

PAPER • OPEN ACCESS

Improving Session-to-session Transfer Performance of Emotion Recognition Using Adaptive Support Vector Machine

To cite this article: Kai Yang *et al* 2020 *J. Phys.: Conf. Ser.* **1601** 042028

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Improving Session-to-session Transfer Performance of Emotion Recognition Using Adaptive Support Vector Machine

Kai Yang¹, Guangcheng Bao¹, Ying Zeng^{2,1}, Li Tong^{1,3}, Jun Shu¹ and Bin Yan¹

¹Henan Key Laboratory of Imaging and Intelligent Processing, PLA Strategic Support Force Information Engineering University, Henan Zhengzhou,

²Key Laboratory for Neuro Information of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Sichuan Chengdu

³E-mail: tttocean@163.com

Abstract. The non-stationarity of electroencephalograph (EEG) signals has been a barrier in real-life application of EEG-based emotion recognition. The features extracted from emotional EEG vary from one session to another and a model trained on a temporally-limited EEG dataset may generalize poorly to data recorded at a different time for the same individual. In this study, the EEG features that stably characterize the differences between emotions are firstly explored. Then the classic progressive transductive support vector machine (PTSVM) is extended to three-classes classifications by a new region labelling rule, furthermore K-nearest neighbour algorithm and iterative process are utilized to improve the confidence of the predicted labels. Experimental results indicate that high gamma band local activation difference features and network features derived from prefrontal lobe, the temporal lobe on both sides, the left occipital-parietal lobe, the right occipital lobe and the posterior occipital lobe are temporally stable features for distinguishing different emotions (positive, neutral and negative). And the proposed session-to-session transfer model achieves an averaged classification accuracy 63.56% for three emotions classification on our dataset recorded from 23 subjects in three different days, which is 4% higher than the existing best adaptive classification model. Hence, the proposed classification model can effectively improve the performance of EEG-based session-to-session emotion recognition.

1. Introduction

Emotion is the psycho-physiological reaction of people to the stimulation from outside world or themselves, which plays an important role in our work and daily life [1]. However, EEG is a non-stationary signal and its statistics vary from session-to-session. As a result, a model trained on a temporally-limited EEG dataset may generalize poorly to data recorded at a different time for the same individual. Among current studies, there are two main methods to solve the challenge. One is to find out temporally stable emotional EEG features which could reflect innate character of emotion [2]. Another is to design an adaptive classifier model which could handle the variations of different days emotional EEG [3].

Researchers have reported several useful and stable EEG features in session-to-session emotion recognition experiments. Steinn et al[4] recorded 10 times resting state EEG of elderly subjects over a



2-month period. They found the most temporally stable feature was power spectral parameter, followed by regularity measures based on entropy and complexity. Lan et al[5] studied 5 peoples' emotion EEG which was recorded in 16 sessions over 8 consecutive days, the stability of four types EEG features was quantified by the intra-class correlation coefficient, the results showed that the 2nd to 6th statistical features were most stable, followed by fractal dimension(FD). These studies illustrate that there are exactly stable features and classification results can be improved using the features.

Among methods proposed to solve the session-to-session emotion recognition problem, adaption is a promising method. Bamdadian et al proposed an adaptive extreme learning machine to update the initial classifier from the calibration session by using EEG data from evaluation session and the classification accuracy was improved[3]. Support vector machine (SVM) which suits small sample size and high feature dimension classification issues also has been a popular classifier in many classification issues[6]. Martin et al proposed an adaptive SVM to increase the classification accuracy of a MEG-based brain-computer interface experiment over 2 sessions [7]. Joachims introduced transductive learning into the field of support vector classification and proposed transductive support vector machine (TSVM) [8]. Afterwards, Chen et al improved the TSVM algorithm and proposed progressive transductive support vector machine (PTSVM)[9], they achieved better testing result than that of TSVM approach.

Although many researchers have proposed many adaptive models for session-to-session classification based on SVM, there is no model suits for three classes classification issue. Here, according to the voting rule of SVM, we proposed a novel region labelling rule which extends PTSVM to a three classes classifier model and the K-nearest neighbor algorithm and iterative process are utilized to improve the confidence of the predicted labels.

2. MATERIALS AND METHOD

2.1. Experiment

2.1.1. Participants. The experiment is performed with 23 local native Chinese undergraduate or graduate students (11 females, mean age 22.3, ranging from 19-24). All participants are right-handed, and had normal or corrected-to-normal vision. Before the experiment, all participants are informed about the experiment and signed an informed consent form. After experiment, all participants would get some allowance.

2.1.2. Stimuli. We select 540 emotional pictures from Chinese Affective Picture System (CAPS) consisting of 180 positive pictures, 180 neutral pictures, and 180 negative pictures based on normative valence and arousal ratings [10]. According to normative ratings of CAPS, three categories pictures are different in valence degree (positive: $M=6.87$, $SD=0.26$; neutral: $M=5.36$, $SD=0.31$; negative: $M=2.59$, $SD=0.47$), as do in arousal degree (positive: $M=5.48$, $SD=0.46$; neutral: $M=4.37$, $SD=0.40$; negative: $M=6.04$, $SD=0.54$).

2.1.3. Experiment procedure. The experiment including three sessions in 3 days, the second session is 3 days after the first session, the third session is one week after the second session. In each session, participants watch 180 pictures, 60 pictures for positive, neutral and negative, respectively. And the 180 pictures are divided into 9 blocks which including 20 pictures of same category, the two same categories of pictures did not display in consecutive blocks. Each block consisted of 20 trails which are started with a "+" followed by an emotional picture display for 5 seconds.

During the experiment, EEG signals are recorded continuously from 62 Ag/Ag-Cl scalp electrodes, using g.Hlamp System (g.tec Medical Engineering, Linz, Austria) with a sample rate of 512 Hz. The Fz electrode and earlobe are used as recording reference, resulting in 61 effective electrodes. The EEGs are filtered by online band-pass filter and notch filter, 0.1-100Hz and 50Hz, respectively.

2.2. Signal pre-processing and features extracting

Before extracting features, some pre-processing procedures are carried out to exclude artifacts. Pre-processing procedures are data segment (500ms before the stimuli onset, 5000ms after), baseline correcting, filtering (0.1-80 Hz) and Fast-ICA. Additionally, a threshold of $\pm 100\mu\text{V}$ were also used to exclude artifacts with high amplitudes. After pre-processing, there were about averaged 52 data samples of each emotion category for one session.

In this paper, we not only extract features which could reflect activation difference from time domain, frequency domain, time-frequency domain, but also calculate brain function network properties which reflect the information propagation and processing patterns in the brain. The features were FD from time domain, power from frequency domain, differential entropy (DE) from time-frequency domain. Four network properties, clustering coefficient (CC), characteristic path length (CPL), local efficiency (Le), global efficiency (Ge) were calculated as whole brain features. These features are all extracted from five sub-frequency-bands, theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), low gamma (30-50Hz), high gamma (50-80Hz) bands.

2.3. Stable emotional EEG feature analysis

There are many feature selection methods, such as Max-Relevance and Min-Redundancy (MRMR)[11], Person correlation, but most of these methods rely on data's real label, while in practical application, the testing samples are unlabelled data. In this work, sequential backward feature selection algorithm (SBFSA) is used to analyse stable EEG features [12]. The main idea of this method is to remove one feature from the current feature set in turn in each iteration, and based on the new feature set, we get a new classification accuracy. According to the accuracy curve, we could find out the features that achieve highest accuracy, i.e. most stable features. Here, the feature set contains features that achieve better performance in same session classification. And the classification process use first and second sessions data as training set, and the third session data as testing set.

2.4. Session-to-session Transfer Model

The classic PTSVM has two drawbacks, firstly, the method only suits for two class classification; secondly, the efficiency of the "pairwise-labelling" method is very low for it only label one or two unlabelled samples in each iteration. To overcome those shortcomings, we propose a three-class PTSVM (TCK-PTSVM) model which suits for three categories classification. The method mainly includes two steps, firstly, according to the voting rule of SVM, the "pairwise-labelling" method is extended to three classes classification and propose the three-class labelling condition; secondly, K-nearest neighbour (KNN) algorithm is introduced to "region label rule" to improve the confidence of predicted labels.

LIBSVM is used as basic classifier [13], for N classes classification problem, LIBSVM would take one-to-one strategy to construct $N(N-1)/2$ binary classifiers, then according to voting principle, the sample would be classified into the class with the most votes. For three classes (class1, class2 and class3) classification, LIBSVM generates three binary classifiers, so each unlabelled sample corresponds to three decision functions i.e. $f_1(x)$ for class 1 and class 2, $f_2(x)$ for class 1 and class 3, $f_3(x)$ for class 2 and class 3. For a testing sample, the predict label will be 1 or 2 or 3, when decision functions meeting the condition (1), (2), (3), respectively.

$$f_1(x) > 0 \ \& \ f_2(x) > 0 \quad (1)$$

$$f_1(x) < 0 \ \& \ f_3(x) > 0 \quad (2)$$

$$f_2(x) < 0 \ \& \ f_3(x) > 0 \quad (3)$$

According to the characteristic of the SVM classifier, the absolute value of the decision function is proportional to the confidence of the predict label of the unlabelled sample. Therefore, the "three-class

region label” rule is given as, the unlabelled sample will be labelled 1 or 2 or 3, when decision functions meet the condition (4), (5), (6), respectively.

$$\max_1 \geq |f_1(x)| + |f_2(x)| \geq \max_1 - b_1 \quad (4)$$

$$\max_2 \geq |f_1(x)| + |f_3(x)| \geq \max_2 - b_2 \quad (5)$$

$$\max_3 \geq |f_2(x)| + |f_3(x)| \geq \max_3 - b_3 \quad (6)$$

And $\max_1$, $\max_2$, $\max_3$ are denoted as follows:

$$\max_1 = \text{Max}(|f_1(x_i)| + |f_2(x_i)|), \text{ s.t. } f_1(x_i) > 0 \& f_2(x_i) > 0 \quad (7)$$

$$\max_2 = \text{Max}(|f_1(x_i)| + |f_3(x_i)|), \text{ s.t. } f_1(x_i) < 0 \& f_3(x_i) > 0 \quad (8)$$

$$\max_3 = \text{Max}(|f_2(x_i)| + |f_3(x_i)|), \text{ s.t. } f_2(x_i) > 0 \& f_3(x_i) > 0 \quad (9)$$

Where b_1, b_2, b_3 are parameters which control the size of the labelled region, and their ranges are $0 < b_1 < \max_1$, $0 < b_2 < \max_2$, $0 < b_3 < \max_3$. When the value of b_1, b_2, b_3 is 0, and the “ $\geq$ ” in (7)(8)(9) changes to “ $=$ ”, the current “region label rule” is similar to “pairwise-labelling”, here we denote this labelling rule as “triple-labelling”. The “triple-labelling” means to label 1-3 unlabelled samples at a time according to the principle that the sum of the absolute values of the decision function is the maximal.

The major steps of TCK-PTSVM are as follows:

Step 1: Specify the penalty factors C and effect factor C^* , training an initial SVM model with all labelled samples, calculating the decision functions of all unlabelled samples.

Step 2: Using the proposed “multi-class region label rule” to label the unlabelled samples, adding the new labelled unlabelled samples to training set and eliminating the corresponding samples from testing set.

Step 3: Retrain the SVM model with the new training set getting from step 2, calculating the decision function values of all the unlabelled samples which added to the training set, remove the inconsistent labelling samples and add the samples to testing set as unlabelled samples.

Step 4: Repeat steps 2 and 3 until there are no unlabelled samples meet the labelling conditions, and use the current classifier to classify the remaining unlabelled samples in the testing set. The algorithm ends and output the classification result.

In order to illustrate the use of KNN can improve the classification results, we denoted another three-class PTSVM (TC-PTSVM) model which doesn’t use KNN, TC-PTSVM has same steps except for KNN process.

3. Results

We firstly use the extracted emotional EEG features in within-session emotion classification to find out features that can better depict the difference of emotions. Then the selected features are combined as fusion features in session-to-session emotion classification to analyse the most stable features. Then the stable features are used to test the performance of the proposed adaptive classification model.

3.1. Within-session classification results

Figure 1 shows the classification results of 31 features for each session. From the classification results, we can find that all extracted features are effective to classify three emotions and the best performing feature was high gamma band DE, it achieved 85.52%, 87.98%, 86.43% on first session, second session and third session EEG data, respectively.

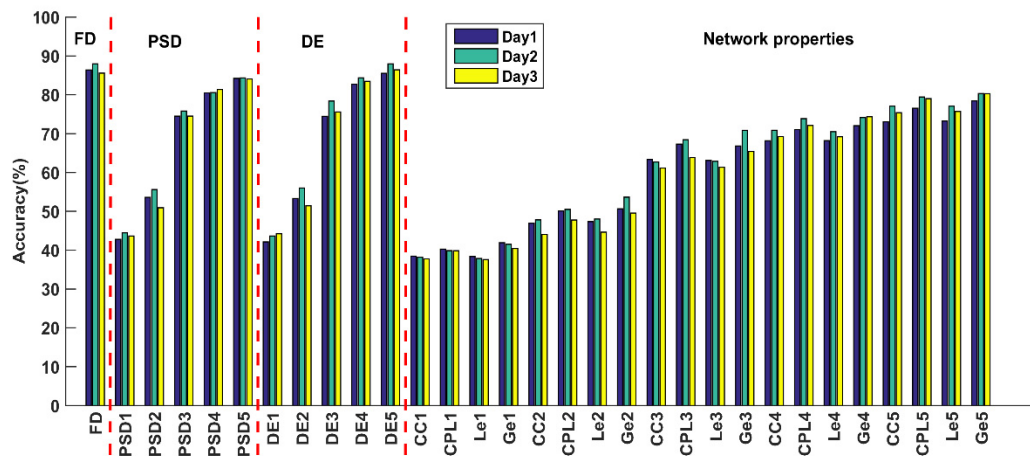


Figure 1. The averaged classification accuracies of 31 features in three sessions. “1,2,3,4,5” represents theta band, alpha band, beta band, low gamma band, high gamma band, respectively.

Figure 2 shows the performance of different dimensionality features selected by SBFSA. The classification accuracy achieves highest accuracy 47.08% when the top 120 features selected by SBFSA were utilized for positive, neutral and negative three emotions recognition. It can be seen that when the dimensionality of feature is smaller than 120, accuracy increases as the dimensionality number of selected features increases, but when the dimensionality number is bigger than 120, the accuracy decreases as the dimensionality number of features increases.

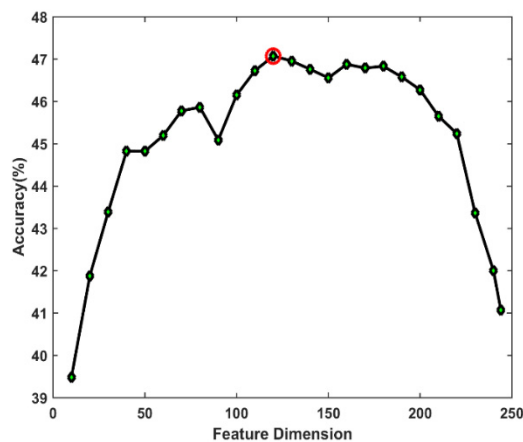


Figure 2. The averaged classification accuracies of different dimension features selected by sequential backward selection algorithm. The red circle marks the accuracy of top 120 features.

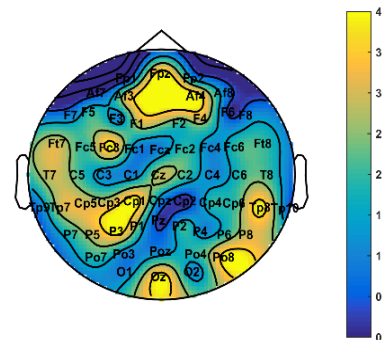


Figure 3. The distribution of top 120 features selected by SBFSA.

We further analyse the composition of top 120 features, they are 33 features of FD, 23 features of high gamma band power, 30 features of high gamma band DE, 34 features of high gamma band Ge. Figure 3 shows the distribution of top 120 features selected by SBFSA. As can be seen from the figure, top 120 features mainly derive from prefrontal, left temporal, right occipital and posterior occipital lobes.

Based on the most stable 120 features selected by SBFSA, three classifier model SVM, TC-PTSVM, TCK-PTSVM are utilized to handle session-to-session transfer emotion recognition. Figure 4 shows the classification accuracies of three classifier models, the accuracies are 47.08%, 58.59%, 63.56% for SVM, T-PTSVM, TC-PTSVM, respectively. Due to the introduction of unlabelled testing

samples information, both the accuracies of TC-PTSVM and TCK-PTSVM are higher than the accuracy of SVM ($p < 0.05$). And the accuracy of TCK-PTSVM is also higher than that of TC-PTSVM ($p < 0.05$).

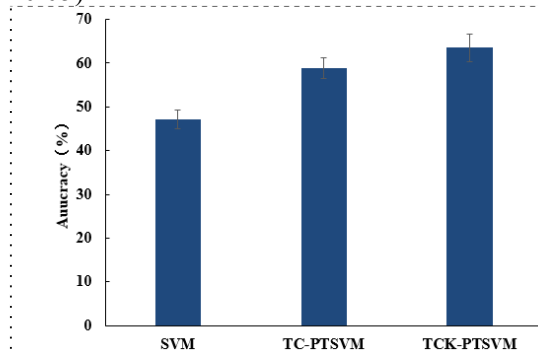


Figure 4. The classification accuracies of three classifier models.

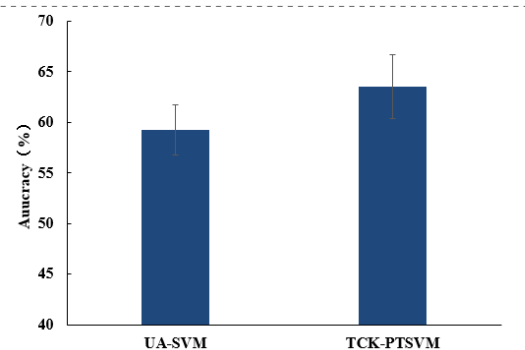


Figure 5. The comparison of classification results between UA-SVM and TCK-PTSVM.

In order to verify the effectiveness of the TCK-PTSVM adaptive classification model proposed in this paper, the unsupervised adaption of the SVM (UA-SVM) classification model proposed by Martin et al[14]. in the time-transfer classification experiment is introduced for performance comparison. Figure 5 shows the comparison of classification results between UA-SVM and TCK-PTSVM, which the classification accuracy of 59.26% and 63.56%, respectively. It can be seen that the classification accuracy of the TCK-PTSVM is on average 4% higher than that of UA-SVM ($p < 0.05$).

4. Conclusion

This work aims to improve the session-to-session emotion recognition performance. We firstly explored the stable features that can stably characterize the differences between emotions. The analysis results show that features from prefrontal lobe, the temporal lobe on both sides, the left occipital-parietal lobe, the right occipital lobe and the posterior occipital lobe are temporally stable features for distinguishing different emotions (positive, neutral and negative). Then, TCK-PTSVM model is proposed for session-to-session emotion recognition. This model extends classic PTSVM to three classes classification and K-nearest neighbor algorithm and iterative process are introduced to improve the confidence of the predicted labels to further improve the classification accuracy. TCK-PTSVM achieves an average accuracy of 63.56%, which is 4% higher than current best adaptive session-to-session model. The results might be useful for real-life application of EEG-based emotion recognition.

Acknowledgments

This work was supported by the National Key R&D Program of China under grant 2017YFB1002502, the National Natural Science Foundation of China (No. 61701089, No.61601518 and No. 61372172) and the National Defense Science and Technology Innovation Zone Project.

References

- [1] Zhang, D.; Wan, B.; Ming, D. Research progress on emotion recognition based on physiological signals. *Journal of Biomedical Engineering* 2015, 32, 229–234.
- [2] Zirui Lan; Olga Sourina. Real-time EEG-based emotion monitoring using stable features. *Visual Computer International Journal of Computer Graphics* 2015, 32, 347–358.
- [3] Bamdadian, A.; Guan, C.; Kai, K.A.; Xu, J. Improving session-to-session transfer performance of motor imagery-based BCI using adaptive extreme learning machine. In *Proceedings of the International IEEE Engineering in Medicine and Biology Society of Conference*; Osaka,

- Japan, 2013; pp. 2188–2191.
- [4] Gudmundsson, S.; Runarsson, T.P.; Sigurdsson, S.; Eiriksdottir, G.; Johnsen, K. Reliability of quantitative EEG features. *Clinical Neurophysiology Official Journal of the International Federation of Clinical Neurophysiology* 2007, 118, 2162–2171.
 - [5] Zirui Lan; Olga Sourina. Real-time EEG-based emotion monitoring using stable features. *Visual Computer International Journal of Computer Graphics* 2015, 32, 347–358.
 - [6] Hai, Y.; He, W.; Fan, L. An incremental learning algorithm for SVM based on voting principle. In *Proceedings of the Advanced Information Management and Service (IMS)*; Seoul, South Korea, 2010.
 - [7] Spüler, M.; Rosenstiel, W.; Bogdan, M. Adaptive SVM-Based Classification Increases Performance of a MEG-Based Brain-Computer Interface (BCI). In *Proceedings of the International Conference on Artificial Neural Networks*; Lausanne, Switzerland, 2012; Vol. 7552, pp. 669–676.
 - [8] JOACHIMS, T. Transductive Inference for Text Classification using Support Vector Machines. In *Proceedings of the Proc. of the 16th International Conference on Machine Learning*; 1999; pp. 200–209.
 - [9] Chen, Y.; Wang, G.; Dong, S. Learning with progressive transductive support vector machine. *Pattern Recognition Letters* 2003, 24, 1845–1855.
 - [10] Lu, B.; Hui, M.A.; Huang, Y.X. The Development of Native Chinese Affective Picture System- A pretest in 46 College Students. *Chinese Mental Health Journal* 2005, 19, 719–722.
 - [11] Ding, C.; Peng, H. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics & Computational Biology* 2005, 3, 185–205.
 - [12] Reeves, S.J.; Zhe, Z. Sequential algorithms for observation selection. *IEEE Transactions on Signal Processing* 1999, 47, 123–132.
 - [13] Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2011, 2, 1–27.
 - [14] Spüler M, Rosenstiel W, Bogdan M. Adaptive SVM-Based Classification Increases Performance of a MEG-Based Brain-Computer Interface (BCI)[A]. *International Conference on Artificial Neural Networks*[C]. Lausanne, Switzerland: 2012, 7552: 669–676.