2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Evaluating the Fitness-to-Drive Using Evoked Visual Responses in Alzheimer's Disease

Ahmad Mitoubsi, Zeyu Liu, Danny Banks, Anahita Khojandi, Michael Oliver, Daniel Cox, and Roberto Fernandez

Abstract— **Alzheimer's Disease (AD) is the sixth leading cause of death in the US. AD causes significant disability due to the devastating impact on the patients' day-to-day living activities and their loss of independence. One such day-to-day activity is driving, a complex task that requires attention, concentration, the ability to follow particular steps, react to stimuli promptly, and the ability to perceive and interpret visual-spatial information, all of which can be impaired in AD. Therefore, to ensure the safety of AD patients and other drivers, it is important to develop accurate and low-cost diagnostic tools to assess patients' fitness-to-drive. In this study, we develop machine learning (ML) models to predict fitness-to-drive using the electroencephalogram (EEG) technique of event-related potential (ERP). Specifically, we develop random forest (RF) models using EEG signals in early-stage AD patients and age-matched controls and conduct numerical experiments to predict fitness-to-drive and other driving performance metrics, collected from driving simulator data. Our results show that RF models predict patients' fitness-to-drive with AUC=0.83 and provide accurate measures of other driving performance metrics. Therefore, ML and ERP offer a valuable approach to assess driving safety for patients with early AD symptoms in the laboratory setting.**

I. Introduction

Alzheimer's Disease (AD) is a neurodegenerative condition that causes progressive impairments in multiple cognitive domains, eventually resulting in dementia. AD has become more prevalent and is the most common form of dementia due to the growing aging population across the globe [1]. It is estimated that between 60-70% out of around 50 million people with dementia (around twice the population of Texas) have AD. In the US, about 5.8 million people aged 65 and older live with AD, of which, around 80% are 75 years old and older [2].

Although a vast number of individuals and their families are affected by AD, a diagnosis of AD at an early stage is often difficult. Usually, a definitive diagnosis is made once cognitive impairment compromises day-to-day living activities. Early signs of the disease include forgetfulness of recent events, word-finding difficulties, and deficits in visual-spatial perception that can present as getting lost or disoriented in familiar places. The steady decline in cognition impairs AD patients' ability to function independently. Consequently, this reduces the life expectancy of AD patients and is ranked the 6th leading cause of death in the U.S. [3].

Although currently there is no cure available for AD, early detection is important as certain medications may help treat symptoms, allowing for retaining functionality and the ability to live independently for a longer period [4]. In addition to standardized neuropsychological testing, which is the gold standard for clinical diagnoses, brain imaging and spinal fluid biomarkers are successfully used in preclinical detection [5]. Recently, non-invasive tools such as electroencephalography (EEG) and the EEG technique of event-related potentials (ERPs) have shown promise in the early detection of AD [5]. These techniques are often leveraged with machine learning (ML) models that allow for eliciting hidden patterns from data. For instance, classification methods such as the support vector machine (SVM) and random forest (RF) have shown great promise in detecting AD from EEG and ERP data [5, 6]. The studies are important as the combination of ML and EEG provides a cost-effective and non-invasive way for early detection of AD [5].

AD negatively impacts patients' day-to-day living activities [7]. One of such day-to-day activities is driving [8], which requires attention, concentration, the ability to follow particular steps and to react to stimuli promptly, all of which can be impaired due to AD. Driving also relies heavily on visual-spatial perception, a domain that is selectively affected in the early stages of the disease and may even precede the onset of memory deficits. Drivers with AD may not recall road regulations and routes. AD patients may also not see or perceive the distance to other cars or infrastructure and may not recognize the pattern of motion called optic flow that allows the determination of distance, speed, and direction of motion. More importantly, the unsafe driving status of AD patients can progress further as the disease develops [9].

Fitness-to-drive for AD patients can be determined using clinical interviews, neuropsychological assessment, and/or driving simulator rides [9]. Hence, as expected, driving simulators, which are safe and accurate methods of testing driver interactions, have shown to be effective in evaluating and quantifying various characteristics of AD**.** For instance,

Ahmad Mitoubsi, Zeyu Liu, Danny Banks, and Anahita Khojandi are with the Department of Industrial and Systems Engineering, University of Tennessee (corresponding authors: Anahita Khojandi, khojandi@utk.edu; Ahmad Mitoubsi, amitoubs@utk.edu)

Michael Oliver is with Psychological Science Department, Belmont University

Daniel Cox is with the Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine

Roberto Fernandez is with Pat Summitt Clinic, University of Tennessee Medical Center

Etienne et. al. [10] evaluates the executive functions associated with the impairment of mental flexibility in the early stages of AD and its impacts on driving activity. The study finds that AD patients have significantly lower performances in tests concerning their reaction times (RT) and the number of errors. Pavlou et. al. [11] indicate that drivers with cerebral diseases drive at lower speeds compared with the control group drivers. The results show that drivers with cognitive impairments significantly deviate from the general population [11]. This is especially the case for AD patients that drive significantly slower than the control patients by 68%. Stinchcombe et. al. [12] compares errors at intersection behavior in a driving simulator among mild AD drivers and healthy control drivers. Results indicate that across all types of intersections, mild AD drivers exhibit a greater number of errors relative to controls. The studies demonstrate that driving simulators can provide reliable measures of the differences between early-stage AD patients and controls, and they can indeed be directly used to accurately stratify the subjects into their corresponding cohorts [13].

To ensure the safety of AD patients and other road users, it is important to develop accurate and low-cost diagnostic tools to assess patients' fitness-to-drive in the laboratory setting. Driving simulators can be costly ranging from $20,000 for a desktop system to $100,000,000 for a full-vehicle simulator [14]. In addition, diving simulators are not widely available. By comparison, EEG is a relatively low-cost and non-invasive tool that is widely available in hospitals, clinics, and research facilities and is highly objective and quantifiable. Therefore, in this study, we aim to leverage the EEG technique of ERPs to predict fitness-to-drive and other driving performance metrics. To do so, we leverage ML in a cohort of AD patients and age-matched controls. Specifically, we collect patients' EEG signals using visual stimuli that simulate the patterns of motion seen by individuals as they make translational movements in the environment during ambulation or driving. We used the EEG technique of ERPs, which allows for the extraction of EEG signals associated with a specific event (e.g., onset of motion). We also assessed the participant's performance scores from a driving simulator. Then, employ an ML model, namely RF, to build the bridge between EEG signals and driving performances. We conduct several numerical experiments to assess the prediction accuracy of the RF model. We are predicting fitness-to-drive without knowing the cognitive status. This is so that the test can be applied to any subject seen in the clinic setting. As such, one of the main practical contributions of our study is to assess driving safety for patients with early AD symptoms in the clinical setting using the low-cost and safe method of ERPs.

II. METHODOLOGY

A. Data Description

In this study, we use two datasets, where the data are collected from a total of 29 subjects, of which 15 have early-stage AD and 14 are age-matched controls. All subjects provided informed consent prior to participation and all study procedures were approved by the University of Virginia Institutional Review Board. The first dataset includes subjects' EEG signals, collected using a 64-channel cap at 1000s/s for sampling, where subjects undergo visual ERP trials. EEG data are filtered using low and high pass filters at 30Hz and 0.1Hz. The temporal resolution of each EEG signal is measured and filtered by one millisecond. Five events are used for the visual ERP paradigm: fixation, pattern onset, motion onset, and change including acceleration and deceleration, and catch [5]. Each subject underwent 360 to 400 trials during approximately a 45-minute evaluation window. The first visual motion stimulus presented is the station fixation point lasting for 1000 milliseconds. The next stimulus is a station of random dots, named pattern onset, lasting for 1500 milliseconds. Following that is the motion onset which shows radial dot motion for 500 milliseconds. Then, there is a change in speed or direction for 1000 milliseconds, randomized at 37.5% of the trials for acceleration or deceleration in speed (each) and 25% of the trials for change in the direction, referred to as catch. At the end of the stimuli, stationary dots follow, namely the motion offset, lasting for 1500 milliseconds. The oddball paradigm was used to measure the reaction of the catch events [5].

The second dataset includes the performance scores of the same subjects from the driving simulator. Data are collected using driving simulations in a mid-range driving simulator at the University of Virginia. The Driver Guidance System (DGS-78) is employed for the driving simulation. It surrounds the patient with a 210° field of view. The driving simulator imitates the inside of a vehicle, including seatbelt, dashboard, steering wheel, mirrors, and all other usual controls [13]. The assessment of driver capability is conducted by 125 operational driving variables that evaluate subjects' visual, motor, cognitive, and executive function abilities. Time-based and score-based measurements are both included in the variables. Finally, the driving simulator provides a composite driving score, where a score less than 70 reflects the incompetence to drive.

B. Data Preprocessing

First, we focus on three events from the ERP trials, namely, acceleration, pattern onset, and catch [5]. This is because based on our past study, EEG data collected from acceleration (right frontotemporal activation), pattern onset (right lateralized and temporal regions), and catch (right-lateralized posteriorly) provide the best model performance in early AD detection [5]. Second, for each patient, we aggregate the data collected across the trials under the same event. For each patient, there are 360 to 400 trials for each event. An average was taken over all trials for each of the 64-channel for acceleration, pattern onset, and catch.

Finally, we extract features from each of the 64 channels, under the three events, for each patient. Features included minimum, maximum, mean, standard deviation, skewness, median, and variance of the EEG time series. Such a feature extraction approach is effective in capturing the temporal changes of time series [15]. The features are calculated from the data collected during 600 milliseconds after the stimulus. In total, we extracted $64 \times 7 = 448$ features from each event for each patient. Then, we select a subset of features to construct the model based on features importance from preliminary experiments.

C. ML Prediction

In this study, we use an RF classifier and regressor to predict patient driving performance. RF combines the prediction results of numerous decision trees and takes the majority vote as the final verdict [16]. As a flexible ML model, RF can be implemented in both classification and regression problems. In addition, RF is less likely to over-fit and can provide feature importance evaluation [16].

In this study, we conduct two experiments. First, we build an RF model on the preprocessed EEG data to classify if a patient is fit-to-drive using the composite driving score. The response variable is fitness-to-drive. We use 1 and 0, respectively, to denote a patient is fit-to-drive (composite driving score of 70 or above) or not. Patients with an RF-estimated probability of 0.5 or above are predicted to be fit-to-drive.

The second experiment in this study looks at other response variables from the driving simulation. Specifically, three response variables are included, namely the 'pothole avoidance steering average,' 'the long duration slow detection,' and 'total correct responses.' The pothole avoidance steering average indicates how well the patient reacts when approaching a pothole during the driving simulator. The long duration slow detection is the total amount of time it takes to slow down in response to the lead car's brake lights activating for 3 seconds. The total correct responses are the total number of times an individual responds appropriately and accurately, e.g., turned to avoid a pothole, and avoided the pothole. These variables are among the strongest contributors to the composite driving score and significant indicators of driving safety[13] [15] [16].

D. Model Evaluation & Metrics

For each event, the RF model uses 70% of the patients to train and 30% to test. The training and testing patients are chosen randomly. The numbers of AD patients and control patients in the training dataset are kept balanced to improve learning. The number of trees in the RF model is selected using the out-of-bag (OOB) error and is set to 100 trees. When testing, we conducted bootstrapping 100 times to calculate the mean performance and 95% confidence interval (CI).

For the experiment, the model is evaluated using six different classification metrics including sensitivity, specificity, precision, accuracy, F1 Score, and area under the curve (AUC). In the second experiment, we use mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) to evaluate the performance of the regression model.

III. RESULTS

Table I shows presents the descriptive statistics of the subjects for the AD and the control groups. As seen in the table, the composite score is overall lower for AD patients, compared with controls.

TABLE I. DESCRIPTIVE STATISTCIS.

Characteristics	Overall	AD	Controls
Subjects (%)	29 (100)	15 (52)	14 (48)
Male (%)	14(48)	6(40)	8(57)
Age (yr.), mean (IQR)	74.6 (61-85)	74.9(61-85)	74.2(63-82)
Composite score, mean (IQR)	58.4(-38, 117)	24.4(-38-91)	94.8(41-117)

Table II presents the performance metrics of the RF classifier in the first experiment. The table shows the mean and 95% CI for acceleration, catch, and pattern onset events. The acceleration classification showed the best results compared with catch and pattern onset, with 81% sensitivity and specificity, and an F1 score, AUC, and accuracy of 0.8, 0.83, and 80%. This suggests that the model built using EEG data collected under the acceleration event can confidently predict the subject's fitness-to-drive.

TABLE II. PERFORMANCE METRICS OF RF CLASSIFIER FOR THE FIRST EXPERIMENT. MEAN AND 95% CI ARE PROVIDED.

Events	Sensitivity	Specificity	Precision
Acceleration	0.81 (± 0.03)	0.81 (± 0.03)	0.80 (± 0.04)
Catch	0.81 (± 0.03)	0.78 (± 0.03)	0.75 (± 0.03)
Pattern Onset	0.81 (± 0.03)	0.75 (± 0.02)	0.73 (± 0.03)
Events	**F1 Score**	**AUC**	**Accuracy**
Acceleration	0.80 (± 0.03)	0.83 (± 0.03)	0.80 (± 0.03)
Catch	0.77 (± 0.02)	0.70 (± 0.03)	0.77 (± 0.02)
Pattern Onset	0.76 (± 0.03)	0.66 (± 0.04)	0.76 (± 0.02)

Fig. 1 presents the receiver operating characteristics curve (ROC) of the acceleration event. When plotting the ROC curve, all testing patients are used without bootstrapping. The ROC curve remains consistent with the results in Table II, with an AUC score of 0.90. The jagged shape of the ROC curve is most likely due to the small number of patients.

Table III provides a confusion matrix under the acceleration event when all testing patients are used without bootstrapping. In Table III, out of five AD patients and four controls in the testing set, one in each class was predicted incorrectly. Therefore, the results are overall consistent across the two cohorts.

The results of the second experiment are shown in Table IV. The results indicate that predicting the pothole avoidance steering average using the acceleration event leads to the best performance, with a minimal 0.0145 MSE, 0.1202 RMSE, 0.0797 MAE, and 0.3585 MAPE. Fig. 2 provides the detailed result for the predicted pothole avoidance steering scores against the corresponding true scores. As seen in figure 2, apart from a few exceptions, most of the predicted scores align very well with the true scores.

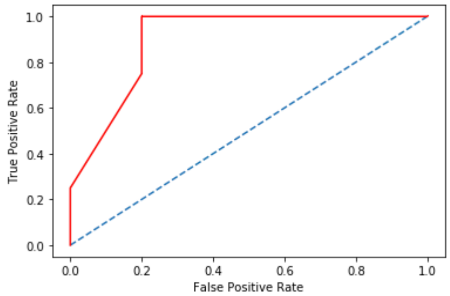

Figure 1. ROC curve for predicting fitness-to-drive using data from the acceleration event. The AUC score is 0.90.

TABLE III. CONFUSION MATRIX OF RF CLASSIFIER FOR ACCELERATION WITHOUT BOOTSTRAPPING.

		True condition	
		fit-to-drive	unfit-to-drive
Predicted condition	fit-to-drive	4	1
	unfit-to-drive	1	3

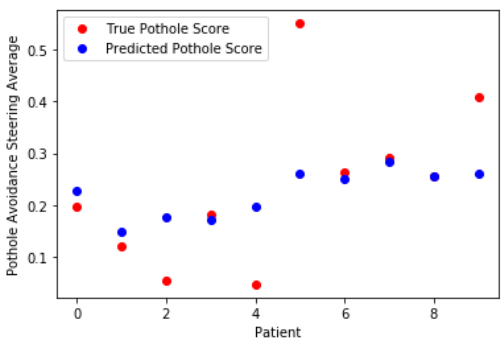

Figure 2: True and predicted score of pothole avoidance. An RF regressor is trained using the acceleration data to make predictions.

TABLE IV. PERFORMANCE METRICS OF RF REGRESSOR FOR THE SECOND EXPERIMENT.

Pothole Avoidance Steering Average				
Events	MSE	RMSE	MAE	MAPE
Acceleration	0.0145	0.1202	0.0797	0.3585
Catch	0.0169	0.1298	0.0945	0.4061
Pattern Onset	0.0147	0.1212	0.0956	0.3814
Long Duration Slow Detection				
Events	MSE	RMSE	MAE	MAPE
Acceleration	0.0280	0.1676	0.1488	0.1781
Catch	0.0274	0.1655	0.1469	0.1755
Pattern Onset	0.0329	0.1815	0.1635	0.2006
Total Correct Responses				
Events	MSE	RMSE	MAE	MAPE
Acceleration	4.4235	2.1032	1.7244	0.1268
Catch	2.5509	1.5971	1.2689	0.0968
Pattern Onset	6.0386	2.4573	2.1512	0.1592

IV. CONCLUSION

In this study, we employ an ML approach to predict the fitness-to-drive for subjects with and without early AD using the EEG technique of ERPs. Specifically, we train RF models with subjects' EEG signals collected during three events and conducted two experiments with various response variables. The results suggest that RF predicts subjects' fitness-to-drive with great accuracy, with an AUC score of up to 0.83. The results are promising, and future research will be needed. There is still more processing that will need to happen before implementing in clinical practice. In summary, using ERPs and ML shows great promise in evaluating fitness-to-drive in AD patients, providing a low-cost and non-intrusive method to ensure their safety. The two experiments are different approaches to determine fitness-to-drive in AD patients.

REFERENCES

[1] Weller, J. and A. Budson, *Current understanding of Alzheimer's disease diagnosis and treatment [version 1; peer review: 2 approved].* F1000 research, 2018. 7: p. 1161.DOI: 10.12688/f1000research.14506.1

[2] *Alzheimer's disease.* 2020; Available from: https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/symptoms-causes/syc-20350447.

[3] *Alzheimer's Disease Fact Sheet.* . 2019; Available from: https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet: .

[4] Kishore, P., C. Usha Kumari, M.N.V.S.S. Kumar, and T. Pavani, *Detection and analysis of Alzheimer's disease using various machine learning algorithms.* Materials today : proceedings, 2020.DOI: 10.1016/j.matpr.2020.07.645

[5] Mollie Turner, T.B., Anahita Khojandi, Michael Oliver, and Roberto Fernandez, A Minimal Set of Evoked Visual Responses Detect Alzheimer's Disease Early

[6] Sabbaghi, N., A. Sheikhani, M. Noroozian, and N. Sabbaghi, *Interval-based features of auditory ERPs for diagnosis of early Alzheimer's disease.* Alzheimer's & dementia : diagnosis, assessment & disease monitoring, 2021. 13(1): p. e12191-e12191.DOI: 10.1002/dad2.12191

[7] Marshall, G.A., R.E. Amariglio, R.A. Sperling, and D.M. Rentz, *Activities of daily living: where do they fit in the diagnosis of Alzheimer's disease?* Neurodegenerative disease management, 2012. 2(5): p. 483-491.DOI: 10.2217/nmt.12.55

[8] Hamdy, R.C., A. Kinser, T. Kendall-Wilson, A. Depelteau, K. Whalen, and J. Culp, *Driving and Patients With Dementia.* Gerontology and geriatric medicine, 2018. 4: p. 2333721418777085-2333721418777085.DOI: 10.1177/2333721418777085

[9] Piersma, D., A.B.M. Fuermaier, D. de Waard, R.J. Davidse, J. de Groot, M.J.A. Doumen, R.A. Bredewoud, R. Claesen, A.W. Lemstra, A. Vermeeren, R. Ponds, F. Verhey, W.H. Brouwer, and O. Tucha, *Prediction of Fitness to Drive in Patients with Alzheimer's Dementia.* PloS one, 2016. 11(2): p. e0149566-e0149566.DOI: 10.1371/journal.pone.0149566

[10] Etienne, V., C. Marin-Lamellet, and B. Laurent, *Mental flexibility impairment in drivers with early Alzheimer's disease: A simulator-based study.* IATSS research, 2013. 37(1): p. 16-20.DOI: 10.1016/j.iatssr.2013.06.005

[11] Pavlou, D., I. Beratis, E. Papadimitriou, C. Antoniou, G. Yannis, and S. Papageorgiou, *Which Are the Critical Measures to Assess the Driving Performance of Drivers with Brain Pathologies?* Transportation Research Procedia, 2016. 14: p. 4393-4402.DOI: 10.1016/j.trpro.2016.05.361

[12] Stinchcombe, A., S. Paquet, S. Yamin, and S. Gagnon, *Assessment of Drivers with Alzheimer's Disease in High Demand Driving Situations: Coping with Intersections in a Driving Simulator.* Geriatrics (Basel), 2016. 1(3): p. 21.DOI: 10.3390/geriatrics1030021

[13] Blanchette, R., A. Khojandi, D. Cox, M. Oliver, and R. Fernandez, *Predicting Alzheimer's Disease Using Driving Simulator Data.* 2020, IEEE. p. 5432-5435.

[14] *What game has the most realistic driving?* How much does a driving simulator cost? 2021; Available from: https://boardgamestips.com/destiny-2/what-game-has-the-most-realistic-driving/.

[15] Liu, Z., A. Khojandi, A. Mohammed, X. Li, L.K. Chinthala, R.L. Davis, and R. Kamaleswaran, *HeMA: A hierarchically enriched machine learning approach for managing false alarms in real time: A sepsis prediction case study.* Computers in biology and medicine, 2021. 131: p. 104255-104255.DOI: 10.1016/j.compbiomed.2021.104255

[16] Breiman, L., *Random Forests.* Machine learning, 2001. 45(1): p. 5-32.DOI: 10.1023/A:1010933404324