Eye-movement event detection meets machine learning

Conference Paper · November 2016

1 author:

Raimondas Zemblys
Smart Eye

Eye-Movement Event Detection Meets Machine Learning

R. Zemblys
Department of Engineering, Siauliai University, Lithuania
E-mail: r.zemblys@tf.su.lt

Introduction. In eye-tracking research, eye-movement event detection plays a major role. The vast majority of all analyses of eye movement data are performed not on raw data, but on events like fixations, saccades, smooth pursuits, etc. Today, event detection is almost exclusively done using computers by applying a detection algorithm to the raw gaze data. Therefore the results produced by these algorithms have a dramatic impact on higher-level analyses and measures used in eye movement research.

For a long time there were two broad types of event detection algorithms – dispersion and velocity threshold based. The first one (I-DT) finds fixations by a spatial criterion on the size of the region in which the recorded gaze points can move and implicitly detect saccades, while other class of algorithms (I-VT) look for saccades by a criterion of the minimal velocity that the eye has to move [101]. In fact, commercial event detection algorithms from major manufacturers like SMI, Tobii and SR-Research even today use the same basic principles with a little bit of heuristic post-processing.

A major drawback of current event detection methods is that user is left with number of parameters which have to be adjusted based on eye movement data quality. Even very recently developed state-of-the-art algorithms, which enable to detect a broader range of events (e.g. post-saccadic oscillations [6, 4]), are not easy to use and need an expert to tweak various parameters. Another drawback is that these new algorithms are designed to solve a specific problem (like smooth pursuit detection [5] or noise resilience [3]) or work on specific type of data. However, new algorithms are nowhere close to at least human level performance [2].

Recent advances in computer science showed that machine learning was the key in reaching and exceeding human level performance in complex tasks such as image classification, natural language processing or robotics. From a stand point of machine learning, eye-movement event detection is a very simple task: there are only 5 basic events and no more than 20-25 other events that exist in the psychological and neurological eye-movement research. Compared to e.g. deep learning models which are able to distinguish between thousands of classes and subclasses of objects in image recognition [7], eye-movement event detection is a good candidate to also meeting machine learning.

This paper presents a comparison of performance of 10 machine learning algorithms in eye-movement event detection task. The goal was to build a universal algorithm, which could work with any type of the eye-tracking data.

Method. Monocular eye movement data used in this study was recorded at Humanities Laboratory, Lund University (Sweden), using Eyelink1000 eye-tracker running at 1000 Hz. A human expert with 9 years of eye-tracking experience manually tagged raw data as belonging to one of 4 classes: fixations, saccades, post-saccadic oscillations and noise (including blinks). To simulate recordings from other eye-trackers, data was augmented by systematically resampling it to sampling rates of commercially available trackers (30 – 1250 Hz). Further, white Gaussian noise was added to the resampled dataset to cover the range from 0.01 to 7 degrees RMS.

Next, for each data sample, a 13-dimensional feature vector was extracted. Features used here are basically the same as the ones in commercial and state-of-the-art event detection algorithms: sampling rate, half of perimeter of bounding box around the data (I-DT), difference between the average/median gaze position preceding and succeeding the sample (used by Tobii), velocity, acceleration (all calculated over 100ms window), Rayleightest (calculated over 22ms window) [101]. In addition, features that describe data quality were added: root-mean-square (RMS), standard deviation, bivariate contour ellipse area (BCEA). The rest of the features are similar to ones used by Tobii, but instead of differences in position, difference between noise measures is calculated. All features were normalized by removing the mean and scaling to unit variance.

To ensure the proper estimation of out-of-the-bag error, all data was randomly split into 2 subsets – training (~6.4 million samples, 75%) and testing (~2.1 million samples, 25%). Classification error was calculated using Cohen's Kappa (K) – a measure which measures inter-rater agreement for categorical data. Cohen's Kappa is a number between -1 and 1, where 1 means perfect agreement and 0 means no agreement between the raters other than what would be expected by chance. The choice of this accuracy measure is motivated by a fact that it was previously used in the eye tracking field to assess the performance of newly developed event detection algorithms [4, 5] and to compare algorithms to manual coding [2] thus allowing to directly compare the results. Moreover, sample level eye-tracking event data is highly unbalanced, as most of the data belong to one class – in this case around 90% of samples belong to the fixations. As a result, if using measures like sensitivity, specificity, F1 score or similar, even majority class model (where all samples are predicted to belong to the majority class) would achieve a considerably high score.

Training data was fed to 10 machine learning algorithms: K nearest Neighbors (K=3), Linear (LDA) and Quadratic (QDA) Discriminant Analysis, Naive Bayes, SVM with linear and RBF kernels, Decision Tree and Random Forest (32 trees), Ada Boost (with 64 Decision Tree estimators) and Gradient Boosting (with 128 estimators) classifiers. In addition, a specialist models were trained on a subset of training data (hereafter *clean dataset*) which included only high quality data, sampled at 500-1000Hz and having an average noise level of up to 0.04 degrees RMS.

Results. Figure 1 shows the classification performance of universal (red bars) and specialist models (blue bars). On the right of Fig. 1, different markers depict the performance of event detection algorithms and human coders, reported in literature. All evaluations in studies listed in legend of Fig.1 were performed on data similar to clean dataset. Results show that specialist models, except LDA, QDA and Naive Bayes, outperform two current state-of-the-art algorithms by quite a large margin. Moreover, even if tested on noisier data, sampled at a broad range of sampling rates, 5 out of 10 machine learning based universal models perform better if compared to hand crafted algorithms.

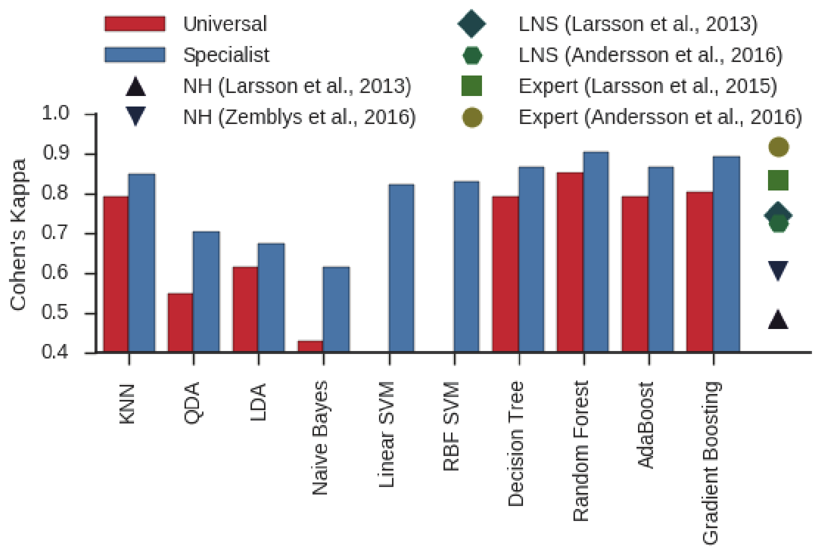

Fig. 1. Performance of universal (red) and specialist (blue) classifiers. Linear and RBF SVM models failed to fit the full dataset. Markers show performance of state-of-the art algorithms and human coders (circle and square) as reported in the literature.

In clean data 7 machine learning based specialist classifiers outperform human coders as reported in [5], while Random Forest almost reaches expert performance reported by Andersson et al. [2].

Conclusions. Results presented in this study suggest that out of 10 tested algorithms, Random Forest is the best choice if building machine learning based event detection algorithm. It is fast, can be easily ran in parallel on multiple cores and does not require any specific hardware, just a regular computer. The drawback of Random Forest and similar classifiers is that they still require hand crafted features to be extracted first and a post-processing step is needed to build meaningful eye-movement events [2].

But the real promise of machine learning is that in the near future we may have a single algorithm that can detect all eye-movement events out there, without a need to adjust any parameters. The future direction of machine learning based event detection is clearly end-to-end deep learning [1]. Such an approach allows to input raw gaze data and get events without the need of intermediate hand crafting. If deep learning works for event detection as it worked in other research areas, it might even surpass human performance.

References

1. Anantrasirichai, N., Gilchrist, I. D., and Bull, D. R. (2016). Fixation identication for low-sample-rate mobile eye trackers. In 2016 IEEE International Conference on Image Processing (ICIP), pages 3126-3130. IEEE.

2. Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., and Nystrom, M. (2016). One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. Behavior Research Methods, pages 1-22.

3. Hessels, R. S., Niehorster, D., Kemner, C., and Hooge, I. T. C. (2016). Noise-robust Fixation in eye-movement data: identification by 2-means clustering (I2MC). Manuscript submitted for publication.

4. Larsson, L., Nystrom, M., and Stridh, M. (2013). Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. IEEE Transactions on Biomedical Engineering, 60(9):2484-2493.

5. Larsson, L., Nystrom, M., Andersson, R., and Stridh, M. (2015). Detection of fixations and smooth pursuit movements in high-speed eye-tracking data. Biomedical Signal Processing and Control, 18(0):145 – 152.

6. Nystrom, M. and Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. Behavior Research Methods, 42(1):188-204.

7. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3):211-252.

1. Salvucci, D. D. and Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the 2000 Symposium on Eye Tracking Research & Applications, ETRA '00, pages 71-78.

2. Zemblys, R., Niehorster, C. D., Komogortsev, O., Holmqvist, K. (2016). Using machine learning to detect events in eye-tracking data. Behavior Research Methods. *In press*.

Eye-Movement Event Detection Meets Machine Learning

R. Zemblys
Department of Engineering, Siauliai University, Lithuania

This paper presents a comparison of 10 machine learning algorithms in eye-movement event detection task. The goal was to build a universal algorithm, which could work with any type of the eye-tracking data. Results show that even if tested on noisy data, sampled at a broad range of sampling rates, 5 out of 10 machine learning based universal models perform better than state-of-the-art algorithms. Even more, 7 machine learning based specialist classifiers, trained to work with high quality data, outperform expert coders as reported by Larsson et al. (2015).