Eye tracking-based Sedentary Activity Recognition with Machine Learning

Julian Meijerink, j.a.m.meijerink@student.tudelft.nl, Supervised by Guohao Lan, Lingyu DU

1. Background

1. Eye based trackers are becoming more popular.
2. Gaze estimation can provide useful context [1], for example attention monitoring while driving, gaming [2].

2. Problem analysis

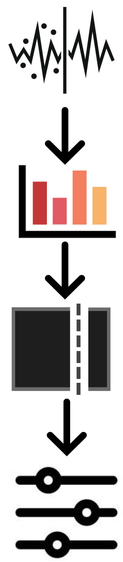
1. There is a gap in research on analyzing impact of using different feature extraction methods on recognizing different sedentary activities (browsing, reading etc)

3. Research Question

1. What are the best features to extract to be used for training conventional machine learning algorithms, e.g. k-NN, SVM, and decision tree?
2. How do the results compare to results obtained by deep learning, or using different data?

4. Method

1. Preprocess by normalizing, filtering outliers
2. Use peak- and median filter to estimate fixations
3. Extract features out of fixations and saccades
4. Split data both user dependent and user independent
5. Use K-NN, SVM and Random Forest Tree to classify
6. Fine-tune parameters and window sizes used for results



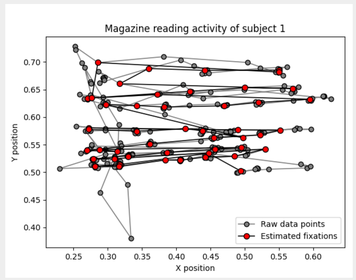
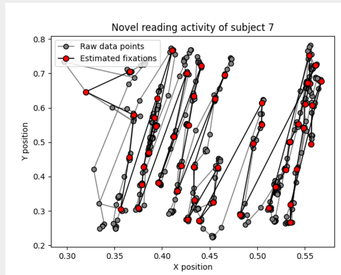


Figure 1: A small excerpt of subject 5 reading a magazine. Fixations are in red over the raw grey data.

Figure 2: Small excerpt of subject 1 reading a novel. The top to bottom reading style is apparent.



	Paper 1		Paper 2		Paper 3	Paper 4		Paper 5			
Algorithm	CNN		LSTM		Conventional Machine Learning						
						Sedentary		Desktop		Reading	
	dep	ind	dep	ind	Classifier	dep	ind	dep	ind	dep	ind
Sedentary	0.99	0.69	0.98	0.67	K-NN	0.77	0.48	0.84	0.54	0.91	0.71
Desktop	0.99	0.39	0.95	0.32	SVM	0.86	0.52	0.95	0.60	0.85	0.75
Reading	0.99	0.67	0.98	0.31	Forest	0.94	0.65	0.82	0.58	0.99	0.67

5. Results

1. Deep learning performed extremely well on user dependent evaluation, but performed less impressive on user dependent evaluation.
2. Conventional machine learning also performed well on user dependent, and quite well on user independent evaluation.

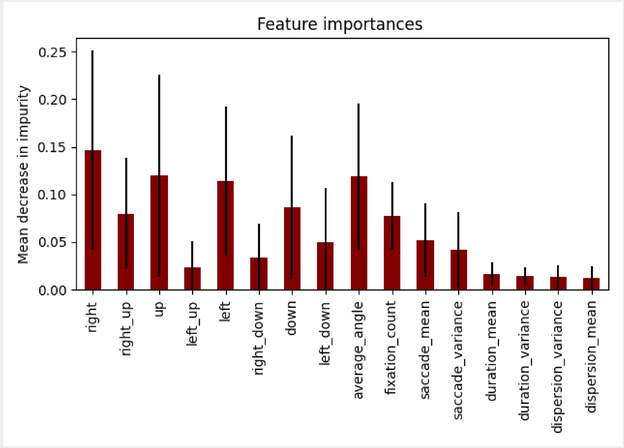


Figure 3: All features and their importance ranked. The first 9 are the direction based-features. Following are the rest ordered by importance

6. Conclusion

1. Deep learning is strong when it comes to classifying user dependent data, and conventional machine learning is somewhat stronger and more consistent classifying user independent data.
2. Direction-based features were the most important. This makes sense as Japanese writing is direction unique.

7. References

1. A. Bulling and T. O. Zander, “Cognition-aware computing,” IEEE Pervasive Computing, vol. 13, no. 3, pp. 80–83, 2014.
2. A. N. Singh, “Practical applications of Eye Tracking Technology,” Medium, 13-May-2019. [Online]. Available: here. [Accessed: 30-Apr-2022].

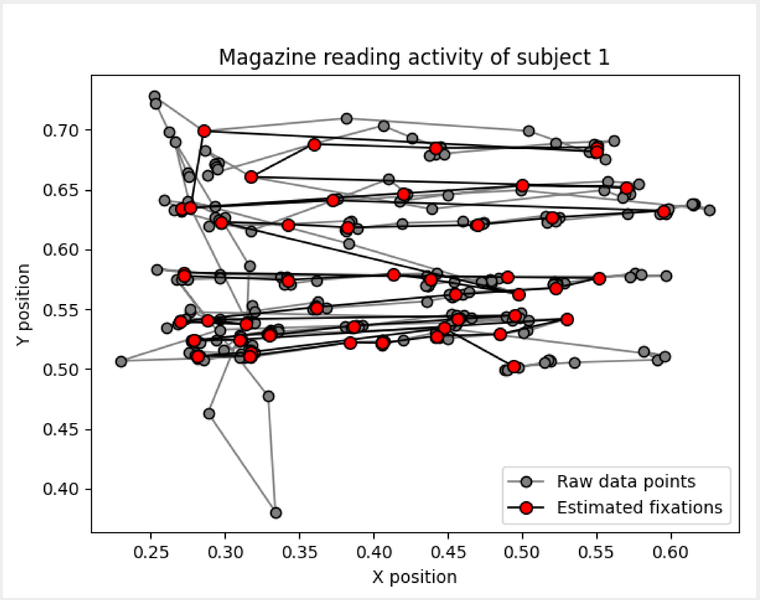


Figure 1: A small excerpt of subject 1 reading a magazine. Fixations are in red over the raw grey data.

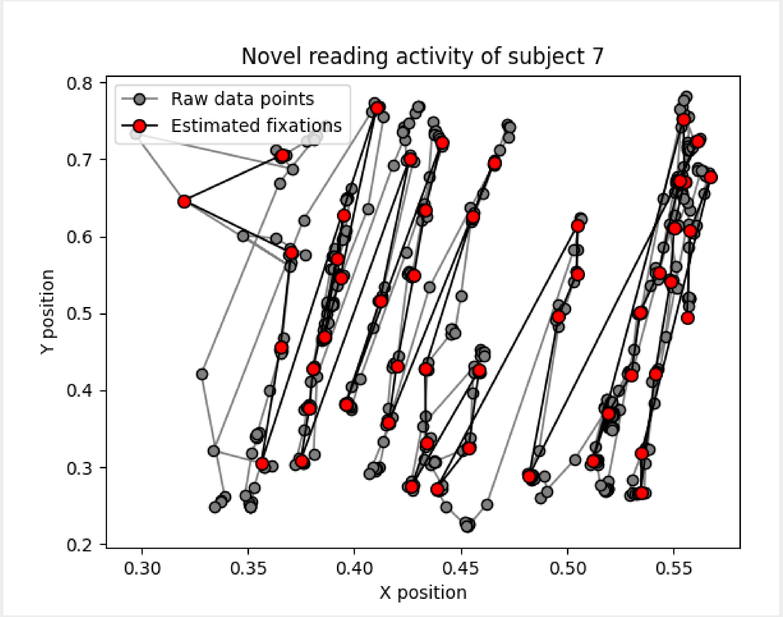


Figure 2: Small excerpt of subject 7 reading a novel. The top to bottom reading style is apparent.

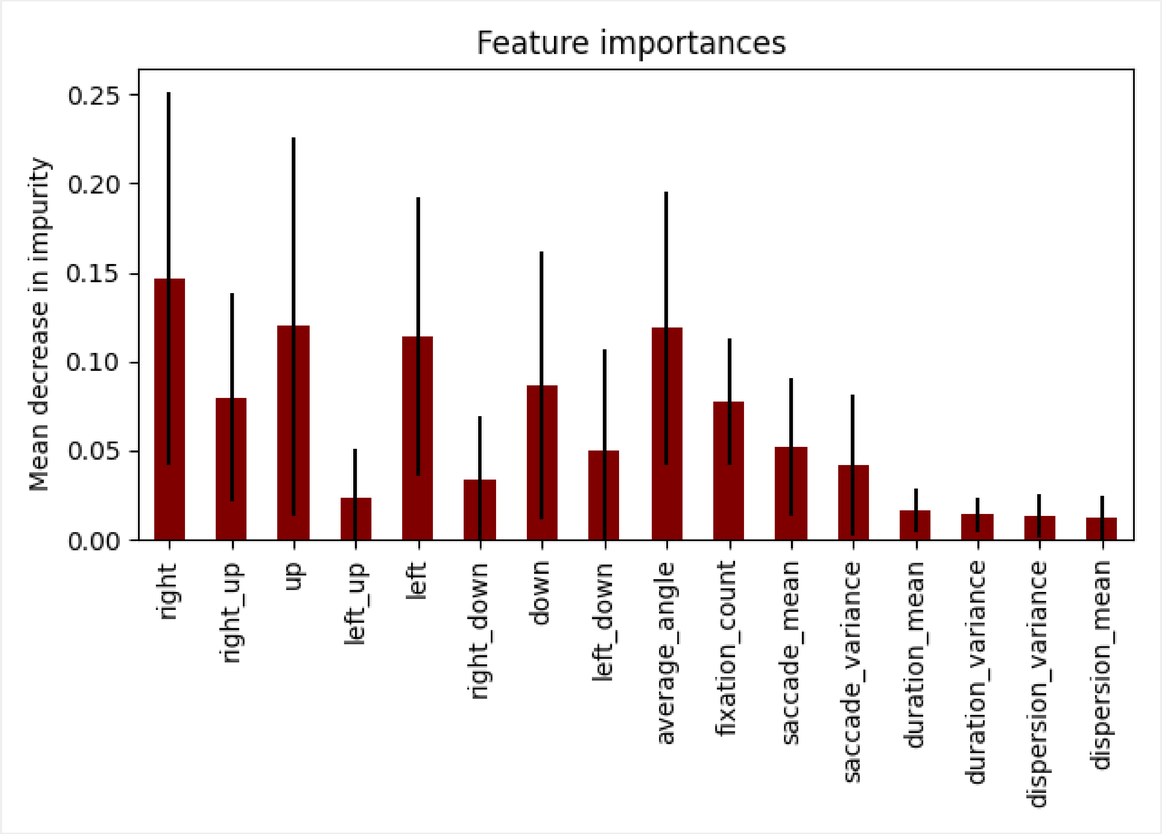


Figure 3: All features and their importance ranked. The first 9 are the direction based-features. Following are the rest ordered by importance