STOP STATE CLASSIFICATION IN INTRACORTICAL BRAIN-MACHINE-INTERFACE

Tze Hui Koh, Camilo Libedinsky, Cuntai Guan, Kai Keng Ang, Rosa Q. So

Abstract—**Invasive brain-machine-interface (BMI) has the prospect to empower tetraplegic patients with independent mobility through the use of brain-controlled wheelchairs. For the practical and long-term use of such control systems, the system has to distinguish between stop and movement states and has to be robust to overcome non-stationarity in the brain signals. In this work, we investigates the non-stationarity of the stop state on neural data collected from a macaque trained to control a robotic platform to stop and move in left, right, forward directions We then propose a hybrid approach that employs both random forest and linear discriminant analysis (LDA). Using this approach, we performed offline decoding on 8 days of data collected over the course of three months during joystick control of the robotic platform. We compared the results of using the proposed approach with the use of LDA alone to perform direct classifications of stop, left, right and forward. The results showed an average performance increment of 22.7% using the proposed hybrid approach. The results yielded significant improvements during sessions where LDA showed a heavy bias towards the stop state. This suggests that the proposed hybrid approach addresses the non-stationarity in the stop state and subsequently facilitates a more accurate decoding of the movement states.**

I. INTRODUCTION

Invasive Brain-Machine-Interfaces (BMI) are able to translate neural data into control signals with good spatial resolution and high signal-to-noise ratio, empowering tetraplegic patients with independent mobility through a BMI wheelchair control system [1]. Such systems have been attempted using electroencephalography (EEG), however performance has been limited due to the poor spatial resolution and low signal-to-noise ratio of EEG [2-4].

For a BMI wheelchair control system to be viable in a daily setting, it is important for the system to be robust and not require constant retraining over long periods of time [5]. Moreover, it is essential that the system has the capacity to distinguish between the stop state and movement state, such that the different movement states may be differentiated more accurately. This isolation of the stop state is necessary due to its inherent instability and tendency to drift across days [6].

In this paper, a decoding algorithm which integrates Random Forest (RF) and Linear Discriminant Analysis (LDA) classifiers is proposed to perform the aforementioned tasks. This approach was applied on offline Neural signals from a macaque trained to move a robotic platform using a 3-direction joystick. Results suggest that in cases where the LDA classifier becomes biased towards the stop class, RF remains unbiased and provides a more accurate distinction between stop and movement.

II. METHODS

A. Data Acquisition

All procedures and experiments were approved and conducted in accordance with the standards of the Singapore Health Services Institutional Animal Care and Use Committee (Singhealth IACUC #2012/SHS/757). Neural signals were acquired from a young male adult macaque (*Macaca fascicularis*) implanted with arrays of intracortical microelectrodes (96 channels) in the hand/arm area of the primary motor cortex. Signals were acquired at a sampling rate of 13 kHz while the macaque performed the task of controlling a robotic platform, upon which it was seated, in the directions of left, right, forward by using a joystick controller, as well holding still for stop, forming four discrete classes. Further details of the experimental setup have been outlined in a prior work [7]. All analyses of neural signals were performed offline.

B. Preprocessing

Raw signals were band-pass filtered between 300 to 3000 Hz and spikes detected via threshold-crossing criterion selected for each channel [8]. No spike sorting was performed as studies have found no significant difference between spike sorting and thresholding [8, 9]. For each channel, firing rates were computed using an equally-weighted moving average with a kernel width of 500ms and time steps of Δt = 100ms. For classification, data collected from 8 days were utilized, spanning over the course of three months. The model was retrained for each day using data from the preceding three days in a causal manner e.g. days 1-3 data were used for training for testing on day 4 data, days 2-4 data were used for training for testing on day 5 data.

C. Classifier Design

To illustrate the necessity of separating the stop class from the rest, a four-class LDA decoder was compared to a three-class LDA decoder, with the stop class data removed

*Research supported by BMRC-EDB and JCO DP grants IAF311022

Tze Hui Koh, is with the Department of Neural and Biomedical Technology, Institute for Infocomm Research, Singapore 138632 (e-mail: kohth@i2r.a-star.edu.sg).

Camilo Libedinsky is with the Department of Psychology, NUS Singapore Institute for Neurotechnology, Singapore 769199 (email: camilo@nus.edu.sg)

Cuntai Guan was with the Department of Neural and Biomedical Technology, Institute for Infocomm Research, Singapore 138632, and is now with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ctguan@ntu.edu.sg).

Kai Keng Ang is with the Department of Neural and Biomedical Technology, Institute for Infocomm Research, Singapore 138632 (email: kkang@i2r.a-star.edu.sg)

Rosa Q. So is with the Department of Neural and Biomedical Technology, Institute for Infocomm Research, Singapore 138632 (corresponding author to provide phone: +6564082096 ; e-mail: rosa-so@i2r.a-star.edu.sg).

978-1-5090-2809-2/17/$31.00 ©2017 IEEE

from the three-class decoder. The resulting model was tested on only movement data. The kappa coefficient measures the agreement between two observers; in this case between the correct output and the decoder output. The kappa coefficient was chosen as a performance measure to account for chance performance [10], a common measure used in BCI [11]. As shown in Fig. 1, a more accurate decoding of left, forward, and right is achieved when the stop class is removed. The same test was conducted with the other three classes, and lower accuracies were achieved, further indicating the poorer performance of the classifier in distinguishing stop from the movement classes. The paired t-test was used to test for statistical significance in the comparison between the two decoders and a significant difference was found on days 2, 3 and 5 (p-value < 0.05) , with the poorest performance on day 3 due to the similar class covariance and mean, resulting in poor separation.

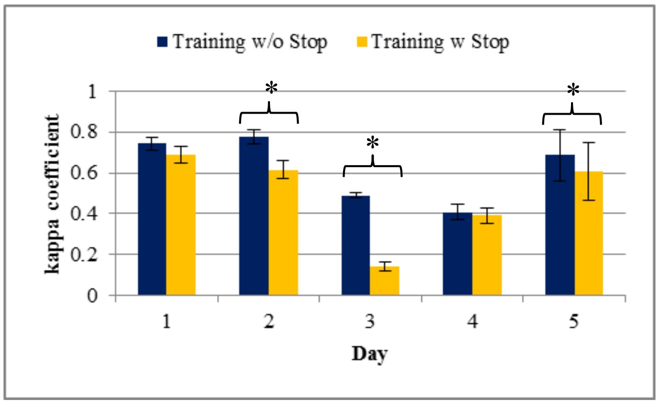

Figure 1. Performance of movement decoding without and with the inclusion of stop data using linear discriminant analysis.

The figure above suggests that it is the stop class that mainly adds non-stationarity, and hence the ability to distinguish between stop and movement classes is important for long-term use of the decoder. In view of this, we propose a hybrid RF+ LDA decoder. For performance comparison, a four-class LDA decoder was used, described in algorithm 1, as has been performed in classification studies of intracortical BMI [7, 12]. An LDA decoder searches for a linear combination of features for the best separation between classes. For the hybrid decoder, RF, which operates by fitting decision trees and averaging to improve accuracy and control over-fitting, was first used as a binary decoder to separate the stop and movement state, following which LDA was performed on data from the predicted movement states mainly to determine the directions left, right, and forward. As the amount of data increases, the heterogeneous nature of the stop data would lead to the biasing of linear decoders, as we can no longer expect the data to be linearly separable. In contrast to LDA, which has low variance and is prone to high bias [13, 14], RF, a non-linear classifier, is capable of achieving both low variance and low bias [14-16]. RF was therefore chosen as the classifier for discriminating between stop and movement. This algorithm is described in algorithm 2.

Algorithm 1: Linear Discriminant Analysis classifier algorithm
Input: A set of training data {X, α}

- $X = \{x_j^i\}_{i=0}^{D}, x_j^i \in \mathbb{R}^{T \times N}$: a set of single session firing rates for each channel, where D is the total number of sessions with N channels and T sample points.
- $\alpha = \{w_i\}_{i=0}^{D}$: true class labels at each sample point i, where $w_i \in \{0,90,180,270\}$.

Output: Predicted class labels P

- $P = \{z_j^i\}_{i=1}^{D}$: predicted class labels at each sample point, where $z_i \in \{0,90,180,270\}$.

The LDA algorithm is briefly described as follows [17]:

- Step 1: Using the training data, a 2Hz threshold is used to identify dropped channels and remove them, and the remaining channels vectorised and used as features.
- Step 2: Scatter between class defined by sample covariance of the estimated class means
- Step 3: Eigenvectors are computed for scatter matrices.
- Step 4: Maximum separation found by transforming firing rates with the eigenvectors and projecting onto a new subspace.

Algorithm 2: Hybrid RF+LDA classifier algorithm
Primary Input: A set of training data {X, α}

- $X = \{x_j^i\}_{i=0}^{D}, x_j^i \in \mathbb{R}^{T \times N}$: a set of single session firing rates for each channel, where D is the total number of sessions with N channels and T sample points.
- $\alpha = \{w_i\}_{i=0}^{D}$: true class labels at each sample point i, where $w_i \in \{0,270\}$.

Primary Output: Predicted class labels P

- $P = \{z_i\}_{i=0}^{D}$: predicted class labels at each sample point i, where $z_i \in \{0,270\}$.

Secondary Input: A set of training data {Y, β}:

- $Y = \{y_j^i\}_{i=0}^{D}, y_j^i \in \mathbb{R}^{S \times N}$: a set of single session firing rates for each channel, where D is the total number of sessions with N channels and S sample points, where S are the sample points with $P = \{0\}$.
- $\beta = \{w_j^i\}_{i=1}^{D}$: true class labels at each sample point within new truncated space, where $w_i \in \{0,90,180,270\}$.

Secondary Output: Predicted class labels Q

- $Q = \{z_i\}_{i=0}^{D}$: predicted class labels at each sample point i, where $z_i \in \{0,90,180,270\}$.

The hybrid algorithm is briefly described as follows:

- Step 1: Using the training data, a 2Hz threshold is used to remove channels with firing rates that fall under the threshold, and the remaining channels vectorised and used as features.

- Step 2: Random forest constructs a multitude of decision trees randomly restricted to be sensitive to particular parameters [18].
- Step 3: Mode of the classes predicted by individual trees is used as the predicted class for each sample point, forming the primary output [18].
- Step 4: Data corresponding to predicted movement classes is used as input for an LDA classifier
- Step 5: Through the steps mentioned in algorithm 1, left, right, forward are predicted with the occasional stop, forming the secondary output.
- Step 6: The primary and secondary outputs are combined to give predictions for all T sample points.

These algorithms were performed in an offline analysis of data collected from the macaque. Preprocessing was done in Matlab (Mathworks Inc, Massachusetts, USA) and classifiers were implemented in Python 3.

III. Results

The results are structured in two parts. First, we show a comparison between RF and LDA in acting as binary classifiers for distinguishing between stop and movement. Second, we present the results of the implemented hybrid RF+LDA decoder, using LDA as the baseline for comparison, demonstrating that in cases where LDA is biased towards the stop class, the addition of RF as a preliminary binary decoder may help to improve performance. Paired t-test was used to test for statistical significance in all comparisons and indicated with an asterisk where significant (p-value < 0.05):

A. *Performance Comparisons of Random Forest and Linear Discriminant Analysis for Binary Classification*

Percentage accuracy of the predictions was used as a performance measure in evaluating RF and LDA as binary classifiers. For comparison the dataset was resampled such that the data for both stop and movement classes were balanced. Chance performance for decoding accuracy is 50%. We found that the performance of RF was comparable or higher than that of LDA across five experimental days, as seen in Fig. 2, suggesting that RF is a better choice for distinguishing between the two classes.

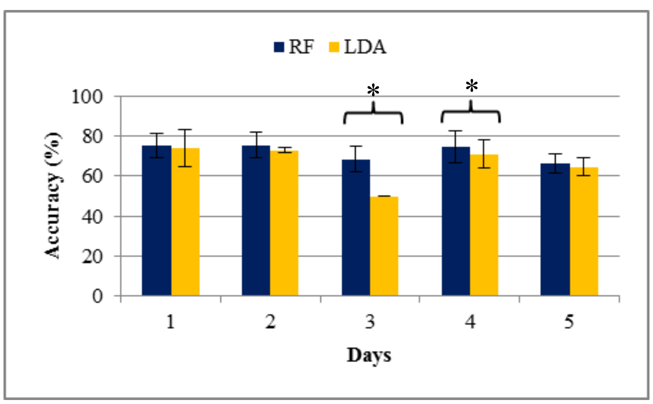

Figure 2. Performance of RF and LDA as binary classifiers distinguishing between stop and movement states.

B. *Performance Comparisons of LDA+RF and LDA as Four-class Decoders*

We further evaluated that a hybrid of LDA+RF could provide a more accurate decoding for entire experimental sessions. The data were resampled such that the four classes were balanced. Chance performance for decoding accuracy is 25%. Across the five experimental days, we found that the hybrid decoder could give a higher performance on experimental day 3, illustrated in Fig.3, with average performance of 59.8% (± 3.13) for the hybrid decoder and 37.1% (± 3.08) for the LDA decoder. This improvement in decoding corresponded with a significant improvement in RF performance in distinguishing between stop and movement (Fig. 2).

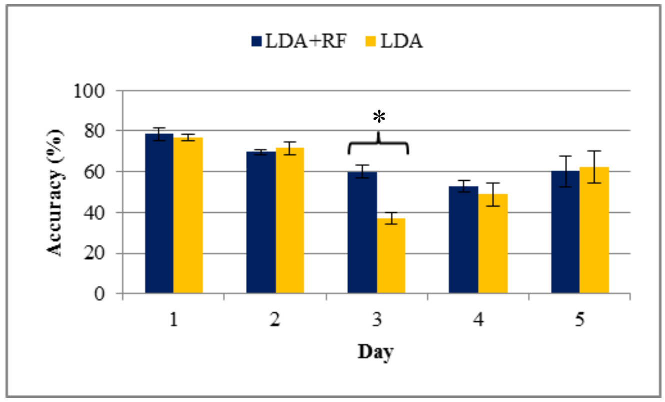

Figure 3. Performance of hybrid random forest + linear discriminant analysis decoder as compared to linear discriminant analysis in four class decoding.

Examining the specific breakdown of decoded directions via normalized confusion matrices, as shown in Fig. 4A and 4B, we found that performance improvements were seen in cases where the LDA was initially biased towards the stop class. In such cases RF remained unbiased and subsequently LDA could distinguish more accurately between left and right in the second step of the hybrid decoder.

However, in cases where the bias did not lie with the stop class, the hybrid decoder performed at a comparable level to LDA, providing no significant improvement, as seen in Fig. 4C and 4D.

IV. Discussion

Our results suggest the binary classification of stop and movement before further classification has the potential to increase the performance of the decoding, when using decoding models built from data collected from previous days. The problem of non-stationarity, especially in the stop class, was addressed in part by using a hybrid decoder approach. This study represents the significance of using a separate decoder for the isolation of the stop class to achieve better decoding for the movement class. In the absence of an overt movement during the stop task, greater diversity was observed in recorded neural signals, leading to drift across experimental sessions, reducing decoding performance [6].

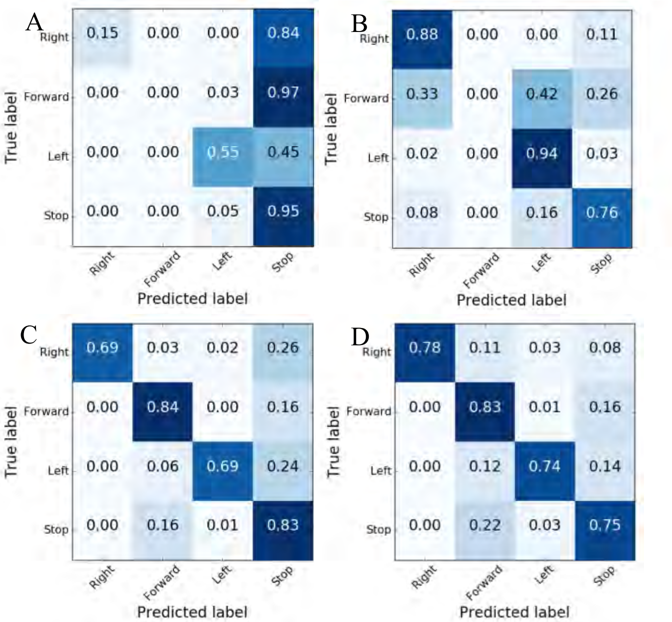

Figure 4. Sample normalized confusion matrices measuring true labels against predicted labels. A) LDA decoding for session 15111601. B) RF+LDA decoding for session 15111601. C) LDA decoding for session 15102606. D) RF+LDA decoding for session 15102606.

A possible reason for the higher accuracy of RF during the binary classification is the non-linearity of the model, and the robustness of the algorithm against bias [15, 16]. The diverse nature of the stop class across sessions and days would lead to a heterogenous dataset that a linear classifier such as LDA may under-fit. This framework of a hybrid decoder may therefore be further refined using non-linear decoders with higher capability of distinguishing between stop and movement states. In an ideal situation, the LDA in the second step would be completely reduced to a three-class classifier with no false negatives for the movement class, resulting in greater improvements in performance, as suggested in Fig. 1.

Improved decoding of the stop vs. movement classes reduces effects of non-stationarity in the stop class, allowing for a lower frequency of retraining, and a more practical implementation of the wheelchair control system in daily life.

V. Conclusion

For a brain-controlled wheelchair to be compatible with activities of daily living, it is essential that the decoding algorithm can distinguish between active and inactive states with high accuracy with low incidence of false positives. To address this need, we report a hybrid method that has the potential to increase the performance of the decoding. This study represents one approach in addressing the non-stationarity of the stop state during classification for a wheelchair application. Further investigations should be carried out on identifying a model which best captures the distinction between movement and stop states, and bringing this method online for testing in a closed-loop experiment, considering the potential added computational complexity.

Acknowledgment

The authors would like to thank Clement Lim for helping with animal training and data collection.

References

[1] J. L. Collinger, S. Foldes, T. M. Bruns, B. Wodlinger, R. Gaunt, and D. J. Weber, "Neuroprosthetic technology for individuals with spinal cord injury," *The Journal of Spinal Cord Medicine,* vol. 36, no. 4, pp. 258-272, 2013.

[2] J. d. R. Millán, F. Galán, D. Vanhooydonck, E. Lew, J. Philips, and M. Nuttin, "Asynchronous non-invasive brain-actuated control of an intelligent wheelchair," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, 2009, pp. 3361-3364: IEEE.

[3] B. Rebsamen *et al.*, "A brain controlled wheelchair to navigate in familiar environments," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 18, no. 6, pp. 590-598, 2010.

[4] I. Iturrate, J. M. Antelis, A. Kubler, and J. Minguez, "A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation," *IEEE Transactions on Robotics,* vol. 25, no. 3, pp. 614-627, 2009.

[5] D. Zhang, Y. Wang, X. Gao, B. Hong, and S. Gao, "An algorithm for idle-state detection in motor-imagery-based brain-computer interface," *Computational Intelligence and Neuroscience,* vol. 2007, pp. 5-5, 2007.

[6] Z. Xu, R. Q. So, K. K. Toe, K. K. Ang, and C. Guan, "On the asynchronously continuous control of mobile robot movement by motor cortical spiking activity," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 3049-3052: IEEE.

[7] C. Libedinsky *et al.*, "Independent mobility achieved through a wireless brain-machine interface," *PloS One,* vol. 11, no. 11, p. e0165773, 2016.

[8] R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural Computation,* vol. 16, no. 8, pp. 1661-1687, 2004.

[9] B. P. Christie *et al.*, "Comparison of spike sorting and thresholding of voltage waveforms for intracortical brain–machine interface performance," *Journal of Neural Engineering,* vol. 12, no. 1, p. 016009, 2014.

[10] A. J. Viera and J. M. Garrett, "Understanding interobserver agreement: the kappa statistic," *Family Medicine,* vol. 37, no. 5, pp. 360-363, 2005.

[11] A. Schlogl, J. Kronegg, J. Huggins, and S. Mason, "Evaluation Criteria for BCI Research," in *Toward Brain-Computer Interfacing*: The MIT Press, 2007, pp. 327-342.

[12] D. Bacher *et al.*, "Neural point-and-click communication by a person with incomplete locked-in syndrome," *Neurorehabilitation and Neural Repair,* vol. 29, no. 5, pp. 462-471, 2015.

[13] L. Breiman, "Bias, variance, and arcing classifiers," 1996.

[14] J. Schiffner, B. Bischl, and C. Weihs, "Bias-variance analysis of local classification methods," in *Challenges at the Interface of Data Analysis, Computer Science, and Optimization*: Springer, 2012, pp. 49-57.

[15] R. Díaz-Uriarte and S. A. De Andres, "Gene selection and classification of microarray data using random forest," *BMC Bioinformatics,* vol. 7, no. 1, p. 3, 2006.

[16] L. Breiman, "Random forests," *Machine Learning,* vol. 45, no. 1, pp. 5-32, 2001.

[17] G. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. John Wiley & Sons, 2004.

[18] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News,* vol. 2, no. 3, pp. 18-22, 2002.