Deep learning techniques for decoding EEG signatures of viewing or refreshing face, scene, and word stimuli

Jacob M. Williams[1], Ashok Samal[1], Matthew R. Johnson[2]

[1]Department of Computer Science and Engineering, University of Nebraska – Lincoln, Lincoln, Nebraska [2]Department of Psychology, University of Nebraska – Lincoln, Lincoln, Nebraska

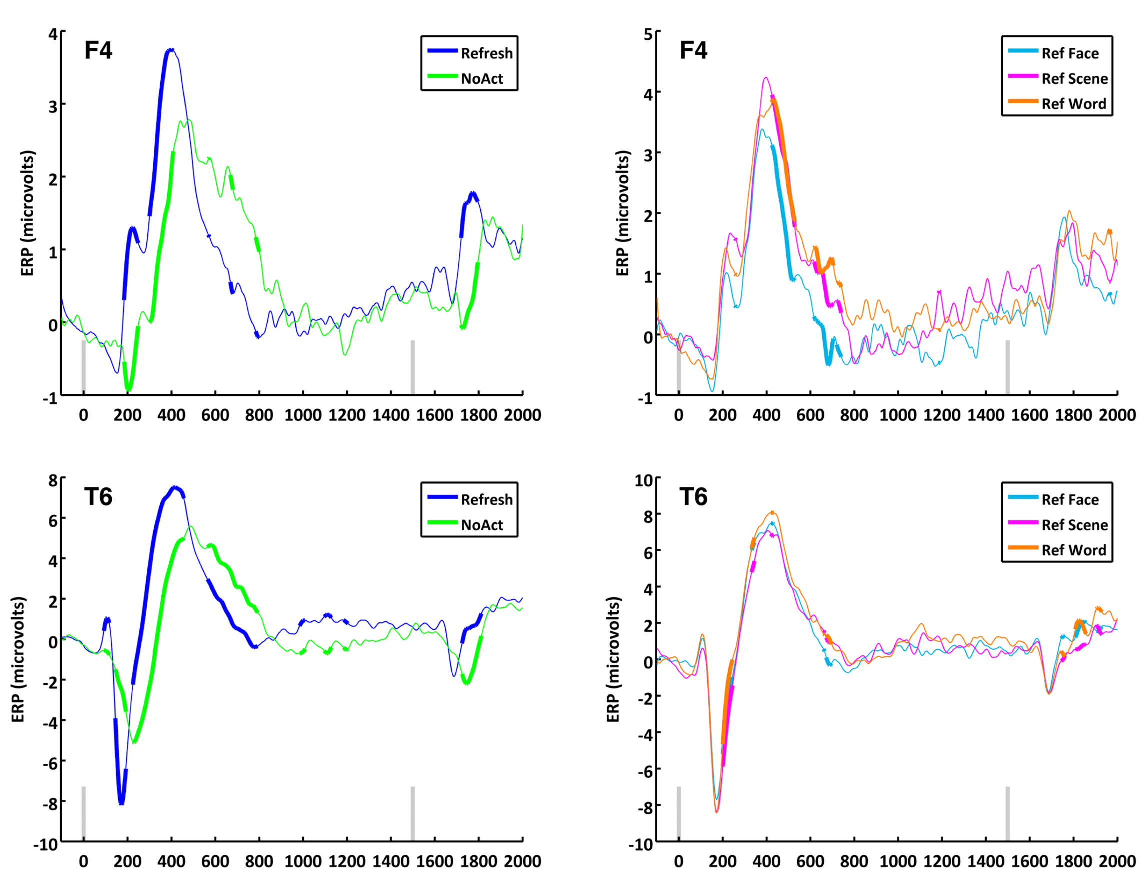
INTRODUCTION

• Contemporary deep learning techniques have been revolutionary for the classification of images, speech signals, and other data types, yet are rarely applied in the analysis of cognitive neuroscience datasets.

• We explored whether deep learning could be fruitfully applied to EEG MVPA in a visual perception and refreshing dataset previously analyzed with Sparse Multinomial Logistic Regression (SMLR).

• How does training an individual model per subject and a single, universal model across subjects influence classification accuracy? What deep learning models are effective, and under what circumstances?

FIGURE 1: PREVIOUS ERP RESULTS



Bold timepoints (left): Significant difference between conditions at FDR-corrected threshold.

Bold timepoints (right): Significant difference between conditions at *uncorrected* threshold. No timepoints survived FDR correction.

ANALYSIS METHODS

• 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.

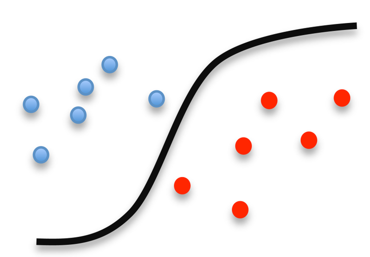
• Time binning for some analyses averaged timepoints in 40-ms bins.

• Classification was over category viewed/refreshed; chance = 33.33%.

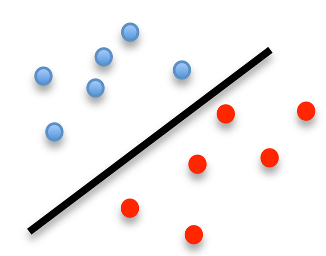
• Traditional MVPA (SMLR, SVM) analyses: Matlab with custom code.

• Deep learning (MLP, CNN, LSTM, LSTM+CNN) analyses: Python using Keras and Theano libraries.

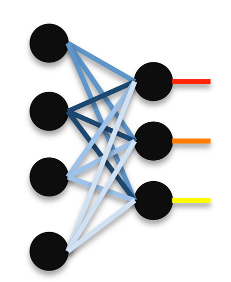
MODELS



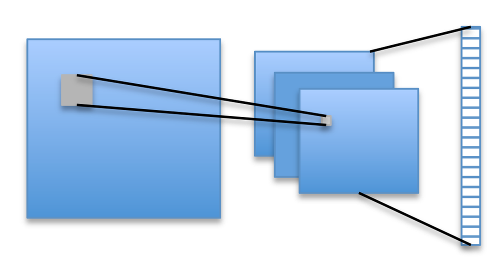
SMLR



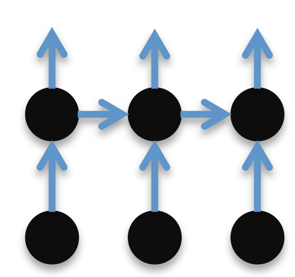
SVM



MLP



CNN



LSTM

SMLR: Sparse Multinomial Logistic Regression
SVM: Support Vector Machine
MLP: Multilayer Perceptron
CNN: Convolutional Neural Network
LSTM: Long Short-Term Memory Network

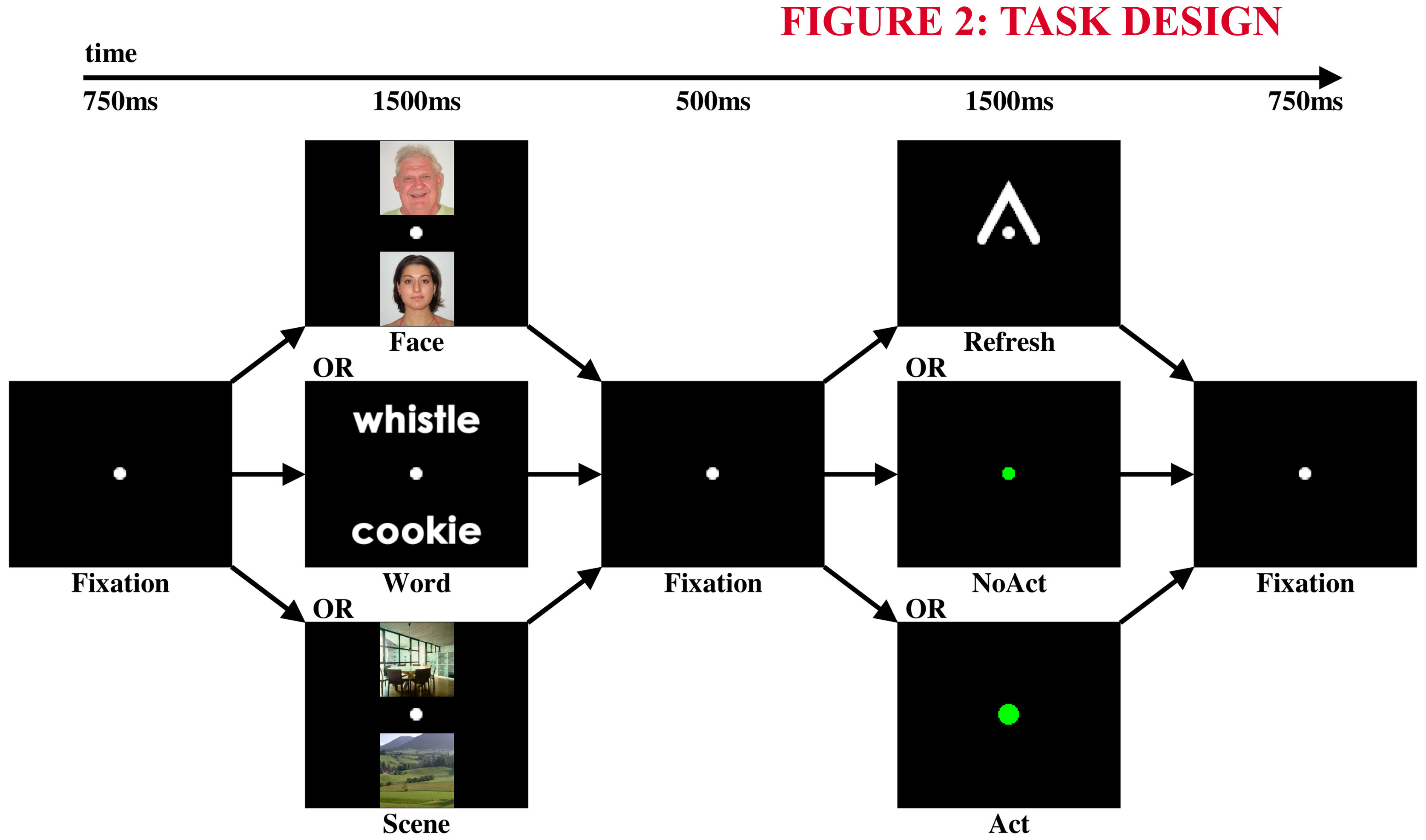
TABLE 1: TRADITIONAL MVPA RESULTS

Model	Subjects	Interval	Time Binned	Accuracy
SMLR	Single	Presentation	Yes	58.69%
SMLR	Single	Presentation	No	59.02%
SMLR	Single	Refresh	Yes	**37.96%**
SMLR	Single	Refresh	No	36.36%
SMLR	Universal	Presentation	Yes	**62.83%**
SMLR	Universal	Presentation	No	57.54%
SMLR	Universal	Refresh	Yes	35.53%
SMLR	Universal	Refresh	No	34.5%
SVM	Single	Presentation	Yes	59.09%
SVM	Single	Presentation	No	57.45%
SVM	Single	Refresh	Yes	37.59%
SVM	Single	Refresh	No	37.04%
SVM	Universal	Presentation	Yes	62.02%
SVM	Universal	Presentation	No	59.51%
SVM	Universal	Refresh	Yes	35.66%
SVM	Universal	Refresh	No	35.63%

TABLE 2: DEEP LEARNING RESULTS

Model	Subjects	Interval	Time Binned	Accuracy
MLP	Universal	Presentation	Yes	54.29%
MLP	Universal	Presentation	No	55.69%
MLP	Universal	Refresh	Yes	34.33%
MLP	Universal	Refresh	No	35.06%
CNN	Universal	Presentation	Yes	60.04%
CNN	Universal	Presentation	No	62.09%
CNN	Universal	Refresh	Yes	34.75%
CNN	Universal	Refresh	No	35.63%
LSTM	Universal	Presentation	Yes	58.44%
LSTM	Universal	Presentation	No	63.61%
LSTM	Universal	Refresh	Yes	35.91%
LSTM	Universal	Refresh	No	**37.72%**
LSTM + CNN	Universal	Presentation	Yes	46.89%
LSTM + CNN	Universal	Presentation	No	**64.36%**
LSTM + CNN	Universal	Refresh	Yes	32.65%
LSTM + CNN	Universal	Refresh	No	34.30%

FIGURE 2: TASK DESIGN



Task/EEG Methods:

- 31 channels low-impedance (<5kΩ)
- Sampled at 250 Hz
- 37 young, healthy subjects
- ***Initial presentation*** interval: 1500ms of a pair of faces, scenes, or words presented onscreen
- ***Refresh*** interval: 1500ms arrow cue directing participants to reflectively attend (think back to) one item
- ***NoAct / Act***: Control conditions for refresh cue (not analyzed here)
- ~200 trials/subject for *initial presentation*
- ~100 trials/subject for *refresh*

RESULTS & CONCLUSIONS

• Consistent with prior results, perceptual categories were substantially more decodable than refreshed categories.

• Single-subject analyses for deep learning models were omitted from the table, as they frequently failed to converge, yielding unstable results.

• Traditional MVPA models (SMLR and SVM) performed best on time-binned data.

• Deep learning models (MLP, CNN, LSTM, LSTM+CNN) performed best on non-time-binned data.

• Neither traditional MVPA nor deep learning models showed evidence of a single model dominating over competing models for all cases.

• Deep learning approaches can offer mild benefits over traditional approaches, given enough data.

• As model complexity increases, so does the negative impact of dimensionality reduction.

FUTURE DIRECTIONS

• Deep learning models have many more configurable parameters than traditional MVPA methods; continue to fine-tune these models.

• Test alternative deep learning architectures; examine other test datasets; employ alternative training paradigms (e.g., transfer learning).

• Explore strategies for data augmentation, which can significantly boost deep learning performance.

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.