FUZZY DIVERGENCE BASED ANALYSIS FOR EEG DROWSINESS DETECTION BRAIN COMPUTER INTERFACES

Tharun Kumar Reddy* *Vipul Arora** *Laxmidhar Behera** *Yu-kai Wang* † *Chin-Teng Lin*†

* Department of Electrical Engineering, Indian Institute Of Technology Kanpur
†University of Technology Sydney, Ultimo, NSW 2007, Australia

ABSTRACT

EEG signals can be processed and classified into commands for brain-computer interface (BCI). Stable deciphering of EEG is one of the leading challenges in BCI design owing to low signal to noise ratio and non-stationarities. Presence of non-stationarities in the EEG signals significantly perturb the feature distribution thus deteriorating the performance of Brain Computer Interface. Stationary Subspace methods discover subspaces in which data distribution remains steady over time. In this paper, we develop novel spatial filtering based feature extraction methods for dealing with nonstationarity in EEG signals from a drowsiness detection problem (a machine learning regression problem). The proposed method: DivOVR-FuzzyCSP-WS based features clearly outperformed fuzzy CSP based baseline features in terms of both RMSE and CC performance metrics. It is hoped that the proposed feature extraction method based on DivOVR-FuzzyCSP-WS will bring in a lot of interest in researchers working in developing algorithms for signal processing, in general, for BCI regression problems.

***Index Terms*—** Brain Computer Interface (BCI), Electroencephalogram (EEG), Stationary Subspace Analysis (SSA), Reaction Time (RT) prediction, Fuzzy Common Spatial Patterns (Fuzzy CSP), Drowsiness

1. INTRODUCTION

Electroencephalogram (EEG) is a bio-signal recorded non-invasively from the scalp. It is a cheaper neuroimaging technique than FMRI and MEG. Due to its high temporal resolution and cost-effective nature, it is in high demand to analyze and decipher the commands in a Brain-Computer Interface (BCI). BCI provides for explicit control command transfer between the brain and a controllable device via signal recordings of brain activity. Motor Imagery (MI) Classification (deciphering thoughts of imagination of left or right hands) is an important research problem which is made use of frequently to control BCIs. A BCI block in general is inclusive of preprocessing, feature extraction and classification or regression modules. EEG signals typically are associated with a reduced signal to noise ratio (SNR) owing to volume conduction and intrinsic non-stationarities. In the preprocessing module, spatial filtering is done to enhance the SNR and further extract features for MI events. Common Spatial Patterns (CSP) [1] method has been one of the widely used routine to execute spatial filtering in a two class MI [2].

Brain Computer Interfaces performance is limited by large variations of the subject state both within and across sessions as well as noise and artefacts. In principle, basic CSP algorithm is not robust to such nonstationarities [3] and they can detrimentally affect the performance of CSP. In [4], an unsupervised method was proposed to draw out stationary parts from EEG for MI classification. Further, in [5], authors performed alignment based on class information (2 MI classes) to optimize stationarity leading to a supervised approach. Later, in [5], a joint optimization of discriminability and stationarity has been accomplished. Later, authors in [6] proposed a unifying optimization framework for spatial filter computation based on divergence maximization. Recently, subspace based methods are robustified to noise and artefacts and presented in [7].

CSP filtering was fundamentally proposed for two class MI. A variety of approaches like One-Versus-Rest (OVR) CSP [8], doublet two class classification succeeded by voting [9], CSP with filter selection [10] and Riemannian approaches [11]were designed for feature extraction and multiclass classification in MI. Stationarity optimization approaches presented in [6, 12] are not derived for multiclass BCI. Recently, in [13, 14], a theoretically ground approach was proposed and validated for multiclass MI BCIs.

Driver drowsiness leading to sleepiness has been recognized by US National Highway Traffic Safety Administration as an integral component of several road accidents [15]. Research [15] has revealed that drowsy driving is more dangerous compared to driving under the influence of a minor amount of alcohol. Early premonition of drowsiness due to fatigue is therefore a pertinent research topic. Accordingly, several nations and state leaders are working towards development and deployment of novel solutions to enhance driving security.

Recently, fuzzy models proved to be promising in several pertinent applications like traffic life cycles [16], networked control systems [17, 18] and BCI [19]. In this work, we fur-

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

ther apply fuzzy modeling approach based on above motivation.

A prominent goal of this work is to pre-train the machine learning blocks by the data recorded from drowsiness simulations in an EEG based Reaction-Time prediction task [20]. All the approaches outlined so far are for EEG classification specific to MI. But, the nonstationarities associated with EEG are also to be dealt with in regression setting. EEG based Reaction-Time (RT) prediction [21, 22, 23] is a practical EEG regression problem. Several variants of CSP like FuzzyCSP [22, 24] are the state of the art spatial filtering based feature extraction frameworks for regression. FuzzyCSP can be reposed as a divergence problem by finding filters optimizing the KL divergence of fuzzy classes [25]. The predictive power of regression is largely influenced by the non-stationarity in the feature distributions [12]. In previous works [26, 27, 28], we demonstrate the effect of nonstationarities on the performance of EEG BCI. In this work, we systematically incorporate non-stationarity into the fuzzy spatial filtering BCI framework for regression.

In this paper, our contributions are multi-fold as indicated below: (1) We extend the divergence based analysis [6] for regression. (2) We formulate a novel stationarity based spatial filtering framework for feature extraction in an EEG regression problem setting. Based on this, we propose DivOVR-FuzzyCSP-WS, a stationarity based filtering method for both binary and multiple fuzzy classes. (3) The method developed optimizes the selection of filters while retaining the stationarity inside the session in addition to enhancing the predictive power of regression.

Rest of the paper is organized as follows. Section 2 discusses the EEG sustained attention task for simulating real-world drowsiness and the associated data format. Section 3 presents the basic formulation of divergence based framework for regression. Section 4 formulates divergence framework for joint optimization of stationarity and predictive power of regression. Section 5 presents the optimization of divergence framework for multiple fuzzy classes within session and documents the results. Section 6 provides implementation details and a brief discussion. Section 7 concludes this work and provides future directions.

2. DATASET FROM EEG BASED DRIVER DROWSINESS EXPERIMENTS

A driving experiment (indicated in fig.2) is selected in this study to investigate the brain dynamics changes connected to human performance in a sustained-attention driving task. Simulated driving experiments are conducted on a virtual reality (VR)-based dynamic driving simulator. A real car frame is mounted on a six degree-of-freedom Stewart motion platform which moved in sync with the driving scene during 'motion' sessions. The motion platform is inactive during 'motionless' sessions. The VR driving scene simulated nighttime cruising (100 km/h) on a straight highway (two lanes in each direction) without other traffic. The computer program generated a random perturbation (deviation onset), and the car started to drift to the left of the right of the cruising lane with equal probability. Following each deviation, subjects are required to steer the car back to the cruising lane as quickly as possible using the steering wheel (response onset), and hold on the wheel after the car returned to the approximate center of the cruising lane (response offset). A lane departure trial is defined as consisting of three events, deviation onset, response onset, and response offset. The next lane-departure trial randomly occurs about 5 to 10 sec after response offset in the current trial. The subject's reaction time (RT) to each lane departure trial is defined as the interval between deviation onset and response onset. If the subject does not respond promptly within 2.5 (1.5) sec, the vehicle will hit the left (right) roadside without a crash and continue to move forward against the curb event, the subject completely ceases to respond. No intervention is made when the subject fell asleep and stopped responding. After reaching lapse period, subjects resumed the task voluntarily and steered the car back to the cruising position at the earliest. The EEG data is recorded from 30 sintered Ag/AgCl EEG active electrode sites (referred to linked mastoids). All the EEG electrodes alluding to the right ear lobe are kept in accordance with a modified International 10–20 system of electrode installation. EEG data is recorded along with the corresponding trialwise reaction time values from 11 subjects in a particular session. This data is already tested in several works [22, 24] as a benchmark for EEG regression. EEG data consists of segmented trials $\mathbf{X} \in \mathbb{R}^{C \times T}$ ((C, T) denote number of channels and Time samples respectively) and their corresponding reaction time values (Y).

3. DIVERGENCE BASED CSP FOR REGRESSION

Samek et al., [6] came up with the divergence based formulation for CSP algorithm for BCI classification problems. For BCI regression problems [22, 24], approaches like Neural Networks [21] and LASSO [24] are applied on features extracted from fuzzy CSP [24]. The tools and techniques developed for the BCI regression tasks are similar to that for classification problems and depend on the EEG signal interpretation as a time series.

Here, in this section, we propose divergence based fuzzy CSP for regression. We formulate a novel objective function for regression using fuzzy covariance matrices. We start with two fuzzy classes and extend it later for as many classes ($M > 2$). The primary assumptions are: the conditional distribution of each fuzzyclass is Gaussian i.e. $\mathcal{N}(\mathbf{0}, \overline{\mathbf{\Sigma}}_1)$ and $\mathcal{N}(\mathbf{0}, \overline{\mathbf{\Sigma}}_2)$ for each of the 2 fuzzy classes respectively ($\overline{\mathbf{\Sigma}}_1$ and $\overline{\mathbf{\Sigma}}_2$ are the fuzzy class covariance matrices).

The KL divergence between two D variate gaussians $p_1 \sim$

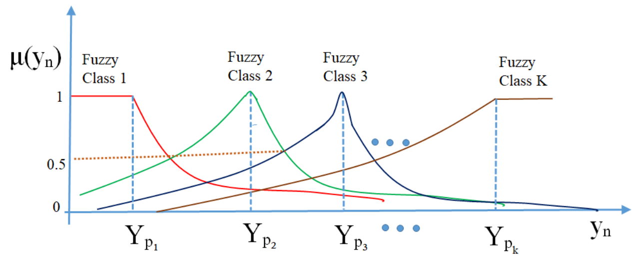

Fig. 1. Fuzzified RT variable.

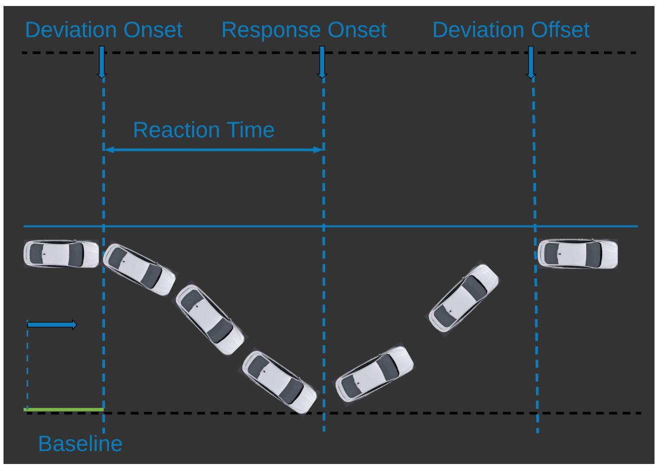

Fig. 2. Paradigm Of Lane-keeping task

$\mathcal{N}(\mu_{\mathbf{0}}, \mathbf{\Sigma}_0)$ and $p_2 \sim \mathcal{N}(\mu_{\mathbf{1}}, \mathbf{\Sigma}_1)$ can be computed as:

$$D_{kl}(p_1 \parallel p_2)) = \frac{1}{2}\left(\log\left(\frac{det(\mathbf{\Sigma_1})}{det(\mathbf{\Sigma_0})}\right) + trace(\mathbf{\Sigma_1}^{-1}\mathbf{\Sigma_0}) + (\mu_{\mathbf{1}} - \mu_{\mathbf{0}})^{\intercal}\mathbf{\Sigma_1}^{-1}(\mu_{\mathbf{1}} - \mu_{\mathbf{0}}) - D \right) \tag{1}$$

Further, an estimate of the class-wise fuzzy covariance matrix $\overline{\mathbf{\Sigma}}_{\mathbf{i}}$ is obtained below. A fuzzy averaged trial matrix $\overline{\mathbf{X}}_i \in \mathbb{R}^{C\times T}$ is obtained for each fuzzy class i. Here, number of channels $C = 30$ and number of samples $T = 250$. Then, the corresponding fuzzy class covariance matrix $\overline{\mathbf{\Sigma}}_i$ is computed.

$$\overline{\mathbf{X}}_i = \frac{\sum_{k=1}^{N_i} \mu_{k,i}\mathbf{X}^k}{N_i} \quad i \in \{1, 2, \cdots M\}\,(k : Y_k = i - 1) \tag{2}$$

In (2), N_i represents the number of trials in the i^{th} fuzzy class, membership function $\mu_{k,i}$ is defined for the k^{th} data trial belonging to the i^{th} fuzzy class.

$$\overline{\mathbf{\Sigma}}_{\mathbf{i}} \quad = \quad \overline{\mathbf{X}}^i\overline{\mathbf{X}}^{i^\intercal} \quad i \in \{1, 2, ..., M\} \tag{3}$$

The obtained class covariance matrices are further normalized using

$$\overline{\mathbf{\Sigma}}_i = \frac{\overline{\mathbf{\Sigma}}_i}{Tr(\overline{\mathbf{\Sigma}}_i)}\ \forall i \in \{1, 2, \cdots M\}\ Tr : \text{denotes trace of a matrix} \tag{4}$$

Assuming the EEG time series trial $\mathbf{X}$, which is to be spatially filtered and one obtains say the spatial filters by some means. For ex: spatial filtering by Fuzzy CSP for regression problems gives $\mathbf{Z}$ where $\mathbf{Y} = \mathbf{Z}^{\intercal}\mathbf{X}$. Thus, the conditional distribution of post spatial filtered signals can be given by: $p_1 = \mathcal{N}(\mathbf{0}, \mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})$ and $p_2 = \mathcal{N}(\mathbf{0}, \mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z})$. One can compute the symmetric KL divergence between two distributions p_1 and p_2 as:

$$\begin{aligned} F(\mathbf{Z}) &= sD_{kl}(p_1 \parallel p_2) \\ &= D_{kl}(p_1 \parallel p_2) + D_{kl}(p_2 \parallel p_1) \\ &= \frac{1}{2}\left[\log\frac{det(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z})}{det(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})} + Tr((\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})^{-1}(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z}))\right] \\ &+ \frac{1}{2}\left[\log\frac{det(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})}{det(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z})} + Tr((\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})^{-1}(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z}))\right] \\ &- d \\ &= \frac{1}{2}\Big[Tr((\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})^{-1}(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z})) \\ &+ Tr((\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z})^{-1}(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z})) - 2d\Big] \end{aligned} \tag{5}$$

Here $F(\mathbf{Z})$ represents the reconstruction objective function for regression and represents the predictive power of regression method. The symmetric KL divergence [6] $sD_{kl}(p_1 \parallel p_2)$ between class conditionals of two fuzzy classes, after being filtered by a spatial filter is thus written as:

$$sD_{kl}(p_1 \parallel p_2) = \frac{1}{2}\left[\frac{\mathbf{z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{z}}{\mathbf{z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{z}} + \frac{\mathbf{z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{z}}{\mathbf{z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{z}} - 2\right] \tag{6}$$

(6) is similar to the CSP cost function in the RHS of (8).

$$\mathbf{Z}_{skl} = \arg\min_{\mathbf{Z}}\ sD_{kl}(\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z} \parallel \mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z}) \tag{7}$$

$$\mathbf{Z}^* = \arg\max_{\mathbf{Z}} \frac{\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_1\mathbf{Z}}{\mathbf{Z}^{\intercal}\overline{\mathbf{\Sigma}}_2\mathbf{Z}} \tag{8}$$

In (8), $\overline{\mathbf{\Sigma}}_1$ and $\overline{\mathbf{\Sigma}}_2$ are fuzzy-class covariance matrices. Here, $\mathbf{Z}^*$ denotes the filters generated from fuzzy CSP [24]. The space spanned by the spatial filters ($\mathbf{Z}_{skl}$) is same as that generated by the span of fuzzy-CSP filters. In other words,

$$span(\mathbf{Z}_{skl}) = span(\mathbf{Z}^*) \tag{9}$$

This can be proved by choosing a scalar function $z + \frac{1}{z} \in z \in \mathbb{R}$ whose minima is attained at $z = 1$. Hence, the filters optimizing sD_{kl} span the space similar to that spanned by multiple fuzzy CSP [24] filters.

4. DIVERGENCE REINFORCING STATIONARITY

Samek et al., [6] incorporated a variety of regularization terms to optimize the stationarity of the EEG signal within session, over the sessions for each subject, and crosswise various subjects. In this section, we incorporate stationarity in the objective function by proposing a novel divergence term for stationarity within session. In this work, we deal with the EEG based driving task, where we mainly analyze the stationarity within session for each subject. Stationarity is being optimized on training data. The regularization term $G(\mathbf{Z})$ is defined such that stationarity is optimized across each fuzzy class.

$$G(\mathbf{Z}) = \frac{1}{N_1 + N_2} \sum_{c=1}^{2} \sum_{j=1}^{N_c} \mu_{j,c} D_{kl}(\mathbf{Z}^\intercal \mathbf{\Sigma_{j,c}} \mathbf{Z} \parallel \mathbf{Z}^\intercal \mathbf{\Sigma_c} \mathbf{Z}) \tag{10}$$

Here, N_c denotes the number of trials per each fuzzy class. In (10), $\mathbf{\Sigma_{j,c}}$ and $\mathbf{\Sigma_c}$ denote the trialwise and classwise fuzzy covariance matrices. Here by, we formulate a composite cost function jointly optimizing both prediction and stationarity objectives.

$$\delta(\mathbf{Z}) = \lambda F(\mathbf{Z}) - (1-\lambda) G(\mathbf{Z}) \tag{11}$$

Here, λ is regularization constant. The optimization routine to be followed is the subspace method using gradient descent on an orthogonal manifold (cf. algorithm 1 in page 5 of [6]). In (11), we subtract the regularization term because we are maximizing the stationarity (which is minimizing divergence term $G(\mathbf{Z})$) and in parallel maximizing the predictive power of model for each fuzzy class (maximizing the divergence term $F(\mathbf{Z})$).

5. DIVERGENCE FRAMEWORK FOR MULTIPLE FUZZY-CLASSES

The above frameworks presented in sections 3 and 4 are applicable for two fuzzy classes. In this work, we use the fuzzy OVR (OneVersusRest) approach [24] and integrate it in divergence framework to optimize stationarity. This allows us to generalize the approach for multiple fuzzy classes.

5.1. Fuzzy OVR CSP in divergence framework

One can formulate a cost function (with reconstruction objective similar to $F(\mathbf{Z})$) for multiple fuzzy classes similar to binary fuzzy classes.

$$F^{OVR}{}_j(\mathbf{Z}) = sD_{kl}\left(\mathbf{Z}^\intercal \overline{\mathbf{\Sigma}}_j \mathbf{Z} \parallel \mathbf{Z}^\intercal \overline{\mathbf{\Sigma}}_{OVR_j} \mathbf{Z}\right) \tag{12}$$

$$\mathbf{Z}^*{}_j = \arg\max_{\mathbf{Z}} F^{OVR}{}_j(\mathbf{Z}) \tag{13}$$

$\mathbf{\Sigma}_{OVR_j} = \frac{1}{M}\sum_{k=1}^{M} \mathbf{\Sigma}_k$, where $M > 2$ is total number of fuzzy classes. We use $F = 3$ where F is the number of filters per each fuzzy class. Final filter matrix is obtained as $\left[\mathbf{Z}_1, ..., \mathbf{Z}_{3M}\right]$. Further, each of $\mathbf{Z}_j$ can be used for feature extraction from fuzzy CSP.

5.2. DivOVR-FuzzyCSP-WS (Within Session)

A framework similar to that in (10) and (11) can be extended for multiple fuzzy classes as below:

$$G(\mathbf{Z}) = \frac{1}{\sum_{r=1}^{M} N_r} \sum_{c=1}^{M} \sum_{j=1}^{N_c} \mu_{j,c} D_{kl}(\mathbf{Z}^\intercal \mathbf{\Sigma_{j,c}} \mathbf{Z} \parallel \mathbf{Z}^\intercal \overline{\mathbf{\Sigma}}_c \mathbf{Z}) \tag{14}$$

Here, N_r is number of trials in r^{th} fuzzy class, $\overline{\mathbf{\Sigma}}_c$ and $\mathbf{\Sigma_{j,c}}$ are fuzzy class covariance and trial covariance matrices respectively and membership function $\mu_{j,c}$ is defined for the j^{th} data trial belonging to the c^{th} fuzzy class. Further we can incorporate regularization too in the OVR framework

$$\mathbf{Z}^*{}_i = \lambda F^{OVR}{}_i(\mathbf{Z}) - (1-\lambda) G(\mathbf{Z}) \tag{15}$$

We calculate the spatial filter column for each OVR model by optimizing (15). λ is the regularization parameter. The combined filter matrix is then constructed from each OVR iteration.

Spatial filters $\mathbf{Z}$ can be optimized using following techniques

- Subspace Technique : A group of filters optimized together
- Deflation technique : Sequential optimization of filters

Filters ($\mathbf{Z}$) decomposed as a product of Whitening matrix ($\mathbf{W}$) and Orthogonal matrix ($\mathbf{R}$), 'd' denotes dimension of stationary subspace which can be tuned using cross validation.

$$\mathbf{Z}^T = \tilde{\mathbf{R}} W \quad \tilde{\mathbf{R}} = \mathbf{I}_d \mathbf{R} \quad \mathbf{Z} \in \mathbb{R}^{D \times d}, \mathbf{W} \in \mathbb{R}^{D \times D}$$
$$\mathbf{W}^T(\mathbf{\Sigma_1} + \mathbf{\Sigma_2})\mathbf{W} = \mathbf{I}$$

Optimization is done on an orthogonal manifold i.e. $\mathbf{R}\mathbf{R}^T = \mathbf{I}$

Objective functions now depend on orthogonal matrix $\mathbf{R}$

$$\Delta(\mathbf{R}) = \underbrace{(1-\lambda)\mathbf{F}^{OVR}{}_j(\mathbf{I}_d \mathbf{R})}_{\text{Fuzzy OVR CSP}} + \underbrace{\lambda G(\mathbf{I}_d \mathbf{R})}_{\text{Stationary}}$$

$d < D$: for the Subspace approach, $d = 1$: for the sequential optimization or Deflation approach

6. EVALUATION RESULTS AND DISCUSSION

8-fold cross-validation is used to compute the regression performance for each of the feature sets corresponding to different spatial filtering approaches $DivOVR-FuzzyCSP-WS$ and $FuzzyCSP$ respectively. Using the weight matrix $\mathbf{Z}$ calculated from various approaches of spatial filtering developed in the previous sections, one can obtain the projected EEG trial matrix:

$$\mathbf{X_{proj}} = \mathbf{ZX}^k;\ \text{here } X_{proj_i} \text{ denotes } i^{th} \text{ row of } \mathbf{X_{proj}} \quad (16)$$

$\mathbf{X}^k$ is the pre-processed EEG trial after passing through PREP [24] pipeline. The spatial filters are the rows of matrix $\mathbf{Z}$, from which we extract the final features, $\mathbf{F}=\begin{bmatrix} F_1 \\ .. \\ F_{FM} \end{bmatrix}$ where each F_i is given by $\log_{10} \frac{\|X_{proj_i}\|^2}{\sum_{j=1}^{FM} \|X_{proj_i}\|^2}$ Combining DivOVR-FuzzyCSP-WS generated feature $\mathbf{F}$ with LASSO regression led to an average RMSE and CC values in plots 3 and 4. In addition, we also computed features from $FuzzyCSP$ baseline and passed through LASSO regression leading to baseline average RMSE and CC values in plots 3 and 4. Plots 3 and 4 show that $DivOVR-FuzzyCSP-WS$ clearly outperformed $FuzzyCSP$ both interms of average RMSE and CC. Also, the effect of varying reglarization parameter λ on the average RMSE and CC is indicated within plot 5. We notice from figures 3 and 4 that proposed DivOVR-FuzzyCSP-WS approach registered lower RMSE and highest CC in comparison to baseline fuzzy CSP [24]. The values being 0.156 and 0.725 seconds respectively. The drowsiness prediction system presented here can estimate the driver reaction time with a mean RMSE error of 0.156 seconds. In other words, it means that the error in the estimated driving distance is around 4.3 meters under the constant speed 100 km/hr. In addition, we analyzed the effect of λ (regularization coefficient) on the performance of regression. We obtained that $\lambda = 0.5$ gave us the optimal performance of mean RMSE and CC. In addition, Student's t-test revealed statistically significant ($p < 0.01$) values for both RMSE and CC of the proposed $DivOVR-FuzzyCSP-WS$ approach over $FuzzyCSP$.

7. CONCLUSION AND FUTURE WORK

In this work, we propose divergence based fuzzy CSP for regression. We formulate a novel objective for regression using fuzzy covariance matrices. In addition, we incorporate stationarity into the objective function by proposing a novel divergence term for stationarity. Further, we generalize the approach for multiple fuzzy classes.

In future, it is interesting to look at the effect of using the affine transform based strategy conjointly with our proposed approach to incorporate both within and between session changes.

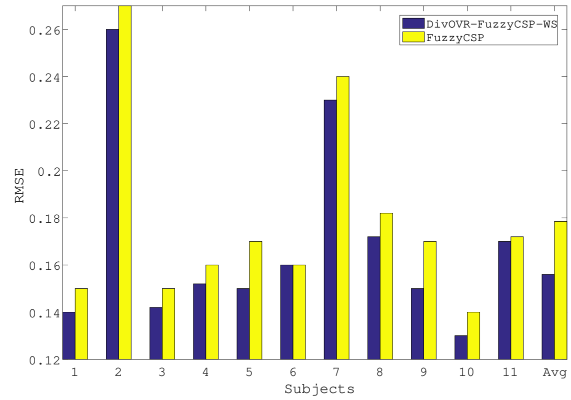

Fig. 3. RMSE of DivOVR-FuzzyCSP-WS on 11 subjects.

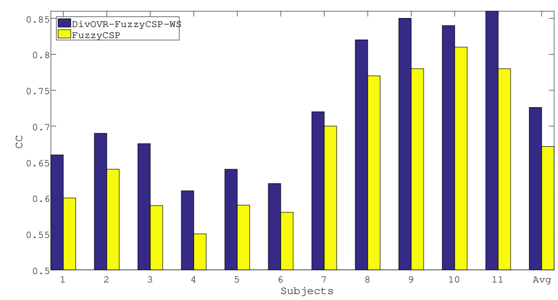

Fig. 4. CC of DivOVR-FuzzyCSP-WS on 11 subjects.

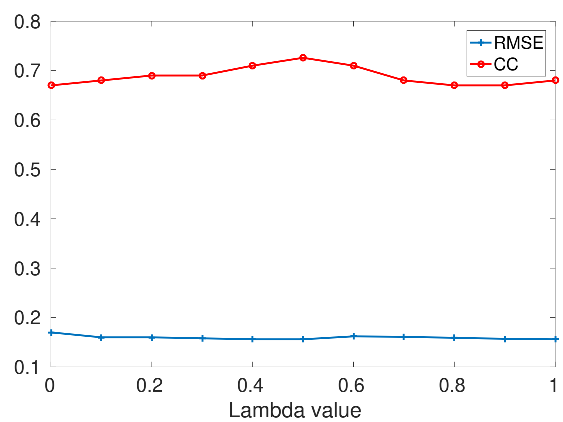

Fig. 5. RMSE and CC of DivOVR-FuzzyCSP-WS on 11 subjects while varying lambda (y-axis is in seconds).

8. REFERENCES

[1] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-Robert Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal processing magazine*, vol. 25, no. 1,

pp. 41–56, 2007.

[2] F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger, "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, pp. 031005, 2018.

[3] Wojciech Wojcikiewicz, Carmen Vidaurre, and Motoaki Kawanabe, "Stationary common spatial patterns: towards robust classification of non-stationary eeg signals," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011, pp. 577–580.

[4] Paul Von Bünau, Frank C Meinecke, Franz C Király, and Klaus-Robert Müller, "Finding stationary subspaces in multivariate time series," *Physical review letters*, vol. 103, no. 21, pp. 214101, 2009.

[5] Wojciech Samek, Klaus-Robert Müller, Motoaki Kawanabe, and Carmen Vidaurre, "Brain-computer interfacing in discriminative and stationary subspaces," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 2873–2876.

[6] Wojciech Samek, Motoaki Kawanabe, and Klaus-Robert Müller, "Divergence-based framework for common spatial patterns algorithms," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 50–72, 2014.

[7] Wojciech Samek and Klaus-Robert Müller, "Tackling noise, artifacts and nonstationarity in BCI with robust divergences," in *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, 2015, pp. 2741–2745.

[8] Benjamin Blankertz, Guido Dornhege, Christin Schafer, Roman Krepki, Jens Kohlmorgen, K-R Muller, Volker Kunzmann, Florian Losch, and Gabriel Curio, "Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 127–131, 2003.

[9] Johannes Müller-Gerking, Gert Pfurtscheller, and Henrik Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clinical neurophysiology*, vol. 110, no. 5, pp. 787–798, 1999.

[10] Moritz Grosse-Wentrup and Martin Buss, "Multiclass common spatial patterns and information theoretic feature extraction," *IEEE transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1991–2000, 2008.

[11] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten, "Multiclass brain–computer interface classification by Riemannian geometry," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 920–928, 2012.

[12] Wojciech Samek, "On robust spatial filtering of EEG in nonstationary environments," *it-Information Technology*, vol. 58, no. 3, pp. 150–154, 2016.

[13] S. Kumar, T. K. Reddy, V. Arora, and L. Behera, "Formulating Divergence Framework for Multiclass Motor Imagery EEG Brain Computer Interface," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1344–1348.

[14] Satyam Kumar, Tharun Kumar Reddy, and Laxmidhar Behera, "Divergence Framework for EEG based Multiclass Motor Imagery Brain Computer Interface," *arXiv preprint arXiv:1901.07457*, 2019.

[15] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, "A Survey on State-of-the-Art Drowsiness Detection Techniques," *IEEE Access*, vol. 7, pp. 61904–61919, 2019.

[16] RP Alvarez Gil, Zsolt Csaba Johanyák, and Tamás Kovács, "Surrogate model based optimization of traffic lights cycles and green period ratios using microscopic simulation and fuzzy rule interpolation," *Int. J. Artif. Intell*, vol. 16, no. 1, pp. 20–40, 2018.

[17] Qing Lu, Peng Shi, Hak-Keung Lam, and Yuxin Zhao, "Interval type-2 fuzzy model predictive control of nonlinear networked control systems," *IEEE Transactions on Fuzzy systems*, vol. 23, no. 6, pp. 2317–2328, 2015.

[18] Radu-Emil Precup, Marius-Lucian Tomescu, and Claudia-Adina Dragos, "Stabilization of Rössler chaotic dynamical system using fuzzy logic control algorithm," *International Journal of General Systems*, vol. 43, no. 5, pp. 413–433, 2014.

[19] Yu-Ting Liu, Yang-Yin Lin, Shang-Lin Wu, Chun-Hsiang Chuang, and Chin-Teng Lin, "Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network," *IEEE transactions on neural networks and learning systems*, vol. 27, no. 2, pp. 347–360, 2015.

[20] Yurui Ming, Yu-Kai Wang, Mukesh Prasad, Dongrui Wu, and Chin-Teng Lin, "Sustained attention driving task analysis based on recurrent residual neural network using EEG data," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2018, pp. 1–6.

[21] Tharun Kumar Reddy, Vipul Arora, Satyam Kumar, Laxmidhar Behera, Yu-Kai Wang, and Chin-Teng Lin, "Electroencephalogram based reaction time prediction with differential phase synchrony representations using co-operative multi-task deep neural networks," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 5, pp. 369–379, 2019.

[22] Tharun Kumar Reddy, Vipul Arora, Laxmidhar Behera, Yukai Wang, and CT Lin, "Multi-class Fuzzy Time-delay Common Spatio-Spectral Patterns with Fuzzy Information Theoretic optimization for EEG based Regression Problems in Brain Computer Interface (BCI)," *IEEE Transactions on Fuzzy Systems*, 2019.

[23] Tharun Kumar Reddy, Rupam Biswas, Vipul Arora, Vinay Gupta, and Laxmidhar Behera, "EEG Reaction Time-prediction with fuzzy Independent phase locking value representations using Lagrangian based Deep Neural Networks," submitted to IEEE Journal on Selected Topics in signal processing, 2020.

[24] Dongrui Wu, Jung-Tai King, Chun-Hsiang Chuang, Chin-Teng Lin, and Tzyy-Ping Jung, "Spatial filtering for EEG-based regression problems in brain–computer interface (BCI)," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 2, pp. 771–781, 2017.

[25] Wojciech Samek, Duncan Blythe, Klaus-Robert Müller, and Motoaki Kawanabe, "Robust spatial filtering with beta divergence," in *Advances in Neural Information Processing Systems*, 2013, pp. 1007–1015.

[26] Lubin Meng, Chin-Teng Lin, Tzyy-Ping Jung, and Dongrui Wu, "White-Box Target Attack for EEG-Based BCI Regression Problems," in *International Conference on Neural Information Processing*. Springer, 2019, pp. 476–488.

[27] Chun-Shu Wei, Yuan-Pin Lin, Yu-Te Wang, Chin-Teng Lin, and Tzyy-Ping Jung, "A subject-transfer framework for obviating inter-and intra-subject variability in EEG-based drowsiness detection," *NeuroImage*, vol. 174, pp. 407–419, 2018.

[28] Tharun Kumar Reddy, Vipul Arora, and Laxmidhar Behera, "HJB-equation-based optimal learning scheme for neural networks with applications in brain–computer interface," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2018.