Selective Transfer Learning for EEG-Based Drowsiness Detection

Chun-Shu Wei, Yuan-Pin Lin, Yu-Te Wang and Tzyy-Ping Jung
Institute for Neural Computation and Institute of Engineering in Medicine
University of California San Diego
La Jolla, California, USA
{cswei, yplin, yute, jung} @sccn.ucsd.edu

Nima Bigdely-Shamlo
Syntrogi, Inc.
San Diego, California, USA
nima.bigdely@syntrogi.com

Chin-Teng Lin
Brain Research Center and Department of Electrical Engineering
National Chiao Tung University
Hsinchu, Taiwan
ctlin@mail.nctu.edu.tw

***Abstract*— On the pathway from laboratory settings to real world environment, a major challenge on the development of a robust electroencephalogram (EEG)-based brain-computer interface (BCI) is to collect a significant amount of informative training data from each individual, which is labor intensive and time-consuming and thereby significantly hinders the applications of BCIs in real-world settings. A possible remedy for this problem is to leverage existing data from other subjects. However, substantial inter-subject variability of human EEG data could deteriorate more than improve the BCI performance. This study proposes a new transfer learning (TL)-based method that exploits a subject's pilot data to select auxiliary data from other subjects to enhance the performance of an EEG-based BCI for drowsiness detection. This method is based on our previous findings that the EEG correlates of drowsiness were stable within individuals across sessions and an individual's pilot data could be used as calibration/training data to build a robust drowsiness detector. Empirical results of this study suggested that the feasibility of leveraging existing BCI models built by other subjects' data and a relatively small amount of subject-specific pilot data to develop a BCI that can outperform the BCI based solely on the pilot data of the subject.**

Keywords-EEG; brain-computer interface; transfer learning

I. Introduction

Recent progress in brain-computer interface (BCI) has been made in a great variety of applications [1]. Electroencephalogram (EEG)-based BCIs have begun to seek real-life applications. For example, several BCI studies have proved the feasibility and practicability of detecting an individual's drowsiness level using spontaneous EEG activities [2-4]. Nevertheless, promising results of the offline prediction were just a first step on the pathway toward real-world applications.

A major challenge in moving BCIs from well-controlled laboratory settings to real-world environments is that most of the BCIs require a significant amount of training data to build an accurate and robust model for each individual, which is labor-intensive and time-consuming and thereby significantly hinders the applications of BCIs in real-world settings.

Taking EEG-based drowsiness estimation as an example, the pilot data collecting session could be very long and tedious because the pilot session for each individual must contain a representative variety of drowsiness levels. It is thus imperative to develop a method to reduce the amount of training data needed from each individual for drowsiness detection BCI. An obvious alternative is to leverage existing data from other subjects. However, substantial inter-subject variability in human EEG could be an enduring obstacle for building a robust and capable EEG-based BCI from other individuals' data [5].

Our previous studies have shown that changes in EEG spectra are highly correlated with minute-scale fluctuations of a global level of drowsiness, indexed by the sustained-attention performance [2-4]. Furthermore, this relationship is stable within individuals across sessions, but variable across individuals [2]. Thus, blindly training a BCI with all existing data from different individuals could deteriorate more than improve its performance. Fortunately, as the relationship between the EEG spectra and the drowsiness level are relatively stable across sessions within an individual, it might be possible to select informative data (or positive samples/sessions) from other individuals based on the EEG-drowsiness correlation seen in the pilot data from the test individual.

This study thus investigates the feasibility of transferring the knowledge of existing data to enhance the performance of EEG-based BCI. Specifically, this study proposes a framework of selective transfer learning (TL) to exploit a test subject's pilot session to select auxiliary models from other subjects to build a more accurate and robust drowsiness-detection BCI. The results of the proposed method are compared to that of a conventional within-subject cross-session validation approach in drowsiness detection and a routine, *i.e.*, non-selective, TL.

II. Material and Methods

A. Experiment and Participants

A lane-keeping driving task [6] was adopted to study the EEG correlates of human cognitive state. During the

978-1-4799-8697-2/15 $31.00 © 2015 IEEE
DOI 10.1109/SMC.2015.560

Authorized licensed use limited to: Jinan University. Downloaded on August 06,2020 at 01:51:52 UTC from IEEE Xplore. Restrictions apply.

experiment, a participant was seated in a driving simulator and was instructed to steer the car back to the original cruising lane as quickly as possible once the participant was aware of the drift from cruising position of the car (*i.e.,* a lane-departure event). The lane-departure event was introduced randomly every 8-12 s after the end of the previous event. Thirty-six voluntary participants with normal or corrected-to-normal vision participated in a total of 77 sessions of the lane-keeping driving task [3]. These data were adopted from four different studies [3, 7-9] in which participants performed the same lane-keeping driving task. The experiment was approved by the Institutional Review Board of the Veterans General Hospital, Taipei, Taiwan. All participants read and signed an informed consent form before the experiments. Fifteen subjects performed the driving task multiple times on different days, resulting in a total of 36 sessions. This study examined the feasibility of the proposed selective transfer learning framework (see below) on these sessions.

B. EEG Data Recording and Preprocessing

The EEG data were collected by a 32-channel Quik-Cap (Compumedical NeuroScan, Inc.) from electrodes placed according to the international 10-20 system, and referenced to the arithmetic mean of the left and right mastoids. The impedance of all electrodes was kept under 5k Ohm during the experiments. The EEG signals were recorded with 16-bit quantization level at the sampling rate of 500 Hz. The EEG data were first processed by a 1-50 Hz band-pass finite impulse response filter to remove low-frequency drifts and high-frequency artifacts. The resultant EEG data were further down-sampled to 250 Hz before further analysis.

C. Estimation of Drowsiness Level

This study first defined the behavioral performance during the lane-keeping driving task based on the reaction time (RT) responded to randomly induced lane-departure events. That is the time interval before the onset of the lane-departure event and the onset of steering wheel. For each session, the measured RT of each lane-departure trial was normalized according to the following equation:

$$Normalized\ RT = \begin{cases} 0 & ,\tau \le \tau_0 \\ (1-e^{-(\tau-\tau_0)})/(1+e^{-(\tau-\tau_0)}) & ,\tau > \tau_0 \end{cases} \quad (1)$$

where τ is the RT to a lane-departure event, and τ_0 is the alert reaction time, which is empirically defined by the 5th percentile of the trials in a session. The resultant normalized RT ranged from zero (fully alert) to one (momentary lapse). The drowsiness level was then estimated by the average of the normalized RTs within a 90-second window before the onset of each trial under study.

D. EEG Feature Extraction

This study explored the relationship between the EEG activities and drowsiness level by correlating the EEG spectra with the putative drowsiness level. For each channel, a 256-point Welch's fast Fourier transform was applied to a 64-point moving window (zero-padded to 256 points) with an overlap of 52 points to calculate the spectral power with a frequency resolution of ~1 Hz. This study focused on the spectra between 0.98 and 30.3 Hz (30 frequency bins). The spectral power of each channel was then converted into a logarithmic scale, and then normalized by the baseline power that was calculated from the average power of the first 60 seconds of a session.

E. EEG-Based Drowsiness Regression Model

This study employed a typical linear regression model to assess the relationship between the EEG spectra and the drowsiness level for each session. Prior to building a model, a dimension-reduction procedure based on principal component analysis (PCA) was applied to the EEG spectra. PCA transferred the 900-dimension EEG spectra (30 channels × 30 frequencies) into a set of principal components (PCs). Only a subset of PCs accounting for 80% of the data variance was retained and used for the regression analysis. The present study used the Pearson correlation coefficient between the actual and predicted RTs as a metric to evaluate the performance of drowsiness estimation.

F. Level of Session Generalizability

In this study, we hypothesized that if the pilot session of an individual provides discriminative information between alertness and drowsiness for the session and for sessions from others, then the information from others might not add any value to improve the model based solely on the pilot session for the individual. Conversely, if the pilot session is not informative enough to model the drowsiness level, and is neither generalizable to predict the drowsiness level of sessions from others, then the model for this individual might benefit from the transfer learning procedure. That is, the individual's model can be improved by leveraging the models based on others' data to better estimate the drowsiness level of an unseen (test) session from the subject. To this end, it was imperative to characterize the extent of the generalizability of each session for each individual. This study thus defined a term, the level of session generalizability (*LSG*), that essentially accounts for both the performance of a given session and the performance of using that session to estimate the sessions from other subjects. The calculation of *LSG* is formularized as follows:

$$LSG_i = \frac{P(i,i)+\overline{P(i,\Phi(i))}}{\overline{P(j,j)+\overline{P(j,\Phi(j))}}\Big|_{j\in\Phi(i)}} \quad (2)$$

where

$P(a,b) \equiv$ performance of session *a*'s model on session *b*

$\Phi(i) \equiv$ assemble of indices of sessions from all other subjects of session *i*

$\overline{P(i,\Phi(i))} =$ the median performance of using session *i*'s model on all other subjects' sessions

The *LSG* of a given session is the summation of its self-prediction performance and its prediction ability to other sessions with respect to the self-prediction and cross-session performance of other subjects' sessions. This study empirically separated the 36 sessions into high- and low-*LSG* groups with the threshold of *LSG*=1. That is, under our hypothesis, the pilot session with *LSG*<1 tended to benefit from the TL procedure to model the test session from the same individual. Otherwise, the TL is not recommended to augment the pilot model for the sessions with *LSG*>1.

Authorized licensed use limited to: Jinan University. Downloaded on August 06,2020 at 01:51:52 UTC from IEEE Xplore. Restrictions apply.

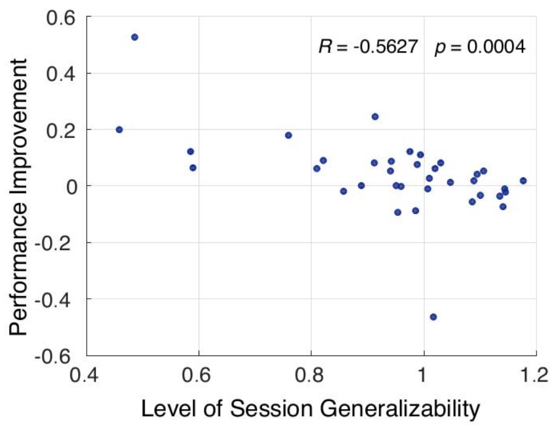

Fig. 1. The performance improvement with transfer learning as a function of the level of session generalizability (*LSG*). The significantly negative correlation implies that a pilot model with lower *LSG* value benefited more from applying transfer learning approach to estmiate the drownsiness level.

G. Selecting Auxiliary Sessions for Transfer Learning

In order to study how many auxiliary sessions were required for a TL procedure to improve the performance of the pilot model, this study systematically incorporated more sessions from other subjects and compared their TL performance. Specifically, for a given pilot session, all of its auxiliary sessions (within-subject sessions were excluded) were ranked by their performance in estimating the drowsiness level of the pilot session. The models from top-ranked auxiliary sessions were recruited first. The output drowsiness level for an unseen session was determined by the mean of the predicted values of all N+1 sessions (N from other subjects, one from the pilot session).

H. Selective transfer Learning and Performance Evaluation

This study proposed a framework of selective transfer learning to test the posed hypothesis that the sessions with lower *LSG* could benefit from leveraging other subjects' models. Thus, only the subjects whose pilot sessions had low-*LSG* would leverage the models from others to build a TL-augmented model. The performance using the proposed selective transfer learning was evaluated and compared to those based on the within-subject cross-session validation, *i.e.*, only using the pilot model from each individual to predict the unseen session (without TL) from the same individual, and routine (non-selective) TL validation, *i.e.*, forcing all individuals to use TL-augmented approach.

III. Results

Fig. 1 shows the performance improvement with transfer learning as a function of the level of session generalizability. As can be seen, the TL improvement was found negatively correlated with the *LSG* values (p=0.0004). The sessions with lower *LSG* values tended to get improved more from the TL procedure that leveraged other subjects' sessions, which evidently supported our hypothesis.

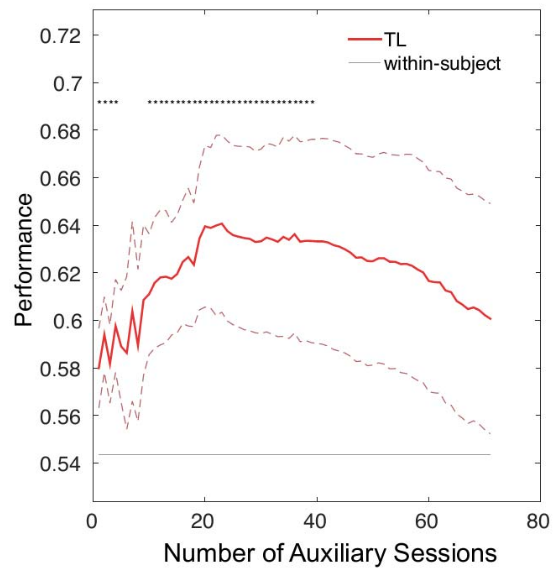

Fig. 2. The performances of TL-augmented sessions (with *LSG*<1) as a function of the number of auxiliary sessions invovled. The performance based on within-subject validation was also provided for comparison. Red bold line shows the mean of the TL performance for 20 low-*LSG* session pairs, whereas red dashed lines represent their standard errors. Gray line indicates the perofrmance based on the within-subject validation(without TL). Asterisk indicates the significant difference between TL and within-subject performances assessed by a paired t-test (p<0.05).

Fig. 2 portrays the performance of the TL-augmented pilot sessions (with *LSG*<1) as a function of the number of auxiliary sessions involved. The performance based on the within-subject cross-session validation (without TL) was also provided for comparison. Note that among the 36 pairs of pilot and test sessions, twenty pilot sessions were regarded as low generalizability and subject to this comparative study. In general, the TL-augmented pilot models (red solid line) significantly outperformed the performance based on the within-subject cross-session validation (gray line, without TL). The TL-augmented performance rose steadily until 24 top-comparable auxiliary models from other subjects were included, but then started declining when adding more auxiliary models. This suggested that naively pooling all auxiliary sessions together might not necessarily lead to better performance for a pilot model. Thus, selecting an optimal set of the auxiliary models for a pilot model was an imperative element for an effective adoption of transfer learning.

Fig. 3 shows the performance using three different approaches: the proposed selective transfer learning, routine transfer learning, and within-subject cross-session validation. Specifically, Fig. 3 (a) shows the comparative results along the individual sessions. As shown in Fig. 3 (a), in the high *LSG* group (*LSG* > 1) where the selective TL directly inherited the performance of within-subject validation (that is, no distinct TL improvement for these subjects). On the low-*LSG* side, the selective TL directly inherited the results from the TL-augmented models (red open circles and blue squares were completely overlapped for the *LSG<1* group 1). Fig. 3 (b)

Authorized licensed use limited to: Jinan University. Downloaded on August 06,2020 at 01:51:52 UTC from IEEE Xplore. Restrictions apply.

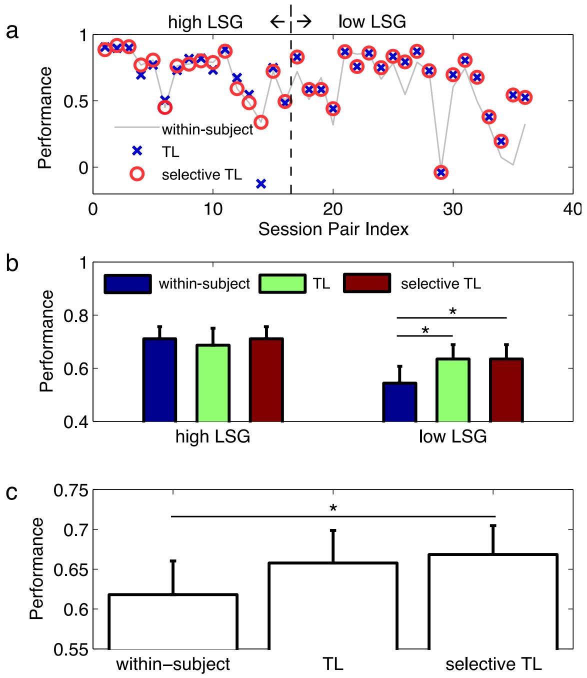

Fig. 3. The performance using the proposed selective TL, TL (non-selective), and within-subject cross-session validation. (a) presents the comparative results along the individual sessions. Sixteen session pairs were regarded as high *LSG*, whereas twenty sessions were categorized as low *LSG*. (b) compares three approaches separately in high and low *LSG* conditions. TL and selective TL both outperformed within-subject under low *LSG* condition (paired *t*-test, $p<0.01$). (c) The selective TL showed consistently significant improvements for all 36 sessions as compared to the with-subject and TL methods using paired *t*-test ($p<0.01$).

compares the average drowsiness-detection performances in high and low *LSG* sessions obtained by different approaches. TL was unable to improve the accuracy of drowsiness detection in the high *LSG* sessions, but could offer significant improvements for the low *LSG* sessions. The overall performance in Fig. 3 (c) suggested that the proposed selective TL could enhance the performance of drowsiness estimation.

IV. Discussions and Conclusion

This study investigated the feasibility of leveraging existing data from other subjects to improve the performance of a drowsiness-detection BCI. The study results showed that exploiting data from other subjects for an individual was not always favorable for estimating the level of drowsiness in a new session for the same individual. As shown in Fig. 1, transfer learning could sometimes deteriorate more than improve the within-subject BCI performance. The results of this study also showed that a pilot session with high *LSG* might not able to take advantage of other subjects' data using the transfer learning method implemented in this paper. Therefore, it might be necessary to formulate a strategy to selectively apply transfer learning under different circumstances.

To appropriately select auxiliary data, this study ranked each session from other subjects by the performance of each session's model testing against the other subject's data. The obtained empirical results showed that the ranking method was effective for selecting informative auxiliary data (*c.f.* Fig. 2), where the transfer learning performance was improved even using only one extra session, and was gradually increased until using ~24 sessions. Fig. 3 shows a comparison study among three approaches: routine TL, selective TL and no-TL. The selective TL evidently outperformed others.

In summary, this study proposed a framework to effectively leverage a large amount of training data from other subjects and a small amount of subject-specific pilot data to improve BCI performance. This framework can be useful to improve BCI performance when collecting a sufficient amount of subject-specific pilot data is difficult or impossible.

Acknowlegments

Research was sponsored in part by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0022, and in part by the Taiwan National Science Council I-RiCE Program under Grant Number NSC-101-2911-I-009-101.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, Jun. 2002.

[2] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski, "Estimating alertness from the EEG power spectrum," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 1, pp. 60–69, Jan. 1997.

[3] C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, and T.-P. Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 12, pp. 2726–2738, Dec. 2005.

[4] C.-H. Chuang, P.-C. Lai, L.-W. Ko, Bor-Chen Kuo, and C.-T. Lin, "Driver's cognitive state classification toward brain computer interface via using a generalized and supervised technology," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 2010, pp. 1–7.

[5] D. Wu, B. J. Lance, and T. D. Parsons, "Collaborative Filtering for Brain-Computer Interaction Using Transfer Learning and Active Class Selection," *PLoS ONE*, vol. 8, no. 2, p. e56624, 21 2013.

[6] R. S. Huang, T. P. Jung, and S. Makeig, "Event-Related Brain Dynamics in Continuous Sustained-Attention Tasks, " *Lecture Notes in Computer Science, Foundations of Augmented Cognition*, vol. 4565, pp. 65-74, 2007.

[7] S.-W. Chuang, L.-W. Ko, Y.-P. Lin, R.-S. Huang, T.-P. Jung, and C.-T. Lin, "Co-modulatory spectral changes in independent brain processes are correlated with task performance," *Neuroimage*, vol. 62, no. 3, pp. 1469–1477, Sep. 2012.

[8] N. R. Pal, C.-Y. Chuang, L.-W. Ko, C.-F. Chao, T.-P. Jung, S.-F. Liang, and C.-T. Lin, "EEG-Based Subject- and Session-independent Drowsiness Detection: An Unsupervised Approach," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, p. 519480, Nov. 2008.

[9] C.-H. Chuang, L.-W. Ko, Y.-P. Lin, T.-P. Jung, and C.-T. Lin, "Independent Component Ensemble of EEG for Brain-Computer Interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 230–238, Mar. 2014.

Authorized licensed use limited to: Jinan University. Downloaded on August 06,2020 at 01:51:52 UTC from IEEE Xplore. Restrictions apply.