Poster: Electrooculography Dataset for Reading Detection in the Wild

Shoya Ishimaru
German Research Center for Artificial Intelligence (DFKI)
Shoya.Ishimaru@dfki.de

Takanori Maruichi
Osaka Prefecture University
maruichi@m.cs.osakafu-u.ac.jp

Manuel Landsmann
Osaka Prefecture University
m_landsman09@gmx.de

Koichi Kise
Osaka Prefecture University
kise@cs.osakafu-u.ac.jp

Andreas Dengel
German Research Center for Artificial Intelligence (DFKI)
Andreas.Dengel@dfki.de

ABSTRACT

Because of the diversity of document layouts and reading styles, detecting reading activities in real life is a challenging task compared to the detection in the laboratory setting. For contributing to the implementation of robust reading detection algorithms, we introduce a dataset which contains 220 hours of sensor signals from JINS MEME electrooculography glasses and corresponding ground truth activity labels. As a baseline study, we propose a statistical feature based reading detection approach and evaluate it on the dataset.

CCS CONCEPTS

• **Computing methodologies → Activity recognition and understanding**.

KEYWORDS

Eyewear computing; Cognitive activity recognition; In-the-wild study; Electrooculography; Internal measurement unit

ACM Reference Format:
Shoya Ishimaru, Takanori Maruichi, Manuel Landsmann, Koichi Kise, and Andreas Dengel. 2019. Poster: Electrooculography Dataset for Reading Detection in the Wild. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), September 9–13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3341162.3343812

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6869-8/19/09…$15.00
https://doi.org/10.1145/3341162.3343812

1 INTRODUCTION

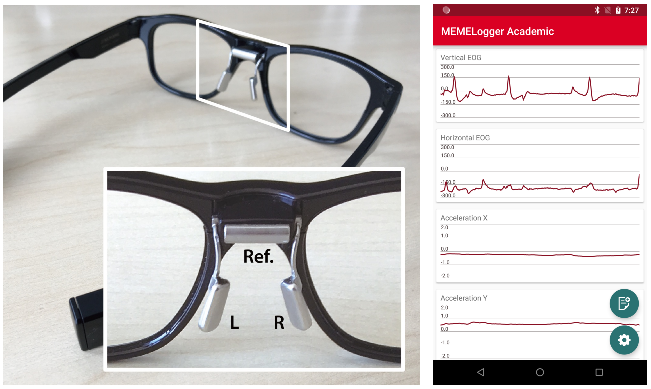
Just as our bodies consist of what we eat, our minds are shaped by the information we obtain. In particular, written text is one of the most important information sources in our lives. Therefore, understanding and improving daily reading habits provides several cognitive benefits, including increased vocabulary and logical thinking [4].

For instance, as people are encouraged to be physically fit by monitoring step counts, tracking the number of words they read in a day has potential to motivate them to read more. The idea of estimating the number of read words has been implemented as *Wordometer* by using mobile eye tracking glasses [11], electrooculography glasses [6], and a remote eye tracker [1]. However, these implementations were evaluated only in the laboratory setting. A finding from our previous in-the-wild study is that readings in real life occur in a variety of situations, and an estimator trained by well-mannered reading data could not cover such natural readings enough [5].

On the basis of several requests relating to the data, we conduct a large-scale recording again solving some issues that appeared in the previous work. We utilize JINS MEME (see Figure 1) as a sensing device. It is equipped with a three-electrode electrooculography (EOG) sensor which measures eye movements and a six-axis internal measurement unit (IMU) which measures head and body movements. A form factor and a long-life battery of the device are designed for in-the-wild studies. A wide variety of approaches have been proposed on it, including recognizing human activities [7], facial actions [13, 14], gaze gestures [15], and internal states [18].

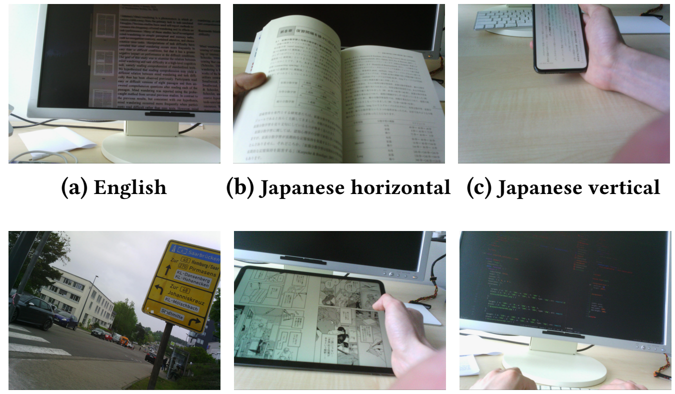
The contributions of our work are two-fold: the EOG dataset and our reading detection method. The dataset, software for the data recording, and sample codes are available on our project repository[1]. This study was conducted with the permission of the Research Ethics Committee of the Graduate School of Engineering, Osaka Prefecture University.

[1] https://github.com/shoya140/ubicomp2019-eog-dataset/



(a) Device overview **(b) Sensor signals**

Figure 1: JINS MEME Electrooculography Glasses and sensor signals visualized on our Android application



(a) English **(b) Japanese horizontal** **(c) Japanese vertical**

(d) Not reading

Figure 2: Examples of pictures taken by Narrative Clip while reading or not reading in the wild

2 BACKGROUND AND RELATED WORK

As summarized in the survey by Lara and Labrador, many physical activities (e.g., walking, running, cycling, sleeping) can be recognized by motion sensors on the body or a smartphone [12]. On the other hand, the recognition of reading activity is considerably restricted because dynamic body movements could not be observed while reading.

One of the interesting approaches this problem is to measure eye movements. Steil and Bulling detected daily activities including readings by using mobile eye tracking glasses [17]. Srivastava *et al.* recognized activities on a computer with several involving reading [16] and Kelton *et al.* classified reading/skimming [9] by using a remote eye tracker attached to a display. These sensors are good at measuring gaze events (e.g., fixations and saccades). But we assume that characteristic eye movements (e.g., frequent horizontal/vertical saccades) have enough potential for the detection.

EOG measures such eye movements from the corneo-retinal standing potential that exists between the front and the back of an eyeball. Traditional setups use four electrodes around an eye [2, 3]. Recent sensing devices have three electrodes on each nose pad and the forehead to identify vertical/horizontal eye movements and blinks [8, 10].

3 DATA RECORDING

We utilized JINS MEME ES_R[2] and Android Nexus 5X for the data recording. We developed an application[3] which enables a user to easily check sensor signals and start/stop recording data. Figure 1 shows a screen capture visualizing three blinks and some vertical and horizontal eye movements. The sampling frequency was set to 100 Hz. Note that we modified kernel codes and built a customized Android operating system for a stable Bluetooth Low Energy connection[4].

To support a ground truth labeling task, we provided Narrative Clip[5], a small life-logging camera which can be attached to clothes and takes a picture every 30 seconds. Examples of the pictures are shown in Figure 2. At the end of each day, every activity in the pictures was annotated by the user into four categories: *reading in English* (EN), *reading in Japanese written horizontally* (JH), *reading in Japanese written vertically* (JV), and *not reading* (NR). The reason we prepared such three labels for reading is that eye movements should be affected by the language (native or non-native) and writing style (vertical or horizontal). Even if characters appeared in a picture, an activity which does not require frequent line breaks was not categorized as reading in this study (e.g., looking at a signboard while walking, reading a comic, writing codes). In order to protect privacies, we collected only activity labels and no pictures.

We recruited ten Japanese college students for two days of data recording. To ensure collecting a minimum amount of valid reading behaviors, we asked them to try to perform each activity (EN, JH, and JV) for at least one hour every day. We held an initial briefing to explain the procedures and the usage of the devices. In addition, all instructions were written in a document and shared with participants. Participants who completed the tasks received 10,000 JPY.

Table 1 shows an overview of the recording durations. In total, our dataset contains 23 hours of English reading, 25 hours of Japanese horizontal reading, 25 hours of Japanese vertical reading, and 146 hours of other activities.

[2]https://jins-meme.com/en/researchers/
[3]https://memelogger.shoya.io/
[4]https://shoya.io/posts/meme-android-connection/
[5]http://getnarrative.com/

Table 1: Recording durations [minutes]

(a) Day 1

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
EN	54	73	74	67	63	82	91	62	59	49
JH	97	70	101	72	73	86	127	63	75	90
JV	74	101	71	60	83	73	65	66	113	65
NR	474	447	490	499	476	379	204	538	476	494

(b) Day 2

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
EN	100	96	61	61	74	115	67	55	58	62
JH	71	74	117	53	64	60	53	58	58	68
JV	89	75	67	60	116	72	74	73	60	75
NR	466	499	429	567	283	409	358	496	307	509

Table 2: Classification accuracies [%]

Condition	User-independent	User-dependent
EN vs JH vs JV vs NR	32	34
(EN + JH) vs JV vs NR	46	45
(EN + JH + JV) vs NR	68	69
EN vs NR	67	66
JR vs NR	68	68
JV vs NR	74	69
EN vs JH vs JV	36	46

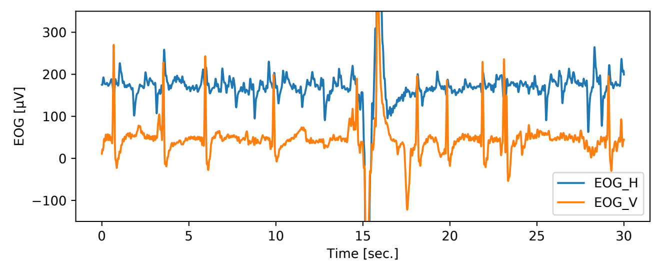


Figure 3: EOG signals while reading an English document

4 READING DETECTION: BASELINE STUDY

We propose a detection method as a baseline. Characteristics of the dataset found from an evaluation are also described.

Approach

Ten features are calculated from one sample (30 seconds window of a data stream): means and variances of the two EOG axes, variances of three accelerometer axes, and variances of three gyroscope axes. Then we utilize a Support Vector Machine to classify the samples. The radial basis function kernel with hyper parameters $C = 1$ and $gamma = 0.125$ were selected experimentally and used for the classifier.

Experimental Conditions

We evaluated the proposed reading detection approach with a user-independent and a user-dependent training. For a user-independent training, we separated data into training and testing in a a *leave-one-participant-out cross-validation* manner. Data of nine participants were used for training a classifier and data of the remaining one participant were used for testing. For a user-dependent training, the classifier was adapted to each participant by his/her data with *leave-one-day-out cross-validation*. One-day was used for training, and another day was used for testing. Since the dataset is unbalanced and our purpose is to detect a minor class, we applied under-sampling for the evaluation.

Results and Discussion

The classification results are shown in Table 2. Chance rates are 50%, 33%, and 25% for two, three, and four-class classifications, respectively. One of the interesting findings from the results is that head and eye movements while reading JV texts are relatively distinctive compared to EN and JH. This may be because the Japanese vertical writing style is often used for well-formatted texts such as novels and newspapers. Reading EN and JH texts can be often performed with other activities, for example, writing and browsing.

As shown in Figure 3, eye blinks (quick up and down eye movements measured in EOG_V) and forward and backward saccades to the reading direction have successfully appeared during a static condition. But dynamic movements of the head or the glasses cause artifacts, which should be removed before calculating features.

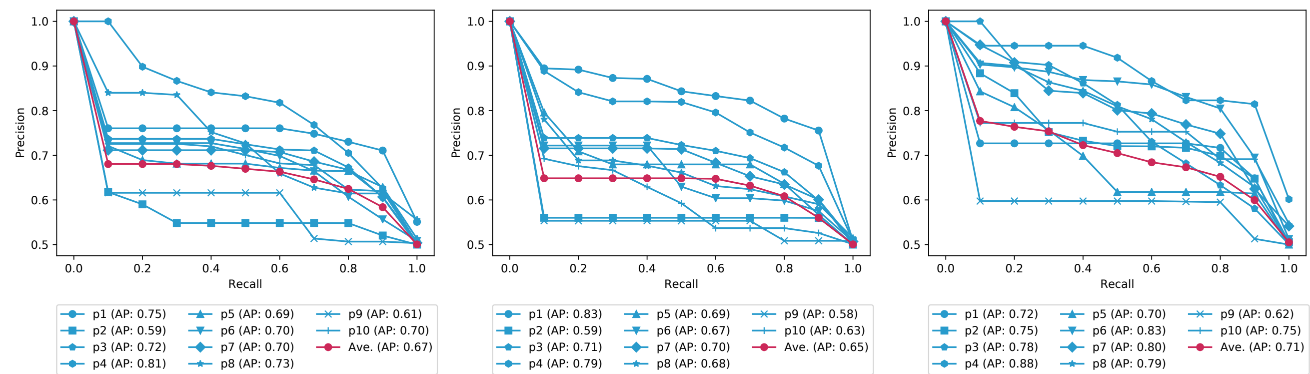
Figure 4 represents 11-point interpolated recall-precision graphs for each activity detection task for each participant. Although there is not much difference between classification accuracies of a user-independent and a user-dependant training, the performances are highly distributed by participants.

5 CONCLUSION

This paper presented a dataset which contains sensor signals of JINS MEME electrooculography glasses and corresponding ground truth activity labels. We demonstrated how challenging the natural reading detection is by evaluating our statistical feature based approach.

ACKNOWLEDGMENTS

This work is supported by JST CREST and JSPS KAKENHI (Grant Numbers: JPMJCR16E1, 17K12728).



(a) English (EN) **(b) Japanese horizontal (JH)** **(c) Japanese vertical (JV)**

Figure 4: Recall-precision curves and average precisions of reading detection with user-independent training

REFERENCES

[1] Olivier Augereau, Charles Lima Sanches, Koichi Kise, and Kai Kunze. 2018. Wordometer Systems for Everyday Life. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 123:1–123:21.

[2] Malcolm Brown, Michael Marmor, Eberhard Zrenner, Mitchell Brigell, Michael Bach, et al. 2006. ISCEV standard for clinical electro-oculography (EOG) 2006. *Documenta Ophthalmologica* 113, 3 (2006), 205–212.

[3] Andreas Bulling, Jamie A Ward, Hans Gellersen, and Gerhard Troster. 2011. Eye movement analysis for activity recognition using electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 4 (2011), 741–753.

[4] Anne E Cunningham and Keith E Stanovich. 1998. What reading does for the mind. *American Educator* 22 (1998), 8–17.

[5] Shoya Ishimaru, Kensuke Hoshika, Kai Kunze, Koichi Kise, and Andreas Dengel. 2017. Towards reading trackers in the wild: detecting reading activities by EOG glasses and deep neural networks. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 704–711.

[6] Shoya Ishimaru, Kai Kunze, Koichi Kise, and Andreas Dengel. 2016. The wordometer 2.0: estimating the number of words you read in real life using commercial EOG glasses. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 293–296.

[7] Shoya Ishimaru, Kai Kunze, Yuji Uema, Koichi Kise, Masahiko Inami, and Katsuma Tanaka. 2014. Smarter eyewear: using commercial EOG glasses for activity recognition. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 239–242.

[8] Shinichiro Kanoh, Susumu Ichi-nohe, Shunsuke Shioya, Kazutaka Inoue, and Ryota Kawashima. 2015. Development of an eyewear to measure eye and body movements. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2267–2270.

[9] Conor Kelton, Zijun Wei, Seoyoung Ahn, Aruna Balasubramanian, Samir R Das, Dimitris Samaras, and Gregory Zelinsky. 2019. Reading detection in real-time. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ACM, 43:1–43:5.

[10] Nataliya Kosmyna, Utkarsh Sarawgi, and Pattie Maes. 2018. AttentivU: Evaluating the Feasibility of Biofeedback Glasses to Monitor and Improve Attention. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, 999–1005.

[11] Kai Kunze, Hitoshi Kawaichi, Kazuyo Yoshimura, and Koichi Kise. 2013. The Wordometer–Estimating the Number of Words Read Using Document Image Retrieval and Mobile Eye Tracking. In *Proceedings of the 12th International Conference on Document Analysis and Recognition*. 25–29.

[12] Oscar D Lara and Miguel A Labrador. 2013. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials* 15, 3 (2013), 1192–1209.

[13] Juyoung Lee, Hui-Shyong Yeo, Murtaza Dhuliawala, Jedidiah Akano, Junichi Shimizu, Thad Starner, Aaron Quigley, Woontack Woo, and Kai Kunze. 2017. Itchy nose: discreet gesture interaction using EOG sensors in smart eyewear. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM, 94–97.

[14] Soha Rostaminia, Alexander Lamson, Subhransu Maji, Tauhidur Rahman, and Deepak Ganesan. 2019. W!NCE: Unobtrusive Sensing of Upper Facial Action Units with EOG-based Eyewear. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 23:1–23:26.

[15] Junichi Shimizu, Thad Starner, Juyoung Lee, Woontack Woo, Murtaza Dhuliawala, Kai Kunze, and Andreas Bulling. 2016. Solar system-smooth pursuit interactions using EOG glasses. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 369–372.

[16] Namrata Srivastava, Joshua Newn, and Eduardo Velloso. 2018. Combining Low and Mid-Level Gaze Features for Desktop Activity Recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 189:1–189:27.

[17] Julian Steil and Andreas Bulling. 2015. Discovery of everyday human activities from long-term visual behaviour using topic models. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 75–85.

[18] Benjamin Tag, Andrew W Vargo, Aman Gupta, George Chernyshov, Kai Kunze, and Tilman Dingler. 2019. Continuous Alertness Assessments: Using EOG Glasses to Unobtrusively Monitor Fatigue Levels In-The-Wild. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 464:1–464:12.