RESEARCH ARTICLE

You Can't Hide Behind Your Headset: User Profiling in Augmented and Virtual Reality

PIER PAOLO TRICOMI[1,2,*], **FEDERICA NENNA**[3,4,*], **LUCA PAJOLA**[1,*], **MAURO CONTI**[1,2], **(Fellow, IEEE), AND LUCIANO GAMBERINI**[3,4]

[1]Department of Mathematics, University of Padova, 35122 Padua, Italy
[2]Chisito SRL, 35129 Padua, Italy
[3]Department of General Psychology, University of Padova, 35122 Padua, Italy
[4]Human Inspired Technology Research Centre HIT, University of Padova, 35121 Padua, Italy

Corresponding author: Pier Paolo Tricomi (tricomipierpaolo@phd.unipd.it)

Pier Paolo Tricomi, Federica Nenna, and Luca Pajola contributed equally to this work.

ABSTRACT Augmented and Virtual Reality (AR and VR), collectively known as Extended Reality (XR), are increasingly gaining traction thanks to their technical advancement and the need for remote connections, recently accentuated by the pandemic. Remote surgery, telerobotics, and virtual offices are only some examples of their successes. As users interact with XR, they generate extensive behavioral data usually leveraged for measuring human activity, which could be used for profiling users' identities or personal information (e.g., gender). However, several factors affect the efficiency of profiling, such as the technology employed, the action taken, the mental workload, the presence of bias, and the sensors available. To date, no study has considered all of these factors together and in their entirety, limiting the current understanding of XR profiling. In this work, we provide a comprehensive study on user profiling in virtual technologies (i.e., AR, VR). Specifically, we employ machine learning on behavioral data (i.e., head, controllers, and eye data) to identify users and infer their individual attributes (i.e., age, gender). Toward this end, we propose a general framework that can potentially infer any personal information from any virtual scenarios. We test our framework on eleven generic actions (e.g., walking, searching, pointing) involving low and high mental loads, derived from two distinct use cases: an AR everyday application (34 participants) and VR robot teleoperation (35 participants). Our framework limits the burden of creating technology- and action-dependent algorithms, also reducing the experimental bias evidenced in previous work, providing a simple (yet effective) baseline for future works. We identified users up to 97% F1-score in VR and 80% in AR. Gender and Age inference was also facilitated in VR, reaching up to 82% and 90% F1-score, respectively. Through an in-depth analysis of sensors' impact, we found VR profiling resulting more effective than AR mainly because of the eye sensors' presence.

INDEX TERMS Augmented reality, machine learning, metaverse, privacy, user profiling, virtual reality.

I. INTRODUCTION

In recent years, the pandemic has increased the need for remote connections, and we have witnessed to mass adoption of virtual technologies, particularly for teamwork. Different platforms have opened up new perspectives for virtual interactions with others, and fostered the already ascending

The associate editor coordinating the review of this manuscript and approving it for publication was Charalambos Poullis.

development of the Metaverse. The Metaverse has been recently defined as a "post-reality universe, a perceptual and persistent multiuser environment merging physical reality with digital virtuality" [1]. While being designed around the human, which constitutes the physical reality of this interplay, digital virtuality relies on immersive technologies that allow spatial and interactive features, namely Augmented Reality (AR) and Virtual Reality (VR), collectively known as Extended Reality (XR). Eventually, these

devices became the core of the fourth wave of computing innovation [2].

Currently, there is an ongoing discussion on the potential protocols that will govern the Metaverse, with a particular focus on the controversial interplay between openness and privacy [1]. The latest virtual devices allow tracking many behavioral data, such as the headset's and controllers' position and rotation, or eye movements. All these data can induce leak of personal information, and even the user's identity (e.g., [3], [4], [5]). While remaining private, this information would help to restrict the use of the headset to specific individuals. For example, it would be possible to allow authentication only to those with appropriate permissions, thus increasing the security of such technologies.

To date, many studies demonstrated the feasibility of user profiling tasks in XR such as authentication [6], [7], users identification [3], [4], [5], [8], and gender inference [8]. Nevertheless, the variety of XR devices and interactions have led researchers to build specific profiling mechanisms for each of their experiments, which were conducted on a single technology and single (or few) actions. Indeed, creating an ad-hoc system for every situation requires significant effort [4], [5]. Moreover, if features are bound to a specific action (e.g., hands distance in a grab action), they will hardly generalize in different scenarios, with the risk of introducing bias. For instance, Miller et al. [3] used the raw Y-axis of the Head Sensor (i.e., roughly the person's height) as a principal descriptor for user identity. However, as pointed out by the authors and recent literature [9], such a feature is not *persistent*. Last, the comprehension of which factors impact profiling in XR technologies is currently limited. Indeed, the literature suggests that profiling performances might depend on the technology (i.e., AR, VR) [10], user actions [4], cognitive workload [11], experimental bias [9], and XR sensors [4], but they were never examined altogether.

A. CONTRIBUTIONS

In this work, we propose a comprehensive study of XR user profiling by leveraging behavioral data obtained through the use of VR and AR headsets. As a first contribution, we introduce a general profiling framework applicable to different virtual devices (e.g., VR, AR), applied fields (e.g., everyday use cases, work scenarios), and types of user behaviors (e.g., walking, searching, pointing). We test our framework on data from our previous works [11], [12], showing the generability of the approach. Since such previous studies revealed gender differences under diverse workload conditions, we additionally investigate the workload impact on profiling. Ultimately, our framework leverages task-independent and free-of-bias features, aiming to become a baseline for XR profiling.

As a second contribution, we implement our framework to study users' profiling at different privacy levels (i.e., identification, personal information), introducing - to the best of our knowledge - the profiling of gender and age in virtual contexts through modern and widespread XR devices (AR Microsoft HoloLens, VR HTC VIVE Pro Eye).[1]

As a third contribution, we explore the impact of device sensors on users' profiling. Precisely, we assess the relevance of the headset's position and rotation, the controllers' position and rotation, and the eye tracker information available in the VR device. Last, we fill a gap in the literature on users' profiling in AR scenarios, which is largely understudied compared to VR. Overall, we summarize our contributions as follows:

- we propose a general profiling framework for XR technologies, which can serve as a generic baseline for future XR profiling studies;
- we examine users' profiling with respect to identification and private information (age and gender) in virtual scenarios, which is novel in the AR context;
- we introduce and explore the role of task workload in user profiling, which is a new concept in the area;
- we conduct extensive studies to assess sensors' importance in the profiling tasks.

B. ORGANIZATION

In Section II, we provide background and review literature on users' profiling. Section III presents the general profiling framework we adopted in our experiments. The dataset and experimental settings are shown in Section IV and Section V, respectively. We report our results in Section VI, and discussion in Section VII. We conclude in Section VIII.

II. BACKGROUND & RELATED WORK

This section describes the importance of security and privacy in XR technologies. Section II-A summarizes the application of virtual technologies in different fields, highlighting benefits deriving from user profiling. Section II-B introduces privacy in XR technologies, while Section II-C analyzes the state of the art in XR user profiling.

A. XR USE-CASES AND BENEFITS OF PROFILING
1) INDUSTRY AND REMOTE WORK

As Industry 4.0 progressed, virtual devices have proven their benefits in many sectors: in the design cycle of products and manufacturing systems [13], for programming machines [14], in the teleoperation industry [15], [16] and also for training novices [17], [18]. In any of these applications, virtual technologies provide the operator with a faithful virtual equivalent of the physical environment. Automatically identifying workers wearing headsets could improve workplace security. For example, authentication could be enabled only for those with appropriate permissions (e.g., site manager). Further, since older workers may prefer a different virtual environment design [19], user profiling could help customize virtual features according to age.

---

[1]Our VR device is a commercially widespread device that comes with an embedded eye-tracker, and as such, can potentially consolidate the findings of previous works based on hand-crafted devices.

2) EDUCATION

Online education through virtual environments is one of the key pillars of Metaverse [1]. Several studies have examined how immersive virtual technologies are successfully integrated into education, as well as how they positively influence learning. Subject-specific benefits include improving skills, living more realistic experiences, and enhancing motivation and interest in learning [20]. Additionally, [21] assessed VR applications for higher education are becoming increasingly popular in engineering, medicine, and computer science education, and are mature enough to teach declarative, procedural, and practical skills. With XR being widely adopted in education, a profiling or identification algorithm comes in handy. For example, teaching methods and content can be tailored based on each student's needs or age.

3) GAMING AND ENTERTAINMENT

While VR games have been popular since the 1990s (e.g., Virtual Reality Gear [22]), AR has been gaining popularity since 2016 with Pokémon Go, Snapchat, Apple's ARKit, and Google.com's ARCore [23]. The sector is expected to grow exponentially, as it encompasses entertainment markets beyond gaming and arcades: the film and music industry, live show sectors and sports are just a few examples [24]. Following the pandemic's devastating effects in these markets, immersive virtual platforms can help support the cinema, music, and live-show industries [22], [25]. Last, the recent proliferation of virtual influencers [26] demonstrates the importance of virtual technologies in both entertainment and marketing. Clearly, user profiling could be used for marketing strategies in this sector (e.g., delivering customized advertising). Further, particularly in gaming platforms, user identification might help detect banned individuals and prevent their access to virtual games.

4) MEDICINE

Both doctors and patients have found virtual technologies to be trustworthy. For instance, VR-simulated surgeries can be beneficial for medical education and training [27], while AR can support surgeries by overlaying salient clinical records or visual aids over the patient's body [28]. Virtual control systems for remote robotic surgery operations are also rising [29]. For patients, VR can help improve cognitive abilities after a traumatic brain injury [30] or increase engagement in Parkinson's motor training [31]. Through identification, detecting whether a user is a surgeon or a student can restrict their rights during an XR surgical procedure. Similarly, profiling patients could allow training customization and automatic recordings of clinical improvements.

5) AR AS A SMART WEARABLE TECHNOLOGY

The latest AR smart glasses are fully wearable devices with computational functions, providing various functionalities by freeing the user's hands [32]. For instance, Vuzix[2]

[2]https://www.vuzix.com/

developed AR smart glasses for navigation in unknown areas, while Zhao et al. developed an AR assistive navigation device [33]. Recently, Facebook has partnered with Ray-Ban and launched their Ray-Ban stories, which have raised important questions about ethical and privacy issues [34]. In the foreseeable future, the next generation of smart glasses will likely allow projecting e-mails and notifications from social networks on the user's field of view. Reliable automatic identification of the user during everyday activities would allow private messages to be viewed only by the owner.

B. PRIVACY IN XR TECHNOLOGIES

The increasing popularity of big data [35] coupled with the rapid adoption of various "smart" devices has resulted in parallel increases in privacy concerns. In today's society, most people consider data collection incessant and believe that the risks outweigh any benefits [36]. To prevent (or at least reduce) the exposure of personal data, current and emerging technologies should support privacy by default [37], following recent legislation such as GDPR [38]. Fortunately, researchers are actively focusing on studying and adding a security and privacy level to XR and, more in general, emerging technologies. For instance, Adams et al. [39] deeply investigated VR security and privacy perceptions from users and developers, outlining a "code of ethics" for developers. Abraham et al. [40] interviewed XR experts from industry and academia to investigate issues relating to security, privacy, and influencing behavior, providing guidelines for future XR devices supporting security and privacy by default. Recent works [41], [42] deeply discussed security and privacy issues arising in the metaverse, allowing a better understanding and a consequent improvement of the technology concerning its users. Similarly, Nair et al. [43], proposed a system to browse metaverse in incognito, protecting their privacy from companies, surveillance agencies, or data brokers. Researchers have also focused on incorporating privacy-preserving measures on daily usage systems, such as authentication [44], and more recently, de-authentication techniques [45].

Besides protecting users' data from unwanted usage or sharing, past literature shows how attackers can use *public* data in unconventional ways to profile users or to infer *private* users' data (e.g., gender, age, personality traits). Examples include video games data [46], Social Networks interactions [47], [48], or online ratings [49]. The results of such studies highlight the high risks connected with public data availability, highlighting the need for further research to enhance user privacy.

C. USERS PROFILING IN AR AND VR APPLICATIONS

Few works discussed user profiling in AR and VR technologies, which are synthesized in Table 1. First, we classified previous works based on the *technology* (AR vs VR), given the diverse level of immersion they provide [10]. Second, we distinguished the *privacy level* they operate, i.e., whether they tackle private data profiling (age and gender),

TABLE 1. State of the art overview. Legend: ◐ = AR, ◑ = VR, ● = AR & VR.

Reference	#Participants	Technology AR	Technology VR	Age	Authentication	Gender	Identification	Head Position	Head Rotation	Eyes	Controller Position	Controller Rotation	Multiple Actions	Workload Impact	Sensors Importance
Roger et al. [5]	20	Google Glass					◐	◐	◐	◐					◐
Li et al. [6]	95	Google Glass		◐				◐							
Mustafa et al. [7]	23		Google Cardboard VR	◑				◑	◑						
Steil et al. [8]	20		Oculus DK2 + Pupil-Lab			◑	◑			◑					
Pfeuffer et al. [4]	22		HTC Vive					◑	◑	◑	◑	◑	◑		◑
Miller et al. [3]	511		HTC Vive					◑	◑		◑	◑	◑		◑
Liebers et al. [9]	16		Oculus Quest HMD					◑	◑		◑	◑	◑		
Our	34 (AR) and 35 (VR)	Microsoft HoloLens	HTC VIVE Pro	●		●	●	●	●	◑	◑	◑	●	●	●

authentication, or identification. We remark that identifying a person (i.e., recognizing a given user among a group of known people) is substantially different than inferring their personal attributes (i.e., age, gender).[3] Third, we considered the *sensors* they adopted for the profiling. Several works [4], [5], [8] built their algorithms on eye trackers, motivated by the connection found between eye movements and personal information [8], [50], [51], [52], [53]. However, researchers have proposed a variety of methods [54], [55], [56] to hide personal identifiers from eye movements, and XR devices integrate a greater number of sensors (e.g., gyroscope, accelerometer) which require additional studies. As we will demonstrate in our experiments (Section VI), eye movements are not strictly necessary for user profiling. Last, we report whether they tested their algorithms on *multiple actions* (i.e., generability) and evaluated the *sensors' importance*, factors that might affect the profiling performances [4]. As a novel point, we introduce the role of *cognitive workload* in profiling, since it affects how users interact with XR technologies [11].

The reader can notice that existing works demonstrated that user profiling in XR technologies is feasible, but to what extent, as well as the required conditions, is currently unclear. We briefly present the limitation of current literature, and how we address such gaps.

1) SINGLE TECHNOLOGY
Previous works focused solely on one technology, AR [5], [6] or VR [3], [4], [7], [8], [9], developing customized and task-dependent algorithms. Given AR and VR both aims to provide an immersive environment and embed similar sensors, future XR studies would highly benefit from a cross-technology profiling framework.

2) LIMITED PRIVACY UNDERSTANDING
Researchers tackled mainly a single privacy level profiling, ignoring other privacy issues associated with XR devices.

For instance, to the best of our knowledge, there are no attempts in the literature to infer users' private data (e.g., age, gender) from modern XR devices. Indeed, many works [39], [40], [50] theorized that private data inference in XR was possible based on eye trackers studies [57], but none of these theories were empirically proven. The only evidence of gender profiling comes from Steil et al. [8], who purposely equipped the VR headset Oculus DK2 (2016) with an eye-tracker (Pupil[4]).

3) RESTRICTED SET OF SENSORS
The most impacting results were gained primarily by leveraging eye-movement features [4], [5]. Others leveraged different behavioral features such as head position and rotation [3], [4], [5], [6], [7], often being prone to experimental bias [9] (see Section III-C). Therefore, it is still unclear how different features contribute to the accuracy of a profiling task, nor if the feature choice should be task-dependent.

4) LACK OF GENERABILITY
Only two works [3], [4] tested their algorithms on multiple actions, questioning their generability. In AR, no works tested generability. We also noticed that no works analyzed the actions' cognitive workload impact, which was demonstrated to be crucial in XR interactions [11].

It follows that testing a general framework, which (1) leverages the same algorithms for profiling users in all XR technologies, (2) systematically considers multiple features, (3) extends to different levels of profiling tasks (identification, private data inference), and (4) works for multiple actions, might be helpful in view of higher generability and broader comprehension of XR user profiling.

III. METHODOLOGY
This section describes our methodology to execute user profiling within virtual technologies. Section III-A motivates the reasons for our investigation. The overview of our

---

[3]For instance, we might identify a person within a population by their surname, which is uncorrelated to their age or gender.

[4]https://pupil-labs.com/products/vr-ar/

proposed framework is presented in Section III-B, while the details are provided in Section III-C.

A. SCOPE OF THE WORK

This study examines whether users can be profiled by leveraging their interaction with AR and VR devices. In particular, we consider two privacy levels of user profiling:

1) *User identification*, where we aim to identify a given user within a known population;
2) *Private information inference*, where we aim to infer users' gender and age.

Thus, we propose a general framework to accomplish both tasks, extendable to infer additional users' information. Further, our framework requires to:

- work across different XR devices and actions;
- reduce the experimental bias by leveraging features uncorrelated from the task.

By satisfying these requirements, our framework can become a simple yet effective baseline to test user profiling over general XR devices and applications. The use of a generic-purpose framework can indeed simplify future research and comparison between multiple applications and devices.

B. INFERENCE FRAMEWORK OVERVIEW

Our goal is to define a generic pipeline that can be adapted and applied to any virtual technology (e.g., AR, VR) context to profile a user, in terms of identification or private information. As shown in Figure 1, the pipeline consists of four steps, starting from the *user* from whom we record the behaviors, to their actual profiling:

1) *Raw Data Acquisition*. In this phase, users' behavioral data are acquired. XR technologies' devices continuously generate data from users' interactions with the virtual environment (i.e., time series). From these data, we can describe users' behavior. The amount and type of information depend on the virtual technology and its devices. For instance, data might come from users' input (e.g., pressing joystick's buttons) and users' movements.
2) *Bias Removal*. This phase aims to remove potential biases from time series that might lead to train erroneous machine learning models.
3) *Time Series Engineering*. This phase aims to extract insightful information from the time series.
4) *Machine Learning Prediction*. This phase aims to infer users' private information from the data elaborated in the previous phase by leveraging machine learning algorithms.

C. FRAMEWORK DETAILED DESCRIPTION
1) RAW DATA ACQUISITION

Users interact with AR and VR applications through devices such as headsets and joysticks. These devices embed several functional sensors to offer users an immersive experience. For example, users move and explore the virtual environment through sensors like accelerometers and gyroscopes embedded in the headset. Thus, by combining information retrievable by each sensor s^i of the equipment, we can trace users activity a at a given time t:

$$\vec{a}_t = [s_t^0, s_t^1, \ldots, s_t^n], \qquad (1)$$

where the subscript denotes the timestamp and the superscript the sensor involved. We call this process *acquisition phase*. The acquisition phase can be repeated over time, resulting in a temporal user-behavioral description. Thus, by acquiring data in $\Delta t = t - t_0$, we obtain a behavioral time series, described as follows:

$$\vec{\mathbf{B}}_{\Delta t} = [\vec{a}_{t_0}, \vec{a}_{t_1}, \ldots, \vec{a}_{t-1}, \vec{a}_t]. \qquad (2)$$

$\vec{\mathbf{B}}_{\Delta t}$ represents an atomic sample of a user action (or task) of duration Δt that we will use in the next phases to infer their private information.

2) BIAS REMOVAL

The acquisition phase might lead to an enormous quantity of raw data. Such data might describe not only users' behavior, but also environmental information strongly correlated to experimental sessions. For example, using the raw headset height to identify users might be erroneous since such information might not be persistent over time (e.g., different shoes, different body position) [3]. The problem of *spurious correlations* in cybersecurity applications is well known [58]. Thus, care must be taken to understand if sensors might lead to erroneous and inconsistent machine learning performance. The process of bias removal depends on the sensors' nature and requires an ad-hoc analysis. We explain in detail our implementation in Section V-B. The de-biasing phase results in a new vector of de-biased actions:

$$\vec{\mathbf{B}}_{\Delta t} = [\vec{d}_{t_0}, \vec{d}_{t_1}, \ldots, \vec{d}_{t-1}, \vec{d}_t], \qquad (3)$$

where d_{t_i} is the de-biased version of the feature a_{t_i}.

3) TIME SERIES ENGINEERING

Raw temporal data should be properly elaborated to extract meaningful information. Moreover, given the vast amount of data, such sequences should be aggregated (i.e., compressed) to limit the computational cost of their analyses. The aggregation strategy can consider the whole sequence of specific features, or just a subpart of it. For example, given a sensor $s_{\Delta t}^i$ and its de-biased values over the time $d_{\Delta t}^i = [d_{t_0}^i, d_{t_1}^i, \ldots, d_{t-1}^i, d_t^i]$, the aggregation of a whole sequence results in a unique number x^i, while the partial aggregation (e.g., a transformation every q times step) in a vector of numbers $[x_0^i, x_1^i, \ldots, x_m^i]$, where $m = t/q$. Note that the subscript does not denote the temporal axis anymore. Popular features derived from the aggregation phase are the mean, standard deviation, min, max [3]. At the end of the process, we obtain, for each participant action or task, an aggregated datapoint $x = [x^0, x^1, \ldots, x^n]$ that will be used by the machine learning models.
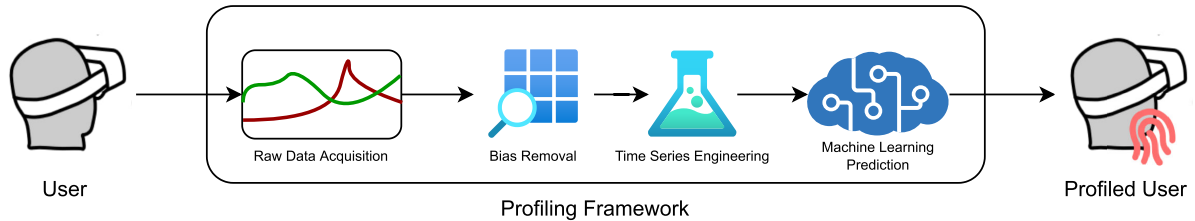
FIGURE 1. Overview of the proposed framework for user profiling in Augmented and Virtual Reality.

4) MACHINE LEARNING

The last phase of the pipeline involves machine learning approaches like Logistic Regression (LR), Decision Tree (DT), and Random Forest (RF). Training a well-performing model requires validation strategies that consider the nature of the inference. For instance, if the aim is to identify a user within a known population, the training, validation, and testing splits should contain samples of the whole population. However, to avoid trial (or session) bias, the three splits should consider samples from different collection trials. Conversely, when inferring information like age and gender, the three splits should contain different sets of users, since we want to infer the characteristics of people not seen at training (and validation) time. Regarding the type of machine learning algorithm, we suggest the use of *inherently interpretable* models (e.g, LR, DT) to better understand models' decisions during inference. Moreover, interpretable models allow a transparent debugging phase to identify the presence of spurious features [59]. Finally, given the unbalanced nature of the problem (i.e, not all the classes are distributed equally), we suggest using performance metrics like F1-score with macro average.

IV. DATASET OVERVIEW

Our previous studies assessed behavioral and workload aspects in individuals using AR while walking outdoor [11], and in users wearing VR for guiding an industrial robotic arm [12], [60]. In the present work, we leverage behavioral and eye-tracking data of both the AR and VR scenarios with the purpose of profiling users. For each technology, we considered tasks and actions to test the generability of our profiling approach and study the conditions or actions which might be more (or less) successful.

A. AR EXPERIMENT
1) OVERVIEW

The AR experiment investigated multitasking effects in participants using AR while walking outdoor [11]. Participants wore the Microsoft HoloLens 1st generation smart glasses, and interacted with the augmented targets both via an Xbox One controller and physical collision with the virtual objects (e.g., walking through an augmented target). They performed: i) a *visual task*, in which they discriminated between different augmented targets presented in their peripheral view, ii) a *navigation task*, in which

they reached a series of augmented landmarks via physical walking outdoor, and iii) the combination of these tasks, i.e., a *dual-task*. For more details about the tasks, please refer to [11]. Figure 2 shows an example of the virtual environment. Each participant performed 80 trials of the visual task, 50 trials of the navigation task, and 50 trials of the dual-task. While the original dataset was composed of 45 participants, we excluded 11 participants whose headset position data were not correctly recorded, and finally run our analyses on 34 participants (10 females age mean = 24.28, SD = 2.22 - 24 males age mean = 24, SD = 2.62). We continuously recorded through the device (60 Hz) the following measures: position (in meters) of the AR headset in the three axes (x, y, z), and rotation of the AR headset in Euler angles.

2) TASKS

From the experimental design, we identified the following tasks:

- *Mental Task (MT).* The mental task corresponds to what [11] call Visual Discrimination Task. Specifically, participants were discriminating between different colored and lateralized augmented objects while standing still.
- *Navigation Task - Low workload (NT-Low).* Participants were looking for augmented targets in their surroundings and then walked through them.
- *Navigation Task - High workload (NT-High).* Participants were executing the navigation task concurrently with the mental task. The concurrent execution of two tasks is known as "dual-task paradigm" and is usually deployed in cognitive science research to create higher mental demand on the participant.

3) ACTIONS

Each task is composed of smaller operations that we named actions. The dataset contains the following actions:

- *Button interaction.* Participants were standing still while discriminating between the lateralized colored targets. Specifically, they were instructed to press specific buttons on the joystick according to the hemifield where the virtual object was displayed.
- *Search.* Participants were engaged in the visual inspection of the surroundings to find a virtual landmark; this action was performed while participants were
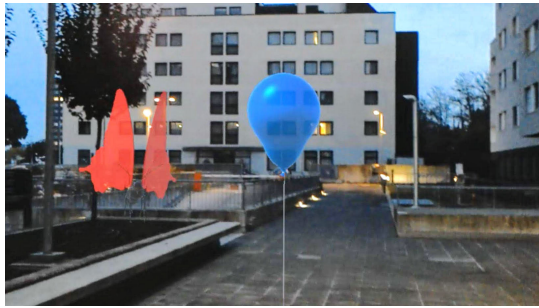
(a) Augmented Reality
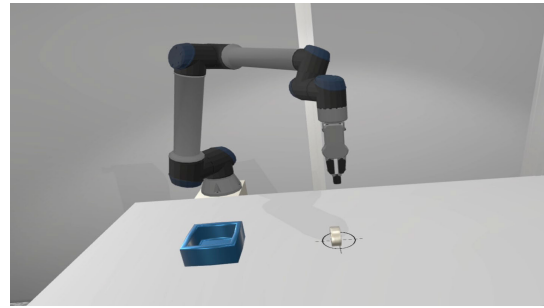
(b) Virtual Reality

FIGURE 2. Virtual environments adopted in the experiments.

standing still and just rotated their heads to inspect the surrounding.

- *Walk.* Participants were physically walking to the identified virtual landmark.

Both the search and walk actions were performed as single-task and concurrently with the secondary mental task (namely, the visual discrimination task). Based on the results obtained in our previous work [11], participants perceived a lower workload in the single-task compared to the dual-task. Therefore, we here refer to the dual-task as the high workload condition, while the single-task is considered as a low workload condition. Furthermore, the button interaction action was categorized as a high workload condition since it entailed high and sustained attentional processes for correctly discriminating the stimuli appearing lateralized to the participant's field of view. Table 2 represents the actions isolated in the AR environment.

TABLE 2. Augmented Reality actions organized per type of action and workload level.

Workload \ Action	Button Interaction	Search	Walk
Low	–	🎈	🎈
High	🔺🔺	🔺🔺 🎈	🔺🔺 🎈

B. VR EXPERIMENT

1) OVERVIEW

As part of the VR experiment, participants guided a virtual replica of an industrial robotic arm (Universal Robot e-Series UR5e) developed in Unity [12]. They were equipped with an HTC VIVE Pro Eye VR device and two VR controllers and guided the robotic arm shown in figure 2b through a pick-and-place, i.e., picking and placing a bolt into a box. They performed the task using two control systems (controller

buttons and physical actions, i.e., moving their hands) and under two levels of workload (single-task and dual-task). In the dual-task, participants also performed simple arithmetic sums, typing the results on a virtual keyboard by pointing the controller. Further details about the task design can be found in [60]. In each condition, the young participants performed 40 trials, while the old participants performed 20 trials. In total, 35 participants executed this task (18 females, age mean = 39.33, SD = 14.21 – 17 males, age mean = 37.75, SD = 16.32). The following measures were continuously recorded through the device (120 Hz): position (in meters) in the three axes (x, y, z), rotation in Euler angles of both the VR headset and its controllers, pupil size (in millimeters), and eye openness (expressed from 0 to 1).

2) TASKS

From the experimental design, we identified the following tasks:

- *Controller-based Task - Low workload (CT-Low).* Participants guided the robot via controller buttons under a low workload;
- *Controller-based Task - High workload (CT-High).* Participants guided the robot via controller buttons under a high workload (i.e., while also calculating sums);
- *Action-based Task - Low workload (AT-Low).* Participants guided the robot via physical actions under a low workload;
- *Action-based Task - High workload (AT-High).* Participants guided the robot via physical actions under a low workload (i.e., while also calculating sums).

3) ACTIONS

From the tasks performed in VR, we extracted the following actions:

- *Idle.* Participants were only looking at the robot while it was executing either a pick or place automation, i.e., were not interacting with the virtual environment;
- *Pointing.* Participants were using the VR controller to point the numbers on the virtual keyboard to report the result of the arithmetical sums;

- *Button Interaction.* Participants guided the virtual robot through the pick-and-place task by only pressing specific buttons on the VR controller;
- *Physical interaction.* Participants physically touched the virtual robot and moved their arms to relocate it over the worktable.

In line with our previous findings [12], the actions performed concurrently with the arithmetic task were categorized under high workload. Differently, when performed without any additional task, they were categorized under the low workload. Table 3 represents the actions isolated in the VR environment.

TABLE 3. Virtual Reality actions organized per type of action and workload level.

Workload \ Action	Idle	Pointing	Button Interaction	Physical Interaction
Low				
High		-		

C. ETHICS

The data for this study come from our previous works, which were approved by the internal ethical committee of the University of Padova, Italy. Participants signed informed consent. The ethical committee approved the possibility of sharing anonymized data with other researchers to foster transparency, reproducibility, and further research.

V. EXPERIMENTAL SETTING

This section describes our experimental settings. Section V-A explains the targets of our profiling, while Section V-B describes the implementation (i.e., de-biasing, feature extraction, model selection).

A. PROFILING TARGETS

In our experiments, we are interested in the identification, age, and gender profiling processes. In each of them, we use the headset's data the user generates when interacting with the XR environment (i.e., behavioral data) to predict a target (i.e., the user identity, gender, and age). These processes will be performed on each task and action presented in Section IV. We now describe in detail the three processes.

1) IDENTIFICATION

This process aims to identify a particular user among a group of known users. In this setting, every user appears in both training, validation, and testing data. Therefore, the training set contains the behavioral data of all the users. First, we train an ML model able to map a user's behavioral data to their identity. Then, when we present the ML model with new (unknown) behavioral data, it identifies the user who generated them.

2) AGE PROFILING

This process aims to infer the user's age starting from their behavioral data. In this setting, users appear only in one of the training, validation, and testing set. Therefore, the training set contains the behavioral data of only a subset of users. First, we train an ML model able to map a user's behavioral data to their age. Then, when we present the ML model with the behavioral data of a new user (unknown), the model infers their age.

3) GENDER PROFILING

This process aims to infer the user's gender starting from their behavioral data. In this setting, users appear only in one of the training, validation, and testing set. Therefore, the training set contains the behavioral data of only a subset of users. First, we train an ML model able to map a user's behavioral data to their gender. Then, when we present the ML model with the behavioral data of a new user (unknown), the model infers their gender.

We remark that age and gender profiling are substantially different from identification. Indeed, in the identification process, the ML model works with data of *known* users, while in age and gender profiling, the aim is to infer the targets of *unknown* users.[5] In other words, identification can be used only when the population is known (e.g., within a family context), while age and gender profiling can be used when the population is unknown (e.g., when a customer wears the device for the first time).

B. IMPLEMENTATION
1) DE-BIASING AND FEATURE EXTRACTION

AR and VR datasets contain different types of raw features acquired from the sensors. We now describe, for each category of sensors, the features and de-biasing techniques we applied.

- Head Position (AR and VR), represented as a 3D coordinate (x, y, z) measuring the relative distance (in meters) of the user from a center point in the virtual environment. This feature might contain both session's and users' static traits (e.g., height). We thus derived different variants of this information, such as the movement, computed as the norm between two points at 5 timestamps of distance, and the vertical oscillation computed as the difference between two height values at 5 timestamps of distance.
- Head Rotation (AR and VR), represented as a 3D value. For each axis, we compute its angular speed by considering points at 5 timestamps of distance. This transformation can remove information related to trials (e.g., specific positioning of objects with respect to the participant).

---

[5]Having the same users in training and test data when performing private data inference causes overfitting, since it degenerates into an identification task [61].

- Eyes (VR), includes data on pupil size (in millimeters) and eye openness (0-1), for both left and right eyes. In order to overcome possible confounding variables [62], [63], it is usually appropriate to pre-process the raw eye data for flattening individual differences. However, as the aim of the present work was specifically to capture individual traits and behaviors for allowing identification/profiling, we opted for not pre-processing eye-tracking data. On the contrary, we leveraged the individual differences in pupil size and eye openness [64], [65], [66] Further, we enhance this set of features by computing the symmetry among the eyes for both pupil dilatation and eye openness. On an applied level, using the raw output of the HTC Vive Pro Eye device speeds up the identification/profiling process and allows higher generability to multiple VR devices.
- Controller Position (VR), represented as 3D coordinates (x, y, z) relative to the virtual environment center point. Similarly to the head position, this feature might contain both sessions and user traits. We thus transform it in the movement, computed as the norm between two points at 5 timestamps of distance.
- Controller Rotation (VR) represented as 3D value. We conduct the same process of head rotation.

Finally, each feature of the previously described families is aggregated with tsfresh.[6] Given a time series, this library extracts more than 100 features, including average, standard deviation, quantile, and entropy. We further refined the features by keeping only the relevant ones.[7] Thus, starting from the raw time series of a single action within a single task performed in a single trial by a single user, we extract a single aggregated data point. The process is repeated for all the users, trials, actions, and tasks, obtaining 9360 datapoints in AR, and 16520 datapoints in VR.

2) MODELS TRAINING AND VALIDATION

In our experiments, we test four different algorithms: logistic regression, ridge classifier, decision tree, and random forest. As a baseline, we defined a Dummy classifier that randomly predicts the outcome based on the training ground-truth distribution. For each experiment presented in Section VI, we adopt a common validation strategy: for each discussed model, we find the best hyper-parameters through a grid-search validation based on training, validation, and testing split of 70%, 10%, and 20% of samples, respectively. For private inferring tasks (i.e., age and gender), the splits contain different sets of users, i.e., users in training are not present in the validation and testing set. Similarly, users in validation are not present in both training and testing sets. Machine learning models are designed as a multiclass classification problem for the user identification task. On the opposite, we considered a binary classification problem for both age

---

[6]https://tsfresh.readthedocs.io/en/latest/index.html
[7]We used tfresh feature_selection function: https://tsfresh.readthedocs.io/en/latest/api/tsfresh.feature_selection.html

(i.e., young and old) and gender (i.e., male and female). Note that the young class correspond to users defined in [19 − 24] (AR) and [23 − 30]; the old class is defined in [25, 29] (AR) and [31 − 69]. We now report the parameter grids involved in the grid search:

- Logistic Regression (LR). C: 0.1, 1, 10.
- Ridge (RI). Alpha: 0.01, 0.1, 1., 10. Fit intercept: *False*, *True*.
- Decision Tree (DT). Max Depth: 3, 5, 7. Min samples leaf: 1, 3, 5.
- Random Forest (RF). N estimators: 50, 100, 150. Max Depth: 3, 5, 7. Min samples leaf: 1, 3, 5.

To provide accurate results, each experiment is repeated five times. We thus report both the mean and standard deviation of the F1-scores (with macro average). We implemented our experiments in Python 3.8.5 and we used Scikit-Learn [67] library for training models and validation algorithms.

VI. RESULTS

In this section, we present the results of our experiments. We present both results for the task and action levels, in sections VI-A and VI-B, respectively. We then conclude with an ablation study to determine the effect of different sensors on models' performance (Section VI-C),

A. TASK-LEVEL

In this section, we present profiling performance at a task-level. In particular, each presented experiment considers distinctly the tasks presented in Sections IV-A and IV-B. In more detail, we train, validate, and test our model only on the task under investigation, predicting each time the identity, age, and gender separately. For instance, we train a specific model to predict gender based only on the Mental Task.

1) IDENTIFICATION

Figure 3 shows the identification results in AR and VR environments. LR and RI achieved the highest (and comparable) performances in AR, whereas LR and RF performed best in VR. In general, all our algorithms outperform the baseline (Dummy). Looking at the results on the Overall Tasks, both in VR (OT-VR) and AR (OT-AR), we immediately notice that in VR identification, the performances remain pretty stable as the number of users increases, while AR degrades significantly. Indeed, AR best algorithms performance goes from nearly 0.90 F1-Score (two users) to slightly above 0.60 F1-Score (30 users). Instead, in VR, LR yields almost perfect prediction on two users, while the F1-Score is above 0.95 when performing identification over 30 users. This might reflect the different amount of sensors available in VR (headset, controller, and eye-related behaviors) compared to those available in AR (only headset-related behaviors). We further discuss the impact of each of the involved sensors in Section VI-C.

When looking at the individual tasks, we can see that the identification algorithm performs even better than the overall task, particularly in AR. For instance, we reached
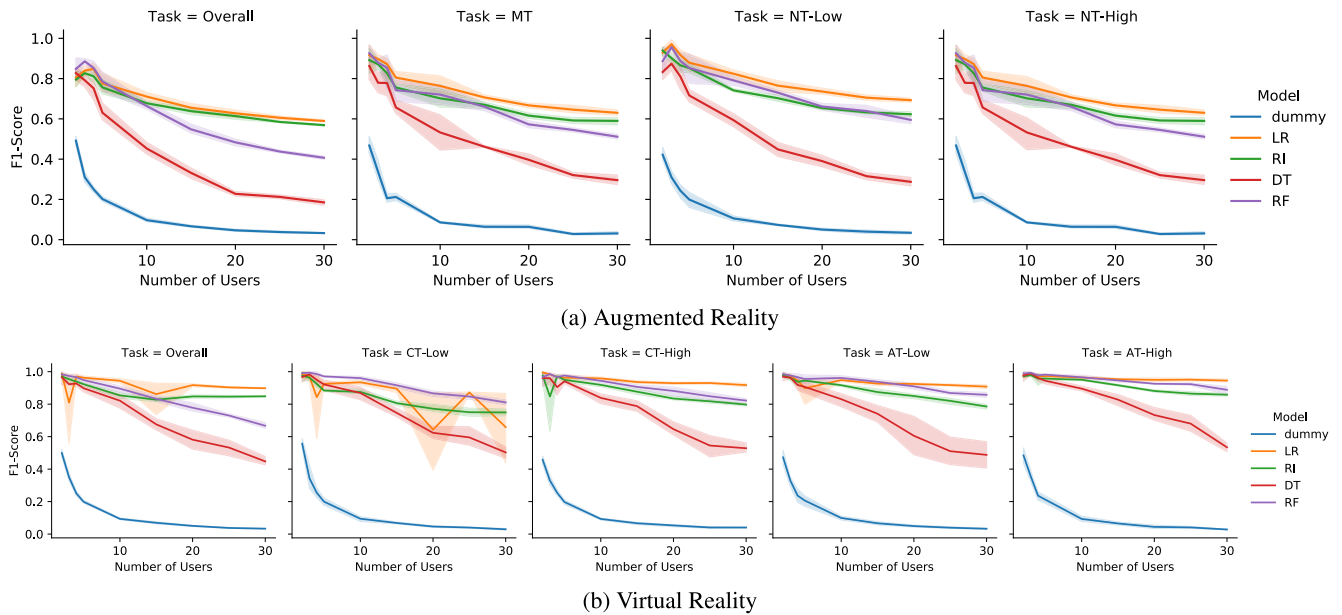
(a) Augmented Reality

(b) Virtual Reality

FIGURE 3. User Identification on task-level.

0.70 F1-Score over 30 users in the NT-Low, which is roughly 0.10 higher than in the OT-AR. One reason for this result might be related to the nature of the performed task: in the NT-Low, participants were actively moving in the surroundings without performing any additional task. Therefore, their movements might have been more linear compared to the situation in which they performed the same task under a high workload (NT-High), thus revealing more identifiable movement patterns. The same does not apply to the VR scenario. Here, when looking at each of the identified actions, the higher the workload the better the performance of the identification algorithm. Indeed, the best performance was obtained at the AT-High and CT-High, where the F1-Score was around 0.95 and 0.97, respectively. Again, possible explanations might be related to the nature of the tasks and the number of sensors embedded in the devices. In the VR scenario, participants were only moving their upper body, and in the high workload conditions they were additionally engaged in a secondary mental task. We know from the literature that a higher workload is related to higher changes in eye behavior [12]. Therefore, the VR-embedded eye-tracker might have had an essential impact on the identification performance, mainly when users were under higher mental strain rather than when performing less demanding tasks (i.e., CT-Low, AT-Low).

2) AGE

Figure 4 shows the age classification results in AR and VR environments at task-level. Results from the age profiling clearly yielded better performance in the VR compared to the AR scenario. While in VR all models performed significantly better than the baseline, in AR the F1-Score was consistently lower than the baseline, in all tasks. This is likely to be related to the low age variability of participants that took part in the AR experiment, or the inadequacy of sensors (see

Section VI-C). While this is a clear limitation of our study, such a result is still valuable, since it suggests that people of similar ages interact similarly with AR devices, meaning that age profiling may not be possible in every circumstance. Therefore, we focus our discussion mainly on age profiling performances in relation to the VR experiment.

In VR, the LR and RF algorithms appear to perform better than the other models in all tasks, but in the OT, where RI produced a higher F1-Score compared to LR. On the task-level, the users' age was profiled with higher accuracy when they performed the pick-and-place task via physical actions (AT-High and AT-Low, in which F1-Score was around 0.90 and 0.85 respectively) compared to controller buttons (CB-High, CB-Low, in which F1-Score was below 0.80 in both cases). A possible interpretation on this point is that the movement patterns of older users might have been quite different from younger users. Also, we know from the literature that robot teleoperation is significantly influenced by age [68]. In this view, our algorithm was particularly successful in detecting users' age during the pick-and-place task only when physical actions were involved.

3) GENDER

Figure 5 shows the gender classification results in AR and VR environments at task-level. When profiling users' gender, we obtained substantially better results in VR compared to AR. Indeed, in VR, all the tested algorithms performed above the baseline (dummy). More specifically, we can observe a better performance obtained through LR and RF, which reached a maximum F1-Score of 0.75. Differently, when detecting users' gender in the AR scenario, our algorithms performed only 0.5-0.10 above the baseline. This discrepancy could be explained by the inadequacy of sensors (see Section VI-C)
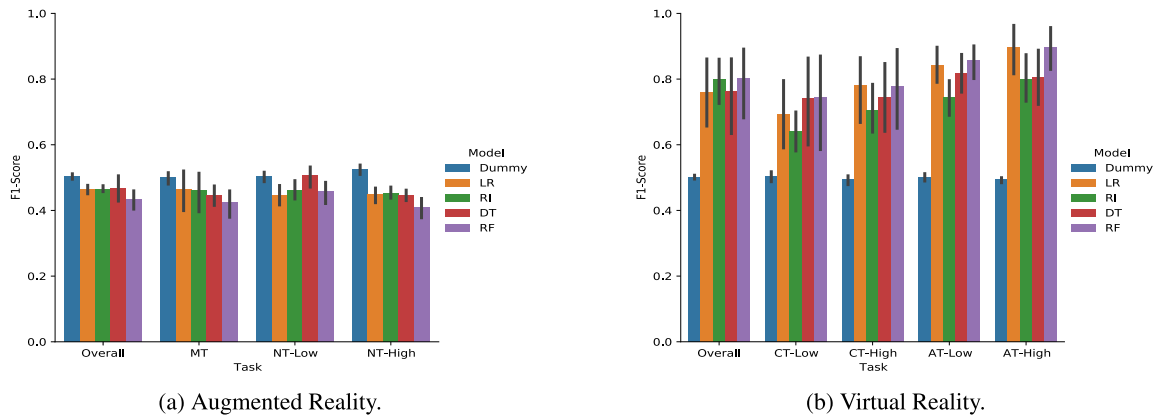
(a) Augmented Reality.

(b) Virtual Reality.

FIGURE 4. Age profiling on task-level.

(a) Augmented Reality.
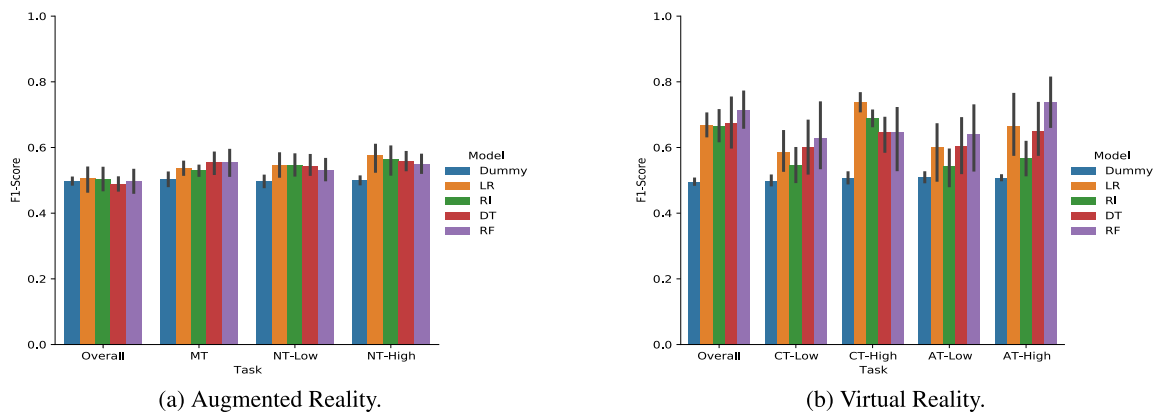

(b) Virtual Reality.

FIGURE 5. Gender profiling on task-level.

In VR, we achieved better performance in tasks involving a higher workload (CT-High, AT-High) than those under a low workload (CT-Low, AT-Low). These results align with recent literature on behavioral gender differences in the VR pick-and-place task. For instance, our previous work [60] demonstrated how men outperformed women in the pick-and-place tasks in terms of task execution time, particularly when using controller buttons. These differences might have been even more marked when performing an additional mental task, thus allowing more precise gender profiling. We observe a similar trend in the AR scenario, where higher workloads (NT-High) result in better performance. This behavior reflects previous findings related to the different walking patterns between men and women [11]. Indeed, on average, the walking velocity of men is significantly higher than women's one, particularly under high workloads. As we recorded the headset shifts in time, the different walking velocities might have been prominent in gender profiling.

B. ACTION-LEVEL

Starting from the results obtained in the overall task, we investigated whether some actions had a particular effect on the identification and profiling performances. Specifically, we opted for leveraging the model that demonstrated better results, which was the Logistic Regression (LR).

Each presented experiment considers distinctly the actions presented in Sections IV-A and IV-B. In more detail, we train, validate, and test our model only on the action under investigation, predicting each time the identity, age, and gender separately. For instance, we train a specific model to predict age based only on Button Interaction with Low Workload.

1) IDENTIFICATION

Table 4 shows the identification results in AR and VR environments at action-level. Previously at task-level we reached an F1-Score of about 0.60 in the AR and above 0.90 in the VR scenario. Looking at the action-level, specifically for AR, we see that the walking action reaches the highest performance (F1-Score is about 0.80 under low workload and 0.78 under high workload), while the search action and button interaction reveal F1-Scores below 0.70. This suggests that the walking action is prominent in identifying users in AR, possibly because the walking pattern is the most singular feature in such a use-case of AR. Differently, in VR, we observe higher F1-Scores for both button and physical interactions, specifically under high workload (F1-Score is about 0.96 in both cases). Also, the pointing action reached a very similar F1-Score (0.96), while the idle time intervals yield lower F1-Scores (below

TABLE 4. User identification on action-level organized per type of operation and workload level. Random guess at 0.03 for both AR and VR tasks. All the measures in F1-Score.

Workload \ Action	Augmented Reality			Virtual Reality			
	Button Interaction	Search	Walk	Idle	Pointing	Button Interaction	Physical Interaction
Low	–	$0.66_{\pm 0.03}$	$0.80_{\pm 0.02}$	$0.78_{\pm 0.02}$	$0.96_{\pm 0.01}$	$0.92_{\pm 0.01}$	$0.93_{\pm 0.02}$
High	$0.61_{\pm 0.02}$	$0.69_{\pm 0.01}$	$0.78_{\pm 0.02}$	$0.86_{\pm 0.01}$	–	$0.96_{\pm 0.00}$	$0.96_{\pm 0.01}$

TABLE 5. Age profiling on action-level organized per type of operation and workload level. Random guess at 0.5 for both AR and VR tasks. All the measures in F1-Score.

Workload \ Action	Augmented Reality			Virtual Reality			
	Button Interaction	Search	Walk	Idle	Pointing	Button Interaction	Physical Interaction
Low	–	$0.40_{\pm 0.03}$	$0.45_{\pm 0.02}$	$0.77_{\pm 0.10}$	$0.88_{\pm 0.06}$	$0.70_{\pm 0.09}$	$0.82_{\pm 0.05}$
High	$0.47_{\pm 0.02}$	$0.44_{\pm 0.01}$	$0.49_{\pm 0.02}$	$0.83_{\pm 0.09}$	–	$0.81_{\pm 0.07}$	$0.90_{\pm 0.05}$

TABLE 6. Gender profiling on action-level organized per type of operation and workload level. Random guess at 0.5 for both AR and VR tasks. All the measures in F1-Score.

Workload \ Action	Augmented Reality			Virtual Reality			
	Button Interaction	Search	Walk	Idle	Pointing	Button Interaction	Physical Interaction
Low	–	$0.50_{\pm 0.02}$	$0.45_{\pm 0.06}$	$0.60_{\pm 0.10}$	$0.82_{\pm 0.09}$	$0.62_{\pm 0.05}$	$0.66_{\pm 0.11}$
High	$0.54_{\pm 0.03}$	$0.58_{\pm 0.03}$	$0.60_{\pm 0.06}$	$0.63_{\pm 0.05}$	–	$0.74_{\pm 0.06}$	$0.66_{\pm 0.08}$

0.80 both under high and low workloads). It seems that the most interactive actions (using controller buttons, pointing, and physically moving the upper body) yield better results compared to periods in which users were passively looking at the virtual surroundings.

2) AGE

Table 5 shows the age classification results in AR and VR environments at action-level. Users' age was profiled with an F1-Score of about 0.50 on the overall task executed in AR, and 0.80 in VR. As the age profiling was unsuccessful in AR, we will not pay close attention to the action-level results of this use case. These results confirm what we observed at task-level (see Figure 4). Regarding the VR scenario, we can note that, under low workload, the pointing (F1-Score = 0.88) and physical interactions (F1-Score = 0.82) were the most crucial in profiling users' age, compared to actions allowing less interactivity with the virtual environment (F1-Scores below 0.80). This might suggest a different movement and interaction pattern shown by older and younger users, especially when greater freedom of movement is allowed. This is also in line with what was observed on task-level. Moreover, this trend becomes even more evident when the physical interactions are performed under a high workload (F1-Score = 0.90), likely reflecting the multitasking and motor difficulties related to age [69].

3) GENDER

Table 6 shows the gender classification results in AR and VR environments at action-level. On task-level, our algorithms reached an F1-Score of about 0.50 in AR and above 0.70 in VR. Even though the gender profiling did not perform sufficiently well in AR, here we can observe that, under high workload, both walk (F1-Score = 0.60) and search (F1-Score = 0.58) had a significant influence in detecting the user gender compared to the same actions performed under the low workload. The button interaction was slightly better than the random classifier (F1-score = 0.54). These results align with task-level results, whereby the gender profiling performed better in the NT-high compared to NT-low. Additionally, we observe how the walking action has the largest influence on the accuracy of gender profiling. Again, it might be related to different walking velocities demonstrated by men and women, particularly under high workload [11].

When looking at the actions performed in VR, the pointing action stands out. With an F1-score of 0.82, it strongly contributes to gender profiling compared to all other actions. This might be related both to a singular movement pattern and/or to gender-related eye parameters' variations. Further, results obtained at task-level on a better performance achieved under high compared to low workload are here confirmed only for button interactions. Indeed, the F1-Score at button interactions is about 0.08 higher when users are under high rather than low workload. Again, this reflects results shown in previous studies demonstrating faster operation times in men compared to women specifically when using controller buttons, but not when acting via physical actions [12]. This suggests that profiling users' gender might be easier during tasks involving button interactions, but not in those allowing higher interactivity with the virtual environment.

C. SENSORS RELEVANCE - ABLATION STUDY

In this section, we conduct an ablation study to understand which sensors contribute the most in our identification, age, and gender predictions. In brief, we trained a Logistic

TABLE 7. Ablation study of sensor importance at task-level in AR. All the measures in F1-Score.

		Identification	Age	Gender
	Guessing	0.03	0.5	0.5
Low W.	**Mental Task**			
	Head Position	0.38	**0.46**	0.51
	Head Rotation	**0.54**	0.40	**0.55**
	Navigation Task			
	Head Position	**0.64**	**0.45**	**0.56**
	Head Rotation	0.46	0.40	0.45
High W.	**Navigation Task**			
	Head Position	**0.65**	**0.45**	0.51
	Head Rotation	0.48	0.44	**0.52**

TABLE 8. Ablation study of sensor importance at task-level in VR. All the measures in F1-Score.

		Identification	Age	Gender
	Guessing	0.03	0.5	0.5
Low Workload	**Controller Based Task**			
	Head Position	0.41	0.68	**0.64**
	Head Rotation	0.45	**0.76**	0.55
	Eyes	**0.83**	0.75	0.59
	Controller Position	0.39	0.69	0.57
	Controller Rotation	0.59	0.69	0.58
	Action Based Task			
	Head Position	0.50	0.76	**0.62**
	Head Rotation	0.51	0.76	0.60
	Eyes	**0.83**	0.74	0.54
	Controller Position	0.51	0.76	0.58
	Controller Rotation	0.68	**0.81**	0.55
High Workload	**Controller Based Task**			
	Head Position	0.48	0.73	0.61
	Head Rotation	0.56	0.68	0.57
	Eyes	**0.88**	**0.79**	**0.69**
	Controller Position	0.45	0.78	0.60
	Controller Rotation	0.64	0.68	0.62
	Action Based Task			
	Head Position	0.55	0.75	0.53
	Head Rotation	0.55	0.80	**0.62**
	Eyes	**0.89**	0.83	0.62
	Controller Position	0.57	0.86	0.50
	Controller Rotation	0.73	**0.87**	0.50

Regression (LR) using only specific subsets of features. In the AR environment, we distinguish between Head Position and Head Rotation features. In VR, we also consider Eyes, Controller Position, and Controller Rotation features. The ablation study was carried out both at Task-Level (Section VI-C1) and Action-Level (Section VI-C2).

1) TASK-LEVEL

Table 7 and Table 8 show the results of the ablation study for AR and VR tasks, respectively. In the AR environment, Head Rotation features are predominant in the Mental Task for identification and gender prediction. Indeed, in this task, participants were standing still and were instructed not to move their heads; however, it was plausible that their heads oscillated in singular ways, which were detected by our algorithm and leveraged for their identification. In opposition, during the navigation task, Head Position had more impact on all the targets, given that it might be associated with walking patterns. Such a pattern was used in the literature to identify people [70], and could help in Age and Gender prediction as well.

In VR, the identification stage seems to be driven mainly by Eyes features, followed by Controller features. Reasonably, eyes blinking patterns and pupils' dilatation can be person-specific [64], [65], [66], and thus act as a biometric feature. The controllers, instead, were the main interface to interact with the virtual world. Thus, it is reasonable that how a person interacts within the environment helps in the identification. This result aligns with recent findings on video games using mice and keyboards to profile users [46]. Therefore, we could expect AR identification to achieve better performances if such sensors are available, particularly eyes trackers, as reasoned before in Section VI-A. In predicting the age, the Controller features yield the best performance. This finding can result from younger people being more familiar with joystick usage. When the workload is high, younger participants may pay more attention to the task rather than how to use the joystick. Moreover, in a low workload scenario, Head and Eyes features contribute similarly. On the other hand, in gender inference, the Head and Eyes features play the most significant role. Indeed, as shown in past literature, gender-based differences exist in how they visually

explore a virtual world [71]. Controller features influence the prediction mainly in high workload controller-based tasks.

2) ACTION-LEVEL

Table 9 and Table 10 report the results of the ablation study for AR and VR Actions, respectively. In AR, the Head Position has more impact than Head Rotation in predicting our target actions, especially for the walk action. This is reasonable given that such a sensor mainly records the users' walking speed. Head Rotation becomes relevant in the Button Interaction action, in which the participants could only rotate their heads, and is quite helpful to distinguish between genders. As in previous results, the age was difficult to predict. The only case in which we surpass the baseline is in the Walk action with a high workload, but the improvement is too tiny to reason about it.

Looking at VR, we notice that Head Position remains relevant to predict the gender, particularly in scenarios with a low workload. However, most of the time, the Eyes features are the main discriminant to predict our targets. In identification, Eyes reached the highest F1-Score in six out of seven actions, suggesting that these features might be the main reason behind the higher identification performances in VR rather than AR. Further, Eyes are predominant in low workload scenarios to predict the users' age. Controller features are pretty helpful in inferring the user's age, especially in high workload actions, while only small differences appear in their usage from people of different genders. Regarding the identification task, the Controller Rotation appears more useful than Controller Position. Last, it is interesting to see how in the idle actions, the Eyes play a significant role, particularly in the high workload scenario, in which we identified a person with 0.81 of F1-Score.

TABLE 9. Ablation study of sensor importance at action-level in AR. All the measures in F1-Score.

		Identification	Age	Gender
	Guessing	0.03	0.5	0.5
Low Workload	**Search**			
	Head Position	**0.60**	0.40	0.52
	Head Rotation	0.51	0.40	**0.58**
	Walk			
	Head Position	**0.77**	0.44	**0.60**
	Head Rotation	0.55	**0.47**	0.49
High Workload	**Button Interaction**			
	Head Position	0.38	**0.46**	0.52
	Head Rotation	**0.56**	0.40	**0.56**
	Search			
	Head Position	**0.62**	0.40	**0.60**
	Head Rotation	0.52	**0.43**	0.57
	Walk			
	Head Position	**0.75**	**0.51**	**0.53**
	Head Rotation	0.55	0.43	0.47

TABLE 10. Ablation study of sensor importance at action-level in VR. All the measures in F1-Score.

		Identification	Age	Gender
	Guessing	0.03	0.5	0.5
Low Workload	**Idle**			
	Head Position	0.41	0.62	**0.62**
	Head Rotation	0.44	0.69	0.59
	Eyes	**0.75**	**0.80**	0.55
	Controller Position	0.38	0.69	0.58
	Controller Rotation	0.55	0.72	0.55
	Pointer			
	Head Position	0.67	0.80	0.57
	Head Rotation	0.73	0.83	0.62
	Eyes	**0.91**	**0.86**	**0.71**
	Controller Position	0.64	0.70	0.59
	Controller Rotation	0.83	0.81	0.51
	Button Interaction			
	Head Position	0.50	0.72	**0.63**
	Head Rotation	0.55	0.73	0.56
	Eyes	**0.85**	**0.78**	0.61
	Controller Position	0.47	0.72	0.58
	Controller Rotation	0.71	0.75	0.60
	Physical Interaction			
	Head Position	0.59	0.75	0.62
	Head Rotation	0.56	0.81	**0.63**
	Eyes	**0.87**	0.80	0.61
	Controller Position	0.63	0.74	0.57
	Controller Rotation	0.75	**0.85**	0.56
High Workload	**Idle**			
	Head Position	0.49	0.75	0.60
	Head Rotation	0.47	0.76	0.55
	Eyes	**0.81**	0.77	**0.63**
	Controller Position	0.46	**0.79**	0.50
	Controller Rotation	0.65	**0.79**	0.49
	Button Interaction			
	Head Position	0.57	0.69	0.56
	Head Rotation	0.65	0.66	0.55
	Eyes	**0.93**	**0.83**	**0.67**
	Controller Position	0.50	0.77	0.61
	Controller Rotation	0.73	0.72	0.61
	Physical Interaction			
	Head Position	0.63	0.82	0.54
	Head Rotation	0.63	0.81	0.62
	Eyes	0.71	0.87	**0.66**
	Controller Position	0.66	**0.90**	0.45
	Controller Rotation	**0.79**	0.86	0.49

VII. DISCUSSION

Literature offers some examples of profiling either in AR or VR, only on specific tasks, and through specific features (motion-based [6], [7], eye-tracking-based [8]). Furthermore, to the best of our knowledge, research work testing gender

and age profiling in immersive technologies is scarce. In our work, we covered these points by combining all the above-mentioned aspects and performing users' identification and profiling in two virtual-based scenarios, one involving AR and the other involving VR. The selected datasets present differences and similarities, offering a wide range of exemplary behaviors that can occur when immersed in XR. Indeed, we specifically aimed to propose a general framework that can accurately profile a user across diverse tasks, actions taken, and scenarios. We thus developed a generic pipeline and analyzed differences between profiling algorithms and features across different tasks. Specifically, we demonstrated to what extent users can be profiled during walking, searching for landmarks in the surroundings, pointing to a virtual keyboard for typing, and operating on a virtual robot both via controller-based interaction and physical actions. Remarkably, both virtual environments simulated highly realistic scenarios, and most of these behaviors were executed under high and low workloads, giving good insights into realistic applications of virtual technologies in the field.

The results show that users can be identified and profiled both in AR and VR, with higher VR accuracy. Specifically, in AR, user identification reached good results within the walking action at a low workload, while in VR, the identification algorithm was particularly successful when users performed more physical actions (i.e., pointing, physically interacting with the virtual robot) under a higher workload. As observed from the ablation study, this was mainly due to the additional eye-tracking sensors embedded in the VR but not in the AR headset. Indeed, while in VR the eye features had the most significant impact, the head movements were most influential on the AR users' identification.

When detecting age, instead, our algorithms were not accurate in AR. This was plausibly related to the low age variability of the tested sample, as the age of participants included in the experiment ranged between 19 and 29. Differently, in VR, we worked on an experimental sample whose age ranged between 23 and 69 years old, resulting in detecting the user's age reasonably accurately. Age detection performed better in most physical actions and interactions than those involving just joysticks and controller buttons, specifically under a higher workload. Interestingly, eye parameters had the most significant impact on age detection in all actions but in the physical interactions, in which the controller position and rotation were more important.

On gender profiling, instead, we observed how the walking activity was again the most prominent in helping detect the user's gender in AR, with the head position being the most influential sensor. Differently, in VR, our algorithms performed better during the pointing action and under actions at a high workload. In this case, the eye-related behaviors demonstrated the most considerable influence on gender detection during both these actions. In agreement with AR findings, the head position is quite relevant. Both findings

align with the literature on the different eye and head movement behaviors between men and women.

A. LIMITATIONS

The intrinsic differences between AR and VR devices, and the different nature of the tasks that our participants executed, prevented us from directly comparing the accuracy of profiling between the two technologies. However, such a comparison was out of the scope of our investigation. In this work, we were mainly interested in building a general framework that could serve to profile users during different tasks and across different technologies and scenarios. Therefore, while such differences generated some methodological limitations, in turn, they also highlighted the generability of the proposed framework. Only secondarily, we leveraged some similarities between the tasks executed in AR and VR (i.e., button interaction, and two levels of workload as generated via dual-tasking) to reason about potential profiling differences in similar actions or workloads for these two technologies. However, we want to stress that, even if we adopted a single framework for XR, they remain substantially different. Future studies could focus on analyzing data coming from the same participants engaged in the same activities in both VR and AR, to assess whether differences related to the technical apparatus of XR affect the ease of profiling.

A second limitation was the narrow age range of AR participants, which resulted in poor classification performance. However, we believe such a result is still valuable since it demonstrates that people of similar ages interact similarly with AR devices; therefore, precise age profiling could require significant effort and might not be feasible in every situation.

VIII. CONCLUSION

In conclusion, our work thoroughly studied users' profiling in XR technologies. We proposed a general profiling framework that can potentially infer any private information in any virtual scenario, and could serve as a simple yet powerful baseline for future works. Our results highlight that VR profiling is more straightforward than AR. Through our ablation study, we found eye sensors to be particularly useful in all our predictions (i.e., identification, age, gender), explaining why AR and VR perform differently. Although we are aware of the technical challenges of accurately detecting eye movements in the real world, our findings highlight the importance of incorporating eye-tracking technologies into AR headsets. Our results strongly impact single application areas of XR technologies, such as VR-based industrial robotics and everyday use of wearable AR devices, but also more generally the fast-growing Metaverse. In fact, we pave the way for further researches in XR privacy, proposing a solid inference framework that can be adapted to different virtual technologies and contexts.

In the future, we plan to conduct more experiments on a higher participant pool, which will permit defining finer targets' granularity, including additional private information (e.g., personality traits). We will also focus on which sensors and activities led to higher risks of profiling, and design privacy-preserving techniques while maintaining data utility. Last, we intend to perform a more precise comparison between AR and VR technologies.

REFERENCES

[1] M. Stylianos, "Metaverse," *Encyclopedia*, vol. 2, no. 1, pp. 486–497, 2022.

[2] K. Kamenov. (2017). *Immersive Experience—The 4th Wave in Tech: Learning the Ropes*. Accessed: Sep. 2022. [Online]. Available: https://www.linkedin.com/pulse/immersive-experience-4th-wave-tech-learning-ropes-kamen-kamenov/

[3] M. R. Miller, F. Herrera, H. Jun, J. A. Landay, and J. N. Bailenson, "Personal identifiability of user tracking data during observation of 360-degree VR video," *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, Oct. 2020.

[4] K. Pfeuffer, M. J. Geiger, S. Prange, L. Mecke, D. Buschek, and F. Alt, "Behavioural biometrics in VR: Identifying people from body motion and relations in virtual reality," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–12.

[5] C. E. Rogers, A. W. Witt, A. D. Solomon, and K. K. Venkatasubramanian, "An approach for user identification for head-mounted displays," in *Proc. ACM Int. Symp. Wearable Comput. (ISWC)*, 2015, pp. 143–146.

[6] S. Li, A. Ashok, Y. Zhang, C. Xu, J. Lindqvist, and M. Gruteser, "Whose move is it anyway? Authenticating smart wearable devices using unique head movement patterns," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2016, pp. 1–9.

[7] T. Mustafa, R. Matovu, A. Serwadda, and N. Muirhead, "Unsure how to authenticate on your VR headset? Come on, use your head!" in *Proc. 4th ACM Int. Workshop Secur. Privacy Anal.*, 2018, pp. 23–30.

[8] J. Steil, I. Hagestedt, M. X. Huang, and A. Bulling, "Privacy-aware eye tracking using differential privacy," in *Proc. 11th ACM Symp. Eye Tracking Res. Appl.*, Jun. 2019, pp. 1–9.

[9] J. Liebers, M. Abdelaziz, L. Mecke, A. Saad, J. Auda, U. Gruenefeld, F. Alt, and S. Schneegass, "Understanding user identification in virtual reality through behavioral biometrics and the effect of body normalization," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2021, pp. 1–11.

[10] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum," *Proc. SPIE*, vol. 2351, pp. 282–292, Dec. 1995.

[11] F. Nenna, M. Zorzi, and L. Gamberini, "Augmented reality as a research tool: Investigating cognitive-motor dual-task during outdoor navigation," *Int. J. Hum.-Comput. Stud.*, vol. 152, Aug. 2021, Art. no. 102644.

[12] F. Nenna, V. Orso, D. Zanardi, and L. Gamberini, "The virtualization of human–robot interactions: A user-centric workload assessment," *Virtual Reality*, vol. 27, no. 1, Jul. 2022. [Online]. Available: https://link.springer.com/article/10.1007/s10055-022-00667-x

[13] A. Berni and Y. Borgianni, "Applications of virtual reality in engineering and product design: Why, what, how, when and where," *Electronics*, vol. 9, no. 7, p. 1064, Jun. 2020.

[14] A. A. Malik, T. Masood, and A. Bilberg, "Virtual reality in manufacturing: Immersive and collaborative artificial-reality in design of human–robot workspace," *Int. J. Comput. Integr. Manuf.*, vol. 33, no. 1, pp. 22–37, Jan. 2020.

[15] C. Linn, S. Bender, J. Prosser, K. Schmitt, and D. Werth, "Virtual remote inspection—A new concept for virtual reality enhanced real-time maintenance," in *Proc. 23rd Int. Conf. Virtual Syst. Multimedia (VSMM)*, Oct./Nov. 2017, pp. 1–6.

[16] J. Xiao, P. Wang, H. Lu, and H. Zhang, "A three-dimensional mapping and virtual reality-based human–robot interaction for collaborative space exploration," *Int. J. Adv. Robotic Syst.*, vol. 17, no. 3, 2020.

[17] F. G. Pratt]icò and F. Lamberti, "Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA," *Comput. Ind.*, vol. 129, Aug. 2021, Art. no. 103446.

[18] J. J. Roldán, E. Crespo, A. Martín-Barrio, E. Peña-Tapia, and A. Barrientos, "A training system for industry 4.0 operators in complex assemblies based on virtual reality and process mining," *Robot. Comput.-Integr. Manuf.*, vol. 59, pp. 305–316, Oct. 2019.

[19] Q. Liu, Y. Wang, Q. Tang, and Z. Liu, "Do you feel the same as i do? Differences in virtual reality technology experience and acceptance between elderly adults and college students," *Frontiers Psychol.*, vol. 11, Sep. 2020, Art. no. 573673.

[20] B. Chavez and S. Bayona, "Virtual reality in the learning process," in *Proc. World Conf. Inf. Syst. Technol.* Midtown Manhattan, NY, USA: Springer, 2018, pp. 1345–1356.

[21] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Comput. Educ.*, vol. 147, Apr. 2020, Art. no. 103778.

[22] S. Z. A. Ansari, V. K. Shukla, K. Saxena, and B. Filomeno, "Implementing virtual reality in entertainment industry," in *Cyber Intelligence and Information Retrieval*. Midtown Manhattan, NY, USA: Springer, 2022, pp. 561–570.

[23] K. Vasista, "Augmented reality vs. virtual reality," *Central Asian J. Math. Theory Comput. Sci.*, vol. 3, no. 3, pp. 1–4, 2022.

[24] M. A. M. Abdelmaged, "Implementation of virtual reality in healthcare, entertainment, tourism, education, and retail sectors," Ph.D. dissertation, Munich Pers. RePEc Arch. (MPRA), Nov. 2021. [Online]. Available: https://mpra.ub.uni-muenchen.de/110491/

[25] S. Sharma and A. Bumb, "Product placement in entertainment industry: A systematic review," *Quart. Rev. Film Video*, vol. 39, no. 1, pp. 103–119, Jan. 2022.

[26] M. Conti, J. Gathani, and P. P. Tricomi, "Virtual influencers in online social media," *IEEE Commun. Mag.*, vol. 60, no. 8, pp. 86–91, Aug. 2022.

[27] A. W. K. Yeung, A. Tosevska, and E. Klager, "Virtual and augmented reality applications in medicine: Analysis of the scientific literature," *J. Med. Internet Res.*, vol. 23, no. 2, 2021, Art. no. e25499.

[28] M. Birlo, P. J. E. Edwards, M. Clarkson, and D. Stoyanov, "Utility of optical see-through head mounted displays in augmented reality-assisted surgery: A systematic review," *Med. Image Anal.*, vol. 77, Apr. 2022, Art. no. 102361.

[29] R. H. Taylor, A. Menciassi, G. Fichtinger, P. Fiorini, and P. Dario, "Medical robotics and computer-integrated surgery," in *Springer Handbook of Robotics*. Midtown Manhattan, NY, USA: Springer, 2016, pp. 1657–1684.

[30] M. G. Maggio and R. S. Calabrò, "Virtual reality and cognitive rehabilitation after traumatic brain injury," in *Diagnosis and Treatment of Traumatic Brain Injury*. Elsevier, 2022, pp. 497–506.

[31] N. M. van der Kolk, N. M. de Vries, R. P. C. Kessels, H. Joosten, A. H. Zwinderman, B. Post, and B. R. Bloem, "Effectiveness of home-based and remotely supervised aerobic exercise in Parkinson's disease: A double-blind, randomised controlled trial," *Lancet Neurol.*, vol. 18, no. 11, pp. 998–1008, Nov. 2019.

[32] D. Kim and Y. Choi, "Applications of smart glasses in applied sciences: A systematic review," *Appl. Sci.*, vol. 11, no. 11, p. 4956, May 2021.

[33] Y. Zhao, E. Kupferstein, B. V. Castro, S. Feiner, and S. Azenkot, "Designing AR visualizations to facilitate stair navigation for people with low vision," in *Proc. 32nd Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2019, pp. 387–402.

[34] M. Z. Iqbal and A. G. Campbell, "Adopting smart glasses responsibly: Potential benefits, ethical, and privacy concerns with ray-ban stories," *AI Ethics*, vol. 2, pp. 1–3, Apr. 2022.

[35] (2022). *Big Data Analytics*. Accessed: Sep. 2022. [Online]. Available: https://www.ibm.com/analytics/big-data-analytics

[36] P. R. Center. (2019). *Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information*. Accessed: Sep. 2022. [Online]. Available: https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

[37] E. Ozturk, "Privacy in emerging technologies," Ph.D. dissertation, Dept. Comput. Sci., UC Irvine, Irvine, CA, USA, 2021.

[38] I. Consulting. (2018). *General Data Protection Regulation*. Accessed: Sep. 2022. [Online]. Available: https://gdpr-info.eu/

[39] D. Adams, A. Bah, C. Barwulor, N. Musaby, K. Pitkin, and E. M. Redmiles, "Ethics emerging: The story of privacy and security perceptions in virtual reality," in *Proc. Fourteenth Symp. Usable Privacy Secur. (SOUPS)*, 2018, pp. 427–442.

[40] M. Abraham, P. Saeghe, M. Mcgill, and M. Khamis, "Implications of XR on privacy, security and behaviour: Insights from experts," in *Proc. Nordic Hum.-Comput. Interact. Conf.*, Oct. 2022, pp. 1–12.

[41] R. Di Pietro and S. Cresci, "Metaverse: Security and privacy issues," in *Proc. 3rd IEEE Int. Conf. Trust, Privacy Secur. Intell. Syst. Appl. (TPS-ISA)*, Dec. 2021, pp. 281–288.

[42] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T. H. Luan, and X. Shen, "A survey on metaverse: Fundamentals, security, and privacy," *IEEE Commun. Surveys Tuts.*, early access, Sep. 7, 2022, doi: 10.1109/COMST.2022.3202047.

[43] V. Nair, G. Munilla Garrido, and D. Song, "Going incognito in the metaverse," 2022, *arXiv:2208.05604*.

[44] M. Barni, T. Bianchi, D. Catalano, M. Di Raimondo, R. Donida Labati, P. Failla, D. Fiore, R. Lazzeretti, V. Piuri, F. Scotti, and A. Piva, "Privacy-preserving fingercode authentication," in *Proc. 12th ACM Workshop Multimedia Secur.*, Sep. 2010, pp. 231–240.

[45] M. Cardaioli, M. Conti, P. P. Tricomi, and G. Tsudik, "Privacy-friendly de-authentication with BLUFADE: Blurred face detection," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2022, pp. 197–206.

[46] M. Conti and P. P. Tricomi, "PvP: Profiling versus player! Exploiting gaming data for player recognition," in *Proc. Int. Conf. Inf. Secur.* Midtown Manhattan, NY, USA: Springer, 2020, pp. 393–408.

[47] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 15, pp. 5802–5805, Apr. 2013.

[48] D. Jurgens, T. Finethy, J. McCorriston, Y. Xu, and D. Ruths, "Geolocation prediction in Twitter using social networks: A critical analysis and review of current practice," in *Proc. Int. AAAI Conf. Web Social Media*, 2015, pp. 188–197.

[49] Y. Zhang, N. Gao, and J. Chen, "A practical defense against attribute inference attacks in session-based recommendations," in *Proc. IEEE Int. Conf. Web Services (ICWS)*, Oct. 2020, pp. 355–363.

[50] D. J. Liebling and S. Preibusch, "Privacy considerations for a pervasive eye tracking world," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput., Adjunct Publication*, Sep. 2014, pp. 1169–1177.

[51] S. Berkovsky, R. Taib, I. Koprinska, E. Wang, Y. Zeng, J. Li, and S. Kleitman, "Detecting personality traits using eye-tracking data," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2019, pp. 1–12.

[52] G. Fernández, F. Manes, L. E. Politi, D. Orozco, M. Schumacher, L. Castro, O. Agamennoni, and N. P. Rotstein, "Patients with mild Alzheimer's disease fail when using their working memory: Evidence from the eye tracking technique," *J. Alzheimer's Disease*, vol. 50, no. 3, pp. 827–838, 2016.

[53] T. Kinnunen, F. Sedlak, and R. Bednarik, "Towards task-independent person authentication using eye movement signals," in *Proc. Symp. Eye-Tracking Res. Appl. (ETRA)*, 2010, pp. 187–190.

[54] J. Steil, M. Koelle, W. Heuten, S. Boll, and A. Bulling, "PrivacEye: Privacy-preserving head-mounted eye tracking using egocentric scene image and eye movement features," in *Proc. 11th ACM Symp. Eye Tracking Res. Appl.*, 2019, pp. 1–10.

[55] E. Bozkir, O. Günlü, W. Fuhl, R. F. Schaefer, and E. Kasneci, "Differential privacy for eye tracking with temporal correlations," *PLoS ONE*, vol. 16, no. 8, Aug. 2021, Art. no. e0255979.

[56] B. David-John, K. Butler, and E. Jain, "For your eyes only: Privacy-preserving eye-tracking datasets," in *Proc. Symp. Eye Tracking Res. Appl.*, Jun. 2022, pp. 1–6.

[57] J. L. Kröger, O. H.-M. Lutz, and F. Müller, "What does your gaze reveal about you? On the privacy implications of eye tracking," in *Proc. IFIP Int. Summer School Privacy Identity Manage.* Midtown Manhattan, NY, USA: Springer, 2019, pp. 226–241.

[58] D. Arp, E. Quiring, F. Pendlebury, A. Warnecke, F. Pierazzi, C. Wressnegger, L. Cavallaro, and K. Rieck, "Dos and don'ts of machine learning in computer security," in *Proc. USENIX Secur. Symp.*, 2022, pp. 1–18.

[59] A. Nadeem, D. Vos, C. Cao, L. Pajola, S. Dieck, R. Baumgartner, and S. Verwer, "SoK: Explainable machine learning for computer security applications," 2022, *arXiv:2208.10605*.

[60] F. Nenna and L. Gamberini, "The influence of gaming experience, gender and other individual factors on robot teleoperations in VR," in *Proc. 17th ACM/IEEE Int. Conf. Human-Robot Interact. (HRI)*, Mar. 2022, pp. 945–949.

[61] P. Paolo Tricomi, L. Facciolo, G. Apruzzese, and M. Conti, "Attribute inference attacks in online multiplayer video games: A case study on Dota2," 2022, *arXiv:2210.09028*.

[62] A. F. Kramer, "Physiological metrics of mental workload: A review of recent progress," in *Multiple-Task Performance*. London, U.K.: Taylor & Francis, 2020, pp. 279–328.

[63] S. Mathôt, "Pupillometry: Psychology, physiology, and function," *J. Cognition*, vol. 1, no. 1, Feb. 2018.

[64] G. Bargary, J. M. Bosten, P. T. Goodbourn, A. J. Lawrance-Owen, R. E. Hogg, and J. D. Mollon, "Individual differences in human eye movements: An oculomotor signature?" *Vis. Res.*, vol. 141, pp. 157–169, Dec. 2017.

[65] C. Fawcett, E. Nordenswan, S. Yrttiaho, T. Häikiö, R. Korja, L. Karlsson, H. Karlsson, and E.-L. Kataja, "Individual differences in pupil dilation to others' emotional and neutral eyes with varying pupil sizes," *Cognition Emotion*, vol. 36, no. 5, pp. 928–942, 2022.

[66] S. Aminihajibashi, T. Hagen, M. D. Foldal, B. Laeng, and T. Espeseth, "Individual differences in resting-state pupil size: Evidence for association between working memory capacity and pupil size variability," *Int. J. Psychophysiol.*, vol. 140, pp. 1–7, Jun. 2019.

[67] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[68] A. Grabowski, J. Jankowski, and M. Wodziński, "Teleoperated mobile robot with two arms: The influence of a human-machine interface, VR training and operator age," *Int. J. Hum.-Comput. Stud.*, vol. 156, Dec. 2021, Art. no. 102707.

[69] K. Z. Li, R. T. Krampe, and A. Bondar, "An ecological approach to studying aging and dual-task performance," in *Cognitive Limitations in Aging and Psychopathology*. Cambridge, U.K.: Cambridge Univ. Press, 2005, pp. 190–218.

[70] R. Katiyar, V. K. Pathak, and K. Arya, "A study on existing gait biometrics approaches and challenges," *Int. J. Comput. Sci. Issues*, vol. 10, no. 1, p. 135, 2013.

[71] B. Abdi Sargezeh, N. Tavakoli, and M. R. Daliri, "Gender-based eye movement differences in passive indoor picture viewing: An eye-tracking study," *Physiol. Behav.*, vol. 206, pp. 43–50, Jul. 2019.

PIER PAOLO TRICOMI is currently pursuing the Ph.D. degree in brain, mind and computer science (BMCS) with the University of Padova, Italy. He is part of the Security and Privacy (SPRITZ) Research Group, under the supervision of Prof. Mauro Conti. His research interests include security and privacy in online social networks and video games, with a particular focus on machine learning and deep learning applications for user profiling, private data inference, and fake news detection. He has recently focused on privacy issues in the Metaverse, authentication, and de-authentication systems. More information can be found at: https://www.math.unipd.it/ tricomi/

FEDERICA NENNA received the M.Sc. degree in cognitive applied psychology from the University of Padova, Italy, in 2018, where she is currently pursuing the Ph.D. degree in brain, mind, and computer science (BMCS), under the supervision of Prof. Luciano Gamberini. Her M.Sc. thesis was conducted at the Berlin Brain/Body Imaging Laboratory (BeMoBIL), TU Berlin. Her research interests include humans interacting with VR-based technologies and their underlying cognitive processes. What characterizes her research is a multimodal approach, combining self-reports with behavioral, eye movements, and brain data collected in applied unconstrained environments.

LUCA PAJOLA received the M.Sc. degree in computer science from the University of Padova, Italy, in 2018, where he is currently pursuing the Ph.D. degree with the School of Brain, Mind and Computer Science. Here, he is part of the SPRITZ Security and Privacy Research Group, under the supervision of Prof. Mauro Conti. He is conducting research in fields, including security and machine learning. More information can be found at: https://www.math.unipd.it/ pajola/
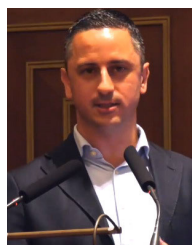
MAURO CONTI (Fellow, IEEE) received the Ph.D. degree from the Sapienza University of Rome, Italy, in 2009. After his Ph.D., he was a Postdoctoral Researcher at Vrije Universiteit Amsterdam, The Netherlands. In 2011, he joined as an Assistant Professor with the University of Padua, Italy, where he became an Associate Professor, in 2015, and a Full Professor, in 2018. Since 2020, he has been the Head of Studies of the Master's Degree in Cybersecurity, University of Padua. He is currently a Full Professor with the University of Padua. He is also affiliated with the TU Delft and the University of Washington, Seattle. His research is also funded by companies, including Cisco, Intel, and Huawei. His main research interest includes security and privacy. In this area, he published more than 400 papers in topmost international peer-reviewed journals and conferences. He is a Senior Member of ACM. He is a member of the Blockchain Expert Panel of the Italian Government. He is a fellow of the Young Academy of Europe. He has been awarded with a Marie Curie Fellowship (2012) by the European Commission and with a Fellowship by the German DAAD (2013). He was the Program Chair of TRUST 2015, ICISS 2016, WiSec 2017, ACNS 2020, and CANS 2021, and the General Chair of SecureComm 2012, SACMAT 2013, NSS 2021, and ACNS 2022. He is the Editor-in-Chief of IEEE Transactions on Information Forensics and Security, an Area Editor-in-Chief of IEEE Communications Surveys and Tutorials, and an Associate Editor of several journals, including IEEE Transactions on Dependable and Secure Computing and IEEE Transactions on Network and Service Management. More information can be found at: https://www.math.unipd.it/ conti/

LUCIANO GAMBERINI received the master's degree in general psychology (cognitive ergonomics) and the Ph.D. degree in experimental psychology (human–computer interaction). He is currently a Full Professor with the University of Padua, the Director of the HIT Research Centre, and the Vice-Coordinator of the Ph.D. School in brain, mind and computer science. Since 2000, he has taught several courses in the area of work and organizational psychology, as well as in the area of social and cognitive ergonomics and mediated communication, in Padua and Trento. He is the author of more than 100 peer-reviewed scientific papers and conference presentations in the area of human–computer interaction. He is a member of the scientific board and the Chair of international conferences and journals, including the International Workshop on Presence, *Persuasive Technology*, *Cybertherapy*, Symbiotic Interaction, ACM CHI, CHItaly ACM SIG, *PsychNology Journal*, *Cyberpsychology, Behavior and Social Networking*, *Cybertherapy and Rehabilitation*, and the Emerging Communication Series by IOS Press.

○ ○ ○