fNIRS as a Quantitative tool to Asses and Predict Surgical Skills

Yuanyuan Gao[1], Pingkun Yan[2], Suvranu De[1,2] and Xavier Intes[2]
1. Center for Modeling, Simulation and Imaging in Medicine, Rensselaer Polytechnic Institute, Troy, New York, USA
2. Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, New York, USA
Author e-mail address: des@rpi.edu, intesx@rpi.edu

Abstract: We report on the application of fNIRs and derived metrics to assess and predict surgical skills within the framework of the Fundamentals of Laparoscopic Surgery.
OCIS codes: 200.4260

1. Introduction

Recently, functional near-infrared spectroscopy (fNIRS) has been shown to monitor human motor performance and broadly utilized to predict motor tasks or workload in various settings [1]. For example, fNIRS data can be used to classify simple and complex finger tapping tasks [2]. fNIRS also showed its capacity to classify left vs. right hand or wrist motor execution or imagery [3, 4], and different motor activities [5]. However, to date, a limited number of preliminary studies have been reported in regards to assessing fine surgical motor skills using fNIRS [6]. Here we will report on our efforts in establishing fNIRS as a robust and quantitative tools to assess and predict surgical skills. Especially, we have recently established that fNIRS enables to retrospectively classify subject surgical skill levels, showing its potential utility of determining the surgical performance within the Fundamentals of Laparoscopic Surgery (FLS) program accredited framework [7]. Additionally, to further investigate the potential of fNIRS as a non-invasive, objective, quantitative methodology for surgical skills assessment in near-real time, we developed a convolutional neural network (CNN) model called "Brain-NET". Based on this model, we further examined which biomarker of fNIRS signal contributed the most to this model using a 'knock-one-out' scheme and demonstrate its potential to predict the FLS-score which is currently employed in surgical certification.

2. Methods and results

fNIRS data were collected on medical students over an eleven-day training period while they were performing the bimanual PC task, which is part of the FLS training program and certification examination. Thirteen healthy subjects (mean age ±SD, 25.2± 3.8, 22 to 34 years old) participated in this study. A written and dated informed consent was obtained from each subject. The selected FLS precision cutting surgical task consists in cutting a gauze following a marked circle as accurately and quickly as possible. The performance score of each trial was calculated using the accredited FLS scoring methodology with consent under a nondisclosure agreement from the FLS Committee. The fNIRS signals were measured simultaneously during each FLS PC trial by a continuous-wave near-infrared spectrometer (CW6 system, Techen Inc., MA, USA) from six locations on the prefrontal cortex (PFC). The acquired optical intensity time courses were preprocessed using the publicly available Homer2 software to obtain the oxy- and deoxy- hemoglobin cortical response (HbO and HbR respectively). We extracted seven features, namely the five first moments (mean, variance, slope, skewness, kurtosis) as well as the range (maximum and minimum) for both these biomarkers [1]. After this procedure, we remained with a 2D input matrix containing fourteen hemoglobin data features from six PFC locations. This data processing flow leaded to a total sample size of 702 for the fNIRS data.

To gauge if these spatially resolved 7 features of HbO and HbR can predict the FLS score, we developed a CNN architecture, Brain-NET, specially designed for this application. The input matrix is expressed as X shaped as ($l \times f$), where l is the dimension of the PFC locations and f is the dimension of the hemoglobin features. In this case $l = 6$ and $f = 14$. There are two kinds of relationships embedded in the input matrix X that we want to extract. The relationships within hemoglobin feature domain and interactions across the PFC locations. Therefore, we apply 1D convolutional kernels along the feature direction and along the PFC location direction hierarchically to learn the relationships, with the kernel size of 1 and the stride of 1. For each convolutional layer, Brain-NET learns 8 filters, and uses ReLU as the activation function. Then we flatten the output derived from the first two CNN layers into a vector and apply two fully connected neural network layers with ReLU activation functions. The output is one value which is the predicted FLS score. To train our Brain-NET model, we use the mean squared error as our objective loss function and stochastic gradient descent as optimization method. The learning rate is set to 0.0001. We select a ten-fold cross validation scheme. We further randomly select 30 samples from training set each time for training performance validation. The value of coefficient of determination (denoted by R^2) is used to quantify the accuracy of the proposed methods.

Next, to investigate if all the seven extracted features are required to regress the motor performance levels, we adopted a "knock-one-out" scheme in which one feature across all locations are removed before training the model. If the R^2 value does not change or increases, it means that the knocked out feature does not contribute or even contributes negatives, which may be noise in this case. On the other hand, if the R^2 value decreases, it means that the knocked out feature contributes to the model. The larger drop of the performance, the higher the contribution of that feature. As depicted in Fig. 1, no matter which of the seven features for HbO and HbR is knocked out, the value of coefficient of determination (denoted by R^2) decreases. Hence, it demonstrates that all the features contribute to the model. Among all the features, knocking out the feature of HbR slope leads to the lowest R^2 value, indicating that it contributes the most to our model. Even though conventionally, HbO is the typical biomarker employed in fNIRS study due to its robust nature, HbR has been shown to be a strong indicator of motor or language stroke diseases [9, 10]. Additionally, the slope feature has also been identified as a discriminatory feature in several prior NIRS-BCI studies [11–13]. Hence, these findings are corroborated by previous studies even though they have been performed in different settings. We will present further data demonstrating the ability of Brain-Net in predicting the FLS score for certification classification purposes (ROC-AUC=0.91).

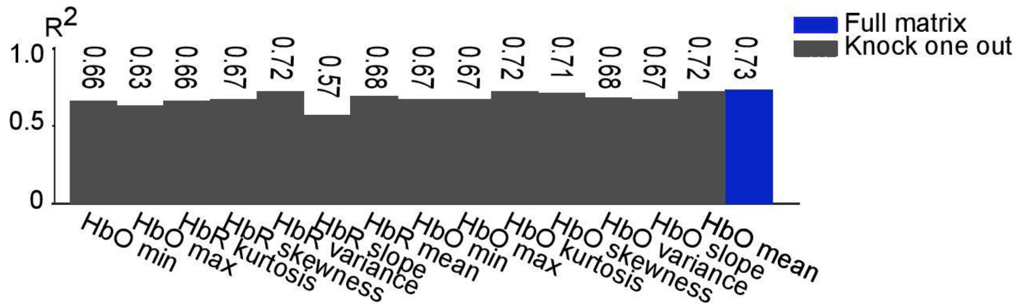

Fig 1 Full matrix of the R2 values when each feature (1 out of 7x2) is knocked out.

3. Funding and Acknowledgement

This work was supported by funding provided by NIH/National Institute of Biomedical Imaging and Bioengineering grants 2R01EB005807, 5R01EB010037, 1R01EB009362, 1R01EB014305, and R01EB019443. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU and Quadra P6000.

4. References

[1] N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: a review," *Front. Hum. Neurosci.*, vol. 9, no. January, pp. 1–15, 2015.

[2] L. Holper and M. Wolf, "Single-trial classification of motor imagery differing in task complexity: a functional near-infrared spectroscopy study," *J. Neuroeng. Rehabil.*, vol. 8, no. 1, pp. 1–13, 2011.

[3] S. Fazli *et al.*, "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *Neuroimage*, vol. 59, no. 1, pp. 519–529, 2012.

[4] N. Naseer and K. S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface," *Neurosci. Lett.*, vol. 553, pp. 84–89, 2013.

[5] J. Shin and J. Jeong, "Multiclass classification of hemodynamic responses for performance improvement of functional near-infrared spectroscopy-based brain–computer interface," *J. Biomed. Opt.*, vol. 19, no. 6, p. 067009, 2014.

[6] H. N. Modi, H. Singh, G. Z. Yang, A. Darzi, and D. R. Leff, "A decade of imaging surgeons' brain function (part I): terminology, techniques, and clinical translation," *Surg. (United States)*, no. part I, pp. 1–10, 2017.

[7] A. Nemani *et al.*, "Assessing bimanual motor skills with optical neuroimaging," *Sci. Adv.*, vol. 4, no. 10, pp. 1–10, 2018.

[8] J. M. Chalovich and E. Eisenberg, "Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling," *Magn Reson Imaging*, vol. 31, no. 3, pp. 477–479, 2013.

[9] K. Sakatani, Y. Xie, W. Lichty, S. Li, and H. Zuo, "Language-activated cerebral blood oxygenation and hemodynamic changes of the left prefrontal cortex in poststroke aphasic patients: a near-infrared spectroscopy study," *Stroke*, vol. 29, no. 7, pp. 1299–1304, 1998.

[10] Y. Murata, K. Sakatani, Y. Katayama, and C. Fukaya, "Increase in focal concentration of deoxyhaemoglobin during neuronal activity in cerebral ischaemic patients," *J. Neurol. Neurosurg. Psychiatry*, vol. 73, no. 2, pp. 182–184, 2002.

[11] S. D. Power, T. H. Falk, and T. Chau, "Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy," *J. Neural Eng.*, vol. 7, no. 2, p. 026002, 2010.

[12] S. D. Power, A. Kushki, and T. Chau, "Towards a system-paced near-infrared spectroscopy brain-computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state," *J. Neural Eng.*, vol. 8, no. 6, p. 066004, 2011.

[13] A. Faress and T. Chau, "Towards a multimodal brain-computer interface: combining fNIRS and fTCD measurements to enable higher classification accuracy," *Neuroimage*, vol. 77, pp. 186–194, 2013.