EyeSeeIdentity: Exploring Natural Gaze Behaviour for Implicit User Identification during Photo Viewing

Yasmeen Abdrabou*, Mariam Hassib†, Shuqin Hu ‡, Ken Pfeuffer §, Mohamed Khamis ¶, Andreas Bulling‖, and Florian Alt**

*Lancaster University, UK, Email: y.abdrabou@lancaster.ac.uk

†Fortiss Research Institute of the Free State of Bavaria, Germany, Email: hassib@fortiss.org

‡LMU Munich, Germany, Email: agnes.hsq@gmail.com

§Aarhus University, Denmark, Email: ken@cs.au.dk

¶University of Glasgow, UK, Email: mohamed.khamis@glasgow.ac.uk

‖University of Stuttgart, Germany, Email: andreas.bulling@vis.uni-stuttgart.de

**University of the Bundeswehr Munich, Germany, Email: florian.alt@unibw.de

***Abstract*—Existing gaze-based methods for user identification either require special-purpose visual stimuli or artificial gaze behaviour. Here, we explore how users can be differentiated by analysing natural gaze behaviour while freely looking at images. Our approach is based on the observation that looking at different images, for example, a picture from your last holiday, induces stronger emotional responses that are reflected in gaze behavioor and, hence, is unique to the person having experienced that situation. We collected gaze data in a remote study ($N = 39$) where participants looked at three image categories: personal images, other people's images, and random images from the Internet. We demonstrate the potential of identifying different people using machine learning with an accuracy of 85%. The results pave the way towards a new class of authentication methods solely based on natural human gaze behaviour.**

I. Introduction

The ongoing pursuit towards balancing usability and security within authentication systems remains a persistent focus in both academic research and industrial applications. Experts in security often pinpoint users as a weak link due to their propensity for creating insecure passwords [1]. In response, researchers have proposed various authentication methods, including biometric authentication [2], implicit authentication [3], and continuous user authentication [4], aiming to enhance both usability and security. Despite these advancements, many authentication techniques encounter vulnerabilities such as shoulder surfing [5], lunchtime attacks [6], thermal attacks [7], smudge attacks [8], or spoofing attacks [9].

One way to address these challenges is adding an identification step to grant access to devices, even when the unlock token is known, referred to as a two/multi-factor authentication. Two general approaches exist: explicit and implicit multi-factor authentication. Regarding explicit approaches, both commercial (e.g., using one-time security token devices or sending email notifications) as well as so far purely academic (e.g., by relying on tangible objects [10]) solutions exist. Approaches to implicit multi-factor authentication include identification from behavioural cues during usage, including the gait cycle [11], browsing history [12], physical location [13], and eye-gaze behaviour during different tasks [14].

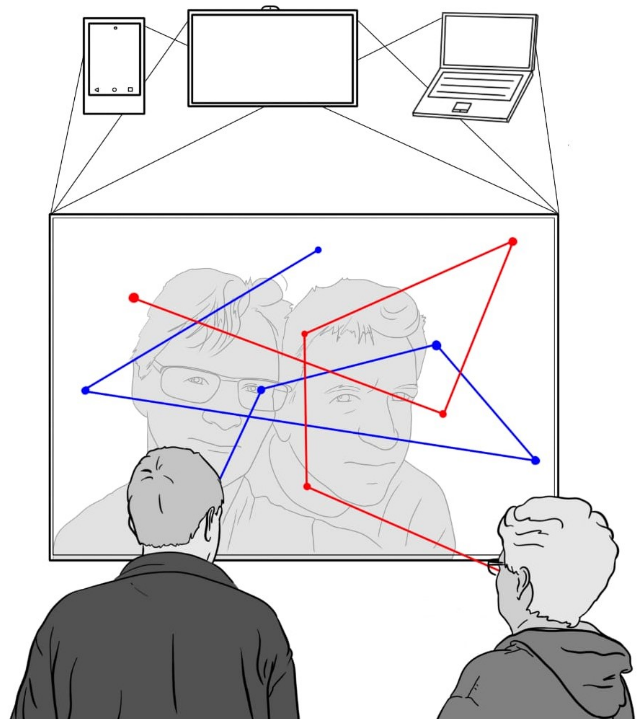

Fig. 1: We propose using eye gaze behaviour during photo viewing as an implicit identification technique, using existing background images on smartphones, tablets, laptops, and PCs. We hypothesise that users' gaze behaviour differs for the same photograph, depending on their relationship with the photo.

While eye-gaze tracking has become increasingly prevalent through smartphone front cameras and laptop webcams,

Symposium on Usable Security and Privacy (USEC) 2024
26 February 2024, San Diego, CA, USA
ISBN 979-8-9894372-5-2
https://dx.doi.org/10.14722/usec.2024.23057
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

its application often necessitates task initiation alongside the identification process, potentially delaying identification. Moreover, gaze-based identification typically relies on suitable stimuli [15] or artificial gaze behaviour [16].

A particularly promising, so far unexplored stimulus is photographs, encapsulating personal memories, events, and moments. People's relationships with these images are diverse, evoking distinct emotions and consequently eliciting different eye movements. These movements carry multifaceted information about individuals [17], including age [18], gender [19], emotional state [20], and even identification [21].

We explore the concept of employing gaze behaviour as an implicit user identification modality during photo viewing. The primary goal is to incorporate this approach as an identification step before authentication, for example, through leveraging the presence of a lock screen photo on personal devices. In our approach, we investigate users' gaze behaviour on several photo categories, namely, 1) personal photos, 2) photos of other unknown individuals, and 3) photos from the internet. Moreover, we study image importance and repetition effect on users' gaze behaviour. Our results show that users' gaze behaviour is significantly influenced by users' personal relationship with the displayed photos. Furthermore, we observed alterations in gaze behaviour with repeated photo exposure, yet individual gaze patterns retained their uniqueness. Our findings enable a new approach for user identification using natural gaze behaviour. To the best of our knowledge, this is the first attempt to investigate user identification from natural gaze behaviour while viewing photos.

Contribution Statement. Through this work, we make the following contributions: First, we analyze eye gaze behavioural data collected in an uncontrolled environment while looking at photographs and images of three different categories, exploring familiarity, repetition, and importance. Second, we contribute person-dependent machine learning classifiers on gaze data and propose directions for future research on user identification.

II. Related Work

Our work draws from prior work on (1) Implicit user identification techniques, (2) Gaze-based Authentication, and (3) Eye Gaze Behaviour and Image viewing.

A. Implicit User Identification

Implicit user identification and authentication approaches were introduced over a decade ago by Jakobsson et al. [3]. They explain that the concept of implicit authentication can be used as a primary or secondary authentication scheme, on any type of device that collects user behavioural and contextual information [3]. Implicit authentication mechanisms do not only authenticate the user *one time*, but also analyze user behaviour collected from sensor and usage data during a specific time span to be able to continuously authenticate the user, and hence reduce re-authentication workload [4], [22]. This concept has been further explored and expanded, and various systems have been built to use information such as typing biometrics [23]–[25], eye gaze tracking [26], gait [11] or a combination of multiple behavioural information sources (e.g., [4]). Notable comprehensive literature reviews of implicit and continuous authentication techniques include Khan et al. [27], Bo et al. [4], and most recently [28].

B. Gaze and User Identification

Different aspects of authentication using eye-gaze were explored over the years [29]. Researchers have introduced several ways where eye-gaze can be used for explicit [16], [30], implicit [26], [31], Biometric [32] multi-factor [33]–[35], and continuous authentication [36]. For example, Kasprowski et al. examined user identification by measuring the eye's reaction to different visual stimuli [37]. In 2005, Bednarik et al. proposed using eye movements as a biometric [38]. Although there are still privacy concerns [39], the identification accuracy is relatively high due to gaze biometric features. Moreover, researchers attempted to integrate gaze-based user identification into daily tasks. For example, Abdulin et al. [40] explored the feasibility and accuracy of using eye movements as biometrics while users are reading [40], Iqbal et al. investigated eye gaze while searching [14], and Eberz et al. explored gaze behaviour during watching videos [41]. Moreover, recent work points to the potential of eye gaze for interventions that increase password strength [42] and reduce password reuse [43].

C. Eye Gaze Behaviour and Images

Gaze behaviour varies depending on the task or activity. For example, Kosch et al. [44] found higher deviations of gaze points for a trajectory during smooth pursuit eye movements when the users are doing an N-back cognitive task. Iqbal et al. [45] showed the possibility of detecting the user's tasks (e.g., reading, mental reasoning) by using eye gaze patterns as each task has a unique signature of eye movements. Visual tasks, such as viewing photographs and images, have also been shown to impact eye movements. Moss et al. showed that eye movements differ between genders during natural image viewing [19]. Cantoni et al. [15] showed that people look at photographs differently, and distinctive features may be extracted while viewing photographs that can be used as a soft biometric. The authors used greyscale face images. They found that users looked differently at the faces, which can be a distinctive feature for user identification. The authors also highlighted that using different images, such as landscapes and abstract images, can improve the recognition rate [15]. Yun et al. showed that a person's gaze behaviour while freely viewing a scene contains a lot of information, not only about users' intent and what they consider as important in the scene but also about the scene's content [46]. Gomez et al. found differences in gaze behaviour based on participants' gender, age, and repeated exposure to the stimuli [47]. Massaro et al. found differences in gaze behaviour during exposure to different art genres, including natural scenes and paintings with human subjects [48]. Researchers also found differences in gaze behaviour when viewing images depending on participants' physical attraction to the photos [49] and image complexity [50].

D. Summary

Implicit authentication and identification techniques show a lot of promising potential for increased security and good usability. Gaze behaviour can be used alone or in combination with other techniques for implicit identification. However, gaze behaviour mostly requires unnatural stimuli. We fill this gap by proposing a concept for using natural eye gaze elicited during viewing photographs and images for implicit authentication.

III. Concept and Research Questions

Leveraging findings from prior work introduced in the previous section that showed that different categories of images elicit different eye movement behaviours, we aim to explore the identification of users through analysing their gaze behaviour while freely viewing photos. This work has two main objectives 1) investigating the various factors influencing users' gaze behaviour—such as photo category, importance, and repetition, and 2) constructing a model for user identification based on these patterns. In addition, an important objective of our study was to conduct it in an ecologically valid remote setting. To this end, we also chose sensors that can be employed in the users' vicinity rather than requiring users to be augmented with on-body sensors.

Our main hypothesis is that individuals exhibit unique gaze behaviour when observing the same photo due to having different relationships or memories with the photos. This was highlighted in prior work as participants felt attached to image elements that had a certain memory with [51]. Throughout our research, we aim to answer two key research questions: *RQ1:* How well can we identify individuals based on their implicit gaze behaviour during photo viewing? We investigate several aspects namely 1) photo category, 2) photo importance, and 3) photo exposure repetition. *RQ2:* Does photo importance and repeated exposure to the same photo change users' gaze behaviour? In the following, we describe our data collection methodology to explore these research questions.

IV. Data Collection

A. Study Design

To investigate users' gaze behaviour while viewing photos, we designed a remote within-subjects study for gaze data collection, consisting of two phases: *Phase 1*–photo collection; and *Phase 2*–recording and acquiring gaze data.

In *phase 1*, we asked participants to upload a selection of 10 personal photos: 'Choose 10 photos that you would like to be included in your annual yearbook, five of which could also be designated for printing out'. The yearbook scenario served as a prime to make participants think of choosing photos for this significant keepsake and historical record, documenting one's events. The printing allowance acts as an indicator of photo importance. Participants were then prompted to rate the importance of each photo on a scale from 1 (least important) to 10 (highly important) and provide reasons, such as connections to special events, individuals, or places.

In *phase 2*, participants were exposed to 30 distinct photos. Each photo was shown three times, based on a Latin square, to study the repetition effect. Images were shown for 5 seconds. The selection of 5 seconds aligns with literature, suggesting that the optimal duration for displaying photos is between 4 to 6 seconds before participants lose focus [52]. During this phase, participants viewed 1) their uploaded photos, 2) others' photos (personal photos provided by the authors), and 3) 10 photos sourced from the Internet (showing worldwide tourist destinations). The study had 3 independent variables: 1) photo category, 2) photo importance 3) repetition; and one dependent variable which is gaze features. We also captured participants' screen resolution and displayed image resolution.

Approval for the study was obtained from the University ethics board. Moreover, the study's call emphasised the importance of participants providing consent to temporarily store their photos on the study server, owned by the University.

B. Recruitment

The recruitment for the study was conducted through multiple channels, including dissemination through the University mailing lists and various social media platforms. The call for participation directed potential participants to a website where they could provide their consent for data collection and storage. Upon consenting, participants could proceed to upload their photos for the study.

C. Apparatus

To obtain highly ecologically valid gaze data, we opted to conduct the study remotely. We utilized the Gazerecorder API[1] with a frame rate of 33 Hz. The GazeRecorder API is specifically designed for Webcam-based eye-tracking, integrated within web browsers. We implemented a website using HTML, CSS, and Javascript, hosted on our University server, where we integrated the eye tracker code, collected the photos, and displayed the tasks and questionnaires.

D. Procedure

Upon clicking the study link, participants were guided to complete the consent and demographics forms. Subsequently, they progressed to phase 1, tasked with uploading 10 photos, assigning importance ratings, and providing reasons for their choices. Transitioning to phase 2, participants began with eye tracker calibration and were then prompted to freely view the displayed photos. Following each photo display, a greyscale photo appeared for 2 seconds, acting as a separator and establishing a baseline for eye gaze data.

Following the photo display phase, participants proceeded to an accuracy test, aimed at assessing the precision of their eye gaze data accuracy upon concluding the study. Subsequently, participants were prompted to complete a post-study questionnaire asking about their experiences with eye trackers and eye fatigue. The study duration was appr. 40 minutes (20 minutes per phase). Participants received acompensation of 15 Euros. Figure 2 depicts the study procedure.

E. Limitations

Our study has several limitations. Although we explicitly specified the requirement for participants to be seated in a well-lit, quiet environment without interruptions, we encountered challenges in ensuring these conditions. To address these uncontrollable factors, during the data cleaning phase, we implemented measures to mitigate their impact (discussed in detail in section VI). Specifically, we assessed the accuracy of eye tracking. Any datasets exhibiting evident discrepancies or incomplete results were consequently removed from our analysis to enhance data integrity and reliability. Another limitation is testing the repetition effect in a short time frame. Future work should look into gaze behaviour over longer periods. Finally, it is important to highlight that the classification

[1] https://gazerecorder.com/, last accessed Nov. 17, 2023

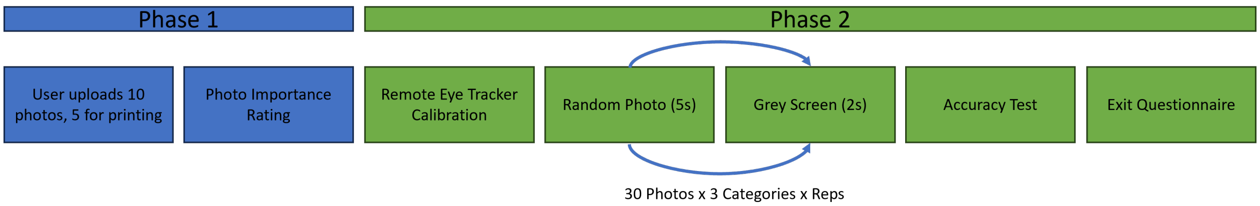

Fig. 2: Study Procedure with the two phases depicted

accuracy is bound to gaze data accuracy which is limited by the camera's framerate. Although eye tracking accuracy is affected by several conditions such as light, we still achieved a relatively good accuracy, however, it could be enhanced by using a high frame rate eye tracker.

V. Features and Classification Approach

We describe our step-by-step process to understand eye gaze behaviour while viewing different photo categories with the ultimate goal of identifying users based on their gaze behaviour. First, we extracted gaze features, required for classification and tested their statistical significance. Second, we built and tested different classifiers based on these features. Finally, we implemented an identification classifier leveraging a comprehensive set of combined features across all photo categories.

A. Feature Extraction

We extracted 10 low-level gaze features from the raw gaze data, inspired by the literature [53], [54] to characterize gaze behaviour during photo viewing effectively. Additionally, we identified Areas of Interest (AOI) by leveraging visually salient regions. To extract these AOIs, we employed the Shreelock et al. [55] algorithm based on Graph-Based Visual Saliency (GBVS) model. This approach enables identifying which can predict fixations on photos with superior performance to the original visual saliency algorithm [51], [56].

- **Fixation count**: Number of fixations performed during single photo viewing.
- **Fixation duration**: Time for which users dwelled with their eyes on the stimuli.
- **Average fixation duration**: The sum of the duration of all the fixations divided by the sum of fixations.
- **Average saccadic duration**: The saccadic duration is calculated by subtracting the timestamps of two consecutive fixations per photo.
- **Average saccadic length**: The average distance between two fixations per photo.
- **AOI Fixation count**: Count of fixations performed during single photo viewing in all salient areas.
- **AOI Fixation duration**: Time for which users dwelled on the stimuli in all salient areas.
- **AOI Average fixation duration**: The sum of the fixations' duration divided by the sum of fixations for all salient areas.
- **AOI Average saccadic duration**: The saccadic duration is calculated by subtracting the timestamps of two consecutive fixations for all salient areas.
- **AOI Average saccadic length**: The average distance between two fixations for all salient areas per photo.

B. Classification Approach

Our classifiers correlate a feature vector computed from a time window of data to construct a user identification classifier. Initially, we built three classifiers to explore potential variations in users' gaze behaviour across distinct, independent variables: 1) photo category (personal, others', and general), 2) photo importance (important vs. unimportant), and 3) repetition (first-time vs. repeated). We report the average and highest AUC.

All classifiers operate on a user-dependent basis, focusing on unique gaze behaviour. We analyzed three types of features: 1) overall photo features, 2) AOI features, and 3) combined features. The comparison involved assessing two commonly utilized classifiers in existing literature [57], [58]: Support Vector Machines (SVM), and Random Forest (RF). We empirically fine-tuned the hyperparameters for optimization using a limited set of values.

The classifiers for photo category, importance, and repetition were individually trained on each user's data. The data was initially cleaned by removing outliers, followed by feature computation for gaze behaviour. Outliers were removed using the z-score algorithm, which eliminates data points beyond three times the standard deviation [59]. The photo importance classifier was only trained on user's personal photos labeled for printing and non printing as indications of importance. Subsequently, the data was split into training and test sets (employing the leave-one-out approach) and underwent 5-fold cross-validation. The resulting best and average AUC scores were recorded for all participants across all classifiers. Finally, the identification classifier is then built on data from all participants and all features.

VI. Results

We present and analyze the collected data from our study. We start with data cleaning and pre-processing. Then, we present a data overview, statistical analysis of the computed gaze features, and finally the classification results.

A. Participants

We recruited 44 participants (20 females) aged between 20 and 42 ($M = 25.9, SD = 4.6$). Participants had different

TABLE I: Friedmann tests for gaze features while participants view photos with different category levels, namely their own photos, others' photos, and general photos from the internet (significant results in bold, $P < .05$).

Gaze Features	Own Photos	Others' Photos	Generic Photos	Friedmann F, P
Fixation Count	M = 6.9, SD = 1.8	M = 7.1, SD = 1.7	M = 6.9, SD = 1.6	1.039, $>.05$
Fixation Duration	**M =1714.2, SD = 638.5**	**M = 1741.2, SD = 528.6**	**M = 1609.2, SD = 488.6**	**5.751, = .008**
Avg Fixation Duration	M = 263, SD = 72	M = 256, SD = 47.3	M = 243, SD =38.8	2.647, $>.05$
Avg Saccadic Duration	M = 1780.8, SD = 477.5	M = 1701.1, SD = 440.1	M = 1834.8, SD = 366.5	1.346, $>.05$
Avg Saccadic Length	M = 552.5, SD = 440.8	M = 537.3, SD = 427.7	M = 525.6, SD = 432.3	2.463, $>.05$
AOI Fixation Count	M = 2, SD = 1.6	M = 1.8, SD = 1.5	M = 1.9, SD = 1.4	.401, $>.05$
AOI Fixation Duration	M = 529.6, SD = 466.8	M = 487.1, SD = 438.2	M = 493.1, SD = 394.2	.365 $>.05$
AOI Avg Fix Duration	M = 99.4, SD = 59.8	M = 100.9, SD = 60.6	M = 102.1, SD = 65.7	.039, $>.05$
AOI Avg Saccadic Duration	M = 616.7, SD = 391.7	M = 572.6, SD = 334.6	M = 629, SD = 401.8	.424, $>.05$
AOI Avg Saccadic Length	M = 133.5, SD = 91.1	M = 135.8, SD = 81.7	M = 151.4, SD = 86.8	.905, $>.05$

nationalities, and the majority were students (35). Our participants had backgrounds in computing science, psychology, business administration, and engineering. Finally, our participants were novice eye tracker users (1.2 on a scale from 1=novice to 5=experienced).

B. Data Cleaning and Pre-Processing

Our data cleaning is based on two aspects: 1) eye-tracking failure and accuracy test results. We observed numerous instances of missing gaze data among participants, possibly indicating interruptions during the study or the presence of an overlay screen that disrupted gaze recording. Consequently, we had to exclude five participants due to insufficient gaze data, leading us to conduct subsequent analyses with a reduced cohort of 39 participants.

For the accuracy test results, following Tobii's guidelines [60] for eye gaze accuracy testing, which shows that the accuracy value should not exceed 1 degree (i.e., the gaze point should not deviate more than 55 pixels from the target center), we conducted accuracy assessments. Detailed calculations are outlined in Abdrabou et al. [57]. All the remaining 39 participants exhibited satisfactory accuracy test results, eliminating the need for further data removal.

In our data pre-processing phase, we eliminated outliers using the z-score algorithm, as explained earlier. Additionally, we implemented data normalization to standardize all features within the same range, thereby reducing data dimensionality [61]. This normalization procedure ensured that all features were brought to a consistent magnitude level for analysis.

C. Data Overview and Gaze Data Analysis

For the data overview, we used webcams as eye trackers. The data collection sampling rate was 33 frames per second. This led to the collection of 2970 frames per user and an overall of 115k eye-tracking data frames from all our participants.

To investigate users' gaze behaviour, we first start with statistical analysis. Below, we reflect on the statistical analysis of gaze behaviour while photo viewing. Our data were non-normally distributed (confirmed by Shapiro-Wilk and Anderson-Darling tests). Hence, unless otherwise stated, we perform non-parametric tests and report on mean values (M).

1) Photo Category Statistical Results: For photo categories, a Friedman test showed a statistically significant effect of the photo category only on users' fixation duration ($\chi^2(37) = 10.595, P < .05$). Using the Wilcoxon pairwise test with Bonferroni correction showed a statistically significant difference between users' fixation duration on their own photos ($M = 21714.30; SD = 638.505$), others' photos ($M = 1741.22; SD = 528.62$), and general photos ($M = 1609.28; SD = 488.67$); see Table I. This indicates that participants spend less time viewing their own photos compared to when they view others' photos or general photos. Utilizing the Wilcoxon signed-rank test with Bonferroni correction, we identified a statistically significant difference in the pairwise comparison between viewing one's own photo ($M = 1714; SD = 638$) and viewing general photos ($M = 1609; SD = 488.6$), with a significance level of $p < .05$. Although other aspects lack statistical significance, they offer insights into user behaviour. The table shows that when users view their own photos, they tend to have fewer but longer fixations. Additionally, they exhibit a pattern of having longer eye movement distances, indicating a more skimming visual behaviour compared to the other photo categories [62].

2) Photo Importance Statistical Results: For photo importance, a Wilcoxon test with Bonferroni correction showed no statistically significant effect of photo importance on eye gaze behaviour; see Table II. From the table, we can see that when observing important photos, participants had longer fixation durations, extended saccadic durations, and cover shorter distances. These patterns suggest a tendency towards scrutinizing the photos more closely [62].

3) Photo Repetition Statistical Results: To illustrate the repetition effects, Figure 3 shows an example of the gaze path of one user while looking at their own image over three repetitions (top three figures), when looking at others' images (middle three figures), and when looking at general images (bottom three figures). As we can see, the user's gaze path is very similar over repetitions when looking at their own images, whereas, when looking at others' images, the user first starts by scanning larger areas of the image in the first repetition, then focusing on particular areas in subsequent repetitions as they become more familiar with it. When looking at general images, the user's gaze is sporadic and the scanpath is different in each repetition. These images however, do not illustrate the variation in the of fixations and only shows the gaze scan path

TABLE II: Wilcoxon signed-rank tests with Bonferroni correction for gaze features while participants viewed photos with different importance levels, namely important vs. unimportant (significant results in bold, $P < .05$).

Gaze Features	**Important Photos**	**Unimportant Photos**	**Wilcoxon Z, P**
Fixation Count	M = 6.9, SD = 1.6	M = 6.9, SD = 1.76	-.039, $>$.05
Fixation Duration	M = 1669.8, SD = 540.1	M = 1739.4, SD = 642.2	-1.195, $>$.05
Avg Fixation Duration	M = 250.7, SD = 41.9	M = 272.3, SD = 103.2	-.549, $>$.05
Avg Saccadic Duration	M = 1793.1, SD = 389.1	M = 1718.8, SD = 522.3	-.196, $>$.05
Avg Saccadic Length	M = 537.8, SD = 448.1	M = 539.2, SD = 438.8	-1.489, $>$.05
AOI Fixation Count	M = 1.9, SD = 1.4	M = 2.4, SD = 1.8	-1.489, $>$.05
AOI Fixation Duration	M = 500.6, SD = 414	M = 644.6, SD = 639.9	-1.705, $>$.05
AOI Avg Fix Duration	M = 101, SD = 55.9	M = 116.1, SD = 90.1	-.803, $>$.05
AOI Avg Saccadic Duration	M = 621, SD = 329.7	M = 697.4, SD = 518.9	-.647, $>$.05
AOI Avg Saccadic Length	M = 141.3, SD = 77.3	M = 157.6, SD = 107.6	-.823, $>$.05

which looks similar from an abstract view, however it differ in the features themselves e.g. fixation duration, fixation count, saccadic length, etc.

Looking at the overall effect of repetitions on all participants, we found a statistically significant effect of the repetition on the overall photo features, meaning that users change their gaze behaviour when seeing the same photo more than once. However, we could not find a statistically significant effect of the repetition on users' gaze behaviour features inside the areas of interest, which can indicate that users tend to have similar eye movements inside saliency areas. However, this needs further investigation; see Table III for statistical results.

The table further demonstrates that participants exhibit a decrease in the number and size of fixations, accompanied by longer saccades and increased gaze distance across repetitions. This shift might suggest a transition in their visual behaviour from scrutinizing to skimming as they become more familiar with the photos over time [62].

D. Classification Results

We compared the performance of two different models: Support Vector Machine (SVM) and Random Forest (RF). We conducted four classifications: 1) photo category classifier, 2) photo importance classifier, 3) photo repetition classifier, and 4) user identification classifier. We ran each of them on 1) photo generic features, 2) saliency areas features, and 3) combined features. Below, we reflect on each. Scores reported below are the average individual scores across all folds.

1) Photo Categories Classifier: Table IV presents the distinct classification outcomes for the photo categories classifier. Our investigation revealed that the random forest model achieved higher accuracy on an individual user level (91%). Given the highly subjective nature of photo perception, we prioritize classifiers delivering consistently high accuracy for individuals. Our findings indicate that, within the random forest classifier, the saliency area features (91%) demonstrated higher classification accuracy compared to both generic photo features (40%) and combined features (33%). Similar results were also found within the SVM classifier accuracies. Note, that all results surpassed the classifier baseline. In our comparison across three categories within the classifier, we adjusted the baseline to 33.3%, and all outcomes exceeded this threshold.

Additionally, Figure 4 showcases a detailed breakdown of classification results per user. This figure emphasizes the AUC (Area Under the Curve) specifically for the random forest classifier utilizing saliency area features due to its better accuracy. It is important to note that due to insufficient data, certain participants' classifiers could not be trained, resulting in the absence of 5 participants in our analysis. The figure further demonstrates the substantial individuality in photo viewing and its impact on classification results. These outcomes vary widely, showcasing instances of high accuracy reaching up to 100%, while in contrast, some, like participants 16 and 22, barely reach the baseline.

2) Photo Importance Classifier: To understand gaze behaviour while viewing (un)important photos, we ran ML classifiers only on the participants' own uploaded photos and not all seen photos. In our assessment of the photo importance classifier, we discovered that random forest outperformed SVM in terms of individual and average accuracy. Upon analyzing the features, we found that generic features had slightly higher average accuracy across participants in contrast to salient areas.

We additionally examine the individual classification accuracy illustrated in Figure 5. Given the high accuracy demonstrated by the random forest model, we reflect on individual classifiers within the RF model specifically focusing on generic features, which exhibited a higher average accuracy. The visual representation in the figure indicates that the classification between important and unimportant is mostly accurate across the majority of participants. Nevertheless, due to individual variability, certain participants, like participant 15, exhibit lower classification accuracy.

3) Photo Repetition Classifier: Looking at the photo repetition classification results (cf., Table VI),the RF provided higher accuracy (88%) than SVM (63%) on individual levels and averages. We also found that within the RF model, saliency features provided higher accuracy than other features. This finding was inconsistent in the SVM model, where generic features yielded higher accuracy. In our comparison across categories within the classifier, we adjusted the baseline to 33.3%

To explore individual variances, Figure 6 displays individual AUC results for the RF classifier utilising saliency features, which yielded the highest accuracy. Note that due to limited data, classifiers for two participants could not be

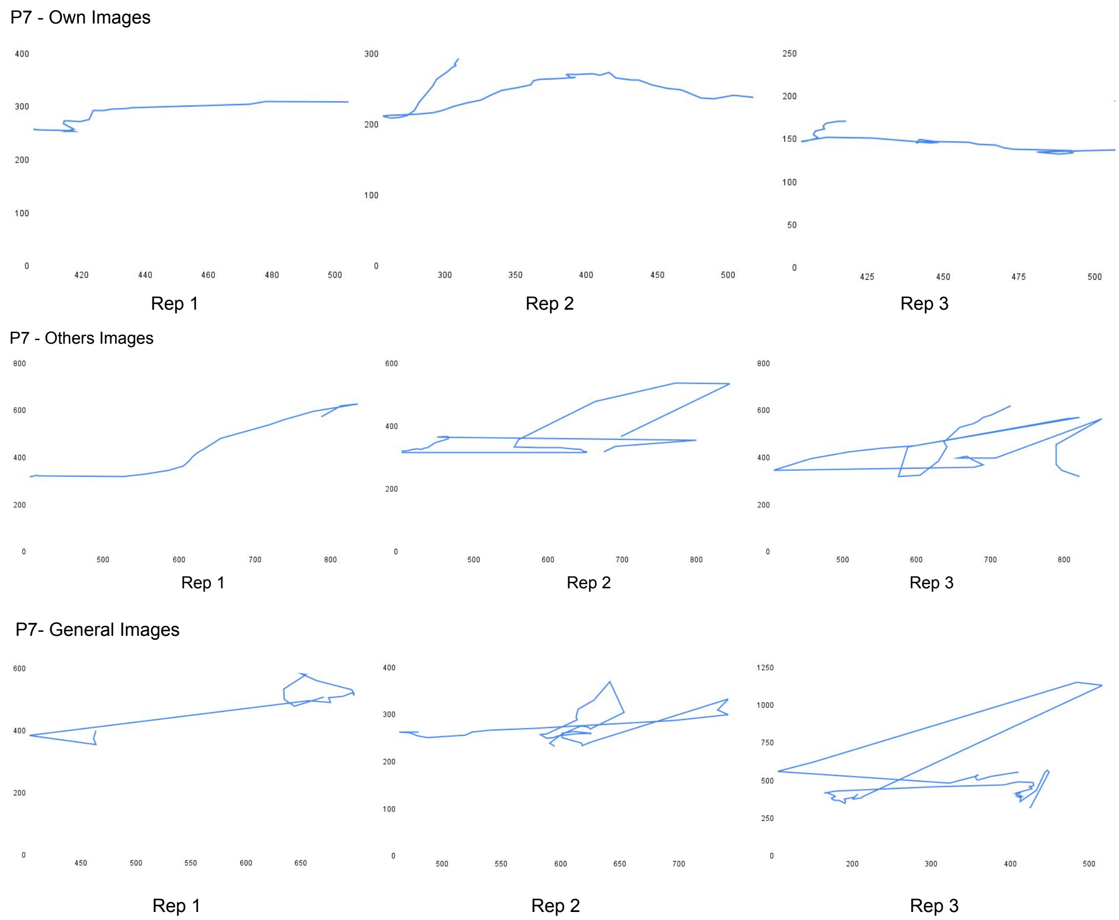

Fig. 3: Repetition Effect on Users' Gaze Movements - Top line shows the gaze path for three repetitions with one of the user's own images - Middle line shows the gaze path for three repetitions with the user looking at others' images - Bottom line shows the gaze path for three repetitions with the user looking at general images.

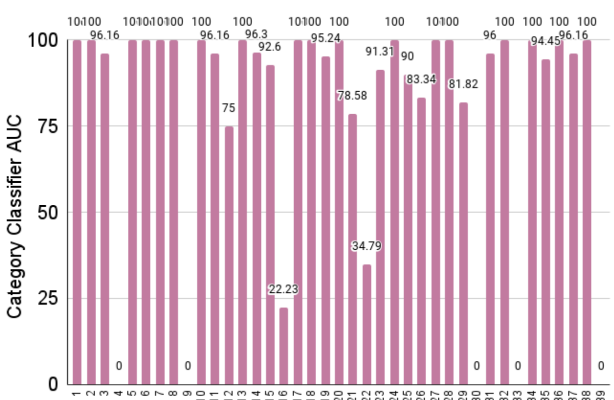

Fig. 4: Photo Category Classifier AUC per User for the Random Forest Classifier on Saliency Areas Features

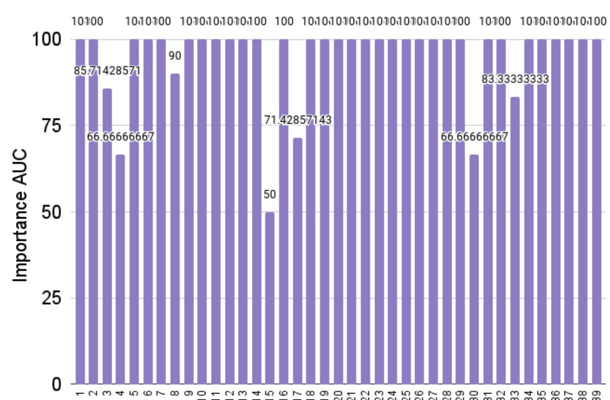

Fig. 5: Photo Importance Classifier AUC per User for the Random Forest Classifier on Generic Features

trained, leading to their exclusion from our analysis. The figure highlights individual disparities in users' gaze patterns concerning photo repetition. For instance, participant 8 exhibits distinct gaze behaviour, indicating a different gaze trajectory across repetitions, presumably due to familiarity with the photo. On the other side, participants 2, 29, and 34 showcase almost identical behaviour in their gaze patterns when viewing the picture repeatedly 3 times.

TABLE III: Friedmann tests for gaze features while viewing photos in different repetitions (significant results in bold, $P < .05$).

Gaze Features	1st Repetition	2nd repetition	3rd Repetition	Friedmann F, P
Fixation Count	M = 8.5, SD = 2.1	M = 6.1, SD = 1.9	M = 5.8, SD = 1.7	**569.3, <.05**
Fixation Duration	M = 1964.9, SD = 699.8	M = 1541.8, SD = 563	M = 1496.4, SD = 546.6	**334.7, <.05**
Avg Fixation Duration	M = 235.6, SD = 58.1	M = 266.1, SD = 73.9	M = 269.9, SD = 60.6	**1119.8, <.05**
Avg Saccadic Duration	M = 1713.5, SD = 456	M = 1769.4, SD = 481.8	M = 1805.5, SD = 476.2	**828.8, <.05**
Avg Saccadic Length	M = 502.8, SD = 419.9	M = 589.4, SD = 474	M = 557.3, SD = 439.8	**59.8, <.05**
AOI Fixation Count	M = 2.5, SD = 1.7	M = 1.6, SD = 1.3	M = 1.5, SD = 1.1	**73.8, <.05**
AOI Fixation Duration	M = 595.7, SD = 498.4	M = 443.9, SD = 385.6	M = 446.8, SD = 340.9	**59.1, <.05**
AOI Avg Fix Duration	M = 93.7, SD = 53.5	M = 99.7, SD = 65	M = 106.1, SD = 59.7	**132.3, <.05**
AOI Avg Saccadic Duration	M = 635.8, SD = 371.5	M = 538.5, SD = 313.2	M = 611.8, SD = 360.1	**149.3, <.05**
AOI Avg Saccadic Length	M = 132.8, SD = 74.3	M = 137, SD = 83.6	M = 146.1, SD = 77.5	**143, <.05**

TABLE IV: Photo categories showing best and (average) classification AUC for SVM and Random Forest classifiers across the different features

Feature/ Classifier	SVM	Random Forest
Generic Photo Features	53.79% (43.6%)	50% (40.40%)
Saliency areas Features	70% (50.41%)	100% (91.76%)
Combined Features	66.54% (45.70%)	69.23% (33.20%)

TABLE V: Photo importance showing best and (average) classification AUC for SVM and Random Forest classifiers across the different features

Feature/ Classifier	SVM	Random Forest
Generic Photo Features	85.71% (56.75%)	100% (95.22%)
Saliency areas Features	100% (50.26%)	100% (73.75%)
Combined Features	76.19% (57.68%) 1	100% (92.11%)

4) User Identification Classifier: We conducted an analysis using an identification classifier that incorporated all available features. Our findings revealed that the random forest classifier achieved the highest accuracy of 85% in identifying users. To have a deeper understanding of the classifier's accuracy, we examined the contributing features using SHAP [63], a tool designed to elucidate a machine learning model's output by assessing each feature's impact on predictions. In Figure 7, we present the feature importance plot generated by the random forest classifier. Notably, our observations indicate that average saccadic length, fixation duration, as well as average fixation and saccadic duration, significantly influence the classifier's accuracy. Additionally, we've included the confusion matrix for all participants in Figure 8, which offers a comprehensive overview of the classifier's performance in distinguishing between different user categories.

VII. Discussion and Future Work

A. Gaze for User Identification

Our approach showed that for many users it was possible to identify the user from their gaze data, collected remotely, using a user-dependent model. However, building a user-independent model proved to be challenging due to the fact that the amount of gaze data collected during the five seconds of exposure using a remote eye tracker is relatively low. This would probably also be the case when using mobile cameras for eye tracking on smartphones, while users are moving, or while they are in inadequate lighting conditions. Throughout our analysis, we found that the most important feature for identifying users from the 10 extracted features is the average saccadic length and fixation duration. This may mean that when looking at familiar, important, or memorable photographs, longer fixations will be found. This suggests that users adapt their visual behaviour, transitioning between skimming and scrutinising based on factors such as photo category, importance, and memorability [62], [64]. Future work should investigate other factors affecting users' gaze behvaiour while looking at photos, and consider using different eye tracekrs with different resolutions and framerates.

TABLE VI: Photo repetition showing best and (average) classification AUC for SVM and Random Forest classifiers across the different features

Feature/ Classifier	SVM	Random Forest
Generic Photo Features	63.89% (37.2%)	66.67% (49.39%)
Saliency areas Features	46.67% (27.42%)	**80% (49.55%)**
Combined Features	63.89% (28.83%)	44.26% (47.79%)

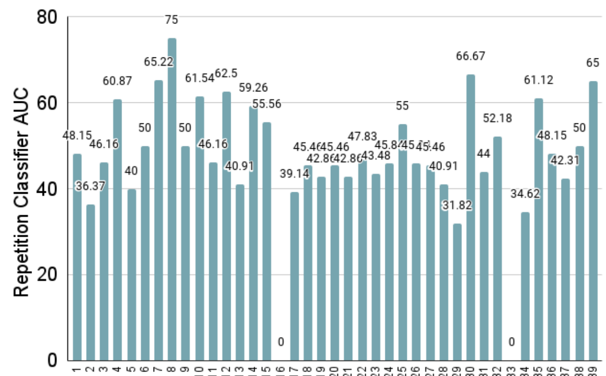

Fig. 6: Repetition Classifier AUC per User for the Random Forest Classifier on Saliency Areas Features

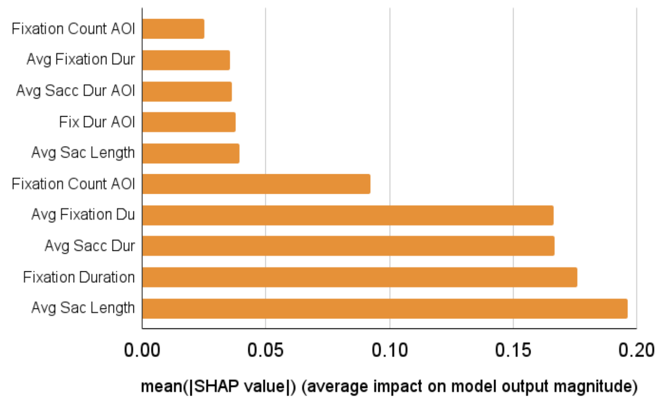

Fig. 7: Results of the feature importance analysis across the tested features for the user identification classifier.

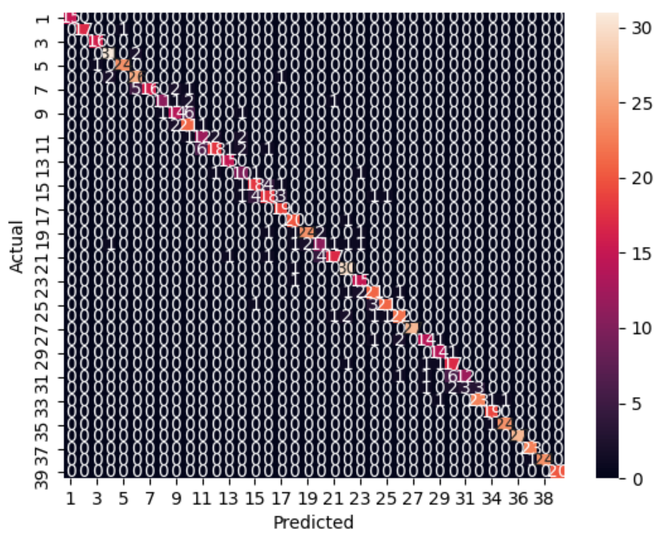

Fig. 8: Confusion Matrix for User Identification Classifier

B. Image Importance and Repetition

In our results section, we provided an analysis of the effect of image repetition and image importance on the users' gaze behaviour. We have seen that the fixation duration feature was found to decrease by repetition during our study, although gaze scanpath may be the same. Hence, when using our technique in the future, photos may need to be altered, for example, by changing their contrast or saturation level, to nudge more fixations or by altering between a set of images with every device unlock trigger. While we have not tested that in our study, we find this to be an interesting direction for future work. Additionally, we would need to test the repetition effect over a longer period of time as well as introduce more variety in the photo collection.

Regarding image importance, we found no significant differences between images labelled by users as highly important vs. unimportant. We hypothesise that users assign varying degrees of importance to photos based on their relationships and associated memories, aspects that were not explored in our study. Our research findings highlight the importance of periodically altering these photos to sustain the efficacy of this approach. This can be achieved by either consistently changing the entire photo, adjusting specific elements, such as saturation, colour mode, or removing parts of the image. We believe that the integration of this strategy is feasible, especially considering that some devices already offer features like automatically changing background images with every unlock or providing live backgrounds with dynamic elements. Consequently, future research should delve into understanding the connection between users and their photos, going beyond solely assessing subjective importance levels.

C. Gaze Behaviour Across Image Categories

Overall, we have explored images in three categories: personal photographs chosen by the participants, photographs of other people, and photos from the internet of tourist locations. Our research findings suggest that using personalised photos as backgrounds is more effective in eliciting unique user behaviour. Reflecting on our age group, our approach might be more suitable for younger generations with high-resolution front cameras and a wide range of photos to choose from. Future work should look into different age groups and user activities, specially since face visibility which is crucial for eye gaze tracking might not fully visible [65]. Moreover, future work could investigate other types of photos or features of photos such as photos with high salience areas, and abstract photos vs photos with more than one person.

D. Integration Into Existing Systems

Identifying users based on photo categories is possible even with a small amount of gaze data (e.g., lower fixation count captured in an unaltered environment with a remote, low-framerate eye tracker). We believe that this concept can be used as a line of defense before authentication/password entry on any device that has a camera by employing a background image on the lock screen and tracking the users' gaze behaviour once the screen is activated. By leveraging the variability in individuals' behaviors when viewing photos, a novel two-factor authentication (2FA) method can be implemented on devices. This involves incorporating photo-based authentication into the lock screen, complementing existing authentication methods for enhanced security. Furthermore, this approach can serve as a continuous means of user identification and authentication during regular phone usage. To advance this concept, we suggest that future research should explore the impact of application layout on users' home screens and its correlation with gaze behaviour. However, using our technique comes with some privacy concerns. Alsaker et al. showed that there are several aspects affecting the wider adoption of eye trackers on smartphones affecting the users themselves and also the bystanders such as the gaze estimation algorithms' transparency and the developers' credibility of what is collected, saved, and shared [66]. Furthermore, nudging users towards using particular images could be further investigated. For example, Abdrabou et al. [67] showed that participants were affected by the background image while choosing alphanumeric passwords reflected in the gaze heatmaps. The authors highlighted a possible threat of manipulating users by creating carefully designed images that guide users; gaze behaviour, which required further investigation.

VIII. Conclusion

In this work, we investigated our novel approach of using photo viewing as an implicit user identification technique leveraging the existence of background photos on most smart devices. We conducted a remote study, where we collected users' eye gaze behaviour while viewing several image categories with different importance levels and on three repetitions. Our results showed that users' gaze behaviour is significantly different across the different photo types. Moreover, gaze behaviour changes over time while viewing the same image multiple times. Our results present a promising avenue for implicit user identification, bypassing the necessity for artificial gaze stimuli and contrived gaze behaviour.

Acknowledgements

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 101021229, GEMINI: Gaze and Eye Movement in Interaction). It was also funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr (Voice of Wisdom). dtec.bw is funded by the European Union – NextGenerationEU. Finally, A. Bulling was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No 801708.

References

[1] A. Hanamsagar, S. S. Woo, C. Kanich, and J. Mirkovic, *Leveraging Semantic Transformation to Investigate Password Habits and Their Causes*. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–12. [Online]. Available: https://doi.org/10.1145/3173574.3174144

[2] D. Bhattacharyya, R. Ranjan, F. Alisherov, M. Choi *et al.*, "Biometric authentication: A review," *International Journal of u-and e-Service, Science and Technology*, vol. 2, no. 3, pp. 13–28, 2009.

[3] M. Jakobsson, E. Shi, P. Golle, R. Chow *et al.*, "Implicit authentication for mobile devices," in *Proceedings of the 4th USENIX conference on Hot topics in security*, vol. 1. USENIX Association, 2009, pp. 25–27.

[4] C. Bo, L. Zhang, T. Jung, J. Han, X.-Y. Li, and Y. Wang, "Continuous user identification via touch and movement behavioral biometrics," in *2014 IEEE 33rd International Performance Computing and Communications Conference (IPCCC)*. IEEE, 2014, pp. 1–8.

[5] M. Eiband, M. Khamis, E. Von Zezschwitz, H. Hussmann, and F. Alt, "Understanding shoulder surfing in the wild: Stories from users and observers," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 4254–4265.

[6] S. Eberz, K. Rasmussen, V. Lenders, and I. Martinovic, "Preventing lunchtime attacks: Fighting insider threats with eye movement biometrics," in *Network and Distributed System Security (NDSS) Symposium*. Internet Society, 2015.

[7] Y. Abdelrahman, M. Khamis, S. Schneegass, and F. Alt, "Stay cool! understanding thermal attacks on mobile-based user authentication," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 3751–3763. [Online]. Available: https://doi.org/10.1145/3025453.3025461

[8] A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens," in *Proceedings of the 4th USENIX Conference on Offensive Technologies*, ser. WOOT'10. USA: USENIX Association, 2010, p. 1–7.

[9] S. Kumar, S. Singh, and J. Kumar, "A comparative study on face spoofing attacks," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 2017, pp. 1104–1108.

[10] K. Marky, M. Schmitz, V. Zimmermann, M. Herbers, K. Kunze, and M. Mühlhäuser, "3d-auth: Two-factor authentication with personalized 3d-printed items," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–12. [Online]. Available: https://doi.org/10.1145/3313831.3376189

[11] X. Zeng, X. Zhang, S. Yang, Z. Shi, and C. Chi, "Gait-based implicit authentication using edge computing and deep learning for mobile devices," *Sensors*, vol. 21, no. 13, p. 4592, 2021.

[12] M. Abramson and D. W. Aha, "User authentication from web browsing behavior." in *FLAIRS conference*, 2013, pp. 268–273.

[13] Y. De Mulder, G. Danezis, L. Batina, and B. Preneel, "Identification via location-profiling in gsm networks," in *Proceedings of the 7th ACM workshop on Privacy in the electronic society*, 2008, pp. 23–32.

[14] S. T. Iqbal and B. P. Bailey, "Using eye gaze patterns to identify user tasks," in *The Grace Hopper Celebration of Women in Computing*, vol. 4, no. 01, 2004, p. 2004.

[15] V. Cantoni, C. Galdi, M. Nappi, M. Porta, and D. Riccio, "Gant: Gaze analysis technique for human identification," *Pattern Recognition*, vol. 48, no. 4, pp. 1027 – 1038, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320314000697

[16] A. De Luca, M. Denzel, and H. Hussmann, "Look into my eyes! can you guess my password?" in *Proceedings of the 5th Symposium on Usable Privacy and Security*, ser. SOUPS '09. New York, NY, USA: Association for Computing Machinery, 2009. [Online]. Available: https://doi.org/10.1145/1572532.1572542

[17] J. L. Kröger, O. H.-M. Lutz, and F. Müller, *What Does Your Gaze Reveal About You? On the Privacy Implications of Eye Tracking*. Cham: Springer International Publishing, 2020, pp. 226–241. [Online]. Available: https://doi.org/10.1007/978-3-030-42504-3_15

[18] M. Erbilek, M. Fairhurst, and M. C. D. C. Abreu, "Age prediction from iris biometrics," in *5th International Conference on Imaging for Crime Detection and Prevention (ICDP 2013)*, 2013, pp. 1–5.

[19] F. J. M. Moss, R. Baddeley, and N. Canagarajah, "Eye movements to natural images as a function of sex and personality," *PLoS One*, vol. 7, no. 11, p. e47870, 2012.

[20] V. Raudonis, A. P. Taraseviciene, G. Dervinis, G. K. Raudone, and A. Vilkauskas, "Evaluation of human emotion from eye motions."

[21] V. Cantoni, M. Musci, N. Nugrahaningsih, and M. Porta, "Gaze-based biometrics: an introduction to forensic applications," *Pattern Recognition Letters*, vol. 113, pp. 54–57, 2018.

[22] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," *IEEE Signal Processing Magazine*, vol. 33, no. 4, pp. 49–61, 2016.

[23] T. Feng, J. Yang, Z. Yan, E. M. Tapia, and W. Shi, "Tips: Context-aware implicit user identification using touch screen in uncontrolled environments," in *Proceedings of the 15th workshop on mobile computing systems and applications*, 2014, pp. 1–6.

[24] B. Draffin, J. Zhu, and J. Zhang, "Keysens: Passive user authentication through micro-behavior modeling of soft keyboard interaction," in *Mobile Computing, Applications, and Services: 5th International Conference, MobiCASE 2013, Paris, France, November 7-8, 2013, Revised Selected Papers 5*. Springer, 2014, pp. 184–201.

[25] J. Gurary, Y. Zhu, N. Alnahash, and H. Fu, "Implicit authentication for mobile devices using typing behavior," in *Human Aspects of Information Security, Privacy, and Trust: 4th International Conference, HAS 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings 4*. Springer, 2016, pp. 25–36.

[26] R. Biedert, M. Frank, I. Martinovic, and D. Song, "Stimuli for gaze based intrusion detection," in *Future Information Technology, Application, and Service: FutureTech 2012 Volume 1*. Springer, 2012, pp. 757–763.

[27] H. Khan, A. Atwater, and U. Hengartner, "A comparative evaluation of implicit authentication schemes," in *Research in Attacks, Intrusions and Defenses: 17th International Symposium, RAID 2014, Gothenburg,*

Sweden, September 17-19, 2014. Proceedings 17. Springer, 2014, pp. 255–275.

[28] R. Ryu, S. Yeom, D. Herbert, and J. Dermoudy, "A comprehensive survey of context-aware continuous implicit authentication in online learning environments," *IEEE Access*, 2023.

[29] C. Katsini, Y. Abdrabou, G. E. Raptis, M. Khamis, and F. Alt, "The role of eye gaze in security and privacy applications: Survey and future hci research directions," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–21. [Online]. Available: https://doi.org/10.1145/3313831.3376840

[30] Y. Abdrabou, M. Khamis, R. M. Eisa, S. Ismail, and A. Elmougy, "Just gaze and wave: Exploring the use of gaze and gestures for shoulder-surfing resilient authentication," in *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, ser. ETRA '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3314111.3319837

[31] A. Boehm, D. Chen, M. Frank, L. Huang, C. Kuo, T. Lolic, I. Martinovic, and D. Song, "Safe: Secure authentication with face and eyes," in *2013 International Conference on Privacy and Security in Mobile Systems (PRISMS)*, 2013, pp. 1–8.

[32] C. Song, A. Wang, K. Ren, and W. Xu, "Eyeveri: A secure and usable approach for smartphone user authentication," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, 2016, pp. 1–9.

[33] B. Hoanca and K. Mock, "Methods and systems for multiple factor authentication using gaze tracking and iris scanning," Jul. 26 2011, uS Patent 7,986,816.

[34] M. Khamis, M. Hassib, E. v. Zezschwitz, A. Bulling, and F. Alt, "Gazetouchpin: Protecting sensitive data on mobile devices using secure multimodal authentication," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ser. ICMI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 446–450. [Online]. Available: https://doi.org/10.1145/3136755.3136809

[35] C. Kumar, D. Akbari, R. Menges, S. MacKenzie, and S. Staab, "Touchgazepath: Multimodal interaction with touch and gaze path for secure yet efficient pin entry," in *2019 International Conference on Multimodal Interaction*, ser. ICMI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 329–338. [Online]. Available: https://doi.org/10.1145/3340555.3353734

[36] Y. Zhang, W. Hu, W. Xu, C. T. Chou, and J. Hu, "Continuous authentication using eye movement response of implicit visual stimuli," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 4, jan 2018. [Online]. Available: https://doi.org/10.1145/3161410

[37] P. Kasprowski and J. Ober, "Eye movements in biometrics," in *Biometric Authentication*, D. Maltoni and A. K. Jain, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 248–258.

[38] R. Bednarik, T. Kinnunen, A. Mihaila, and P. Fränti, "Eye-movements as a biometric," in *Image Analysis*, H. Kalviainen, J. Parkkinen, and A. Kaarna, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 780–789.

[39] J. Steil, I. Hagestedt, M. X. Huang, and A. Bulling, "Privacy-aware eye tracking using differential privacy," *CoRR*, vol. abs/1812.08000, 2018. [Online]. Available: http://arxiv.org/abs/1812.08000

[40] E. R. Abdulin and O. V. Komogortsev, "Person verification via eye movement-driven text reading model," in *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2015, pp. 1–8.

[41] S. Eberz, K. B. Rasmussen, V. Lenders, and I. Martinovic, "Looks like eve: Exposing insider threats using eye movement biometrics," *ACM Trans. Priv. Secur.*, vol. 19, no. 1, Jun. 2016. [Online]. Available: https://doi.org/10.1145/2904018

[42] Y. Abdrabou, A. Shams, M. O. Mantawy, A. Ahmad Khan, M. Khamis, F. Alt, and Y. Abdelrahman, "Gazemeter: Exploring the usage of gaze behaviour to enhance password assessments," in *ACM Symposium on eye tracking research and applications*, 2021, pp. 1–12.

[43] Y. Abdrabou, J. Schütte, A. Shams, K. Pfeuffer, D. Buschek, M. Khamis, and F. Alt, ""'your eyes tell you have used this password before": Identifying password reuse from gaze and keystroke dynamics," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491102.3517531

[44] T. Kosch, M. Hassib, P. W. Woundefinedniak, D. Buschek, and F. Alt, "Your eyes tell: Leveraging smooth pursuit for assessing cognitive workload," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–13. [Online]. Available: https://doi.org/10.1145/3173574.3174010

[45] S. T. Iqbal and B. Bailey, "Using eye gaze patterns to identify user tasks," 2004.

[46] K. Yun, Y. Peng, D. Samaras, G. Zelinsky, and T. Berg, "Exploring the role of gaze behavior and object detection in scene understanding," *Frontiers in Psychology*, vol. 4, p. 917, 2013. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyg.2013.00917

[47] P. Gomez, A. von Gunten, and B. Danuser, "Eye gaze behavior during affective picture viewing: Effects of motivational significance, gender, age, and repeated exposure," *Biological psychology*, vol. 146, p. 107713, 2019.

[48] D. Massaro, F. Savazzi, C. Di Dio, D. Freedberg, V. Gallese, G. Gilli, and A. Marchetti, "When art moves the eyes: a behavioral and eye-tracking study," *PloS one*, vol. 7, no. 5, p. e37285, 2012.

[49] T. H. Charlotte Hall and K. Guo, "Differential gaze behavior towards sexually preferred and non-preferred human figures," *The Journal of Sex Research*, vol. 48, no. 5, pp. 461–469, 2011, pMID: 20967669. [Online]. Available: https://doi.org/10.1080/00224499.2010.521899

[50] B. Bonev, L. Chuang, and F. Escolano, "How do image complexity, task demands and looking biases influence human gaze behavior?" *Pattern Recognition Letters*, vol. 34, no. 7, pp. 723–730, 2013, scene Understanding and Behaviour Analysis. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865512001687

[51] F. Alt, M. Mikusz, S. Schneegass, and A. Bulling, "Memorability of cued-recall graphical passwords with saliency masks," in *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 191–200. [Online]. Available: https://doi.org/10.1145/3012709.3012730

[52] I. Rigas, G. Economou, and S. Fotopoulos, "Biometric identification based on the eye movements and graph matching techniques," *Pattern Recogn. Lett.*, vol. 33, no. 6, p. 786–792, Apr. 2012. [Online]. Available: https://doi.org/10.1016/j.patrec.2012.01.003

[53] G. E. Raptis, C. Katsini, M. Belk, C. Fidas, G. Samaras, and N. Avouris, "Using eye gaze data and visual activities to infer human cognitive styles: Method and feasibility studies," in *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, ser. UMAP '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 164–173. [Online]. Available: https://doi.org/10.1145/3079628.3079690

[54] R. J. Jacob and K. S. Karn, "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises," in *The mind's eye*. Elsevier, 2003, pp. 573–605.

[55] "Implementation of graph based visual saliency algorithm by j. harel, c. koch, and p. perona," https://github.com/shreelock/gbvs, [Online; accessed 19-September-2020].

[56] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Advances in neural information processing systems*, vol. 19, 2006.

[57] Y. Abdrabou, E. Karypidou, F. Alt, and M. Hassib, "Investigating user behaviour towards fake news on social media using gaze and mouse movements," in *Proceedings of the Usable Security Mini Conference*, 2023.

[58] Y. Abdelrahman, A. A. Khan, J. Newn, E. Velloso, S. A. Safwat, J. Bailey, A. Bulling, F. Vetere, and A. Schmidt, "Classifying attention types with thermal imaging and eye tracking," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, sep 2019. [Online]. Available: https://doi.org/10.1145/3351227

[59] V. Aggarwal, V. Gupta, P. Singh, K. Sharma, and N. Sharma, "Detection of spatial outlier by using improved z-score test," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 788–790.

[60] T. T. AB, "Accuracy and precision test method for remote eye trackers," 2011.

[61] B. M. S. Hasan and A. M. Abdulazeez, “A review of principal component analysis algorithm for dimensionality reduction,” *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 20–30, 2021.

[62] Q. Lohmeyer, S. Matthiesen, M. Mussgnug, and M. Meboldt, “Analysing visual behaviour in engineering design by eye tracking experiments,” in *International Symposium on Tools and Methods of Competitive Engineering (TMCE 2014), Budapest, Hungary*, 2014.

[63] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[64] Q. Lohmeyer, S. Matthiesen, M. Meboldt *et al.*, “Task-dependent visual behaviour of engineering designers-an eye tracking experiment,” in *DS 77: Proceedings of the DESIGN 2014 13th International Design Conference*, 2014, pp. 549–558.

[65] M. Khamis, A. Baier, N. Henze, F. Alt, and A. Bulling, “Understanding face and eye visibility in front-facing cameras of smartphones used in the wild,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–12. [Online]. Available: https://doi.org/10.1145/3173574.3173854

[66] N. Alsakar, Y. Abdrabou, S. Stumpf, and M. Khamis, “Investigating privacy perceptions and subjective acceptance of eye tracking on hand-held mobile devices,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 7, no. ETRA, pp. 1–16, 2023.

[67] Y. Abdrabou, M. Asbeck, K. Pfeuffer, Y. Abdelrahman, M. Hassib, and F. Alt, “Empowering users: Leveraging interface cues to enhance password security,” in *IFIP Conference on Human-Computer Interaction*. Springer, 2023, pp. 359–368.