An Evaluation of Gaze-Based Person Identification with Different Stimuli

Wolfgang Fuhl
wolfgang.fuhl@uni-tuebingen.de
University Tübingen
Tübingen, Baden-Württemberg
Germany

Dennis Grueneberg
dennis.grueneberg@student.uni-tuebingen.de
University Tübingen
Tübingen, Baden-Württemberg
Germany

Abdullah Yalvac
abdullah.yalvac@student.uni-tuebingen.de
University Tübingen
Tübingen, Baden-Württemberg
Germany

ABSTRACT

In eye tracking, the most promising approaches for person identification are based on the iris. Therefore, the iris of a subject is extracted and compared against a database of stored iris templates or directly classified by a deep neural network. While this approach is robust and has found its way into many practical applications like security access devices, it has some limitations, since it is possible to fake the iris. This can be done either by presenting a rendered image or by using a contact lens. Our research focuses on the gaze behavior of a subject for person identification. Therefore, we present different stimuli as well as moving dots. In past research, only a static dot has been used so far. We show that, especially, the combination of different stimuli is a promising approach and could be used as an additional security layer in combination with iris recognition.

CCS CONCEPTS

• **Computing methodologies** → *Classification and regression trees*; **Neural networks**; **Learning latent representations**; *Regularization*; *Feature selection.*

KEYWORDS

Eye Tracking, Person Identification, Behaviour Analysis, Different Stimuli

ACM Reference Format:
Wolfgang Fuhl, Dennis Grueneberg, and Abdullah Yalvac. 2024. An Evaluation of Gaze-Based Person Identification with Different Stimuli. In *2024 Symposium on Eye Tracking Research and Applications (ETRA '24), June 04–07, 2024, Glasgow, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3649902.3655651

1 INTRODUCTION

User identification is a vital process in various industrial domains. It serves to protect data and access to networks or premises in its conventional form. However, the demand for user identification has increased, especially in the online environment, which involves both personalized advertising and product placement, as well as online banking or external access to corporate networks [Das 2009; Vosecky et al. 2009; Zafarani et al. 2015]. For security-critical applications, such as online banking [Hammood et al. 2020] or external access to corporate networks [Pevný et al. 2012], user IDs and passwords are the standard methods. These applications also employ additional security measures, such as generated PINs or SMS prompts, when using security-critical functionalities [Das 2009]. In online advertising and product placement, companies aim to identify a user without accessing sensitive personal data [Fuhl et al. 2021]. This is achieved in the modern world by cookies, which require the user's consent, or stateless approaches that only use browser statistics. However, this method has a drawback: the statistics can effectively identify a computer, but not a specific user in the case of multiple users on the same computer [Fuhl et al. 2021; Nnamoko et al. 2022; Yang 2010]. The password and user recognition procedure also has disadvantages. For instance, an attacker can gain access if they know the identification and password of a user.

In this work, we examine new approaches to measuring personal visual behaviors. The underlying idea is that a user can be identified by their gaze signals or visual behavior, which have been proven to be distinctive. We want to explore if we can improve the person's identification based on different or moving stimuli. Therefore, we conducted a battery of experiments that ranged from a single dot to an image and also included moving dots as well as dots with different sizes. We also propose to use transfer learning in combination with a metric loss function for this purpose, since it worked best in our evaluations.

ETRA '24, June 04–07, 2024, Glasgow, United Kingdom

ACM ISBN 979-8-4007-0607-3/24/06
https://doi.org/10.1145/3649902.3655651

2 RELATED WORK

The most prominent challenge in this domain is the Eye Movements Verification and Identification Competition (EMVIC) [Kasprowski et al. 2012]. Due to the space limitation, we will only go over the state-of-the-art approaches, which we used in our evaluation, and mention other approaches briefly. The first scanpath classification approaches are based on statistics [Goldberg and Helfman 2010] computed on eye movements [Fuhl et al. 2023]. Here, fixation duration, saccade velocity, and so on are used to compute statistical moments like the mean and standard deviation, which are used for classification. An extension of those approaches is based on areas of interest that are used to compute statistics additionally [Fuhl et al. 2018b; Kübler et al. 2017]. Other features used for scanpath classification are the heatmap (HEAT) [Fuhl et al. 2021] or the histogram of oriented velocities (HOV) [Fuhl et al. 2018a]. There are also more modern deep learning-based approaches like DeepEyedentification [Jäger et al. 2020] which uses the sequence of angular velocities as input. Other approaches are Multimatch [Wagner et al. 2019], Ferns on saccade angles [Fuhl et al. 2019b], different input representations [Fuhl et al. 2019a] and many more. We selected only a subset of all approaches, since we are only interested in the effectiveness of the stimuli on person identification itself.

3 METHODS AND RESULTS

For our approach we used a ResNet-18 [He et al. 2016] which was trained on the datasets Gaze [Dorr et al. 2010], WherePeopleLook [Judd et al. 2009], DOVES [Rajashekar et al. 2009], and ETRAChalleng [McCamy et al. 2014] with the metric loss function from [Hoffer and Ailon 2015]. The learning task was to minimize the distance of two samples from the same recording and to maximize the distance of samples from different recordings. After training of the model for approximately one week on an NVIDIA 3090 RTX we used it as a feature extractor in a transfer learning fashion [Tan et al. 2018]. This means, that we replaced the last two fully connected stages and added a softmax loss function for classification. For evaluation, we trained the last two fully connected layers on the given training set with the features of ResNet-18.

For our evaluation, we conducted 2 studies with 7 different stimuli. The first study had 55 subjects, and the second study had 15 subjects. The three tasks in the first study where DOT, which is a single dot as a stimulus; Multi DOT, which is multiple dots on an image as stimuli; and Image, which is a normal image as a stimulus. For the second study, we recorded 5 tasks. The first task was again DOT, the second task was moving dots, the third task was moving dots with different speeds, the fourth task was differently sized dots, and the fifth task was differently sized dots with different speeds. The task for each dataset was to classify which person belongs to the scanpath.

For our evaluation, we did a two-fold cross-validation. This means that we split the data from each dataset randomly into two folds, where each fold contains 50% of the recordings. Each fold is once the training and once the testing data. This way, we can report the F1 score over the entire datasets from our studies. For the state-of-the-art approaches, which are not based on deep neural networks, we performed a grid search for the parameters of the machine learning method's support vector machine (SVM), two-layered neural network, and ensemble of decision trees and reported only the best results. In addition, we performed a grid search for the optimal parameters for the features HEAT and HOV. This was always done using 20% of the training data as a validation set. For the deep learning approaches, we searched for the optimal learning rate, initial learning rate, and weight decay using a grid search. The approach Deep Eyedentification [Jäger et al. 2020] received the same pretraining as our approach and we rebulid the model.

Table 1 shows the results of our evaluation. Note that Deep Eyedentification [Jäger et al. 2020] received the same pretraining as our approach. As can be seen for the first experiment, the image seems to work best for user identification, but the multi-dot approach also leads to an improvement compared to the single dot used in the EMVIC challenge [Kasprowski et al. 2012]. The approach with moving dots and differently sized dots improves the classification, but it seems, that the improvement is not as good as for the image or the multi-dot approach. Those recordings are from the second study and difficult to compare directly. The most interesting part is that the combination of size and moving dots brings the best improvement.

Table 1: We evaluated different approaches for scanpath classification namely HEAT [Fuhl et al. 2021], HOV [Fuhl et al. 2018a], Statistics [Goldberg and Helfman 2010] on eye movements (EM) [Fuhl et al. 2023], DeepEyedentification [Jäger et al. 2020] and our transfer learning approach. As metric we used the F1 score and selected always the best performing classifier as well as the best performing parameters for the state-of-the-art approaches which where determined with a grid search.

Subjects	Method	Stimulus	F1 Score
55	HEAT	DOT	0.19
	HOV		0.31
	Statistics on EM		0.23
	DeepEyedentification		0.54
	Transfer (Proposed)		0.61
	HEAT	Multi DOT	0.27
	HOV		0.26
	Statistics on EM		0.24
	DeepEyedentification		0.65
	Transfer (Proposed)		0.71
	HEAT	Image	0.41
	HOV		0.36
	Statistics on EM		0.33
	DeepEyedentification		0.68
	Transfer (Proposed)		0.74
15	HEAT	DOT	0.20
	HOV		0.29
	Statistics on EM		0.24
	DeepEyedentification		0.51
	Transfer (Proposed)		0.53
	HEAT	Moving	0.41
	HOV		0.46
	Statistics on EM		0.34
	DeepEyedentification		0.57
	Transfer (Proposed)		0.62
	HEAT	Fast	0.49
	HOV		0.53
	Statistics on EM		0.48
	DeepEyedentification		0.66
	Transfer (Proposed)		0.68
	HEAT	Size	0.34
	HOV		0.45
	Statistics on EM		0.37
	DeepEyedentification		0.63
	Transfer (Proposed)		0.67
	HEAT	Size+Fast	0.52
	HOV		0.67
	Statistics on EM		0.56
	DeepEyedentification		0.82
	Transfer (Proposed)		0.89

4 CONCLUSIONS AND FUTURE WORK

Since the combination of moving dots with different sizes got the best result, we will further investigate additional combinations and also increase the amount of participants. It also has to be noted, that the classification is not perfect, and therefore, a combination with other identification approaches has to be evaluated too. One example here could be an online login where the username and password were stolen by a third party. During the usage of the online content, we want to try to identify if it is really the correct user or not. This could be an additional security layer for online banking, for example, and also requires longer recordings. Additionally, we want to conduct a study on browser-based user identification, if multiple users use the same computer or smartphone. This could help to detect if a smartphone was stolen or to automatically adapt the environment to the user without a login.

ACKNOWLEDGMENTS

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 508330921

REFERENCES

Manik Lal Das. 2009. Two-factor user authentication in wireless sensor networks. *IEEE transactions on wireless communications* 8, 3 (2009), 1086–1090.

Michael Dorr, Thomas Martinetz, Karl R Gegenfurtner, and Erhardt Barth. 2010. Variability of eye movements when viewing dynamic natural scenes. *Journal of vision* 10, 10 (2010), 28–28.

Wolfgang Fuhl, Efe Bozkir, Benedikt Hosp, Nora Castner, David Geisler, Thiago C Santini, and Enkelejda Kasneci. 2019a. Encodji: encoding gaze data into emoji space for an amusing scanpath classification approach. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. 1–4.

Wolfgang Fuhl, Nora Castner, and Enkelejda Kasneci. 2018a. Histogram of oriented velocities for eye movement detection. In *Proceedings of the Workshop on Modeling Cognitive Processes from Multimodal Data*. 1–6.

Wolfgang Fuhl, Nora Castner, Thomas Kübler, Alexander Lotz, Wolfgang Rosenstiel, and Enkelejda Kasneci. 2019b. Ferns for area of interest free scanpath classification. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. 1–5.

Wolfgang Fuhl, Anne Herrmann-Werner, and Kay Nieselt. 2023. The Tiny Eye Movement Transformer. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*. 1–2.

Wolfgang Fuhl, Thomas C Kübler, Thiago Santini, and Enkelejda Kasneci. 2018b. Automatic Generation of Saliency-based Areas of Interest for the Visualization and Analysis of Eye-tracking Data.. In *VMV*. 47–54.

Wolfgang Fuhl, Nikolai Sanamrad, and Enkelejda Kasneci. 2021. The gaze and mouse signal as additional source for user fingerprints in browser applications. *arXiv preprint arXiv:2101.03793* (2021).

Joseph H Goldberg and Jonathan I Helfman. 2010. Visual scanpath representation. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. 203–210.

Waleed A Hammood, Ruzaini Abdullah, Omar A Hammood, Salwana Mohamad Asmara, Mohammed A Al-Sharafi, and Ali Muttaleb Hasan. 2020. A review of user authentication model for online banking system based on mobile IMEI number. In *IOP Conference Series: Materials Science and Engineering*, Vol. 769. IOP Publishing, 012061.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

Elad Hoffer and Nir Ailon. 2015. Deep metric learning using triplet network. In *Similarity-Based Pattern Recognition: Third International Workshop, SIMBAD 2015, Copenhagen, Denmark, October 12-14, 2015. Proceedings 3*. Springer, 84–92.

Lena A Jäger, Silvia Makowski, Paul Prasse, Sascha Liehr, Maximilian Seidler, and Tobias Scheffer. 2020. Deep Eyedentification: Biometric identification using micro-movements of the eye. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2019, Würzburg, Germany, September 16–20, 2019, Proceedings, Part II*. Springer, 299–314.

Tilke Judd, Krista Ehinger, Frédo Durand, and Antonio Torralba. 2009. Learning to predict where humans look. In *2009 IEEE 12th international conference on computer vision*. IEEE, 2106–2113.

Paweł Kasprowski, Oleg V Komogortsev, and Alex Karpov. 2012. First eye movement verification and identification competition at BTAS 2012. In *2012 IEEE fifth international conference on biometrics: theory, applications and systems (BTAS)*. IEEE, 195–202.

Thomas C Kübler, Colleen Rothe, Ulrich Schiefer, Wolfgang Rosenstiel, and Enkelejda Kasneci. 2017. SubsMatch 2.0: Scanpath comparison and classification based on subsequence frequencies. *Behavior research methods* 49 (2017), 1048–1064.

Michael B McCamy, Jorge Otero-Millan, Leandro Luigi Di Stasi, Stephen L Macknik, and Susana Martinez-Conde. 2014. Highly informative natural scene regions increase microsaccade production during visual scanning. *Journal of neuroscience* 34, 8 (2014), 2956–2966.

Nonso Nnamoko, Joseph Barrowclough, Mark Liptrott, and Ioannis Korkontzelos. 2022. A behaviour biometrics dataset for user identification and authentication. *Data in Brief* 45 (2022), 108728.

Tomáš Pevnỳ, Martin Rehák, and Martin Grill. 2012. Identifying suspicious users in corporate networks. In *Proceedings of workshop on information forensics and security*. 1–6.

Umesh Rajashekar, Lawrence K Cormack, Alan C Bovik, and Ian van der Linde. 2009. DOVES: a database of visual eye movements. *Spatial Vision* 22, 2 (2009), 161–177.

Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. 2018. A survey on deep transfer learning. In *Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27*. Springer, 270–279.

Jan Vosecky, Dan Hong, and Vincent Y Shen. 2009. User identification across multiple social networks. In *2009 first international conference on networked digital technologies*. IEEE, 360–365.

Adina S Wagner, Yaroslav O Halchenko, and Michael Hanke. 2019. multimatch-gaze: The MultiMatch algorithm for gaze path comparison in Python. *Journal of Open Source Software* 4, 40 (2019), 1525.

Yinghui Catherine Yang. 2010. Web user behavioral profiling for user identification. *Decision Support Systems* 49, 3 (2010), 261–271.

Reza Zafarani, Lei Tang, and Huan Liu. 2015. User identification across social media. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 10, 2 (2015), 1–30.