Driver Distraction Detection Based on EEG Feature Fusion Using Random Forest

Xin Zuo, Chi Zhang, Jian Zhao, Timo Hämäläinen and Fengyu Cong

July 10, 2022

Driver Distraction Detection Based on EEG Feature Fusion Using Random Forest

Xin Zuo[1,4], Chi Zhang[1,2], Jian Zhao[3], Timo Hämäläinen[4], Fengyu Cong[1,4]
[1]*School of Biomedical Engineering, Dalian University of Technology,* Dalian, China
[2]*Liaoning Key Laboratory of Integrated Circuit and Biomedical Electronic System,* Dalian, China
[3]*School of Automotive Engineering, Dalian University of Technology,* Dalian, China
[4]*Faculty of Information Technology*, *University of Jyväskylä,* Jyväskylä, Finland
Email: zuoxin93@foxmail.com, {chizhang, jzhao, cong}@dlut.edu.cn, timo.t.hamalainen@jyu.fi

Abstract*—Driver distraction has been one of the primary causes of traffic accidents. Electroencephalography (EEG), a record of the electric potential from the scalp, is considered as a reliable indicator of brain activities. It has been widely used to detect driver distraction. Previous studies have analyzed driver distraction based on time and frequency domain features of EEG. However, challenges still exist in manifesting the distraction information of EEG which contains a large amount of complex information about driver distraction in realistic driving scenarios from the perspective of complexity. In this paper, we propose a driver distraction detection framework using Random Forest (RF) based on the complexity feature fusion of EEG in real driving environment. Five complexity-based features of EEG are firstly extracted with a sliding window. Then, an RF classifier is trained with the extracted features to detect driver distraction. Our results show that differential entropy (DE) with an accuracy of 72.9% achieves the best result while single type feature is applied to detect distraction. The classifier's accuracy is further increased by about 7% using fused multiple features compared with the highest accuracy obtained by single type feature. In terms of feature contribution, we found that the feature with the best distraction detection result by using single type features may not contribute the most when using fused multiple features.

Keywords—EEG, driver distraction, feature fusion, entropy, random forest

I. Introduction

Driver distraction has been considered to be one of the main causes of car accidents, as it can reduce drivers' ability to manipulate cars and their awareness of potentially dangerous surroundings [1]. The National Highway Traffic Safety Administration (NHTSA) reported that about 2,800 people died and 400,000 were injured in traffic accidents involving driver distraction in 2018 while it rose to 3142 deaths and about 424,000 injuries in 2019 [2], [3]. There is a large number of factors diverting drivers' attention away from driving safely and thus leading to driver distraction, such as mobile phones, passengers, in-vehicle infotainment facilities, and so on. The factor of mobile phone usage ranked even the first among all possible factors [4].

To avoid potentially high-risk situations and prevent the happening of accidents caused by driver distraction, it is important to detect whether a driver is distracted or not. Many methods have been utilized in the literature to study driver distraction. Most of them divide driver distraction into four types (i.e., visual distraction, manual distraction, cognitive distraction, and audio distraction) and mainly focus on one type of distraction [5]-[8]. For example, Le *et al.* [9] designed an n-back digit recall experiment in both simulated and naturalistic driving scenarios to induce cognitive workload. The result showed that high level of distraction would be caused by tasks with high cognitive demand. Although these kinds of experiments can to some extent make contributions to the study of driver distraction, it is usually a combination of two or more distraction types in real driving scenarios. In Le's digit recall experiment, drivers firstly need to hear the voice instructions and bear in mind and then give responses when the same instruction appears. It actually induces both audio distraction and cognitive distraction, which is because driver distraction is caused by the interactions among driver, vehicle, and environment and it usually appears in a form of mixed types in real driving. Hence, challenges still exist in detecting driver distraction efficiently in real traffic.

Many kinds of data have been used to detect driver distraction, such as visual data, physiological data, and vehicle behavioral data in current research [10]-[12]. Physiological signals can provide more reliable information than other data types as they are reflections of the driver's actual internal state. Among all the physiological signals, electroencephalography (EEG) is used more frequently to estimate driver states with the superior performance of representing brain information [13]. For instance, Fan *et al.* [14] collected the EEG data in a simulated driving environment and proposed a time-series ensemble learning method to detect fatigue and distraction based on EEG features. Quantities of research have been done analyzing EEG data from the perspectives of time domain and frequency domain to study driver distraction. Yang *et al.* [15] extracted frequency domain EEG features like power spectral density and log-transformed power of four EEG frequency bands and used them to detect driver distraction. Wang *et al.* [16] utilized the frequency domain features, time domain features as well as time-frequency features to predict the duration of the distraction period and reached a satisfying result. However, EEG signals record electrical activities in the brain regions between pairs of electrodes on the scale. It not only reflects the temporal and spatial information of brain activity but also contains a large amount of complexity information [17]. The traditional most commonly used features may be not enough to manifest the useful complexity information to some degree. Recently, it has been demonstrated that the entropy based methods can explore the complex human state information contained in EEG in many research fields (e.g., sleep staging, disease detection, and mental stress detection). Su *et al.* [18] presented a sleep stage classification system with log energy entropy of EEG and found that the system has high generality which is consistent with the polysomnography records. Wang *et al.* [19] proposed a novel algorithm to predict the preictal state of seizure based on wavelet packet based entropy features of EEG and compared the results with traditional statistical features. The result showed that it reaches a higher classification rate than traditional features. Azami *et al.* [20] extracted the multi-scale entropy (MSE) feature of EEG to observe the dynamic complex brain activity information of Alzheimer's patients and found that MSE could mine for the dynamic EEG changes in an obvious way. Sharma *et al.* [21] extracted sample entropy (SE) and Renyi entropy (RE) at different frequency bands and used them to detect mental stress. Their results show the potential for reliable and timed detection of stress. Zheng *et al.* [22] trained a Deep Belief Network to recognize different emotions with differential entropy (DE) extracted from different brain regions. They found that DE can possess the useful information of EEG and achieve a high emotion classification accuracy.

Although entropy based methods have shown advantages in detecting human states, there is still a challenge in EEG analysis using complexity features. Different features reveal the implicit information of EEG from different aspects [23]-[25]. How the information compensation between different features happens in feature fusion step still needs to be studied. Hence, it is vital to evaluate the importance of different features to improve the classification performance.

A wide range of machine learning methods has been adopted to detect driver distraction in the literature. Random Forest (RF) proposed by Breiman [26] in 2001 is widely applied to classification tasks. It is an algorithm that integrates multiple decision trees according to the idea of ensemble learning. With the superior features of running fast on large databases and estimating variables' importance in classification, it has been used in many fields for classification, feature and channel selection, and so on. Zhang *et al.* [27] presented an advanced RF classifier to select informative features and classify motor imagery EEG with higher accuracy than prevailing approaches. Wang *et al.* [28] proposed an automatic epileptic seizure detection framework using an advanced RF model based on the time-frequency features of EEG and achieved high accuracy in detecting seizures.

In this paper, we propose a driver distraction detection framework based on complexity feature fusion using RF classifier. Non-intrusive wearable EEG sensors are firstly used to gather EEG signals in real driving scenarios. Then, different kinds of complexity based features in a sliding window are calculated to extract the complex distraction information in EEG. After that, the EEG features are fed into RF classifier to detect driver states and to estimate the importance of different features. The results of different kinds of features are finally compared.

The remaining part of the paper is structured as follows. Section II explains the designed experiment. Section III describes the adopted methodologies. The results are shown in Section IV and discussed in Section V. Section VI concludes the paper.

II. Experiment Design

This study was reviewed and approved by the Ethics Committee, Dalian University of Technology. An experiment was conducted on a real straight road at Dalian University of Technology. The Mangold-10 Bluetooth enabled wireless polygraph, a wearable and non-intrusive data acquisition headband, was used to collect EEG data. As the occipital brain region has been demonstrated to be related to driver mental state in previous studies, we put the headband's electrodes on O1 and O2 according to the International 10-20 System. The sample rate was set as 256 Hz.

We recruited six experienced right-handed drivers to participate in the experiment. All of them are mental health and have normal vision and auditory. Besides, they are also required to be experienced in using smartphones. In addition, all subjects are banned from smoking, and consuming drinks containing caffeine and alcohol the day before the experiment. Prior to participating in the experiment, we verified each subject's qualification and obtained the informed consent from them. What's more, written and oral instructions about the experiment were illustrated to all subjects.

The experiment contains one normal driving trail lasting for about 6 seconds and five distracted driving trials with a duration of about 20 seconds. In the normal trial, the subjects were supposed to pay full attention to driving while there were distracting factors in the distracted driving trials. In these trials, they would firstly focus on driving, then they would receive cellphone messages from the experimenter few seconds later. After that, subjects were asked to check the message for at least three seconds. Finally, they need to react to the obstacles that appeared on the road at the end of the trail. The EEG data was gathered from the car starting to stopping.

III. Methodology

It can be divided into three steps to analyze the EEG data including preprocessing, artifacts removal, and feature extraction.

A. Preprocessing

The EEG segments of each trial were extracted from the raw EEG signals at first. Alpha frequency band was then obtained using wavelet decomposition method, since alpha rhythm has been proved to be correlated highly with distraction [29].

Wavelet transform is widely used to extract sub-bands of EEG with the character of multi-resolution. To decompose the signal, a mother wavelet $\psi(t)$ is firstly utilized. Then, the signal can be expressed according to scaled and shifted versions of $\psi(t)$ and a corresponding scaling function $\phi(t)$ [30]. The discrete $\psi(t)$ can be expressed as

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi\left(2^{-j} t - k\right), \qquad k, j \in \mathrm{Z} \tag{1}$$

The signal $S(t)$ then is defined as

$$S(t) = \sum_{k} s_j(k) \phi_{j,k}(t) + \sum_{k} d_j(k) \psi_{j,k}(t) \tag{2}$$

where $s_j(k)$ and $d_j(k)$ are the approximate and detailed coefficients at level j.

B. Artifacts Removal

The obtained alpha frequency band contains artifacts like blinks that need to be removed in this step. A wavelet-based method was applied in this paper. According to large coefficients usually generated at the places where artifacts appear, we can decrease these large coefficients by thresholding technique [31], [32]. The threshold can be defined as

$$T_j = mean(C_j) + 2 \times \mathrm{std}(C_j) \tag{3}$$

This work was supported in part by the National Natural Science Foundation of China under Grant 61703069 and in part by the Fundamental Research Funds for the Central Universities under Grant DUT18RC(4)035.

where C_j is the wavelet coefficient at jth level of decomposition. If the value of any coefficient is greater than the defined threshold, it will be halved. A new set of wavelet coefficients are then reconstructed to obtain the artifacts removed signal.

C. Feature Extraction

We extracted five entropy based features to explore the distraction information of alpha frequency band including approximate entropy (AE), fuzzy entropy (FE), SE, DE, and MSE. As the first three features are more frequently used in the literature than the other two features, we mainly introduce the algorithm of DE and MSE in this part.

1) Differential entropy

DE is an extension of Shannon entropy so that it can be used to reflect the complexity of continuous variables [33]. It has been validated that DE is more accurate than features like energy spectrum and asymmetrical features in recognizing different emotion types based on EEG [34]. The calculation formula of DE is

$$\mathrm{DE} = -\int_a^b f(x)\log\left(f(x)\right)\mathrm{d}x \tag{4}$$

where $f(x)$ represents the probability density function of the continuous variable and $[a,b]$ shows the taking value interval. If the variable obeys Gaussian distribution $N(\mu,\sigma^2)$ approximately, its DE can then be expressed as

$$\mathrm{DE} = -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{2\sigma^2}}\right)\mathrm{d}x = \frac{1}{2}\log 2\pi e\sigma^2 \tag{5}$$

2) Multi-scale entropy

MSE can mine for the complexity information of signals in different time scales [20]. It involves two steps in MSE feature calculation: the coarse-graining process and SE calculation. The algorithm can be detailed as follows:

In the first step, for EEG signal $\{x_1, \ldots, x_i, \ldots, x_N\}$, a consecutive coarse-grained time series $\{y^{(\tau)}\}$ should be constructed corresponding to the scale factor τ. The coarse-grained time series $\{y^{(\tau)}\}$ is defined as

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \le j \le N/\tau \tag{6}$$

In the second step, The SE of time series $\{y^{(\tau)}\}$ is then calculated according to the following sub-steps.

- An m dimension vector $Y^m(i)$ can be made up firstly for time series $\{y_1, \ldots, y_j, \ldots, y_n\}$,

$$Y^m(i) = [y(i), y(i+1), \ldots, y(i+m-1)],\ 1 \le i \le n-m \tag{7}$$

- Define d as the absolute value of the maximum difference between the corresponding elements in vectors $Y^m(i)$ and $Y^m(j)$,

$$d = \max[|\, y(i+k) - y(j+k)\,|], \quad 0 \le k \le m-1, i \ne j, 1 \le j \le n-m \tag{8}$$

- Then count the number of $d < r$ for each i where r is the given threshold and $B_i^m(r)$ can be expressed as

$$B_i^m(r) = \frac{\{\text{the number of } d < r, i \ne j\}}{(n-m-1)} \tag{9}$$

- The set of $B_i^m(r)$ are then averaged and the average value $B^m(r)$ is defined by

$$B^m(r) = \frac{1}{n-m}\sum_{i=1}^{n-m} B_i^m(r) \tag{10}$$

- Add the dimension by 1 and repeat the above process, then $B^{m+1}(r)$ is obtained. After all the steps, SE is calculated by

$$\mathrm{SE}(m,r) = \lim_{n\to\infty}\left[-\ln\frac{B^{m+1}(r)}{B^m(r)}\right] \tag{11}$$

D. Random Forest Classifier

Algorithm 1

Input: T the training set
N the number of decision trees to be built
M the number of variables chosen for splitting at each node
Training: for each i=1:N **do**
1. Draw a bootstrap sample Bs from T.
2. Build tree Tr_b on bootstrap sample Bs.
3. Randomly select M candidate sets at each node of tree Tr_b, and find the best split among M sets.
4. Build tree Tr_b without pruning.
end for
Output: the ensemble of trees $\{Tr_b\}_1^N$
X the testing set
Classification: Assume $C_b(X)$ is the classification result of each tree. Then the result of RF
$C(X)$=majority vote $\{C_b(X)\}_1^N$

RF is a typical bagging model integrating multiple decision trees according to the idea of ensemble learning. In order to ensure the generalization ability of the model, the principles of random data and feature selection are followed while building each tree [35]. It works as follows [26]: bootstrap sample Bs is selected from the training set T at first, and decision trees Tr_b can then be built on the bootstrap samples. During this step, there is one-third of the samples are left called out-of-bag (OOB) data to calculate the classification error and to get estimates of variable importance in the classification step. After that, M variable candidate sets are randomly selected from the whole variable set at each split. Then select the best splitting way from M candidate sets and split at the node. To ensure a low bias, each tree is grown to the largest extent without pruning. After this step, the RF tree will repeat the above steps recursively until it is large enough to obtain the minimum classification error and then all decision tress $\{Tr_b\}_1^N$ are obtained. Finally, the trained RF classifier can be used to classify the testing set by voting for all trees' results. The pseudo-code of RF is shown in Algorithm 1.

IV. RESULTS

After obtaining the five entropy based features of alpha frequency band from all subjects, an RF model was trained using Algorithm 1. The data of five subjects was selected as the training set and the remained data was used as the testing set. RF classifier adopted the feature matrixes and corresponding label vectors of the training set to optimize the model parameters and then output the binary classification results of the testing set. In this paper, to compare the performance of RF using single type feature with that of multiple features, we trained classifiers for each type of feature and fused multiple features, respectively. The results of different features are shown in Table I. "ALL" stands for all the five entropy features of EEG.

TABLE I. THE MEAN ACCURACIES OF DIFFERENT FEATURES (%)

Feature	AE	DE	FE	MSE	SE	ALL
RF	58.88	72.9	65.42	68.22	63.55	79.51

As for the results of single type EEG features, it is shown in Table I that the mean accuracy of DE reaches 72.9%, which is obviously higher than the results of the other four entropy features. MSE, followed by FE and SE, ranks second with a classification accuracy of 68.22%. The AE feature leads to the lowest accuracy of the RF model, which is only 58.88%. the model performance increases significantly when using multi-features to detect driver distraction peaking at 79.51%. Furthermore, with RF's ability to output the importance of different features during classification, we also estimated each feature's importance in the feature fusion distraction detection process. The results are shown in Fig. 1.

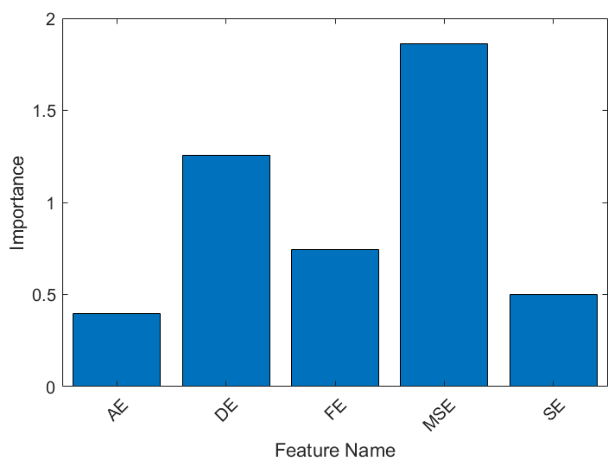

Fig. 1. The importance of different feature in distraction detection based on multiple features.

We can see clearly from Fig. 1 that there are significant differences in the importance of different EEG features for detecting driver distraction. MSE shows greater importance on the feature fusion distraction detection than the other four features. DE following MSE is the second important feature to detect distraction, which is contrary to the accuracy results of classification based on single type feature. Besides, AE still shows the lowest importance among all features and the importance of FE and SE lies between the importance of AE and DE.

V. DISCUSSION

Driver distraction has drawn a growing concern in recent years with the widespread usage of smartphones and advanced in-vehicle infotainment facilities [7]. An RF model to detect driver distraction is trained using five kinds of complexity based EEG features. The results are compared not only among different single type features but also between fused multiple features and single type features.

The results of driver distraction detection in Table I indicate that the mean accuracy of DE achieves the highest than other single type entropy based features. It is consistent with the results in [22], as DE has been proved to be a better feature to recognize human mental states. Moreover, the classification accuracy of multiple features, with an accuracy of 75.91%, is notably greater than that of any single type feature. Since different features can compensate for the inadequacy of each other [23], the RF model can learn more sufficient information from several different kinds of EEG features and then has a better performance in detecting driver distraction. By estimating the importance of different features in distraction detection using multiple features, we can know from Fig. 1 that MSE ranks the first among all features and DE is apparently less important than MSE on the classification results. The finding is not accordant with the classification results utilizing single type features in this paper but it corresponds to the results in our previous work. The BiLSTM model achieved the highest accuracy based on the MSE feature of EEG in [36], which might be because of the advantages of BiLSTM to learn the bidirectional long and short-term dependency of EEG. In this case, the MSE feature of EEG may reveal more contextual information in EEG and thus leading to the highest importance in the feature fusion distraction detection. Furthermore, the results in Table I and Fig. 1 also shows that single type of feature with which the best classification result is obtained may be not necessarily the feature with the most contribution after feature fusion.

VI. CONSLUSION

In this paper, we propose a driver distraction detection framework applying the RF classifier based on fused complexity features of EEG. It proves that DE feature is the best choice to explore the complex distraction information in EEG than other entropy features used in the literature while detecting driver distraction based on single type feature. Besides, the classifier's performance is greatly enhanced by fusing different EEG features, which demonstrates that different features can provide complementary distraction information of EEG. Additionally, MSE contributes the most among all features to detect driver distraction by fused features. It confirms that a feature achieving the best distraction detection result while using single type features may not contribute the most for driver distraction detection utilizing multiple features. Our work provides a machine learning method to detect driver distraction in real driving situations from the perspective of complexity features of EEG signals. It is useful for mining the complex dynamic brain activity information and driver distraction detection systems in real traffic.

ACKNOWLEDGMENT

We gratefully acknowledge the financial support from the National Natural Science Foundation of China (grant number: 61703069 and 62001312) and the Fundamental Research

Funds for the Central Universities (grant number: DUT21GF301).

References

[1] B. Darrow, “Distracted driving is now an epidemic in the U.S.,” Fortune, Sep. 14, 2016, [online], Available: http://fortune.com/2016/09/14/distracted-driving-epidemic/

[2] "Distracted driving 2018," National Highway Traffic Safety Administration, Washington, USA. Apr. 2020. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812926

[3] "Distracted driving 2019," National Highway Traffic Safety Administration, Washington, USA. Apr. 2021. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813111

[4] "Research on distracted driving in China," Ford Motor Company, and Tsinghua University, Beijing 2017.

[5] D. Gerónimo, A. M. López, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 7, pp. 1239-1258, Jul. 2010.

[6] J. D. Lee, "Dynamics of driver distraction: The process of engaging and disengaging," Ann Adv Automot Med, vol. 58, pp. 24-32, Mar. 2014.

[7] A. Gaffar and S. M. Kouchak, "Quantitative driving safety assessment using interaction design benchmarking," in Proc. IEEE SmartWorld Ubiquitous Intell. Comput. Adv. Trusted Comput. Scalable Comput. Commun. Cloud Big Data Comput. Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 2017, pp. 1-8.

[8] N. Lavie, "Attention, distraction, and cognitive control under load," Curr. Dir. Psychol., vol. 19, no. 3, pp. 143-148, 2010.

[9] A. S. Le, T. Suzukic, and A. Hirofumi, "Evaluating driver cognitive distraction by eye tracking: From simulator to driving," Transp. Research Interdiscip. Persp., vol. 9, pp. 100307, 2021.

[10] M., Gjoreski, M., Ž Gams, M., Luštrek, P., Genc, J., U Garbas, and T., Hassan, "Machine learning and end-to-end deep learning for monitoring driver distractions from physiological and visual signals," IEEE Access, vol. 8, pp. 70590-70603, 2020.

[11] M. Taherisadr, P. Asnani, S. Galster, and O. Dehzangi, "ECG-based driver inattention identification during naturalistic driving using Mel-frequency cepstrum 2-D transform and convolutional neural networks," Smart Health, vol. 9-10, pp. 50-61, 2018.

[12] I. Jegham, A. Ben Khalifa, I. Alouani, and M. A. Mahjoub, "A novel public dataset for multimodal multiview and multispectral driver distraction analysis: 3MDAD," Signal Process.: Image Commun., vol. 88, pp. 115960, Aug. 2020.

[13] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," IEEE Trans. Intell. Transp. Syst., vol. 12, no. 2, pp. 596-614, Jun. 2011.

[14] C. Fan et al., "Detection of train driver fatigue and distraction based on forehead EEG: a time-series ensemble learning method," IEEE Trans. Intell. Transp. Syst., pp. 1-11, Nov. 2021.

[15] L. Yang, W. Guan, R. Ma, and X. Li, "Comparison among driving state prediction models for car-following condition based on EEG and driving features," Accid. Anal. Prev., vol. 133, pp. 105296, 2019.

[16] S. Wang, Y. Zhang, C. Wu, F. Darvas, and W. A. Chaowalitwongse, "Online prediction of driver distraction based on brain activity patterns," IEEE Trans.. Intell. Transp., vol. 16, no. 1, pp. 136-150, 2015.

[17] Z. Gao et al., "An adaptive optimal-Kernel time-frequency representation-based complex network method for characterizing fatigued behavior using the SSVEP-based BCI system," Knowl.-Based Syst., vol. 152, pp. 163-171, 2018.

[18] C. Su, L. Ko, J. Juang, and C. Hsu, "Hybrid sleep stage classification for clinical practices across different polysomnography systems using frontal EEG," in Processes. vol. 9, 2021.

[19] Y. Wang, J. Cao, X. Lai, and D. Hu, "Epileptic state classification for seizure prediction with wavelet packet features and random forest," in 2019 Chinese Control And Decision Conference (CCDC), pp. 3983-3987, 2019.

[20] H. Azami, D. Abásolo, S. Simons, and J. Escudero, " Univariate and multivariate generalized multiscale entropy to characterise EEG signals in Alzheimer’s disease," Entropy. vol. 19, 2017.

[21] L. D. Sharma, V. K. Bohat, M. Habib, A. M. Al-Zoubi, H. Faris, and I. Aljarah, "Evolutionary inspired approach for mental stress detection using EEG signal," Expert Syst. Appl., vol. 197, p. 116634, 2022.

[22] W. Zheng, J. Zhu, Y. Peng, and B. Lu, "EEG-based emotion classification using deep belief networks," in IEEE Int. Conf. Multimed. Expo, Chengdu, China, 2014, pp. 1-6.

[23] H. Liu and L. Yu, “Toward integrating feature selection algorithms for classification and clustering,” IEEE Trans. Knowl. Data En., vol. 17, pp. 491-502, April 2005.

[24] R. Desai, P. Porob, P. Rebelo, D. R. Edla, and A. Bablani, "EEG data classification for mental state analysis using wavelet packet transform and Gaussian process classifier," Wireless Pers. Commun., vol. 115, no. 3, pp. 2149-2169, 2020.

[25] M. J. Hasan and J. Kim, "A hybrid feature pool-based emotional stress state detection algorithm using EEG signals," in Brain Sciences. vol. 9, 2019.

[26] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5-32, 2001.

[27] R. Zhang et al., "A new motor imagery EEG classification method FB-TRCSP+RF based on CSP and random forest," IEEE Access, vol. 6, pp. 44944-44950, 2018.

[28] X. Wang, G. Gong, N. Li, and S. Qiu, "Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization," Front. Hum. Neurosci., vol. 13, 2019.

[29] J. Murphy, C. Devue, P. M. Corballis, and G. M. Grimshaw, " Proactive control of emotional distraction: evidence from EEG alpha suppression," Front. Hum. Neurosci., vol. 14, 2020.

[30] O. Rioul and M. Vetierli, "Wavelets and signal processing," IEEE Signal Proc. Mag., vol. 8, no. 4, pp. 14-38, Oct. 1991.

[31] C. Zhang et al., "Network entropy for the sequence analysis of functional connectivity graphs of the brain," Entropy, vol. 20, no. 5, pp. 311, Apr. 2018.

[32] C. Zhang, L. Sun, F. Cong, and T. Ristaniemi, "Spatio-temporal dynamical analysis of brain activity during mental fatigue process," IEEE Trans. Cogn. Dev. Syst., vol. 13, no. 3, pp. 593-606, 2021.

[33] A. Feutrill and M. Roughan, "A review of shannon and differential entropy rate estimation," Entropy. vol. 23, 2021.

[34 R. Duan, J. Zhu, and B. Lu, "Differential entropy feature for EEG-based emotion classification," in 2013 6th International IEEE/EMBS Confer. on Neural Engineering (NER), 2013, pp. 81-84.

[35] W. Chen, Y. Wang, G. Cao, G. Chen, and Q. Gu, "A random forest model based classification scheme for neonatal amplitude-integrated EEG," Biomed. Eng. Online, vol. 13, no. 2, p. S4, 2014.

[36] X. Zuo, C. Zhang, F. Cong, J. Zhao, and T. Hämäläinen, "Driver distraction detection using bidirectional long short-term network based on multiscale entropy of EEG," IEEE Trans. Intell. Transp., pp. 1-14, 2022.