Paired Trial Classification: A Novel Deep Learning Technique for MVPA

Jacob M. Williams, Ashok Samal, Prahalada K. Rao, Matthew R. Johnson

INTRODUCTION

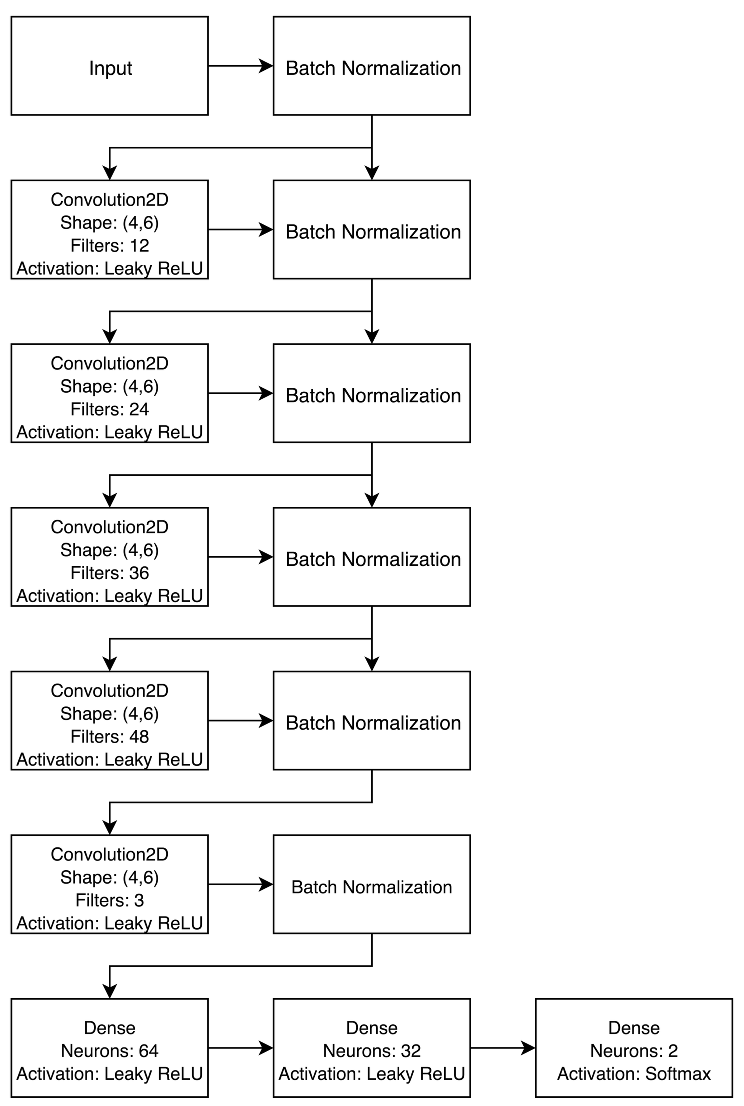
• Most traditional MVPA techniques attempt to directly classify trials into one of two or more classes

• We propose an alternative: Use a neural network to classify *pairs* of EEG trials as belonging to either the same or different classes, called Paired Trial Classification (PTC)

• Potential advantages: Can be used as a classification decision or a similarity metric; pairing samples allows for inherent data augmentation; facilitates other novel analysis strategies

• Can still be used to perform traditional multiclass classification by using a dictionary-based technique

PTC NETWORK ARCHITECTURE



ANALYSIS METHODS

• Data evaluated either as *unaveraged* single trials or sets of 20 *averaged* trials. Dictionary similarly composed of averaged or unaveraged

• Models evaluated:

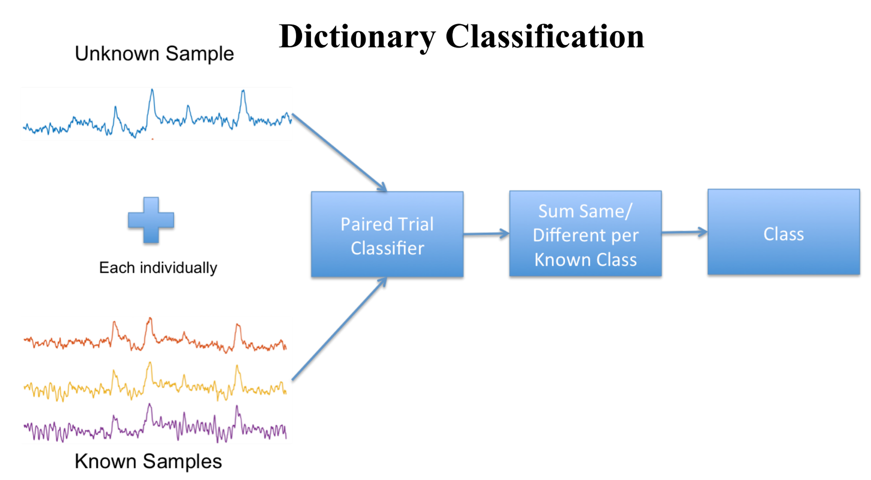
• Support Vector Machine (SVM) via PyMVPA

• Sparse Multinomial Logistic Regression (SMLR) via PyMVPA

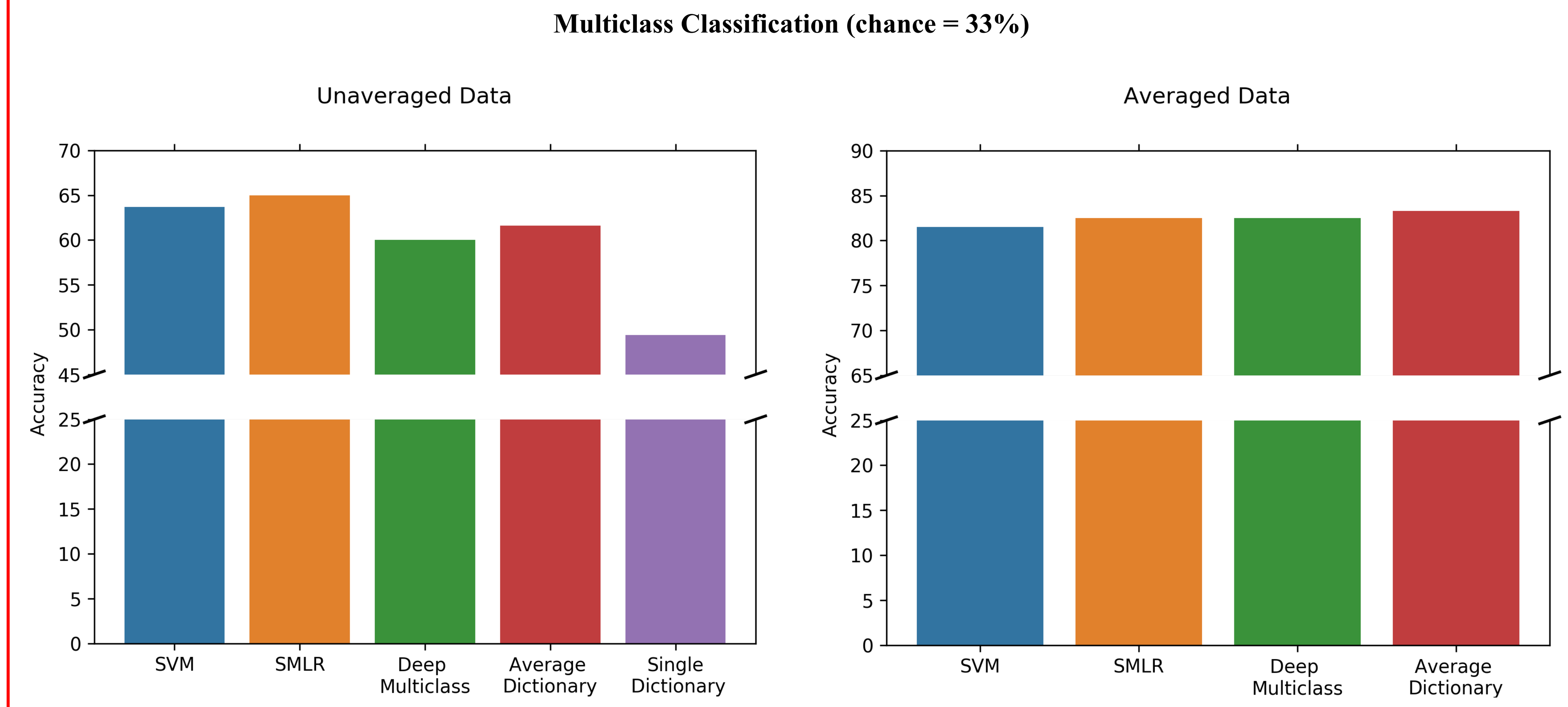
• Deep Multiclass Network via Keras (shares PTC architecture)

• PTC via Keras (architecture on left)

• Dictionary: Used PTC to compare each test trial to a set of known trials for classification

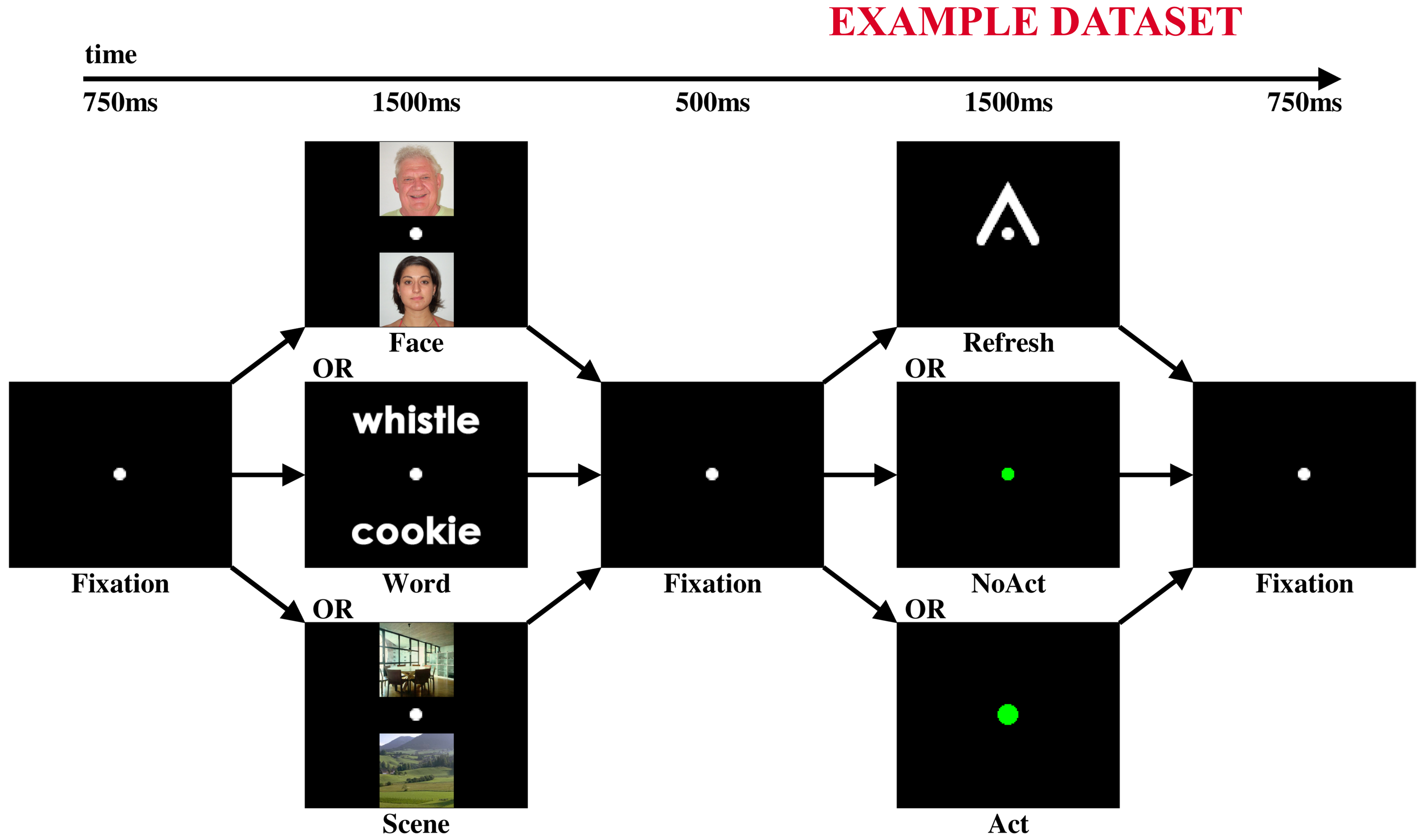


RESULTS



Paired Trial Classification (chance = 50%)		
Unaveraged to Unaveraged	**Unaveraged to Averaged**	**Averaged to Averaged**
56.1%	71.3%	86.2%

EXAMPLE DATASET



Task/EEG Methods:

•31 channels low-impedance (<5kΩ)

•Sampled at 250 Hz, time binning averaged in 40 ms bins

•37 young, healthy subjects

•***Initial Presentation*** interval: 1500ms of a pair of faces, scenes, or words presented onscreen

•Followed by one of several subsequent cues (but cue interval not analyzed here)

•Only ***Initial Presentation*** used for all analyses: ~200 trials/subject

•Bandpass filter of .01–100Hz during acquisition. Trials rejected with peak-to-peak amplitude > 150μV; EOG signal regressed out and each trial linearly detrended. Pre-cue baseline (100ms) average subtracted.

RESULTS & CONCLUSIONS

• Paired Trial Classification performed above chance

• Dictionary based classification is comparable to other methods

• Averaging significantly improves results for all techniques

• At these values, we may be hitting the noise ceiling of the dataset

• Still hard to beat tuned SVM/SMLR + simple dimensionality reduction

•.Convolutional network structure necessary for successful PTC; in preliminary tests, other models could not perform PTC

• Deep multiclass network overfits due to model complexity

• Very high variance among subjects, e.g. in all averaged analyses some subjects achieve up to 99% accuracy while a few sit at about 45%, even after trial rejection

FUTURE DIRECTIONS

• Evaluate on more datasets

• Improve method of selecting entries in dictionary

• Explore applications to outlier detection and clustering

• See poster D104 for applications to an fMRI dataset

REFERENCES & ACKNOWLEDGEMENTS

Chollet F. Keras. 2017. GitHub repository: https://github.com/fchollet/keras

Johnson MR, McCarthy G, Muller KA, Brudner SN, Johnson MK. 2015. *Journal of Cognitive Neuroscience* 27: 1823-1839.

The Theano Development Team. 2016. arXiv E-Prints: http://arxiv.org/abs/1605.02688

Supported by NVIDIA GPU grant to JMW/MRJ and NSF/EPSCoR grant #1632849 to MRJ and colleagues.