can provide credible, trustworthy, and accurate degrees to their students, thereby sustaining their credibility in society. The COVID-19 pandemic has highlighted the importance of promoting best practices and learning from both positive and negative experiences during this period of intensified distance learning.

HEIs should prioritize the deployment of secure, trustworthy, and credible continuous student identity management solutions that can adapt to various online educational models (synchronous *vs.* asynchronous) and teaching approaches (structured *vs.* unstructured), taking into account the diverse needs and preferences of students and teachers. Relevant opportunities include the integration of such technologies in LMSs and addressing the limitations of existing proctoring systems (see "Intelligent online proctoring systems" section).

Lack of solutions that combine multiple inputs under an agile system integration model

Several literature works propose solutions based on the analysis of users' interaction behavior analysis on both desktop computers and smartphones (Buschek et al., 2015; Gascon et al., 2014), physiological data analysis including body signals (heart rate, skin conductance, etc.) Ometov et al. (2018); Rui and Yan (2018), electrocardiographic data (Silva et al., 2011), face biometrics (Dabbah et al., 2007), and eye gaze analytics (Jain et al., 2004). Various organizations and companies, such as Acceptto and Veridiumid, are adopting continuous user identification. However, these are far from fully practical solutions as they do not combine multiple sources of input (e.g., face, voice, interaction behavior) within an agile system integration model. Instead, existing solutions are mostly dedicated and favor a certain user feature within certain interaction systems. So far, to the best of our knowledge, there is a lack of frameworks that address continuous student identity management for HEIs, combining face, voice, and interaction behavior of users for continuous identification (see "Technologies for continuous user identification" section). Such frameworks would build a more credible and comprehensive user model while enabling continuous user identification under a unified, agile system integration model, bootstrapped on synchronous and asynchronous online teaching and learning activities.

Break away from traditional user authentication, which compromises student's continuous identification

The current state-of-the-art online education LMS of HEIs currently compromises students' continuous identification by relying on traditional user authentication methods (e.g., passwords). Innovative and credible identity management methods for continuously identifying students during online learning activities are needed in order to adopt more secure institutional and instructional strategies for continuous student identification management and develop new competencies in novel methods. In practical terms, this requires reliable and secure solutions for processing and storing biometric data with high levels of security and privacy (see "Data privacy-preservation issues" section).

Opportunities include detecting fraudulent student activities after the single entry-point of authentication has been performed and removing the need for instructors to manually confirm individual students' identification, a practice that is unfortunately still common in critical online learning activities nowadays.

Trade-off between local-based and online-based identification systems

It is clear from the literature that intelligent biometrics, based on a combined analysis of face, voice, and interaction behavior data analytics, will advance the state-of-the-art for seamlessly identifying users. Popular methods, particularly those based on machine learning, and specifically deep learning, are consistently improving the accuracy of user detection and recognition every day. However, these methods are also data-hungry and computationally expensive. In practical terms, the deployment of continuous user identification systems needs to be grounded on a system architecture that is carefully designed and which does not compromise usability. Offline recognition solutions require the availability of computational resources on the end-users workstation, while improving privacy and keeping infrastructure costs low. On the other hand, online verification solutions, such as cloud-based architectures, allow continuous identification of users on virtually any terminal, e.g., smartphones, but raise concerns about the transmission of sensitive data over the network, increased costs, and scalability issues. Research opportunities exist in studying and testing hybrid architectures that aim to exploit the advantages of both online and offline identification systems.

Data protection barriers for adoption

Collecting biometric data from students raises various data protection concerns. For instance, in the European Union, HEIs need to comply with the General Data Protection Regulation (GDPR) when implementing continuous user identification systems. Impact assessment of privacy risks must be carried out to safeguard sensitive data, and obtaining individual student consent is currently a minimal prerequisite to make such systems a reality. However, regulations become a barrier when no consent is granted. In addition, other sensitive scenarios must be accounted for, such as the case of an online examination, where the student should not be denied access due to the absence or a malfunction of a web camera. Alternatives must be provided within the existing regulations. The need to legislate more clearly the use of biometrics in HEIs is a clear opportunity, so that guidelines from national data protection commissions may allow that in some situations biometrics could be used without the student's consent. Encouragingly, promising preliminary groundwork is being done, such as exceptions in biometric registration control for HEIs while strictly adhering to Article 9 of the GDPR.

Conclusion

In this work, we have addressed continuous user identification from a technological perspective, focusing on the unique requirements of distance learning. To achieve this, we have provided an overview of the prevailing intelligent online proctoring systems and automated identification methods based on image, voice and interaction analysis. Furthermore, relevant points, such as the use of biometrics in higher education and data privacy-preservation issues have been highlighted in order to elicit research gaps, open issues and prospects for the advancement of the continuous student identification systems of the future.

We strongly believe that this study paves the way for the **design of an innovative framework for student identity management**. This study will utilize privacy-preserving techniques for face-, voice- and interaction-based continuous user identification. The ultimate goal is to deploy this framework in HEIs using a unified, agile system integration model. Future endeavors include the development of the mentioned system following a User-Centered Design (UCD) methodology grounded on case studies validation in three distinct Universities in Europe. Goals include the demonstration of the technology through dissemination activities, and designing guidelines to aid the institutional, personal and technological transition towards more sophisticated online student management solutions, making the system accessible via an open-source software toolkit.

We anticipate that our work can contribute to fostering trust in HEIs that pursue an online academic strategy. In fact, one of the most important missions and strategic objectives of HEIs is to verify that each single graduate has gone through a credible academic process (e.g., laboratories, examinations, class attendance, etc.) and has therefore acquired the necessary knowledge and competence in order to provide their services to society. Therefore, we argue that a continuous user identification system fills an important gap in the current working dynamics between HEIs, their students and society.

The expected impact of a continuous user identification solution is that HEIs will enhance their digital readiness by offering inclusive, trustworthy, and credible online education activities through the provision of innovative and open-source solutions for continuous student identification and presence awareness. The open-source policy allows HEIs to customize the solutions according to their specific requirements and needs, thereby increasing the sustainability of the results and the tools produced.

As an indirect consequence, this allows HEIs to conduct a self-assessment of their current institutional strategy for online student identification and engage in self-reflection of current practices to identify areas for improvement and adapt them to their needs and requirements.

Within online learning contexts, nearly every student owns several password-protected accounts. Clearly, the identity management market represents one of the largest Information Technology (IT) markets to day, and future plans in this area present promising opportunities for exploitation.

Abbreviations

AES	Advanced encryption standard
AI	Artificial intelligence
AIPS	AI-based proctoring systems
ASR	Automatic speech recognition
CNN	Convolutional neural network
COVID-19	Coronavirus disease 2019
DCNN	Deep convolutional neural network
DFFNN	Deep feed-forward neural network
DTW	Dynamic time warping (DTW)
FF MLP	Feed forward multi-layered perceptron
FLF	Feature level fusion
GDPR	General data protection regulation
GMM	Gaussian mixture model
GPU	Graphics processing unit
HEI	Higher Education Institution
HMM	Hidden Markov model
ICT	Information and communications technology

ID	Identity
ISR	Interactive speaker recognition
IT	Information technology
LBP	Local binary patterns
LDA	Linear discriminant analysis
LF-MMI	Lattice-free maximum mutual information
LPP	Locality preserving projection
LMS	Learning management system
MDP	Markov decision process
MFCC	Mel-frequency cepstral coefficient
MLLR	Maximum likelihood linear regression
MOOC	Massive open online course
OPS	Online proctoring systems
PCA	Principal component analysis
PS	Proctoring system
RL	Reinforcement learning
SI	Speaker identification
SMC	Secure multiparty computation
SMOWL	Smiley owl
STT	Speech-to-text
SV	Speaker verification
SVM	Support vector machine
TOTP	Time-based one time passwords
TAN	Transaction authentication numbers
UCD	User-centered design
W3C	World wide web consortium
WER	Word error rates
ZKP	Zero-knowledge proof

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, analysis and first draft was written by DP, JNF significantly contributed to the writing. MB, PM, AC and AP also contributed to the writing. All authors read and approved the final manuscript.

Funding

This research work has been supported by the European project TRUSTID - Intelligent and Continuous Online Student Identity Management for Improving Security and Trust in European Higher Education Institutions (Grant agreement no: 2020-1-EL01-KA226-HE-094869), which is funded by the European Commission within the Erasmus+ 2020 Programme and the Greek State Scholarships Foundation I.K.Y. David Portugal also acknowledges the Portuguese Science Agency "Fundação para a Ciência e a Tecnologia" (FCT) for support under the Scientific Employment Stimulus 5th Edition, contract reference 2022.05726.CEECIND.

Availability of data and materials

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors

Competing interests

The authors declare that they have no conflict of interest.

Received: 9 May 2023 Accepted: 16 July 2023
Published online: 26 July 2023

References

Acquah, M. A., Chen, N., Pan, J.-S., Yang, H.-M., & Yan, B. (2020). Securing fingerprint template using blockchain and distributed storage system. *Symmetry, 12*(6), 951.

Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(12), 2037–2041.

Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2017). Examining the effect of proctoring on online test scores. *Online Learning, 21*(1), 146–161.

Alexandru, A., Tirziu, E., Tudora, E., & Bica, O. (2015). Enhanced education by using intelligent agents in multi-agent adaptive e-learning systems. *Studies in Informatics and Control, 24*(1), 13–22.

Amin, T.B., & Mahmood, I. (2008) Speech recognition using dynamic time warping. In *2008 2nd international conference on advances in space technologies*, pp. 74–79 . IEEE

Arkorful, V., & Abaidoo, N. (2014) The role of e-learning, the advantages and disadvantages of its adoption in Higher Education. CRC Publications

Asha, S., & Chellappan, C. (2008) Authentication of e-learners using multimodal biometric technology. In *2008 international symposium on biometrics and security technologies*, pp. 1–6. IEEE

Ashibani, Y., & Mahmoud, Q.H. (2018) A behavior profiling model for user authentication in iot networks based on app usage patterns. In *IECON 2018-44th annual conference of the IEEE industrial electronics society*, pp. 2841–2846. IEEE

Ashibani, Y., & Mahmoud, Q.H. (2019) A behavior-based proactive user authentication model utilizing mobile application usage patterns. In *Canadian conference on artificial intelligence*, pp. 284–295. Springer

Atal, B. S. (1969). Automatic speaker recognition based on pitch contours. *The Journal of the Acoustical Society of America, 45*(1), 309.

Atoum, Y., Chen, L., Liu, A. X., Hsu, S. D., & Liu, X. (2017). Automated online exam proctoring. *IEEE Transactions on Multimedia, 19*(7), 1609–1624.

Bailey, K. O., Okolica, J. S., & Peterson, G. L. (2014). User identification and authentication using multi-modal behavioral biometrics. *Computers and Security, 43*, 77–89.

Bakia, M., Shear, L., Toyama, Y., & Lasseter, A. (2012). *Understanding the implications of online learning for educational productivity*. US Department of Education: Office of Educational Technology.

Baltrušaitis, T., Robinson, P., & Morency, L.-P. (2016) Openface: an open source facial behavior analysis toolkit. In *2016 IEEE winter conference on applications of computer vision (WACV)*, pp. 1–10 . IEEE

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(7), 711–720.

Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation, 15*(6), 1373–1396.

Benesty, J., Sondhi, M.M., & Huang, Y. et al.: (2008) Springer handbook of speech processing vol. 1. Springer.

Bergadano, F., Gunetti, D., & Picardi, C. (2002). User authentication through keystroke dynamics. *ACM Transactions on Information and System Security (TISSEC), 5*(4), 367–397.

Beust, P., Duchatelle, I., & Cauchard, V. (2018) Exams taken at the student's home. In *Online, Open and Flexible Higher Education Conference, EADTU 2018*

Bhalla, A. (2020). The latest evolution of biometrics. *Biometric Technology Today, 2020*(8), 5–8.

Bhargav-Spantzel, A., Squicciarini, A.C., Xue, R., & Bertino, E. (2010) Multifactor identity verification using aggregated proof of knowledge. In *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40(4), 372–383

Blanchard, N., Bixler, R., Joyce, T., & D'Mello, S. (2014) Automated physiological-based detection of mind wandering during learning. In *International conference on intelligent tutoring systems*, pp. 55–60. Springer

Boddeti, V.N. (2018). Secure face matching using fully homomorphic encryption. In *2018 IEEE 9th international conference on biometrics theory, applications and systems (BTAS)*, pp. 1–10. IEEE

Boles, A., & Rad, P. (2017). Voice biometrics: Deep learning-based voiceprint authentication system. In *2017 12th system of systems engineering conference (SoSE)*, pp. 1–6 . IEEE

Bringer, J., Chabanne, H., & Patey, A. (2013). Privacy-preserving biometric identification using secure multiparty computation: An overview and recent trends. *IEEE Signal Processing Magazine, 30*(2), 42–52.

Buschek, D., De Luca, A., & Alt, F. (2015) Improving accuracy, applicability and usability of keystroke biometrics on mobile touchscreen devices. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pp. 1393–1402

Cai, H., & King, I. (2020) Education technology for online learning in times of crisis. In *2020 IEEE international conference on teaching, assessment, and learning for engineering (TALE)*, pp 758–763 . IEEE

Cao, Q., Shen, L., Xie, W., Parkhi, O.M., & Zisserman, A. (2018) Vggface2: A dataset for recognising faces across pose and age. In *Proceedings of the IEEE International conference on automatic face and gesture recognition (FG 2018)*, pp. 67–74

Cavoukian, A., Stoianov, A., & Carter, F. (2008) Keynote paper: Biometric encryption: Technology for strong authentication, security and privacy. In *Policies and Research in Identity Management*, pp. 57–77. Springer.

Chakroborty, S., & Saha, G. (2009). Improved text-independent speaker identification using fused MFCC & IMFCC feature sets based on Gaussian filter. *International Journal of Signal Processing, 5*(1), 11–19.

Chang, T.-Y., Tsai, C.-J., & Lin, J.-H. (2012). A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices. *Journal of Systems and Software, 85*(5), 1157–1165.

Chirumamilla, A., Sindre, G., & Nguyen-Duc, A. (2020). Cheating in e-exams and paper exams: The perceptions of engineering students and teachers in Norway. *Assessment and Evaluation in Higher Education, 45*(7), 940–957.

Chua, S.S., Bondad, J.B., Lumapas, Z.R., & Garcia, J.D.L. (2019) Online examination system with cheating prevention using question bank randomization and tab locking. In *2019 4th international conference on information technology (InCIT)*, pp. 126–131. IEEE

Chun, H., Elmehdwi, Y., Li, F., Bhattacharya, P., & Jiang, W. (2014)Outsourceable two-party privacy-preserving biometric authentication. In *Proceedings of the 9th ACM symposium on information, computer and communications security*, pp. 401–412

Clarke, N. L., & Furnell, S. M. (2007). Authenticating mobile phone users using keystroke analysis. *International journal of information security, 6*(1), 1–14.

Coghlan, S., Miller, T., & Paterson, J. Good proctor or " big or brother" ? Ethics of online exam supervision technologies. *Philosophy and Technology*, 1–26 (2021)

Coman, C., Tîru, L. G., Meseşan-Schmitz, L., Stanciu, C., & Bularca, M. C. (2020). Online teaching and learning in higher education during the coronavirus pandemic: Students' perspective. *Sustainability, 12*(24), 10367.

Constantinides, A., Faria, J., Sousak, T., Martins, P., Portugal, D., Belk, M., Pitsillides, A., & Fidas, C. (2023). TRUSTID: Intelligent and Continuous Online Student Identity Management in Higher Education. In *Adjunct proceedings of the 31st ACM conference on user modeling, adaptation and personalization*, pp. 110–114

Constantinides, A., Fidas, C., Belk, M., Pietron, A. M., Han, T., & Pitsillides, A. (2021). From hot-spots towards experience-spots: Leveraging on users's sociocultural experiences to enhance security in cued-recall graphical authentication. *International Journal of Human-Computer Studies, 149*, 102602.

Dabbah, M., Woo, W., & Dlay, S. (2007). Secure authentication for face recognition. In *2007 IEEE symposium on computational intelligence in image and signal processing*, pp. 121–126 IEEE

Deng, J., Dong, W., Richard, S., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *Proceedings of the IEEE conference on computer vision and pattern recognition*

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*

Deng, L., Hinton, G., & Kingsbury, B. (2013) New types of deep neural network learning for speech recognition and related applications: An overview. In *2013 IEEE International conference on acoustics, speech and signal processing*, pp. 8599–8603 . IEEE

Dimauro, G., Di Nicola, V., Bevilacqua, V., Caivano, D., & Girardi, F. (2017). Assessment of speech intelligibility in Parkinson's disease using a speech-to-text system. *IEEE Access, 5*, 22199–22208.

Dornaika, F., Assoum, A., & Ruichek, Y. (2015) Graph optimized laplacian eigenmaps for face recognition. In *Proceedings of the intelligent robots and computer vision XXXII: Algorithms and techniques*, pp. 91–100

Draffin, B., Zhu, J., & Zhang, J. (2013) Keysens: Passive user authentication through micro-behavior modeling of soft keyboard interaction. In: International Conference on Mobile Computing, Applications, and Services, pp. 184–201 . Springer

Dzulkifly, S., Aris, H., & Janahiraman, TV. (2020) Enhanced continuous face recognition algorithm for bandwidth constrained network in real time application. In *Proceedings of the 2020 The 9th international conference on informatics, environment, energy and applications*, pp. 131–135

El Bachari, E., Abelwahed, E., & El Adnani, M. (2010) An adaptive learning model using learner's preference. In *International conference on models of information and communication systems*

Faria, J.N., Portugal, D., Martins, P., Belk, M., Constantinides, A., Pitsillides, A., & Fidas, C. (2023) Image-based Face Verification for Student Identity Management-the TRUSTID Case Study. In *Adjunct proceedings of the 31st ACM conference on user modeling, adaptation and personalization*, pp. 66–71

Fenu, G., Marras, M., & Boratto, L. (2018). A multi-biometric system for continuous student authentication in e-learning platforms. *Pattern Recognition Letters, 113*, 83–92.

Fidas, C. A., Belk, M., Constantinides, A., Portugal, D., Martins, P., Pietron, A. M., Pitsillides, A., & Avouris, N. (2023). Ensuring academic integrity and trust in online learning environments: A longitudinal study of an ai-centered proctoring system in tertiary educational institutions. *Education Sciences, 13*(6), 566.

Fidas, C., Belk, M., Portugal, D., & Pitsillides, A. (2021) Privacy-preserving biometric-driven data for student identity management: Challenges and approaches. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 368–370

Foster, D., & Layman, H. (2013) Online proctoring systems compared. Online verfügbar unter https://ivetriedthat.com/wp-content/uploads/2014/07/Caveon-Test-Security.pdf

Frank, M., Biedert, R., Ma, E., Martinovic, I., & Song, D. (2012). Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE Transactions on Information Forensics and Security, 8*(1), 136–148.

Gascon, H., Uellenbeck, S., Wolf, C., & Rieck, K. (2014). Continuous authentication on mobile devices by analysis of typing motion behavior. Sicherheit 2014–Sicherheit, Schutz und Zuverlässigkeit

Genereux, R., & McLeod, B. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education, 36*(6), 687–704. https://doi.org/10.1007/BF02208251

Goel, A., Agarwal, A., Vatsa, M., Singh, R., & Ratha, N. (2019) Securing cnn model and biometric template using blockchain. In *2019 IEEE 10th international conference on biometrics theory, applications and systems (BTAS)*, pp. 1–7 . IEEE

González-González, C. S., Infante-Moro, A., & Infante-Moro, J. C. (2020). Implementation of e-proctoring in online teaching: A study about motivational factors. *Sustainability, 12*(8), 3488.

Gonzalez-Manzano, L., Fuentes, J. M. D., & Ribagorda, A. (2019). Leveraging user-related internet of things for continuous authentication: A survey. *ACM Computing Surveys (CSUR), 52*(3), 1–38.

Grajek, S. (2021) Educause covid-19 quick poll results: Grading and proctoring. Educause review https://er.educause.edu/blogs/2020/4/educause-covid-19-quickpoll-results-grading-and-proctoring. Accessed **18**

Gruber, I., Hlavac, M., Zelezny, M., & Karpov, A. (2017) Facing face recognition with resnet: Round one. In *Proceedings of the international conference on interactive collaborative robotics*, pp. 67–74

Guillén-Gámez, F.D., García-Magariño, I., & Palacios-Navarro, G. (2018) Comparative analysis between different facial authentication tools for assessing their integration in m-health mobile applications. In *World conference on information systems and technologies*, pp. 1153–1161 . Springer

Gunasinghe, H., & Bertino, E. (2017). Privbiomtauth: Privacy preserving biometrics-based and user centric protocol for user authentication from mobile phones. *IEEE Transactions on Information Forensics and Security, 13*(4), 1042–1057.

Guo, G.-D., & Zhang, H.-J. (2001) Boosting for fast face recognition. In *Proceedings of the IEEE ICCV workshop on recognition, analysis, and tracking of faces and gestures in real-time systems*

Hadian, H., Sameti, H., Povey, D., & Khudanpur, S. (2018). Flat-start single-stage discriminatively trained hmm-based models for ASR. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 26*(11), 1949–1961.

Hafed, Z. M., & Levine, M. D. (2001). Face recognition using the discrete cosine transform. *International Journal of Computer Vision, 43*(3), 167–188.

Hargreaves, W. A., & Starkweather, J. A. (1963). Recognition of speaker identity. *Language and Speech, 6*(2), 63–67.

Harish, S., Rajalakshmi, D., Ramesh, T., Ram, S. G., & Dharmendra, M. (2021). New features for webcam proctoring using python and opencv. *Revista Geintec-Gestao Inovacao E Tecnologias, 11*(2), 1497–1513.

Hasan, M. R., Jamil, M., Rahman, M., et al. (2004). Speaker identification using MEL frequency cepstral coefficients. *Variations, 1*(4), 565–568.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*

He, X., Yan, S., Hu, Y., Niyogi, P., & Zhang, H.-J. (2005). Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(3), 328–340.

Heisele, B., Ho, P., & Poggio, T. (2001) Face recognition with support vector machine: Global versus component-based approach. In *Proceedings of the IEEE international conference on computer vision*

Helms, N. (2021) Against Proctoring Software. https://colab.plymouthcreate.net/2021/04/07/against-proctoring-software/. [Online; accessed 05-November-2021]

Henry, J. V., & Oliver, M. (2021). Who will watch the watchmen? the ethico-political arrangements of algorithmic proctoring for academic integrity. *Postdigital Science and Education* , 1–24.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine, 29*(6), 82–97.

Hu, H., & Zahorian, S.A. (2010). Dimensionality reduction methods for hmm phonetic recognition. In *2010 IEEE international conference on acoustics, speech and signal processing*, pp. 4854–4857. IEEE

Huang, X., Alleva, F., Hon, H.-W., Hwang, M.-Y., Lee, K.-F., & Rosenfeld, R. (1993). The SPHINX-II speech recognition system: An overview. *Computer Speech and Language, 7*(2), 137–148.

Huang, X., & Lee, K.-F. (1993). On speaker-independent, speaker-dependent, and speaker-adaptive speech recognition. *IEEE Transactions on Speech and Audio processing, 1*(2), 150–157.

Huang, R., Liu, D., Tlili, A., Yang, J., & Wang, H. et al.: (2020) Handbook on facilitating flexible learning during educational disruption: The Chinese experience in maintaining undisrupted learning in COVID-19 outbreak. *Smart Learning Institute of Beijing Normal University*, 1–54

Huang, G.B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007) Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst (2007).

Hussein, M. J., Yusuf, J., Deb, A. S., Fong, L., & Naidu, S. (2020). An evaluation of online proctoring tools. *Open Praxis, 12*(4), 509–525.

Jain, A. K., Nandakumar, K., & Ross, A. (2016). 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recognition Letters, 79*, 80–105.

Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 14*(1), 4–20.

Jindal, A.K., Chalamala, S.R., & Jami, S.K. (2019) Securing face templates using deep convolutional neural network and random projection. In *2019 IEEE international conference on consumer electronics (ICCE)*, pp. 1–6 . IEEE

Jindal, A.K., Shaik, I., Vasudha, V., Chalamala, S.R., Rajan, M., & Lodha S. (2020) Secure and privacy preserving method for biometric template protection using fully homomorphic encryption. In *2020 IEEE 19th international conference on trust, security and privacy in computing and communications (TrustCom)*, pp. 1127–1134. IEEE

Kaur, N., Prasad, P., Alsadoon, A., Pham, L., & Elchouemi, A. (2016) An enhanced model of biometric authentication in e-learning: Using a combination of biometric features to access e-learning environments. In *2016 International conference on advances in electrical, electronic and systems engineering (ICAEES)*, pp. 138–143. IEEE

Khlifi, Y., & El-Sabagh, H. A. (2017). A novel authentication scheme for e-assessments based on student behavior over e-learning platform. *International Journal of Emerging Technologies in Learning, 12*(4), 62.

Kim, K. I., Jung, K., & Kim, H. J. (2002). Face recognition using kernel principal component analysis. *IEEE Signal Processing Letters, 9*(2), 40–42.

Kim, S.-I., & Lee, K. (2020). A study on the operation of smart remote lecture-focusing on cisco webex meeting. *Journal of Digital Convergence, 18*(9), 317–322.

Kinnunen, T., & Li, H. (2010). An overview of text-independent speaker recognition: From features to supervectors. *Speech Communication, 52*(1), 12–40.

Klašnja-Milićević, A., Vesin, B., Ivanović, M., Budimac, Z., & Jain, L. C. (2016). *E-learning systems: Intelligent techniques for personalization* (Vol. 112). Springer.

Kremic, E., & Subasi, A. (2016). Performance of random forest and SVM in face recognition. *International Arab Journal of Information Technology, 13*(2), 287–293.

Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012) Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pp. 1097–1105

Kulaglić, S., Mujačić, S., Serdarević, I.K., & Kasapović S. (2013) Influence of learning styles on improving efficiency of adaptive educational hypermedia systems. In *2013 12th International conference on information technology based higher education and training (ITHET)*, pp 1–7 . IEEE

Kumar Jindal, A., Chalamala, S., & Kumar Jami, S.(2018). Face template protection using deep convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 462–470

Labati, R. D., Genovese, A., Muñoz, E., Piuri, V., Scotti, F., & Sforza, G. (2016). Biometric recognition in automated border control: A survey. *ACM Computing Surveys (CSUR), 49*(2), 1–39.

Labayen, M., Vea, R., Flórez, J., Aginako, N., & Sierra, B. (2021). Online student authentication and proctoring system based on multimodal biometrics technology. *IEEE Access, 9*, 72398–72411.

Lamiche, I., Bin, G., Jing, Y., Yu, Z., & Hadid, A. (2019). A continuous smartphone authentication method based on gait patterns and keystroke dynamics. *Journal of Ambient Intelligence and Humanized Computing, 10*(11), 4417–4430.

Leaton Gray, S. (2017) Biometrics in schools: The role of authentic and inauthentic social transactions. *British sociological association (BSA) conference*

Li, F., Clarke, N., Papadaki, M., & Dowland, P. (2014). Active authentication for mobile devices utilising behaviour profiling. *International Journal of Information Security, 13*(3), 229–244.

Li, F., Clarke, N., Papadaki, M., & Dowland, P. (2011) Behaviour profiling for transparent authentication for mobile devices. In *European conference on information warfare and security (ECIW)*. Academic Publishing Ltd.

Li, X., Chang, K.-m., Yuan, Y., & Hauptmann, A. (2015) Massive open online proctor: Protecting the credibility of moocs certificates. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 1129–1137

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, CL. (2014). Microsoft coco: Common objects in context. In *Proceedings of the European conference on computer vision*

Liu, C., & Wechsler, H. (2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing, 11*(4), 467–476.

Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017) Sphereface: Deep hypersphere embedding for face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 212–220

Machuletz, D., Laube, S., & Böhme, R. (2018) Webcam covering as planned behavior. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13

Mady, H., & Hilles, S.M.S. (2018) Face recognition and detection using random forest and combination of lbp and hog features. In *Proceedings of the international conference on smart computing and electronic enterprise*

Maniar, S., Sukhani, K., Shah, K., & Dhage, S. (2021) Automated proctoring system using computer vision techniques. In *2021 International conference on system, computation, automation and networking (ICSCAN)*, pp. 1–6 . IEEE

Mare, S., Baker, M., & Gummeson, J. (2016). A study of authentication in daily life. In *Twelfth symposium on usable privacy and security (SOUPS 2016)*, pp. 189–206

Martin, M. (2005). Seeing is believing: The role of videoconferencing in distance learning. *British Journal of Educational Technology, 36*(3), 397–405.

Massner, C. K. (2021) The use of videoconferencing in higher education. Communication Management

Meng, J., Zhang, J., & Zhao, H. (2012) Overview of the speech recognition technology. In *2012 Fourth international conference on computational and information sciences*, pp. 199–202 . IEEE

Metzger, R., & Maudoodi, R. (2020) Using access reports and api logs as additional tools to identify exam cheating. In *Society for information technology and teacher education international conference*, pp. 294–299. Association for the Advancement of Computing in Education (AACE)

Moini, A., & Madni, A. M. (2009). Leveraging biometrics for user authentication in online learning: A systems perspective. *IEEE Systems Journal, 3*(4), 469–476.

Monaco, J.V., Stewart, J.C., Cha, S.-H., & Tappert, C.C. (2013) Behavioral biometric verification of student identity in online course assessment and authentication of authors in literary works. In *2013 IEEE sixth international conference on biometrics: theory, applications and systems (BTAS)*, pp. 1–8. IEEE

Mondal, S., & Bours, P. (2013) Continuous authentication using mouse dynamics. In *2013 International conference of the BIOSIG special interest group (BIOSIG)*, pp. 1–12. IEEE

Montebello, M. (2018) Ai injected e-learning. Springer International Publishing (745). Online verfügbar unter https://link.springer.com/content/pdf/10.1007/978-3-319-67928-0.pdf, zuletzt geprüft am 19, 2018

Musambo, L. K., & Phiri, J. (2018). Student facial authentication model based on openCV's object detection method and QR code for Zambian higher institutions of learning. *International Journal of Advanced Computer Science and Applications, 9*(5), 1–7.

Nagrani, A., Chung, J. S., Xie, W., & Zisserman, A. (2020). Voxceleb: Large-scale speaker verification in the wild. *Computer Speech and Language, 60*, 101027.

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. *IEEE Access, 7*, 19143–19165.

Navarro, P., & Shoemaker, J. (2000). Performance and perceptions of distance learners in cyberspace. *American Journal of Distance Education, 14*(2), 15–35.

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021) A systematic review on ai-based proctoring systems: Past, present and future. *Education and Information Technologies*, 1–25

Norris, M. (2019) University online cheating–How to mitigate the damage. *Research in Higher Education Journal, 37*

O'Reilly, G., & Creagh, J. (2016) A categorization of online proctoring. In: *Global Learn*, pp. 542–552. Association for the Advancement of Computing in Education (AACE)

Okmawati, M. (2020). The use of google classroom during pandemic. *Journal of English Language Teaching, 9*(2), 438–443.

Omar, A., & Abdul Razak, S. (2020) Remote learning in the time of covid-19: an interactive learning calculus ii for engineers (mat235) by using microsoft teams digital platform. In *Virtual Symposium on Teaching and Learning (VSTL)*.

Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., & Koucheryavy, Y. (2018). Multi-factor authentication: A survey. *Cryptography, 2*(1), 1.

Othman, A., & Callahan, J. (2018). The horcrux protocol: a method for decentralized biometric-based self-sovereign identity. In *2018 international joint conference on neural networks (IJCNN)*, pp. 1–7 IEEE

Pagnin, E., & Mitrokotsa, A. (2017). Privacy-preserving biometric authentication: Challenges and directions. *Security and Communication Networks, 2017*, 7129505.

Pandey, R.K., Zhou, Y., Kota, B.U., & Govindaraju, V. (2016) Deep secure encoding for face template protection. In *2016 IEEE conference on computer vision and pattern recognition workshops (CVPRW)*, pp. 77–83. IEEE

Parkhi, O.M., Vedaldi, A., & Zisserman, A. (2015) Deep face recognition. In *Proceedings of the British machine vision conference*, pp. 41–14112

Partners, I. (2021) Online exam proctoring market forecast to 2027 . https://www.marketwatch.com/press-release/online-exam-proctoring-market-size-and-growth-2021-2027-major-key-players-analysis-changing-trends-size-share-industry-development-opportunities-and-challenges-includes-covid-19-impact-analysis-2021-08-19. Accessed 2021-11-05

Peacocke, R.D., & Graf, D.H. (1995) An introduction to speech and speaker recognition. In *Readings in human–computer interaction*, pp. 546–553. Elsevier.

Poddar, A., Sahidullah, M., & Saha, G. (2018). Speaker verification with short utterances: A review of challenges, trends and opportunities. *IET Biometrics, 7*(2), 91–101.

Prakash, A., Krishnaveni, R., & Dhanalakshmi, R. (2020). Continuous user authentication using multimodal biometric traits with optimal feature level fusion. *International Journal of Biomedical Engineering and Technology, 34*(1), 1–19.

Prathish, S., Narayanan, A., & Bijlani, K. (2016)An intelligent system for online exam monitoring. In *2016 International conference on information science (ICIS)*, pp. 138–143. https://doi.org/10.1109/INFOSCI.2016.7845315

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77*(2), 257–286.

Ratha, N. K., Chikkerur, S., Connell, J. H., & Bolle, R. M. (2007). Generating cancelable fingerprint templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(4), 561–572.

Rathgeb, C., Pöppelmann, K., & Gonzalez-Sosa, E. (2020) Biometric technologies for elearning: State-of-the-art, issues and challenges. In *2020 18th International conference on emerging elearning technologies and applications (ICETA)*, pp. 558–563. IEEE

Ravanelli, M., & Bengio, Y. (2018) Speaker recognition from raw waveform with SINCNET. In *2018 IEEE spoken language technology workshop (SLT)*, pp. 1021–1028. IEEE

Reddy, D. R. (1976). Speech recognition by machine: A review. *Proceedings of the IEEE, 64*(4), 501–531.

Rouhani, S., & Deters, R. (2019) Blockchain based access control systems: State of the art and challenges. In *IEEE/WIC/ACM international conference on web intelligence*, pp. 423–428

Rui, Z., & Yan, Z. (2018). A survey on biometric authentication: Toward secure and privacy-preserving identification. *IEEE Access, 7*, 5994–6009.

Sambur, M. (1975). Selection of acoustic features for speaker identification. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 23*(2), 176–182.

Sandnes, F.E., & Zhang, X. (2012) User identification based on touch dynamics. In *2012 9th international conference on ubiquitous intelligence and computing and 9th international conference on autonomic and trusted computing*, pp. 256–263. IEEE

Sarier, N.D. (2018). Privacy preserving biometric identification on the bitcoin blockchain. In *International symposium on cyberspace safety and security*, pp. 254–269. Springer

Sayed, B., Traoré, I., Woungang, I., & Obaidat, M. S. (2013). Biometric authentication using mouse gesture dynamics. *IEEE Systems Journal, 7*(2), 262–274.

Schleicher, A. (2021) The state of education - one year into COVID. https://oecdedutoday.com/state-of-education-one-year-into-covid/. [Online; Accessed Oct-01-2021]

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815–823

Selwyn, N., O'Neill, C., Smith, G., Andrejevic, M., & Gu, X. (2021) A necessary evil? the rise of online exam proctoring in Australian universities. Media International Australia, 1329878–211005862

Seurin, M., Strub, F., Preux, P., & Pietquin, O. (2020) A machine of few words–interactive speaker recognition with reinforcement learning. arXiv preprint arXiv:2008.03127

Shahzad, M., & Singh, M. P. (2017). Continuous authentication and authorization for the internet of things. *IEEE Internet Computing, 21*(2), 86–90.

Shen, C., Cai, Z., Guan, X., Du, Y., & Maxion, R. A. (2012). User authentication through mouse dynamics. *IEEE Transactions on Information Forensics and Security, 8*(1), 16–30.

Shen, C., Chen, Y., Guan, X., & Maxion, R. A. (2017). Pattern-growth based mining mouse-interaction behavior for an active user authentication system. *IEEE Transactions on Dependable and Secure Computing, 17*(2), 335–349.

Shen, C., Zhang, Y., Cai, Z., Yu, T., & Guan, X. (2015). Touch-interaction behavior for continuous user authentication on smartphones. In *2015 International conference on biometrics (ICB)*, pp. 157–162 . IEEE

Shen, C., Zhang, Y., Guan, X., & Maxion, R. A. (2015). Performance analysis of touch-interaction behavior for active smartphone authentication. *IEEE Transactions on Information Forensics and Security, 11*(3), 498–513.

Shen, C., Cai, Z., & Guan, X. (2012) Continuous authentication for mouse dynamics: A pattern-growth approach. In *IEEE/IFIP International conference on dependable systems and networks (DSN 2012)*, pp. 1–12. IEEE

Shi, E., Niu, Y., Jakobsson, M., & Chow, R. (2010) Implicit authentication through learning user behavior. In *International conference on information security*, pp. 99–113. Springer

Silva, H., Lourenço, A., Fred, A., & Filipe, J. (2011). Clinical data privacy and customization via biometrics based on ecg signals. In *Symposium of the Austrian HCI and usability engineering group*, pp. 121–132. Springer

Simonyan, K., & Zisserman, A. (2015) Very deep convolutional networks for large-scale image recognition. In *Proceedings of the international conference on learning representations (ICLR)*.

Sinha, P., & Yadav, A. (2020). Remote proctored theory and objective online examination. *International Journal of Advanced Networking and Applications, 11*(06), 4494–4500.

Slusky, L. (2020). Cybersecurity of online proctoring systems. *Journal of International Technology and Information Management, 29*(1), 56–83.

Stern, B. S. (2004). A comparison of online and face-to-face instruction in an undergraduate foundations of American education course. *Contemporary Issues in Technology and Teacher Education, 4*(2), 196–213.

Stolcke, A., Shriberg, E., Ferrer, L., Kajarekar, S., Sonmez, K., & Tur, G. (2007) Speech recognition as feature extraction for speaker recognition. In *2007 IEEE workshop on signal processing applications for public security and forensics*, pp. 1–5. IEEE

Taigman, Y., Yang, M., Ranzato, M.A., & Wolf, L. (2014) Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*

Teoh, A. B., Goh, A., & Ngo, D. C. (2006). Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(12), 1892–1901.

Tian, Y., Li, Y., Liu, X., Deng, R.H., & Sengupta, B. (2018) Pribioauth: Privacy-preserving biometric-based remote user authentication. In *2018 IEEE conference on dependable and secure computing (DSC)*, pp. 1–8 . IEEE

Tran, Q. N., Turnbull, B. P., & Hu, J. (2021). Biometrics and privacy-preservation: How do they evolve? *IEEE Open Journal of the Computer Society, 2*, 179–191.

Tran, Q. N., Turnbull, B. P., Wu, H.-T., Silva, A., Kormusheva, K., & Hu, J. (2021). A survey on privacy-preserving blockchain systems (PPBS) and a novel PPBS-based framework for smart agriculture. *IEEE Open Journal of the Computer Society, 2*, 72–84.

Truong, H. M. (2016). Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior, 55*, 1185–1193.

Tse, K.-W., & Hung, K. (2019). Behavioral biometrics scheme with keystroke and swipe dynamics for user authentication on mobile platform. In *2019 IEEE 9th Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, pp. 125–130 . IEEE

Turk, M., & Pentland, A. (1991a). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3*(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

Turk, M., & Pentland, A. (1991b) Face recognition using eigenfaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 586–591

Ullah, A., Xiao, H., & Barker, T. (2016). A classification of threats to remote online examinations. In *2016 IEEE 7th annual information technology, electronics and mobile communication conference (IEMCON)*, pp. 1–7 . IEEE

Uludag, U., Pankanti, S., Prabhakar, S., & Jain, A. K. (2004). Biometric cryptosystems: Issues and challenges. *Proceedings of the IEEE, 92*(6), 948–960.

Vajpai, J., & Bora, A. (2016). Industrial applications of automatic speech recognition systems. *International Journal of Engineering Research and Applications, 6*(3), 88–95.

Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: patterns and parameters part I: Recognition of backward voices. *Journal of Phonetics, 13*(1), 19–38.

Vicens, P. (1969). Aspects of speech recognition by computer. PhD thesis, Stanford University.

Villa, M., Gofman, M., Mitra, S., Almadan, A., Krishnan, A., & Rattani, A. A survey of biometric and machine learning methods for tracking students' attention and engagement. In *2020 19th IEEE international conference on machine learning and applications (ICMLA)*, pp. 948–955 (2020). IEEE

Waibel, A., & Lee, K.-F. (1990) Readings in speech recognition. Morgan Kaufmann.

Wan, V., & Carmichael, J. (2005) Polynomial dynamic time warping kernel support vector machines for dysarthric speech recognition with sparse training data. In *Ninth European conference on speech communication and technology*

Wang, A. I., & Tahir, R. (2020). The effect of using kahoot! for learning—A literature review. *Computers and Education, 149*, 103818. https://doi.org/10.1016/j.compedu.2020.103818

Wang, F., Xiang, X., Cheng, J., & Yuille, A.L. (2017). Normface: L2 hypersphere embedding for face verification. arXiv preprint arXiv:1704.06369

Wen, Y., Zhang, K., Li, Z., & Qiao, Y. (2016) A discriminative feature learning approach for deep face recognition. In *Proceedings of the European conference on computer vision*, pp. 499–515

Wright, J., Yang, A., Ganesh, A., Sastry, S. S., & Yi, M. (2009). Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31*(2), 210–227.

Yang, M., Zhang, L., Yang, J., & Zhang, D. (2011) Robust sparse coding for face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 625–632

Yi, D., Lei, Z., Liao, S., & Li, S.Z. (2014) Learning face representation from scratch. arXiv preprint arXiv:1411.7923

Yu, D., & Deng, L. (2016) Automatic speech recognition. Springer.

Zafar, S., Lai, Y., Sexton, C., & Siddiqi, A. (2020). Virtual reality as a novel educational tool in pre-clinical paediatric dentistry training: Students' perceptions. *International Journal of Paediatric Dentistry, 30*(6), 791–797. https://doi.org/10.1111/ipd.12648

Zhang, L., Yang, M., & Feng, X. (2011) Sparse representation or collaborative representation: Which helps face recognition? In *Proceedings of the IEEE international conference on computer vision*, pp. 471–478

Zhang, R., Xue, R., & Liu, L. (2019). Security and privacy on blockchain. *ACM Computing Surveys (CSUR), 52*(3), 1–34.

Zhang, Y., & Liu, L. (2018). Using computer speech recognition technology to evaluate spoken English. *Educational Sciences: Theory & Practice, 18*(5).

Zhang, Z., Zhang, M., Chang, Y., Esche, S.K., & Chassapis, C. (2016). A virtual laboratory system with biometric authentication and remote proctoring based on facial recognition. In *2016 ASEE Annual conference and exposition*

Zheng, N., Paloski, A., & Wang, H. (2016). An efficient user verification system using angle-based mouse movement biometrics. *ACM Transactions on Information and System Security (TISSEC), 18*(3), 1–27.

Zheng, J., Chen, J.-C., Bodla, N., Patel, V.M., & Chellappa, R. (2016) Vlad encoded deep convolutional features for unconstrained face verification. In *Proceedings of the IEEE international conference on pattern recognition*

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

David Portugal completed his Ph.D. degree on Robotics and Multi-Agent Systems at the University of Coimbra in Portugal, in March 2014. His main areas of expertise are cooperative robotics, multi-agent systems, simultaneous localization and mapping, field robotics, human-robot interaction, sensor fusion, metaheuristics, and graph theory. After his Ph.D., he spent 4 years outside academia, and he is currently working as an Assistant Researcher at the Institute of Systems and Robotics and as an Invited Assistant Professor at the University of Coimbra. He has been involved in several local and EU-funded research projects in Robotics and ICT, such as CHOPIN, TIRAMISU, Social Robot, CogniWin, GrowMeUp, STOP, CORE, SEMFIRE, WoW, 5GSmartFact, and TRUSTID. He has co-authored over 90 research articles included in international journals, conferences and scientific books.

José N. Faria received his MSc. degree in Engineering Physics at the University of Coimbra in February 2022 after developing a Wireless IoT Smart Bed System for his dissertation within the scope of the WoW CMU Portugal Large-Scale Collaborative project. He has been working as a Research Engineer since December 2021 on image recognition and face verification for Online Student Identity Management, as part of the ERASMUS+ TRUSTID project, at the Institute of Systems and Robotics (ISRUC). His main research interests focus on Internet of Things, Computer Networks, Image Processing, Machine Learning and Robotics.

Marios Belk is CEO of Cognitive UX GmbH. He completed a Ph.D. degree in Computer Science, with focus on usable security and human-centered computing (2015) at the Department of Computer Science, University of Cyprus. His research work has been published in accredited scientific journals and conferences, such as ACM Transactions on Computing for Healthcare, ACM CHI, ACM IUI. His publications include a best paper award at the SouthCHI 2013 conference and a best paper nomination award and honorable mention at the UMAP 2014 conference. He has professional experience for over ten years as a principal investigator, work package leader, technical manager and researcher in EU and national projects (Erasmus+ 2020 TRUSTID, Erasmus+ 2020 CREAMS, H2020 Serums, H2020 GrowMeUp, AAL Success, AAL CogniWin, FP7 Miraculous-Life, Marie Sklodowska-Curie SocialRobot).

Pedro Martins received both his M.Sc. and Ph.D. degrees in Electrical Engineering from the University of Coimbra, Portugal in 2008 and 2012, respectively. Currently, he is a senior researcher at the Institute of Systems and Robotics (ISR) at the University of Coimbra (UC), Portugal since March 2019. His main research areas include computer vision, image processing, computer graphics, pattern recognition, human computer interaction, machine learning, visual tracking and facial biometrics with a special focus on non-rigid image alignment (deformable models), dynamic facial biometrics, face tracking and facial expression recognition. Pedro is also an Invited Assistant Professor at the University of Coimbra, and he has co-authored more than 30 research articles included in international journals and scientific conferences

Argyris Constantinides is a Research Associate at the Department of Computer Science, University of Cyprus, and CTO and co-founder of Cognitive UX Ltd. He completed a Ph.D. degree in Computer Science, with focus on usable security (2022) at the Department of Computer Science, University of Cyprus. He holds a M.Sc. in Web Science and Big Data Analytics from the University College London (2015), and a B.Sc. in Computer Science from the University of Cyprus (2014). He has published scientific journals and papers on accredited conferences (ACM Transactions on Computing for Healthcare, International Journal of Human-Computer Studies, ACM IUI, ACM UMAP, ACM MobileHCI, IEEE/WIC/ACM Web Intelligence). Constantinides has professional experience as a researcher and software engineer in EU research projects (Erasmus+ 2020 TRUSTID, Erasmus+ 2020 CREAMS, H2020 Serums, EUAAL MEMENTO) and National research projects. In the past, he has worked as a software engineer at Playtech BGT Sports, London, UK (2015-2017).

Anna Pietron is a Project Manager at SHE Informationstechnologie AG. She holds an M.A. in Design from the Shanghai Jiao Tong University, China, and she has professional experience of over ten years in various fields, that gave her an opportunity to gain wide range of soft and hard skills in marketing, branding, business analysis, and user experience optimization. In the past, she has worked as a business development specialist in Germany and Poland including the Deutsche Bank, Trinity Management Systems GmbH, Poczta Polska and Cognitive UX GmbH. She has professional experience as project manager and researcher in EU projects (Erasmus+ 2020 TRUSTID)

Andreas Pitsillides is a Professor at the Department of Computer Science, University of Cyprus, codirector of the Networks Research Laboratory, and appointed Visiting Professor at the University of the Witwatersrand, School of Electrical and Information engineering, Johannesburg, South Africa. He has published over 300 refereed papers in flagship journals, international conferences, and book chapters, coauthored 2 books (1 edited), participated as principal or co-principal investigator in over 40 European Commission and locally funded research projects with over 6.6 million Euro, received awards, including best paper, presented keynotes, invited lectures at major research organizations, short courses at conferences and industry.

Nikolaos Avouris is a Professor at the University of Patras, Greece. His research interests revolve around Software Technology in relation to Industrial, Educational and Environmental Applications, Human-Computer Interaction, Interactive Systems Design, Distributed Intelligent Systems, Machine Learning, application of Knowledge-based techniques in Computer-Supported Collaborative Learning, educational, industrial and environmental fields. He has research and teaching experience in industry and academia for over 20 years. He was a key researcher in many national and international funded research projects in the frame of IST, ESPRIT, Environment, PENED, YPER etc. He has also served as the Editor of two international volumes and has published over 100 scientific papers and technical reports in the above research areas. He is also member of the Technical Chamber of Greece (1979), Greek Electrical Engineers Association (1979), Greek Computer Society (1992), IEEE Computer Society (1995), founding member of the Greek Artificial Intelligence Association (EETN) and Hellenic Association of Computer and Communication Technologies in Education

Christos A. Fidas received the Diploma degree in electrical and computer engineering and the Ph.D. degree in information systems from the University of Patras, Greece. He is a currently a Faculty Member with the Department of Electrical and Computer Engineering, University of Patras. His research interests include information systems, with an emphasis on human-computer interaction, usable-security, and cultural heritage. He has an extensive publication record in reputable scientific journals and conferences, has been granted a patent, and has received several scientific awards and research grants for his contributions to the field. For more information visit the link (http://cfidas.info).