

A Single-Channel EEG-Based Approach to Detect Mild Cognitive Impairment via Speech-Evoked Brain Responses

Saleha Khatun¹, Student Member, IEEE, Bashir I. Morshed¹, Member, IEEE, and Gavin M. Bidelman¹

Abstract—Mild cognitive impairment (MCI) is the preliminary stage of dementia, which may lead to Alzheimer's disease (AD) in the elderly people. Therefore, early detection of MCI has the potential to minimize the risk of AD by ensuring the proper mental health care before it is too late. In this paper, we demonstrate a single-channel EEG-based MCI detection method, which is cost-effective and portable, and thus suitable for regular home-based patient monitoring. We collected the scalp EEG data from 23 subjects, while they were stimulated with five auditory speech signals. The cognitive state of the subjects was evaluated by the Montreal cognitive assessment test (MoCA). We extracted 590 features from the event-related potential (ERP) of the collected EEG signals, which included time and spectral domain characteristics of the response. The top 25 features, ranked by the random forest method, were used for classification models to identify subjects with MCI. Robustness of our model was tested using leave-one-out cross-validation while training the classifiers. Best results (leave-one-out cross-validation accuracy 87.9%, sensitivity 84.8%, specificity 95%, and F score 85%) were obtained using support vector machine (SVM) method with radial basis kernel (RBF) ($\sigma = 10/\text{cost} = 10^2$). Similar performances were also observed with logistic regression (LR), further validating the results. Our results suggest that single-channel EEG could provide a robust biomarker for early detection of MCI.

Index Terms—Electroencephalography, event-related potential, mild cognitive impairment, speech-evoked brain responses.

I. INTRODUCTION

MEMORY impairment due to aging is common. However, if the level of impairment progresses beyond what is expected in normal aging, mild cognitive impairment (MCI) can ensue [1]. MCI is a prodromal state of cognitive aging

Manuscript received June 9, 2018; revised October 14, 2018, December 14, 2018 and March 19, 2019; accepted April 6, 2019. Date of publication April 18, 2019; date of current version May 9, 2019. The work of S. Khatun was supported by the Herff College of Engineering, University of Memphis, through the Herff Graduate Fellowship. (Corresponding author: Saleha Khatun.)

S. Khatun and B. I. Morshed are with the Department of Electrical and Computer Engineering (EECE), The University of Memphis, Memphis, TN 38152 USA (e-mail: misti.buet@gmail.com; bmorshed@memphis.edu).

G. M. Bidelman is with the School of Communication Sciences and Disorders (CSD), The University of Memphis, Memphis, TN 38152 USA, and also with the Institute of Intelligent Systems (IIS), The University of Memphis, Memphis, TN 38152 USA (e-mail: gmbdlman@memphis.edu).

Digital Object Identifier 10.1109/TNSRE.2019.2911970

between changes deriving from natural/normal aging and dementia [2], [3]. Alzheimer's disease (AD) is the most prevalent dementia [4]–[7] among elderly population [1], [3] in many countries, and it is accompanied by progressively worsening memory, reasoning, and other aspects of cognition [7], [8]. As the MCI is a preliminary stage of cognitive impairment, most often it is not treated properly despite the fact that there is up to 54% chance that MCI may lead to AD or related dementias [8]. The cost of providing care for the AD patients in the US was \$200 billion in 2012 and it is estimated to grow to \$1.1 trillion per year by 2050 [9]. Therefore, preventing this disease is of great importance for better healthcare as well as for the national financial interest. Early MCI detection may play a critical role to enhance the management of the AD and dementia care as the root cause of this neurodegenerative disease is still unclear.

Detection and characterization of MCI is an active field of research. Various physiological data such as resting-state functional magnetic resonance imaging (rs-fMRI) [4], structural magnetic resonance imaging (sMRI), diffusion tensor imaging (DTI) [5], positron emission tomography (PET) [10], fluorodeoxyglucose positron emission tomography (FDG-PET) [11], cerebrospinal fluid (CSF) [11], [12], and magnetoencephalography (MEG) [13] are being investigated to study the effect of MCI and its underlying biomarkers. Although these physiological data provide multi-dimensional information about the brain, these methods are costly and also impractical in terms of portability. Low cost spontaneous speech data had been used by researchers to detect MCI or AD [14]–[17]. Relatively low-cost electroencephalography (EEG) is also being investigated to detect [2], [7], [18]–[23] and classify [21] MCI from other diseases that affect the cognitive state. Multi-channel EEG data have been used to characterize MCI or AD using EEG signals in various literature such as: (i) mismatch negativity (MMN) and auditory P300 component from 256-channel EEG [2], (ii) features extracted by recurrence quantification analysis (RQA) and cross recurrence quantification analysis (CRQA) from 14-channel EEG [21], (iii) spectral features extracted from 19-channel EEG [7], (iv) auditory P2 component computed from 64-channel EEG [22], (v) the auditory mismatch negativity (MMN) from 19-channel EEG [23], (vi) event-related potential (ERP) amplitude and latency from 256-channel

EEG [24], (vii) accuracy and response time during low and high working memory conditions of memory task using 32-channel EEG [19] etc. Non-ERP based multi-channel EEG approach has also been investigated to detect MCI, e.g. using data from a 19-channel EEG system [20]. Although multiple studies based on multi-channel EEG data are available in the literature, to the best of our knowledge, none of the studies attempted to use single-channel EEG data for the MCI detection and classification.

In this study, we have used single-channel EEG data from Fpz (near forehead) as this location is considered optimal for analyzing auditory evoked potential [25]–[27] to classify between MCI and normal functioning individuals using five sound stimuli for better sensitivity in MCI detection [24], [28]. This kind of system can be integrated into the wearable headband (i.e. MUSE™ headband with a mobile application), which will allow the assessment of cognitive strength. It is important to mention that EEG data may be affected by different types of artifacts such as ocular, muscle, electrode artifacts, which can be removed using either multi-channel [29] or single-channel [30], [31] EEG data. Therefore, studying EEG data obtained from a single-channel is not hindered by the artifacts that might degrade the quality of the EEG signal. Initial findings of this study were reported elsewhere [32]. In this paper, we provide our complete and comprehensive study results.

II. METHOD

We designed our experiments with five contrastive speech sounds along a vowel continuum (/a/ vs. /u/) [33]. This vowel continuum were played sequentially, and the subjects were requested to identify the sounds while their EEG data was collected. We obtained 590 features from the ERP extracted from the EEG signal, which included time domain and spectral domain characteristics from the windowed ERP data. The top 25 features ranked by the random forest algorithm were used in several classification models that are widely used in the literature e.g. Support Vector Machine (SVM), Logistic Regression (LR). We described the methodology of this study below, which is consisted of six steps: Subjects, Experimental Design, Data Collection, ERP Processing, Features Extraction and Ranking, and Classification.

A. Subjects

The physiological data used in this study was collected from an experiment, where twenty-three older adults (age ranges from 52-86 years; mean \pm standard deviation: 70.2 \pm 7.2 yrs) participated. All the subjects were strongly right-handed [28] with no known history of psychiatric or neurological illness. The cognitive state of the participants was evaluated by the well-established Montreal Cognitive Assessment (MoCA) test [34]. In this assessment, among all the participants, fifteen older adults (8 male, 7 female) were normal (MoCA score: 26-30), and eight participants' (4 male, 4 female) were found to have MCI (MoCA score: 22-25). It is worth mentioning that patients having Alzheimer's disease or severe dementia generate MoCA score within 11.4 to 21 [35]. The age of the

MCI group and the control group was 74.6 \pm 3.3 years, and 67.5 \pm 8.2 years, respectively ($t_{21} = 2.34$, $p = 0.03$). The chi-square test for the gender of the two groups gave a p-value of 0.81. The total years of formal education of the MCI, and normal group was 14.6 \pm 3.2 years, and 17.4 \pm 3.75 years, respectively ($t_{21} = -1.68$, $p = 0.11$). Participants in the study were compensated for their time and they gave their written consent under the protocol approved by the Baycrest Centre Ethics Committee (REB #06-31). The study was designed to observe the aging effect on the auditory system and to evaluate the cognitive state of the subjects. While recruiting the participants, exclusion criteria were based on age, musical training, handedness, and hearing loss [33].

B. Experiment Design

The phonetic continuum was generated by varying the first formant (F1) frequency within 430 Hz and 730 Hz over five equal steps. Fundamental (F0), second (F2), and third formant (F3) frequencies were kept the same for all five sound tokens. The values of F0, F2, and F3 were 100 Hz, 1090 Hz, and 350 Hz, respectively. The five-stepped vowel continuum (vw1-vw5) was constructed in a way so that each sound token of 100 ms would differ minimally acoustically, still be perceived categorically [36], [37]. We were aware that hearing loss due to aging may alter cortical auditory evoked potentials [25], and may influence the response; however, audiometric testing showed that hearing thresholds did not differ between groups at octave frequencies between 250 and 4000 Hz [33], which is well beyond the bandwidth of the stimuli.

C. Data Collection

Data collection technique and response evaluation were homogeneous to the studies reported in [25], [26], [32], [33], [36]. During EEG recording, participants went through 200 trials for each sound token. Each participant took part in the data collection twice. The experiment was conducted in an electroacoustically shielded chamber (Industrial Acoustics, Inc.). Subjects experienced the stimuli through earphones (ER-3A, Etymotic Research) in both ears at an intensity of 83 dB SPL. To eliminate electromagnetic stimulus artifact from corrupting neurophysiological responses, extended acoustic tubing (50 cm) was used [27], [33], [36]. Sound token came to subjects randomly, and they were requested to rapidly categorize them with a binary response ("u" or "a") by pressing specific buttons on the keyboard. In this study, their response does not matter as we were only interested to see how the ERP changes with different stimulus. However, we wanted them to focus while the stimulus was applied. This was done to make sure that the resulting ERP is produced only due to the applied stimulus. Each sound token was 100 ms duration with 10 ms of rise and fall time to reduce the spectral splatter [25]. After the participants' response, an inter-stimulus interval (ISI) followed randomly between 400 and 600 ms (20-ms steps, rectangular distribution) to avoid subjects anticipating subsequent

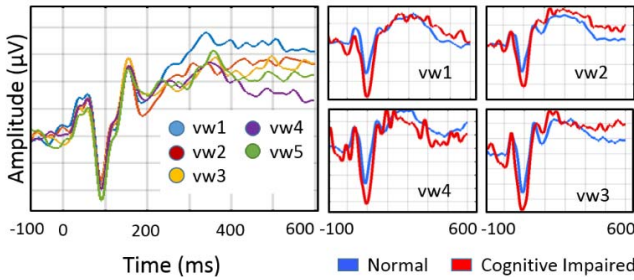


Fig. 1. Visualization of the ERP at different auditory stimuli. Left: individual response for the auditory stimulus vw1, vw2, vw3, vw4, and, vw5 of a normal subject; right: comparison between the ERPs of two subjects that belong to the Normal and the MCI group. ERP responses in the case of four auditory stimulus vw1, vw2, vw3, vw4 are visualized.

stimuli [38]. SynAmps RT EEG amplifiers (Compumedics Neuroscan, Charlotte, NC, USA) were used to capture EEG data. EEGs were recorded differentially between an electrode placed on the high forehead at the hairline referenced to linked mastoids. To record auditory evoked potentials from the cortical origin, this montage (\sim Fpz-A1/A2) is considered optimal [36], [39], [40]. Throughout the duration of the experiment, contact impedances were maintained below 3 k Ω , and the EEG signal was captured at 20 kHz sampling rate and then filtered by a band pass filter having passband within 0.05 Hz – 3500 Hz.

D. Event-Related Potential Processing

EEG data were processed using ERPLAB, an open source toolbox, which runs in the MATLAB environment [41]. An interval of 700 ms (100 ms pre-stimulus and 600 ms post-stimulus) constituted an EEG epoch as shown in Fig. 1 [32]. The left part of Fig. 1 represents five grand average ERP from a normal subject due to five auditory stimuli. The right part of Fig. 1 contains the comparison between the ERPs of Normal and MCI group. The pre-stimulus region was used for baseline correction, where the subtraction method [42] was used. Trials exceeded $\pm 50 \mu\text{V}$ were excluded from the analysis as they were probably contaminated by different artifacts such as eye-blinks, eye-movements etc as mentioned in [30], [31]. For each auditory stimulus, artifact-free epochs were used to calculate the grand average ERP. Finally, the grand average ERP is bandpass filtered from within (0-30) Hz because of a priori knowledge of the ERP bandwidths and the stimuli [26], [32], [33], [36], [40], [43]. It is to be noted that data from the different subjects do not go into the ERP calculation, rather each subject has its own ERP, which is calculated from the 200 trials of two sessions.

E. Features Extraction and Ranking

We extracted total 590 candidate features from the ERP prominent points and the time and spectral domain characteristics, and used top 25 features in the classification models, which were ranked by the random forest algorithm. In the cortical auditory evoked responses, prominent ERP points seem to have discriminatory power between normal

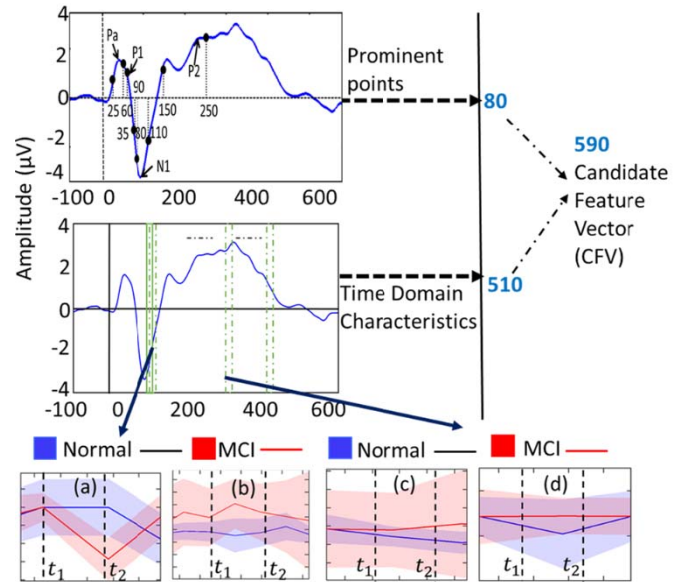


Fig. 2. Schematic of the feature extraction process: (a) and (c) are slopes and (b) and (d) are covariances for two different time windows from timestamp t_1 to t_2 .

and MCI stage [22] in the older adults. For that reason, we included the ERP prominent points in the candidate feature vector (CFV). The ERP prominent points such as Pa, P1, N1, and P2 were defined as the peak points between the intervals [25 ms - 35 ms], [60 ms - 80 ms], [90 ms - 110 ms], and [150 ms - 250 ms] respectively [32] as shown in Fig. 2. The peak amplitudes and their respective latencies of these prominent points, and the mean amplitudes of the intervals containing the prominent points were also included in the CFV because of their known importance in separating groups in the experiments involving evoked responses [44]. Relative powers in the EEG bands [i.e. delta (0-4)Hz / theta (4-7)Hz / alpha (8-12)Hz / beta (12-30)Hz] were also useful in classifying normal, mild cognitive impaired, and Alzheimer’s disease group [45]; thereby, included in the CFV. We used band-power.m function of MATLAB to compute the relative power. Total 16 features were calculated from the ERP prominent points from each stimulus, which resulted in total 80 feature points in the CFV. As the ERP changes rapidly with time within the window over which the stimulus is applied, it was important to track the variation of the time-domain characteristics in high-resolution. In order to accomplish that, we applied a 25ms window, with a 50% overlap through the entire ERP signal (i.e. 700ms EEG epoch) as depicted in Fig. 2.

The sliding window allowed us to observe the variation of the time domain properties, which has shown significance in classifying MCI and normal subject in earlier studies [33]. Total 107 time-domain characteristics such as signal statistics, correlation properties, entropies, etc. from each window were calculated using the opensource software package “HCTSAtool” [46], [47] written in Matlab. For the details of the extracted time domain characteristics, please see [46], [47]. The authors of [46], [47] also provided the documentation of the HCTSA tool [50] for the detail explanation of the time-domain characteristics. Among available properties in the

TABLE I
MODELS SELECTED FOR OBSERVING AGGREGATED SOUND FEATURES' PERFORMANCE

Model Name	Combination	Type	
SVMD _n C _m	SVM (degree = n, cost = m)	Polynomial kernel	degree = {2, 3, 4} cost = {10 ⁻² , 10 ⁻¹ , 1, 10 ¹ , 10 ² }
SVMSigma _n C _m	SVM (sigma = n, cost = m)	Radial basis kernel	sigma = {10 ⁻⁵ , 10 ⁻⁴ , 10 ⁻³ , 10 ⁻² , 10 ⁻¹ , 1, 10} cost = {10 ⁻² , 10 ⁻¹ , 1, 10 ¹ , 10 ² }
LRL1C _m	LR (regularization = 1, cost = m)	L1 regularization	cost = {10 ⁻² , 10 ⁻¹ , 1, 10 ¹ , 10 ² }
LRL2C _m	LR (regularization = 2, cost = m)	L2 regularization	cost = {10 ⁻² , 10 ⁻¹ , 1, 10 ¹ , 10 ² }

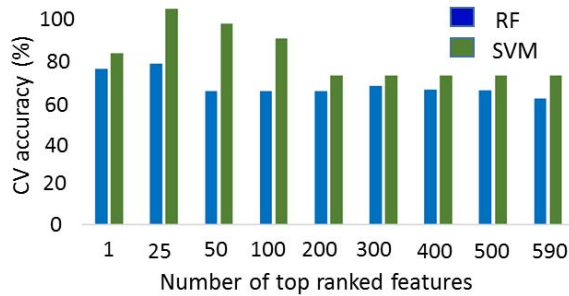


Fig. 3. Feature ranking and feature subset selection.

HCTSA tool, we have considered four categories such as: (i) Distribution, (ii) Correlation, (iii) Entropy and information theory, and (iv) Basics statistics and trend [other groups are not applicable for the EEG signal].

To calculate the variation of the time-domain characteristics over time, we computed the slope and the coefficient of variation (CV), which constituted two feature points at each time-stamp for each stimulus. The feature points that had intra-class similarity and inter-class variability (judged by visual inspection) were included in the CFV. For example, Fig. 2(a & b) show the slope and CV of one of the time domain characteristics (computed by the HCTSAtool known as “proportion of data within two-standard deviation of mean [46]–[48]”) for the normal and the MCI group, respectively. Here we observed that this feature point was different for two classes and the shaded regions depict that there was an intra-class similarity, but this feature was different between the classes (i.e. interclass variability), thereby selected as a feature point in the CFV. Similarly, the slope and CV shown in Fig. 2 (c&d) at a different timestamp also have intra-class similarity and inter-class variability, which fulfill their requirement to be in the CFV. Among all the time-domain feature points, 510 such feature points met the condition as mentioned above and were included in the CFV.

All 590 features of the CFV (i.e. 80 from the prominent points and 510 from the time-domain characteristics) were ranked by the random forest algorithm [49], and top 25 features were used in the classification models. The reason for taking top 25 features can be explained by Fig. 3. We obtained the highest cross-validation accuracy while using top 25 features and then it decreased at a higher number of features. The random forest algorithm was implemented using the “randomForest” package of R. In the random forest algorithm,

total 590 trees were constructed and at each split, 50 randomly sampled features were considered. To demonstrate the efficacy of the feature ranking process by random forest, we also demonstrated the results obtained by SVM at different number of features in Fig. 3. The SVM algorithm was implemented using “fitcsvm” module of MATLAB and radial basis kernel was used ($C = 10^2$, $\sigma = 10^3$). We observed that the SVM also gave the highest cross-validation accuracy while using top 25 features. Among 25 features, there are four features related with the prominent points of the ERP and rest of them are related with the statistical properties of the ERP extracted with the opensource software package “HCTSAtool” (Fulcher, Little, & Jones, 2013) written in Matlab. In the statistical properties related features, seven are related with entropy, twelve are related with the mean, and one is related with mean absolute deviation. Feature ranking helped us to reduce the number of features in final implementation and it helped to keep the computation cheap and also to reduce the probability of overfitting the data.

F. Classification

In this study, we used support vector machine (SVM), and logistic regression (LR) as they have already been used in the study related with EEG/ERP [49] data, and at the same time inference from these algorithms were computationally efficient. We used “fitcsvm” module of MATLAB to implement the support vector machine (SVM), and “liblinear” package of Weka 3.7 to implement the logistic regression (LR). In the SVM grid search, we implemented polynomial and radial basis kernel, and varied the cost, $C = \{10^{-2}, 10^{-1}, 1, 10^1, 10^2\}$.

The degree and sigma used in the case of polynomial and RBF kernel, respectively are reported in TABLE I. Similar approach was followed while LR grid search was implemented. For the detail grid search parameters, please see TABLE I.

It was necessary to find a model that balances between bias and variance to prevent overfitting and ensure generalization. Preventing overfitting is especially challenging if the sample size is small. To overcome this challenge, we followed the approach of [26] that means we used leave-one-out cross-validation set fixed for all classifiers (SVM, LR). In leave-one-out cross validation, the model is trained with $N-1$ samples and tested by the N^{th} sample. This process is repeated for all N samples. The experimental setup, EEG pre-processing, feature extraction, and classification method are summarized through a flowchart in Fig. 4.

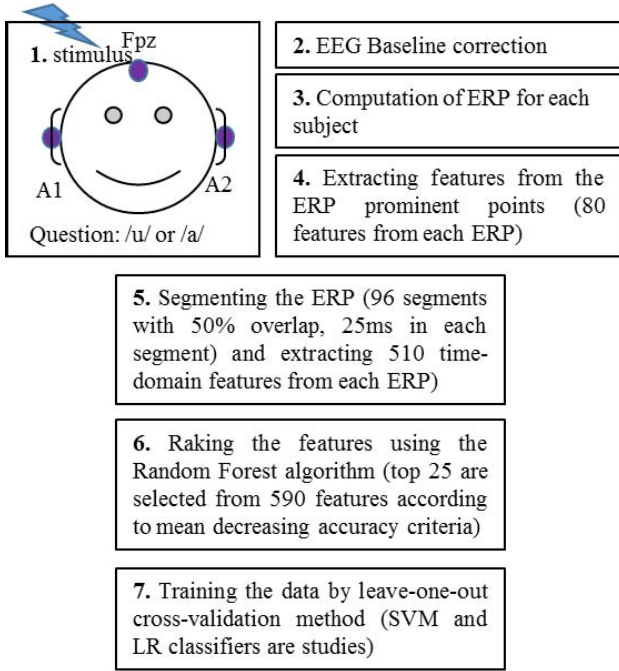


Fig. 4. Flow diagram of the algorithm.

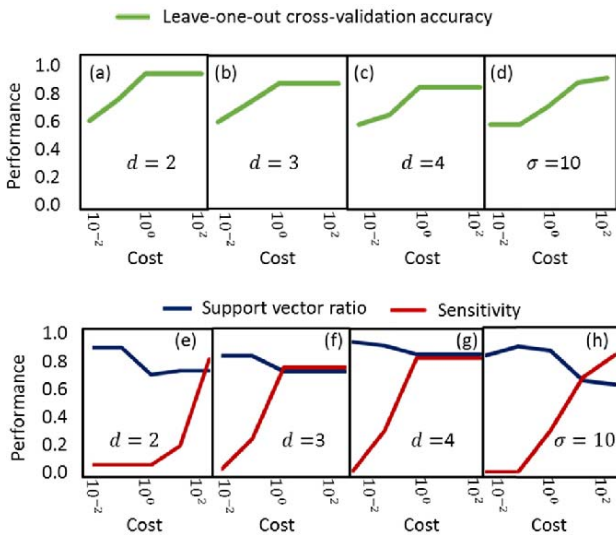


Fig. 5. Performance evaluation of SVM models: Leave-One-Out Cross Validation accuracy of Polynomial kernel (a) $d = 2$, (b) $d = 3$, (c) $d = 4$, Leave-One-Out Cross Validation accuracy of RBF kernel (d) $\sigma = 10$, Support vector ratio and Sensitivity of Polynomial kernel (e) $d = 2$, (f) $d = 3$, (g) $d = 4$, Support vector ratio and Sensitivity of RBF kernel (h) $\sigma = 10$.

III. RESULTS

To evaluate the classification models, we observed the key performance attributes such as leave-one-out cross-validation accuracy, sensitivity of the positive class (i.e. MCI group), and support vector ratio (SVR) (when applicable) in our SVM/LR grid search. In Fig. 5, we demonstrate the variation of the performance metrics in the case of the SVM for both the polynomial ($d = 2, 3$, and 4), and the RBF ($\sigma = 10$) kernels.

The top row of Fig. 5 shows the variation of the leave-one-out cross-validation accuracy with respect to the cost

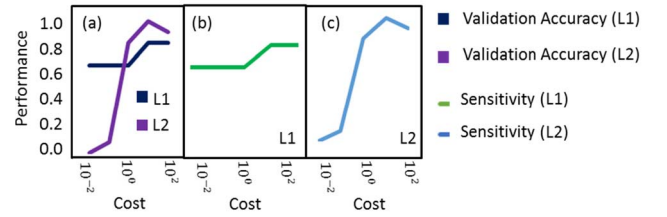


Fig. 6. Performance evaluation of LR models: (a) Sensitivity of L1 and L2 regularization, Leave-One-Out Cross-Validation accuracy of (b) L1 regularization, (c) L2 regularization.

TABLE II

PERFORMANCE OF SVMSIGMA₁₀C₁₀₀ MODEL SELECTED FOR OBSERVING AGGREGATED SOUND FEATURES' PERFORMANCE

Sen.	Spec.	Prec.	F	ROC
0.85	0.95	0.92	0.88	0.91

TABLE III
STIMULUS RANK

Stimulus	Rank
vw1	1
vw2	4
vw3	5
vw4	2
vw5	3

parameters, $C = \{10^{-2}, 10^{-1}, 1, 10^1, 10^2\}$. The results show that with the increase of C , both the leave-one-out cross-validation accuracy initially increases and reaches at the maximum for the optimal values of C . In Fig. 5(a), Fig. 5(b) and Fig. 5(c), the cross-validation accuracy reaches the maximum point and did not change later with the increment in C . The optimal C also provides the highest sensitivity and the lowest SVR as shown in the bottom row of Fig. 5 (exception for Fig. 5(e), the sensitivity increased after the optimal C). Lower support vector ratio ensures the model's stability with respect to the overfitting. The variation of the sensitivity, and, leave-one-out cross-validation accuracy with respect to the cost parameter, $C = \{10^{-2}, 10^{-1}, 1, 10^1, 10^2\}$ in the case of LR for both the L1 and L2 regularization is shown in Fig. 6. Likewise, SVM, the variation of the leave-one-out cross-validation accuracy, and sensitivity follow similar trend with C , and becomes maximum for the optimal value of C .

We observed that SVMSigma₁₀C₁₀₀ works best among all the models considered here. The performance attributes for the best classification models for different cases are summarized in TABLE II. These observations suggest that the selected feature vector is robust to multiple classification models in classifying between the normal and the MCI group.

In order to investigate the impact of different sounds used in the experiment, we used features collected from each sound in our classification models. The results obtained with the best models are summarized in Fig. 7. The sounds were ranked according to the leave-one-out cross-validation accuracy, sensitivity for MCI, and F score in TABLE III. We observed that

TABLE IV
LITERATURE SUMMARY

Paper	Source of Data	Focus	Classification Method	Accuracies (%)	Sensitivity (%)	Specificity (%)	Stimuli
Zhang <i>et al</i> [11].	sMRI, FDG-PET, CSF	Classification AD vs HC, MCI vs HC	Linear SVM	93.2, 76.4	93.0, 81.8	93.3, 66.0	N/A
Sui <i>et al.</i> [12]	fMRI, DTI	Classification MCI vs. HC	Multi kernel SVM	96.3	100	94.1	N/A
Suk <i>et al.</i> [10]	MRI, PET	Classification AD vs. HC, MCI vs. HC, AD vs. MCI	Deep learning	98.8, 90.7, 83.7	N/A	N/A	N/A
Ruzzoli <i>et al</i> [23]	Multi-ch. EEG	Classification MCI vs. HC	LR	N/A	76.9	73.3	Auditory
Kashefpoor <i>et al</i> [20]	Multi-ch. EEG	Classification HC vs MCI	Neurofuzzy system and KNN	88.9	100	83.3	N/A
Ahmed <i>et al.</i> [5]	sMRI, DTI	Classification AD vs. HC, MCI vs. HC, AD vs. MCI	Multiple kernel learning	90.2, 79.4, 76.6	82.9, 71.6, 65.5	94.6, 84.7, 81.3	N/A
Toth <i>et al.</i> [15]	Speech	Classification HC vs. MCI	Naïve Bayes, Linear SVM, Random Forest	71.4	79.2	61.1	Visual
This Study	Single ch. EEG	Classification HC vs MCI	SVM, LR	87.9	84.8	95.0	Auditory

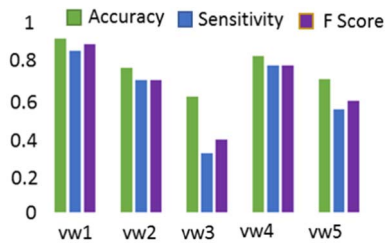


Fig. 7. Comparison of performances among auditory stimulus.

the clear sounds (vw1: /u/, vw4, and vw5: /a/) are perceived differently by the normal and impaired individuals and thereby are a better candidate to be used for future study related to MCI detection.

IV. DISCUSSION

Multi-channel EEG data with auditory stimulus is proven to work to classify between MCI and HC [20], [23]. From this study, we observed that even single-channel EEG data with auditory stimulus has the potential to classify between the MCI and HC. It was reported in previous studies that there are differential properties in the ERP responses between the MCI and HCs due to auditory stimulus. For example, Pekkonen *et al.*, (1994) [50] reported that dementia patients have faster decay in auditory sensory memory than the age-matched controls, which gets reflected in the mismatch negativity (MMN) (a component measured from ERP). In our previous study (Bidelman *et al.*, (2017) [33]), we observed that prefrontal dysfunction via efferent connections or abnormalities within the ascending auditory pathways may be related to MCI. Due to these reasons. We believe that the EEG signal collected from the frontal lobe (Fpz channel) has proven to be useful in the MCI detection. Additionally, we observed that the choice of auditory signal is also important. We found that ambiguous stimulus has less performance than non-ambiguous stimulus in the classification.

We tabularized our result with the existing work in the literature in TABLE IV. Expensive sMRI, FDG-PET, CSF, fMRI, DTI, PET data achieved more than 90% accuracy in discriminating MCI vs HC in multiple studies [5], [10]–[12].

Low- cost speech data also showed promising performance in this area [15]. Relatively low-cost multi-channel EEG data also demonstrated comparable accuracy ([20], [23]) obtained around 88.9% accuracy in differentiating MCI and HC). Our single-channel EEG-based MCI classifier, which obtained 87.9% accuracy, further demonstrates the significance of the EEG data in MCI study.

Our results are comparable to those observed with multi-channel EEG, and fMRI-based techniques in terms of classification accuracy, sensitivity, and specificity. This suggests that our single- channel based method may provide an alternative way of MCI detection, which is easy-to-use, and cost-effective.

The primary goal of this work is to investigate the use of the single-channel EEG data in detecting the early cognitive impairment to support the wearable technology. As we mentioned earlier in the discussion section, multi-channel EEG and other physiological data have proven to be a reliable source for the MCI detection; however, the idea of using the single-channel EEG data in determining complex neurological phenomena is relatively new. We believe that our work will motivate further study in this area.

In MRI technique, a patient needs to lie in the MRI scanner (a very large, strong magnet) and a radio wave is used to send signals to the part of the body of interest and receive them back. A computer attached to the scanner converts the returning signal into images. In PET technique, a patient swallows, inhales, or gets injected by radioactive tracer and then lies under a big PET scanner for generating PET image for diagnosis. fMRI uses the same basic principles as MRI. However, MRI scans anatomical structure whereas fMRI scans metabolic function. DTI, and sMRI also have similarities with MRI technique. High-density multichannel EEG setup requires a patient to sit in a quiet environment for a certain time determined by the study for data collection. All the techniques discussed so far here can only be employed in a hospital environment but not in wearable or portable devices. Conversely, the technique discussed in this paper can be implemented in a home environment.

V. CONCLUSION

In this study, we targeted to find a solution in the MCI detection using minimalistic, non-invasive, and low-cost approach –via ERP responses from the scalp EEG data. We selected existing algorithms such as SVM, LR and extracted features from time and frequency domain responses during speech processing responses reflected in the single-channel EEG data obtained from Fpz location. We observed that the top 25 ranked features ranked by the random forest method performed well with most of the classification models. Based on the best performances from the SVM model, we can predict MCI with 87.9% leave-one-out cross-validation accuracy, 84.8% sensitivity, and 95% specificity. In the future, this study can be expanded to real-time implementation of the system with the hardware-software implementation.

REFERENCES

- [1] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen, "Mild cognitive impairment: Clinical characterization and outcome," *Arch. Neurol.*, vol. 56, no. 3, pp. 303–308, 1999.
- [2] A. Tsolaki *et al.*, "Brain source localization of MMN and P300 ERPs in mild cognitive impairment and Alzheimer's disease: A high-density EEG approach," *Neurobiol. Aging*, vol. 55, pp. 190–201, Jul. 2017.
- [3] H. V. Vinters, "Emerging concepts in Alzheimer's disease," *Annu. Rev. Pathol., Mech. Disease*, vol. 10, pp. 291–319, Jan. 2015.
- [4] H. Ni *et al.*, "Network analysis in detection of early-stage mild cognitive impairment," *Phys. A, Stat. Mech. Its Appl.*, vol. 478, pp. 113–119, Jul. 2017.
- [5] O. B. Ahmed, J. Benois-Pineau, M. Allard, G. Catheline, and C. B. Amar, "Recognition of Alzheimer's disease and mild cognitive impairment with multimodal image-derived biomarkers and multiple kernel learning," *Neurocomputing*, vol. 220, pp. 98–110, Jan. 2017.
- [6] R. B. Knowles *et al.*, "Plaque-induced neurite abnormalities: Implications for disruption of neural networks in Alzheimer's disease," *Proc. Nat. Acad. Sci.*, vol. 96, no. 9, pp. 5274–5279, 1999.
- [7] L. R. Trambaiolli, N. Spolaôr, A. C. Lorena, R. Anghinah, and J. R. Sato, "Feature selection before EEG classification supports the diagnosis of Alzheimer's disease," *Clin. Neurophysiol.*, vol. 128, pp. 2058–2067, Oct. 2017.
- [8] C.-L. Tsai, M.-C. Pai, J. Ukropec, and B. Ukropcova, "The role of physical fitness in the neurocognitive performance of task switching in older persons with mild cognitive impairment," *J. Alzheimer's Disease*, vol. 53, no. 1, pp. 143–159, 2016.
- [9] D. De Venuto, V. F. Annese, and G. Mezzina, "Remote neuro-cognitive impairment sensing based on P300 spatio-temporal monitoring," *IEEE Sensors J.*, vol. 16, no. 23, pp. 8348–8356, Dec. 2016.
- [10] H. II Suk, S. W. Lee, and D. Shen, "Latent feature representation with stacked auto-encoder for AD/MCI diagnosis," *Brain Struct. Function*, vol. 220, no. 2, pp. 841–859, 2015.
- [11] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, and I. Alzheimer's, "Multimodal classification of Alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 55, no. 3, pp. 856–867, 2011.
- [12] J. Sui, R. Huster, Q. Yu, J. M. Segall, and V. D. Calhoun, "Function—Structure associations of the brain: Evidence from multimodal connectivity and covariance studies," *NeuroImage*, vol. 102, no. 15, pp. 11–23, 2014.
- [13] M. E. López *et al.*, "MEG beamformer-based reconstructions of functional networks in mild cognitive impairment," *Frontiers Aging Neurosci.*, vol. 25, no. 9, p. 107, 2017.
- [14] A. König *et al.*, "Use of speech analyses within a mobile application for the assessment of cognitive impairment in elderly people," *Current Alzheimer Res.*, vol. 15, no. 2, pp. 120–129, 2018.
- [15] L. Toth *et al.*, "A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech," *Current Alzheimer Res.*, vol. 15, no. 2, pp. 130–138, 2018.
- [16] G. Gosztolya *et al.*, "Detecting mild cognitive impairment from spontaneous speech by correlation-based phonetic feature selection," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, Apr. 2016, pp. 107–111.
- [17] J. Weiner, C. Herff, and T. Schultz, "Speech-based detection of Alzheimer's disease in conversational German," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, May 2016, pp. 1938–1942.
- [18] A. K. Kaiser, M. Doppelmayr, and B. Iglseider, "EEG beta 2 power as surrogate marker for memory impairment: A pilot study," *Int. Psychogeriatrics*, vol. 29, no. 9, pp. 1515–1523, 2017.
- [19] R. A. López Zunini *et al.*, "Event-related potentials elicited during working memory are altered in mild cognitive impairment," *Int. J. Psychophysiol.*, vol. 109, pp. 1–8, Nov. 2016.
- [20] M. Kashefpoor, H. Rabbani, and M. Berekatain, "Automatic diagnosis of mild cognitive impairment using electroencephalogram spectral features," *J. Med. Signals Sensors*, vol. 6, no. 1, p. 25, Jan. 2016.
- [21] 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.
- [22] J. J. Lister, A. L. H. Bush, R. Andel, C. Matthews, D. Morgan, and J. D. Edwards, "Cortical auditory evoked responses of older adults with and without probable mild cognitive impairment," *Clin. Neurophysiol.*, vol. 127, no. 2, pp. 1279–1287, Feb. 2016.
- [23] M. Ruzzoli, C. Pirulli, V. Mazza, C. Miniussi, and D. Brignani, "The mismatch negativity as an index of cognitive decline for the early detection of Alzheimer's disease," *Sci. Rep.*, vol. 6, Sep. 2016, Art. no. .33167.
- [24] C. D. Papadaniil, V. E. Kosmidou, A. Tsolaki, M. Tsolaki, I. Y. Kompatsiaris, and L. J. Hadjileontiadis, "Cognitive MMN and P300 in mild cognitive impairment and Alzheimer's disease: A high density EEG-3D vector field tomography approach," *Brain Res.*, vol. 1648, pp. 425–433, Oct. 2016.
- [25] G. M. Bidelman, J. W. Villafuerte, S. Moreno, and C. Alain, "Age-related changes in the subcortical—Cortical encoding and categorical perception of speech," *Neurobiol. Aging*, vol. 35, no. 11, pp. 2526–2540, 2014.
- [26] G. M. Bidelman and C. Alain, "Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception," *J. Neurosci.*, vol. 35, no. 3, pp. 1240–1249, Jan. 2015.
- [27] S. J. Aiken and T. W. Picton, "Envelope and spectral frequency-following responses to vowel sounds," *Hearing Res.*, vol. 245, nos. 1–2, pp. 35–47, Nov. 2008.
- [28] R. C. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.
- [29] R. Mahajan and B. Morshed, "Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 158–165, Jan. 2015.
- [30] S. Khatun, R. Mahajan, and B. I. Morshed, "Comparative study of wavelet-based unsupervised ocular artifact removal techniques for single-channel EEG data," *IEEE J. Transl. Eng. Health Med.*, vol. 4, no. 9, Mar. 2016, Art. no. 2000108.
- [31] S. Khatun, R. Mahajan, and B. I. Morshed, "Comparative analysis of wavelet based approaches for reliable removal of ocular artifacts from single channel EEG," in *Proc. IEEE Int. Conf. ElectroInf. Technol. (EIT)*, May 2015, pp. 335–340.
- [32] S. Khatun, B. I. Morshed, and G. M. Bidelman, "Single channel EEG time-frequency features to detect mild cognitive impairment," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2017, pp. 437–442.
- [33] G. M. Bidelman, J. E. Lowther, S. H. Tak, and C. Alain, "Mild cognitive impairment is characterized by deficient brainstem and cortical representations of speech," *J. Neurosci.*, vol. 37, no. 13, pp. 3610–3620, Mar. 2017.
- [34] Z. S. Nasreddine *et al.*, "The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment," *J. Amer. Geriatrics Soc.*, vol. 53, no. 4, pp. 695–699, 2005.
- [35] D. M. C. Doerflinger, "Mental status assessment in older adults: Montreal cognitive assessment: MoCA Version 7.1 (original version)," *Clinical Neuropsychol.*, vol. 25, no. 1, pp. 119–126, 2012.
- [36] G. M. Bidelman, S. Moreno, and C. Alain, "Tracing the emergence of categorical speech perception in the human auditory system," *Neuroimage*, vol. 79, no. 1, pp. 201–212, 2013.
- [37] D. B. Pisoni, "Auditory and phonetic memory codes in the discrimination of consonants and vowels," *Perception Psychophys.*, vol. 13, no. 2, pp. 253–260, Jun. 1973.
- [38] S. J. Luck, *An Introduction to the Event-related Potential Technique*. Cambridge, MA, USA: MIT Press, 2005.

- [39] A. Krishnan, J. T. Gandour, and G. M. Bidelman, "The effects of tone language experience on pitch processing in the brainstem," *J. Neurolinguistics*, vol. 23, no. 1, pp. 81–95, Jan. 2010.
- [40] G. Musacchia, D. Strait, and N. Kraus, "Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians," *Hearing Res.*, vol. 241, nos. 1–2, pp. 34–42, 2008.
- [41] J. Lopez-Calderon and S. J. Luck, "ERPLAB: An open-source toolbox for the analysis of event-related potentials," *Frontiers Hum. Neurosci.*, vol. 8, p. 213, Apr. 2014.
- [42] L. Hu, P. Xiao, Z. G. Zhang, A. Mouraux, and G. D. Iannetti, "Single-trial time—Frequency analysis of electrocortical signals: Baseline correction and beyond," *Neuroimage*, vol. 84, pp. 876–887, Jan. 2014.
- [43] G. M. Bidelman, "Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials," *J. Neurosci. Methods*, vol. 241, no. 15, pp. 94–100, Feb. 2015.
- [44] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components—A tutorial," *Neuroimage*, vol. 56, no. 2, pp. 814–825, 2011.
- [45] K. van der Hiele *et al.*, "EEG and MRI correlates of mild cognitive impairment and Alzheimer's disease," *Neurobiol. Aging*, vol. 28, no. 9, pp. 1322–1329, Sep. 2007.
- [46] B. D. Fulcher and N. S. Jones. (2016). "Automatic time-series phenotyping using massive feature extraction." [Online]. Available: <https://arxiv.org/abs/1612.05296>
- [47] B. D. Fulcher, M. A. Little, and N. S. Jones, "Highly comparative time-series analysis: The empirical structure of time series and their methods," *J. Roy. Soc. Interface*, vol. 83, no. 10, 2013, Art. no. 20130048.
- [48] Accessed: Apr. 23, 2019. [Online]. Available: <https://hctsa-users.gitbook.io/hctsa-manual/list-of-included-code-files>.
- [49] P. Bashivan, M. Yeasin, and G. M. Bidelman, "Single trial prediction of normal and excessive cognitive load through EEG feature fusion," in *Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB)*, Dec. 2016, pp. 1–5.
- [50] E. Pekkonen, V. Jousmäki, M. Könönen, K. Reinikainen, and J. Partanen, "Auditory sensory memory impairment in Alzheimer's disease: An event-related potential study," *Neuroreport*, vol. 5, no. 18, pp. 2537–2540, 1994.