Using Gaze Behavior and Head Orientation for Implicit Identification in Virtual Reality

Conference Paper · October 2021

DOI: 10.1145/3489849.3489880

5 authors, including:

Jonathan Liebers
University of Duisburg-Essen

Uwe Gruenefeld
University of Duisburg-Essen

Stefan Schneegaß
University of Duisburg-Essen

Using Gaze Behavior and Head Orientation for Implicit Identification in Virtual Reality

Jonathan Liebers
University of Duisburg-Essen
Essen, Germany
jonathan.liebers@uni-due.de

Patrick Horn
University of Duisburg-Essen
Essen, Germany
patrick.horn@stud.uni-due.de

Christian Burschik
University of Duisburg-Essen
Essen, Germany
christian.burschik@stud.uni-due.de

Uwe Grünefeld
University of Duisburg-Essen
Essen, Germany
uwe.gruenefeld@uni-due.de

Stefan Schneegass
University of Duisburg-Essen
Essen, Germany
stefan.schneegass@uni-due.de

ABSTRACT

Identifying users of a Virtual Reality (VR) headset provides designers of VR content with the opportunity to adapt the user interface, set user-specific preferences, or adjust the level of difficulty either for games or training applications. While most identification methods currently rely on explicit input, implicit user identification is less disruptive and does not impact the immersion of the users. In this work, we introduce a biometric identification system that employs the user's gaze behavior as a unique, individual characteristic. In particular, we focus on the user's gaze behavior and head orientation while following a moving stimulus. We verify our approach in a user study. A hybrid post-hoc analysis results in an identification accuracy of up to 75 % for an explainable machine learning algorithm and up to 100 % for a deep learning approach. We conclude with discussing application scenarios in which our approach can be used to implicitly identify users.

CCS CONCEPTS

• **Human-centered computing** → *Interaction techniques*; *Virtual reality*; • **Security and privacy** → *Usability in security and privacy*.

KEYWORDS

virtual reality, gaze-based authentication, eye tracking, implicit identification

ACM Reference Format:
Jonathan Liebers, Patrick Horn, Christian Burschik, Uwe Grünefeld, and Stefan Schneegass. 2021. Using Gaze Behavior and Head Orientation for Implicit Identification in Virtual Reality. In *27th ACM Symposium on Virtual Reality Software and Technology (VRST '21), December 8–10, 2021, Osaka, Japan.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3489849.3489880

VRST '21, December 8–10, 2021, Osaka, Japan

ACM ISBN 978-1-4503-9092-7/21/12…$15.00
https://doi.org/10.1145/3489849.3489880

1 INTRODUCTION

As Virtual Reality (VR) headsets proliferate, their sensing capabilities also improve. While first generations of devices mainly used hand-held controllers for input, head-mounted displays (HMD), nowadays, increasingly include eye tracking capabilities (e. g., the HTC Vive Pro Eye or the Pico Neo 2 Eye) as an additional input modality. Moreover, eye tracking is added to HMDs to enable foveated rendering, an approach that renders less detail outside of the area that users fixate their gaze on to save computational resources [34]. Although games are currently the main driving factor for VR headsets, more serious applications are gaining importance such as education [36], training [40], or collaboration [6]. This multitude of applications results in devices being shared among multiple users. Thus, understanding who is currently using the device becomes an important challenge. This can be either used to authenticate users and permit access to confidential information or to identify users and personalize the experiences.

The use of behavioral biometrics is a promising method for both use cases since they can be applied implicitly in the background without interrupting the user through actions they would "carry out anyway" [17]. While body movements have been in the focus of previous research on behavioral biometrics in VR [22, 25, 35, 41], gaze has not yet received considerable attention. Virtual Reality bears unique benefits for gaze-based identification, such as highly controllable brightness within the enclosed HMD and less diversion, as the eyes are shielded from the outside world.

In this work, we present a novel gaze-based behavioral biometric method that enables implicit identification in VR. We utilize the individual eye gaze behavior that occurs when users follow an accelerating stimulus with their eyes. In particular, we use the velocities and characteristics of *how* the user switches between different gaze types, as well as head orientation characteristics.

We evaluate our approach in a lab study (N=12) and were able to identify users with a cross-validated accuracy of 75 % in an explainable machine learning approach and with an accuracy of up to 100 % in a deep learning approach. We envision that this approach can be integrated in games and serious applications to identify the current user. As this identification scheme can be implemented in an implicit manner, it also establishes a zero-cost layer of additional security when used for authentication.

The contribution of this work is two-fold. First, we present a new type of behavioral biometric that exploits differences in users' gaze behaviors for identification. Second, we report on a user study ($N = 12$) that evaluates the general feasibility of the approach.

2 RELATED WORK

We discuss identification and authentication approaches utilizing gaze and taking place in Virtual Reality.

2.1 Identification and Authentication with Gaze

Gaze-based authentication is an established topic that has been in the focus of research for the last 20 years. While early approaches in this field have enabled knowledge-based authentication with gaze as an input modality, successfully preventing shoulder-surfing attacks (e. g., dwell-time based PIN entry [5, 21]), more recent approaches tend to utilize the actual behavior of the gaze. For instance, the complex patterns and characteristics of fixations and saccades form a biometric trait [14]. This can be used, for example, to derive information about a user's identity during a reading activity, as the captured information is highly individual [13]. In a more abstract way, the reflexive movement of the eyes as a reaction to a moving stimulus is also usable for the purpose of identification and authentication [42]. Rigas et al. have shown that saccadic eye movements are also individual, as their vigor and acceleration differ from user to user [37]. Furthermore, recent research has looked into the concept of using the eye tracker's calibration data for the means of identification [18].

For Virtual Reality, Khamis et al. employed smooth pursuit movements as an interaction technique in virtual reality to enter a knowledge-based component such as a PIN [20]. Alike, Mathis et al. have developed "RubikAuth", a fast and secure authentication mechanism in Virtual Reality that employ gaze, head pose and controller tapping [29, 30]. Current research directions and an extensive overview of the state-of-the-art is also provided by Katsini et al. in their work on the role of eye gaze in security and privacy applications where they highlight implicit gaze-based authentication to be one key research direction [19]. One of the proposed research directions was explored by Niitsu and Nakayama by measuring effects based on time and presentation size for biometric identification [32]. However, while gaze-based identification and authentication has been explored rather frequently, to our knowledge, the feasibility of the approaches in Virtual Reality is an underexplored topic, as only a small number of approaches has been validated in VR.

It is important to note that, in gaze-based authentication, only the approaches that are associated with biometrics can be employed in an implicit authentication. Implicit authentication is a form of authentication that does not interrupt the user's interaction and is performed through actions that the user would carry out anyway [17]. They can, in general, add a zero-cost layer of additional security but their error rate might make it necessary that another form of authentication should be requested from the user.

2.2 Identification and Authentication in VR

Nowadays, most methods for authentication in VR still use a handheld controller-based PIN, password entry, or pattern lock, all of which are derived from mobile devices [47]. However, George et al. have shown that such traditional authentication methods are inherently insecure for usage in VR, as up to 18% of the authentication attempts can be breached by someone shoulder surfing the controller input motion [12]. As the VR user is immersed in the virtual world, they cannot perceive their environment or the potential shoulder surfer that observes the input motion. Moreover, these approaches, similar to most knowledge-based methods form explicit authentication schemes that place the workload on the user to remember their PIN or password and which require time to be entered [12].

This conflicting situation substantiates the need for new approaches to identification and authentication in VR. Pfeuffer et al. have shown that body motion and body relations can be used for authentication [35], which is a pure biometric method. Analogously, the head and movement patterns that can be collected from the HMD also form means of biometric identification [31, 41]. Furthermore, knowledge-based methods exist for VR, such as 3D passwords [2] and multi-modal techniques that employ gaze and head orientation to improve the usability and observation resistance of such 3D passwords [11]. More complex task behaviors, such as motions associated with certain interactions (e. g., throwing a ball at a target), also bear a large amount of individual information [22] and it has been shown that these principles can be extended to other interactions [25].

Equipping an HMD with eye tracking provides further opportunities. Luo et al. have shown that electrooculography can be integrated into an HMD, using the complete human visual system as a physiological biometric [28]. In addition, Boutros et al. utilized the images that the cameras of the integrated eye tracking unit of the HMD had collected in order to generate iris and periocular biometrics [4].

Summarizing, one can state that the field of applicable authentication schemes in VR is very diverse, ranging from pure knowledge-based methods [47] to multimodal methods that employ biometrics besides knowledge-based components [11] to pure biometric methods [22]. Nevertheless, the unique benefits of implicit authentication such as being transparent to the user and reducing the user's workload [17] can primarily be associated with pure biometric authentication schemes. Here, in particular, gaze-based authentication seems to increase in importance [19], as eye trackers are more frequently included in HMDs and the stable and controllable brightness conditions within the HMD form a suitable environment.

3 CONCEPT

The core concept of the presented system is to exploit differences in gaze behavior, in particular the gaze velocity. The velocity of gaze is directly dependent upon the speed of a stimulus that is being followed with the eyes (cf., Figure 1). A fixation occurs when a user stares at an unmoving stimulus [44]. When the stimulus is slowly moving, but it can still be easily followed with the eyes, it induces a smooth pursuit movement [3]. Once the target stimulus exceeds a certain speed, the user can no longer follow it using smooth pursuit eye movements. The eyes then switch to a saccadic movement [3]. Once a certain threshold of velocity is surpassed, or when the stimulus leaves the field of view, a rotation of the head becomes necessary to keep track of the stimulus.

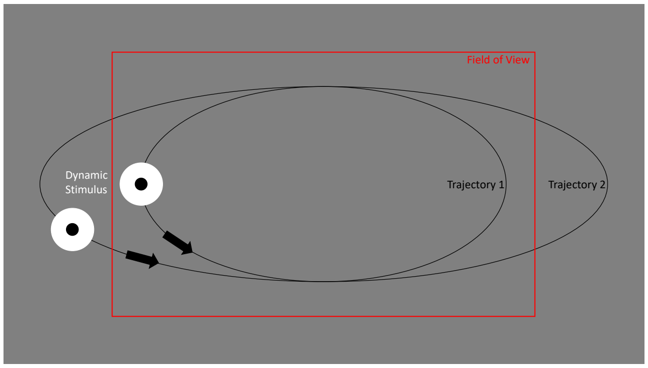


Figure 1: The two stimuli and their trajectories. The outer trajectory exceeds the field of view of the HMD, requiring a head rotation to follow it, whilst the inner trajectory stays within these bounds at all times.

Besides the head rotation, we exploit the individual gaze behavior as a biometric trait that is dependent on a visual stimulus. The behavior of eye gaze is a highly individual trait and differences in gaze exist across persons, up to a level of being able to predict personality traits from gaze behavior [15]. Previous work has shown that the response of the human body to a given stimulus can result in highly different outcomes, even if the applied stimulus is the same across different persons [7, 38]. This allows to distinguish different users by employing the concept of "functional biometrics" [26]. In case that no implicit identification is possible for any reason, an escalation to an explicit method always is a possibility.

To demonstrate the usefulness of our approach, we outline three application scenarios that show how the presented concept can be applied to real world scenarios. The stimulus within each scenario elicits the required gaze behavior for implicit identification.

- **360° Movies.** Movies and videos often contain objects that accelerate, which can be used as stimuli for identification. Here, an example might be nearly any kind of ball sport in which the ball moves on the screen. Similarly, any object that moves within a scene might easily be utilized (e.g., planes, birds).
- **Games.** Games in VR contain objects that move in the virtual world. One example is an object (e. g., a plane or a bird) in a virtual environment that moves across the sky at either varying speed or at an oblique viewing angle. When the user looks at these objects, the perceived change in motion can induce gaze behavior which could be employed to continuously identify the user.
- **Applications.** User interfaces for virtual reality can be designed in such a way that they include moving stimuli. For instance, sending an email could involve a visual display of the message being folded and placed in an envelope that then flies away to communicate the sending progress. Loading bars that are followed by the user's eyes could also be employed for the same purpose, to implicitly elicit the wanted gaze behavior.

4 EVALUATION

We conducted a lab study to verify our proposed approach in Virtual Reality using gaze and head orientation behavior. To validate the approach, we implemented a K-Nearest-Neighbors-based algorithm based on summary statistics to classify the participant's gaze data and a deep learning model operating on raw data generated by the eye tracker.

4.1 Stimulus

In this work, we evaluate two stimuli that are represented by a white sphere that has a second black sphere at its core (cf., Figure 1), a composition which follows the design of Lohr et al. [27]. The VR users are located in a virtual space that is surrounded with light-grey walls to avoid any distraction and the stimulus is located exactly one virtual meter in front of them. To avoid any change in distance, the positional tracking of the HMD is disabled, and thus, only rotational tracking is available (i. e., three degrees of freedom are available).

4.1.1 Stimulus 1. The first stimulus consists of the sphere moving counterclockwise on an elliptical path. The path of this sphere always remains within the field of view of the HMD and is defined by a major axis of 3 meters and a minor axis of 1 meter. Over a time span of 40 seconds, the sphere constantly accelerates. The first completion of the path takes 5 seconds (72 deg/s), while the last completion of the path takes only 1.4 seconds (257 deg/s).

4.1.2 Stimulus 2. The second stimulus is identical to the first stimulus except for the fact that the elliptical path is modified. Here, it has a major axis of 4 meters, and thus, it does not fit within the HMD's field of view. To follow this stimulus at all times, a rotation must be imposed on the HMD, as the stimulus will otherwise leave the field of view. The maximum speed remains unchanged, but the acceleration is adapted to the increased length of the path in comparison to stimulus 1.

4.2 Participants

We invited 12 participants through our university's mailing list (2 female and 10 male, aged from 21 to 28 with a median age of 25). Out of the 12 participants, 5 stated that they had normal vision and 7 stated that they had corrected to normal vision. One participant used contact lenses during the study, while six others used prescription glasses. The other participants stated that their vision would not hinder them in their participation or interfere with the study.

Out of the 12 participants, only 11 participated in the full procedure, as one participant had to abort the study due to acute cyber sickness. We exclude this participant's data from further processing since the recording is incomplete.

4.3 Ethics

One of the purposes of the proposed system is identification. If this system is employed in practice, it would be highly important to inform the users of the system of the implicit collection of gaze and head orientation data and to obtain the users' consent. Furthermore, it is recommended that all VR applications follow ethical development standards [1].

To protect the privacy of our study participants, we have assigned their data pseudonyms at the time of recording. After finishing the analysis of the data we have deleted the corresponding mapping of the participants' true identity to the given pseudonyms so that their true identity cannot be derived from elements in the data set.

4.4 Study Design and Procedure

After the participants arrived in our lab, we obtained their informed written consent, explained the procedure and fully answered all questions. Participants could abort their participation at any time without any detriments. Our study design follows a within-subject, repeated-measures paradigm. Each recording session began with a manual adjustment of the eye tracker and its associated parameters. We set 400×400 px as the eye tracker's resolution and 120 hz as the eye tracker's sampling rate. We then manually adjusted the region of interest to where the pupil was located and performed the vendor's calibration routine. After the eye tracker was set and ready for recording, we conducted three repetitions of stimulus 1 followed by three repetitions of stimulus 2.

The study procedure took approximately 20 minutes in total. The participants were asked to sit on a chair in a 3×6 m room. The light in the room was dimmed and the blinds were closed, although the HMD itself already shields the participant from light sources within the environment. With the help of the experimenter, the participants placed the HMD on their heads and adjusted the straps and interpupillary distances.

4.5 Apparatus

We used the HTC Vive Pro as our HMD. It is equipped with one display per eye, where each display has a resolution of 1440×1600 px. The field of view is 110° and offers a display refresh rate of 90 hz. We equipped the HTC Vive with the binocular eye tracking add-on from Pupil Labs[1]. This add-on consists of two infrared cameras mounted near the HMD's lenses. It operates at an adjustable sampling rate of 120 to 200 hz and a resolution of 192×192 px to 400×400 px, where the highest resolution can only be chosen at the lowest sampling rate, which we did for our study. We connected the HMD to a workstation that was equipped with Intel Core i9 9900k CPU, 32 GB RAM, and an Nvidia 2080 GPU. We used Unity3D to implement and execute the stimuli. To calibrate the eye tracker, we used the default 16-point calibration routine of the "hmd-eyes" Unity plugin provided by Pupil Labs[2].

4.5.1 Gaze Classification. The Pupil Labs software stack provides us with capabilities to binocularly record the gaze of the user that is elicited as a response to the stimuli. Each recording consists of the gaze point coordinates in a normalized coordinate frame together with a confidence value. The confidence value is reported by the Pupil Labs software and provides information about the degree of confidence that the eye tracker has correctly estimated the gaze point (i. e., a blink results in low confidence).

To generate features for classification from the given stream of gaze data, we utilize REMoDNaV [9], which is an open-source software library written in Python for classifying a stream of gaze data and extracting high level features from it. With REMoDNaV, we classify three types of gaze: fixations, smooth pursuits, and saccades. We then collect six features for each: the onset time, the duration, the amplitude, the peak velocity, the median velocity and the average velocity of the user's gaze. Before supplying REMoDNaV with the captured data, we filter the raw gaze data by a confidence value of 0.4. We remove all data below this threshold by setting it to"NaN" ("Not a Number", e. g., data that was captured during blinks).

4.5.2 Head Orientation. Besides the user's gaze, we also log the rotational coordinates (x, y and z in Euler angles) from the HMD. The data is captured as its change in value over time. Given the rotational coordinates, we calculate the rotational peak velocity, median velocity and average velocity to match the gaze features captured by the eye tracker.

4.6 Classification

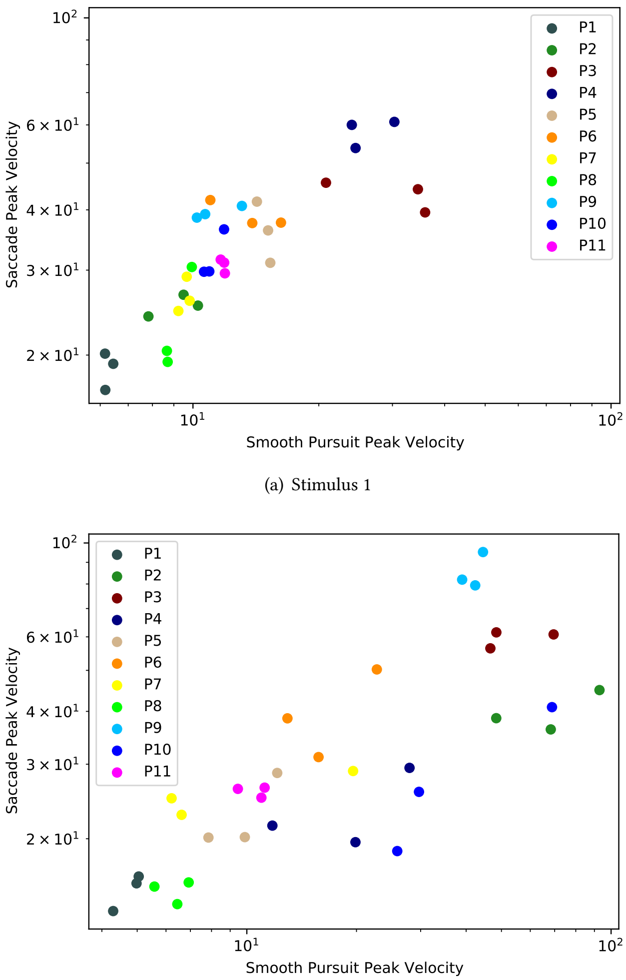
We use a K-Nearest-Neighbors method with $k = 1$ ("1NN") as an explainable machine learning classifier to classify the identities of our participants in a post-hoc analysis. For the 1NN, we employ summary statistics to aggregate the data into different features per gaze type and for the head orientation. Furthermore, we also classify the raw gaze and rotation data with deep neuronal networks ("DNN").

4.6.1 Data and Features. First, given the output of our gaze classifier, REMoDNaV, and our head orientation feature extractor, we create an aggregation of the data that consists of 21 features for the remaining 11 participants. A specific gaze behavior can, across all participants' data recordings, appear in varying amounts per recording (e. g., one participant's recording can have more detected saccades than a recording from another participant). Yet, to create a classifier, we need to unify the shape of the data so that all recordings are data-wise comparable. We use the arithmetic mean function as a summary statistics for the 1NN-approach to aggregate the given features per data recording for each gaze feature.

To evaluate our approach, we strictly separate the first and second stimuli and evaluate each on their own. Furthermore, from the extensive set of 21 features, we only pick four features for further analysis, following an univariate feature selection. These correspond to the peak velocities of the four gaze and head movements (i. e., fixations, smooth pursuits, and saccades, as well as the head orientation in the case of the second stimulus).

We then split the data into a training set and a validation set. The training data set consists of two repetitions of each participant per stimulus and the validation data set consists of one repetition. This results in a 2/3 training to 1/3 validation data split, where the former consists of the first two repetitions and the latter of the last repetition. We chose this split because we expect a learning effect and fatigue to degrade the quality of the last recording. For the 1NN-approach we decide to add one person to the "unknown identity" class, i. e., designate the person as a negative that should be rejected by the system. We cross-validate this person by picking each person once as a true negative to simulate an open-set classification. For the

[1]HTC Vive Add-On. https://docs.pupil-labs.com/vr-ar/htc-vive, last followed on October 12, 2021.
[2]HMD Eyes. https://github.com/pupil-labs/hmd-eyes, last followed on October 12, 2021.



(a) Stimulus 1

(b) Stimulus 2

Figure 2: Visualization of the KNN that classifies the user's identities for both stimuli. For better visibility, we only depict the smooth pursuit and saccade peak velocities (px/deg) on a log scale.

deep learning evaluation, we on the other hand opt for a traditional closed-set classification and do not assign an unknown identity, as only little literature exists with relation to deep learning-based open-set classification and biometrics [33].

4.6.2 K-Nearest-Neighbors. At the core of our system, we employ a modified K-Nearest-Neighbors algorithm with $k = 1$ (1NN) for classifying the elicited data from the participants and to distinguish their identities. Our modification of the KNN consists of the inclusion of a threshold to enable the KNN to determine negative matches (i. e., true negatives and false negatives). This threshold is a positive number that acts as the upper bound for the distance that may exist between two neighbors in order for them to be recognized as neighbors and form a positive match (i. e., true positives and false positives). If the distance between two close neighbors falls below this threshold, they form a positive match (i. e., an acceptance by the system). If the distance between two close neighbors is equal to or greater than the threshold, they form a negative match (i. e., a rejection by the system). Therefore, in our system, an acceptance is met when the equation $euclidean\_distance(P1, P2) < threshold$ is true, where $P1$ and $P2$ are two neighboring points. A rejection applies in all other cases (i. e., when the equation results in false).

To determine the optimal threshold, we tested each distance within the KNN as a threshold and calculated the accuracy for each once we validated our data set. In the end, we chose the threshold that yielded the highest accuracy.

4.6.3 Deep Learning. For the classification of raw gaze and head orientation data through deep learning we first combine the raw gaze data with the head orientation data. As the eye tracker captures data at 120 hz but the Unity prototype runs at 90 hz and thus the head orientation is also captured at 90 hz, we first resample the head orientation data from 90 hz to 120 hz to match the gaze data. Next, as we do not employ summary statistics but raw data here, we interpolate the gaze data linearly when its confidence drops below 0.4 (e. g., at blinks or tracking-losses) to keep the time series and the relation to the head orientation data intact.

In the next step, we center the mean of the head orientation data to +0.5 to fit it into the interval of $[0, 1]$. We also clip the eye tracker's captured coordinates ("norm_pos_x" and "norm_pos_y") to the same interval. This way, we obtain time-series samples that consist of five dimensions over time ("norm_pos_x", "norm_pos_y", "rot_x", "rot_y" and "rot_z"). Then we unify its shape through the application of a zero-based prepadding.

The last step of the preprocessing of the data is implemented through the "window slicing technique", a popular method used in time-series classification [23]. The deep learning models are trained and validated on slices of the actual raw gaze and head orientation data that are obtained from a window that has a certain size and that is moved across the stream of data in the dimension of its time axis at a given stride. Through a parameter search we determined a window size of 300 and a window stride of 150 as optimal values. Each sample (i. e., each recording from the study) of 40 s is therefore sliced into 60 new samples, where each sample has a new length of 2.5 s of data. An overlap between each predecessor and success of 50 % is given by the chosen stride. The original classified identity can be reconstructed by applying a majority vote to the classified slices of data.

The actual classification is then performed through ten deep learning models that are proposed by Fawaz et al. in their review paper on time series classification [10]. We employ their exact model definitions and Keras-based implementation[3], as well as their chosen hyperparameters such as learning rate, optimizer and other parameters to enable seamless comparability and validate the training in a 3-fold cross-validation, where each gaze recording is used for validation once.

The utilized model architectures are a Time Convolutional Neural Network ("Time-CNN") [48], Multi Layer Perceptron ("MLP") [46], Fully Convolutional Neural Network ("FCN") [46], Residual

[3]Deep Learning for Time Series Classification. https://github.com/hfawaz/dl-4-tsc, last followed on October 12, 2021.

Network "ResNet" [46], "Encoder" [39], Multi-scale Convolutional Neural Network ("MCNN") [8], Time Le-Net "t-LeNet" [24], Multi Channel Deep Convolutional Neural Network ("MCDCNN") [49], Time Warping Invariant Echo State Network ("TWIESN") [43] and InceptionTime ("Inception") [16] following the implementation of Fawaz et al. [10]. The full training and evaluation of all models of the deep learning evaluation took approx. 48 hours on three NVIDIA A40 GPUs in a concurrent execution.

4.7 Results

The results of the explainable 1NN-classifier and the deep learning approach are presented in the following. While the deep learning operating on raw data surpasses the summary-statistics-based 1NN in terms of accuracy, the latter is able to provide valuable insights due to its explainability (cf., Figure 2). Furthermore, the elicited data set is available online[4].

4.7.1 Results for K-Nearest-Neighbors. The results that we obtained from our 1NN-classifier are listed in Table 1. Figure 2 depicts two of the features for all participants and recordings (saccade peak velocity and smooth pursuit peek velocity). For stimulus 1, that corresponds to trajectory 1 (cf., Figure 1), where the stimulus does not leave the field of view. We reach a mean accuracy of 45 %, following an 11-cross-validation of the unknown person. In this cross validation, each participant's data once was marked as a negative. At best, we can report an accuracy of 55 %, consisting of 6 true positives (TP) and 5 false positives (FP) at a threshold value of 7.77 for stimulus 1.

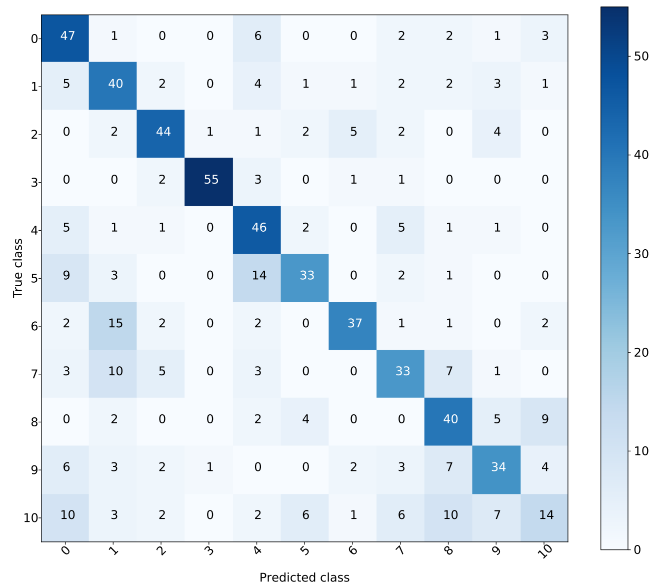
For stimulus 2, which follows trajectory 2 and exceeds the HMD's field of view, we reach an 11-cross-validated mean accuracy of 75 %. In addition to the features of stimulus 1, we add the head's peak rotational velocity to the feature set. Here, we can at best report an accuracy of 82 %, which is the result of 8 TP, 1 FP, 1 TN and 1 FN at a threshold value of 17.78.

4.7.2 Results for Deep Learning. Table 2 lists the results for the deep learning classifiers. With exception of the MCNN and t-LeNet, all models have converged and yielded high accuracy ratings of the classified data. The best model for the first stimulus is the "Encoder", resulting in a cross-validated mean of 0.9697 of the classified identites after performing a majority vote (cf., Figure 3). For the second stimulus, the "Encoder" model is able to generate a cross-validated mean accuracy of 1.0, as it met an accuracy of 1.0 in each of the cross-validation cycles.

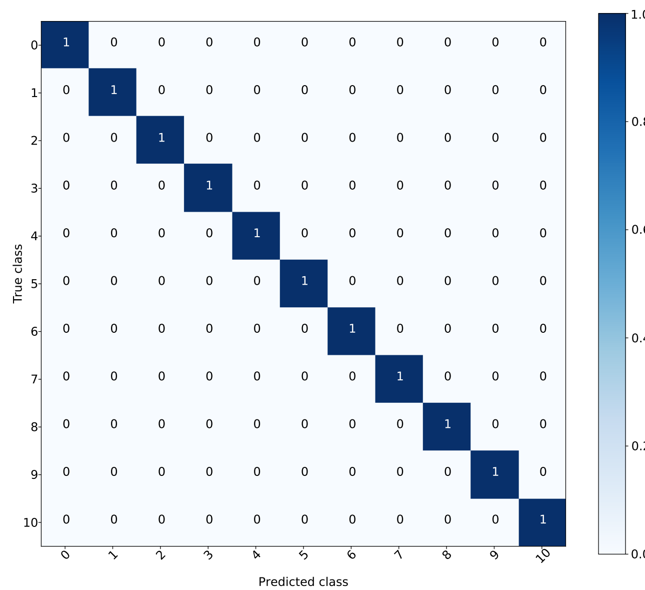
5 DISCUSSION

From the given results, it is apparent that gaze contains a high amount of user-specific information. By achieving a cross-validated accuracy of up to 75% following the 1NN and 100 % through deep learning in the second stimulus, we show that our approach can be used to identify users.

First, it is apparent that the inclusion of the head orientation data in the second stimulus resulted in an increase in the identification rate for both classifiers (45 % to 75 % for the 1NN and 96 % to 100 % for deep learning). This is expected, as previous work exists that is only based on head orientation in VR [31, 41]. Nevertheless, as



(a)



(b)

Figure 3: Confusion Matrices for the first cross-validation of Stimulus 1 utilizing the "Encoder" deep learning model architecture. (a) The resulting classification of the slices of gaze and head orientation data obtained from the window slicing approach. (b) The application of a majority vote to the slices of gaze and head orientation data leads to a classification accuracy of 1.0.

both classifiers show the same trend of an increasing accuracy, we can conclude that the underlying data is more meaningful and individual with the head orientation data being present. This can

[4] The data set of this study is publicly available at https://research.hcigroup.de.

Table 1: Results of the 11-cross-validation of the negative for the 1NN. The feature set consists of the mean peak velocity (mpv) for the fixations (F), pursuits (P) and saccades (S). In the case of stimulus 2, we also add the head (H) orientation peak velocity. We report the mean accuracy (acc.) gained through the 11-fold cross validation, as well as the best and worst accuracy ratings encountered during the validation.

Stimulus	Feature Set	Best acc.	Worst acc.	Mean acc.
Stimulus 1	$F_{mpv}, P_{mpv}, S_{mpv}$	0.55	0.36	0.45
Stimulus 2	$F_{mpv}, P_{mpv}, S_{mpv}, H_{mpv}$	0.82	0.73	0.75

Table 2: Deep learning accuracies per model architecture per validation cycle in the 3-cross-validation and the obtained mean accuracy (acc.). The accuracies for MCNN and t-LeNet are not listed as their accuracy did not exceed the random chance of 1/11 for each validation and the best cross-validated accuracy is marked in italic.

Stim.	k-Fold	Time-CNN	Encoder	FCN	Inception	MCDCNN	MLP	ResNet	TWIESN
St. 1	1	0.9091	1.0000	0.9091	0.9091	0.9091	0.7273	0.9091	0.9091
	2	1.0000	0.9091	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	3	0.9091	1.0000	0.7273	0.9091	0.9091	1.0000	0.9091	0.9091
	mean acc.	0.9394	*0.9697*	0.8788	0.9394	0.9394	0.9091	0.9394	0.9394
St. 2	1	0.8182	1.0000	0.9091	1.0000	1.0000	0.7273	1.0000	1.0000
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	3	0.8182	1.0000	1.0000	1.0000	0.9091	0.7273	1.0000	1.0000
	mean acc.	0.8788	*1.0000*	0.9697	*1.0000*	0.9697	0.8182	*1.0000*	*1.0000*

be visually confirmed from Figure 2, where it is apparent that for the second stimulus, the per-class clusters are less overlapping than for the first stimulus. Further, an increase in velocity can be seen in Figure 2(b) which is expected due to the enlarged trajectory of stimulus 2.

5.1 Validation Accuracy in Deep Learning

The high classification accuracy of the deep learning classifiers comes at the cost of explainability of the approach as well as the immense required computational power that is necessary to automatically craft the features that distinguish the users. In contrast, the 1NN operates on summary statistics but does so in an explainable manner. The summary statistics that were applied to transform the data for usage with the 1NN comes at the cost of losing information, as they reduce the amount of information from each raw gaze recording file (each having a size of approx. 4 megabytes) into a set of three or four floating point numbers, in case of stimulus 1 and 2 respectively. This reduction comes at the cost of a loss of information that is not present for the deep learning algorithms. Nevertheless, both approaches work on the same underlying data and a per-person similarity can already be visually derived from Figure 2, as it is apparent, that local, consistent clusters per person exist. In particular, $k = 1$ for the K-Nearest-Neighbors classifier indicates that, as the classification of a point takes place by the nearest neighbor only, the data is already strongly correlated and clusters exist per participant, as otherwise the accuracy could not live up to a rate of 75 %.

5.2 Limitations

We acknowledge the following limitations to our work. First, the number of participants that took part in the study which resulted in the presented results could have been higher, to determine the true upper boundary of the deep learning algorithms for stimulus 2. While related work used similar sample sizes for their initial investigations (e.g., 10 participants [38] or 13 participants [7]), a larger sample size would have been useful to better understand when the system starts confusing users with each other. As we determined a mean accuracy of 1.0 for certain models, we are unable to state when this accuracy starts to degrade, i. e., which number of participants is necessary to mislead the deep learning models for stimulus 2. Second, our data was only elicited during one day per participant, similar to other explorations [7, 38]. Therefore, we cannot conclude how "stable" this biometric is across days, as for example fatigue has a strong influence on gaze [45]. On the other hand we can clearly state that the data that was captured during the study bears highly individual information, allowing to distinguish the participants for the purpose of identification.

5.3 Explicit vs. Implicit Identification

Although we deem the investigated interaction to be suitable for an inclusion within an implicit identification scheme, we have evaluated it in an explicit way during a controlled lab study as this work acts as a foundation to understand the feasibility of employing such gaze behavior in Virtual Reality. Nevertheless, it is important to investigate in future studies, to what extend the necessary time frame for identification can be shortened at any given error rate and how it would affect the individual components within the gaze data. Also, by employing the approach in less controlled virtual environments (such as in the described application scenarios), one would need to keep track of the error that might be induced by each participant, by, for example, simply not looking and focusing at the desired stimulus. An employment as an implicit identification nevertheless has the unique opportunities of designing the mechanism in a way that is transparent to the user and could be executed in a continuous manner, removing all workload from the user during the process.

6 CONCLUSION

In this work, we have presented a novel approach for identifying the users of VR systems based on their gaze behavior elicited by two moving stimuli, where one stimulus moves on an elliptical path within the field of view of the user and the other stimulus moves beyond that area. After capturing the user's gaze on an HTC Vive Pro with the Pupil Labs eye tracking add-on, we extracted the peak velocity for the three most important types of gaze behavior by applying the REMoDNaV gaze classifier [9] to the data. We furthermore include the head orientation as another feature for the trajectory in which the stimulus leaves the field of view and classify the data by an explainable Nearest Neighbor classifier. Moreover, we feed the raw data to in total ten different deep learning model architectures. Here, the models can fully utilize their capability of engineering the relevant features from the data on their own.

We have verified our approach in a user study (N=12) with 11 remaining usable data recordings. Following an 11-cross-validation of the negative, we reached a mean accuracy of 75 % through a modified K-Nearest-Neighbor classification algorithm with $k = 1$, showing that implicit user identification is feasible and that the data has a strong, individual component per participant. This accuracy then is surpassed by the deep learning evaluation, which yields a cross-validated accuracy of up to 100 % for multiple model architectures taken from previous work.

We have explored the first steps towards a novel, implicit type of identification that exploits differences in the user's gaze and head orientation which are captured in Virtual Reality. We show that these individual components within the data are on one hand explainable and on the other hand highly individual, given the acquired accuracy rates of the classifiers. We show that the users' gaze behavior and in particular the individual velocities the specific types of eye movements have, is well suited for the purpose of identification. To facilitate future research, we moreover release the data set and we believe that the future of gaze-based identification in Virtual Reality lies within the domain of implicit interactions due to their unique properties of being transparent to the user.

ACKNOWLEDGMENTS

The presented work was funded by the German Research Foundation (DFG) under project no. 426052422.

REFERENCES

[1] Devon Adams, Alseny Bah, Catherine Barwulor, Nureli Musaby, Kadeem Pitkin, and Elissa M. Redmiles. 2018. Ethics Emerging: the Story of Privacy and Security Perceptions in Virtual Reality. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. USENIX Association, Baltimore, MD, 427–442. https://www.usenix.org/conference/soups2018/presentation/adams

[2] F. A. Alsulaiman and A. El Saddik. 2008. Three-Dimensional Password for More Secure Authentication. *IEEE Transactions on Instrumentation and Measurement* 57, 9 (2008), 1929–1938.

[3] Graham R Barnes. 2011. Ocular pursuit movements. In *The Oxford Handbook of Eye Movements*. Oxford University Press, Oxford, England, UK, 115–132.

[4] Fadi Boutros, Naser Damer, Kiran Raja, Raghavendra Ramachandra, Florian Kirchbuchner, and Arjan Kuijper. 2020. Iris and periocular biometrics for head mounted displays: Segmentation, recognition, and synthetic data generation. *Image and Vision Computing* 104 (2020), 104007. https://doi.org/10.1016/j.imavis.2020.104007

[5] Andreas Bulling, Florian Alt, and Albrecht Schmidt. 2012. Increasing the Security of Gaze-Based Cued-Recall Graphical Passwords Using Saliency Masks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 3011–3020. https://doi.org/10.1145/2207676.2208712

[6] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Immersive Insights: A Hybrid Analytics System ForCollaborative Exploratory Data Analysis. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 9, 12 pages. https://doi.org/10.1145/3359996.3364242

[7] Yuxin Chen, Zhuolin Yang, Ruben Abbou, Pedro Lopes, Ben Y. Zhao, and Haitao Zheng. 2021. User Authentication via Electrical Muscle Stimulation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 6, 15 pages. https://doi.org/10.1145/3411764.3445441

[8] Zhicheng Cui, Wenlin Chen, and Yixin Chen. 2016. Multi-Scale Convolutional Neural Networks for Time Series Classification. arXiv:1603.06995 [cs.CV]

[9] Asim H. Dar, Adina S. Wagner, and Michael Hanke. 2021. REMoDNaV: robust eye-movement classification for dynamic stimulation. *Behavior Research Methods* 53, 1 (01 Feb 2021), 399–414. https://doi.org/10.3758/s13428-020-01428-x

[10] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. 2019. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* 33, 4 (March 2019), 917–963. https://doi.org/10.1007/s10618-019-00619-1

[11] Ceenu George, Daniel Buschek, Andrea Ngao, and Mohamed Khamis. 2020. GazeRoomLock: Using Gaze and Head-Pose to Improve the Usability and Observation Resistance of 3D Passwords in Virtual Reality. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso De Paolis and Patrick Bourdot (Eds.). Springer International Publishing, Cham, 61–81.

[12] Ceenu George, Mohamed Khamis, Emanuel von Zezschwitz, Henri Schmidt, Marinus Burger, Florian Alt, and Heinrich Hußmann. 2017. Seamless and Secure VR: Adapting and Evaluating Established Authentication Systems for Virtual Reality. In *Proceedings 2017 Workshop on Usable Security*. Internet Society, San Diego, CA, USA. https://doi.org/10.14722/usec.2017.23028

[13] Corey Holland and Oleg V. Komogortsev. 2011. Biometric identification via eye movement scanpaths in reading. In *2011 International Joint Conference on Biometrics (IJCB)*. IEEE, Washington, DC, USA, 1–8. https://doi.org/10.1109/IJCB.2011.6117536

[14] Corey D. Holland and Oleg V. Komogortsev. 2013. Complex eye movement pattern biometrics: Analyzing fixations and saccades. In *2013 International Conference on Biometrics (ICB)*. IEEE, Madrid, Spain, 1–8. https://doi.org/10.1109/ICB.2013.6612953

[15] Sabrina Hoppe, Tobias Loetscher, Stephanie A. Morey, and Andreas Bulling. 2018. Eye Movements During Everyday Behavior Predict Personality Traits. *Frontiers in Human Neuroscience* 12 (2018), 105. https://doi.org/10.3389/fnhum.2018.00105

[16] Hassan Ismail Fawaz, Benjamin Lucas, Germain Forestier, Charlotte Pelletier, Daniel F. Schmidt, Jonathan Weber, Geoffrey I. Webb, Lhassane Idoumghar, Pierre-Alain Muller, and François Petitjean. 2020. InceptionTime: Finding AlexNet for time series classification. *Data Mining and Knowledge Discovery* 34, 6 (01 Nov 2020), 1936–1962. https://doi.org/10.1007/s10618-020-00710-y

[17] Markus Jakobsson, Elaine Shi, Philippe Golle, and Richard Chow. 2009. Implicit Authentication for Mobile Devices. In *Proceedings of the 4th USENIX Conference on Hot Topics in Security* (Montreal, Canada) *(HotSec'09)*. USENIX Association, USA, 9.

[18] Pawel Kasprowski. 2020. Identifying Users Based on Their Eye Tracker Calibration Data. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, Article 38, 2 pages. https://doi.org/10.1145/3379157.3391419

[19] Christina Katsini, Yasmeen Abdrabou, George E. Raptis, Mohamed Khamis, and Florian Alt. 2020. *The Role of Eye Gaze in Security and Privacy Applications: Survey and Future HCI Research Directions*. Association for Computing Machinery, New York, NY, USA, 1–21. https://doi.org/10.1145/3313831.3376840

[20] Mohamed Khamis, Carl Oechsner, Florian Alt, and Andreas Bulling. 2018. VRpursuits: Interaction in Virtual Reality Using Smooth Pursuit Eye Movements. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces* (Castiglione della Pescaia, Grosseto, Italy) *(AVI '18)*. Association for Computing Machinery, New York, NY, USA, Article 18, 8 pages. https://doi.org/10.1145/3206505.3206522

[21] Manu Kumar, Tal Garfinkel, Dan Boneh, and Terry Winograd. 2007. Reducing Shoulder-Surfing by Using Gaze-Based Password Entry. In *Proceedings of the 3rd Symposium on Usable Privacy and Security* (Pittsburgh, Pennsylvania, USA) *(SOUPS '07)*. Association for Computing Machinery, New York, NY, USA, 13–19. https://doi.org/10.1145/1280680.1280683

[22] Alexander Kupin, Benjamin Moeller, Yijun Jiang, Natasha Kholgade Banerjee, and Sean Banerjee. 2019. Task-Driven Biometric Authentication of Users in Virtual Reality (VR) Environments. In *MultiMedia Modeling*, Ioannis Kompatsiaris, Benoit Huet, Vasileios Mezaris, Cathal Gurrin, Wen-Huang Cheng, and Stefanos Vrochidis (Eds.). Springer International Publishing, Cham, 55–67.

[23] Arthur Le Guennec, Simon Malinowski, and Romain Tavenard. 2016. Data Augmentation for Time Series Classification using Convolutional Neural Networks. In *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*.

Riva Del Garda, Italy. https://halshs.archives-ouvertes.fr/halshs-01357973

[24] Arthur Le Guennec, Simon Malinowski, and Romain Tavenard. 2016. Data Augmentation for Time Series Classification using Convolutional Neural Networks. In *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*. Riva Del Garda, Italy. https://halshs.archives-ouvertes.fr/halshs-01357973

[25] Jonathan Liebers, Mark Abdelaziz, Lukas Mecke, Alia Saad, Jonas Auda, Uwe Gruenefeld, Florian Alt, and Stefan Schneegass. 2021. Understanding User Identification in Virtual Reality Through Behavioral Biometrics and the Effect of Body Normalization. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 517, 11 pages. https://doi.org/10.1145/3411764.3445528

[26] Jonathan Liebers and Stefan Schneegass. 2020. Introducing Functional Biometrics: Using Body-Reflections as a Novel Class of Biometric Authentication Systems. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3334480.3383059

[27] Dillon Lohr, Samuel-Hunter Berndt, and Oleg Komogortsev. 2018. An Implementation of Eye Movement-Driven Biometrics in Virtual Reality. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 98, 3 pages. https://doi.org/10.1145/3204493.3208333

[28] Shiqing Luo, Anh Nguyen, Chen Song, Feng Lin, Wenyao Xu, and Zhisheng Yan. 2020. OcuLock: Exploring Human Visual System for Authentication in Virtual Reality Head-mounted Display. In *27th Annual Network and Distributed System Security Symposium NDSS*. The Internet Society, San Diego, California, USA, 15. https://dx.doi.org/10.14722/ndss.2020.24079

[29] Florian Mathis, John Williamson, Kami Vaniea, and Mohamed Khamis. 2020. RubikAuth: Fast and Secure Authentication in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382827

[30] Florian Mathis, John H. Williamson, Kami Vaniea, and Mohamed Khamis. 2021. Fast and Secure Authentication in Virtual Reality Using Coordinated 3D Manipulation and Pointing. *ACM Trans. Comput.-Hum. Interact.* 28, 1, Article 6 (Jan. 2021), 44 pages. https://doi.org/10.1145/3428121

[31] Tahrima Mustafa, Richard Matovu, Abdul Serwadda, and Nicholas Muirhead. 2018. Unsure How to Authenticate on Your VR Headset? Come on, Use Your Head!. In *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics* (Tempe, AZ, USA) *(IWSPA '18)*. Association for Computing Machinery, New York, NY, USA, 23–30. https://doi.org/10.1145/3180445.3180450

[32] Yudai Niitsu and Minoru Nakayama. 2021. Effects of Measurement Time and Presentation Size Conditions on Biometric Identification Using Eye Movements. In *ACM Symposium on Eye Tracking Research and Applications*. Association for Computing Machinery, New York, NY, USA, Article 23, 5 pages.

[33] Guansong Pang, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. 2021. Deep Learning for Anomaly Detection: A Review. *ACM Comput. Surv.* 54, 2, Article 38 (March 2021), 38 pages. https://doi.org/10.1145/3439950

[34] Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn. 2016. Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics (TOG)* 35, 6 (2016), 179.

[35] Ken Pfeuffer, Matthias J. Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 2019. Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300340

[36] Johanna Pirker, Andreas Dengel, Michael Holly, and Saeed Safikhani. 2020. Virtual Reality in Computer Science Education: A Systematic Review. In *26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 8, 8 pages. https://doi.org/10.1145/3385956.3418947

[37] Ioannis Rigas, Oleg Komogortsev, and Reza Shadmehr. 2016. Biometric Recognition via Eye Movements: Saccadic Vigor and Acceleration Cues. *ACM Trans. Appl. Percept.* 13, 2, Article 6 (Jan. 2016), 21 pages. https://doi.org/10.1145/2842614

[38] Stefan Schneegass, Youssef Oualil, and Andreas Bulling. 2016. *SkullConduct: Biometric User Identification on Eyewear Computers Using Bone Conduction Through the Skull*. Association for Computing Machinery, New York, NY, USA, 1379–1384. https://doi.org/10.1145/2858036.2858152

[39] Joan Serrà, Santiago Pascual, and Alexandros Karatzoglou. 2018. Towards a universal neural network encoder for time series. arXiv:1805.03908 [cs.LG] https://arxiv.org/pdf/1805.03908.pdf

[40] Christian Sinnott, James Liu, Courtney Matera, Savannah Halow, Ann Jones, Matthew Moroz, Jeffrey Mulligan, Michael Crognale, Eelke Folmer, and Paul MacNeilage. 2019. Underwater Virtual Reality System for Neutral Buoyancy Training: Development and Evaluation. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 29, 9 pages. https://doi.org/10.1145/3359996.3364272

[41] M. Sivasamy, V. N. Sastry, and N. P. Gopalan. 2020. VRCAuth: Continuous Authentication of Users in Virtual Reality Environment Using Head-Movement. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, Coimbatore, India, 518–523. https://doi.org/10.1109/ICCES48766.2020.9137914

[42] Ivo Sluganovic, Marc Roeschlin, Kasper B. Rasmussen, and Ivan Martinovic. 2016. Using Reflexive Eye Movements for Fast Challenge-Response Authentication. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (Vienna, Austria) *(CCS '16)*. Association for Computing Machinery, New York, NY, USA, 1056–1067. https://doi.org/10.1145/2976749.2978311

[43] Pattreeya Tanisaro and Gunther Heidemann. 2016. Time Series Classification Using Time Warping Invariant Echo State Networks. In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, Anaheim, CA, USA, 831–836. https://doi.org/10.1109/ICMLA.2016.0149

[44] Nicholas J Wade and Benjamin W Tatler. 2011. Origins and applications of eye movement research. In *The Oxford Handbook of Eye Movements*. Oxford University Press, Oxford, England, United Kingdom, 17–43.

[45] Yan Wang, Rui Huang, and Lei Guo. 2019. Eye gaze pattern analysis for fatigue detection based on GP-BCNN with ESM. *Pattern Recognition Letters* 123 (2019), 61–74. https://doi.org/10.1016/j.patrec.2019.03.013

[46] Zhiguang Wang, Weizhong Yan, and Tim Oates. 2017. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, Anchorage, AK, USA, 1578–1585. https://doi.org/10.1109/IJCNN.2017.7966039

[47] Z. Yu, H. Liang, C. Fleming, and K. L. Man. 2016. An exploration of usable authentication mechanisms for virtual reality systems. In *2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*. IEEE, Jeju, South Korea, 458–460. https://doi.org/10.1109/APCCAS.2016.7804002

[48] Bendong Zhao, Huanzhang Lu, Shangfeng Chen, Junliang Liu, and Dongya Wu. 2017. Convolutional neural networks for time series classification. *Journal of Systems Engineering and Electronics* 28, 1 (2017), 162–169. https://doi.org/10.21629/JSEE.2017.01.18

[49] Yi Zheng, Qi Liu, Enhong Chen, Yong Ge, and J. Leon Zhao. 2014. Time Series Classification Using Multi-Channels Deep Convolutional Neural Networks. In *Web-Age Information Management*, Feifei Li, Guoliang Li, Seung-won Hwang, Bin Yao, and Zhenjie Zhang (Eds.). Springer International Publishing, Cham, 298–310.