Robustness of Eye Movement Biometrics against Varying Stimuli and Varying Trajectory Length

Christoph Schröder[1] Sahar Mahdie Klim Al Zaidawi[2]
Martin H.U. Prinzler[2] Sebastian Maneth[2] Gabriel Zachmann[1]

Contact: schroeder.c@uni-bremen.de

[1]*University of Bremen, Germany, Institute for Computer Graphics and Virtual Reality*
[2]*University of Bremen, Germany, AG Datenbanken*

CHI, May 2020

Eye Movement Biometrics

- Identify people
- Liveness
- Seamless and Continuous
- Applicable in HMDs

Previous Work

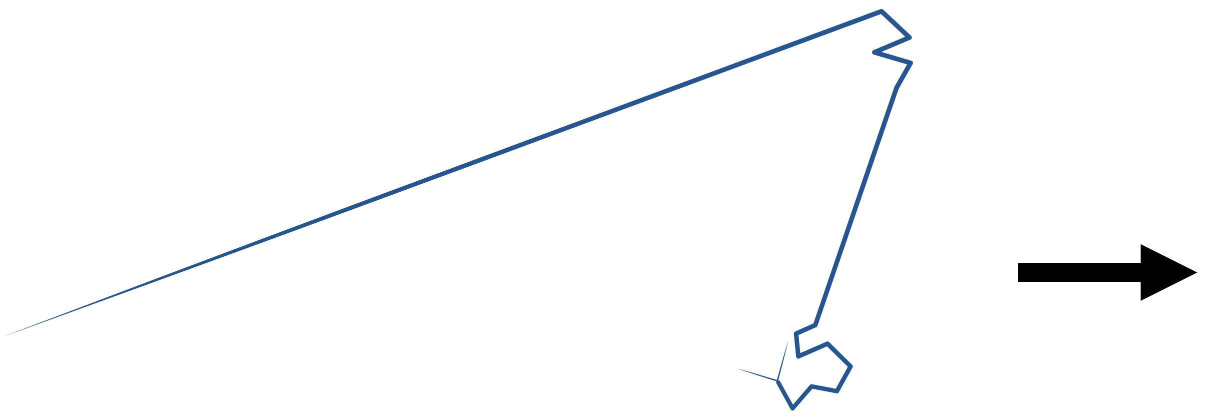
- Bio-eye competition [Rigas et al. 2017]
  - 82%-84%, 153 Users [George et al. 2016]
- Task independent person authentication
  - Gaussian Mixture Models, 29%-47% with 17 users [Kimmunen et al. 2010]
  - Random Forest, 37% with 17 users [Darwish et al. 2013]
- Multi-modal biometrics
  - Support Vector Machine, 24% (only view based features) with 18 users [Pfeuffer et al. 2019]

Our Contributions

1. Two extensions to the state-of-the-art classifier [George et al. 2016]
   - More features
   - Different classifier (Random Forest)
2. Analysis of stimulus (in-)dependence
   - Different stimuli for training and testing
   - Comparison of 4 configurations
3. Influence of tracking duration on identification accuracy
   - Varying training and testing sample size
   - Evaluation on less artificial dataset

Identification Pipeline

Trajectory



Classifiers

- Radial Basis Function Networks (RBFN)
  - Reimplementation of George et al. 2016
  - $C \cdot K$ neurons; C = number of users, $K = 32$
  - $\varphi(x) = e^{-\beta\|x-\mu\|^2}$ with $\beta = \frac{1}{2\sigma^2}$
  - μ: K-Means clustering for each user
  - σ: Mean euclidian distance to cluster
- Random Decision Forest (RDF)
  - 400 trees, no depth limit, min. 2 samples for split, $\sqrt{F}$ features

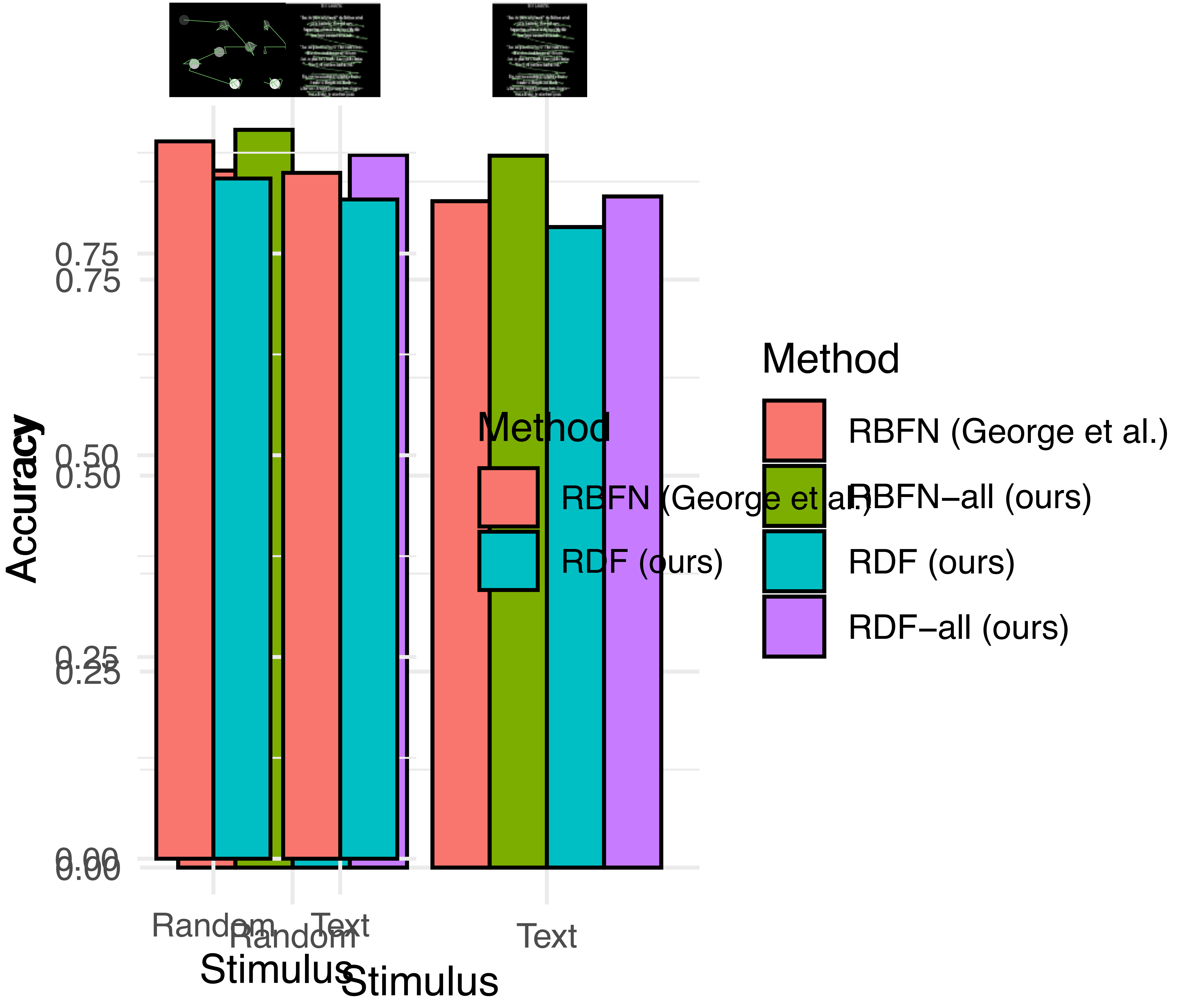
Features

- George et al. 2016
  - Iterative feature selection
  - Fixation 9 Text / 9 Random
  - Saccade 43 Text / 40 Random
- Ours: all combined
  - 52 unique features
  - Stimulus independent

Fixation duration
Standard deviation (X)
Standard deviation (Y)
Path length
Angle with previous fixation
Distance from the last fixation
Skewness (X)
Skewness (Y)
Kurtosis (X)
Kurtosis (Y)
Dispersion
Average velocity

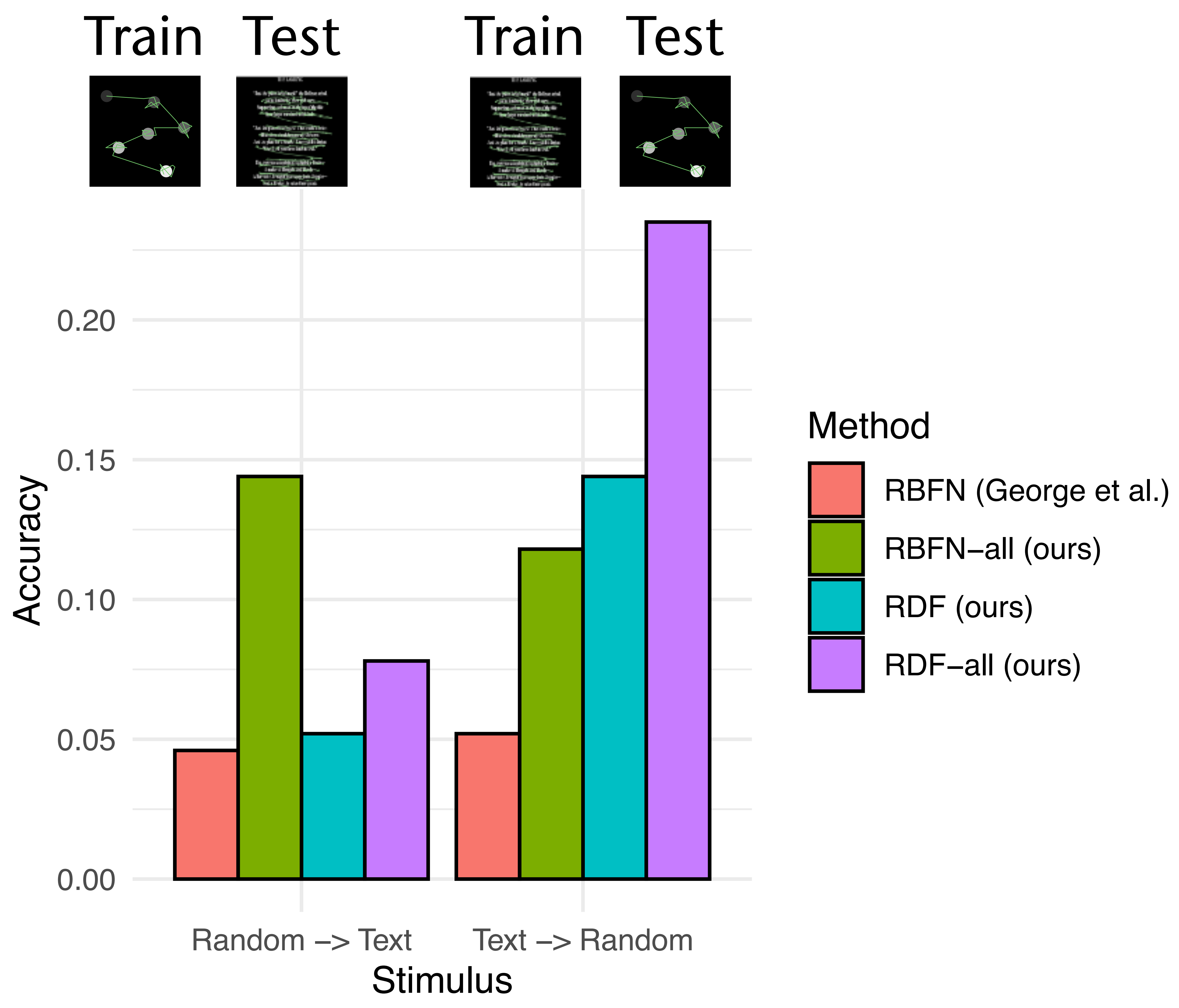
Stimulus Dependent Results

- Training and evaluation on same stimulus
- RBFN always better than RDF
- Increased accuracy without feature selection



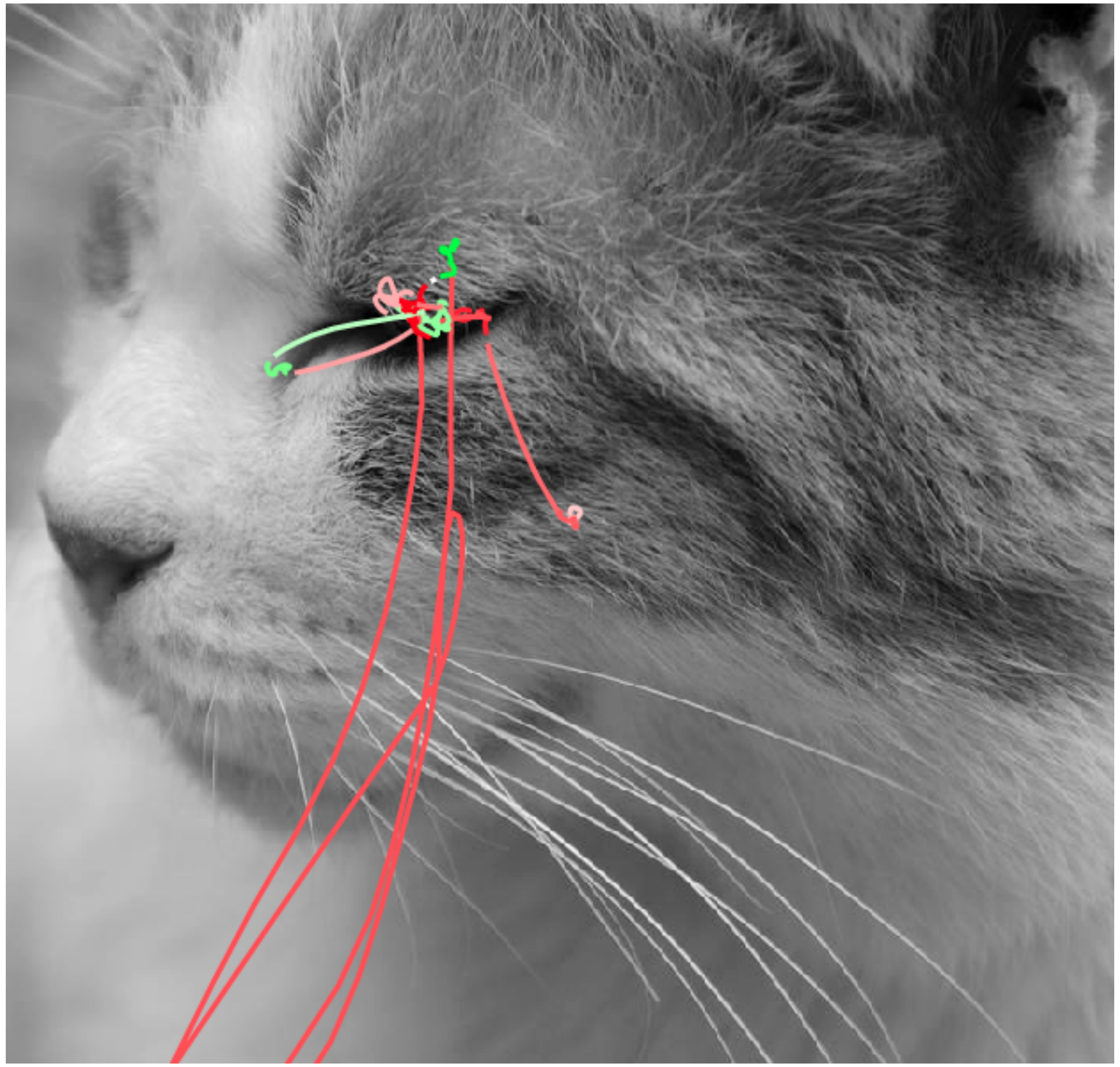
Stimulus In-Dependent Results

- Training on one and testing on another stimulus
- Maximum accuracy drops from 94.1% to 23.5%
- RDF generalises better

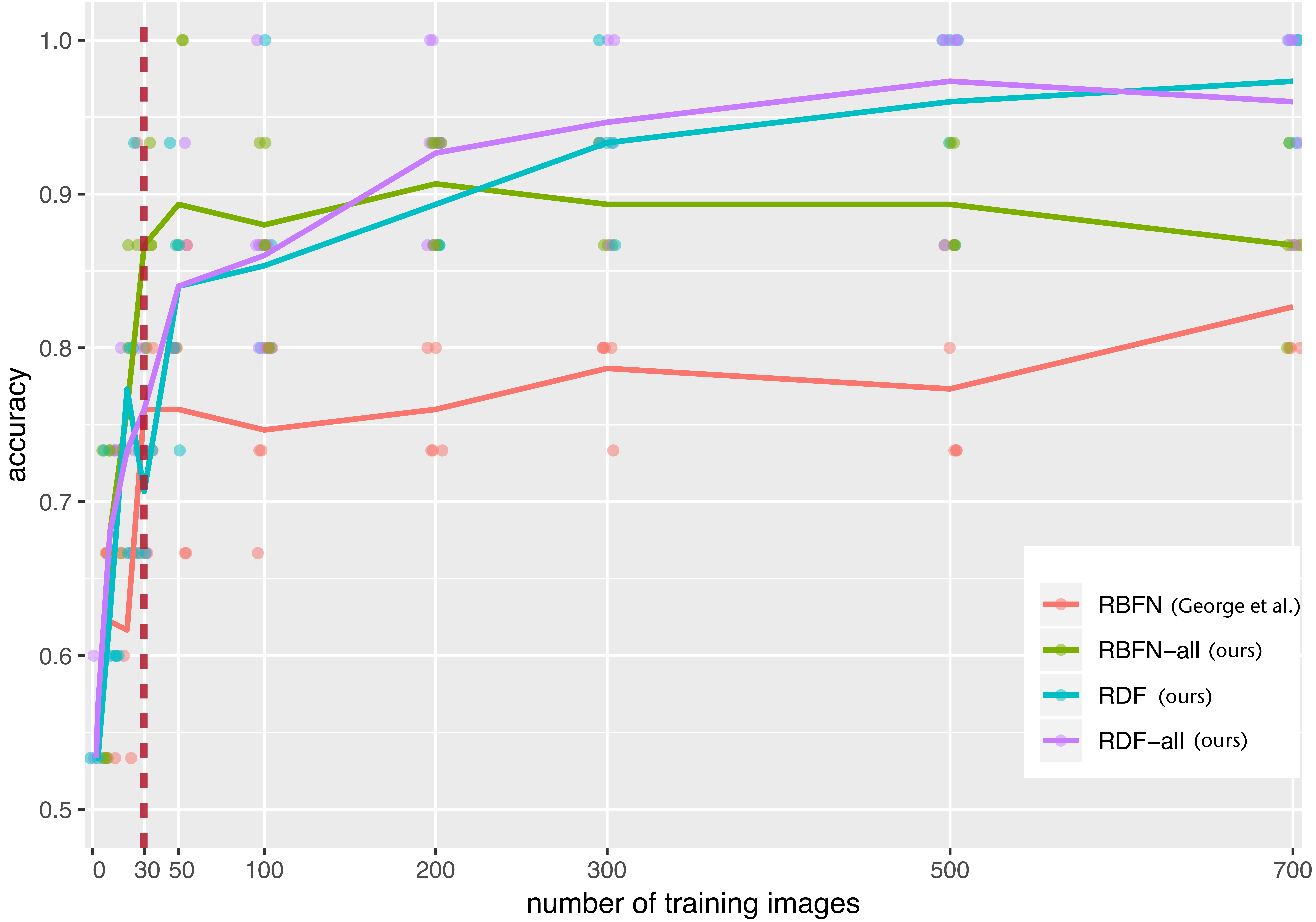


Optimal Trajectory Length

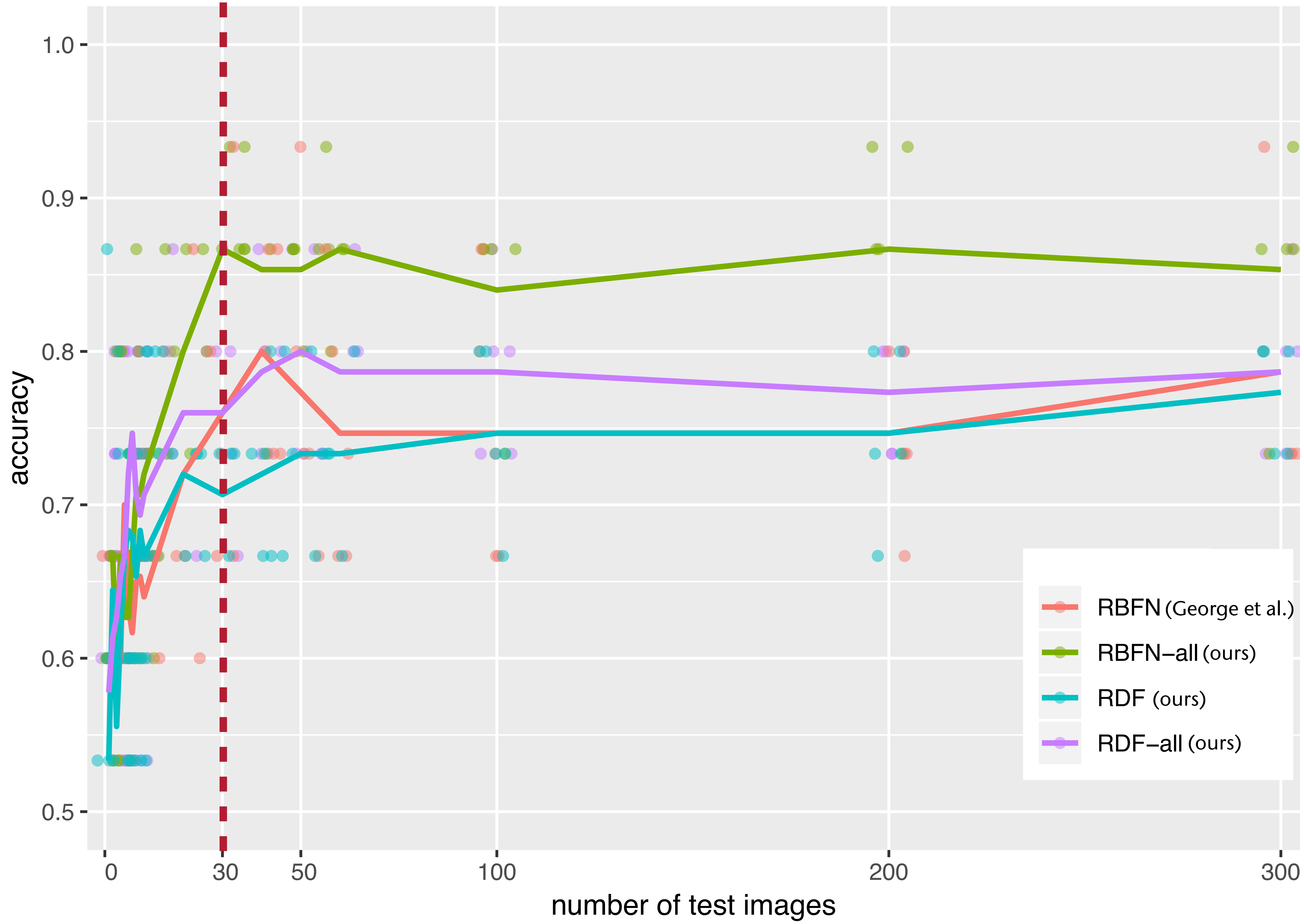
- Research question
  - Influence of train/test size
  - Weakly task independent performance
- MIT dataset [Judd et al. 2009]
  - 39 users
  - 3 seconds per image (sample)
  - 50 minutes per user
  - Task-agnostic



Influence of Number of Training Samples



Influence of Number of Test Samples



Conclusion

1. Identification accuracy improved by 5.2pp over the state-of-the-art (86.0% to 92.5%)
2. Task-independent identification still a challenge
   - Asymmetrical performance
   - Text ⇒ Ran 3x better than the other way around
3. Applicable for weakly task-independent identification with only 90s of trajectory data
   - 86.7% accuracy with 90 seconds training data
   - 94.7% with 900 seconds training data

- Our code is available here: https://cgvr.cs.uni-bremen.de/research/smida_ml/

Limitations and Future Work

- Reduce trajectory length for identification
- Weighting of saccades and fixations
- Recurrent Neural Networks

Thank You.

https://cgvr.cs.uni-bremen.de/research/smida_ml