Single Channel EEG Time-Frequency Features to Detect Mild Cognitive Impairment

Saleha Khatun
Electrical and Computer Engineering
The University of Memphis
Memphis, TN, USA
skhatun@memphis.edu

Bashir I. Morshed
Electrical and Computer Engineering
The University of Memphis
Memphis, TN, USA
bmorshed@memphis.edu

Gavin M. Bidelman
Institute for Intelligent Systems and Communication Sciences & Disorders
The University of Memphis
Memphis, TN, USA
gmbdlman@memphis.edu

***Abstract*— Detection of Mild Cognitive Impairment (MCI), a possible biomarker for the Alzheimer disease (AD), has a huge importance for the management of the elderly's healthcare. As the treatment of AD is very costly and the early discovery is beneficial, a low-cost, early detection mechanism is needed. In this study, Electroencephalography (EEG) data from seventeen subjects was used to determine if brain activity could be used to distinguish MCI with cognitively normal individuals and compare classification performance. Event-related brain potentials (ERP) were recorded in response to auditory speech stimuli. We extracted spectral and temporal features of the ERPs and built a MCI detection process with Support vector machine (SVM), Logistic Regression (LR), and Random Forest (RF). We have compared behavioral response against EEG based time domain features, frequency domain features, and top-ranked features of time-frequency domains. Four feature groups of our study demonstrate that the ranked time and frequency domain features of the EEG perform better than behavioral features and other EEG/ERP response metrics. Our results demonstrate the performance of the detection of MCI with a cross-validation accuracy of 87.9%, sensitivity 85%, specificity 90%, and F-score 94%. Ability to objectively and reliably detect MCI at early might lead to efficacious treatment of AD and related disorders.**

Keywords*— *Electroencephalography; Event related potential; Mild cognitive impairment; Support vector machine

I. Introduction

In 2012, the cost of providing care for the Alzheimer patients in the US was $200 billion and it is projected to grow to $1.1 trillion per year by 2050 [1]. Therefore, preventing this disease is of great importance for better healthcare as well as for national interest. Earlier detection of MCI is critical in order to manage Alzheimer's or dementia care. Memory impairment occurs naturally due to aging. However, if the level of impairment goes beyond the expected level, that condition is defined as mild cognitive impairment (MCI) [2]. MCI is the preliminary stage early cognitive decline, and if it is not treated properly, there is a 10-54% and 10-15% chance that MCI may lead to dementia, or Alzheimer's, respectively [3].

Detection of MCI is an active field of research in which people are investigating different physiological data such as structural magnetic resonance imaging (SMRI) and positron emission tomography (PET) data with deep learning [4], diffusion tensor imaging (DTI) and resting functional magnetic resonance imaging (fMRI) data with multi-kernel SVM [5]. Although the data used in these approaches give multi-dimensional information about the brain, the setups to obtain them can be impractical and costly. Thus, people are also investigating relatively low cost physiological data such as EEG to detect [6-9], and classify [10] MCI from other diseases affecting cognitive state. ERP extracted from multi-channel EEG data have been used to characterize MCI. The auditory P2 component computed from 64-channel EEG [6], the auditory mismatched negativity (MMN) from 19- channel EEG [7], ERP amplitude and latency from 256 channels EEG [8], and accuracy and response time during low and high working memory conditions of memory task using 32-channel EEE [9] have been investigated.

A non-ERP based direct EEG approach has also been carried out to detect MCI from 19-channel EEG data [10], which is the only study intended to classify the MCI vs. normal subject using multi-channel EEG to the best of our knowledge. Although multi-channel EEG has been studied rigorously, single-channel EEG based MCI detection and classification have not been investigated yet, which is important in order to facilitate minimalistic and portable devices for continuous patient monitoring. Single channel based disease detection and classification is of growing interest [11-13], and we also reported a single channel based artifact removal algorithm in [14-15]. In this study, we have used single channel EEG data to classify between MCI and normal functioning individuals.

We have included time and spectral domain features extracted from the Fpz channel of the EEG data (neural features), and subject related behavioral features in our study. Support vector machine (SVM), logistic regression (LR), and random forest (RF) have been investigated to classify MCI vs. normal people. We observed that neural features along with the SVM classifier provide the best performance such as 87.9% leave-one-out cross-validation accuracy with sensitivity 85%, specificity 90%, and F-score 94% while using SVM for classification. Our result is comparable with the multi-channel EEG and fMRI based techniques.

II. EXPERIMENT

A total of seventeen subjects took part in this study out of which ten persons (age [mean ± standard deviation]: 66.4 ± 7.5 years) were control and seven persons (age [mean ± standard deviation]: 76.6 ± 5.1 years) had MCI. All the subjects were right-handed and had similar levels of education. Both male and female had taken part in this study. A brief cognitive screening tool called Montreal Cognitive Assessment (MOCA) [16-18] was used to screen participants' cognitive function. MOCA scores <= 26 (out of 30) is indicative of MCI. All the participants were monolingual speakers of English [17].

Continuous EEG data was collected using SynAmps RT EEG amplifiers (Compumedics Neuroscan, Charlotte, NC, USA). EEG were recorded differentially between an electrode placed on the high forehead at the hairline referenced to linked mastoids. This montage is optimal for recording auditory evoked potentials [19-21]. Throughout the experiment, contact impedances were maintained below 3 kΩ. EEGs were digitized at a 20 kHz sampling rate and bandpass filtered online between 0.05 and 3500 Hz.

A synthetic 5-step vowel continuum (vw1-vw5) was constructed such that the first sound token sounded as /u/ and last one sounded as /a/ (details described in [8]). During ERP recording, the participants heard each sound token 200 times in a randomly ordered way and they were asked to categorize them with a binary response as quickly as possible ("u" or "a"). Each sound token was 100 ms in duration and included 10 ms of rise and fall time to reduce spectral platter [17]. After the participant's behavioral response, the interstimulus interval (ISI) was jittered randomly between 400 and 600 ms (20-ms steps, rectangular distribution) to avoid response anticipation.

III. MCI DETECTION PROCESS

The proposed MCI detection process includes following the steps a) ERP extraction, b) neural/behavioral feature extraction, c) feature ranking, d) classification (support vector machine/logistic regression/random forest). In this study, we only considered response to the sound "u" for the analysis.

A. Event Related Potential (ERP)

Trials which exceeded ± 50 µV were excluded as ocular artifacts [17]. EEG epochs were created from 100 ms before the stimulus to 600 ms after the stimulus. Epochs were baselined according to their pre-stimulus interval (-100 -0 ms). The grand average ERP was then computed as the average of all artifact-free epochs. Filtering was performed from 0-30 Hz. via ERPLAB [22]. The grand average ERPs of both groups are shown in Fig. 1.

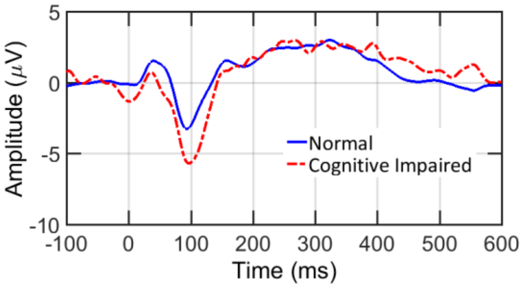

Fig. 1: Grand average ERP for sound "u" for both groups.

B. Neural Feature Extraction

Maximum amplitude and latency for the prominent waves of the ERPs (Pa, P1, N1, and P2) were measured at specific intervals. The intervals were taken as (25-35) ms, (60-80) ms, (90-110) ms, and (150-250) ms respectively [17] (Fig. 2). Also, the mean amplitudes of these specific intervals were taken on a basis that they might give meaningful information [23]. In addition to these time domain features, we measured the percentage of total power in delta (0-4) Hz, theta (4-8) Hz, alpha (8-16) Hz, and beta (16-30) Hz, respectively [24], to quantify salient spectral features of the neural responses. From the spectrogram of the ERP, the band powers of delta, theta, alpha, and beta and the power of the ERP were calculated. Then the percentages of the ERP power in the four bands were calculated.

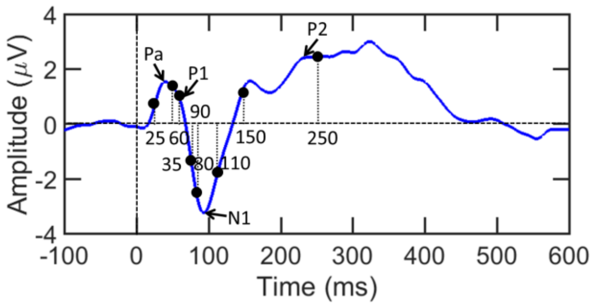

Fig. 2: Temporal neural features extracted from the ERPs

C. Behavioral Feature Extraction

For each subject, reaction times (RTs) for every trial were logged to quantify the speed of listeners' speech categorization. The acceptable limit for RTs considered here was 200 – 1500 ms [8-9]. For every subject, the RTs are T_1, T_2, T_3, T_4, T_5... T_k, respectively (Fig. 3) where k is the trial number. The maximum, minimum, mean and standard deviation of the RTs of the trials per subject were calculated and considered as behavioral features.8888

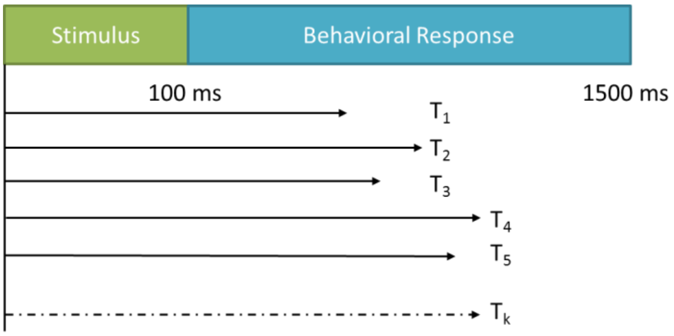

Fig. 3: Behavioral feature extraction process

D. Classification

We have investigated SVM, LR, and RF classifiers using four combinations of the extracted features: a) time domain features, b) frequency domain features, c) top-ranked features from time-frequency domain, and d) behavioral features. In the later part of our discussion, we will address the upper four approaches as approach a, b, c, and d. For feature ranking, we have used "rank.correlation" method from package "Fselector" of software R. The algorithm uses Spearman's rank correlation coefficient formula [25]. In addition to feature ranking from the R package, we have also conducted non-parametric statistics (Mann-Whitney U test/ Wilcoxon rank-sum test) to rank features. In the statistical test, the lower p-value the feature has, the better discrimination power it has (considered). The mean amplitude of the interval (90-110) ms and the amplitude of N1 have the p-value 0.021 and 0.026 respectively. The ranking result of the neural features are given in Table I.

TABLE I. FEATURE RANKING OF NEURAL FEATURES

Ranking	Feature Description
1	Mean of the amplitude of the interval (90-100) ms
2	Amplitude of N1
3	Latency of P1
4	Amplitude of P1
5	Mean of the amplitude of the interval (60-80) ms
6	Amplitude of Pa
7	Latency of Pa
8	% of total power in Theta Band
9	Mean of the amplitude of the interval (10-45) ms
10	% of total power in Beta Band
11	Latency of P2
12	Mean of the amplitude of the interval (150-200) ms
13	Latency of N1
14	Amplitude of P2
15	% of total power in Delta Band
16	% of total power in Alpha Band

In SVM, parameter search is performed using grid-search approach (kernel = {polynomial, rbf}, C = {0.01, 0.1, 1, 10, 100}, degree, D = {2, 3, 4}, $\gamma = \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$) with ten-fold cross validation. After choosing the top few parameters, classification has been performed and support vector ratio has been calculated with other performance metrics as a measure of the degree of overfitting [26]. L2-regularized logistic regression, a binary classifier, have been used in this study. Optimal parameter C was selected from set $\{10^{-2}, 10^{-1}, 1, 10, 100, 1000\}$. Random forest, an ensemble method, has been used in this study. The number of trees used in this method is 50. For the convenience of the readers, all the classification methods considered in this study and their acronyms are given in Table II. For the SVM implementation, MATLAB 2016a version has been used. For LR and RF implementation, Weka 3.8 has been used. Liblinear package has been used to implement LR.

TABLE II. CLASSIFICATION PARAMETERS

Model Name	Combination
CubC1	SVM (kernel = poly, C=1, D = 3)
CubC10	SVM (kernel = poly, C=10, D = 3)
CubC100	SVM (kernel = poly, C=100, D = 3)
QuadC1	SVM (kernel = poly, C=1, D = 4)
QuadC10	SVM (kernel = poly, C=10, D = 4)
QuadC100	SVM (kernel = poly, C=100, D = 4)
LRl2	Logistic Regression (l2 regularization, C = 1000)
RF	Random Forest

IV. RESULTS

Performance of the different feature approaches and classifiers which are considered in this study are presented in the Table III. We report sensitivity (SN), specificity (SP), and F score (F) for all the individual classes. We have also presented the ratio of support vector (Ratio of SV/ SVR), a performance measure for measuring the degree of overfitting, for the SVM approaches. Generally, the higher the SN, SP, and F, the better the performance. Whereas the lower the ratio of SV, the better the technique is (as it has low overfitting, so it has more generalization). For measuring performance, we have used ten-fold cross-validation. As MCI detection is a priority, we chose a method that has a higher sensitivity (SN) for MCI identification. Among all of the approaches, approach c) top-ranked features from time-frequency domains performs better than others in sensing MCI with a sensitivity 85% (SN0) while using CubC1. This model also generates 33.3% ratio of support vector, which denotes a better score to fight with overfitting. The SP for MCI detection is also high for this combination (SP0 = 90%) which depicts that this model has high true negative rate (Table III). F score is another metric which shows the balance between precision and recall of a model. It is observed from the Table III that the highest F score for MCI (F0) is 94% (while using CubC1) and for Normal (F1) is 96% (while using CubC10/ QuadC10).

Leave-one-out cross-validation (LOOCV) accuracy is calculated for all the feature approaches and models as LOOCV gives a view of overall performance of a classification approach. The result is given in Fig. 4. Approach c (combined top ranked) generates the highest LOOCV accuracy for almost all the models except RF. In RF, approach a (time domain neural) performs better than other approaches but the accuracy is 84.8% which is lower than the overall highest accuracy which is 90.9%. The overall highest accuracy has been generated by approach c by model CubC100 or QuadC10. In this study, our priority is to detect MCI and as this has been obtained by CubC1 model, this becomes our model of interest. The LOOCV accuracy of CubC1 model is 87.9%. We have compared the feature approaches' performances while CubC1 model is used in Fig. 5. Sensitivity, specificity, and F score are

TABLE III. OVERALL PERFORMANCES OF FEATURE APPROACHES AND CLASSFIERS

Feature Approach	**Combination**	**Ratio of SV**	**SN0**	**SN1**	**SP0**	**SP1**	**F0**	**F1**
Neural Feature From Time Series	Cubic(C=1)	24(72.7%)	55	50	50	55	65.7	57.5
	Cubic(C=10)	19(57.6%)	50	55	55	50	56	57.4
	Cubic(C=100)	19(57.6%)	55	70	70	55	69.2	70.4
	Quadratic(C=1)	23(69.7%)	45	80	80	45	68.6	68
	Quadratic(C=10)	18(54.5%)	50	60	60	50	64.3	64.6
	Quadratic(C=100)	18(54.5%)	80	65	65	80	74.3	77.9
	Logistic Regression(l2 regularized, C = 1000)	-	46.2	70	70	46.2	48	68.3
	Random Forest	-	76.9	90	90	76.9	80	87.8
Neural Feature From Spectral characteristics	Cubic(C=1)	21(63.6%)	20	65	65	20	54.2	56.7
	Cubic(C=10)	20(60.6%)	25	65	65	25	56.7	53.7
	Cubic(C=100)	17(51.5%)	55	65	65	55	75.2	63.3
	Quadratic(C=1)	21(63.6%)	35	40	40	35	52.2	54.4
	Quadratic(C=10)	21(63.6%)	55	50	50	55	59.3	61.7
	Quadratic(C=100)	21(63.6%)	60	45	45	60	62.6	73.9
	Logistic Regression(l2 regularized, C = 1000)	-	30.8	85	85	30.8	40	73.9
	Random Forest	-	30.8	85	85	30.8	40	73.9
Combined and Ranked from above	Cubic(C=1)	11(33.3%)	85	90	90	85	94	90
	Cubic(C=10)	8(24.2%)	75	90	90	75	93	96
	Cubic(C=100)	6(18.2%)	65	90	90	65	91	93
	Quadratic(C=1)	10(30.3%)	70	90	90	70	88	93
	Quadratic(C=10)	8(24.2%)	75	90	90	75	93	96
	Quadratic(C=100)	6(18.2%)	75	85	85	75	88	92
	Logistic Regression(l2 regularized, C = 1000)	-	61.5	75	75	61.5	61.5	75
	Random Forest	-	30.8	85	85	30.8	40	73.9
Behavioral Feature	Cubic(C=1)	24(72.7%)	50	65	65	50	63.3	68.3
	Cubic(C=10)	24(72.7%)	45	55	55	45	54	65.7
	Cubic(C=100)	27(81.8%)	35	75	75	35	63.9	62.3
	Quadratic(C=1)	18(54.5%)	50	55	55	50	58.8	60.4
	Quadratic(C=10)	17(51.5%)	35	75	75	35	65.3	62
	Quadratic(C=100)	19(57.6%)	50	75	75	50	66.7	69.6
	Logistic Regression(l2 regularized, C = 1000)	-	28.6	80	80	28.6	36.4	69.6
	Random Forest	-	21.4	45	45	21.4	21.4	45

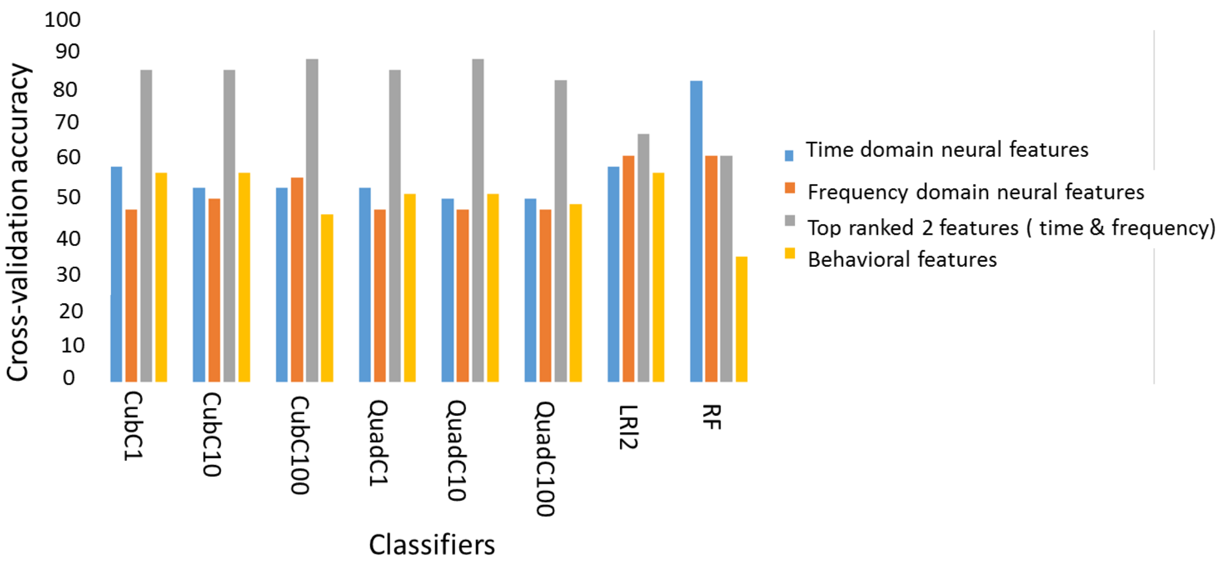

Fig. 4. Cross-Validation Accuracy of four feature appr

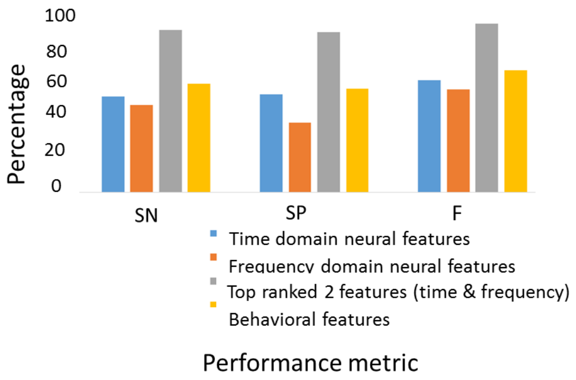

Fig. 5. Overall performaces of CubC1 model for four feature approaches

used in the comparison. In Fig. 5, the weighted sensitivity, specificity, and F score of all the CubC1 model have been presented. In terms of sensitivity, specificity, and F score, it is observed that the approach c (top-ranked features from time-frequency domains) dominates the others.

V. DISCUSSION

The classifiers used in this study are standard ERP based classification approaches. Although our data represent only a single channel (Fpz) of EEG from a small number subjects (Normal: 10, MCI: 7), we have found a model which can be trained with a small number of features to distinguish between MCI and cognitively normal individuals. Our model achieves a good SVR (support vector ratio) which shows that the proposed model can deal with overfitting. This study serves as a proof of concept for the later stage of our pilot study. We obtained accuracy of 87.9%, which is comparable with the accuracy of multi-channel EEG system (88.9%) [10], DTI, and fMRI based system (96.3%) [5]. In the future, we plan to focus on improving accuracy and sensitivity by exploring more features, and complex classifiers like deep neural networks.

Since EEG data are contaminated with artifacts like ocular, muscle artifacts, we have rejected the trials that exceed a certain threshold for artifacts. It might be possible to utilize artifact removal algorithm to preserve neuronal information during artifact zones [14-15] instead of the rejection of trials.

VI. CONCLUSIONS

In this study, our target was to detect MCI—a common precursor of Alzheimer's—via ERP responses from the scalp EEG. We selected existing algorithms such as SVM, LR, and RF and extracted features from neural (single channel EEG) and behavioral responses during speech processing tasks. We have observed that the approach utilizing top-ranked features from time-frequency domain of neural activity performs the best among other approaches considered here. Based on our classification accuracy (87.9%), we can predict that single channel based neuro-monitoring has vast potential to be used in efficient patient monitoring with ease and cost effectiveness. This work presents our early analysis from one stimulus-response. Future research directions include MCI detection performance with multiple sound tokens. This study can also be expanded to multichannel connectivity analysis, and real-time implementation of the system with hardware-software implementation.

REFERENCES

[1] D. D. Venuto, A. V. Francesco, and M. Giovanni, "Remote neuro-cognitive impairment sensing based on P300 spatio-temporal monitoring," in *IEEE Sensors Journal*, vol. 16, no. 23, pp.8348-8356, Dec. 2016.

[2] 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," in *Archives of Neurology*, vol. 56, no. 3, pp. 303–308, 1999.

[3] C. L. Tsai, M. C. Pai, J. Ulkropec, and B. Ukropcová, "The role of physical fitness in the neurocognitive performance of task switching in older persons with mild cognitive impairment," in *Journal of Alzheimer's Disease*, vol. 53, no. 1, pp.143-159, Jan. 2016.

[4] H. II Suk, S. W. Lee, D. Shen, and Alzheimer's Disease Neuroimaging Initiative, "Latent feature representation with stacked auto-encoder for AD/MCI diagnosis," in *Brain Structure and Function*, vol. 220, no. 2, pp. 841-859, Mar. 2015.

[5] 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," in *Neuroimage*, vol. 102, pp. 11-23, Nov 2014.

[6] J. J. Lister, A. L. H. Bush, R. Andel, C. Matthews, D. Morgan, J. D. Edwards, "Cortical auditory evoked responses of older adults with and without probable mild cognitive impairment," in *Clinical Neurophysiology*, vol. 127, no. 2, pp. 1279-1287, Feb. 2016.

[7] 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," in *Scientific Reports*, vol. 6, 2016.

[8] C. D. Papadaniil, V. E. Kosmidou, A. Tsolaki, M. Tsolaki, I. Kompatsiaris, L. J. Hadjileontiadis, "Cognitive MMN and P300 in mild cognitive impairment and Alzheimer's disease: A high density EEG-3D vector field tomography approach," in *Brain Research*, vol. 1648, pp. 425-433, Oct. 2016.

[9] R. A. L. Zunini, F. Knoefel, C. Lord, F. Dzuali, M. Breau, L. Sweet, R. Goubran, and V. Taler, "Event-related potentials elicited during working memory are altered in mild cognitive impairment," in *International Journal of Psychophysiology*, vol. 109, pp. 1-8, Nov. 2016.

[10] M. Kashefpoor, H. Rabbani, and M. Barekatain, "Automatic diagnosis of mild cognitive impairment using electroencephalogram spectral features," in *Journal of medical signals and sensors*, vol. 6, no. 1, pp. 25-33, Jan. 2016.

[11] S. P. Cho, J. Lee, H. D. Park, and K. J. Lee, "Detection of arousals in patients with respiratory sleep disorders using a single-channel EEG," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the* (pp. 2733-2735). IEEE.

[12] M. Bachmann, J. Lass, and H. Hinrikus, "Single channel EEG analysis for detection of depression," in *Biomedical Signal Processing and Control*, vol 31, pp. 391-397, Jan. 2017.

[13] M. Li, W. Chen, and T. Zhang, "Automatic epileptic EEG detection using DT-CWT based non-linear features," in *Biomedical Signal Processing and Control*, vol. 34, pp. 114-125, Apr. 2017.

[14] S. Khatun, R. Mahajan, and B. Morshed, "Comparative study of wavelet-based unsupervised ocular artifact removal techniques for

single-channel EEG data," in *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 4, pp. 1-8, 2016.

[15] 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 *Electro/Information Technology (EIT)* , pp. 335-340, May 2015.

[16] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Colline, J. L. Cummings, and H. Chertkow, "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," in *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695-699, Apr. 2005.

[17] G. M. Bidelman, J. W. Villafuerte, S. Moreno, and C. Alain, "Age-related changes in the subcortical-cortical encoding and categorical perception of speech," in *Neurobiology of aging*, vol. 35, no. 11, Nov. 2014.

[18] G. M. Bidelman, J. Lowther, S. Tak, and C. Alain, (under review). Mild cognitive impairment is characterized by deficient hierarchical speech coding between auditory brainstem and cortex.

[19] G. M. Bidelman, S. Moreno, and C. Alain, "Tracing the emergence of categorical speech perception in the human auditory system," in *Neuroimage*, vol. 79, pp. 201-212, Oct. 2013.

[20] A. Krishnan, J. T. Gandour, and G. M. Bidelman, "The effects of tone language experience on pitch processing in the brainstem," in *Journal of Linguistics*, vol. 23, no. 1, pp. 81-95, Jan. 2010.

[21] G. Musacchia, D. Strait, and N. Kraus, "Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians," in *Hearing Research*, vol. 241, no. 1, pp. 34-42, Jul. 2008.

[22] J. Lopez-Calderon, and S. J. Luck, "ERPLAB: an open-source toolbox for the analysis of event-related potentials," in *Frontiers in human neuroscience*, vol. 8, pp. 213, Apr. 2014.

[23] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K. Müller, "Single-trial analysis and classification of erp components – a tutorial," in *Neuroimage*, vol. 56, no. 2, pp. 814-825, May 2011.

[24] V. J. Samar, A. Bopardikar, R. Rao, and K. Swartz, "Wavelet analysis of neuroelectric waveforms: a conceptual tutorial," in *Brain and Language*, vol. 66, no. 1, pp. 7-60, Jan. 1999.

[25] C. Spearman, "The proof and measurement of association between two things," in *The American Journal of Psychology*, vol. 15, no. 1, pp. 72-101, Jan 1904.

[26] P. Bashivan, M. Yeasin, and G. M. Bidelman, "Single trial prediction of normal and excessive cognitive load through EEG feature fusion," in Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1-5, 2015.