Task Learning Over Multi-Day Recording via Internally Rewarded Reinforcement Learning Based Brain Machine Interfaces (2020)

Shen, X., Zhang, X., Huang, Y., Chen, S. and Wang, Y.,

Sachin Salim

BME 517: Paper Presentation
04 Apr 2023

Introduction

Introduction

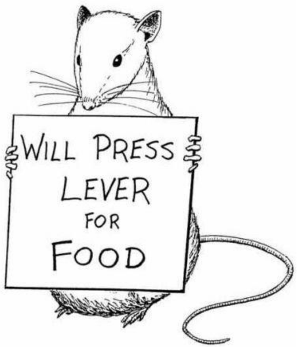
- **Motivation:** to develop an autonomous BMI that enables paralyzed individuals to control external devices using their neural activity
- I'm interested in the potential of BMIs to improve the lives of people, and in the use of RL-based decoders to interpret neural activity
- A novel approach that utilizes the brain's internal reward mechanism to replace external guidance and enable autonomous task learning

Methods

Methods

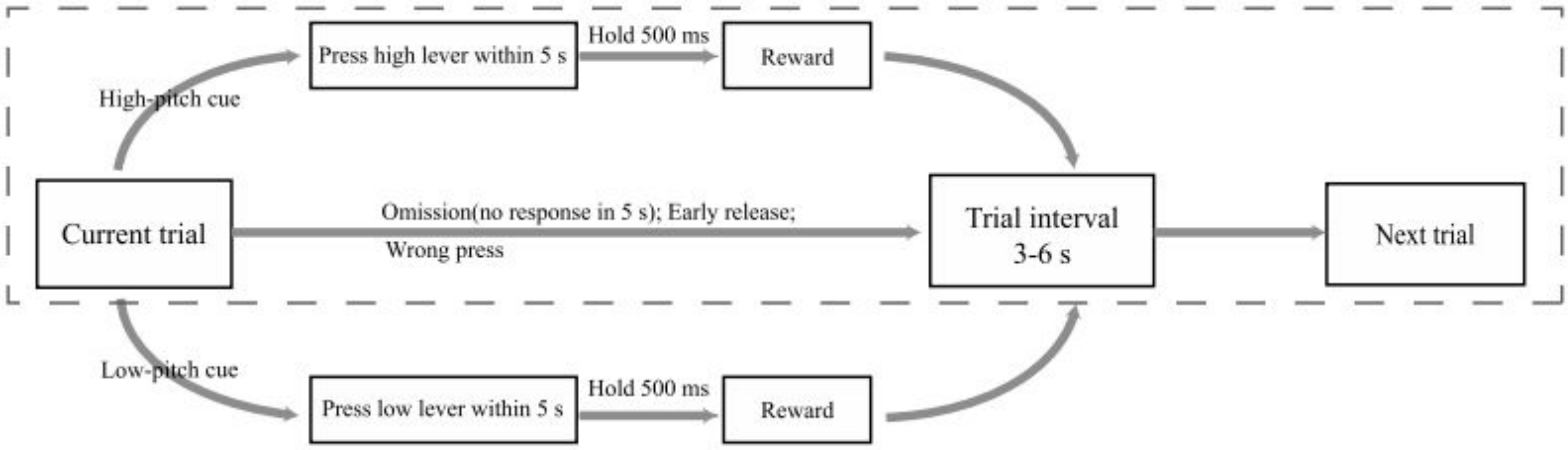
Behavioral Task and Data Collection

- The experiment was conducted with six male rats which were trained to perform a one-lever-press task before learning a two-lever-press discrimination task
- Neural data was collected from two 16-channel microelectrode arrays in both the M1 and mPFC (medial prefrontal cortex)



Methods

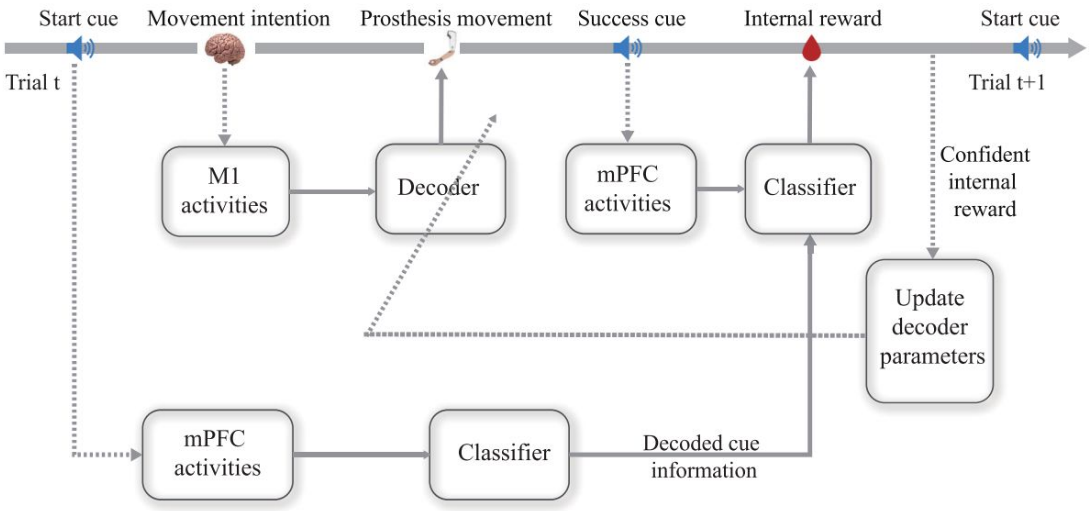
Behavioral Task and Data Collection



Two-lever-press discrimination task diagram

Methods

The Internally Rewarded Decoding Framework



1. mPFC signals are decoded after cue onset to first classify the cue information
2. M1 activity post cue onset is decoded to generate an action
3. post-action mPFC activity is used to extract the internal reward information given the classified cue
4. classified reward or non-reward results are used to update the decoder parameters before the next trial starts

Methods

Internal Critic from mPFC Activity - SVM

- SVM is used to classify cue information and reward/non-reward trials using mPFC neural activity

$$\hat{r} = h\left(\omega^T \phi(x) + b\right)$$

- The discrimination result using mPFC activity is used to replace the explicit reward in the RL structure

- We learned about supervised classifier algorithms in Lab 7
- In contrast to Naive Bayes and LDA, SVM is a nonlinear method that uses a kernel function to transform the data into a higher dimensional space

Methods

RL Action Decoding with Internal Reward

- Attention-gated reinforcement learning (AGREL) neural network is used to decode actions from M1 neural activity with the internal critic from mPFC activity.
- It has three perceptron elements in the output layer:
  - high lever press
  - low lever press
  - no lever press

- We learned about continuous decoders in Labs 8-9
- Similar to Kalman filters, RL decoder is suitable for online data as they can be updated in real-time as new data becomes available

Results

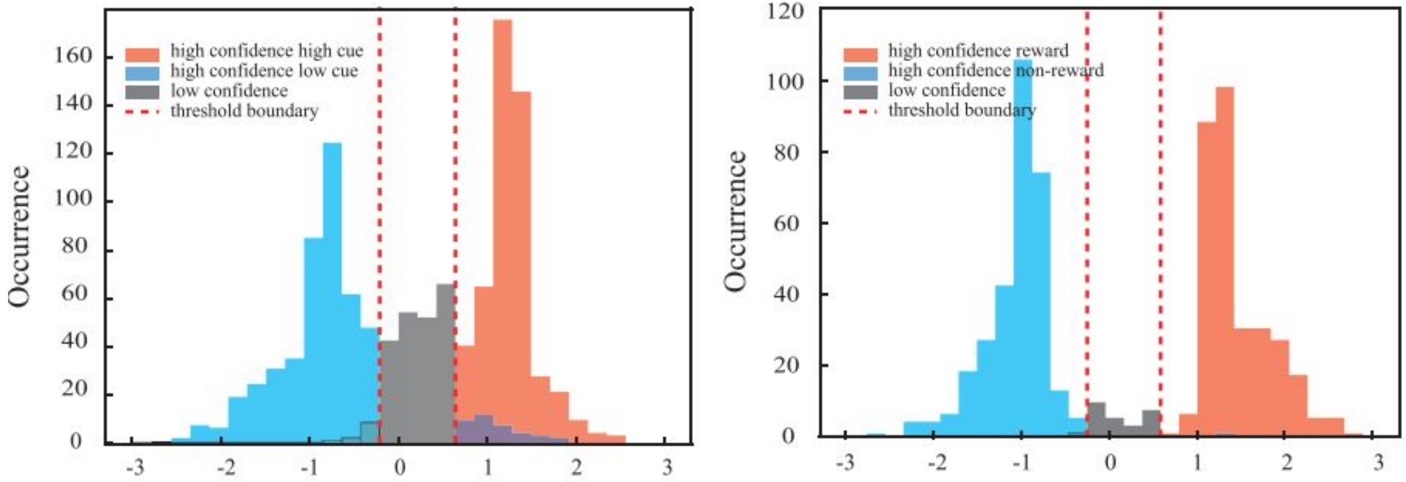
Results

Internal Critic from mPFC Activity - SVM Classifier

- mPFC activity encodes both cue and reward information and is a better internal critic than M1 activity
- Average classification accuracy for cue information is **83.9%** once the task is mastered
- Using mPFC activity, the average reward discrimination accuracy is **87.5%**, which is higher than using M1 activity (**62.2%**)

Results

Internal Critic from mPFC Activity - SVM Classifier



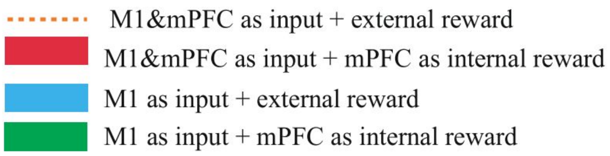
(a) Confidence level of SVM decision values (cue classification)

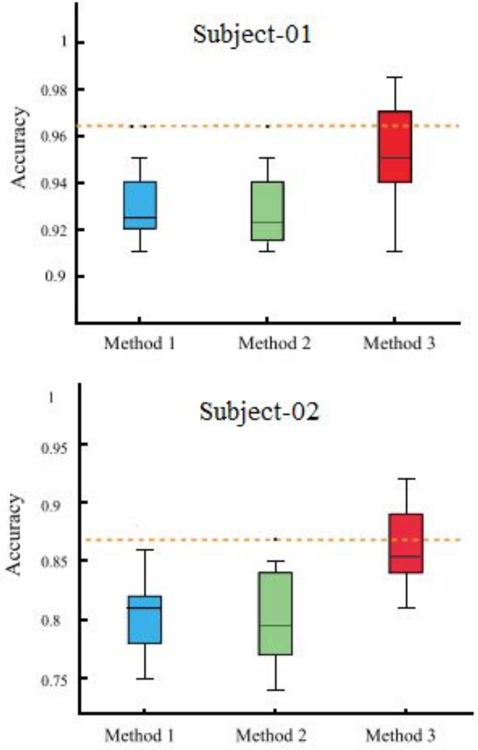
(b) Confidence level of SVM decision values (reward classification)

Results

RL Action Decoding with Internal Reward

- Results show that using the mPFC activity as an internal critic in RL achieves good performance, which is very close to that using the external reward
- The multi-site recording using both M1 and mPFC as input achieves better decoding performance than using M1 as input only





Results

RL Action Decoding with Internal Reward

TESTING ACCURACIES OF DECODING MODELS ACROSS SUBJECTS.

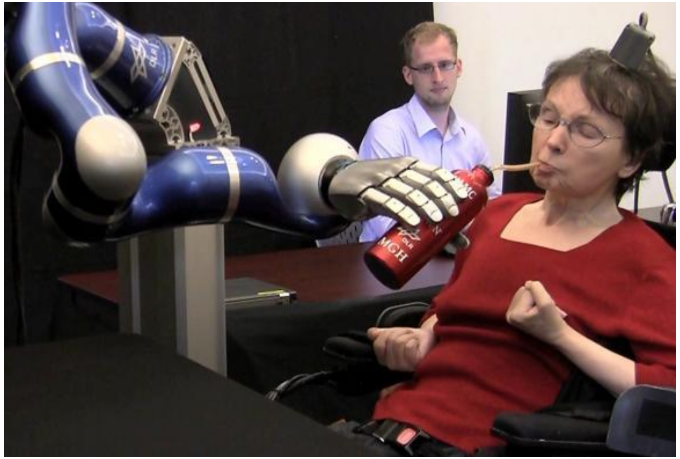
Decoding model		Subject	
Input	Reward type	Subject-01	Subject-02
M1	Internal reward from mPFC	91.7%±1.2%	79.2%±3.8%
M1	External water reward	92.9%±1.5%	80.4%±3.4%
M1+mPFC	Internal reward from mPFC	94.7%±1.1%	85.6%±3.5%

Discussion

Discussion

Implications

- The proposed decoder can be implemented in clinical trials, allowing the neuroprostheses to autonomously learn a new task
- The framework can address the challenges of time-variant neural activity and maintain high decoding accuracy



https://news.brown.edu/articles/2012/05/braingate2

References

Shen, X., Zhang, X., Huang, Y., Chen, S., & Wang, Y. (2020).
Task learning over multi-day recording via internally rewarded reinforcement learning based brain machine interfaces.
IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(12), 3089-3099.

Questions?