A Multi-Modal Classification Method for Early Diagnosis of Mild Cognitive Impairment and Alzheimer's Disease Using Three Paradigms With Various Task Difficulties

Sheng Chen, Chutian Zhang, Hongjun Yang, Liang Peng, Haiqun Xie, Zeping Lv, and Zeng-Guang Hou, *Fellow, IEEE*

***Abstract*— Alzheimer's Disease (AD) accounts for the majority of dementia, and Mild Cognitive Impairment (MCI) is the early stage of AD. Early and accurate diagnosis of dementia plays a vital role in more targeted treatments and effectively halting disease progression. However, the clinical diagnosis of dementia requires various examinations, which are expensive and require a high level of expertise from the doctor. In this paper, we proposed a classification method based on multi-modal data including Electroencephalogram (EEG), eye tracking and behavioral data for early diagnosis of AD and MCI. Paradigms with various task difficulties were used to identify different severity of dementia: eye movement task and resting-state EEG tasks were used to detect AD, while eye movement task and delayed match-to-sample task were used to detect MCI. Besides, the effects of different features were compared and suitable EEG channels were selected for the detection. Furthermore, we proposed a data augmentation method to enlarge the dataset, designed an extra ERPNet feature extract layer to extract multi-modal features and used domain-adversarial neural network to improve the performance of MCI diagnosis. We achieved an average accuracy of 88.81% for MCI diagnosis and 100% for AD diagnosis. The results of this paper suggest that our classification method can provide a feasible and affordable way to diagnose dementia.**

***Index Terms*— Dementia, multi-modal, machine learning, domain-adversarial neural network.**

Manuscript received 14 October 2023; revised 23 January 2024; accepted 15 March 2024. Date of current version 4 April 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFC3603700, and Grant 2018YFC2001700; in part by the National Natural Science Foundation of China under Grant U1913601, Grant 62073319, and Grant U20A20224; in part by the Beijing Natural Science Foundation under Grant L212034; and in part by the Beijing Nova Program under Grant 20220484175. *(Corresponding authors: Hongjun Yang; Zeng-Guang Hou.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Institute of Automation, Chinese Academy of Sciences under Approval No. IA-201944.

Sheng Chen, Hongjun Yang, Liang Peng, and Zeng-Guang Hou are with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: chensheng2016@ia.ac.cn; hongjun.yang@ia.ac.cn; liang.peng@ia.ac.cn; zengguang.hou@ia.ac.cn).

Chutian Zhang is with the Department of Engineering Science, Faculty of Innovation Engineering, Macau University of Science and Technology, Macao, China (e-mail: 2109853pmi30005@student.must.edu.mo).

Haiqun Xie is with the First People's Hospital of Foshan, Foshan 528000, China (e-mail: haiqunx@foxmail.com).

Zeping Lv is with the Rehabilitation Hospital Affiliated to National Research Center for Rehabilitation Technical Aids, Beijing 100176, China (e-mail: lvzeping@163.com).

Digital Object Identifier 10.1109/TNSRE.2024.3379891

I. Introduction

CURRENTLY, more than 55 million people have dementia worldwide, this number is expected to reach 139 million by 2050 [1]. Alzheimer's Disease (AD) accounts for 60% - 70% of all dementia cases [2], it can cause irreversible cognitive damage such as memory, orientation and reasoning, and even interfere with daily life activities. Currently, AD can be well treated only if the diagnostic and treatment envelop fall back to the early prodromal stage [3], [4]. The neurological state prodromal to AD is known as Mild Cognitive Impaired (MCI). From the point of MCI diagnosis, 10% - 15% of patients will convert to AD per year [5]. Early diagnosis of MCI has become a critical factor in predicting the occurrence of AD in the long run [6]. However, the diagnosis of MCI requires various examinations, including examination based on scale measures, such as Mini-mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA) [7], which are comparably subjective and requires a high degree of professional knowledge of the doctor; and examinations based on neuroimaging, such as Functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET), which require expensive equipments. Due to the high examination cost and the lack of medical resources, many people who suffer from dementia cannot be diagnosed in time and intervened effectively, which makes it easier for mild symptoms to become worse. Therefore, it is necessary to deploy a more affordable method to detect dementia.

Electroencephalogram (EEG) is by far the most common non-invasive method for measuring electrical brain activity [8], and has become an important clinical tool for understanding brain activity and diagnosing brain disorders. Many researches used EEG to diagnose dementia, in which different features of EEG data and different paradigms for collecting EEG signals were used. Some researches [9], [10], [11], [12], [13] collected EEG signals in resting-state with eyes open or eyes closed. Siuly et al. [14] introduced piecewise aggregate approximation for compressing massive volumes of EEG data, and used permutation entropy and auto-regressive model features to detect MCI. Mitsukura et al. [15] used the ratio of power spectrum at each frequency to discriminate among the three classes of dementia and achieved an accuracy of 88.89%; Ieracitano et al. [10] evaluated the average time-frequency map for each channel, extracted some statistical coefficients from the main five conventional EEG sub-bands, and used 5 different classifiers to distinguish AD. Ieracitano et al. [9] developed a multi-modal features extraction methodology based on continuous wavelet transform and bispectrum analysis for classifying AD, MCI and healthy controls. Durongbhan et al. [13] extracted the frequency-based and time-frequency-based features in resting-state EEG with eyes open and closed and developed a classification framework for AD detection. Others [7], [16], [17] collected EEG signals with specific experimental paradigms. Khatun et al. [16] collected single-channel EEG data when the participants were stimulated by sound, extracted and ranked 590 features from event-related potential, achieving 87.9% accuracy using Support Vector Machine (SVM) classifier. Zeng et al. [7] collected EEG when the participants were doing a task-state experiment, and proposed a workload index combining the EEG data of subjects at α and β frequency bands to assess the mental cognitive status of MCI. San-Martin et al. [18] recorded EEG data during N-back working memory tasks, they performed frequency domain analysis of current source density (CSD) patterns in task-related EEG and extracted spectral CSD features using SVM classifier to distinguish MCI and healthy elderly. Timothy et al. [19] collected EEG data using short term memory task and eyes closed resting-state, and classified MCI by combining complexity and synchronization features based on quantifiers from the common platform of recurrence based analysis, the results reveal that the short term memory task performed better than eyes closed condition.

There are some limitations in existing researches. Most existing studies include 20 - 50 subjects, the accuracy varies from 76% to 99% in different studies due to the effect of a small sample size. Many studies did not introduce the details about how their data samples are defined, and data leakage can easily occur when dividing datasets. Besides, due to the great individual differences in EEG signals, the model is easy to overfit, the results of the same dataset but divided into different train sets and test sets may vary a lot, many studies did not divide the dataset multiple times to verify the generalization of the model. Furthermore, most studies used only single-modal EEG signals to feed to models, while different modalities data can represent information about different aspects of the subject, so using multi-modal data may be more effective than spending a lot of effort tuning the parameters of the model.

In this paper, we propose a classification method fusing EEG data, eye movement data and behavioral data to detect AD and MCI patients. The major contributions of this work are as follows:

- Data from different paradigms and different modalities are analyzed to obtain the basis for feature extraction, and the performances of different features are compared to obtain better features for the diagnosis of dementia.
- A data augmentation method is proposed to enlarge the dataset, an extra ERPNet feature extract layer is designed to extract multi-modal features and domain-adversarial neural network is used to improve the performance of MCI diagnosis, an average accuracy of 88.81% is achieved for detecting MCI.
- The dataset is divided into different train sets and test sets, and multiple experiments are conducted in different dataset constructions to explore the generalization of the model.

II. Methodology

A. Data Recording

In this paper, 69 participants were recruited at the First People's Hospital of Foshan, Guangdong Province, China, consisting of 10 AD patients, 32 MCI patients and 27 healthy controls (HC). The Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-5) [20] and the 2018 Guidelines for Diagnoses and Treatments of Dementias and Cognitive Impairments in China [21] were followed to formulate the diagnosis of AD, MCI or HC. Before the experiment, all the participants were informed about the research and signed the Informed Consent Form. The EEG signals were recorded by a dry electrode EEG device (DSI-24, Wearable Sensing) with 19 channels, sampling at a frequency of 300 Hz, and the eye movement signals were recorded by an eye tracker (Eye Tracker 5L, Tobii).

1) Resting-State Task: The Resting-state task was chosen to distinguish AD from non-AD (including healthy elderly and MCI patients) because it is more difficult for AD patients to understand and implement complicated paradigms, and the resting-state task is the paradigm that requires the least cooperation. In this task, participants were asked to sit in a chair to relax in a quiet room for 3 - 4 minutes with their eyes closed, but remain awake during the process.

2) Delayed Match-to-Sample (DMS) Task: The DMS task was chosen to distinguish MCI from HC based on previous work [22]. The schematic of the DMS task is shown in Fig. 1. In this task, participants sat in front of the computer screen with each hand on one button, respectively. At the beginning of each block, a target image with a green border appeared on the screen for 5 s, participants were asked to remember the target image, and after a period of delay, 5 test images that each lasted for 2 s appeared in sequence, during this time, participants needed to judge whether these appearing images were the same as the target image, if the test image was the same as the target image (match stimulus), the participant

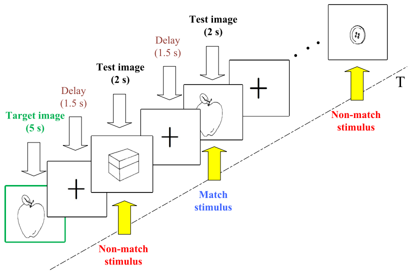

Fig. 1. Schematic of the delayed match-to-sample (DMS) task.

needed to press the button on the right, if not (non-match stimulus), press the button on the left. The test image and the target image were the same or different were regarded as two different stimuli to the participants.

3) Eye Movement Task: In this task, participants were asked to sit in a chair and perform the following tasks as required, four eye tracking experiments were performed:

- Fixation: The participants focus their gaze on the target as accurately as possible.
- Pursuit: The participants follow a moving target as closely as possible with their eyes.
- Pro-saccade: The participants first fixate on the fixation point, and then when the visual stimulus point appears, the participants need to make a rapid saccade towards the visual stimulus point.
- Anti-saccade: The participants first fixate on the fixation point, and then when the visual stimulus point appears, the participants need to make a rapid saccade to the mirror position of the visual stimulus point.

B. Feature Extraction

Feature extraction can reduce the dimension of the raw data while maintaining the integrity of the description of the data. In this paper, we extracted the following features from different modalities of data to train the model:

1) Eye Movement Features (EF): In the eye movement task, gaze point, eye position and pupil diameter were recorded by the eye tracker, and some metrics of specific events were counted in each eye movement experiment. The Identification by Velocity Threshold (IVT) method [23], [24] was used to detect saccade. In the IVT model, the velocity value is computed for every eye position sample, the eye position sample was marked as a saccade if the sampled velocity was greater than the manually set threshold. The numbers of saccades were counted in the fixation and pursuit experiments. Besides, the latencies of saccade in the saccade experiment and antisaccade experiment were also used as extracted features. The latency [25] refers to the time difference between the visual stimulus point starting to move and the point of the participant's gaze starting to move, it can be defined as:

$$\text{latency} = t - t_{\text{onset}} \tag{1}$$

where t denotes the time when the first saccade occurred after the event began; t_{onset} denotes the time the visual stimulus point started to move (in the saccade experiment and anti-saccade experiment). In the anti-saccade experiment, some participants made a saccade towards the visual stimulus point or paused too long at the fixation point, the antisaccade error rate was calculated to reflect how well the participant completed the experiment, the error rate was the proportion of times that the participant fail to saccade in the opposite direction of the visual stimulus point. Thus, the EF are normalized number of saccades in the fixation and pursuit experiments, the latencies in the saccades and antisaccadic experiment, and the error rates in the antisaccadic experiment.

2) Power Spectral Density Features (PSDF): Power spectral density describes how the power of a signal is distributed with frequency [26]. In this paper, Welch method was used to estimate the PSD of each frequency band. Fast Fourier Transform (FFT) was performed on all epochs, and the average magnitudes of all FFT coefficients in a specific frequency band were used as PSD features, it can be written as:

$$PSD_i = \frac{\sum_{k=fs_i}^{fe_i} |Y_k|}{fe_i - fs_i} \tag{2}$$

where fs_i and fe_i is the start and end frequency in the i-th frequency band respectively, and Y_k are the FFT coefficients. In this paper, frequency-domain features are extracted according to 4 conventional frequency bands.

3) Continuous Wavelet Transform Features (CWTF): The wavelet transform represents the signal in terms of both time and frequency, and can well express the changes of non-stationary signals [27]. It can be defined as

$$C_{i,j} = C(\tau_i, a_j) = \frac{1}{\sqrt{a_j}} \sum_{-\infty}^{\infty} x(t)\psi^*(\frac{t-\tau}{a_j})dt \tag{3}$$

where $C_{i,j}$ are the wavelet transform coefficients at sample τ_i and scale a_j; $x(t)$ is the signal, and ψ denotes mother wavelet. Morlet wavelet is used as the mother wavelet in this paper.

The average magnitudes of all wavelet transform coefficients in a specific frequency band over the signal length were used as CWT features. it can be written as:

$$WF_i = \frac{\sum_{j=ts}^{te} \sum_{k=fs}^{fe} C_{i,j}}{(te - ts)(fe_i - fs_i)} \tag{4}$$

where ts is the start of epoch in the resting-state task or the time when the target image appears in the DMS task, and te is the end of epoch in resting-state task or the time when the next image appears in the DMS task.

4) ERP Features (ERPF): Event-related potentials (ERPs) are obtained by extracting multiple time periods from an ongoing EEG that define the event of interest and then averaging them [28], [29]. In the experiment of this paper, the DMS task was an ERP-triggered task, EEG was triggered differently in match and non-match stimuli, and literature [30], [31], [32] have showed that stable grand-average ERPs can be obtained with 15 or fewer trials. We made differences between the ERP waveforms of match stimuli and non-match stimuli at specific time points, and extracted the following 6 time-domain features within the period of 300 - 600 ms. The reason for extracting these features is based on the data analysis.

- The mean of the maximum/minimum values between the match stimuli and non-match stimuli.

- The difference of the maximum/minimum values between the match stimuli and non-match stimuli.
- The ERP latency differences of the maximum/minimum values between the match stimuli and non-match stimuli.

5) Behavioral Features (BF): In the DMS task, the participant's degree of concentration and ability to judge correctly in the task also reflected important information. The information about when each participant pressed the button at each epoch and whether the participant made the correct action as the task required were recorded by E-Prime software. BF are the correct/error rate, missing rate (participant doesn't press any button within 2 s of the test image appearing) and the average reaction time (RT) of the participant calculated by these information.

C. Dataset Construction and Evaluation Methods

There are two ways to define samples, different sample definition ways will affect the size of the data set.

A conventional way is to regard each subject as a sample, in this way, the sample size is equal to the number of the subjects. Because of the small sample size, the model can only be trained with traditional classifiers with few parameters. To use more samples when training the model, leave-one-out method was used to evaluate the model. Thus, we built a number of models equal to the number of subjects, and each model was either correct or incorrect on its corresponding test set. The ratio of the number of correctly classified models to the total number of models was used as the evaluation index of a specific classifier.

Another way of data augmentation is to regard each segment signal of each subject as a sample, but in this case, each sample is not independent, the different segments of the same subject are strongly correlated (they are all from the same subject), so it is necessary to restrict all samples of the same subject to be in one of the train set or test set at the same time to avoid data leakage from the train set to the test set. This restriction will also ensure that the feature extraction method becomes independent, thereby preventing the parameters of the feature extraction module from being influenced by segments of the same subjects in the test set. In this paper, all the subjects were divided into 5 parts, and each time one part was used as the test set, and the rest were used as the train set, we will gain 5 models at the end. It is worth noting that the accuracy directly calculated by the model is based on the segment in this case, and they needed to be converted into the subject-level accuracy at the end. The predicted label of a subject was determined by majority voting of segment labels of the same subject and compared with the real label of the subject to obtain subject-level accuracy. The accuracy of the classifier is defined as the average accuracy of the 5 models.

Data augmentation was only implemented in the EEG modality in this paper, since the eye movement features and behavioral features were fixed for the same subject. The features of the sample in this way included the features extracted from an EEG segment, the eye movement features and behavior features extracted from the subject, the label of the sample was the subject's label. In the resting-state task,

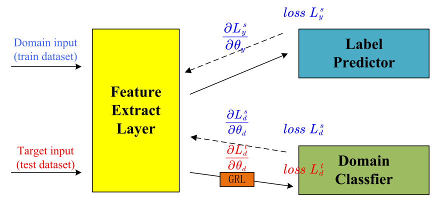

Fig. 2. Structure of domain-adversarial neural network.

each 2 s interval EEG epoch is regarded as a segment. In the DMS task, 30 epochs of the same stimuli (match stimuli or non-match stimuli) were randomly selected to superimpose to get the ERP waveforms as a segment. In this way, each segment reflected information about the subject, and all input samples were the real data on the subjects.

D. EEG Channel Selection

Feature selection is very important in machine learning, it can reduce the computation cost and improve the performance. Since the dimension of the features is high in the experiment, using channel selection as a feature selection method can not only eliminate some irrelevant interference, but also reduce the requirements for EEG device, so cheaper EEG devices that only measure a few channels can suffice for identifying dementia. Channel selection is also physiologically reasonable, because only some brain regions may better reflect the changes in the EEG signal during the task, while the EEG signals in other brain regions are more caused by other interference factors. In this paper, random forest was used to rank all the extracted features and get the weight of each feature, and then the weight of each channel was calculated according to the features contained in each channel. Several channels with larger weights and spatially closer locations were chosen to reduce the input data dimension.

E. Classification Model

In this paper, different classifiers were used for different dataset construction:

1) Support Vector Machine (SVM): SVM is a classifier that transforms the features into a higher dimensional space and then constructs an optimal separating hyperplane in the transformed space to separate the two classes [14], [33]. In the case of regarding each subject as a sample, due to the small sample size, it is advisable to utilize a classification model with a limited number of parameters. Consequently, SVM is well-suited for such scenarios. This study explored both the linear kernel and the radial basis function (RBF) kernel as options for SVM. To identify the optimal hyperparameters associated with the kernel, the grid search method was employed to select the best hyperparameters corresponding to the kernel. The accuracy of the validation set was also presented as a reference for the accuracy of the test set.

2) Domain Adversarial Neural Networks (DANN): In the case of data augmentation, the sample size is enough for us to use some complex models. DANN [34] was used as the classification model to distinguish MCI patients from healthy elderly. The structure of DANN is shown in Fig. 2. It consists of

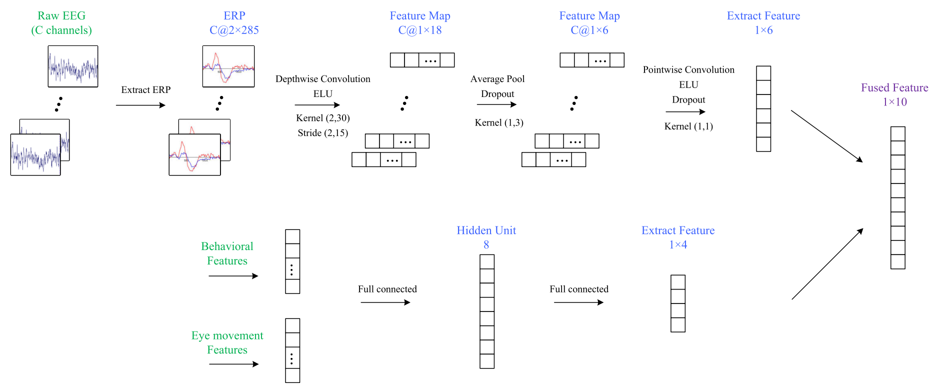

Fig. 3. Structure of ERPNet feature extraction layer.

three parts: feature extraction layer, label predictor and domain discriminator. The function of the feature extraction layer is to learn another representation of the source domain and target domain features to confuse the domain discriminator. Domain discriminator is used to judge whether the features come from the source domain or the target domain. Label predictor is used to classify the sample label. The optimization strategy is to minimize the loss of the label predictor and maximize the loss of the domain discriminator, making the distribution of the train set and the test set as consistent as possible. The total loss function is defined as:

$$L = \frac{\text{batchsize}^s}{\text{batchsize}^t} * L_d^t + L_d^s + \beta * L_y^s$$

where L_d^s is the loss of domain classifier using the source domain data; L_t^s is the loss of domain classifier using the target domain data; L_y^s is the loss of label predictor using the source domain data, and β is a parameter to adjust the weight of domain classifier and label predictor.

There are two ways to feed EEG signals as inputs into the classifiers: one is to extract various features in advance through the feature extraction method, and then feed the extracted features to the model; the other is to directly feed the original EEG signal or ERP waveform to the model, and use an extra feature extract layer to extract features. The latter way is to train the feature extractor through a neural network. The features extracted by the neural network have no practical meaning compared with the time-domain or frequency-domain features extracted based on theoretical knowledge, but it would not lose some information about the original signals and can extract the features of all aspects of the original signal to improve the performance. The results of both ways of feeding input data will be presented for comparison in this paper.

An ERPNet module was designed to extract EEG features and fused with other features in the DMS task, the structure of ERPNet is shown in Fig. 3. Since the behavioral features and eye movement features are not suitable for convolution, they were concatenated with the convolved EEG features to get the fused features. In the ERPNet, the ERPs of the match stimuli and non-match stimuli were extended to another dimension, and a depthwise convolution was used on the ERPs during the 950 ms post-stimulus period. The kernel size of the first dimension was set to 2 so that the difference between the match stimuli ERP and the non-match stimuli ERP can be obtained. The kernel size of the second dimension was set to 30 so that the ERP information of every 10 ms can be captured during the convolution process (the device samples at 300 Hz). Depthwise convolution can ignore the relationship between the ERPs of different channels, focus more on the information of two different stimuli in the specific channel, and reduce the number of trainable parameters to fit. The purpose of average pooling operation is to smooth the feature map, reduce dimension, and eliminate the influence of noise. Pointwise convolution can learn how to optimally mix the features of different channels and reduce the dimension. The main difference between ERPNet and EEGNet [35] is that the object of convolution in ERPNet is the ERPs of different stimuli, and the convolution operation emphasizes time-domain features and the difference between different stimuli, and reduces the number of parameters by depthwise convolution in the channel dimension. While the object of convolution in EEGNet is EEG signal at different band-pass frequencies, the convolution operation reduces the number of parameters by depthwise convolution in frequency dimension.

III. Results

A. Data Analysis

1) Resting-State Task: In the resting-state task, a frequency domain analysis of the AD group and non-AD group is performed, the result is shown in Fig. 4. It can be found that the EEG power of the AD group was lower than that of the non-AD group in the frequency band of α and β, but higher than that of the non-AD group in the frequency band of δ. The conclusion is similar to the study [36], but the same conclusion that θ rhythm is increased in the AD group was not reached in this paper. No similar differences were found between HC group and MCI group.

2) Eye Movement Task: In the eye movement task, EF were calculated, the results are shown in Table. I. It can be found that the number of saccades in fixation experiment was more frequent as dementia becomes more severe, Two-Sample T-test revealed that significant differences ($p<0.05$) were found in the number of saccades in fixation experiment between the AD and HC groups. No significant differences were found in

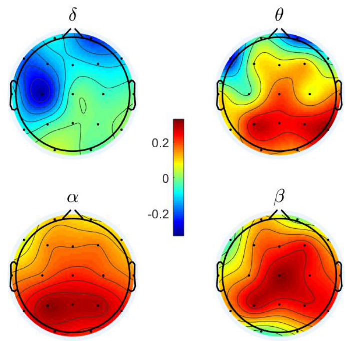

Fig. 4. Differences in spectrograms maps of four frequency bands between AD group and non-AD group.

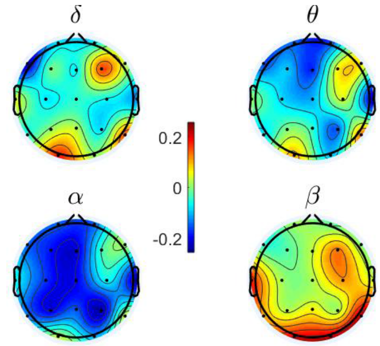

Fig. 5. Differences in spectrograms maps of four frequency bands between MCI group and HC group.

TABLE I
MEAN VALUE OF EYE MOVEMENT DATA IN THREE GROUPS

Group	Fixation	Pursuit	Latency[1]	Latency[2]	Error Rate
HC	5.0	28.6	295.1	505.7	6.3%
MCI	7.9	29.3	270.6	609.1	9.8%
AD	9.1	27.8	301.3	581.9	22.7%

Fixation: the number of saccades in fixation experiment;
Pursuit: the number of saccades in pursuit experiment;
Latency[1]: the latency of saccade in saccade experiment;
Latency[2]: the latency of antisaccade in antisaccade experiment;
Error Rate: error rate in antisaccade experiment.

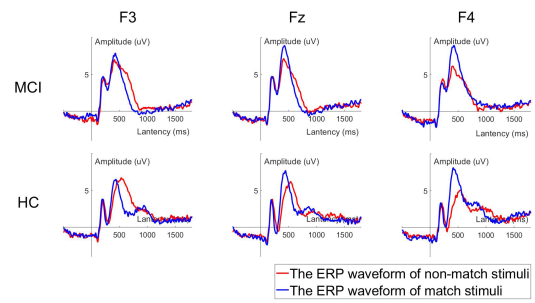

Fig. 6. ERP waveforms under match stimuli and non-match stimuli in MCI group and HC group.

TABLE II
BEHAVIORAL DATA IN THREE GROUPS IN DMS TASK

Group	Correct Rate	Error Rate	Missing Rate	RT (ms)
HC	0.98(0.02)	0.01(0.01)	0.01(0.01)	720(81)
MCI	0.96(0.03)	0.02(0.02)	0.01(0.01)	767(107)
AD	0.59(0.15)	0.23(0.16)	0.18(0.13)	1089(187)

the number of saccades in pursuit experiment between any two groups. In addition, AD group had a higher error rate in antisaccade experiment, and significant differences ($p<0.05$) were found in the error rate between the AD and HC groups. Furthermore, significant differences ($p<0.01$) were found in the latency in antisaccade experiment between the MCI group and the HC group, but the same result was not found between AD group and HC group, this might be because AD group has a higher error rate in antisaccade experiment, which leads to the fact that AD group did not do the same experiment as other group.

3) DMS Task: In the DMS task, a frequency domain analysis of MCI group and HC group was performed, the result is shown in Fig. 5. It can be found that the EEG powers of the MCI group were higher than or equal to that of HC group at the frequency band of α on all channels. On the contrary, the EEG powers of the MCI group were lower than or equal to that of HC group at the frequency band of β on all channels. The phenomenon suggests that MCI is likely due to insufficient power at the frequency band of β on F4-C4-T6-O1-O2 channels, while the power at the frequency band of α on the other channels makes up for this deficiency. This finding is consistent with the conclusion in [37], [38], and [39].

Fig. 6 shows the ERP waveforms of the MCI group and the HC group under the match stimuli and the non-match stimuli. It could be seen that the ERP waveforms under both match and non-match stimuli tend to be stable at about 1000 ms after the test image appears. But the maximum amplitude of the match stimuli was higher than that of the non-match stimuli, and the difference was largest in the right brain area, it is consistent with the conclusion in [40] that the HC group has a typical P300 match enhancement phenomenon. The ERP waveforms' amplitude of the HC group dropped more than that of MCI group at about 300 ms after the test image appeared, and the latency of the non-match stimuli was slower. The occurrence time of the extremely large and extremely small amplitudes was concentrated around 300 - 600 ms. Thus, these differences were used to extract features during this period.

The behavioral data of the participants in the DMS task are shown in Table. II. It can be seen that the error rates and reaction times increase as dementia becomes more severe. Two-Sample T-test revealed that significant differences ($p<0.001$) were found either in the correct rate, error rate, non-response rate and reaction time between AD and non-AD groups. Although the average error rate and reaction time of the MCI group were higher than those of the HC group, we did not find any features that were significantly different between HC and MCI groups.

B. The Result of EEG Channel Selection

When random forest was used to calculate the weight of channels, due to the different selection of random seeds, the weight order of channels was not fixed, but after averaging the weights calculated by the random forest more than 10 times,

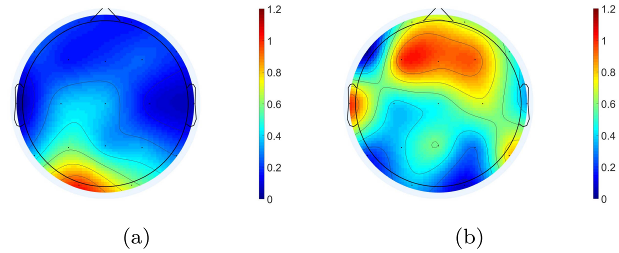

Fig. 7. Results of EEG channel selection: (a) The weights of each channel in the resting-state task; (b) The weights of each channel in the DMS task.

the order became stable and the channels with the largest weight tended to be the fixed channels.

1) Resting-State Task: The weights computed by random forest for each channel in the resting-state task are plotted in a topographic map in Fig. 7(a). It can be seen that the weights of the occipital lobe were relatively higher, which means that the channels in the occipital lobe region had a larger role in distinguishing AD, it may be because the occipital lobe is related to visual function [41], and AD patients have poor compliance and cannot help but open their eyes occasionally after closing their eyes for a long time. In this paper, the EEG signals of O1-O2 channels were used in the resting-state to distinguish AD.

2) DMS Task: The weights of each channel in the DMS task are plotted in a topographic map in Fig. 7(b). Unlike the resting-state task, the channels with relatively larger weights in the DMS task were concentrated in the frontal lobes, which is consistent with existing research [42], [43], [44] showing that the frontal area is related to the cognitive function of the brain. In this paper, the EEG signals of F3-Fz-F4 channels were used in the DMS task to distinguish MCI.

C. The Result of Distinguishing Between AD and Non-AD

1) The Effect of Different Features Without Data Augmentation: Table III shows the results of different features to distinguish AD from non-AD. The accuracy can reach 92.21% when using only eye movement features, accompanied by high specificity and low sensitivity. The accuracy obtained by using the EEG data in the resting-state can reach 94.44%, with higher sensitivity than using eye movement features. Moreover, to reduce the cross-subject differences in EEG signals, the ratios of frequency domain features at each frequency band were tried in this paper, and it was found that the performance of the ratio of frequency domain features performed better than directly using frequency domain features. It can also be seen that the accuracy of using only BF features in the DMS task is 98.48%, but DMS task is more complicated than eye movement task and resting-state task. It is more efficient to use eye movement data and resting-state EEG data to distinguish AD, and an accuracy of 100% was achieved after fusing $CWTF^2$ with EF.

2) The Results With Data Augmentation: Table IV shows the results of regarding an epoch as a sample. In this case, the ratio of extracted features ($CWTF^2$) still had better results than commonly used extracted features ($CWTF^1$). The accuracy was improved when it was finally converted into subject-level accuracy compared to epoch-level accuracy. Compared with the results of no data augmentation, the results became worse. This means that after data augmentation, the model did not necessarily gain more useful information, and the correlation between different segments was still too strong, so the purpose of improving performance hasn't been achieved. The reason for the higher accuracy of regarding each subject as a sample might be that the EEG signal itself is unstable, and the EEG signal obtained after the average superposition of multiple epochs is more accurate in that case.

TABLE III
RESULTS OF DIFFERENT FEATURES (AD VS NON-AD)

Features	Channels	Sen (%)	Spe (%)	Vali Acc (%)	Test Acc (%)
EF	\	**72.73**	**95.45**	**88.70**	**92.21**
$PSDF^1$	O1-O2	80.00	92.30	90.00	88.89
$PSDF^2$	O1-O2	90.00	88.46	88.30	88.89
$CWTF^1$	O1-O2	90.00	88.46	91.74	88.89
$CWTF^2$	O1-O2	**100.0**	**92.30**	**92.61**	**94.44**
BF	\	**100.0**	**98.18**	**98.43**	**98.48**
$CWTF^2$+EF	O1-O2	**100.0**	**100.0**	**99.18**	**100.0**

$PSDF^1$, $CWTF^1$: commonly used PSDF or CWTF features at each frequency band;
$PSDF^2$, $CWTF^2$: the ratio of extracted PSDF or CWTF features at each frequency band.

TABLE IV
RESULTS WITH DATA AUGMENTATION (AD VS NON-AD)

Level	Features	Sen (%)	Spe (%)	Vali Acc (%)	Test Acc (%)
epoch-level	$CWTF^1$	62.86	85.46	87.09	83.05
	$CWTF^2$	63.07	94.76	89.77	87.67
subject-level	$CWTF^1$	70.00	92.66	\	86.07
	$CWTF^2$	70.00	100	\	91.42

$CWTF^1$: commonly used CWTF features at each frequency band;
$CWTF^2$: the ratio of extracted CWTF features at each frequency band.

D. The Results of Distinguishing Between MCI and HC

1) The Results Without Data Augmentation: Table V shows the results of different features to distinguish MCI from HC without data augmentation. It can be seen that eye movement features and behavioral features were not effective in distinguishing MCI. From the analysis in the previous section, it can also be seen that there were no significant differences between these features between the MCI group and the HC group, so lower accuracy rates for these two features were to be expected. The performances were much better using features extracted from EEG, especially with ERPF, the accuracy reached 74.58%.

By comparing the validation set accuracy and test set accuracy in Table V, it can be seen that the accuracy of the validation set was not always higher than the accuracy of the test set. This means that the hyperparameters adjusted on the validation set were more suitable for the test set in these cases. The reason for this phenomenon is that the feature

TABLE V
RESULTS OF DIFFERENT FEATURES WITHOUT DATA AUGMENTATION (MCI VS HC)

Features	Sen (%)	Spe (%)	Vali Acc (%)	Test Acc (%)
EF	74.07	46.90	58.75	59.32
BF	66.67	43.75	59.31	54.24
PSDF	84.38	44.44	63.80	66.10
CWTF	62.50	62.97	66.15	62.71
ERPF	**87.50**	**59.23**	**74.55**	**74.58**
ERPF+PSDF	87.50	62.97	76.04	76.27
ERPF+PSDF+BF	87.50	62.97	78.20	76.27
ERPF+PSDF+BF+EF	**87.50**	**74.07**	**78.20**	**81.35**

TABLE VI
RESULTS USING EEG SINGLE-MODAL DATA WITH DATA AUGMENTATION (MCI VS HC)

Level	Features	Division number (%)					Average (%)
		1	2	3	4	5	
epoch	ERPF	75.11	80.10	72.19	65.78	63.63	71.36
	PSDF	75.57	69.23	66.84	63.10	65.21	68.00
	CWTF	71.04	60.63	74.87	63.64	78.61	69.76
	ERPF+PSDF	80.10	75.57	74.33	70.05	64.17	72.84
	CWTF+PSDF	83.25	64.25	79.67	63.63	75.93	73.35
	ERPF+CWTF	66.52	73.30	69.52	70.06	68.98	69.68
	ERPNet	78.73	76.47	80.21	63.63	75.40	74.89
subject	ERPF	**92.30**	**92.30**	**72.72**	**63.63**	**81.81**	**80.55**
	PSDF	76.92	69.23	72.72	63.63	63.63	69.23
	CWTF	69.23	61.53	81.82	63.64	90.91	73.43
	ERPF+PSDF	100	92.30	63.64	72.72	63.63	78.46
	CWTF+PSDF	84.61	61.54	90.91	63.63	72.72	74.68
	ERPF+CWTF	69.23	84.61	72.72	63.63	72.72	72.58
	ERPNet	**76.92**	**76.92**	**90.91**	**63.63**	**81.81**	**78.04**

TABLE VII
RESULTS OF DIFFERENT EEG CHANNELS USING MULTI-MODAL DATA WITH DATA AUGMENTATION (MCI VS HC)

Level	Methods	Channels	Division number (%)					Average (%)
			1	2	3	4	5	
epoch	DANN	F3-Fz-F4	77.82	80.10	70.05	69.52	77.54	75.01
	DANN	All	77.82	69.68	70.05	63.10	59.35	68.00
	ERPNet-DANN	F3-Fz-F4	78.73	86.87	76.47	76.47	92.51	82.21
	ERPNet-DANN	All	80.54	91.85	85.03	64.70	73.79	79.18
subject	DANN	F3-Fz-F4	76.92	84.61	81.81	81.81	81.81	81.39
	DANN	All	84.61	61.53	72.73	63.63	63.63	69.23
	ERPNet-DANN	F3-Fz-F4	**84.61**	**92.30**	**81.81**	**81.81**	**100.0**	**88.81**
	ERPNet-DANN	All	84.61	100	90.91	72.72	72.72	84.19

distribution of validation set was inconsistent with that of the test set due to the large individual differences of EEG signal. Among these features, the accuracy of validation set and the accuracy of the test set were the closest when using ERPF, which means ERPF was affected least by individual differences among these features.

It can also be seen that the performance improves by fusing different features, fusing ERPF with PSD can improve the accuracy, which indicates that although both ERPF and PSDF were features extracted from EEG signals, they had good complementarity, because ERPF focused on time domain information, and PSDF focused on frequency domain information. In addition, EF and BF can also provide some help for classification. After fusing the ERPF, PSDF, EF and BF features, the model achieved the highest average accuracy of 81.35%.

2) The Results Using EEG Single-Modal Data With Data Augmentation: In the case of data augmentation, DANN was used as the classifier of the model. Table VI shows the results of different features to distinguish MCI from HC with data augmentation. For the way of directly inputting the extracted features, it can be seen that ERPF is still the most effective feature for a single feature, the average accuracy at epoch-level was 71.36%, and the accuracy at subject-level was 80.55%. Although the average accuracy using PSDF or CWTF was still not as good as ERPF, it was better than using SVM model, which was because the feature extract layer of DANN makes the distributions of the training set features and the test set features tend to be consistent, reducing the individual difference.

For the way of using ERPNet as the extra feature extraction module, the average accuracy at epoch-level increased to 74.89%. That means the features extracted from ERPNet modules can also represent the difference between MCI group and HC group well. Comparing the results with and without data augmentation, it was found that the results were better with data augmentation, which is contrary to the previous conclusion that the accuracy of resting-state with data augmentation was lower than that without data augmentation. The phenomenon can be explained as that the "epoch" in the DMS task was not an original EEG segment, but an ERP waveform was obtained by superimposing a certain number of segments, which played a role in denoising. Conversely, the "epoch" in the resting-state task was simply a segment of the EEG, so the noise was larger and the signal was more unstable, making the extracted features inaccurate. It is worth noting that the average accuracy using ERPF was higher than using the ERPNet module at subject-level, but with lower accuracy at epoch-level, which shows the performance at epoch-level and the performance at subject-level were not completely uniform. The performance at the epoch-level can better reflect the quality of the model or features since the optimization goal of the model is at the epoch-level, and the performance at the subject-level can better reflect the possibility of practical application.

From the results of the 5 experiments, it can be seen that there are some differences between different experiments, the results of the 4th experiment performed poorly under any feature or feature extraction module. This shows that the selection of the train set and test set had a great influence on the accuracy of the model. The lower accuracy of the 4th experiment may be due to the poor EEG signals of specific subjects, or in other words, the EEG signals of some subjects may not be suitable to be used to distinguish whether he/she is MCI or HC.

3) The Results Using Multi-Modal Data With Data Augmentation: Table VII shows the results using multi-modal data. It can be seen that in each experiment, the accuracy of ERPNet-DANN was higher than that of using DANN directly. That means the features extracted by ERPNet module were indeed more effective than the features we extracted in advance. Furthermore, after fusing various modal features, none of the 5 experiments had a particularly poor performance. Compared with using single-modal EEG data, the accuracy of the 4th experiment had been greatly improved, which shows that the features of other modalities play a role in the 4th experiment and improve the performance. We can conclude that by using multi-modal features in various paradigms, the model has better generalization ability and performs better.

It can also be seen that the channel selection step is beneficial to the model. No matter which model was used, the accuracy of only using 3 channels was higher than that of using all channels. For different models, it can be seen that channel selection had less impact on ERPNet-DANN, but can greatly improve the performance on DANN. It is because the pointwise convolution operation in ERPNet-DANN reduced the dimension of the EEG channels, while the input of DANN were the features of all channels. which expanded the dimension of EEG signal features, after being fused with the features of other modalities, the effectiveness of other modal features would be greatly "diluted", and at the same time, the increase of the feature dimension did not provide more useful information, resulting in feature redundancy, so the model performance dropped a lot on DANN.

IV. Discussion

In this work, eye movement task and resting-state task are used to distinguish AD patients, because they are not complicated and do not require too much cooperation from the participants, they are particularly suitable for AD patients with severe symptoms. Referring to the results of this paper, almost all AD patients can be screened out. The detection of MCI requires participants to do another DMS task. By fusing the data collected in the DMS task and eye movement task, we achieved a reference accuracy of 88.81%. Compared with other studies, the accuracy of detection MCI in this paper is not the most outstanding, but data leakage was avoided in the dataset construction and model evaluation, and the final accuracy was close to the unbiased performance estimates. Although the EEG device had 19 channels in this paper, only 3 channels of the EEG data were used in the ERPNet-DANN model, the scenarios in which we collected the data are well suited for community screening.

The results show that EEG data had great advantages in detecting MCI compared to other modalities. Although the features extracted from other modalities had poor performance, they can indeed contribute to the detection of MCI after fusing with EEG data. Besides, we found that feeding only specific channels of EEG signals to the model could improve the performance, in future work, a cheaper device with fewer EEG channels can be used.

The results also show that the performance of the same dataset with different train sets and test sets differed a lot. This is a problem caused by small size data. So in this paper, the dataset was divided multiple times, and the average accuracy was used to evaluate the performance of the classification method. After using multi-modal data, the effect of different divide methods on the result became smaller, which alleviated this overfitting phenomenon.

Some prior analyses for feature extraction and feature selection were done in this paper, the features were selected based on this prior knowledge. The final results verified that the ERPF had the best classification performance and were least affected by individual differences. Besides, the results show that data augmentation helped improve the MCI classification model, and the structure of ERPNet feature extract layer we proposed can also extract EEG features and fuse them with other modality features well to detect MCI.

All above, this paper improved performance mainly in three ways:

- Find features that are suitable for the classification demand and less affected by individual differences, and discard irrelevant features at the same time.
- Use a more complex classification model. But a complex model means a larger sample size, the solution is usually to use a reliable and effective data augmentation method.
- Use multi-modal data. Some patients may not be sensitive to certain modality data, so using multi-modal data can make the classification model more stable.

A shortcoming of this work is that when we distinguished MCI from HC with data augmentation, the optimization goal of the model was inconsistent with the practical goal. The goal of the model training process was to find the model with the best epoch-level accuracy, but actually, the selected model was not optimal at subject-level, which made the model not necessarily the best model for classifying MCI patients. Since the accuracy at epoch-level is strongly correlated with the accuracy at subject-level, the final model also achieved ideal results.

V. Conclusion

In this work, a novel dementia classification method based on multi-modal data was proposed. The purpose of this work is to find a non-invasive, low-cost and more accurate solution for dementia diagnosis. Three paradigms of various task difficulties (eye movement task, resting-state task, and DMS task) were used to collect the data of the subjects. We achieved an accuracy of 100% for distinguishing AD patients using eye movement task and resting-state task, and an accuracy of 88.81% for distinguishing MCI patients using eye movement task and DMS task. Different features were analyzed and their performances were compared, we found that ERP in the DMS task had an advantage in distinguishing MCI patients. Based on this finding, we designed an ERPNet module to extract the features of ERP in the DMS task and fused it with other modality features. The results show that using multi-modal data can improve the model performance and the ERPNet module can extract features effectively to distinguish MCI patients. SVM and DANN were used as classification models under different dataset construction. The performance was improved using specific EEG channels, we found that O1-O2 were suitable for resting-state task to detect AD, while F3-Fz-F4 were suitable for DMS task to detect MCI. Our work provides a potential for the future diagnosis of dementia using EEG and eye movement signals. In the future, we will try different paradigms with more features of different modalities for dementia diagnosis.

References

[1] J.-H. Shin, "Dementia epidemiology fact sheet 2022," *Ann. Rehabil. Med.*, vol. 46, no. 2, pp. 53–59, Apr. 2022.

[2] A. Association, "2018 Alzheimer's disease facts and figures," *Alzheimer's Dementia*, vol. 14, no. 3, pp. 367–429, Mar. 2018.

[3] T. Karakaya, F. Fußer, J. Schroder, and J. Pantel, “Pharmacological treatment of mild cognitive impairment as a prodromal syndrome of Alzheimer’s disease,” *Current Neuropharmacol.*, vol. 11, no. 1, pp. 102–108, Jan. 2013.

[4] L. Robinson, E. Tang, and J.-P. Taylor, “Dementia: Timely diagnosis and early intervention,” *BMJ*, vol. 350, p. 3029, Jun. 2015.

[5] A. Breton, D. Casey, and N. A. Arnaoutoglou, “Cognitive tests for the detection of mild cognitive impairment (MCI), the prodromal stage of dementia: Meta-analysis of diagnostic accuracy studies,” *Int. J. Geriatric Psychiatry*, vol. 34, no. 2, pp. 233–242, Feb. 2019.

[6] S. Yang, J. M. S. Bornot, K. Wong-Lin, and G. Prasad, “M/EEG-based bio-markers to predict the MCI and Alzheimer’s disease: A review from the ML perspective,” *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2924–2935, Oct. 2019.

[7] H. Zeng et al., “EMCI: A novel EEG-based mental workload assessment index of mild cognitive impairment,” *IEEE Trans. Biomed. Circuits Syst.*, vol. 16, no. 5, pp. 902–914, Oct. 2022.

[8] G. R. Müller-Putz, “Electroencephalography,” in *Handbook of Clinical Neurology*, vol. 168. Amsterdam, The Netherlands: Elsevier, 2020, pp. 249–262.

[9] C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito, “A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia,” *Neural Netw.*, vol. 123, pp. 176–190, Mar. 2020.

[10] C. Ieracitano, N. Mammone, A. Bramanti, S. Marino, A. Hussain, and F. C. Morabito, “A time-frequency based machine learning system for brain states classification via EEG signal processing,” in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8.

[11] M. Şeker, Y. Özbek, G. Yener, and M. S. Özerdem, “Complexity of EEG dynamics for early diagnosis of Alzheimer’s disease using permutation entropy neuromarker,” *Comput. Methods Programs Biomed.*, vol. 206, Jul. 2021, Art. no. 106116.

[12] S. Ruiz-Gómez et al., “Automated multiclass classification of spontaneous EEG activity in Alzheimer’s disease and mild cognitive impairment,” *Entropy*, vol. 20, no. 1, p. 35, Jan. 2018.

[13] P. Durongbhan et al., “A dementia classification framework using frequency and time-frequency features based on EEG signals,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 826–835, May 2019.

[14] S. Siuly et al., “A new framework for automatic detection of patients with mild cognitive impairment using resting-state EEG signals,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 9, pp. 1966–1976, Jul. 2020.

[15] Y. Mitsukura, B. Sumali, H. Watanabe, T. Ikaga, and T. Nishimura, “Frontotemporal EEG as potential biomarker for early MCI: A case-control study,” *BMC Psychiatry*, vol. 22, no. 1, pp. 1–6, Dec. 2022.

[16] S. Khatun, B. I. Morshed, and G. M. Bidelman, “A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 1063–1070, May 2019.

[17] J. Li et al., “A cognitive electrophysiological signature differentiates amnestic mild cognitive impairment from normal aging,” *Alzheimer’s Res. Therapy*, vol. 9, no. 1, pp. 1–10, Dec. 2017.

[18] R. San-Martin, E. Johns, G. Q. Mamani, G. Tavares, N. A. Phillips, and F. J. Fraga, “A method for diagnosis support of mild cognitive impairment through EEG rhythms source location during working memory tasks,” *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102499.

[19] L. T. Timothy, B. M. Krishna, and U. Nair, “Classification of mild cognitive impairment EEG using combined recurrence and cross recurrence quantification analysis,” *Int. J. Psychophysiol.*, vol. 120, pp. 86–95, Oct. 2017.

[20] J. B. Williams and M. First, “Diagnostic and statistical manual of mental disorders,” in *Encyclopedia of Social Work*. Oxford, U.K.: Oxford Univ. Press, 2013.

[21] Neurologist Branch of the Chinese Medical Doctor Association, “2018 China guidelines for the diagnosis and treatment of dementia and cognitive impairment (V): Diagnosis and treatment of mild cognitive impairment,” *Chin. Med. J.*, vol. 98, no. 17, p. 8, 2018.

[22] S. Chen, H. Xie, H. Yang, C. Fan, Z. Hou, and C. Zhang, “A classification framework based on multi-modal features for detection of cognitive impairments,” in *Proc. China Intell. Robot. Annu. Conf.* Singapore: Springer, 2022, pp. 349–361.

[23] D. D. Salvucci and J. H. Goldberg, “Identifying fixations and saccades in eye-tracking protocols,” in *Proc. Symp. Eye Tracking Res. Appl. (ETRA)*, 2000, pp. 71–78.

[24] O. V. Komogortsev, D. V. Gobert, S. Jayarathna, D. H. Koh, and S. M. Gowda, “Standardization of automated analyses of oculomotor fixation and saccadic behaviors,” *IEEE Trans. Biomed. Eng.*, vol. 57, no. 11, pp. 2635–2645, Nov. 2010.

[25] M. Startsev and R. Zemblys, “Evaluating eye movement event detection: A review of the state of the art,” *Behav. Res. Methods*, vol. 55, no. 4, pp. 1653–1714, Jun. 2022.

[26] S. A. Unde and R. Shriram, “Coherence analysis of EEG signal using power spectral density,” in *Proc. 4th Int. Conf. Commun. Syst. Netw. Technol.*, Apr. 2014, pp. 871–874.

[27] P. Ghorbanian, D. M. Devilbiss, T. Hess, A. Bernstein, A. J. Simon, and H. Ashrafiuon, “Exploration of EEG features of Alzheimer’s disease using continuous wavelet transform,” *Med. Biol. Eng. Comput.*, vol. 53, no. 9, pp. 843–855, Sep. 2015.

[28] S. Sur and V. Sinha, “Event-related potential: An overview,” *Ind. Psychiatry J.*, vol. 18, no. 1, p. 70, 2009.

[29] M. D. Rugg, “Event-related potentials (ERPs),” in *Encyclopedia of Human Behavior* (ERPs), 2nd ed., V. Ramachandran, Ed. San Diego, CA, USA: Academic, 2012, pp. 79–85.

[30] M. A. Boudewyn, S. J. Luck, J. L. Farrens, and E. S. Kappenman, “How many trials does it take to get a significant ERP effect? It depends,” *Psychophysiology*, vol. 55, no. 6, Jun. 2018, Art. no. e13049.

[31] K. S. Button et al., “Power failure: Why small sample size undermines the reliability of neuroscience,” *Nature Rev. Neurosci.*, vol. 14, no. 5, pp. 365–376, May 2013.

[32] A. G. Fischer, T. A. Klein, and M. Ullsperger, “Comparing the error-related negativity across groups: The impact of error- and trial-number differences,” *Psychophysiology*, vol. 54, no. 7, pp. 998–1009, Jul. 2017.

[33] D. A. Pisner and D. M. Schnyer, “Support vector machine,” in *Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 101–121.

[34] Y. Ganin et al., “Domain-adversarial training of neural networks,” *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, Apr. 2016.

[35] V. Lawhern, A. Solon, N. Waytowich, S. M. Gordon, C. Hung, and B. J. Lance, “EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces,” *J. Neural Eng.*, vol. 15, no. 5, p. 056013, 2018.

[36] P. Monllor et al., “Electroencephalography as a non-invasive biomarker of Alzheimer’s disease: A forgotten candidate to substitute CSF molecules?” *Int. J. Mol. Sci.*, vol. 22, no. 19, p. 10889, Oct. 2021.

[37] J. Jeong, “EEG dynamics in patients with Alzheimer’s disease,” *Clin. Neurophysiol.*, vol. 115, no. 7, pp. 1490–1505, Jul. 2004.

[38] L. A. Coben, W. L. Danziger, and L. Berg, “Frequency analysis of the resting awake EEG in mild senile dementia of Alzheimer type,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 55, no. 4, pp. 372–380, Apr. 1983.

[39] L. A. Coben, W. Danziger, and M. Storandt, “A longitudinal EEG study of mild senile dementia of Alzheimer type: Changes at 1 year and at 2.5 years,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 61, no. 2, pp. 101–112, Aug. 1985.

[40] C. Guo, A. L. Lawson, Q. Zhang, and Y. Jiang, “Brain potentials distinguish new and studied objects during working memory,” *Human Brain Mapping*, vol. 29, no. 4, pp. 441–452, Apr. 2008.

[41] A. Rehman and Y. Al Khalili. (2023). *Neuroanatomy, Occipital Lobe*. [Online]. Available: http://europepmc.org/books/NBK544320

[42] E. K. Braithwaite, E. J. H. Jones, M. H. Johnson, and K. Holmboe, “Dynamic modulation of frontal theta power predicts cognitive ability in infancy,” *Develop. Cognit. Neurosci.*, vol. 45, Oct. 2020, Art. no. 100818.

[43] D. Cui et al., “Analysis of the neural mechanism of spectra decrease in MCI by a thalamo-cortical coupled neural mass model,” *J. Neural Eng.*, vol. 19, no. 6, Dec. 2022, Art. no. 066043.

[44] S. J. Chung et al., “Frontal atrophy as a marker for dementia conversion in Parkinson’s disease with mild cognitive impairment,” *Human Brain Mapping*, vol. 40, no. 13, pp. 3784–3794, Sep. 2019.