



# A Review of Transfer Learning for EEG-Based Driving Fatigue Detection

Jin Cui<sup>1,2</sup>, Yong Peng<sup>1</sup>, Kenji Ozawa<sup>2</sup>, and Wanzeng Kong<sup>1</sup>(✉)

<sup>1</sup> School of Computer Science and Technology,

Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China

{cuijin50057, yongpeng, kongwanzeng}@hdu.edu.cn

<sup>2</sup> Department of Computer Science and Engineering, University of Yamanashi, Kofu, Japan

ozawa@yamanashi.ac.jp

**Abstract.** Driver mental state detection has been playing an increasingly significant role in safe driving for decades. Electroencephalogram (EEG)-based detection methods have already been applied to improve detection performance. However, numerous problems still have not been addressed in practical applications. Specifically, most of the existing traditional methods require a large number of training data, caused by differences in cross-subject samples and cross-time of the same subject, resulting in enormous calculations and time consumption. To overcome the above limitations, transfer learning, which applies data or knowledge from the source domain to the target domain, has been widely adopted in EEG processing. This article reviews the current state of mainstream transfer learning methods and their application based on driver mental state detection. To the best of our knowledge, this is the first comprehensive review of transfer learning methods for driving fatigue detection.

**Keywords:** Transfer learning · Electroencephalogram · Driver mental state detection

## 1 Introduction

Fatigue driving refers to the physiological and psychological disorders caused by the driver's long continuous driving [1–4]. Due to lack of sleep, long continuous driving, and other factors, the driver is easily fatigued, in turn resulting in inattentive driving, decreased judgment, improper driving operation, and increasing the potential for traffic accidents [5, 6]. Effective driver mental state detection reduces the probability of unsafe driving and property loss [7].

Currently, there are three main driver mental state detection methods: computer vision-based [8–10], human physiological signal-based [11, 12], and information integration technology-based [13, 14] methods. Previous methods typically collect images of the driving process, establishing appropriate criteria for judgement and using image processing techniques to analyze the driver's facial expressions to determine if the driver is fatigued. Human physiological signal-based methods usually collect data such as

the subject's electroencephalogram (EEG), eye motion, electrocardiogram (ECG), heart rate, and blood pressure to analyze the driver's physical state and predict its mental state. Information integration technology-based methods mainly establish an integration model of a variety of factors that may contribute to the driver's mental state, so as to perform the driver mental state detection and analysis. Among them, EEG [15], as an objective signal, is capable of quickly reflecting the process of human physiological and mental changes and is widely considered as the most readily available and effective driver mental state detection method.

However, the following problems are inevitably encountered when using EEG signals for analysis.

1. The EEG signal is particularly weak and susceptible to noise.
2. EEG is spontaneous and highly individualized, with varying data distributions across subjects and time periods.
3. EEG data samples are precious, and the collection of large amounts of data entails high time and financial costs.

Accordingly, it is extremely important to obtain desirable detection results from a small number of cross-subject sample features. Transfer learning [16] can address the problem of sparse data labels by transferring knowledge from the learned source domain to an unlabeled target domain. It differs from traditional machine learning methods in two aspects: (1) Transfer learning approaches forbid the premise that data from different domains obey the same distribution and are applicable to cases where data distributions are inconsistent; (2) Transfer learning approaches attempt to solve unsupervised problems based on only a few samples.

A growing number of studies have used transfer learning for driver state detection as transfer learning continues to progress in EEG data processing. Among current review articles [17–19] addressing driver mental state and fatigue, to the best of our knowledge, this is the first review article that focuses on the transfer learning approach.

The remainder of this paper is organized as follows. Section 2 provides a simple introduction to transfer learning. Section 3 describes the transfer learning-based driver fatigue detection methods. Finally, our conclusions are outlined in Sect. 4.

## 2 Transfer Learning

Transfer learning is the ability to systematically identify and apply knowledge and skills learned in a previous domain to a new domain.

There are two very important terms in transfer learning: domain and task. In EEG-based driver fatigue state detection, a domain usually represents the EEG observations obtained when a subject performs the same learning task. The EEG observations of different subjects under the same task are defined as their own domains. Furthermore, the domain can be divided into the source domain ( $D_S$ ) and the target domain ( $D_T$ ), depending on whether the domain has knowledge or not. Specifically, the source domain with label information is defined as  $\{X_S, Y_S\}$ , and the target domain without label information is recorded as  $\{X_T\}$ . Additionally, a learning task consists of labels and the corresponding

function, in which labeled spaces are represented by  $Y$  and the function is represented by  $f(\cdot)$ . For instance, sentiment analysis and driver state analysis are two different tasks.

## 2.1 Definition of Transfer Learning

**Definition 1.** As described in [16], there are two important parts in domain  $D$ , namely, feature space  $X$  and marginal probability distribution  $P(X)$ ,  $X = \{x_1, x_2, \dots, x_n\}$ . Obviously, various domains have various feature spaces or obey various marginal probability distributions. Furthermore, given a source domain  $D_S$ , a learning task  $T_S$ , a target domain  $D_T$ , and a learning task  $T_T$ , the goal of transfer learning is to improve the learning performance of the target prediction function  $f_T(\cdot)$  in  $D_T$  by effectively employing the knowledge learned by  $D_S$  and  $T_S$ , in which the following formula must be obeyed:  $D_S \neq D_T$  or  $T_S \neq T_T$ .

The condition  $D_S \neq D_T$  implies that the source and target domain instances are various, i.e.,  $X_S \neq X_T$ , or that the source and target domain marginal probability distributions are various,  $P_S(X) \neq P_T(X)$ . Specifically, for each task,  $T = \{Y, f(\cdot)\}$ , where  $f(\cdot)$  represents the conditional probability distribution  $P(Y|X)$ . Furthermore,  $T_S \neq T_T$  means that the source labels are unequal to target domain labels, i.e.,  $Y_S \neq Y_T$ , or that the source and target conditional probability distributions are unequal,  $P(Y_S|X_S) \neq P(Y_T|X_T)$ . It is worth noting that the current problem becomes a traditional machine learning problem if the source and target domains are equal,  $D_S = D_T$ , and the source and target tasks are also equal,  $T_S = T_T$ .

**Domain Adaptation.** As described in [20], Given a marked source domain  $D_S = \{X_i, Y_i\}_{i=1}^n$  and an unmarked target domain  $D_T = \{X_j\}_{j=1}^m$ , it is assumed that their eigenspaces are the same, i.e.,  $X_S = X_T$ , and their class spaces are also the same, i.e.,  $Y_S = Y_T$ . However, the marginal and conditional distributions of both domains are different, i.e.,  $P_S(X) \neq P_T(X)$  and  $P(Y_S|X_S) \neq P(Y_T|X_T)$ . Then, the goal of transfer learning is to use marked data  $D_S$  to train a classifier  $f : x_S \rightarrow y_S$  to classify the label  $y_T \in Y_T$  of the target domain  $D_T$ .

## 2.2 A Brief Introduction of Transfer Learning Methods

**Instance-Based Transfer Learning.** To efficiently use the similarity with the target domain, some source domain data samples are reused according to weight generation rules to carry out the transfer learning [21–24]. The instance-based weight method has rich theoretical achievements and is easily deduced and used. However, this type of method is usually effective only when the distribution difference between fields is small.

**Feature-Based Transfer Learning.** This type of method mine the correlation between the source and target domains through feature transformation, so as to decline the variation between the two domains [25]; or merge the data features of the source and target domains into a unified feature space, so that the improved traditional method is able to complete the related task [26–31]. Currently, this is the most important and common transfer learning method, which is extensively employed for the cross-subject transfer of EEG in fatigue driving.

**Model (Parameter)-Based Transfer Learning.** Let the data in the source and target domains share certain model parameters, then the goal of transfer learning is to discover the shared relevant parameters from the source and target domains. Presently, methods involved in deep neural networks [32–35] are also among the model-based transfer learning methods. In addition, models combining domain adaptation and deep neural networks are commonly used methods, which are not only model-based but also feature-based methods.

**Relationship-Based Transfer Learning** [36]. Most of this type of method emphasize on identifying the relationship between the source and target domain samples, but with only a few related studies. Furthermore, most of them are based on Markov logic net to mine the commonality between various domains.

Over the past decades, to tackle the non-linear, unstable, and high-dimensional characteristics of EEG data, an abundance of effective EEG feature extraction methods further upgraded the extraction effectiveness. In the current EEG-based application, instance-based, feature-based, and model (parameter)-based methods have mostly drawn the attention of researchers. Therefore, Sect. 3 focuses on introducing the development of these three methods.

### 3 Transfer Learning-Based Driver Fatigue Detection Methods

Transfer learning has been proposed to address the small sample problem and the adaptation of different domains. Transfer learning is capable of providing more effective solutions to the EEG transfer classification problem across subjects. Accordingly, it has been applied in driver fatigue detection.

In the application of transfer learning in driver fatigue detection, most existing detection systems (framework) have the same construction processes, in which several general characteristics can be summarized. Therefore, the first part of this chapter summarizes the characteristics of driver fatigue detection systems based on transfer learning methods. Presently, the transfer learning pattern recognition methods for driver fatigue detection are mainly feature- and model-based, and to a lesser proportion instance-based. These methods are summarized next.

#### 3.1 System Features

Driver fatigue detection systems collect EEG signals during driving, process EEG signals either online or offline, feedback and control the results.

A complete driver fatigue detection system [18, 37] usually includes signal acquisition, signal preprocessing, feature extraction, pattern recognition, and feedback.

**Signal acquisition:** The weak EEG signal is detected by an electrode placed on the subject's scalp, then amplified and digitized, and finally recorded by the matching recording system. It is mainly collected by the EEG cap and other EEG devices, including 64- and 32-electrode EEG caps.

**Signal preprocessing:** Used to remove common noise, interference, and artifacts, in order to improve signal quality. Commonly used preprocessing methods are digital filtering and independent component analysis (ICA).

**Feature extraction:** Used to reduce the dimensionality of EEG data and extract relevant features for pattern recognition. Common feature extraction methods are in temporal, frequency, and spatial domains as well as a combined analysis of the two in three domains. Frequency domain analysis: autoregressive (AR) model and power spectrum estimation (PSD); temporal-frequency analysis: wavelet transform and wavelet packet transform; spatial domain analysis: principal component analysis (PCA) and common spatial pattern (CSP).

**Pattern recognition model (PRM):** The existing transfer learning pattern recognition driver fatigue detection model is mainly based on classification (C) and regression (R) models. The classification model usually sets the category information for the driver's mental state from the EEG data according to a certain threshold. Then, the transfer learning model outputs the fatigue category, which is a discrete value. The regression model predicts the specific sleepiness state through the model, and the output is a continuous value.

**Explore transferability in transfer learning:** Many existing methods [38–40] consider how to select the optimal auxiliary source domains in order to further reduce the transfer learning cost and error. The appropriate auxiliary data is often more effective for transfer learning with less effort.

In the literature, feature extraction methods and pattern recognition models have some characteristics that can be summarized. In the following sections we summarize them according to different pattern recognition methods.

### 3.2 Instance-Based Transfer Learning Methods

The processing of instance-based transfer learning method is simple. In order to fully utilize the existing source domain data in EEG-based fatigue detection, it is necessary to perform similarity matching with the target domain according to weight generation rules to complete the data alignment and transfer learning. Table 1 summarizes this approach.

Wu *et al.* [38] proposed an online weighted adaptation regression regularization (OwARR) algorithm used to decrease the amount of data required for a given subject calibration. A source domain selection (SDS) method was also proposed to reduce the computational cost of OwARR by 50%. The online classification/regression task means that there is not enough labeled data for calibration. In the literature, each subject performing the same driving task is considered a distinct source domain. Initially, OwARR is applied to each source domain, then the final regression model is constructed as a weighted mean of these basic models. Together, the final regression models are applied to future unlabeled data. Specifically, SDS is employed to reduce the clustering error of multiple source domains before domain adaptation. By selecting the best first  $Z$  source domains, SDS maintains model performance with less computation cost. On average, the training time for OwARR-SDS is approximately half of that for OwARR.

**Table 1.** Instance-based transfer learning methods.

Ref.	Year	Task	Dataset	Signal	Feature extraction	Pattern recognition method	Performance	PRM
Wu <i>et al.</i> [38]	2017	Simulated driving	[41]	EEG	Power spectral density (PSD)	OwARR	Calibration time reduced by ~50%	R
Wei <i>et al.</i> [39]	2015	Lane-keeping driving	–	EEG	Welch's fast Fourier transform	Selective transfer learning	–	R
Wei <i>et al.</i> [40]	2018	Lane-keeping driving	–	EEG	Multi-channel EEG powers	Hierarchical cluster analysis and Subject-transfer framework	Calibration time reduced by 90%	R

As EEG correlations between individuals are stable, there are existing auxiliary subject data to improve EEG performance. Wei *et al.* [39] proposed a framework for selective transfer learning that effectively utilizes the large amount of training data from other subjects to improve the recognition efficiency of unlabeled target domain data. This theoretical finding is a good reference for cross-subject transfer.

In an effort to improve system performance with minimal personalized calibration data, Wei *et al.* [40] used hierarchical clustering methods to evaluate inter- and intra-subject variability in a wide-scale EEG dataset of a simulated driving task. In addition, based on the existing data collected from the source subject, a model source pool was constructed. Furthermore, the framework carries out the design of an adjustment mechanism for ordering and fusing the source models of each target subject. In terms of time cost, the calibration time of the self-decoding (SD) method was 89.91 min, and that of the subject-transfer (ST) method was 1.48 min, the calibration time required for new users was reduced by 90%.

### 3.3 Feature-Based Transfer Learning Methods

The largest proportion of EEG-based driver fatigue detection methods are feature-based transfer learning methods due to its better feature alignment effect. This method tends to find source and target domain data based on two common mapping spaces, either optimization based on probability distribution alignment, or a combination of both. It is not difficult to train and the training effect is significant. Furthermore, compared to deep learning methods, training time and training data costs are low, therefore, it is extensively used in EEG-based tasks. Table 2 summarizes this approach.

An online multi-view and transfer Takagi–Sugeno–Kang (TSK) fuzzy system [42] is proposed to estimate the driver's sleepiness, which represents the source and target domain characteristics from multiple perspectives. In this algorithmic framework, the domain EEG data are characterized in terms of multiple perspectives. The multi-angle setting is injected into the transfer learning framework to enhance the consistency of the

**Table 2.** Feature-based transfer learning methods.

Ref.	Year	Task	Dataset	Signal	Feature extraction	Pattern recognition method	Performance	PRM
Jiang <i>et al.</i> [42]	2020	Lane-keeping driving	–	EEG	Time, frequency and time-frequency features	O-MV-T-TSK-FS	–	R
Chen <i>et al.</i> [43]	2019	Real driving	Dataset A [49]	EEGEOG	Time, frequency wavelet domains features.	Domain transfer learning (ARTL, TCA, JDA, TJM)	94.15% on dataset A	C
			Dataset B [50]				88.43% on dataset B	
Zhang <i>et al.</i> [44]	2017	Simulated driving	–	EEG	Fuzzy mutual information-based wavelet- packet transform	KSR and TDDR (3 classes)	KSR+TDDR+SVM/KSR+TDDR+KNN achieved the same highest G-mean of 79.00%	C
Liu <i>et al.</i> [45]	2020	Simulated driving	–	EEG	Spectral band power	MIDA	73.01% of thirty channels	C
Zhang <i>et al.</i> [48]	2015	Simulated driving	–	EEG	DE	TCA	TCA-SVM (77.02%)	C

different angles. This online fuzzy system is more flexible and controllable than offline training.

Chen *et al.* [43] proposed an automatic detection system based on cross-subject feature selection and transfer classifiers to identify different driving mentalities. Considering the negative effects of noise and irrelevant information on transfer learning, they designed the class separation and domain fusion (CSDF) and utilized a hybrid feature choice methodology to combine different types of filtering methods in one framework. Additionally, they adopted a common adaptation regularization-based transfer learning (ARTL) as the pattern recognition method, which simultaneously optimizes the structural risk, the joint distribution, and the manifold consistency of two domains. This optimization method is based on the structural risk minimization principle and regularization theory.

The kernel spectral regression (KSR) with transformable discriminant dimension reduction (TDDR) method was proposed by Zhang *et al.* [44]. This method uses the reduced feature vector dimensionality to achieve the transfer of the classifier model cross-subjects. However, considering only low-dimensional source space discrimination is undesirable, as this would poorly generalize to the target domain of traditional dimensions. In this work, knowledge transfer using TDDR rewards the separation of domain merge data and penalizes the distance between the source and target domains by defining an objective function that rewards domain merge data. A low-dimensional latent space can be found, ensuring both discriminability and transferability, which addresses the problem of traditional dimension reduction methods only considering low dimension recognition. Furthermore, KSR is capable of overcoming the linear discriminant analysis (LDA) limitation to detect nonlinear components when reducing the EEG feature dimension. In the literature, detection results on two datasets show that the framework improves the performance of multi-class and multi-bandwidth identification.

Liu *et al.* [45] proposed a transfer learning-based cross-subject EEG fatigue recognition algorithm without correction. They also explored the influence of the number of EEG signal channels on algorithm accuracy and compared single and multi-channel situations. Specifically, the random forest algorithm was used to select the channel with the highest characteristic resolution. Their experimental results demonstrated that the occipital lobe channel has a better effect when considering only one channel. In this paper, two classical transfer learning strategies, namely, transfer component analysis (TCA) and maximum independence domain adaptation (MIDA) [46], are used. Among them, TCA is employed to alleviate the classification accuracy decline problem resulting from the distribution mismatch between the source and target data. The goal of TCA is to seek a potential mapping subspace where the maximum mean difference (MMD) between the source and target data is reduced in the Reproducing Kernel Hilbert Space (RKHS) [47]. The distance between these measures is empirically averaged. MIDA enables data from different domains into a potential domain invariant space, where the projected samples are independent of domain features. The accuracy was determined to be 73.01% for all thirty channels using MIDA and 68.00% for one selected channel using TCA, which was better than the baseline and deep learning methods.



### 3.4 Model-Based Transfer Learning Methods

A parametric model-based transfer learning method in driver fatigue detection usually addresses how to find a common parameter or prior distribution between the spatial model of the source and target domains in order to transfer knowledge through further processing. Deep transfer learning methods also belong to this category. Table 3 summarizes this approach.

Wu *et al.* [51] proposed a combined method based on transfer learning, active class selection (ACS), and a mean squared difference user-similarity heuristic and selects the best sample. Specifically, collaborative filtering is used to combine training data from a solitary subject with external training data from other similar subjects. In addition, in order to improve learning performance by combining a limited number of training samples with a substantial number of supplementary training samples from other similar topics, ACS optimizes class selection to generate individual user-specific training samples. It can boost recognition accuracy by not increasing the number of training samples.

Wu *et al.* [41] proposed an online EEG-based sleep estimation method based on adaptive model fusion. In this framework, only a few subjects require correction to achieve satisfactory results. Specifically, for each domain in  $Z$  auxiliary source domains, it combines with the target domain to implement the ridge regression-based domain adaptation operation and  $Z$  different models are obtained, which are fused into the final model.

In [52], a deep neural network-based transfer learning driver fatigue detection system is proposed, which increased system availability by relying solely on EEG channels. First, the signal is preprocessed and filtered, then transformed into two-dimensional spectrum. Then, the two-dimensional spectrum is classified by using AlexNet, the final normal and fatigue classification is carried out by using a transfer learning method. The FP1 and T3 channels have been shown experimentally to be the most effective channels for reflecting the driver's fatigue state. Furthermore, with the improved AlexNet convolutional neural network (CNN) model, an efficient driver fatigue detection system can be obtained using only one channel. This method makes the driver fatigue detection system flexible, which is a major advantage.

In [53], two kinds of domain adaptive neural network (DaNN) and adverse discriminative domain adaptation (ADDA), based on the SEED-VIG dataset [50], are used to classify electrooculogram (EOG) and EEG signals. Compared with traditional domain adaptation methods, this method significantly improves the data. The experimental results show that the Pearson correlation coefficients of both domain adaptation networks are improved by more than 10% compared to the baseline. Therefore, the use of adversarial networks for EEG driver fatigue classification is a promising experiment.

Due to the continuous development of deep networks, EEG data could also be processed using model-based transfer learning methods. Moreover, parameter-based methods could be combined with feature-based methods to achieve better experimentation performance. However, there are not many deep network-based methods for fatigue detection and classification, which are necessary to further adopt efficient methods and achieve better experimental results.

Table 3. Model-based transfer learning methods.

Ref.	Year	Task	Dataset	Signal	Feature extraction	Pattern recognition method	Performance	PRM
Wu <i>et al.</i> [41]	2015	Lane-keeping driving	–	EEG	PSD	DAMF	–	R
Wu <i>et al.</i> [51]	2013	Virtual Reality Stroop Task	Virtual reality Stroop task [54]	EEG, EDA, RSP, ECG	Spectral power	Combining transfer learning (TL) and Active Class Selection (ACS)	56.76%–95.82% for one subject	C
Shalash <i>et al.</i> [52]	2019	Simulated driving	–	EEG	–	Modified AlexNet CNN model	T3: 91% FPI: 90%	C
Li <i>et al.</i> [53]	2018	Simulated driving	SEED-VIG [50]	EEG, EOG	ICA	Adversarial Discriminative Domain Adaptation (ADDA)	PCC: 0.8442	R

## 4 Conclusion

With the continuous improvement of EEG acquisition devices, EEG-based driver mental state detection methods have become objective and accurate. Presently, traditional machine learning- and deep learning-based methods effectively achieve remarkable results on inter-subject experiments. However, EEG data distribution is complex and unstable. In practice, samples are precious, and more powerful models are needed to address the problem of monitoring cross-subject and cross-time EEG signals. The cross-subject problem may be addressed more effectively with the on-going transfer learning research. However, there are still certain limitations, which could be overcome in the future with the development of a large number of transfer learning algorithms.

**Acknowledgment.** This work was supported by National Key R&D Program of China for Intergovernmental International Science and Technology Innovation Cooperation Project (2017YFE0116800), National Natural Science Foundation of China (U20B2074, U1909202), Science and Technology Program of Zhejiang Province (2018C04012), and supported by Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province (2020E10010).

## References

1. Coetzer, R.C., Hancke, G.P.: Eye detection for a real-time vehicle driver fatigue monitoring system. In: 2011 IEEE Intelligent Vehicles Symposium (IV), pp. 66–71 (2011)
2. Lal, S.K., Craig, A.: A critical review of the psychophysiology of driver fatigue. *Biol. Psychol.* **55**, 173–194 (2001)
3. Sikander, G., Shahzad, A.: Driver fatigue detection systems: a review. *IEEE Trans. Intell. Transp. Syst.* **20**, 2339–2352 (2019)
4. May, J.F., Baldwin, C.L.: Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transp. Res. Part F: Traffic Psychol. Beh.* **12**(3), 218–224 (2009)
5. Lal, S.K.L., Craig, A.: Driver fatigue: electroencephalography and psychological assessment. *Psychophysiology* **39**(3), 313–321 (2002)
6. Lewis, I., Watson, B., Tay, R., White, K.M.: The role of fear appeals in improving driver safety: a review of the effectiveness of fear-arousing (threat) appeals in road safety advertising. *Int. J. Behav. Consultation and Therapy* **3**(2), 203 (2007)
7. Azam, K., Shakoor, A., Shah, R.A., Khan, A., Shah, S.A., Khalil, M.S.: Comparison of fatigue related road traffic crashes on the national highways and motorways in Pakistan. *J. Eng. Appl. Sci.* **33**(2), 47–54 (2014)
8. Chellappa, R., Charles, L.W., Saad, S.: Human and machine recognition of faces: a survey. *Proc. IEEE* **83**(5), 705–741 (1995)
9. Mandal, B., Li, L., Wang, G.S., Lin, J.: Towards detection of bus driver fatigue based on robust visual analysis of eye state. *IEEE Trans. Intell. Transp. Syst.* **18**(3), 545–557 (2016)
10. Yao, K.P., Lin, W.H., Fang, C.Y., Wang, J.M., Chang, S.L., Chen, S.W.: Real-time vision-based driver drowsiness/fatigue detection system. In: 2010 IEEE 71st Vehicular Technology Conference, pp. 1–5 (2010)
11. Eoh, H.J., Chung, M.K., Kim, S.H.: Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. *Int. J. Ind. Ergon.* **35**(4), 307–320 (2005)
12. Kong, W., Zhou, Z., Jiang, B., Babiloni, F., Borghini, G.: Assessment of driving fatigue based on intra/inter-region phase synchronization. *Neurocomputing* **219**, 474–482 (2017)

13. Jung, S.J., Shin, H.S., Chung, W.Y.: Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel. *IET Intel. Transport Syst.* **8**(1), 43–50 (2014)
14. Brandt, T., Stemmer, R., Rakotonirainy, A.: Affordable visual driver monitoring system for fatigue and monotony. In: 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583), vol. 7, pp. 6451–6456 (2004)
15. Chai, R., et al.: Improving EEG-based driver fatigue classification using sparse-deep belief networks. *Frontiers Neurosci.* **11**, 103 (2017)
16. Pan, S.J., Qiang, Y.: A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359 (2010)
17. Wang, P., Lu, J., Zhang, B., Tang, Z.: A review on transfer learning for brain-computer interface classification. In: 2015 5th International Conference on Information Science and Technology (ICIST) (2015)
18. Monteiro, T.G., Skourup, C., Zhang, H.: Using EEG for mental fatigue assessment: a comprehensive look into the current state of the art. *IEEE Trans. Hum. Mach. Syst.* **49**(6), 599–610 (2019)
19. Lotte, F., et al.: A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *J. Neural Eng.* **15**(3), (2018)
20. Patel, V.M., Gopalan, R., Li, R., Chellappa, R.: Visual domain adaptation: a survey of recent advances. *IEEE Signal Process. Mag.* **32**, 53–69 (2015)
21. Khan, M.N., Heisterkamp, D.R.: Adapting instance weights for unsupervised domain adaptation using quadratic mutual information and subspace learning. In: 2016 23rd International Conference on Pattern Recognition (ICPR), vol. 21, pp. 1560–1565 (2016)
22. Dai, W., Yang, Q., Xue, G.R., Yu, Y.: Boosting for transfer learning. In: Proceedings of the 24th International Conference on Machine Learning, pp. 193–200 (2007)
23. Zhao, Z., Chen, Y., Liu, J., Shen, Z., Liu, M.: Cross-people mobile-phone based activity recognition. In: Twenty-Second International Joint Conference on Artificial Intelligence (2011)
24. Wei, Y., Zhu, Y., Leung, C., Song, Y., Yang, Q.: Instilling social to physical: co-regularized heterogeneous transfer learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 30, no. 1 (2016)
25. Hu, D., Yang, Q.: Transfer learning for activity recognition via sensor mapping. In: International Joint Conference on IJCAI. DBLP (2011)
26. Pan, S.J., Tsang, I.W., Kwok, J.T., Yang, Q.: Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **22**(2), 199–210 (2010). <https://doi.org/10.1109/tnn.2010.2091281>
27. Long, M., Wang, J., Ding, G., Sun, J., Yu, P.S.: Transfer feature learning with joint distribution adaptation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2200–2207 (2013)
28. Satpal, S., Sarawagi, S.: Domain adaptation of conditional probability models via feature subsetting. In: Kok, J.N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenić, D., Skowron, A. (eds.) PKDD 2007. LNCS (LNAI), vol. 4702, pp. 224–235. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-74976-9\\_23](https://doi.org/10.1007/978-3-540-74976-9_23)
29. Gong, M., Zhang, K., Liu, T., Tao, D., Glymour, C., Schölkopf, B.: Domain adaptation with conditional transferable components. In: International Conference on Machine Learning, pp. 2839–2848 (2016)
30. Fernando, B., Habrard, A., Sebban, M., Tuytelaars, T.: Unsupervised visual domain adaptation using subspace alignment. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2960–2967 (2013)
31. Sun, H., Shuai, L., Shilin, Z.: Discriminative subspace alignment for unsupervised visual domain adaptation. *Neural Process. Lett.* **44**(3), 779–793 (2016)

32. Long, M., Wang, J., Cao, Y., Sun, J., Philip, S.Y.: Deep learning of transferable representation for scalable domain adaptation. *IEEE Trans. Knowl. Data Eng.* **28**(8), 2027–2040 (2016)
33. Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C.: A survey on deep transfer learning. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds.) *ICANN 2018*. LNCS, vol. 11141, pp. 270–279. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01424-7\\_27](https://doi.org/10.1007/978-3-030-01424-7_27)
34. Litjens, G., et al.: A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017)
35. Ganin, Y., et al.: Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* **17**(1), 2030–2096 (2016)
36. Mihalkova, L., Mooney, R.J.: Transfer learning by mapping with minimal target data. In: *Proceedings of the AAAI-2008 Workshop on Transfer Learning for Complex Tasks* (2008)
37. Wu, D., Xu, Y., Lu, B.: Transfer learning for EEG-based brain-computer interfaces: a review of progresses since 2016. *arXiv preprint arXiv:2004.06286* (2020)
38. Wu, D., Lawhern, V.J., Gordon, S., Lance, B.J., Lin, C.T.: Driver drowsiness estimation from EEG signals using online weighted adaptation regularization for regression (OwARR). *IEEE Trans. Fuzzy Syst.* **25**(6), 1522–1535 (2016)
39. Wei, C.S., Lin, Y.P., Wang, Y.T., Jung, T.P., Bigdely-Shamlo, N., Lin, C.T.: Selective transfer learning for EEG-based drowsiness detection. In: *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3229–3232 (2015)
40. Wei, C.S., Lin, Y.P., Wang, Y.T., Lin, C.T., Jung, T.P.: A subject-transfer framework for obviating inter-and intra-subject variability in EEG-based drowsiness detection. *NeuroImage* **174**, 407–419 (2018)
41. Wu, D., Chuang, C.H., Lin, C.T.: Online driver’s drowsiness estimation using domain adaptation with model fusion. In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 904–910 (2015)
42. Jiang, Y., Zhang, Y., Lin, C., Wu, D., Lin, C.T.: EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system. *IEEE Trans. Intell. Transp. Syst.* **22**, 1752–1764 (2020)
43. Chen, L.L., Zhang, A., Lou, X.G.: Cross-subject driver status detection from physiological signals based on hybrid feature selection and transfer learning. *Expert Syst. Appl.* **137**, 266–280 (2019)
44. Zhang, J., Wang, Y., Li, S.: Cross-subject mental workload classification using kernel spectral regression and transfer learning techniques. *Cogn. Technol. Work* **19**(4), 587–605 (2017)
45. Liu, Y., Lan, Z., Cui, J., Sourina, O., Müller-Wittig, W.: Inter-subject transfer learning for EEG-based mental fatigue recognition. *Adv. Eng. Inform.* **46**, (2020)
46. Yan, K., Kou, L., Zhang, D.: Learning domain-invariant subspace using domain features and independence maximization. *IEEE Trans. Cybern.* **48**(1), 288–299 (2017)
47. Gretton, A., Borgwardt, K., Rasch, M., Schölkopf, B., Smola, A.: A kernel method for the two-sample-problem. *Adv. Neural. Inf. Process. Syst.* **19**, 513–520 (2006)
48. Zhang, Y.Q., Zheng, W.L., Lu, B.L.: Transfer components between subjects for EEG-based driving fatigue detection. In: Arik, S., Huang, T., Lai, W.K., Liu, Qingshan (eds.) *ICONIP 2015*. LNCS, vol. 9492, pp. 61–68. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-26561-2\\_8](https://doi.org/10.1007/978-3-319-26561-2_8)
49. Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* **6**(2), 156–166 (2005)
50. Zheng, W.L., Lu, B.L.: A multimodal approach to estimating vigilance using EEG and forehead EOG. *J. Neural Eng.* **14**(2), (2017)
51. Wu, D., Lance, B.J., Parsons, T.D.: Collaborative filtering for brain-computer interaction using transfer learning and active class selection. *PLoS ONE* **8**(2), (2013)

52. Shalash, W.M.: Driver fatigue detection with single EEG channel using transfer learning. In: 2019 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 1–6 (2019)
53. Li, H., Zheng, W.L., Lu, B.L.: Multimodal vigilance estimation with adversarial domain adaptation networks. In: 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (2018)
54. Parsons, T.D., Courtney, C.G., Dawson, M.E.: Virtual reality stroop task for assessment of supervisory attentional processing. *J. Clin. Exp. Neuropsychol.* **35**(8), 812–826 (2013)