Correction to: “Using machine learning to detect events in eye-tracking data”

Raimondas Zemblys¹ · Diederick C. Niehorster² · Kenneth Holmqvist^{3,4,5,6}

Published online: 24 September 2018
© Psychonomic Society, Inc. 2018

Correction to: Behav Res

DOI 10.3758/s13428-017-0860-3

It has come to our attention that the section “Post-processing: Labeling final events” on page 167 of “Using Machine Learning to Detect Events in Eye-Tracking Data” (Zemblys, Niehorster, Komogortsev, & Holmqvist, 2018) contains an erroneous description of the process by which post-processing was performed. Specifically, the sentence “Removal of a saccade, PSOs or fixations means that the sample is marked as unclassified, a fourth class.” should be replaced by “Removal of a saccade, PSO, or fixation means that the probability of a sample belonging to that class is downvoted—that is, set to 0. Removal of an event thus, in effect, entails labeling the affected samples as the next most likely event.”

Furthermore, we have come to realize that a more thorough description of the post-processing approach is needed so that the reader can correctly appreciate what happens. We therefore add the following paragraph to the end of the post-processing section that this erratum concerns:

It is important to realize that in a machine-learning context, it is natural to perform post-processing by downvoting probabilities of samples that violate heuristic rules. This probabilistic approach, using the heuristic rules listed above, was selected so as to retain classification even when some of the events detected as most likely did not meet predefined criteria. For example, just removing short saccades would result in a larger number of short undefined events. However, if the probabilities of samples marked as saccades that are found to be too short are downvoted, these samples can then be made to belong to the next most likely class. In our case, the only other classes are the fixation and PSO classes, but in future uses of this approach, a user might want to train the algorithm with, for instance, smooth pursuit. The “too short saccade” samples could then be downvoted to become either fixation or pursuit samples, or any other event, depending on which of these has the next highest probability.

In addition, we have also noticed an inaccurate reference to one of the biometric datasets that was used in the study. In the article we stated that we used two databases—the Eye Movement Biometric Database, version 1 (Komogortsev, 2011), and Eye Movement Biometric Database, version 2 (Komogortsev, 2016). However, it turned out that instead of version 1, another database was actually used that is not publicly available. The reference to version 1 should therefore be removed.

Finally, it has come to our attention (Friedman, Rigas, Abdulin, & Komogortsev, 2018) that on unseen data (not belonging to the training or validation set when the identification by random forest [IRF] algorithm was constructed), some of the events output by the IRF event detector were erroneously not removed or reclassified as other events, despite the fact that they violated the heuristic post-processing rules listed in the “post-processing” section of the article. This may have led to a slight

The online version of the original article can be found at <https://doi.org/10.3758/s13428-017-0860-3>

✉ Raimondas Zemblys
r.zemblys@tf.su.lt

¹ Siauliai University, Siauliai, Lithuania

² Lund University Humanities Laboratory and Department of Psychology, Lund University, Lund, Sweden

³ Department of Psychology, Regensburg University, Regensburg, Germany

⁴ Universiteit van die Vrystaat, Bloemfontein, South Africa

⁵ UPSET, NWU Vaal, Vanderbijlpark, South Africa

⁶ Faculty of Arts, Masaryk University, Brno, Czech Republic

increase in the number of erroneous events included in the evaluation of IRF's performance reported in Zemblys et al. (2018), dragging down IRF's performance slightly. In response to these discoveries, we have updated the post-processing code to minimize such erroneous behavior. In addition, we have implemented a routine that checks the output of the probabilistic post-processing steps and that, if required, as a final step performs "hard" post-processing with deterministic rules—that is, completely removing offending events. We have furthermore collected and hand-coded a new eye-tracking dataset and replicated the main results of the original study with this updated version of the IRF algorithm. The details of our replication, as well as all code and input data needed to use our algorithm, are available at <https://github.com/r-zemblys/irf>. The trained model used in this replication is furthermore available at <https://doi.org/10.5281/zenodo.1343920>.

References

- Friedman, L., Rigas, I., Abdulin, E., & Komogortsev, O. V. (2018). A novel evaluation of two related and two independent algorithms for eye movement classification during reading. *Behavior Research Methods*. Advance online publication. <https://doi.org/10.3758/s13428-018-1050-7>
- Komogortsev, O. V. (2011). Eye Movement Biometric Database (Version 1). San Marcos, TX: Texas State University. Retrieved from https://userweb.cs.txstate.edu/~ok11/emdb_v1.html
- Komogortsev, O. V. (2016). Eye Movement Biometric Database (Version 2). San Marcos, TX: Texas State University. Retrieved from https://userweb.cs.txstate.edu/~ok11/emdb_v2.html
- Zemblys, R., Niehorster, D. C., Komogortsev, O., & Holmqvist, K. (2018). Using machine learning to detect events in eye-tracking data. *Behavior Research Methods*, 50, 160–181. <https://doi.org/10.3758/s13428-017-0860-3>