ARTICLE IN PRESS

Methods xxx (xxxx) xxx

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Subject matching for cross-subject EEG-based recognition of driver states related to situation awareness

Ruilin Li [a,b], Lipo Wang [a,*], Olga Sourina [b,*]

[a] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore
[b] Fraunhofer Singapore, Singapore 639798, Singapore

ARTICLE INFO

Keywords:
Situation awareness
Electroencephalography (EEG)
Transfer learning
Machine learning
Classification

ABSTRACT

Situation awareness (SA) has received much attention in recent years because of its importance for operators of dynamic systems. Electroencephalography (EEG) can be used to measure mental states of operators related to SA. However, cross-subject EEG-based SA recognition is a critical challenge, as data distributions of different subjects vary significantly. Subject variability is considered as a domain shift problem. Several attempts have been made to find domain-invariant features among subjects, where subject-specific information is neglected. In this work, we propose a simple but efficient subject matching framework by finding a connection between a target (test) subject and source (training) subjects. Specifically, the framework includes two stages: (1) we train the model with multi-source domain alignment layers to collect source domain statistics. (2) During testing, a distance is computed to perform subject matching in the latent representation space. We use a reciprocal exponential function as a similarity measure to dynamically select similar source subjects. Experiment results show that our framework achieves a state-of-the-art accuracy 74.32% for the Taiwan driving dataset.

1. Introduction

Endsley [1] defined Situation Awareness (SA) as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”. In recent years, not only dynamic system design requires measurement of SA [2], but also SA is indispensable to evaluating and training the operators in dynamic systems [3]. Therefore, SA has received increasing attention in recent years. Researchers are concerned about how to measure SA efficiently and precisely in practice. Different from conventional evaluation methods, such as SAGAT, SPAM, etc. [4], electroencephalography (EEG) is one of the most commonly used signals to record brain activities [5]. Moreover, EEG has become one of the most favored ways to assess SA, due to its non-intrusiveness and objectivity. Recently, researchers adopted machine learning algorithms to perform recognition based on EEG [6–8] (or based on facial expression [9–12]). However, EEG-based SA recognition is restricted by the subject variability problem, which also occurs commonly in the related fields of workload [13] and motor imaginary [14] recognition.

A solution to the subject variability problem is to find subject-invariant features. Like previous works [15,16], we also treat this problem as a “domain shift” problem which was first proposed in image processing [17]: each subject constitutes a domain himself and EEG data distribute differently across different domains. The “domain shift” problem in EEG can be attributed mainly to four reasons: a) Individual differences in human brain functional and anatomical connections [18], b) misregistration during data collection from different skull shapes across subjects, c) changes of environment and sensor states in different experiment sessions and days, and d) variations in subjects’ other mental states, emotional conditions, and task-irrelevant brain activity disturbances [19].

Up to now, in EEG processing, state-of-the-art approaches dealing with the “domain shift” problem mainly employ unsupervised domain adaptation [20], a transfer leaning technique that transfers knowledge from the source domain to the target domain. Domain adaptation techniques have made great progress in image processing [21–23] and have been applied in EEG processing, for example, Transferable Component Analysis (TCA) [24] and Maximum Independence Domain Adaptation (MIDA) [25]. Domain adaptation exploits the labeled source domain data and unlabeled target domain data to perform feature alignment. However, target subjects are usually unknown during training. Domain generalization (DG), another popular branch of

* Corresponding authors.
E-mail addresses: elpwang@ntu.edu.sg (L. Wang), EOSourina@ntu.edu.sg (O. Sourina).

https://doi.org/10.1016/j.ymeth.2021.04.009
Received 2 February 2021; Received in revised form 5 April 2021; Accepted 7 April 2021

transfer learning, is more suitable when dealing with unknown domains. Traditional machine learning assumes that training and testing data are identically and independently distributed (i.i.d.). However, this assumption may not hold in applications like multi-subject EEG. Domain generalization (DG) aims at generalizing from source domains to unseen target domains with different data distributions from the source domains. In computer vision, various domain generalization frameworks are designed to generalize from source domains to target domains [26–28].

In EEG processing, most of the previous works using transfer learning considers training data from different subjects a whole source domain and considers the test subject data the target domain. In deep learning, batch normalization (BN) layer [29] uses statistics of training data to normalize test data. However, if we combine the different subjects' data with different distributions as a whole training set, the obtained statistics will not precisely represent any subject. This combination could lead to the potential risk of not obtaining desired compatible features. Therefore, we consider the problem in an opposite direction: leveraging on the subject-specific information. Wei et al. [30] and Wang et al. [31] exploited the small part of test subject data with label information to select source subjects for training. However, in practice, the test subject is totally unknown during training. [32] did not use label information and transformed the resting-state of the EEG signals into a frequency domain. The power spectral density (PSD) features were computed. The authors then computed the similarity by applying the cosine distance on the obtained features. However, extracting and selecting proper features for various tasks are complicated and time-consuming. In our work, we exploit domain-specific batch normalization (DSBN). DSBN uses specific

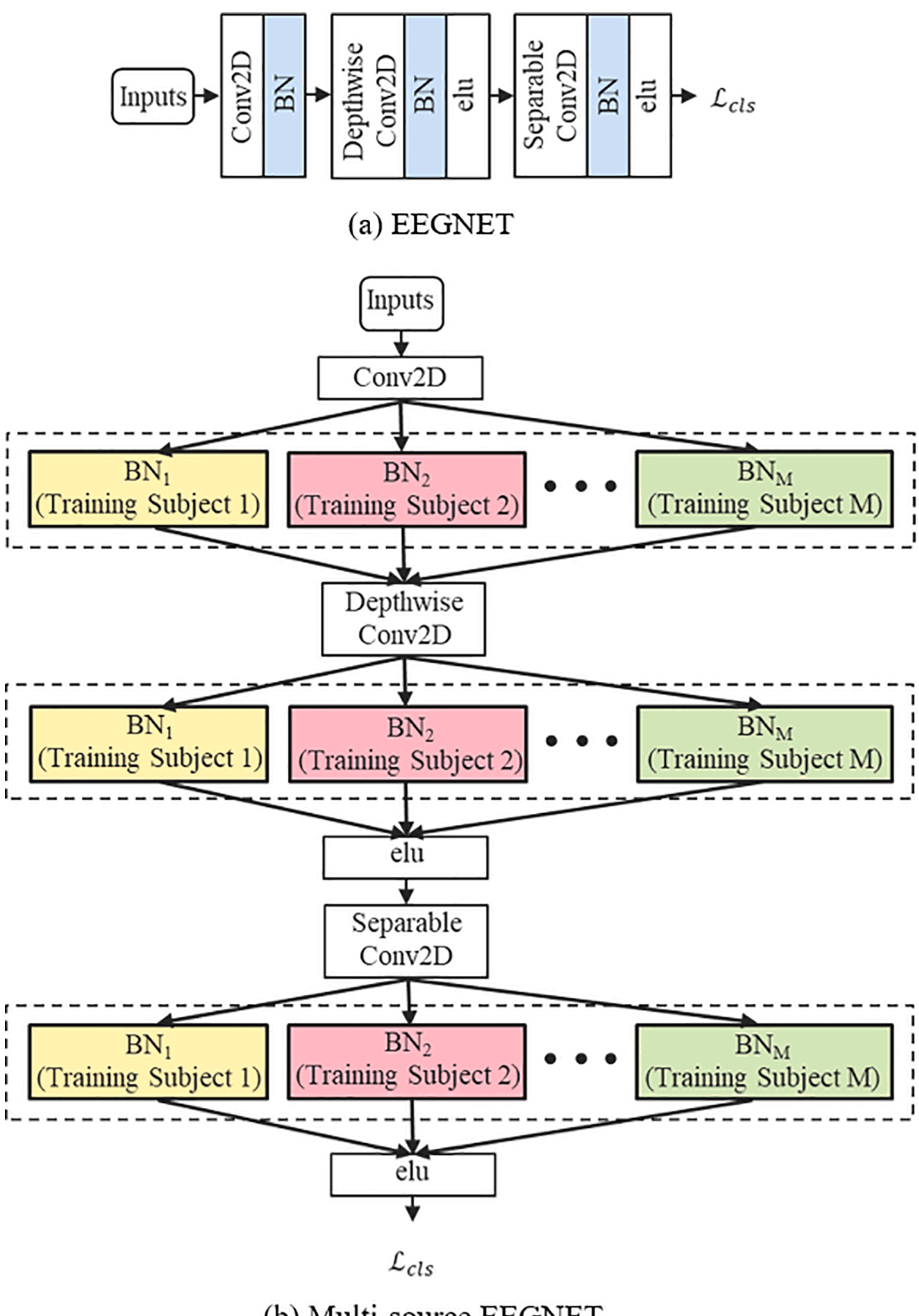

Fig. 1. Multi-source EEGNET architecture. (a) and (b) illustrates the replacement of the original BN layer by a multi-source domain alignment (mDA) layer (in the dashed boxes).

batch normalization statistics for each known domain to independently align source and target distributions, which was first introduced in image processing [33,34]. Multi-source DSBN has recently been exploited in domain adaptation [35] and domain generalization [36,37]. In domain generalization, the relationship between BN and instance normalization (IN) statistics is exploited. When performing domain adaptation between domains with different styles, BN and IN are combined into a new normalization parameter or used to map the source and target domains in a shared latent space. In our work, we find that BN can effectively reduce subject variability.

We propose a simple but efficient subject matching framework which exploits subject-specific statistics and selects the similar subjects' network outputs dynamically and adaptively. Specifically, we replace the BN layer with a multi-source domain alignment layer and the remaining network shares the weights. During training, the domain-specific BN statistics (BNS) of source domains are extracted and stored. During testing, same BN statistics of the target domain is computed. Then, a reciprocal exponential function is applied to the distance between BNS of both domains to obtain the similarity. Finally, we use the similarity to weight the output from each source domain to obtain the recognition result of the target domain. The contributions of our work are summarized as follows:

- We propose a novel subject matching framework that automatically match the test subject to the most similar training subjects.
- We find that BN statistics is reliable to reduce subject-variability and remove incompatible information of different subjects.
- We use a reciprocal exponential function to select similar training subjects during testing.

2. Materials and methods

2.1. Problem formulation and notation

Let $\mathscr{X}$ be the input space (e.g., EEG signals) and $\mathscr{Y}$ the output space (e.g., EEG signal categories) of our learning task. In domain generalization, we have M source domains: $\mathscr{S} = \{s_i\}_{i=1}^{M}$ that are identified via probability distributions $p_{xy}^{s_i} = p(y|x,s_i)$, defined over $\mathscr{X} \times \mathscr{Y}$, while the target domain is unknown. During training, we are provided with each source domain dataset $s_i = \{(x_1^{s_i}, y_1^{s_i}), \ldots, (x_m^{s_i}, y_m^{s_i})\}$ of i.i.d. observations from $p_{xy}^{s_i}$. We leverage multiple source domains to learn a mapping ($\mathscr{X}$, $\mathscr{S}$)$\rightarrow\mathscr{Y}$ that can generalize to the target domain.

Our main goal is to align the distributions of the target domain to source domains, i.e., correctly classifying the EEG signals in the target domain. When testing, the target dataset $\mathscr{T} = \{x_1^t, \ldots, x_n^t\}$ of i.i.d. observations is from the marginal p_x^t. We compute similarity between the target domain and each source domain. Then based on the conditional distribution $\{p_{xy}^{s_i}\}_{i=1}^{M}$ of different branches, we combine the outputs as the final output for the target domain.

We denote $F(\cdot)$ as the output of a forward pass in the model. For the source domain, the classification model is trained using the standard supervised loss [22]:

$$\mathscr{L}_{cls} = \mathbb{E}_{(x^s,y^s)\ (\mathscr{X},\mathscr{Y})}\left(-\sum_{i=1}^{m} \mathbb{1}_{[i=y^s]} logF(x^s)\right) \tag{1}$$

2.2. Multi-source domain alignment layer

BN [29] was originally designed to alleviate the issue of internal covariate shifting which is a common problem while training a very deep neural network. The function of a BN layer can be described as: For activations within a mini-batch of N samples, it first performs whitening among activations, and then learns affine parameters γ and β, which transform the inherent mean and variance to trainable variables. Denote input of a BN layer of each channel as $x \in \mathbb{R}^{N\times H\times W}$, where $H \times W$ is the size of a 2D feature map. The BN layer is expressed as:

$$BN(x;\gamma,\beta) = \gamma\cdot\widehat{x} + \beta, \tag{2}$$

where

$$\widehat{x} = \frac{x-\mu}{\sqrt{\sigma^2+\varepsilon}}, \tag{3}$$

where $\varepsilon > 0$ is a small constant to avoid numerical issues in case of zero variance. We use $x[n,h,w]$ to express the value of a single element of the feature map. The mean and variance of activations within a mini-batch, μ and σ^2, are computed by:

$$\mu = \frac{1}{N\times H\times W}\sum_n \sum_{h,w} x[n,h,w] \tag{4}$$

$$\sigma^2 = \frac{1}{N\times H\times W}\sum_n \sum_{h,w} (x[n,h,w]-\mu)^2 \tag{5}$$

During training, the BN layer also estimates the mean and variance of the entire activations by using an exponentially weighted moving average (EWMA) [38]. Formally, in t^{th} step, the moving mean $\widetilde{\mu}$ and moving variance $\widetilde{\sigma}^2$ are updated by using an attenuation factor α:

$$\widetilde{\mu}^{t+1} = (1-\alpha)\widetilde{\mu}^t + \alpha\mu^{t+1} \tag{6}$$

$$\left(\widetilde{\sigma}^{t+1}\right)^2 = (1-\alpha)(\widetilde{\sigma}^t)^2 + \alpha\left(\sigma^{t+1}\right)^2 \tag{7}$$

Here, some relationship between the attenuation factor and the number of update steps is introduced: when $\alpha = 0.9$, the moving average can be approximated to the weighted average of the latest 10 updates. $\alpha = 0.98$ corresponds to latest 50 updates, while $\alpha = 0.99$ corresponds to latest 100 updates.

In our work, a multi-source domain alignment layer [35,39] is employed, which is to separate the single BN layer to a set of BN branches corresponding to different source domains. Fig. 1 shows the architecture of our model and illustrates the replacement of the BN layers by multi-source domain alignment layers. We adopt EEGNET [40] as the backbone network of our framework. Firstly, a 2D convolutional layer is fitted with the filter length to be half of the sampling rate, which allows for capturing frequency information at 2 Hz and above. Secondly, depthwise convolution is used to learn a spatial filter. Then, a separable convolution is exploited, which is a combination of depthwise convolution and pointwise convolutions. Depthwise convolution learns a temporal summary for each feature map individually, followed by a pointwise convolution which learns how to optimally mix the feature maps together. Separable convolution reduces the number of parameters to fit. In our ensemble model, every network shares all the weights except the multi-source domain alignment layers. Then, each domain has its own specific affine parameters and BN statistics. Formally, we reformulated the Eqs. (2)–(7) as:

$$BN_s(x_s;\gamma_s,\beta_s) = \gamma_s\cdot\widehat{x}_s + \beta_s \tag{8}$$

where x_s is the input activations at each channel of the branch of source domain s.

$$\widehat{x}_s = \frac{x_s-\mu_s}{\sqrt{\sigma_s^2+\varepsilon}} \tag{9}$$

Domain-specific mean and variance is obtained by:

$$\mu_s = \frac{1}{N\times H\times W}\sum_n \sum_{h,w} x_s[n,h,w] \tag{10}$$

$$\sigma_s^2 = \frac{1}{N\times H\times W}\sum_n \sum_{h,w} (x_s[n,h,w]-\mu_s)^2 \tag{11}$$

Domain-specific moving mean and moving variance is computed by:

$$\widetilde{\mu}_s^{t+1} = (1-\alpha)\widetilde{\mu}_s^t + \alpha\mu_s^{t+1} \tag{12}$$

$$\left(\widetilde{\sigma}_s^{t+1}\right)^2 = (1-\alpha)\left(\widetilde{\sigma}_s^t\right)^2 + \alpha\left(\sigma_s^{t+1}\right)^2 \tag{13}$$

In conclusion, after training using the source domain dataset, for each domain-specific branch of the BN layer, we have individual affine parameters (γ_s, β_s) and BN statistics (BNS) $(\widetilde{\mu}_s, \widetilde{\sigma}_s^2)$. Moreover, with the replacement of the BN layers by multi-source domain alignment layers, the model $F(\cdot)$ is transformed to $\{F_s(\cdot)\}_{s\in\mathscr{S}}$.

The multi-source domain alignment layer extracts BN statistics of each source domain which remove the influence of the incompatible information and normalize the target domain using specific statistics of each branch. This variant BN layer can make the target domain data clearly align to each source domain, avoiding the alignment to the whole source domain which may have poor classification performance. Therefore, a model with the multi-source domain alignment layer is a suitable structure to alleviate the subject variability problem. When testing, the target domain data will separately flow through every branch and be combined to a final output. However, how the target domain is related to multiple source domains and how do we deal with possible noisy samples? Based on the BN statistics, we will give a detailed introduction in next subsection.

2.3. Subject matching

There are a lot of differences in EEG signals, especially across different subjects, because of the influence of other mental states, emotional condition, and task-irrelevant brain activity disturbance mentioned in Section 1. The data distributions of some source subjects are more similar to the test subject compared to the other source subjects. Then the subjects with poor similarity should be declined. The specific process is as follows.

2.3.1. Computing similarity

The multi-source domain alignment layer is derived from the original BN layer. As opposed to BN, the mDA layer computes the domain specific distribution $p^s_{x\to y}$, separating the BN layer of integral $F(\cdot)$ to a branched BN with $\{F_s(\cdot)\}_{s\in\mathscr{S}}$. Although this ensemble model can explicitly and concisely embody the distribution of each source domain, the generalization ability of each branch or even the entire network to an unknow target domain is still a question. During testing, we employ the 2-Wasserstein distance to compute the similarity between each source domain and the target domain. We select source domains by exploiting the reciprocal exponential function. The resulting target domain classification probabilities is a weighted mixture of the output of more similar source domains branches.

Let $l \in \mathscr{B} = \{1, 2, ..., L\}$ represents the different BN layers in the network. Then we define a latent representation space $\mathscr{H}$ across source domains. Specifically, we have $\mathscr{H}^l$ in l^{th} layer, which map the population statistics of target domain $(\widetilde{\mu}_t^l, (\widetilde{\sigma}_t^l)^2)$ to the counterpart of source domain $(\widetilde{\mu}_s^l, (\widetilde{\sigma}_s^l)^2)$. The BNS of all BN layers in the source domains are combined as:

$$BNS_s = \left[\left(\widetilde{\mu}_s^1, \widetilde{\sigma}_s^1\right), \left(\widetilde{\mu}_s^2, \widetilde{\sigma}_s^2\right), \cdots, \left(\widetilde{\mu}_s^L, \widetilde{\sigma}_s^L\right)\right] \tag{14}$$

During testing, we first pass the testing data through the network, and then extract the population statistics of the target domain from each BN layer. Finally, we combine them as:

$$BNS_t = \left[\left(\widetilde{\mu}_t^1, \widetilde{\sigma}_t^1\right), \left(\widetilde{\mu}_t^2, \widetilde{\sigma}_t^2\right), \cdots, \left(\widetilde{\mu}_t^L, \widetilde{\sigma}_t^L\right)\right] \tag{15}$$

Then we evaluate the similarity between BNS_s and BNS_t by computing the distance between the two points projected from both the source domain and the target domain. We assume that the input activations of each BN layer follow the Gaussian distribution: $z \sim \mathscr{N}(\widetilde{\mu}^l, \widetilde{\sigma}^l)$ and the similarity measure can be specified as the distance between two multivariate Gaussian distributions. We compute the 2-Wasserstein distance between $q_s \sim \mathscr{N}(\widetilde{\mu}_s^l, \widetilde{\Sigma}_s^l)$ and $q_t \sim \mathscr{N}(\widetilde{\mu}_t^l, \widetilde{\Sigma}_t^l)$, corresponding to the activation distributions of the source domain and the target domain, respectively, where $\widetilde{\Sigma}_t^l$ represents the covariance matrix.

$$W_2(q_s, q_t) = \|\widetilde{\mu}_s^l - \widetilde{\mu}_t^l\|_2^2 + Tr\left(\widetilde{\Sigma}_s^l + \Sigma_t^l - 2\left[\left(\widetilde{\Sigma}_s^l\right)^{1/2}\Sigma_t^l\left(\widetilde{\Sigma}_s^l\right)^{1/2}\right]^{1/2}\right) = \|\widetilde{\mu}_s^l - \widetilde{\mu}_t^l\|_2^2 + Tr\left(\left(\widetilde{\Sigma}_s^l\right)^{1/2} - \left(\widetilde{\Sigma}_t^l\right)^{1/2}\right)^2 \tag{16}$$

where $Tr(\cdot)$ is the trace of the matrix.

2.3.2. Source subject selection

Now, we want to select appropriate source subjects and eliminate the impact of subjects with large distances to the distribution of the target subject. Therefore, we propose to use reciprocal of exponential function to transform the distance to similarity. The result of the reciprocal exponential function with a large distance approaches zero, which can effectively eliminate the negative contribution. Once we obtain the distance measure of both domains, the similarities are computed by applying the exponential on the distance and we can obtain e^{d_s}, where $d_s = W_2(q_s, q_t)$. We define the similarity $r_{s,t}$ between source domain s and target domain t as:

$$r_{s,t} = \frac{\frac{1}{e^{d_s}}}{\sum_{s\in\mathscr{S}}\frac{1}{e^{d_s}}} \tag{17}$$

The exponential function can effectively detect the subjects with large distance. To avoid the exponential calculation of some large distances, we set 50 as the distance threshold. If distances are larger than 50, we set the result of reciprocal exponential function to zero. If all distances are larger than 50, we only use the most similar (1-nearest) subject data to compute the final result. After the selection of useful subjects, we can obtain the output distribution of the model on the target domain by weighted combination of the outputs of all branches.

$$p_t = \sum_{s\in\mathscr{S}} r_{s,t} p_{xy}^s \tag{18}$$

3. Experiments

3.1. Experiment settings

3.1.1. Data preparation

We use an open driving dataset in our experiments, which was collected during 2005–2012 and released in 2019 [41]. The dataset comprises 62 EEG datasets of 27 subjects (aged between 22 and 28) who were students or staff at National Chiao Tung University in Taiwan. The EEG signals were recorded in 32 channels (30 valid channels plus 2 reference channels), with a sampling frequency of 500 Hz. We further down-sample the data to 128 Hz.

Table 1
Situation awareness dataset content.

Subject ID	Subject Index	Number of Samples	
		High SA	Low SA
1	0	94	94
5	1	66	66
22	2	75	75
31	3	74	74
35	4	112	112
41	5	83	83
42	6	51	51
43	7	132	132
44	8	157	157
45	9	54	54
53	10	113	113
Total		1011	1011

In the experiment, lane-departure events were randomly induced to make the car drift from the original cruising lane towards the left or right sides (deviation onset). Each participant was instructed to quickly compensate for this perturbation by steering the wheel (response onset) to cause the car to move back to the original cruising lane (response offset). A complete trial included events with deviation onset, response onset, and response offset.

In our study, fatigue-related SA is analysed. We follow the fatigue labels of the data set because the fatigue and non-fatigue data were labelled according to the reaction time on the repeated lane-departure event that corresponds to the definition of high and low SA [42,43]. Specifically, we extract 3 s EEG data prior to the deviation onset as a measure of subject's SA before the start of the trial. Since the subjects' states cannot be specified before the response onset, that is, the data may mix both high and low SA, we did not use the EEG data between deviation onset and response onset. We apply the method described in [41] to extract SA related data as follows. We set SA labels based on the reaction time (RT), which is the length of the interval between the deviation onset and response onset. Additionally, global RT was defined as the average of local RTs across all epochs within a 90-second window before the deviation onset. The baseline alert-RT was defined as the 5th percentile of local RT in the entire session. The label process is as follows. When both the local and global RT are shorter than 1.5 times the alert-RT, the corresponding extracted EEG data is labelled as "low SA", and when both the local and global RT are longer than 2.5 times the alert-RT, the data is labelled as "high SA". Transitional states with moderate performance are excluded and the neutral state is not considered in this work. To ensure sufficient samples of data for training the model, we filter the datasets such that the dataset of each subject should have at least 50 samples of both states. For the subjects that have multiple datasets, we select the most balanced one to perform the filter operation. Finally, we obtain a whole balanced SA dataset which includes 11 subjects' 1674 samples data. The data size of one sample is 30 (channels) $\times$ 384 (sample points). The number of samples for each subject is shown in Table 1.

3.1.2. Implementation details

We follow [40] to set the parameters of EEGNET. Unlike the most previous works [19,44] which inputted the extracted features to the network with the purpose of reducing the impact of noise, we use only the raw data as the input. The main difference is that extracted features limit EEG information, for instance, PSD features limit the EEG information to only the frequency domain but discard certain temporal information. The input data of the model is in the form 30 (channels) $\times$ 384 (sample points). We randomly choose one subject as the validation set, one subject as the test set and all remaining subjects as the training set. The batch size is set to 50 EEG samples for each source domain. Adam [45] is used to optimize the network parameters with $\beta_1 = 0.9$, $\beta_2 = 0.999$. The learning rate is set to 0.001. For BN blocks in multi-source domain alignment layers, based on the relationship between the attenuation factor and the number of update steps introduced in Section 2.1, we set $momentum = 0.9$.

Table 3
Confusion matrix for 2-way classification using TCA.

		True Class	
		High SA	Low SA
Predicted Class	High SA	742	256
	Low SA	269	755

Table 4
Confusion matrix for 2-way classification using MIDA.

		True Class	
		High SA	Low SA
Predicted Class	High SA	764	269
	Low SA	247	742

Table 5
Confusion matrix for 2-way classification using our subject matching method.

		True Class	
		High SA	Low SA
Predicted Class	High SA	838	348
	Low SA	173	663

3.2. Cross-subject SA recognition results

In the experiment, we evaluate the classification accuracy and other performance metrics using leave-one-subject-out cross-validation. First, we compare our method with a traditional model, i.e., the support vector machine (SVM) [46], as well as state-of-the-art algorithms, including EEGNET, Domain Adversarial Neural Network (DANN) [21], TCA [24] and MIDA [47]. For SVM, TCA and MIDA, we extract the PSD features as inputs. We also include AdaBN [34] in the comparisons. To the best of our knowledge, this is the first time to apply AdaBN on EEG-based SA recognition. Secondly, we evaluate our approach based on the confusion matrix. The overall confusion matrix is obtained by adding the confusion matrix of each single validation. Precision, Sensitivity, Specificity and F1 score are further computed.

Table 2
Leave-one-subject-out cross-validation accuracy. Methods with * use PSD features as inputs, whereas other methods use raw data as inputs (%).

Subject Index	0	1	2	3	4	5	6	7	8	9	10	Avg.	Std.
EEGNET [40]	57.08	59.16	59.07	56.24	57.57	55.31	58.01	54.16	59.12	73.72	59.51	59.00	5.18
DANN [21]	67.41	68.79	68.03	69.62	62.00	68.50	66.38	68.36	69.71	74.13	**72.29**	68.66	2.97
SVM* [46]	79.26	**71.97**	66.67	65.54	82.59	74.10	60.78	65.15	**88.22**	71.30	62.83	71.67	8.31
TCA* [24]	**90.43**	53.03	70.67	72.97	79.91	78.31	63.73	68.94	81.85	83.33	61.50	73.15	10.97
MIDA* [25]	80.32	59.09	80.67	77.03	82.14	76.51	51.96	71.97	87.26	**85.19**	55.31	73.40	12.35
AdaBN [34]	73.68	55.56	**83.33**	**90.00**	82.29	79.41	**66.67**	**75.47**	85.71	54.54	67.39	74.00	11.81
Subject matching	78.72	68.18	79.33	68.24	**85.27**	**83.73**	64.71	57.20	78.03	82.41	71.68	**74.32**	8.94

Table 6
Comparison of performance measures with state-of-the-art methods.

	Precision (%)	Sensitivity (%)	Specificity (%)	F1 Score
TCA [24]	74.35	73.39	74.68	0.7387
MIDA [25]	73.96	75.57	73.39	0.7476
Subject matching	70.66	82.89	65.58	**0.7628**

The comparison results are presented in Table 2. We apply one-way ANOVA [48] to analyse the significance of the differences in the results. Significant difference over the accuracy of models is observed: $F(6,70) = 3.81,\ p < 0.005$. Overall, our method achieves the best average accuracy among these methods. From Table 2, we can observe that the replacement of BN layers by multi-source domain alignment layers significantly improves the performance of the backbone EEGNET by approximately 15%. Compared with AdaBN, a domain adaptation technique which uses target domain statistics to update the BN statistics, our method has a comparable accuracy and no significant difference is observed. The comparison results demonstrate that the subject-specific normalization and subject matching can indeed be beneficial for reducing subject variability. The mixture of all training subjects as source domain could mislead the model to learn biased statistics of each activation and cannot precisely represent any subject.

We compare other performance metrics with state-of-the-art methods: MIDA and TCA. Tables 3–5 show the confusion matrices of TCA, MIDA and subject matching method, respectively. The performance measures based on the given tables are shown in Table 6. Better F1 score (0.7628) demonstrates that the proposed subject matching framework achieves the best overall performance among these three methods. Superior sensitivity result reveals that our approach can

Table 7
Comparison of different similarity measure functions.

Subject Index	0	1	2	3	4	5	6	7	8	9	10	Avg.
Arithmetic Mean	71.81	67.42	72.00	72.97	85.26	83.13	**64.71**	66.67	73.89	77.78	65.93	72.87
Reciprocal	72.87	**75.00**	62.67	**77.70**	71.43	**86.14**	62.75	**70.08**	**86.31**	77.78	69.47	73.83
Reciprocal Exponential	**78.72**	68.18	**79.33**	68.24	**85.27**	83.73	**64.71**	57.20	78.03	**82.41**	**71.68**	**74.32**

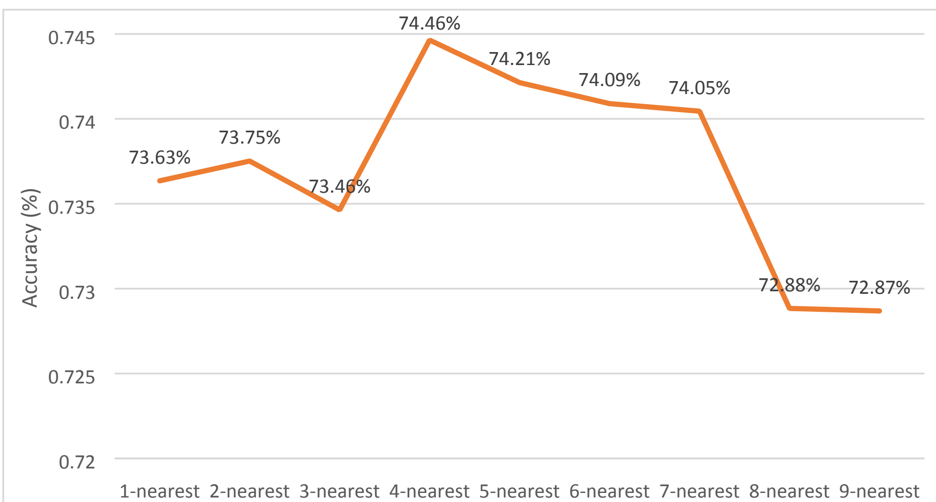

Fig. 2. Arithmetic average results of different number of nearest subjects.

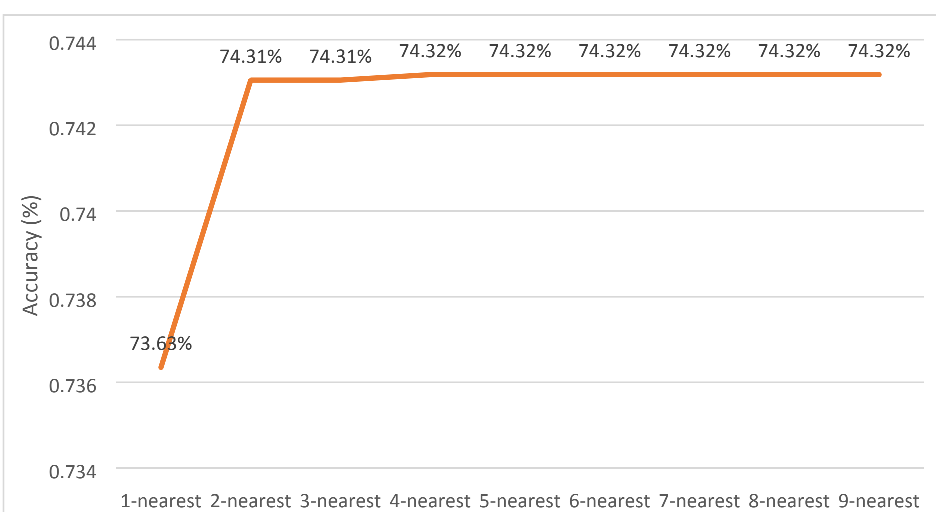

Fig. 3. Subject matching results of different number of nearest subjects.

effectively capture samples with good performance and is useful for SA training. In practice, we prefer to train operators to maintain a high SA state.

3.3. Ablation study

We compare our similarity measure function with other functions. The results are presented in Table 7. For arithmetic mean function, similarity $r_{s,t} = \frac{1}{M}$, where M is the number of source domains. For reciprocal function, similarity $r_{s,t} = \frac{\frac{1}{d_s}}{\sum_{s\in\mathscr{S}}\frac{1}{d_s}}$. Then we can obtain the results of using these functions by applying Eq. (18). The reciprocal exponential function is proved to be a more reliable choice.

According to the ablation experiment results shown in Table 7, the arithmetic mean of multi-source EEGNET achieves an accuracy of 72.87% which is better than the accuracy of EEGNET (59%), which demonstrates the effectiveness of the subject-specific batch normalization. On the other hand, the accuracy of using the arithmetic mean as the similarity measure function is 72.87%, while the accuracy of our subject matching method can achieve 74.31%. Although simple reciprocal can also reduce the impact of the subjects with negative contribution, there are still negative contributes from them. The comparisons demonstrate the validity of the proposed subject matching method.

3.4. Effect of different number of the nearest subjects

We perform a comparison experiment on our framework to analyse the effect of different number of nearest subjects. Fig. 2 shows the arithmetic average results of different number of nearest subjects. The overall results show an increasing trend until using 4-nearest outputs, then decreases to 72.87% finally. Although the distances between source subjects' data and target subject data are obtained, it is still a question that the proper number of the nearest source subjects that can achieve a good performance. For ours, subject matching results of different number of nearest subjects are shown in Fig. 3. We can observe that the results show an upward trend. This comparison demonstrates that the use of reciprocal exponential function benefits to reserve the useful branches output. In addition, our computed similarity can directly show the useful number of the nearest source subjects (the number of branches with non-zero similarities).

4. Conclusion

We have presented a study on the effectiveness of our subject matching framework on the driving dataset using leave-one-subject-out cross validation. There are two main points that benefits to the performance: 1) Subject-specific consideration. In previous EEG processing works in related fields, researchers usually considered different training subjects as a whole source domain, while ignoring the subject-specific characteristics. In our work, we tackle this problem by employing a multi-source domain alignment layer to exploit the information which are usually overlooked. Based on the experiment results, features normalized by statistics from similar source subjects can be classified more accurately. 2) We find that batch normalization is a reliable way in EEG processing to alleviate subject variability effect based on the observations in our experiments. However, limiting the number of source domain subjects could restrict the generalization ability of the model trained only on the source domain to some extent.

The proposed subject matching is a convenient and efficient technique to evaluate SA and obtain similarity between target subjects and source subjects. Similar subjects' data is useful and beneficial in EEG-based recognition tasks [30–32]. The similarity information can be further exploited to find subject-invariant representatives and improve the recognition performance. In practice, our approach can be applied for operator's training in dynamic systems. The operator's SA patterns can be analysed efficiently by simply applying the pre-trained model on the trained operators.

In this paper, we present a study of subject matching for cross-subject recognition of driver state related to situation awareness. Domain generalization technique is applied in our work. Firstly, the BN layers in the model are replaced with multi-source domain alignment layers which help to generalize the model trained on the source domain to the target domain by performing the domain-specific normalization. Secondly, during testing, we compute the similarities between the target subject and each source subject by exploiting BN statistics. The source subjects similar to the target subject are found by applying a reciprocal exponential function. Our proposed framework outperforms the state-of-the-art techniques on a popular driving dataset. The framework is designed to adaptively select better matched subjects in the source domains for the target subject based on statistics, which can be applied in any EEG-based cross-subject model with BN layers. In the future, more generalized algorithms based on BN statistics can be investigated, which would be beneficial to the improvement of cross-subject EEG-based mental states recognition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is supported by the National Research Foundation, Singapore under its International Research Centres in Singapore Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

References

[1] M.R. Endsley, Measurement of situation awareness in dynamic systems, Hum. Fact. 37 (1) (1995) 65–84.
[2] H. Wei, D. Zhuang, X. Wanyan, Q. Wang, An experimental analysis of situation awareness for cockpit display interface evaluation based on flight simulation, Chin. J. Aeronaut. 26 (4) (2013) 884–889.
[3] C.M. Muehlethaler, C.P. Knecht, Situation awareness training for general aviation pilots using eye tracking, IFAC-PapersOnLine 49 (19) (2016) 66–71.
[4] T. Nguyen, C.P. Lim, N.D. Nguyen, L. Gordon-Brown, S. Nahavandi, A review of situation awareness assessment approaches in aviation environments, IEEE Syst. J. 13 (3) (2019) 3590–3603.
[5] J. Luo, Z. Feng, J. Zhang, N. Lu, Dynamic frequency feature selection based approach for classification of motor imageries, Comput. Biol. Med. 75 (2016) 45–53.
[6] R.L. Li, Z.R. Lan, J. Cui, O. Sourina, L. Wang, in: EEG-based Recognition of Driver State Related to Situation Awareness Using Graph Convolutional Networks, IEEE, 2020, pp. 180–187.
[7] D.J. Hemanth, J. Anitha, L.H. Son, Brain signal based human emotion analysis by circular back propagation and Deep Kohonen Neural Networks, Comput. Electr. Eng. 68 (2018) 170–180.
[8] J.A.M. Saucedo, J.D. Hemanth, U. Kose, Prediction of electroencephalogram time series with electro-search optimization algorithm trained adaptive neuro-fuzzy inference system, IEEE Access 7 (2019) 15832–15844.
[9] M.K. Chowdary, D.J. Hemanth, Emotion recognition using feature extraction techniques, Inform. Technol. Intell. Transport. Syst. 323 (2020) 71.
[10] M. K. Chowdary, D. J. Hemanth, A. Angelopoulou, and E. Kapetanios, "Feature extraction techniques for human emotion identification from face images," in 9th International Conference on Imaging for Crime Detection and Prevention (ICDP-2019), London, UK, 2019, pp. 86–92.
[11] M. Flynn, D. Effraimidis, A. Angelopoulou, E. Kapetanios, D. Williams, J. Hemanth, and T. Towell, "Assessing the Effectiveness of Automated Emotion Recognition in Adults and Children for Clinical Investigation," Frontiers in human neuroscience, vol. 14, 2020.
[12] M.K. Chowdary, D.J. Hemanth, Human emotion recognition using intelligent approaches: a review, Intell. Decis. Technol. 13 (4) (2019) 417–433.
[13] W.L. Lim, O. Sourina, L. Wang, in: Cross Dataset Workload Classification Using Encoded Wavelet Decomposition Features, IEEE, 2018, pp. 300–303.
[14] S.U. Amin, M. Alsulaiman, G. Muhammad, M.A. Mekhtiche, M.S. Hossain, Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion, Future Gener. Comput. Syst. 101 (2019) 542–554.

[15] J. Li, S. Qiu, C. Du, Y. Wang, and H. He, "Domain Adaptation for EEG Emotion Recognition Based on Latent Representation Similarity," IEEE Transactions on Cognitive and Developmental Systems, 2019.
[16] Y. Luo, S.-Y. Zhang, W.-L. Zheng, and B.-L. Lu, "WGAN domain adaptation for EEG-based emotion recognition," in International Conference on Neural Information Processing, 2018, pp. 275-286: Springer.
[17] S. Ben-David, J. Blitzer, K. Crammer, and F. Pereira, "Analysis of representations for domain adaptation," in Advances in neural information processing systems, 2007, pp. 137-144.
[18] J. Gu, R. Kanai, What contributes to individual differences in brain structure? Front. Hum. Neurosci. 8 (2014) 262.
[19] B.-Q. Ma, H. Li, Y. Luo, and B.-L. Lu, "Depersonalized Cross-Subject Vigilance Estimation with Adversarial Domain Generalization," in 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1-8: IEEE.
[20] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in International conference on machine learning, 2015, pp. 1180-1189: PMLR.
[21] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, J. Mach. Learn. Res. 17 (1) (2016) 2096–12030.
[22] E. Tzeng, J. Hoffman, K. Saenko, T. Darrell, Adversarial discriminative domain adaptation, in: in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7167–7176.
[23] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in Advances in neural information processing systems, 2018, pp. 1640-1650.
[24] Z. Lan, O. Sourina, R. Lipo Wang, Scherer, G.R. Müller-Putz, Domain adaptation techniques for EEG-based emotion recognition: a comparative study on two public datasets, IEEE Trans. Cognit. Devel. Syst. 11 (1) (2018) 85–94.
[25] Y. Liu, Z. Lan, J. Cui, O. Sourina, W. Müller-Wittig, Inter-subject transfer learning for EEG-based mental fatigue recognition, Adv. Eng. Inf. 46 (2020), 101157.
[26] F. Qiao, L. Zhao, X. Peng, Learning to learn single domain generalization, in: in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 12556–12565.
[27] D. Li, Y. Yang, Y.-Z. Song, T. Hospedales, in: in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, p. no. 1..
[28] F.M. Carlucci, A. D'Innocente, S. Bucci, B. Caputo, T. Tommasi, Domain generalization by solving jigsaw puzzles, in: in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2229–2238.
[29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in: International conference on machine learning, 2015, pp. 448-456: PMLR.
[30] W. Wei, S. Qiu, X. Ma, D. Li, B. Wang, H. He, Reducing Calibration Efforts in RSVP Tasks With Multi-Source Adversarial Domain Adaptation, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (11) (2020) 2344–2355.
[31] F. Wang, W. Zhang, Z. Xu, J. Ping, and H. Chu, "A deep multi-source adaptation transfer network for cross-subject electroencephalogram emotion recognition," Neural Computing and Applications, pp. 1-13.
[32] E. Jeon, W. Ko, H.-I. Suk, in: Domain Adaptation with Source Selection for Motor-Imagery based BCI, IEEE, 2019, pp. 1–4.
[33] F.M. Carlucci, L. Porzi, B. Caputo, E. Ricci, S.R. Bulo, in: Autodial: Automatic domain alignment layers, IEEE, 2017, pp. 5077–5085.
[34] Y. Li, N. Wang, J. Shi, X. Hou, J. Liu, Adaptive batch normalization for practical domain adaptation, Pattern Recogn. 80 (2018) 109–117.
[35] W.-G. Chang, T. You, S. Seo, S. Kwak, B. Han, Domain-specific batch normalization for unsupervised domain adaptation, in: in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7354–7362.
[36] S. Seo, Y. Suh, D. Kim, J. Han, B. Han, Learning to optimize domain specific normalization for domain generalization, in: A. Vedaldi, H. Bischof, T. Brox, JM. Frahm (Eds.), Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science, 12367, Springer, Cham, 2020.
[37] M. Segù, A. Tonioni, and F. Tombari, "Batch Normalization Embeddings for Deep Domain Generalization," arXiv preprint arXiv:2011.12672, 2020.
[38] M. Nounou, B.R. Bakshi, B. Walczak, Multiscale methods for denoising and compression, Elsevier Science BV, 2000.
[39] M. Mancini, L. Porzi, S. Rota Bulò, B. Caputo, E. Ricci, Boosting domain adaptation by discovering latent domains, in: in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3771–3780.
[40] V.J. Lawhern, A.J. Solon, N.R. Waytowich, S.M. Gordon, C.P. Hung, B.J. Lance, EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces, J. Neural Eng. 15 (5) (2018), 056013.
[41] Z. Cao, C.-H. Chuang, J.-K. King, C.-T. Lin, Multi-channel EEG recordings during a sustained-attention driving task, Sci. Data 6 (1) (2019) 1–8.
[42] M.R. Endsley, Automation and situation awareness, Automat. Hum. Perform. Theory Appl. 20 (1996) 163–181.
[43] L. G. Yeo, H. Sun, Y. Liu, F. Trapsilawati, O. Sourina, C.-H. Chen, W. Mueller-Wittig, and W. T. Ang, "Mobile EEG-based situation awareness recognition for air traffic controllers," in 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, pp. 3030-3035: IEEE.
[44] R. F. Rojas, E. Debie, J. Fidock, M. Barlow, K. Kasmarik, S. Anavatti, M. Garratt, and H. Abbass, "Encephalographic Assessment of Situation Awareness in Teleoperation of Human-Swarm Teaming," in International Conference on Neural Information Processing, 2019, pp. 530-539: Springer.
[45] D.P. Kingma, J. Ba, "Adam: A method for stochastic optimization," in: 3rd International Conference on Learning Representations, 2015.
[46] J.A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett. 9 (3) (1999) 293–300.
[47] Y. Liu, Z. Lan, J. Cui, O. Sourina, W. Müller-Wittig, in: EEG-Based Cross-Subject Mental Fatigue Recognition, IEEE, 2019, pp. 247–252.
[48] R.G. Lomax, Statistical concepts: a second course, Law. Erlbaum Assoc. Publ. (2007).