2016 ICSEE International Conference on the Science of Electrical Engineering

Enhancement Of BCI Classifiers Through Domain Adaptation

Hadas Benisty*[1], Daniel Furman†[1], Talor Abramovich*[2], Amir Ivry*[2] and Hillel Pratt†
* Electrical Engineering Department
†Faculty Of Medicine
Technion, Israel Institute of Technology
Technion City, Haifa 32000, Israel
Email: {hadasbe@tx, dfurman@tx, talor@t2, sivry@t2, hillel@tx }.technion.ac.il

***Abstract*—Clinical Brain-Computer Interface (BCI) systems seek to enable paralyzed individuals to operate devices with their brain activity. Non-invasive systems based on electroencephalographic (EEG) signals are popular since they avoid risks associated with invasive procedures, but unfortunately EEG signals are inherently noisy, making effective classifiers challenging to develop. Commonly, new classifiers are benchmarked on signals from healthy subjects executing physical movements, under the assumption that the performance will transfer to clinical cases where only imagined movements are possible. Here, we show in contrast that classifiers trained on signals associated with actual movements perform erratically when applied to signals associated with imagined movements. We suggest that this is because the signals lay in different domains. Then, to exploit the different statistical distributions, we apply a domain adaptation technique, Frustratingly Easy Domain Adaptation (FEDA), improving classifier performance accuracy by a third on a discrimination task that simulates the clinical condition.**

I. Introduction

The translation of brain activity directly into command signals for electronic devices through brain-computer interfaces (BCIs) [1], [2], [3] can provide substantial improvements to the quality of life of paralyzed individuals by ameliorating communicative deficits, navigational difficulties, and facilitating many other professional and recreational activities [4], [5]. At present, however, there remain several obstacles to the widespread adoption of BCIs for everyday use by those in need.

One significant obstacle is the inability to meaningfully transfer classifiers, both from one population to another (for example, healthy individuals to patients), and between sessions (for example, first day using a BCI to the second day). Despite much effort dedicated to the extraction of stationary features [6], [7], [8], and the design of classifiers for these features alone, the challenge remains at the forefront, since inter- and intra- subject signal variability continues to be problematic for current algorithmic approaches [9], [10]. Long and laborious calibration and training stages are thus currently required for each new user of a BCI system, as well as at each new session, and this reflects an emerging bottleneck in the field, as the accumulation of data and understanding has a limited potential to productively enhance clinical systems directly.

In this paper, we suggest that one reason for the BCI field's inability to design robust and reliable classifiers is a common and inaccurate assumption: the idea that the brain activity encoding actual movements and imagined movements is equivalent. While it has been observed that certain principles of movement-related brain activity are preserved during both actual and imagined movement [11], [12], these findings are often crudely (and perhaps, conveniently) used as support for the assumption of equivalence which has guided many methodological inquiries in clinically ineffectual directions. Though partially overlapping, the cortical and subcortical sources responsible for actual movement and imagined movement are well-understood by neuroscientists to be distinct [13], [14], [15], and associated scalp-recorded electrophysiology signals are likewise distinguishable, through differences in event-related potentials as well as in the frequency domain, where the propagation of beta (15-30Hz) and gamma (30-100Hz) activity varies both in timing and topography [16], [17], [18].

Despite this understanding however, signals corresponding to actual and imagined movement continue to be conflated in the BCI literature, with typical research resembling the sequence: collect data from healthy individuals performing real or imagined movements, create and evaluate a classifier, and then claim, due to data equivalence, that the classifier can help paralyzed individuals. Here we simulate the process of transferring laboratory results to the clinical case differently, and show that a BCI classifier's performance degrades when trained on brain signals corresponding to actual movements (simulating typical, healthy subjects) and tested on brain signals corresponding to imagined movements (simulating the clinical goal, a paralyzed subject). Furthermore, we show particularly significant degradation in the accuracy rate pertaining to brain signals of subjects who were not included in the training set (simulating new subjects). This degradation is so severe that the classifier's accuracy approaches random selection.

After establishing that the brain activity encoding actual movements exists in a different statistical domain than the

[12] These two authors contributed equally to this work.

978-1-5090-2152-9/16/$31.00 ©2016 IEEE

brain activity encoding imagined movements, in order to exploit this difference to strengthen BCI classifiers, we look to an established technique in the machine-learning literature: Domain Adaptation (DA), also known as Transfer Learning [19]. DA has previously been successfully employed in the BCI field to assist in subject-to-subject classifier transfer [20], as well as session-to-session transfer [21], and most recently a new transfer learning framework has been proposed that addresses both transfer objectives [22].

Here, we contribute to the growing DA-BCI literature by demonstrating the usefulness of transfer learning within a new scenario, that in which actual movement data is used to train a classifier which is afterwards adapted to the imagined domain. We use imagined movement data to adapt the classifier to the imagined data domain, and hence demonstrate that data collected from healthy subjects executing real movements, which is more practical for a variety of reasons, can feasibly be used to train a classifier that is applicable to a clinical population.

We show that while a naive classifier training protocol, which treats actual and imagined data as equivalent, leads to almost random accuracy rates in a four-class condition (25%) when the classifier is tested on unknown subjects, our proposed approach raises the accuracy to 33%, a value that is 1.32 times chance accuracy for this task. This improved accuracy nears that achieved by highly specialized classifiers that employ much more extensive feature extraction stages while performing related discrimination tasks, such as 1.85 [23], 1.71 [24], and 1.52 [25] chance accuracies reported in other publications. We emphasize that this current paper is meant to explore a new method for improving a classifier's accuracy, and as such, simple features were selected for computational efficiency. In the future, applying the current approach to more advanced classification schemes may be a fruitful strategy for achieving the highest absolute classification accuracy.

Our proposed approach is both immediately applicable as a strategy for improving an existing classifier's performance, as well as a platform upon which future research programs can be built, since it suggests that easily acquired data, from physically-capable subjects, may be used to systematically and reliably enhance the performance of BCIs in a clinical case.

II. Experiments Setup

1) Subjects: Fifteen (15) right-handed, female and male volunteers, 18-30 years old, with normal hearing and motor function, and without a history of neurological or psychiatric disease, participated in the experiment which is described in more depth in [23].

2) Paradigm: Subjects sat upright in an electrically shielded and acoustically isolated room, holding a custom-built spherical response device that contained four buttons. During the experiment, auditory stimuli (delivered using 'Presentation' software by Neurobehavioral Systems, Inc) were used to cue the subject to perform finger flexion and release movements. These movements were either actual or imagined, dependent upon the experimental condition. In the actual condition, the subject physically pressed and released the appropriate button on the response device using one of four fingers, two on each hand, while in the imagined condition the subject solely thought about performing the appropriate finger flexion. The auditory stimuli in both conditions were identical, and delivered with equivalently pseudo-random timing and order. All experimental events were marked on the EEG record with sub-millisecond timing resolution.

3) Data acquisition: A 61-channel EEG cap (Electro-Cap International Inc.) with electrodes embedded according to the 10/20 and intermediate electrode locations system was used for data acquisition. Each electrode well was filled with conducting gel (Electro-Cap International Inc) to improve signal quality. An additional 9-mm disc electrode was placed on the middle of each subject's chin as a reference, and another clip-electrode on the left earlobe served as the ground. Three additional electrodes were used to record the subjects horizontal and vertical eye movements (EOG); two placed bilaterally near the outer canthus of each eye, and one below the right eye. The cap-mounted and additional electrodes were connected to an EEG recording system (MicroMed). Potentials from all channels were amplified at 0.15-134.4 Hz, digitized with a 16-bit A/D converter, and sampled at 512 Hz. The impedance of each electrode was maintained below 5 kOhm.

4) Data processing: EEG data is vulnerable to numerous noise sources, and thus, cleaning of EEG data is an essential prerequisite for conducting a proper analysis of brain activity. Cleaning began with segmentation of the continuous EEG into epochs, removal of DC bias, and subsequently, second-order infinite impulse response (IIR) Butterworth band-pass filtering (0.1-24 Hz, 6dB/octave slopes). Automatic voltage $\pm$75uV and spectral thresholding (0-2Hz at $\pm$50dB, 20-40Hz at $\pm$100-$\pm$25dB) was implemented to guide subsequent visual inspection and removal of trials irreparably affected by movement artifacts as well as miscellaneous electrical noise. Following manual removal of artifact-contaminated trials, Independent Component Analysis (ICA) was applied to facilitate removal of artifact-related components [26]. Evaluation and elimination of independent components was aided by the ADJUST algorithm [27], a processing heuristic that uses spatial and temporal features to identify components associated with artifacts such as eye-movements, electrocardiogram, muscle activity and mains AC 50Hz. For the purposes of this present study, the cleaned data was segmented into 1000ms epochs, with 0ms corresponding to the delivery of auditory cues to move or imagine movement. Data from scalp electrodes was used exclusively, amounting to a total of 61 channels in each epoch. Feature extraction was performed by computing the cross-channel covariance matrix of each epoch, yielding a total of 3721 features. The feature matrix was labeled according to condition (actual or imagined) and finger moved.

III. ACTUAL DOMAIN VS. IMAGINED DOMAIN

In this section we show that signals related to actual and imagined movements lay in different domains, and predict that cross-training and testing without sensitivity to the presence of these domains damages classifier performance. A common practice to examine whether two types of signals lay in different domains is to examine how well they are separated using a binary classifier. In our case, as an initial step, we trained a binary classifier to distinguish between the actual and imagined feature vectors, using an available Support Vector Machine (SVM) toolkit [28].We used the actual and imagined feature vectors extracted from all 15 subjects for a 10-fold cross-validation where at each fold 80% of the data was used for training, 10% was used for parameter tuning and the remaining 10% for testing. Results show a clear separation of actual data and imagined data, as the mean accuracy rate on the test set data was 98.94% using a linear classifier and 98.8% using an RBF kernel. The high (almost perfect) classification accuracy rate indicates that actual and imagined signals lay in two, almost distinct domains. To investigate the influence of the difference between the actual and imagined domains on the classifier performance, we trained multi-class classifiers on data from four types of single finger movements and evaluated classification performance. In all experiments, 14 subjects participated in the training process while the 15th subject was used for testing. We examined linear, polynomial and radial basis function (RBF) kernels. In all our experiments, the RBF kernel led to higher accuracy rates than the others, and as such, all classification results reported hereafter correspond to results from this kernel. Table I shows the mean accuracy rates and standard deviations (averaged over all 15 subjects) for all four possible permutations of training and testing on actual and imagined data. As predicted, when testing on actual data, higher accuracy rate is achieved by training on actual data than by training on imagined data. Testing on imagined data, however, leads to inconsistent results, as the accuracy rate is comparable to random selection: 25%. Moreover, training on imagined data does not raise the accuracy rate above random selection for testing on either actual or imagined data. This implies that the imagined data is far less stable than the actual data, and it cannot be used alone for training a reasonable classifier. A naive solution for this setup would be to use both data sets for training. Table. II presents the mean accuracy rate (and standard deviation) for testing on imagined movements obtained by a classifier trained by data related to actual movements as well as data related to the imagined movements. Still, using this naive approach leads to accuracy rate which is comparable to random selection.

In the next sections we propose a classification scheme based on a Domain Adaptation (DA) technique called Frustratingly Easy Domain Adaptation (FEDA) [29]. Using FEDA, we significantly improve the accuracy rate by utilizing actual data for training a stable classifier and the imagined data for adapting it to the imagined domain.

TABLE I
MEAN ACCURACY RATES [%] FOR CLASSIFIERS TRAINED BY ACTUAL AND/OR IMAGINED SIGNALS.

	Testing Data	
Training Data	Actual	Imagined
Actual	30.0 ± 4.9	25.5 ± 2.6
Imagined	26.8 ± 2.9	25.9 ± 2.8

TABLE II
MEAN ACCURACY RATE AND STANDARD DEVIATION [%] OBTAINED BY A CLASSIFIER TRAINED USING A NAIVE APPROACH (USING BOTH DATA TYPES) AND TESTED ON DATA RELATED TO IMAGINED SIGNALS.

Training Strategy	**Mean Accuracy**
A Naive Approach	25 ± 3

IV. DOMAIN ADAPTION (DA)

Domain Adaption (DA), also known as Transfer Learning, is a method for training a classifier using one data set from a source domain together with a typically smaller data set from a target domain for the purpose of classifying samples related to the target domain.

Denote $\mathcal{D}^s$ as a probability space of a source domain and $\mathcal{D}^t$ as the probability space of a target domain, where $\mathcal{D}^s \not\sim \mathcal{D}^t$, meaning that the probability spaces are distributed differently. In DA applications, the amount of target samples available for training, M, is typically smaller than the amount of source samples, N, available for training. The goal of DA is to utilize the source samples for training a robust classifier, and the target samples for adaptation to the target domain.

A. FEDA

FEDA [29] is a DA algorithm which relies on feature augmentation. The feature vector's space is notated by $\mathcal{X} \subseteq \mathbb{R}^F$ (for some $F > 0$) and the output space by $\mathcal{Y} = \{1, ..., K\}$. For example, in our case of learning a classifier for movements of 4 fingers, $K = 4$. The adapted classifier is learned using two stages: augmentation and training. At the augmentation stage, the training sets related to the source and target domains are transformed to an augmented space using $\Phi \colon \mathcal{X} \to \tilde{\mathcal{X}}$, where $\tilde{\mathcal{X}} \subseteq \mathbb{R}^{3F}$, defined by:

$$\Phi(\mathbf{x}) = \begin{cases} (\mathbf{x}, \mathbf{x}, \mathbf{0}) & \mathbf{x} \sim \mathcal{D}^{\mathbf{s}} \\ (\mathbf{x}, \mathbf{0}, \mathbf{x}), & \mathbf{x} \sim \mathcal{D}^{\mathbf{t}} \end{cases} \tag{1}$$

where $\mathbf{x} \in \mathcal{X}$ represents the feature vector of a certain data sample and $\mathbf{0} = (0, 0, \dots, 0) \in \mathbb{R}^F$. The first F elements of $\Phi(\mathbf{x})$ represent a general domain which captures common features of the two domains, the following F elements represent the source domain, and the last F elements of $\Phi(\mathbf{x})$ represent the target domain. At the training stage, the classifier is learned using the augmented training sets.

During testing, each test sample $\mathbf{x}$ is augmented according to its domain and then classified using the learned classifier.

B. BCI Classification Using FEDA

As stated above, actual data is more stable and more easily obtained than imagined data. However, imagined data is closer to a real-life setup where potential users suffer from some degree of paralysis. In this section, we propose a BCI-classification scheme based on FEDA which utilizes the availability and stability of the actual data for training a robust classifier, and the available imagined data for adaptation to the imagined data domain. Since our goal is to classify imagined data, we define the actual and imagined domains as source and target domains, respectively. We augment the feature vectors and train an adapted classifier as described above. Given a test sample, we augment it according to its domain and apply the learned classifier.

V. FEDA Results

In this section we present classification results for imagined movements, obtained by the proposed approach. We followed the same protocol as in Sec. III of training on 14 subjects, and testing on the 15th, for all 15 subjects, and also used the same 10-fold cross-validation. Again, we used data related to actual movements and data related to imagined movements for training the classifier. Table III presents the mean accuracy rate and standard deviation obtained by the proposed FEDA-based classifier. For comparison, it also presents the naive classifier trained without DA (presented above in Table II), using both data sets as if they were equivalent and yielding the random selection rate. The naive approach leads to an inconsistent classifier that has a 25% accuracy rate, while our proposed approach of applying FEDA significantly improves the classification accuracy rate to 33%. This improvement of 8% to the total classification accuracy indicates an enhanced classifier.

TABLE III
Mean accuracy rates and standard deviations [%] obtained by the naive classifier and the proposed approach of using FEDA, tested on data related to imagined signals.

Training Strategy	Mean Accuracy
A Naive Approach	25 ± 3
FEDA	$\mathbf{33 \pm 9}$

VI. Conclusion

Brain-Computer Interfaces (BCIs) aim to improve the quality of life for paralyzed individuals by enabling them to control electronic devices through their thoughts alone. Although it is well-established in basic neuroscience that actual and imagined movements correspond to different neural activity signatures, signals from these action types are still often treated as equivalent in many BCI research programs and prototype systems. In this paper, our findings reinforce the basic neuroscience understanding, by showing that signals recorded while subjects performed actual and imagined movements lay in different statistical domains. We simulate a real-world BCI setup, where the classification objective is to discriminate between imagined movements of an unknown subject, and show that inconsistent classification results are obtained when the classifier is trained on brain signals acquired during actual movements. Moreover, exclusively training on data from imagined movements also leads to inconsistent results due to the high variability of this domain of signals. To address both shortcomings, we propose here a classification scheme based on a domain adaptation technique called FEDA. We utilize the availability and stability of signals acquired during actual movements to train a robust classifier, and signals acquired during imagined movements for adapting it to the imagined data domain. Since fingers are humans' primary effectors for interacting with modern technologies, and EEG is the most commonly used signal acquisition modality for BCIs, we chose to evaluate this proposed domain adaptation approach using EEG data corresponding to four classes of finger flexions. Compared to the naive approach of training a classifier (without adaptation) using both actual and imagined signal types, our approach enhances the classifier in the clinically-relevant domain by a third, from 25% to 33%. This initial result suggests that applying domain-adaptation techniques to data recorded from physically-capable subjects can facilitate meaningful improvements in classifiers developed for clinical purposes.

Acknowledgment

The authors would like to thank the Evoked Potentials Laboratory and the Rappaport Family Institute for Research in the Medical Sciences for funding and support in conducting the electrophysiology experiment. Thanks to the Signal and Image Processing Laboratory (SIPL), where this project was conducted, under the academic head Prof. David Malah, and the chief engineer, Nimrod Peleg.

References

[1] J. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: principles and practice*. Oxford University Press, 2012.

[2] D. J. McFarland, C. W. Anderson, K. Muller, A. Schlogl, and D. J. Krusienski, "BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, p. 135, 2006.

[3] J. R. Wolpaw, "Brain–computer interfaces as new brain output pathways," *The Journal of physiology*, vol. 579, no. 3, pp. 613–619, 2007.

[4] P. Rigby, S. E. Ryan, and K. A. Campbell, "Electronic aids to daily living and quality of life for persons with tetraplegia," *Disabil Rehabil Assist Technol*, vol. 6, no. 3, pp. 260–267, 2011.

[5] A. Folan, L. Barclay, C. Cooper, and M. Robinson, "Exploring the experience of clients with tetraplegia utilizing assistive technology for computer access," *Disabil Rehabil Assist Technol*, vol. 10, no. 1, pp. 46–52, Jan 2015.

[6] W. Samek, F. C. Meinecke, and K.-R. Muller, "Transferring subspaces between subjects in brain–computer interfacing," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 8, pp. 2289–2298, 2013.

[7] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 2, pp. 355–362, 2011.

[8] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, and G. Pfurtscheller, "Nonstationary brain source separation for multiclass motor imagery," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 2, pp. 469–478, 2010.

[9] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, “Co-adaptive calibration to improve bci efficiency,” *Journal of Neural Engineering*, vol. 8, no. 2, p. 025009, 2011.
[10] A. Vuckovic, “Non-invasive bci: How far can we get with motor imagination?” *Clinical neurophysiology*, vol. 120, no. 8, pp. 1422–1423, 2009.
[11] G. Pfurtscheller and F. L. Da Silva, “Event-related eeg/meg synchronization and desynchronization: basic principles,” *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
[12] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. L. Da Silva, “Mu rhythm (de) synchronization and eeg single-trial classification of different motor imagery tasks,” *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.
[13] E. Gerardin, A. Sirigu, S. Lehéricy, J.-B. Poline, B. Gaymard, C. Marsault, Y. Agid, and D. Le Bihan, “Partially overlapping neural networks for real and imagined hand movements,” *Cerebral cortex*, vol. 10, no. 11, pp. 1093–1104, 2000.
[14] T. Hanakawa, M. A. Dimyan, and M. Hallett, “Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional mri,” *Cerebral cortex*, vol. 18, no. 12, pp. 2775–2788, 2008.
[15] P. Dechent, K.-D. Merboldt, and J. Frahm, “Is the human primary motor cortex involved in motor imagery?” *Cognitive Brain Research*, vol. 19, no. 2, pp. 138–144, 2004.
[16] C. Kranczioch, S. Mathews, P. Dean, and A. Sterr, “Task complexity differentially affects executed and imagined movement preparation: evidence from movement-related potentials,” *PloS one*, vol. 5, no. 2, p. e9284, 2010.
[17] S. Jankelowitz and J. Colebatch, “Movement-related potentials associated with self-paced, cued and imagined arm movements,” *Experimental brain research*, vol. 147, no. 1, pp. 98–107, 2002.
[18] J. Ginter Jr, K. Blinowska, M. Kamiński, and R. Kuś, “Propagation of brain electrical activity during real and imagined motor task by directed transfer function,” in *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on.* IEEE, 2005, pp. 105–108.
[19] S. J. Pan and Q. Yang, “A survey on transfer learning,” *Knowledge and Data Engineering, IEEE Transactions on*, vol. 22, no. 10, pp. 1345–1359, 2010.
[20] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, “Subject-independent mental state classification in single trials,” *Neural networks*, vol. 22, no. 9, pp. 1305–1312, 2009.
[21] M. Krauledat, M. Tangermann, B. Blankertz, and K.-R. Müller, “Towards zero training for brain-computer interfacing,” *PloS one*, vol. 3, no. 8, p. e2967, 2008.
[22] V. Jayaram, M. Alamgir, Y. Altun, B. Schölkopf, and M. Grosse-Wentrup, “Transfer learning in brain-computer interfaces,” *arXiv preprint arXiv:1512.00296*, 2015.
[23] D. Furman, R. Reichart, and H. Pratt, “Finger flexion imagery: EEG classification through physiologically-inspired feature extraction and hierarchical voting,” *to Appear In: Brain-Computer Interface (BCI), 2016 4th International Winter Conference on*, 2016.
[24] F. Quandt, C. Reichert, H. Hinrichs, H.-J. Heinze, R. Knight, and J. Rieger, “Single trial discrimination of individual finger movements on one hand: A combined meg and eeg study.” *NeuroImage*, vol. 59, no. 4, pp. 3316–3324, 2012.
[25] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, “Decoding individual finger movements from one hand using human eeg signals,” *PLoS ONE*, vol. 9, no. 1, p. e85192, 01 2014.
[26] S. Makeig, A. J. Bell, T.-P. Jung, T. J. Sejnowski *et al.*, “Independent component analysis of electroencephalographic data,” *Advances in neural information processing systems*, pp. 145–151, 1996.
[27] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, “Adjust: An automatic eeg artifact detector based on the joint use of spatial and temporal features,” *Psychophysiology*, vol. 48, no. 2, pp. 229–240, 2011.
[28] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.
[29] H. Daumé III, “Frustratingly easy domain adaptation,” *arXiv preprint arXiv:0907.1815*, 2009.