2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

Improved Grip Force Prediction Using a Loss Function that Penalizes Reward Related Neural Information

Jaganth Nivas Asok Kumar, Joseph Thachil Francis, *Member, IEEE*

Abstract—**Neural activity in the sensorimotor cortices has been previously shown to correlate with kinematics, kinetics, and non-sensorimotor variables, such as reward. In this work, we compare the grip force offline Brain Machine Interface (BMI) prediction performance, of a simple artificial neural network (ANN), under two loss functions: the standard mean squared error (MSE) and a modified reward penalized mean squared error (RP_MSE), which penalizes for correlation between reward and grip force. Our results show that the ANN performs significantly better under the RP_MSE loss function in three brain regions: dorsal premotor cortex (PMd), primary motor cortex (M1) and the primary somatosensory cortex (S1) by approximately 6%.**

I. INTRODUCTION

BMI's for the restoration of movement decode neural activity, usually from the sensorimotor cortices, to drive external end effectors, such as robotic prosthetics, or computer cursors [1, 2, 3]. However, it has been shown by many groups that the activity in the sensorimotor cortices are not only related to motor variables, such as direction and force, but also non-movement variables, such as reward [4], reward prediction error [5], arousal [6], motivation [7] etc.

This reward modulation of the sensorimotor cortices has been hypothesized to occur because of dopaminergic projections that connect known reward centers in the mid and deep brain such as the Ventral Tegmental Area (VTA) and Substantia Nigra (SNc) to the cortex [4]. Single unit (SU) activity [8, 9, 10, 11], local field potentials (LFP) [12,13] and pairwise correlation [14] between units have all been shown to be modulated by reward expectation.

The goal of intention based BMIs is to translate movement intention from neural activity to actual movement in a robust and consistent manner. Since any action a subject makes, is linked to some sense of reward, either explicit or implicit, in the present or in the future, it makes intuitive sense to remove the effects of reward, and associated affect, to obtain a "pure" movement signal BMI decoding.

To this end, we hypothesized that the prediction of force from the neural activity can be improved if we remove the effect of non-kinematic or non-kinetic variables. The aim of this work is to test whether the performance of a simple ANN is significantly improved when the correlation between the predicted grip force and reward level is penalized in the loss function.

This research was supported by National Institutes of Health (NIH) grants NIH 1R01NS092894-01, NSF IIS-1527558, and DARPA N66001-10-C-2008.

Jaganth Nivas Asok Kumar is a with the Biomedical Engineering department at the University of Houston, TX-77004, Ph:832-679-0697; e-mail: jasokkumar@uh.edu).

II. METHODS

A. Surgery

Two non-human primates (NHPs), one male rhesus macaque (NHP S) and one female bonnet macaque (NHP P) were implanted with chronic 96-channel platinum microelectrode arrays (Utah array, 10 × 10 array separated by 400 µm, 1.5 mm electrode length, ICS-96 connectors, Blackrock Microsystems). The hand and arm region of M1 contralateral to their dominant hand was implanted with the same technique as our previous work [8,9]. All surgical procedures were conducted in compliance with guidelines set forth by the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the State University of New York Downstate Institutional Animal Care and Use Committee.

B. Neural Data Recording

As explained in our previous work [8], after a 2–3 week recovery period, spiking activity was recorded with a multichannel acquisition processor system (MAP, Plexon Inc.) while the subjects performed the experimental task (described below). Neural signals were amplified and band-pass filtered between 170 Hz and 8 kHz to isolate single and multi-unit activity and sampled at 40 kHz, and each channel manually thresholded to detect single units. Single and multi-units were sorted offline based on their waveforms using principal component (PC)-based methods in offline sorter (Plexon Inc.). The number of units used for analysis after sorting were (87, 57, 81) and (87, 58, 85) for PMd, M1 and S1 in block 1 and block 2 respectively for NHP S. Similarly, for NHP P, the number of units were (140, 77, 75) and (83, 67, 78).

C. Behavioral Task

The NHP's were trained to perform a reach-grasp-transport-release task [15,16] where aspects of a simulated anthropomorphic robotic arm were controlled using a physical grip-force transducer.

The task consisted of 6 scenes (as shown in Figure 1). 1) Cue Display: A colored cue was displayed intimating the animal about how much reward it was going to receive on successful completion of the trial, 2) Reaching: The simulated arm automatically reached for the target, 3) Grasping: The NHP squeezed the physical manipulandum to exert force. The amount of force exerted by the animal was shown visually using a rectangle which changed its area in proportion to the force exerted. The amount of force that the NHP should exert and maintain was shown using the blue rectangles which appeared on opposite sides of the red rectangle. If the animal exerts more force than necessary, the trial was considered failed. 4) Transporting: The simulated arm automatically moved to a target destination, seen as a pink target.
It is this part of the trial where the animal needed to maintain its exerted force within the blue rectangles. 5) Releasing: The animal released its grip on the force transducer when the robotic arm reached the pink target. 6) Reward: The animal received 0,1,2 or 3 juice delivery periods, each 0.5 s in duration, depending on the number of green squares shown during the cuing period.

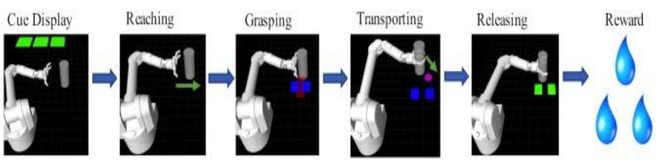

Figure 1: Visual depiction of the 6 stages of the reach-grasp-transport-release task.

D. Neural Network Architecture

The neural network used to predict grip force from neural spiking activity utilized a simple feedforward architecture with one hidden layer with 50 units [17]. The input to the network was the neural activity binned at 50 ms and fully connected to the hidden layer. The activation function used was the Rectified Linear Unit (ReLU) [18]. The output was the predicted grip force from the grip-force task. Since the aim was to predict the grip force, we only use the neural activity corresponding to the part of the trial when the NHP exerts force on the force sensor, i.e. between scenes 3 and 4.

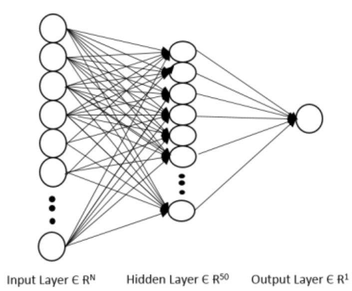

Figure 2: Architecture of a simple fully connected neural network with input layer size of N where N is the number of units. The hidden layer is of size 50 and an output layer with the output force at a time bin.

E. Reward structures

As described previously, the animal received a reward of 0, 1, 2 or 3 juice delivery periods depending on what was initially shown to it during the Cue Display. To denote the reward level of a particular bin of data in a trial, we use two kinds of reward structures 1) Non-Linear, where the reward values were [-10, 10, 20, 30], 2) Linear, with reward values [10, 20, 30, 40] corresponding to number of juice delivery periods that would be obtained by the NHP. The reason behind using a Linear structure is very straightforward as each additional delivery period might corresponds to a linear increase in subjective reward-value, however the rationale behind using the Non-Linear reward was to separate the Non-Reward condition (0 juice delivery periods) from the other conditions, where at least one drop of juice would be delivered, as R0 might represent punishment.

F. Loss Functions

We compare the performance of the neural network under two loss functions defined as

$$MSE = (X-D)^2 \qquad (1)$$

$$RP\_MSE = (X-D)^2 + \beta\,(D.\boldsymbol{R})^2 \qquad (2)$$

Where X is an array corresponding to the actual grip force, D is the predicted grip force, $\boldsymbol{R}$ is an array with the reward level values associated with the corresponding grip force values in D and β is a constant. *MSE* is the standard mean squared error function and *RP_MSE* is the reward correlation penalized loss function where ($D.\boldsymbol{R}$) refers to the correlation coefficient between the predicted grip force and the reward level.

III. RESULTS

The datasets consisted of neural activity binned at 50 ms (the firing rate of the unit within a 50 ms period) and the corresponding grip force value exerted by the NHP which is a normalized value. The neural data was analyzed from three brain regions: Dorsal Premotor Cortex (PMd), Primary motor cortex (M1), and Primary Somatosensory cortex (S1). The grip force was also predicted from all the regions combined. Each of the datasets was divided into training, validation and testing sets with the percentage ratio of 70:15:15.

The performance of the two loss functions was compared using the R^2 scores in the test set. Specifically, the neural network was run 25 times for each β value from 0 to 10000 and the median R^2 scores on the test set were compared using the Wilcoxin signed rank test to check for significant improvement in prediction. It should also be noted that cross-validation was performed for each run such that all data points were used in the test set and the median R^2 score was

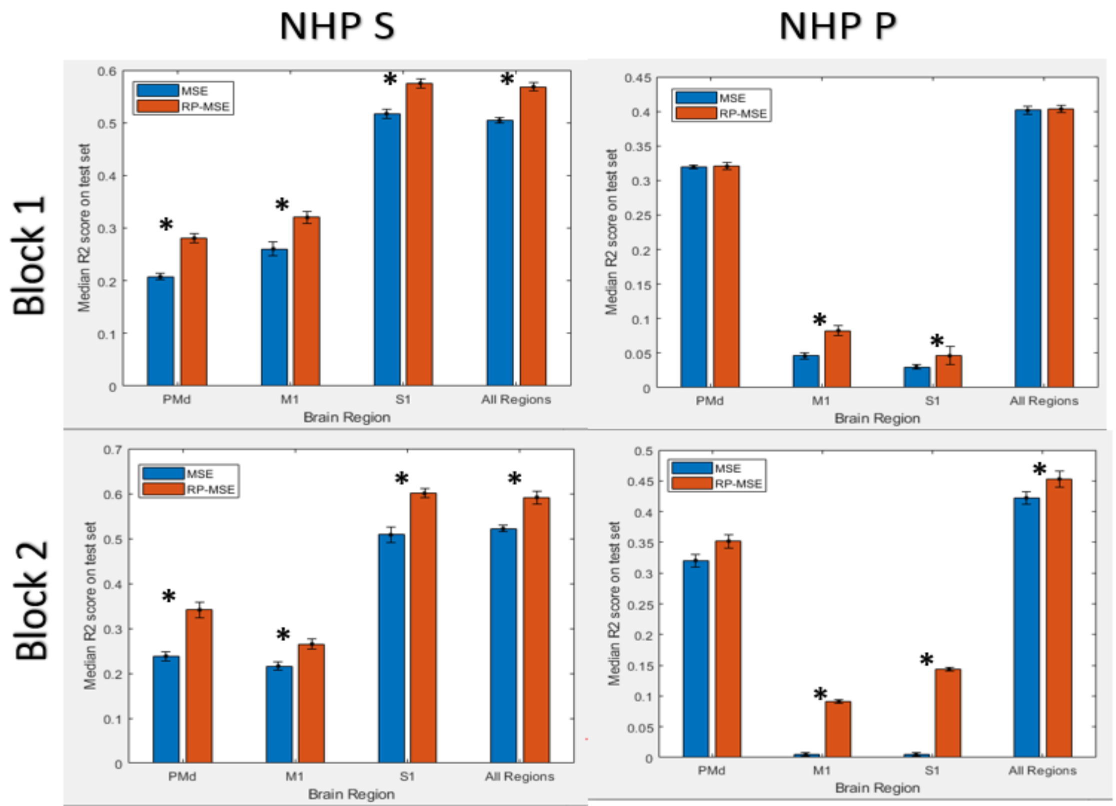

Figure 3: Comparison of median R^2 scores on test set using mean squared error (MSE) and reward penalized mean squared error (RP-MSE) loss functions. The top and bottom figures refer to Block 1 and 2 respectively. Block refers to one session and the NHP gets a break of 5 mins between blocks. The median R^2 scoreds for RP-MSE are from the best β values ([500,1000,2000,500] for NHP S and [1000,2000,500,500] for NHP P where each value refers to the best β value for PMd, M1 and S1 respectively). The black error bars are the absolute deviations around the median R^2 scores.

chosen as the metric of comparison between the two loss functions. The performance of RP_MSE with the best β value is plotted in Figure 2. Our results clearly show that 1) The neural network with reward correlation penalization significantly ($p<0.01$) outperformed the standard MSE by approximately 6% in all three brain regions across both NHP's except for PMd in NHP P. 2) S1 shows the best performance among the three regions for NHP S and PMd shows the highest performance for NHP P. 4) Combining all the regions together to predict yielded slightly higher performance than the best predicting regions. 5) There are block level differences in performances in both NHP's. 6) Both linear and non-linear reward structures show similar performance.

IV. DISCUSSION

We tested and proved our hypothesis that a simple neural network can predict grip force significantly better from neural activity when the reward-correlation is penalized. There are however some limitations in our work that need to be addressed.

First, the linear and non-linear reward structures that we used may not be the way the brain encodes reward values. Our previous work has shown the existence of a divisive normalized relationship between neural activity and reward [16]. Also, the dopamine system in the brain has been shown to represent a belief state system [19]. A belief state system could also explain the subject wise differences in this study as well as our previous studies [8,15,16]. Once such a belief state can be incorporated to a decoding model, optimal reward structures can be elucidated and subsequently penalized in a similar way as shown in our work hopefully improving the performance of a decoding network.

Second, in our work, each time bin is assumed to be independent of the others. However, it has been shown that neural activity evolves in a time dependent manner and temporal sequence of reward history plays an important role in predicting current reward [20]. This time dependent activity can also occur across different timescales [21]. Complex neural network architectures such as LSTM's [22] automatically account for the time dependencies between the input and output and may boost the performance of the neural network.

Third, it has been shown that a given area of the brain rarely acts in isolation and that the different brain areas interact with one another during vision, movement, and other cognitive tasks [13,14]. Interactions between different brain regions is likely to be modulated by reward. This reward modulation of the interaction could further help predict grip force more accurately.

Although our analysis has shown only a 6% improvement in performance, it is still a step towards designing decoders for BMI’s that utilize ANNs along with the reward encoding capacity of the sensorimotor cortices (PMd, M1 and S1). Such a paradigm shift in BMI decoder design has potential for greater clinical relevance in the future [15].

V. CONCLUSION

From our work, it can reasonably be concluded that reward affects the prediction of grip force from neural activity and that this effect can be minimized by penalizing the reward correlation. Future work should incorporate belief state about rewards, time dependencies in neural activity and the interaction between different brain areas to improve predictive performance of neural networks from a BMI perspective.

ACKNOWLEDGEMENT

The authors thank Dr. Yao Zhao for collection of the datasets used in this analysis.

REFERENCES

[1] Wessberg, J., Stambaugh, C. R., Kralik, J. D., Beck, P. D., Laubach, M., Chapin, J. K., ... & Nicolelis, M. A. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature, 408(6810), 361-365.doi:10.3389/fnins.2018.00579W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[2] Carmena, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., ... & Segev, I. (2003). Learning to control a brain–machine interface for reaching and grasping by primates. PLoS biology, 1(2), e42.

[3] Rouse, A. G., Stanslaski, S. R., Cong, P., Jensen, R. M., Afshar, P., Ullestad, D., ... & Denison, T. J. (2011). A chronic generalized bi-directional brain–machine interface. Journal of neural engineering, 8(3), 036018.

[4] Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. Science 275, 1593–1599. doi: 10.1126/science.275.5306.1593

[5] Schultz, W. (2017). Reward prediction error. Current Biology, 27(10), R369-R371.

[6] Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron, 29(2), 537-545.

[7] Lee, W., & Reeve, J. (2017). Identifying the neural substrates of intrinsic motivation during task performance. Cognitive, Affective, & Behavioral Neuroscience, 17(5), 939-953.

[8] Marsh, B. T., Tarigoppula, V. S., Chen, C., and Francis, J. T. (2015). Toward an autonomous brain machine interface: integrating sensorimotor reward modulation and reinforcement learning. J. Neurosci. 35, 7374–7387. doi: 10.1523/JNEUROSCI.1802-14.2015

[9] Tarigoppula, V. S. A., Choi, J. S., Hessburg, J. H., McNiel, D. B., Marsh, B. T., and Francis, J. T. (2018). Primary motor cortex encodes a temporal difference reinforcement learning process. bioRxiv [Preprint]. doi: 10.1101/257337

[10] Ramkumar, P., Dekleva, B., Cooler, S., Miller, L., and Kording, K. (2016). Premotor and motor cortices encode reward. PLoS One 11:e0160851. doi: 10.1371/journal.pone.0160851

[11] Ramakrishnan, A., Byun, Y. W., Rand, K., Pedersen, C. E., Lebedev, M. A., and Nicolelis, M. A. L. (2017). Cortical neurons multiplex reward-related signals along with sensory and motor information. Proc. Natl. Acad. Sci. U S A 114, E4841–E4850. doi: 10.1073/pnas.1703668114

[12] An, J., Yadav, T., Hessburg, J. P., and Francis, J. T. (2019). Reward expectation modulates local field potentials, spiking activity and spike-field coherence in the primary motor cortex. eNeuro 6:ENEURO.0178-19.2019. doi: 10.1523/ENEURO.0178-19.2019

[13] An, Junmo, et al. "Near perfect neural critic from motor cortical activity toward an autonomously updating brain machine interface." 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2018Babayan, Benedicte M., Naoshige Uchida, and Samuel J. Gershman. "Belief state representation in the dopamine system." Nature communications 9.1 (2018): 1-10.

[14] Moore, B., Khang, S., & Francis, J. T. (2020). Noise-Correlation Is Modulated by Reward Expectation in the Primary Motor Cortex Bilaterally During Manual and Observational Tasks in Primates. Frontiers in Behavioral Neuroscience, 192.

[15] Zhao Y, Hessburg JP, Asok Kumar JN, Francis JT. Paradigm Shift in Sensorimotor Control Research and Brain Machine Interface Control: The Influence of Context on Sensorimotor Representations. Front Neurosci. 2018;12:579. Published 2018 Sep 10.

[16] Yao, Zhao, John P. Hessburg, and Joseph Thachil Francis. "Normalization by valence and motivational intensity in the sensorimotor cortices (PMd, M1, and S1)." Scientific reports 11.1 (2021): 1-19.

[17] Caselli, M., et al. "A simple feedforward neural network for the PM10 forecasting: comparison with a radial basis function network and a multivariate linear regression model." Water, Air, and Soil Pollution 201.1 (2009): 365-377.

[18] Agarap, Abien Fred. "Deep learning using rectified linear units (relu)." arXiv preprint arXiv:1803.08375 (2018).

[19] Babayan, Benedicte M., Naoshige Uchida, and Samuel J. Gershman. "Belief state representation in the dopamine system." Nature communications 9.1 (2018): 1-10.

[20] Bayer, Hannah M., and Paul W. Glimcher. "Midbrain dopamine neurons encode a quantitative reward prediction error signal." Neuron 47.1 (2005): 129-141.

[21] Williams, Alex H., et al. "Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis." Neuron 98.6 (2018): 1099-1115.

[22] Tseng, Po-He, et al. "Decoding movements from cortical ensemble activity using a long short-term memory recurrent network." Neural Computation 31.6 (2019): 1085-1113