IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 July 2023, accepted 9 August 2023, date of publication 28 August 2023, date of current version 5 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3309000

RESEARCH ARTICLE

User Authentication by Eye Movement Features Employing SVM and XGBoost Classifiers

PIOTR ODYA, (Member, IEEE), FRANCISZEK GÓRSKI, (Member, IEEE), AND ANDRZEJ CZYŻEWSKI, (Member, IEEE)

Department of Multimedia Systems, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, 80-233 Gdańsk, Poland

Corresponding author: Piotr Odya (pioodya@pg.edu.pl)

This work was supported by the Polish National Centre for Research and Development (NCBR) funded the Research Project entitled: ''BIOPUAP—a biometric cloud authentication system'' from the European Regional Development Fund under Grant POIR.01.01.01-0092/19.

ABSTRACT Devices capable of tracking the user's gaze have become significantly more affordable over the past few years, thus broadening their application, including in-home and office computers and various customer service equipment. Although such devices have comparatively low operating frequencies and limited resolution, they are sufficient to supplement or replace classic input interfaces, such as the keyboard and mouse. The biometric application we researched verifies a user's identity based on parameters acquired with a low-cost eye tracker. The use of the eye-tracking device in bank booths has many advantages, including the fact that eye trackers are contactless devices, which, especially in the light of the Covid pandemic, has increasing importance, in addition to providing a solution for confirming the liveness of the user. This paper describes an experiment in which 20 features extracted from eye movement data related mainly to saccades and fixations are used as a complementary biometric modality to authenticate clients at banking kiosks. Data were collected from 39 subjects while operating a banking system using engineered biometric kiosk prototypes. Authentication performance employing eye-movement tracking and parameterizing was compared for two classifiers: Support Vector Machine (SVM) and eXtreme Gradient Boosting (XGBoost). The results showed that the XGBoost-based classifier outperformed the SVM-based one regarding equal error rates (6.76% to 8% vs. 16.21 to 18.78%). Similar differences were obtained for true acceptance rates at different false acceptance rates (0.1 and 0.01), where the SVM-based classifier achieved a maximum of 81.08% and the XGBoost-based achieved 98.65%. Finally, prospects for the broader application of eye movement tracking as a biometric modality are discussed.

INDEX TERMS Biometrics, gaze tracking, machine learning, support vector machines, XGBoost.

I. INTRODUCTION

Efforts are constantly being made to improve biometric authorization technologies and to attempt to implement them into everyday life. In contrast to needing to remember numerous PINs and passwords, biometrics appears to be the best method of verifying identity from the ordinary user's perspective. Besides the popular fingerprint and face recognition, there are multiple proposals for using new biometric modalities. These include eye movement parameters acquired using eye trackers [1].

It should also be noted that despite earlier predictions that eye-tracking devices would become standard in laptops and smartphones, this has not happened so far. Consequently, with practical applications in mind, one has to reckon with planning rational scenarios, as it is doubtful that the ordinary user will seek to facilitate authentication by equipping themself with separate eye-tracking attachments. Following this thought, the experiments described here assume a particular field of application: banking kiosks. Therefore, the

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

authors of this paper gathered practical experience in this domain resulting from long-term cooperation between the Gdansk University of Technology and the largest Polish bank [2].

Certain assumptions about systems for biometric authentication are universal, regardless of the modality or biometric data acquisition method adopted, and this also applies to eye-tracking systems. Thus, achieving the lowest possible FRR (False Rejection Rate) and FAR (False Acceptance Rate) values is generally desirable. The system must also be resistant to phishing attempts – e.g., in the form of providing the system input with data that has been previously registered without the customer's knowledge. Furthermore, the authentication process should not require the customer to possess unique knowledge and skills. It is also expected that the user authentication process should be as fast as possible.

The objective of the experiments described in this article was to verify whether a low-cost eye tracker could fulfill the above assumptions. Usually, such devices are used as an additional interface for controlling a computer using the eye. Such a non-contact interface has become a useful solution, especially in the face of the Covid-19 pandemic, without requiring special preparation by the user. The continuous analysis of the user's gaze provided by eye trackers can also be used to improve security at the banking stand – such as by simply verifying the user's liveness. However, we wanted to evaluate the possibility of employing eye movement analysis of the user as an additional modality for user authentication. Using several biometric verification algorithms in a single system ensures higher accuracy and adapts the system operation to the existing conditions [3]. Using multiple biometric modalities makes it possible to drop selected modalities if the customer, for example, cannot use them.

The work described in the manuscript involved developing an optimal way to collect biometric features using an eye tracker, followed by their parameterization (i.e., saving the personal characteristics in the form of a mathematical description, explicitly not describing the identity and making it impossible to return to the state of personal features) and classification (i.e., comparing samples recorded in the database with a verification sample).

II. RELATED WORK

The simplest way to use an eye-tracking device to authenticate a user seems to be an interface for entering credentials (similar to a mouse or keyboard). In the simplest case using eye gaze, it is possible to enter a PIN using a keyboard displayed on the screen or select a desired image from a specific set of images displayed on the screen. However, this approach does not use any actual biometric information. It only makes it more challenging to view the login credentials and thus provides a slightly higher level of security than the classic approach. Even worse, installing a simple web camera registering the user's eye movements will allow, for example, the PIN entered in this way to be captured [4], [5]. Kasprowski and Harezlak presented an interesting approach in [6]. They started from the principle that the eye tracker requires calibration, which usually involves gazing at several evenly spaced points on the screen. Using this, they demonstrated impostor detecting by analyzing how much the mouse cursor's position differs from the gaze's location. The fusion of data acquired from eye and mouse movements has also been investigated by Fuhl et al. [7]. They noted that due to the strong correlation between mouse and eye movements, using the two methods together is not particularly efficient; though applied separately, it is reasonable, bringing comparable results.

An additional advantage of employing an eye tracker is quickly verifying whether a living person is present in front of the computer (so-called liveness checking). The device will not work if it does not detect the user's moving eyeballs. Even the use of a mask is problematic in such a situation, as it may slightly obscure the eyes, which will cause the eye tracker to work incorrectly [8].

The aspect of a person's viability is also crucial in the facial biometrics being developed by the authors of this paper for the banking sector. In this context, the issue to be verified is whether the eye-related biometric modality serves as an additional functionality or could provide results comparable to voice, face, and hand vein biometrics so that it can fall within the scope of multimodal biometric fusion [9].

The remainder of this section will provide a rationale for system assumptions for visual authentication.

A. GAZE TRACKER REQUIREMENTS

When formulating the objectives of an authentication system using eye tracker data, it is essential to also consider the cost and technical capabilities of the device. Until recently, eye trackers were expensive (costing several thousand euros). This significantly limited the possibility of its use in real life and restricted this modality to scientific research only. The operating frequency of this type of eye tracker reaches hundreds of hertz, providing accurate observation of saccadic movements. In a formerly published article, Holland and Komogortsev [10] recommended that the eye tracker operate at 250 Hz at least for biometric applications. They also noted that the minimum operating frequency for this type of application is 30 Hz, which is the frequency achieved by "in-home" class eye trackers (costing no more than a few hundred euros). However, such eye trackers' parameters (not only operating frequency) are considerably lower than those of scientific ones.

Furthermore, so-called "low-frequency" eye trackers are usually missing vital functionality: they do not measure pupil size. In the literature, one can often find information showing the high effectiveness of systems using data related to the pupil or iris size in the authentication process [11], [12], [13], [14]. On the other hand, the disadvantages of such an approach – resulting, for example, from the strong dependence of the pupil size on the illumination intensity or the person's psychophysical state [15], [16] – are equally often pointed out.

A drawback of pupil-related parameters is their susceptibility to changes resulting from a person's aging. Studies show that these parameters may help to effectively assess a person's age [17], [18]. However, this feature is not recommended for biometric systems since it may reduce their effectiveness. According to the literature [17], parameters related to fixations do not indicate an age dependence.

Nevertheless, research shows that despite certain limitations, "low-frequency" eye trackers can be used to identify users. The work of Lyamin and Cherepovskaya [13], [36] and Cantoni et al. [11] confirm this. They employed a 30 Hz eye tracker in their experiments.

Studies conducted by Lohr et al. [19] also confirmed the usefulness of low-frequency eye trackers in biometrics. In the experiments, the authors used data recordings from an eye tracker operating at 1000 Hz, which were then downsampled, obtaining frequencies of 500, 250, 125, 50, and 31.25 Hz. An exponentially dilated convolutional neural network was employed as the classifier. As the frequency decreased, the EER value increased to about 0.3 for 31.25 Hz. In addition, the authors noticed that an interval of up to three years in the acquisition of successive eye movement data did not affect the accuracy of the system performance. It is an important observation for the practical use of this type of data in biometrics. Following this observation, the authors of this article assumed that a tracked eye, delivering new data every 30 ms (frequency around 33.3 Hz), should be sufficient to perform this kind of biometry. However, it must be considered that the downsampled data may differ from the data obtained with the low-speed 30 Hz eye tracker.

B. VISUAL STIMULI

When operating a banking transaction system, different types of content can appear on the screen, i.e., alphanumeric, graphic, and variable background textures. Therefore, if the biometric solution is to be applied practically, one must either view that the selection of stimulation material is not essential [10] or choose a specific phase of the user's interaction with the system during which biometric authentication will occur.

Moreover, how the person looks around the screen will depend on the type of task – whether the person has to memorize an element or find something in the picture. The more "high-level" the task to be performed by the user, the greater the possibility that, with subsequent authentication, the data extracted will be different from the previous patterns. This is due to the way humans analyze different images. For example, eye movements (e.g., fixation points and trajectories) may be completely different when a subject looks at displayed content for the first time and when the same subject views it the next time. Also, a subject's age, gender, or health conditions can affect their eye movements [17], [20], [21], [22]. Thus, such system effectiveness is doubtful [12], [19], [23], [24], [25].

Hence, the more straightforward the task and stimulus, the better the authentication results that can be expected. In [26], the authors included a table comparing different proposals for authentication systems based on eye movement analysis, distinguishing between high-level and low-level. It shows that better results (lower ERR) are obtained with the latter.

The results of the eye movement verification and identification competitions held in 2012 and 2014 also confirm the above observations [27], [28]. The first competition employed data collected using low-level stimuli (so-called "jumping points"), while the second one was based on data related to high-level tasks (viewing pictures of faces). The identification results differ significantly from the disadvantage of the second approach. The best result achieved by the winner of the 2nd competition was about 40%, while for the first, it was 97.7%.

Low-level tasks can include, for example, reading text displayed on a computer screen or watching a point moving on a screen [29]. Most commonly, the mentioned "jumping points" are used as stimuli. The points are displayed sequentially on a screen at different coordinates. The user's task is to focus their gaze on successive points. In this way, it is possible to collect both parameters related to fixations (focusing the gaze on a point) and saccades (when the user moves the eye from point to point) [29]. The "jumping points" approach was adopted by many authors, usually proving to be effective [26], [27], [31].

Naturally, this does not mean other stimuli are not employed in the research. In the experiment described in [30] (in which 322 subjects took part), the authors point to a better performance of reading text excerpts on the effectiveness of user recognition systems. In the experiment, they compared three types of visual stimuli: dots, text, and video. Another conclusion of the study was improved performance thanks to information on saccadic vigor and acceleration.

Also, imagining completely different types of visual stimuli is possible. For example, in [25] and [33], the authors used squares moving along different trajectories. A recent paper by Yin et al. [32] cites about 20 other papers in which stimulus materials are used to elicit eye movement behaviors, including designed jump dots, random jump dots, text, video, and others. Moreover, the authors found that most of the previous works used eye movement recordings with a duration of more than 60 s and times in the range of 5–12 s, as the subject's concentration decreased over time.

C. EYE MOVEMENT FEATURES

Generally, eye movement features applicable to user authorizing can be divided into three categories: frequency domain-based, statistical-based, and spatial-based methods [32]. Moreover, features can be extracted in the time domain or from the spectral representation (discrete Fourier transform

(DFT) and discrete wavelet transform (DWT) are primarily used in the latter case).

Solutions close to the subject of this article exploit information concerning the size of the pupil, the number of fixations and their durations, the parameters of saccadic movements, etc., to determine the user's identity [11], [12], [13], [14], [19], [27], [32]. For example, in the paper [33], the feature vector contained 60 features (out of the 73 collected). They were related to the statistical values determined for, i.a., fixation durations, distance between fixations, saccade speed, pupil diameter and differences between left and right pupil diameters.

However, it can be noted that a group of parameters related to eye movement is used in almost all studies dealing with user authentication. These include:

- saccade duration (saccades discerned at the speed of gaze >300 deg/s);
- fixation duration (fixation detected at the speed of gaze <100 deg/s);
- latency of the saccades (concerning stimulus onset);
- speed and acceleration during saccades;
- speed and acceleration during fixation;
- fixation density.

Similar parameters have been used in many research experiments [11], [19], [26], [31], [35], including low-frequency retrieval eye trackers [13], [36]. Moreover, gaze angles, spatial changes of adjacent fixation points, and saccade distribution maps were used in more recent studies [32].

D. CLASSIFICATION METHODS

Approaches to classifying eye tracker data have evolved over the years. Holland and Komogortsev compared the distribution of fixations and saccades with statistical techniques such as the two-sample t-test, the Ansari-Bradley test, the two-sample Kolmogorov-Smirnov test, and the two-sample Cramér-von Mises test [8].

Lyamin and Cherepovskaya achieved the minimum ERR value of 15.44% for the k-Nearest Neighbors classification method and 16.88% for the algorithm based on Naïve Bayes [13], [36]. They employed a 30 Hz eye tracker, and subjects were asked to follow the mazes visible on the screen with their gaze.

Several studies also used methods inspired by another modality, i.e., text-independent speaker recognition techniques. These include Gaussian mixture model (GMM) or normalizing the speaker models concerning a so-called universal background model (UBM) [37], [38].

Researchers have often used classification performed using a Support Vector Machine (SVM). In [26] and [31], Sluganovic et al. applied (SVM) with the Radial Basis function kernel to classify users employing 16 features extracted from eye tracking. The achieved EER was 6.3%. Furthermore, the authors mentioned the short authentication time (5 s) as a major advantage of the proposed method. In turn, in the paper [33], the authors analyzed various classification algorithms (Support Vector Machine, Nu-Support Vector Machine, Random Forest, and Multi-Layer Perceptron), achieving an accuracy of about 80% and an ERR of less than 10%, depending on the features set, its dimensions, and the classifier. SVM exhibited the best results.

Cantoni et al. employed Naïve Bayes, Random Forest, neural network, and AdaBoost as classifiers in [11]. The FAR values ranged from 6.14% to 23.13%, and the FRR values ranged from 4.11% to 15.20%. Subjects were classified using 42 features and divided into eight groups, and their task was to enter six-digit PINs using the on-screen keypad.

Lohr and Komogortsev employed the DenseNet neural network algorithm for relatively small data sets [35]. In this context, it is worth noting that the sampling rate of 31.5 Hz, for which an EER error value of 23% was achieved, was obtained by downsampling the data stream from a high-speed device.

Recently, Yin et al. reported applying a convolutional neural network (CNN) as a feature extraction tool. The branch network that processes the saccade distribution map (SDM) contains four 2D-CNNs, and each 2D-CNN has 64 filters, a kernel size of (5,5), and a dilation of 1 [32]. The experimental results of their method show reaching 12.48% and 10.62% EER when using eye movement recordings with a duration of 5 s and 12 s, respectively.

In recent years, it has become increasingly common to observe attempts to employ deep learning approaches in classifying eye movement data [39], [40], [41]. The results obtained in this way can be surprisingly good (e.g., EER less than 1%). However, the assumptions made often make it challenging to use such a system in real life. Among others, Jia et al. require the participant to attend a data collection session lasting about 30 minutes [39].

III. MATERIAL AND METHODS

Based on the information gathered in Section II, it is apparent that there is a range of possible approaches to the user authentication process using the data collected with an eye tracker. For the system we developed, we decided to make the following assumptions.

Firstly, we designed a separate procedure to collect visual biometric samples. This procedure will use jumping points as the stimuli since many researchers indicated this method is highly effective (e.g., [26], [27], [31]). In addition, such stimulation reduces the influence of demographic attributes that could interfere with the performance of the biometric system. It also seems straightforward for ordinary users to understand, although we were aware that the time required to collect the data might be a drawback [32].

The choice of low-cost, low-frequency eye trackers automatically imposed some limitations on the parameters that could be collected, practically restricting them to fixations and saccades. Features related to pupil parameters were not available. However, the results of earlier studies ([11], [13], [19], [26], [31], [35], [36]) demonstrated that this approach does not preclude obtaining decent quality performance of such a biometric system. The complete list of features employed in the experiment is provided in section III-B.

The choice of classifier had to reflect that our goal was user authentication instead of person recognition. In other words, the algorithm was only supposed to confirm a person's identity, so it did not identify people within groups. Therefore, the authentication task analyzes the records of a specific person, not represented by a large data set.

The SVM and XGBoost classifiers, considered efficient machine learning algorithms (different from neural networks), were chosen by the authors of this paper. The SVM classifier was used for comparison purposes. The advantages of gradient-enhanced decision trees, such as eXtreme Gradient Boosting (XGBoost), were known through the authors' practical experience. Some benefits of using the XGBoost algorithm include improved accuracy, faster training, and improved scalability. It can also help with feature selection and feature engineering since it can identify the essential features for a given problem. Additionally, XGBoost can handle many features and missing values, making it suitable for many tasks. Finally, it is faster and more efficient than other boosting algorithms, allowing it to handle large datasets quickly and efficiently. It is important to note that the XGBoost algorithm, unlike, for example, neural networks or profound neural networks, can be trained efficiently on relatively small data sets. This is precisely the situation that arises in the case of biometric confirmation of a specific person's identity. Furthermore, with cross-validation, one can confirm, relatively quickly and easily, that the model is not overfitting or underfitting the training dataset.

A. EXPERIMENTAL SETUP

Our experiments were conducted in a room similar to the typical small banking branch space. A special test stand was developed (Fig. 1). Subsequently, ten booths or "kiosks" identical to the one shown in Fig. 1, equipped with 2D and 3D cameras, a voice processing track, a hand-vein scanner, and eye-tracking devices, were manufactured and installed in bank branches. Since a real-life bank branch is not a suitable place to conduct research, the developed software is based on experiments obtained under laboratory conditions, where the biometric kiosk was installed at the Gdańsk University of Technology. The engineered stand (kiosk) was equipped with a 24-inch screen with a resolution of 1920 × 1080 pixels. A 30 Hz eye tracker was mounted at the bottom of the monitor (at the height of 125 cm from the floor). The subjects stood in front of the test stand at a distance of about 55–70 cm from the monitor during the study. Distances translate into angles, hence many parameters are calculated with a mapping to angles rather than pixels.

The gaze data collection began with calibrating the eye tracker for the particular subject's way of looking. The calibration (lasting 15 s) used blue dots as visual objects, while data acquisition for biometric authentication was based on red dots (jumping points) displayed on the vertical screen.

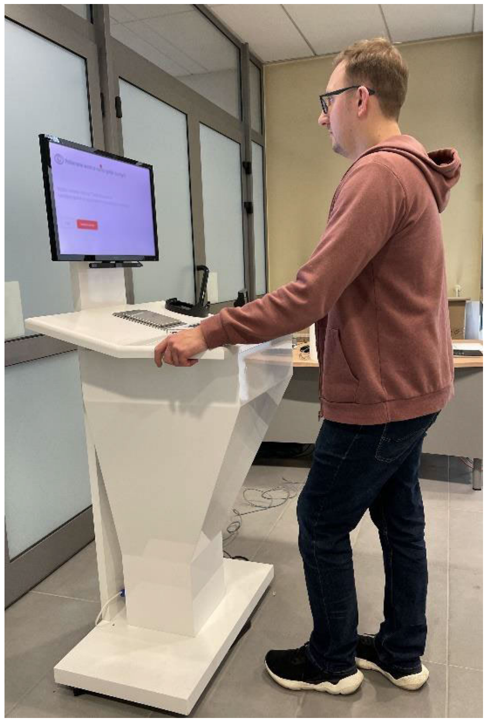

FIGURE 1. View of the test stand ("Biometric kiosk").

Next, a screen appeared (for 2 s) with a message telling the subject to focus their gaze on the red dots displayed on the monitor. Then a starting point was displayed in the center of the screen with a radius of 25 pixels. Each subsequent point was displayed for 1.5 seconds and had a radius of 10 pixels (+3 s for displaying the start point + 3 s for the endpoint). The test procedure ended with displaying the final point, with the same size and location as the starting point. Hence, the total observation time for visual patterns was $27 \times 1.5 + 2 \times 3 = 46.5$ s. Hence, the global experiment time for each subject was 50 seconds. We used 27 test points, so 27 parameter vectors were created. Each point was represented by one vector of 20 parameters. Once the data acquisition was completed, the parameterization process of the collected data was automatically started. In addition, the raw data acquired from the eye tracker was also saved for possible further data analysis.

In total, 39 volunteers took part in the experiment (33 males and six females). The male-to-female ratio was similar to that at the Faculty of Electronics, Telecommunications, and Informatics, where the study was conducted (2,486 students, including 442 women). Signed informed consent for the use of biometric data was obtained from all subjects involved in the experiment. Most participants (31) were university students. In addition, 9 participants were wearing glasses during the data collection, and no issues regarding the glasses were observed.

For 35 participants, data were collected twice. The sessions were spaced at least several minutes apart. For four participants, data were collected only once (for some technical reasons).

All the software used in the experiment (for data collection, parameterization, and verification) was implemented in Python (version 3.9).

B. DATA PARAMETRIZATION

The process of parameterizing the collected data was carried out in two steps. The first can be referred to as preprocessing. In this step, erroneous values of gaze point coordinates smaller than zero and larger than the horizontal or vertical screen resolution were removed. In addition, gaze movement data only for the test points were left – values for the start and end points were removed. The following parameters were then determined for each captured gaze point:

- X coordinate (in pixels) of the gaze point;
- Y coordinate (in pixels) of the gaze point;
- normalized time of data capture (starting from 0);
- time between successive data acquisitions: $\Delta t = t_n - t_{n-1}$;
- distance between eyes and screen for a given point (necessary to transform to angular values): $dist_{mean} = (dist_{Leye} + dist_{Reye}): 2$, where: $\mathrm{dist}_{\mathrm{Leye}}$ – distance for the left eye and $\mathrm{dist}_{\mathrm{Reye}}$ – distance for the right eye;
- distance between two consecutive points on the X axis (in pixels): $\Delta x = x_n - x_{n-1}$;
- distance between two consecutive points on the Y axis (in pixels): $\Delta y = y_n - y_{n-1}$;
- angular distance between two consecutive points on the X-axis: $\Delta x_{deg} = \Delta x \cdot k_{Xpix-deg}$, where:

$$k_{Xpix-deg} = tan^{-1} \frac{\frac{screen\_size_X}{screen\_resolution_X}}{dist_{mean}} \tag{1}$$

$screen\_size_X$ represents X-axis screen size,

$screen\_resolution_X$ represents X-axis screen resolution;

- angular distance between consecutive points on Y the axis: $\Delta y_{deg} = \Delta y \cdot k_{Ypix-deg}$, where:

$$k_{Ypix-deg} = tan^{-1} \frac{\frac{screen\_size_Y}{screen\_resolution_Y}}{dist_{mean}} \tag{2}$$

$screen\_size_Y$ represents Y-axis screen size,

$screen\_resolution_Y$ represents Y-axis screen resolution;

- angular distance between subsequent points:

$$\Delta s_{deg} = \sqrt{\Delta x_{deg}^2 + \Delta y_{deg}^2} \tag{3}$$

- angular speed of movement between consecutive points on the X axis: $v_X = \Delta x_{deg} : \Delta t$;
- angular speed of movement between consecutive points on the Y axis: $v_Y = \Delta y_{deg} : \Delta t$;
- angular speed of movement between consecutive points: $v = \Delta s_{deg} : \Delta t$;
- angular acceleration of movement between consecutive points on the X axis: $a_X = \Delta v_X : \Delta t$;
- angular acceleration of movement between consecutive points on the Y-axis: $a_Y = \Delta v_Y : \Delta t$;
- angular acceleration of movement between consecutive points: $a = \Delta v : \Delta t$;
- distance between eyes (eye-spacing) in [mm].

After performing the calculations described above, erroneous readings were once again filtered out:

- results for which the distance between the eyes is less than 40 [mm] were removed;
- results for which the angular speed of eye movement between consecutive points was higher than 1000 [degrees/s] were removed – because, as is known from the literature, physiological conditions make it impossible to obtain such speeds [42].

During the second step of calculating the gaze movement parameters, based on the time dependencies between when the test points are displayed and eye movements, it is possible to divide the collected data into two groups:

- saccades – for which the angular speed of movement between points was greater than or equal to 25 [degrees/s]; these are associated with rapid eye movements, which move the gaze from one test point to another. Employing a 33 Hz eye tracker, some errors resulting from the limited temporal resolution must be expected.
- fixations – for which the angular speed of movement between points was less than 25 [deg/s] and the displacement between points on the X and Y axis was also small (below 0.15 [deg]), they are related to focusing the gaze on a particular test point.

For the case of saccades, the latency of saccadic movements for each test point (the subject's reaction time to the presentation of a new test point) is also calculated. In addition, for each of the 27 fixation areas, the median of the coordinates on the X and Y axes is calculated. The points for which the distance from the center of the area thus calculated is greater than the assumed value are filtered out. Then the fixation areas are approximated with polygons, and the area and perimeter of each polygon are calculated. Fig. 2 contains visualizations of the detected fixation regions for sample data, and Fig. 3 shows examples illustrating the approximation of focus areas with polygons.

As a result, 20 parameters are obtained for each of the 27 test points (equations based on standard definitions of statistical measures are omitted in the following summary):

A. delay of saccadic movements;
B. average speed of saccadic movements;
C. maximum saccadic speed;
D. standard deviation of saccade speed;
E. mean acceleration (absolute value) of saccadic motion;
F. standard deviation of acceleration (absolute value) of saccadic motion;
G. maximum acceleration (absolute value) of saccadic motion;
H. duration of saccadic motion;

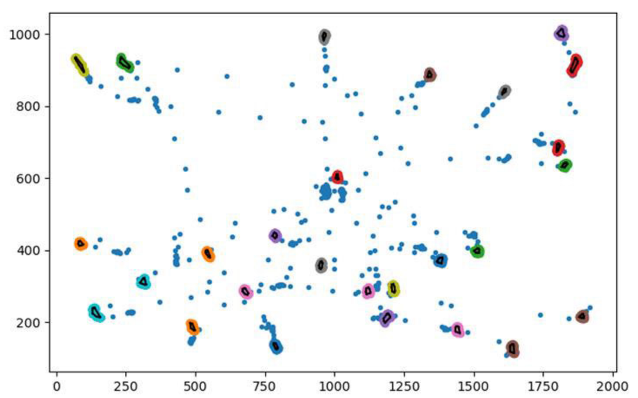

FIGURE 2. Visualization of gaze fixation areas for individual test points (sample data).

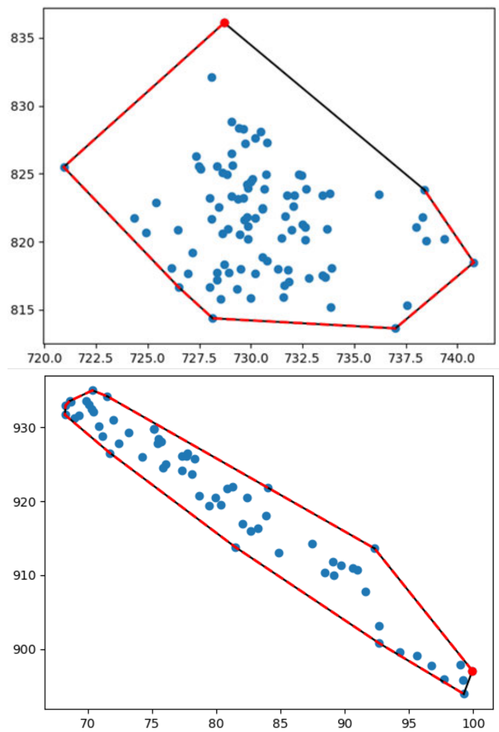

FIGURE 3. Polygon approximations of sample gaze areas (example data).

I. mean distance between eyes for saccadic movements;
J. mean speed of movements during fixation;
K. standard deviation of the speed of movements during fixation;
L. maximum speed of movements during fixation;
M. mean acceleration (absolute value) of movements during fixation;
N. standard deviation of acceleration;
O. maximum acceleration (absolute value) of movements during fixation;
P. duration of fixation;
Q. mean distance between eyes during fixation;
R. area of the polygon containing fixations;
S. fixation density (calculated as the area of the polygon divided by the number of points contained in the polygon);
T. perimeter of the polygon containing fixations.

The basic unit of data is one sample, i.e., a set of 20 parameters associated with a single test point. Samples acquired from the same person are grouped into sets containing 27 samples. Since we have the calculated values of individual parameters, it might be possible to try to visualize them, but the tests carried out have shown that little can be deduced from such visualizations. Attempts to project parameter values onto planes have demonstrated that their two-dimensional representations are shuffled with each other. An algorithm based on machine learning should cope with their separation because the problem must be solved in a multidimensional parameter hyperspace.

C. CLASSIFIERS

Support Vector Machine (SVM) and Gradient Boosted Trees (GBT) were used as the classifiers in our experiment. It was decided to employ a class SVC (Support Vector Classification) from the Scikit-learn library. The model was trained as a binary classification model, which means that it outputted only two classes:

- 1 for genuine (positive) samples;
- 0 for impostors (negative) samples.

The model was finetuned with 3 hyperparameters:

- Kernel – rbf kernel;
- Regularization parameter C, values [5e-1, 1e-1, 5e-2, 55e-3] were tested;
- Kernel coefficient gamma values [1], [2], [5], [7], [10] were also tested.

Before the main training, hyperparameter finetuning was performed, which searched for the best parameters C and gamma values. The Scikit-learn class GridSearchCV was used for that task. It is a function that evaluates a given model in (by default) a 5-fold cross-validation manner, testing each combination of hyperparameters. In our case, it took one argument – the make_pipeline function from sklearn, which represents three arguments to the function:

- estimator, which is an SVC with balanced weights of classes;
- param_grid, which is a dictionary of hyperparameters for finetuning; in our case, they were C and gamma parameters;
- and score, which measures the performance of a model; in our case, we used ROC AUC.

At the end of the pipeline, a proper model was trained with values of the hyperparameters that gave the highest ROC AUC score during GridSearchCV.

The result of the verification was a vector containing 0 and 1. Its length equals the number of samples provided to the model input. The ratio of "1" to "0" was the result of the final verification.

The second verification method relies on the Gradient Boosted Trees (GBT) model. The term "Gradient Boosting" originates from the paper [43]. Gradient Boosted Trees is an ensemble supervised learning method of the gradient boosting type. It consists of learning the model as a sequence of successive classification or regression trees so that each successive tree in its training aims to minimize the errors received by the previous tree. During training, the weights of the model are optimized. This optimization was done using the gradient from the previously received tree and minimizing the error received by it.

The open-source XGBoost library and its implementation of boosted trees were used for this purpose. From this library, the xboost booster class was used and returned by the xgboost.train function. This function adopts the following arguments to train a booster:

- maximum tree depth: 2;
- eta – step size shrinkage used in the update to prevent overfitting; after each boosting step, eta shrinks the feature weights to make the boosting process more conservative;
- objective function: binary-logistic, which means logistic regression for binary classification;
- number of threads running in parallel for tree processing: 4;
- evaluation metric: Area Under Curve.

All the above metrics should be passed to the function as a dictionary. Apart from them, there are a few other arguments to pass:

- dtrain: data to be trained, which must be stored as an object of DMatrix class;
- num_boost_round: number of boosting iterations;
- evals: list of validation sets for which metrics will be evaluated during the training.

The model verification process resulted in a vector of values between <0, 1>, with a length equal to the size of the input data vector. The final classification was determined according to the following formula:

$$y = \vdash \begin{matrix} 1 & m > \frac{1}{2}n \\ 0 & m \leq \frac{1}{2}n \end{matrix} \tag{4}$$

where:

n – number of samples in a given vector;

m – number of samples whose predicted value is greater than the given threshold of t = 0.9.

The threshold of t = 0.9 was used as a result of the tests conducted because the achieved TAR (True Acceptance Rate) level for a given sample set allowed a high degree of separability between real users and so-called impostors. The 0.5 thresholds for sample classification, adopted as standard in UM models, was also used in pilot tests, but the model achieved less satisfactory results.

D. DATA PROCESSING

During the experiments, 35 people had their results collected twice, four only once. Since the model was created for a particular person, samples of other users input into the model are treated as impostors. The order of the samples also does not play a role because, in the course of analyzing the results, the model takes into account the separate responses to individual samples and then averages the result. This approach corresponds to real-world conditions because it takes into account that the user may look at things differently at the beginning of the study (when they are learning the rules of the study). They may also respond differently at the end (when they are already a bit fatigued and have diminished attention).

The analyses were performed for the following variants:

A. model created based on 27 samples, verification also based on 27 samples;
B. model created based on 27 samples, and verification based on 20 samples.
C. model created based on 27 samples, verification based on 15 samples;
D. model created based on 20 samples, verification also based on 20 samples.

Where fewer than 27 samples were used, the first 20/15 samples were picked.

The main idea was to investigate whether reducing the number of test points (and consequently shortening the data collection process) significantly affects the accuracy of the system.

Regardless of the classifier used, only samples from the first sessions were used for the training. The model was trained for a given participant by employing:

- 27/20 samples of a given subject, i.e., class 1 data;
- 30 samples from a group of selected subjects (3 samples each from 10 subjects), treated as class 0 data (impostors).

For impostors, samples 6–8 from the first session were used. The goal was to select random samples and keep the user's attention as high as possible [32]. The first five samples were skipped to avoid errors from the subject becoming familiar with the test procedure.

Each created model was then tested using samples collected for all participants and all sessions. This means that the model for a given subject was tested using samples from both sessions (or only the first session – when the person did not attend the second session) and samples from all other users (from the first and, if available, second sessions).

Thus, for each classifier and each variant, we obtained a total of:

- 74 authentication results when samples belonging to the owner of the model were provided to the model input (35 participants with two series and 4 participants with only one series: $35 \times 2 + 4 \times 1$ representing the "genuine" group;
- 2,847 authentication results in situations where impostors' samples were submitted to the model input (72

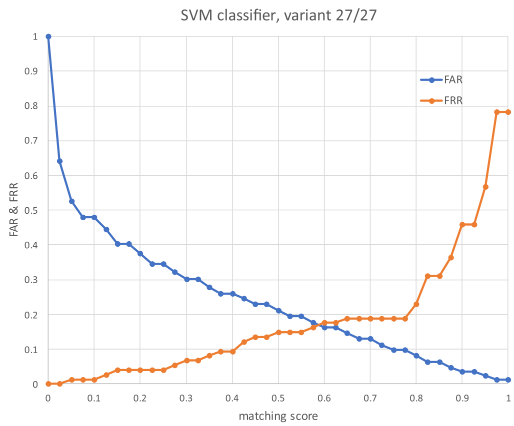

FIGURE 4. Exemplary FAR and FRR curves (SVM classifier employed, variant 27/27).

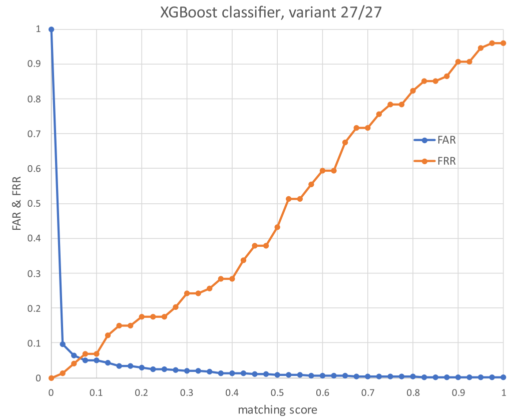

FIGURE 5. Exemplary FAR and FRR curves (XGBoost classifier employed, variant 27/27).

TABLE 1. EER values.

classifier/variant	27/27	27/20	27/15	20/20
SVM	0.1757	0.1835	0.1878	0.1621
XGBoost	0.0676	0.0576	0.0782	0.0800

results each for 35 participants and 73 results each for 4 participants: $72 \times 35 + 73 \times 4$).

IV. RESULTS

Experiments were carried out to evaluate the accuracy of the proposed biometric system based on eye tracker data analysis for two different classifiers (SVMs and XGBoots) and four variations in the number of samples used for training and testing the model. FAR (False Accept Rate), FRR (False Reject Rate), and TAR (True Accept Rate) values were determined to evaluate the accuracy of the proposed biometric system. This way, it was possible to determine the EER values and plot the ROC (Receiver Operating Characteristic) curves. The estimation of EER values required adopting an approximation method – due to the way the verification result was determined, the obtained values of FAR and FRR were discrete (not continuous). As a consequence, the FAR and FRR curves have a stepped shape, as illustrated in Figs. 4 and 5. Thus, the ERR values were determined by finding the minimum gap between the FAR and FRR values, and then the higher value of the two parameters was read. The results obtained are summarized in Table 1 and shown in Fig. 6.

The EER values are noticeably higher for the SVM classifier, while the differences between the values for the variants are minimal. Thus, it is difficult to select or reject any variants unambiguously.

Figs. 7 and 8 show the ROC curves obtained for both classifiers. The curves for each variant are marked with different colors in both figures. The shape of the curves demonstrates the greater accuracy of participant verification using the XGBoost classifier.

The last parameters analyzed were the TAR values for FAR = 0.1 and FAR = 0.01. Unfortunately, for the SVM

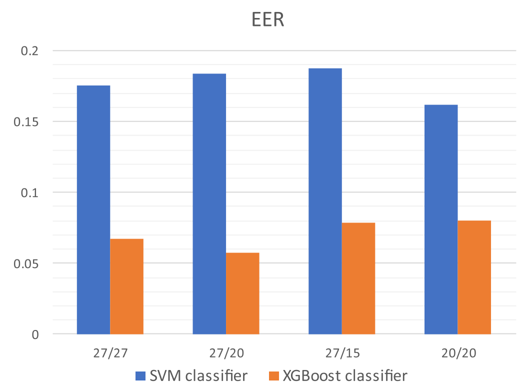

FIGURE 6. EER values for both classifiers.

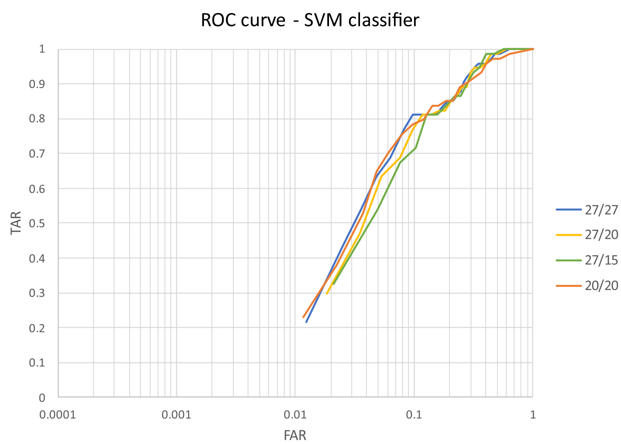

FIGURE 7. ROC curves for SVM classifier (all variants).

classifier and variants 27/20 and 27/15, it was not possible to obtain a TAR value for FAR = 0.01 (the FAR parameter did not reach such a value). The determined values are provided in Tables 2 and 3 and presented in Fig. 9 and 10.

The results of the experiments showed that the classifier based on XGBoost outperformed the SVM-based one. This is valid for all calculated parameters used to assess the quality of

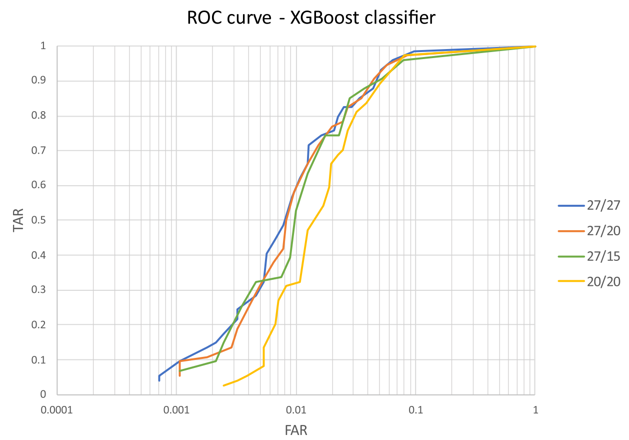

FIGURE 8. ROC curves for XGBoost classifier (all variants).

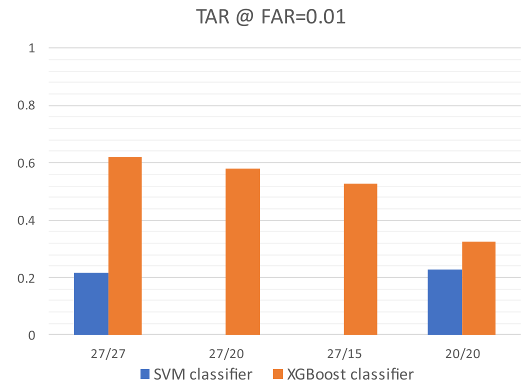

FIGURE 10. TAR values for FAR = 0.01.

TABLE 2. TAR @ FAR = 0.1.

classifier/variant	27/27	27/20	27/15	20/20
SVM	0.8108	0.7703	0.7162	0.7838
XGBoost	0.9865	0.9865	0.9865	0.9595

TABLE 3. TAR @ FAR = 0.01.

classifier/variant	27/27	27/20	27/15	20/20
SVM	0.2162	-	-	0.2297
XGBoost	0.6216	0.5811	0.5270	0.3243

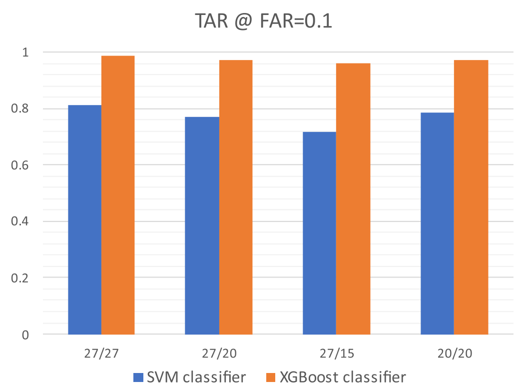

FIGURE 9. TAR values for FAR = 0.1.

biometric systems. The XGBoost classifier achieved satisfying results, although the TAR values for FAR = 0.01 slightly differ from the expected ones.

The accuracy of the system's performance was also verified when the number of samples used for both the training and testing processes was reduced. The differences between the achieved EER, FAR, FRR, and TAR values are relatively small. However, they might provide a rationale for reducing the number of test points displayed to the user during the verification process. Eliminating a single test point display minimizes the duration of the procedure by 1.5 seconds. Therefore, it seems that the number of test points displayed (and thus the number of samples corresponding to the number of rows of the parameter matrix) during the collection of reference data should not be reduced. However, large-scale experiments would be needed to confirm this opinion.

V. LIMITATIONS

In the experiments described in this paper, the applicability of using a low-frequency, inexpensive eye tracker was investigated to verify a person's identity. Our approach thus focused on a *1:1* authentication case rather than a *1:many* (recognition case) authentication process.

The most significant limitation of the created system is related to the capabilities of the eye tracker used. Its operating frequency of 30 Hz is the minimum for this type of application. It was also insufficient to use pupil-related features. The specifics of the eye tracker also made it practically impossible to use publicly available datasets to compare our results with other studies. Our priority was to utilize the limited capabilities of the eye tracker comprehensively. Hence, for example, we used the distance between the eyes as a feature. It is not present in other datasets (e.g. [17], [18], [19], [27], [28]). Furthermore, such incompatibility with different approaches provides additional protection against fraud attempts.

From the user's point of view, the relatively long time taken to collect biometric data can be tiresome, which is especially noticeable when the person wants to authenticate. This can result in a loss of concentration, incorrect completion of the task, and, consequently, erroneous results. The results achieved (for a reduced number of test points) show that there is an option to reduce the number of test points and, thus, the duration of the procedure. Therefore, changing the data collection procedure and making the number of points displayed dependent on the verification results achieved seems advisable. It means setting a threshold beyond which the person's identity gets confirmed (or rejected) and thus terminating the data acquisition procedure.

Another problem that may affect our results is the relatively small number of participants willing to submit their biometric data. The short interval between test sessions that was chosen

TABLE 4. Comparison to existing systems.

ref	number of participants	stimulus	eye tracker operating frequency [Hz]	pupil-related parameters	classifier	EER [%]	notes
[11]	45	PIN	30	yes	Naïve Bayes	NA	FRR 8.24% FAR 10.72%
					Random Forest	NA	FRR 7.14% FAR 12.48%
					Neural Network	NA	FRR 4.55% FAR 6.35%
					AdaBoost	NA	FRR 4.1% FAR 6.14%
[13]	45	maze	30	yes	k-Nearest Neighbors	15.44	
					Naïve Bayes	16.18	
[26]	30	jumping points	500	no	SVM with Radial Basis	6.3	
[33]	44	moving objects	150	yes	SVM	12.46	
					NuSVM	14.17	
					Random Forest	9.53	
					Multi-Layer Perceptron	14.60	
[19]	59	reading task	1000	yes	exponentially dilated convolutional neural network	14.20	dataset used
			31.25	yes		26.66	dataset used, downsampled data
[35]	59	reading task	500	yes	DenseNet-based convolutional neural network	5.66	dataset used
			31.25	yes		23.37	dataset used, downsampled data
[32]	31	horizontal saccade task, random squinting task, reading task	1000	yes	convolutional neural network (CNN)	5.25	dataset used
this research	39	jumping points	30	no	SVM	17.57	
					XGBoost	6.76	

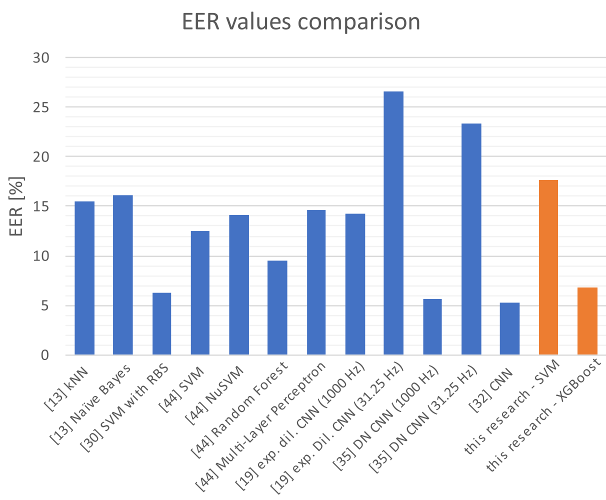

FIGURE 11. EER values comparison to existing systems.

for organizational reasons was also a limiting factor. Taking this into account, it is planned to carry out another series of experiments in more realistic conditions involving a larger and more diverse population of participants. The results will be expected to give a conclusive answer on the applicability of the developed system in real life.

Still, we believe that the proposed solution can be used in real-life applications – not necessarily as a completely independent biometric solution, but, for example, as a component of a multimodal biometric system using data fusion for a person's authentication. Furthermore, it was possible to detect the absence of looking at the screen or a failure in tracking consecutive points, which can provide means to verify the "liveness" of the subject, a valuable feature supporting other biometric modalities.

VI. CONCLUSION

Eye-tracking technology has shown promising results in recent studies. With EER errors lower than 10%, it can be considered a prospective method for biometric authentication, especially in multimodal systems, whose development is a current trend. The exact EER level would depend on many factors, such as the gaze tracking device selection, the applied methodology, the dataset used, and the specific conditions and environments in which the authentication is performed. In Table 4 and in Fig. 11, we have summarized the results obtained by different researchers under a range of conditions with the results of our experiment. It should be noted that the EER for the XGBoost-based classifier is comparable to the values obtained in studies using considerably more sophisticated eye trackers, which is positive evidence of the approach we adopted. Furthermore, it can be noted that the number of participants in our experiment does not differ significantly from the number of participants in other studies, including when using more comprehensive datasets.

The findings of the presented research are as follows:

- It has been shown that it is relatively easy to collect biometric patterns and perform authentication in solutions where an eye tracker is used to visually control the stand;
- It was found that the requirements for the user's positions during data enrollment are not restrictive, since the system performed well with the different heights of participants and spontaneous changes in their location. This is crucial information since, in the majority of studies, subjects were allowed minimal freedom of movement of the head (e.g. [30], [33]), which seriously limits the practical applications;
- It was demonstrated that a low-frequency eye tracker, refreshing data with a time resolution of 30 ms, is sufficient to achieve an EER in the range of 6.76%–8% with adequate feature extraction and an XGBoost classifier employed. Therefore, the results achieved are comparable in efficiency to those obtained by other researchers. However, we also found that the XGBoost classifier performance was significantly higher than those of the SVMs previously used by many researchers, e.g., [21], [26], [27], and [33].

Since gathering eye movement data, automatically parameterizing it, and verifying the user's identity is possible, the solution developed in the study can be used in real-life conditions. This raises prospects for the broader application of eye movement tracking as a biometric modality, mainly operating in controlled conditions, such as kiosks, booths, and consoles.

REFERENCES

[1] S. Zhu, J. Qi, J. Hu, and S. Hao, "A new approach for product evaluation based on integration of EEG and eye-tracking," *Adv. Eng. Inform.*, vol. 52, Apr. 2022, Art. no. 101601, pp. 1–15, doi: 10.1016/j.aei.2022.101601.

[2] A. Czyżewski, P. Bratoszewski, P. Hoffmann, M. Lech, and M. Szczodrak, "The project IDENT: Multimodal biometric system for bank client identity verification," in *Multimedia Communications, Services and Security* (Communications in Computer and Information Science), vol. 785, A. Dziech and A. Czyżewski, Eds. Cham, Switzerland: Springer, 2017, pp. 16–34, doi: 10.1007/978-3-319-69911-0_2.

[3] J. A. Unar, W. C. Seng, and A. Abbasi, "A review of biometric technology along with trends and prospects," *Pattern Recognit.*, vol. 47, no. 8, pp. 2673–2688, Aug. 2014, doi: 10.1016/j.patcog.2014.01.016.

[4] A. Bulling, F. Alt, and A. Schmidt, "Increasing the security of gaze-based cued-recall graphical passwords using saliency masks," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Austin, TX, USA, May 2012, pp. 3011–3020, doi: 10.1145/2207676.2208712.

[5] M. Mehrubeoglu and V. Nguyen, "Real-time eye tracking for password authentication," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Las Vegas, NV, USA, Jan. 2018, pp. 1–4, doi: 10.1109/ICCE.2018.8326302.

[6] P. Kasprowski and K. Harezlak, "Protecting from lunchtime attack using an uncalibrated eye tracker signal," in *Proc. ACM Symp. Eye Tracking Res. Appl.*, New York, NY, USA, Jun. 2020, pp. 1–5, doi: 10.1145/3379156.3391348.

[7] W. Fuhl, D. Weber, and S. Eivazi, "The gaze and mouse signal as additional source for user fingerprints in browser applications," 2021, *arXiv:2101.03793*.

[8] S. S. Ketab, N. L. Clarke, and P. S. Dowland, "A robust e-invigilation system employing multimodal biometric authentication," *Int. J. Inf. Educ. Technol.*, vol. 7, no. 11, pp. 796–802, 2017, doi: 10.18178/ijiet.2017.7.11.975.

[9] P. Szczuko, A. Harasimiuk, and A. Czyżewski, "Evaluation of decision fusion methods for multimodal biometrics in the banking application," *Sensors*, vol. 22, no. 6, p. 2356, Mar. 2022, doi: 10.3390/s22062356.

[10] C. D. Holland and O. V. Komogortsev, "Complex eye movement pattern biometrics: The effects of environment and stimulus," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp. 2115–2126, Dec. 2013, doi: 10.1109/TIFS.2013.2285884.

[11] V. Cantoni, T. Lacovara, M. Porta, and H. Wang, "A study on gaze-controlled PIN input with biometric data analysis," in *Proc. 19th Int. Conf. Comput. Syst. Technol.*, Ruse, Bulgaria, Sep. 2018, pp. 99–103, doi: 10.1145/3274005.3274029.

[12] F. Deravi and S. Guness, "Gaze trajectory as a biometric modality," in *Proc. Int. Conf. Bio-Inspired Syst. Signal Process.*, Rome, Italy, 2011, pp. 335–341, doi: 10.5220/0003275803350341.

[13] A. V. Lyamin and E. N. Cherepovskaya, "An approach to biometric identification by using low-frequency eye tracker," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 4, pp. 881–891, Apr. 2017, doi: 10.1109/TIFS.2016.2639342.

[14] S. Sarhan, S. Alhassan, and S. Younis, "Multi-biometric systems: Comparative study," *Arabian J. Sci. Eng.*, vol. 41, no. 2, pp. 443–457, Feb. 2017, doi: 10.1007/s13369-016-2241-0.

[15] S. Eberz, G. Lovisotto, K. B. Rasmussen, V. Lenders, and I. Martinovic, "28 blinks later: Tackling practical challenges of eye movement biometrics," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, London, U.K., Nov. 2019, pp. 1187–1199, doi: 10.1145/3319535.3354233.

[16] I. Griswold-Steiner, Z. Fyke, M. Ahmed, and A. Serwadda, "Morph-a-Dope: Using pupil manipulation to spoof eye movement biometrics," in *Proc. 9th IEEE Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, New York, NY, USA, Nov. 2018, pp. 543–552, doi: 10.1109/UEMCON.2018.8796625.

[17] C. Bisogni, L. Cascone, and F. Narducci, "Periocular data fusion for age and gender classification," *J. Imag.*, vol. 8, no. 11, p. 307, Nov. 2022, doi: 10.3390/jimaging8110307.

[18] L. Cascone, C. Medaglia, M. Nappi, and F. Narducci, "Pupil size as a soft biometrics for age and gender classification," *Pattern Recognit. Lett.*, vol. 140, pp. 238–244, Dec. 2020, doi: 10.1016/j.patrec.2020.10.009.

[19] D. Lohr, H. Griffith, and O. V. Komogortsev, "Eye know you: Metric learning for end-to-end biometric authentication using eye movements from a longitudinal dataset," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 4, no. 2, pp. 276–288, Apr. 2022, doi: 10.1109/TBIOM.2022.3167633.

[20] N. Sammaknejad, H. Pouretemad, C. Eslahchi, A. Salahirad, and A. Alinejad, "Gender classification based on eye movements: A processing effect during passive face viewing," *Adv. Cognit. Psychol.*, vol. 13, no. 3, pp. 232–240, Sep. 2017, doi: 10.5709/acp-0223-1.

[21] D. Lagun, C. Manzanares, S. M. Zola, E. A. Buffalo, and E. Agichtein, "Detecting cognitive impairment by eye movement analysis using automatic classification algorithms," *J. Neurosci. Methods*, vol. 201, no. 1, pp. 196–203, Sep. 2011, doi: 10.1016/j.jneumeth.2011.06.027.

[22] Y. Mao, Y. He, L. Liu, and X. Chen, "Disease classification based on synthesis of multiple long short-term memory classifiers corresponding to eye movement features," *IEEE Access*, vol. 8, pp. 151624–151633, 2020, doi: 10.1109/ACCESS.2020.3017680.

[23] M. S. Castelhano, M. L. Mack, and J. M. Henderson, "Viewing task influences eye movement control during active scene perception," *J. Vis.*, vol. 9, no. 6, pp. 1–15, Mar. 2009, doi: 10.1167/9.3.6.

[24] V. Cantoni, M. Musci, N. Nugrahaningsih, and M. Porta, "Gaze-based biometrics: An introduction to forensic applications," *Pattern Recognit. Lett.*, vol. 113, pp. 54–57, Oct. 2018, doi: 10.1016/j.patrec.2016.12.006.

[25] M. Porta and A. Barboni, "Strengthening security in industrial settings: A study on gaze-based biometrics through free observation of static images," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Zaragoza, Spain, Sep. 2019, pp. 1273–1277, doi: 10.1109/ETFA.2019.8868961.

[26] I. Sluganovic, M. Roeschlin, K. B. Rasmussen, and I. Martinovic, "Using reflexive eye movements for fast challenge-response authentication," in *Proc. ACM CCS*, Vienna, Austria, 2016, pp. 1056–1067, doi: 10.1145/2976749.2978311.

[27] P. Kasprowski, O. V. Komogortsev, and A. Karpov, "First eye movement verification and identification competition at BTAS 2012," in *Proc. IEEE 5th Int. Conf. Biometrics, Theory, Appl. Syst. (BTAS)*, Arlington, VA, USA, Sep. 2012, pp. 195–202, doi: 10.1109/BTAS.2012.6374577.

[28] P. Kasprowski and K. Harężlak, "The second eye movements verification and identification competition," in *Proc. IEEE Int. Joint Conf. Biometrics*, Clearwater, FL, USA, Sep./Oct. 2014, pp. 1–6, doi: 10.1109/BTAS.2014.6996285.

[29] C. Galdi and M. Nappi, "Eye movement analysis in biometrics," in *Biometrics Under Biomedical Considerations*, A. Nait-Ali, Ed. Singapore: Springer, Dec. 2018, pp. 171–183, doi: 10.1007/978-981-13-1144-4_8.

[30] I. Rigas, O. Komogortsev, and R. Shadmehr, "Biometric recognition via eye movements: Saccadic vigor and acceleration cues," *ACM Trans. Appl. Perception*, vol. 13, no. 2, pp. 1–21, Mar. 2016, doi: 10.1145/2842614.

[31] I. Sluganovic, M. Roeschlin, K. B. Rasmussen, and I. Martinovic, "Analysis of reflexive eye movements for fast replay-resistant biometric authentication," *ACM Trans. Privacy Secur.*, vol. 22, no. 1, pp. 1–30, Feb. 2019, doi: 10.1145/3281745.

[32] J. Yin, J. Sun, J. Li, and K. Liu, "An effective gaze-based authentication method with the spatiotemporal feature of eye movement," *Sensors*, vol. 22, no. 8, p. 3002, Apr. 2022, doi: 10.3390/s22083002.

[33] M. Porta, P. Dondi, N. Zangrandi, and L. Lombardi, "Gaze-based biometrics from free observation of moving elements," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 4, no. 1, pp. 85–96, Jan. 2022, doi: 10.1109/TBIOM.2021.3130798.

[34] J. Z. Lim, J. Mountstephens, and J. Teo, "Eye-tracking feature extraction for biometric machine learning," *Frontiers Neurorobotics*, vol. 15, pp. 1–21, Feb. 2022, doi: 10.3389/fnbot.2021.796895.

[35] D. Lohr and O. V. Komogortsev, "Eye know you too: Toward viable end-to-end eye movement biometrics for user authentication," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 3151–3164, 2022, doi: 10.1109/TIFS.2022.3201369.

[36] A. V. Lyamin and E. N. Cherepovskaya, "Biometric student identification using low-frequency eye tracker," in *Proc. 9th Int. Conf. Appl. Inf. Commun. Technol. (AICT)*, Rostov on Don, Russia, Oct. 2015, pp. 191–195, doi: 10.1109/ICAICT.2015.7338544.

[37] T. Kinnunen, F. Sedlak, and R. Bednarik, "Towards task-independent person authentication using eye movement signals," in *Proc. Symp. Eye-Tracking Res. Appl. (ETRA)*, Austin, TX, USA, Jan. 2010, pp. 187–190, doi: 10.1145/1743666.1743712.

[38] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digit. Signal Process.*, vol. 10, nos. 1–3, pp. 19–41, Jan. 2000, doi: 10.1006/dspr.1999.0361.

[39] S. Jia, D. H. Koh, A. Seccia, P. Antonenko, R. Lamb, A. Keil, M. Schneps, and M. Pomplun, "Biometric recognition through eye movements using a recurrent neural network," in *Proc. IEEE Int. Conf. Big Knowl. (ICBK)*, Singapore, Nov. 2018, pp. 57–64, doi: 10.1109/ICBK.2018.00016.

[40] A. Abdelwahab and N. Landwehr, "Deep distributional sequence embeddings based on a Wasserstein loss," 2019, *arXiv:1912.01933*.

[41] L. A. Jäger, S. Makowski, P. Prasse, S. Liehr, M. Seidler, and T. Scheffer, "Deep eyedentification: Biometric identification using micro-movements of the eye," in *Machine Learning and Knowledge Discovery in Databases*, U. Brefeld, E. Fromont, A. Hotho, A. Knobbe, M. Maathuis, and C. Robardet, Eds. Cham, Switzerland: Springer, Apr. 2020, pp. 299–314, doi: 10.1007/978-3-030-46147-8_18.

[42] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, J. Halszka, and J. van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, Great Britain: Oxford Univ. Press, 2015.

[43] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.

PIOTR ODYA (Member, IEEE) received the M.Sc. degree from the Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Poland, in 1999, and the Ph.D. degree from the Gdańsk University of Technology, in 2007. His thesis was related to the problem of sound quality improvement in contemporary broadcasting studios. The goal of his Ph.D.'s thesis concerned methods and algorithms for correcting stuttering.

He is currently an Assistant Professor with the Department of Multimedia Systems and the Deputy Head of the Department of Multimedia Systems. From 2007 to 2014, he coordinated two projects to develop new multimodal interfaces. He is the author or coauthor of more than 70 conference papers and journal articles. His research interests include audio signals processing, auditory-visual signals perception, telemedicine, video compression algorithms, and multichannel sound systems.

FRANCISZEK GÓRSKI (Member, IEEE) received the B.E. degree in computer science from the Gdańsk University of Technology, in 2021, where he is currently pursuing the degree in informatics (specialization in machine learning). He is a Researcher with the Department of Multimedia Systems, Gdańsk University of Technology, and is collaborating with the Centre of Biostatistics and Bioinformatics Analysis, Medical University of Gdańsk. His research interests include applying machine learning and deep learning algorithms to solve problems in biology and medicine. He is the Co-Chairperson of the Gradient Science Club, GUT, whose mission is to popularize knowledge about deep learning among students.

ANDRZEJ CZYŻEWSKI (Member, IEEE) received the M.Sc. degree in sound engineering from the Gdańsk University of Technology, in 1982, and the Ph.D. and D.Sc. degrees from the Academy of Mining and Metallurgy, Kraków, in 1987 and 1992, respectively.

In December 1999, the President of Poland granted him the title of Professor. He is currently the Head of the Department of Multimedia Systems, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology. He has authorized more than 600 scientific articles in international journals and conference proceedings. He has led more than 30 research and development projects funded by the Polish Government and participated in seven European projects. He is the author of 30 Polish patents and five international patents. He has extensive experience in soft computing algorithms and sound and image processing for surveillance, human–computer interaction, telemedicine, and education applications.