Multivariate Models for Decoding Hearing Impairment using EEG Gamma-Band Power Spectral Density

Md Sultan Mahmud*, Faruk Ahmed*, Mohammed Yeasin*, Claude Alain † ‡ §, and Gavin M. Bidelman¶ ‖ **
Email: mmahmud, mfahmed, myeasin@memphis.edu, calain@research.baycrest.org, gmbdlman@memphis.edu
*Department of Electrical and Computer Engineering, University of Memphis, TN, USA
†Rotman Research Institute–Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada
‡University of Toronto, Department of Psychology, Toronto, Ontario, Canada
§University of Toronto, Institute of Medical Sciences, Toronto, Ontario, Canada
¶School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA
‖Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA
**University of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology, Memphis, TN, USA

Abstract*—Speech-in-noise (SIN) comprehension decreases with age, and these declines have been related to social isolation, depression, and dementia in the elderly. In this work, we build models to distinguish the normal hearing (NH) or mild hearing impairment (HI) using the different genres of machine learning. We compute band wise power spectral density (PSD) of source-derived EEGs as features in building models using support vector machine (SVM), k-nearest neighbors (KNN), and AdaBoost classifiers and compare their performance while listeners perceived clear or noise-degraded sounds. Combining all frequency bands features obtained from the whole-brain, the SVM registered the best performance. The group classification accuracy was found to be 94.90% [area under the curve (AUC) 94.75%; F1-score 95.00%] perceived the clear speech, and for noise-degraded speech perception, accuracy was found to be 92.52% (AUC 91.12%, and F1-score 93.00%). Remarkably, individual frequency band analysis on whole-brain data showed that γ frequency band segregated groups with a best accuracy of 96.78%, AUC 96.79% for clear speech data and noise-degraded speech data yielded slightly less accuracy of 93.62% with AUC 93.17% by using SVM. A separate analysis using the left hemisphere (LH) and right hemisphere (RH) data showed that the LH activity is a better predictor of groups compared to RH. These results are consistent with the dominance of LH in auditory-linguistic processing. Our results demonstrate that spectral features of the γ-band frequency could be used to differentiate NH and HI older adults in terms of their ability to process speech sounds. These findings would be useful to model attentional and listening assistive devices to amplify a more specific pitch than others.

I. Introduction

Hearing impairment (HI) is the top fifth leading disability worldwide [1], and the third most common chronic disease behind heart disease and arthritis [2], [3]. It is one of the key contributors to the growing disability problem in the United States [4]. Among other sensory organs, hearing is an essential function that allows people to communicate properly. In older adults, HI has been associated with poor cognitive health and contributes to social isolation as well as loneliness [5]. Age-related HI (i.e., presbycusis) is thought to occur due to a myriad of changes in both peripheral and central aspects of the auditory pathway [6].

Speech-in-noise (SIN) perception is problematic for older adults with and without hearing loss. The neurophysiological factors that influence SIN recognition are not well understood. Regardless of age, the audiogram –the conventional behavioral test of hearing– fails to always predict speech perception skills, especially in background noise [7]. Event-related potentials (ERPs), phase-locking value, power spectral density (PSD), and connectivity analysis are commonly used for understanding the brain functionality and identify normal and disorders status [8]. PSD is useful for investigating cognitive status [9] and complex processes such as working memory [10], attention [11], [12], and language processing [13]. In addition, PSD of electroencephalogram (EEG) signal offers a non-invasive means for various clinical diagnostics including epileptic seizure [14], [15] and Alzheimer's disease (AD) [16].

There is a growing interest in using machine learning (ML) techniques to process neuroimaging data such as EEG, magnetoencephalography (MEG), functional magnetic resonance imaging, and positron emission tomography. Evidence suggests that ML can help distinguish healthy and abnormal states of the brain. ML is a branch of artificial intelligence that *"learns a model"* from the past data to predict the future [17]. Support vector machine (SVM), k-nearest neighbors (KNN), and AdaBoost classifiers are extensively used as powerful tools to recognize subtle patterns, complex datasets, and classification in various fields, including the neurosciences.

EEG signals can be divided into different frequency bands (e.g., δ, θ, α, β, γ). Different bands are associated with different brain processes. Previous studies [18], [19] found that ERPs differed between NH and HI groups and between the clear and noise-degraded stimulus conditions. Bidelman et al. used graph theoretic and ML analyses on EEG data [20] to show that NH listeners could be distinguished from HI with 85.71% accuracy. Mahmud et al. decoded the hearing impairment by using the ERP features and a sliding window basis [18]. They found hearing loss could be identified with 81.50% accuracy as early as $\sim$ 50 ms after speech onset. Such model performances are inadequate for clinical practice. Beside, it is little known which frequency bands are associated with mild hearing loss.

In this research, we used EEG recordings to understand the neurodynamics of speech processing and spectral changes that are associated with age-related hearing impairment. EEGs

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

provide a means to detect subclinical and speech understand problems that may go undetected by standard audiometric testing. To explore these possibilities, we used source-level EEG and developed multivariate models to assess changes in the neural encoding of speech in older adults.

To improve model performance; we investigated how well spectral features (e.g., PSD) can segregate NH from HI; which frequency bands (e.g., θ, α, β, γ) could segregate the groups best; which brain hemisphere is dominant in speech processing in terms of spectral feature analysis? We present a comprehensive study in developing multivariate models distinguishing the NH and HI from band wise PSD. We conducted empirical analyses using all frequency bands, individual frequency bands both at the whole-brain level, and at hemispheric-specific using the different genres of machine learning.

Our main contributions are as follows:

- We developed multivariate models to examine age-related hearing impairment. It was found that NH and HI could be segregated above 95% accuracy from whole-brain data. We found that SVM achieved the best segregation of groups (e.g., NH and HI) over KNN and AdaBoost.
- We investigated using hemispheres data and found that LH measure is a better predictor of both the NH and HI.
- Finally, we compared different frequency bands and identified that the PSD of γ-band frequency provided the best performance among other frequency bands.

The rest of the paper is organized as follows. We review the related research in Section II and describes the methodology in Section III. Subsequently, Section IV presents the results, and Section V concludes the paper with the key findings.

II. Related Work

A plethora of research has used spectral features and ML to investigate normal and disordered states of brain activity. Many studies suggested that PSD is useful for exploring cognitive brain function, including Alzheimer's disease (AD), Schizophrenia, and Parkinson's disease (PD) [21], [22]. Related to aging and hearing assessment, Alain et al. [19] showed that older adults with mild hearing loss had higher cortical neuromagnetic evoked responses than the normal hearing. Bidelman et al. [20] demonstrated that graph-theoretic features of the brain's speech networks (estimated using EEG) could be used to differentiate adults with and without mild hearing losses.

Other researchers use the ERPs, band power, and other relevant statistical features (e.g., mean, variance, kurtosis) to classify healthy and disorder populations via SVM [23], [24], [25], KNN, and AdaBoost [8]. These classifiers are widely used for classification (e.g., face detection, speech recognition, text mining, etc.) and regression analysis. These are used as a prognostic (e.g., heart disease, cancer detection, epileptic seizure detection, and Alzheimer's classification) in neuroscience with better accuracy than traditional classifiers [23].

III. Materials and Methods

This work represents a new analysis of data that were originally collected by the Rotman Research Institute in Toronto [26], [20]. Following EEG preprocessing (outlined below), the power spectral density of the source level neural oscillations was used as input feature vectors to classifiers with the objective for segregating NH and HI older adults. The working flowchart of this analysis is presented in Fig. 1.

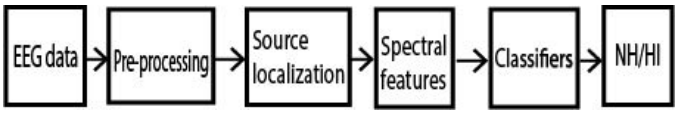


Fig. 1: Working flowchart.

A. Participants and Data Description

Thirty-two older adults (13 NH and 19 HI; aged 52 to 72 years) were recruited in the greater Toronto area to investigate the aging auditory system. All provided written informed consent in accordance with a protocol approved by the Baycrest Research Ethics Review Board. None had reported neurological or psychiatric diseases. Puretone audiometry was conducted at octaves frequencies in the range 250 to 8000 Hz. The listeners were grouped into two cohorts based on their hearing thresholds (Fig. 3). As standard hearing threshold [27], participants who had the hearing thresholds 25 dB or better were considered as NH, and those who had poorer than 25 dB were included into the HI cohort.

Neural activity of the brain was recorded at standard 10-20 locations on the scalp surface using 32 Ag/AgCl electrodes. EEGs were recorded during a rapid speech detection task. Three consonant-vowel (CV) speech tokens */ba/*, */pa/*, */ta/* were used as stimuli during the EEG recording. These tokens were presented in a pseudo-random manner, where */ba/*, */pa/* were presented frequently and */ta/* infrequently. The stimulus set comprised 3000 */ba/*, 3000 */pa/* and 210 */ta/*. The duration of each token was 100 ms and interstimulus interval 250 ms. A total of 6210 CVs was presented in a clear (i.e., no noise) and noise-degraded block in which speech sounds were presented concurrent with +10 dB SNR of speech babble noise. Participants were asked to respond via a button press on the computer when they detected the infrequent targets (i.e., */ta/ tokens*). The reaction time (RT) and percentage of correct detections (%) were recorded. Because of their limited number, infrequent (i.e., */ta/*) responses were not analyzed. The stimulus presentation scenario is depicted in Fig. 2. The details data description is given in [20].

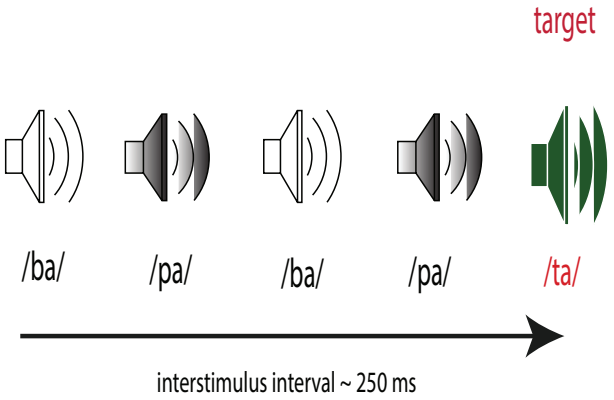


Fig. 2: Stimulus presentation scenario.

B. Signal Pre-processing and Source localization

EEGs were digitized using Neuroscan SynAmps RT amplifiers at a 20 kHz sampling rate. Subsequent pre-processing

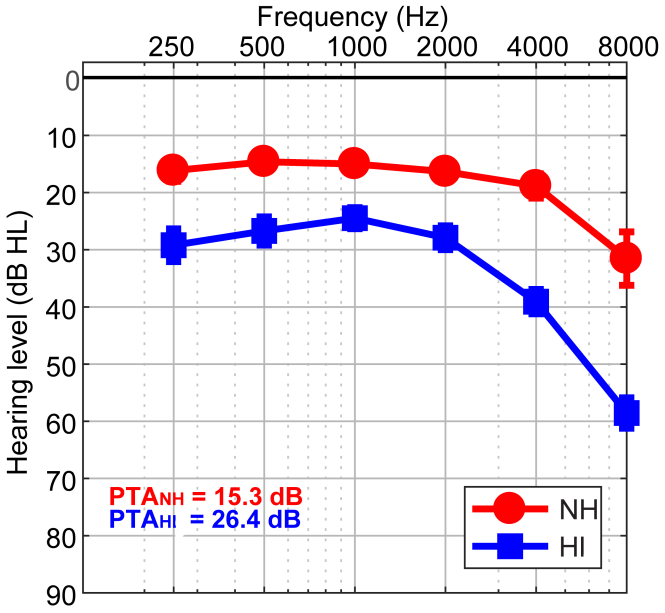


Fig. 3: Behavioral results (audiogram of NH and HI). NH, normal hearing group; HI, hearing impaired group; PTA, puretone average threshold.

was performed in BESA Research (v6.1) and Brainstorm software [28]. The ocular artifacts (e.g., eye-blinks) were corrected in the continuous EEG using principal component analysis and then filtered (1-100 Hz; notched filtered 60 Hz). EEGs were epoched into the single-trial (-10 to 200 ms) per condition (clear and noise). We localized the sources of the scalp recorded EEG data by performing a distributed source analysis. We conducted source localization by using a realistic boundary element head model (BEM) volume conductor and standard low-resolution brain electromagnetic tomography (sLORETA) as an inverse solution with Brainstorm [28] environments. A BEM model has less spatial errors than other existing head models (e.g., concentric spherical head model). We used the default setting parameters of Brainstorm's tool (SNR=3.00, regularization noise covariance = 0.1). The localization error of sLORETA for 32 channels is $1.5\times$ less accurate than 64 channels [29]. However, the mean localization error of sLORETA for 32 is estimated at 1.45 mm [30]. Once the source localization was done, then from each single trial sLORETA volume, we extracted the time-courses within the 68 regions of interest (ROIs) across the left and right hemispheres defined by the Desikan-Killiany (DK) atlas [31] (LH: 34 ROIs and RH: 34 ROIs). Single-trial data were baseline corrected to the epoch's pre-stimulus interval then used to compute the PSD.

C. Spectral Feature Extraction

Our dataset comprised $\sim$ 6000 trials per subject and condition (e.g., clear and noise). Previous studies on this dataset [18], empirically found that averaged over 100 trials ERPs provided the best group classification (accuracy 81.5%). Thus, we measured ERPs averaged over randomly chosen 100 trials without replacement. We then computed the band average PSD for each ROI across the entire epoch by using 'pwelch' Matlab function [32] and extracted in several frequency bands including of theta (θ: 5-8 Hz), alpha (α: 9-13 Hz), beta (β: 14-30 Hz), and gamma (γ: 31-45 Hz) bands [33]. We computed these spectral features for each participant, group, and condition (clear and noise). These spectral features were used as input to the classifiers. The data were z-score normalized before submitting to the classifiers to ensure all features were on a common scale range [34].

D. Classifiers (SVM, KNN, and AdaBoost)

In neuroscience data, a goal is to find relationships and patterns of data that can be useful for monitoring the progression of the disease. A benefit of classical machine learning techniques is that they reveal such relationships with reasonable performance with small sample sizes, which is common in human neuroimaging studies. Here, we used support vector machines (SVMs), k-nearest neighbor (KNN,) and AdaBoost classifiers to investigate distinguishing listeners' group membership (e.g., NH and HI) from spectral attributes of their EEG. The machine learning models learned from the training data that comprised of features (e.g., PSD) and corresponding class labels (e.g., NH and HI). Once the model learned, then the unseen test data was used for prediction. We randomly split the data into training and test sets 80%, and 20%, respectively [14], [35], [36]. Classification performance (accuracy, F1-score, and area under the curve (AUC)) was calculated by using the standard formulas [37] from the predicted and true class labels. AUC demonstrates the degree to which a model is capable of distinguishing between the classes. An excellent model has AUC close to 1, meaning it has a good separability. On the other hand, a bad model has AUC near to 0, meaning it has no separability.

SVM: SVM classifier performance is greatly affected by the kernel function, among other factors. The tunable parameters (e.g., C, γ) also reflect the performance [38]. Determining which kernel provides better performance depends on the nature of the problem. As such, a grid search approached was performed to find the optimal kernel, C, and γ values. During the training phase, we fine-tuned the C, and γ parameters to find the optimal values; so that the classifier could accurately distinguish HI from the test data. In the grid search approach, we used five-folds cross-validation [39], kernels = 'RBF', and fine-tuned 20 different values of (C, γ) in the following range for the C = [2^{-1} to 2^{10}] , and γ = [2^{-5} to 2^{3}]. The SVM learned the support vectors from the training data that comprise of the spectral attributes (e.g., PSD) and class labels (e.g., NH and HI). The resulting hyperplanes were fixed with maximum margin (e.g., maximum separation between the two classes) and used for predicting the unseen test data (only by providing the attributes but no class labels). We selected the best model then predicted the group membership from the unseen features attributes of the test data.

KNN: We also used a parameter optimized KNN classifier [40]. The value of k plays an important role in the performance of the KNN classifier [41]. During the training phase, we tuned the value of k from 1 to 30 for achieving maximum accuracy. In our data, we found an optimal value of $k = 5$ that provided the best classification accuracy.

AdaBoost: We used AdaBoost classifier with a base estimator Decision Tree Classifier. During the training phase, we set the hyperparameters, the number of estimator = 50, and algorithm ='SMME.R', which has typically less error and is faster [8]. Another hyper-parameter (e.g., learning rate) of

TABLE I: SVM, KNN, and AdaBoost classifiers' performance metrics (%) for distinguishing hearing status.

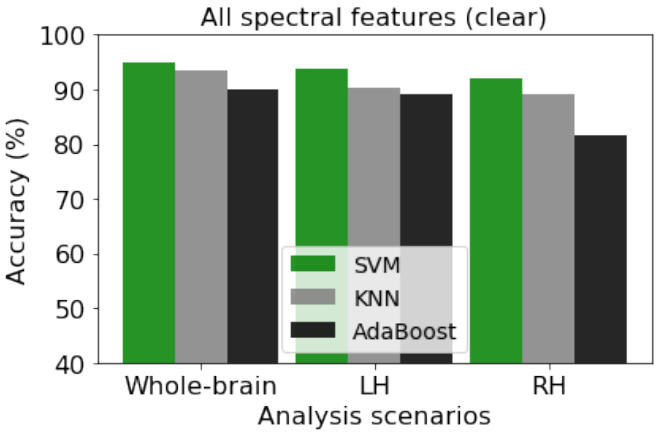
Stimulus	Classifiers name	Average measure(%)	Whole-brain features	LH's features	RH's features
Clear	SVM	Accuracy	94.90	93.83	91.95
		AUC	94.75	93.41	91.61
		F1-score	95.00	94.00	92.00
		Precision	95.00	94.00	92.00
		Recall	95.00	94.00	92.00
	KNN	Accuracy	93.56	90.34	91.95
		AUC	93.18	90.26	91.61
		F1-score	94.00	90.00	89.00
		Precision	94.00	90.00	89.00
		Recall	94.00	90.00	89.00
	Adaboost	Accuracy	90.08	89.00	81.50
		AUC	89.52	87.94	80.48
		F1-score	90.00	89.00	81.00
		Precision	90.00	89.00	81.00
		Recall	90.00	89.00	81.00
Noise	SVM	Accuracy	92.52	91.96	89.75
		AUC	91.12	90.86	88.29
		F1-score	93.00	92.00	90.00
		Precision	93.00	92.00	90.00
		Recall	93.00	92.00	90.00
	KNN	Accuracy	92.16	91.68	87.81
		AUC	90.89	91.07	87.98
		F1-score	92.00	92.00	88.00
		Precision	92.00	92.00	88.00
		Recall	92.00	92.00	88.00
	Adaboost	Accuracy	84.48	85.59	81.44
		AUC	84.28	83.91	79.28
		F1-score	85.00	86.00	81.00
		Precision	85.00	86.00	81.00
		Recall	84.00	86.00	86.00

TABLE II: SVM performance metrics (%) for distinguishing hearing status by using the γ frequency band spectral feature.

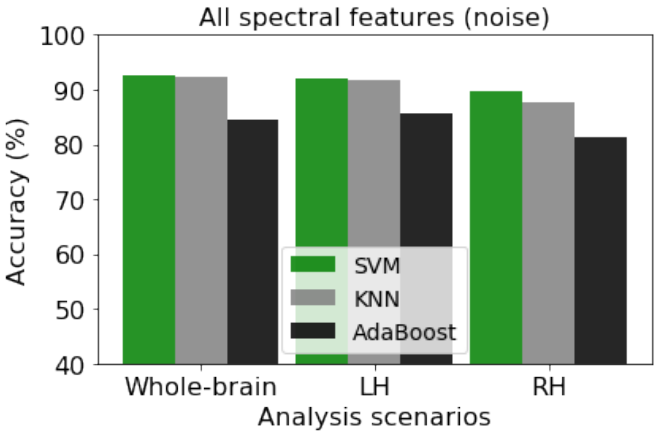
Stimulus	Classifiers name	Average measure(%)	Whole-brain features	LH's features	RH's features
Clear	SVM	Accuracy	96.78	93.03	91.96
		AUC	96.79	92.87	91.53
		F1-score	97.00	93.00	92.00
		Precision	97.00	93.00	92.00
		Recall	97.00	93.00	92.00
Noise	SVM	Accuracy	93.62	92.24	90.02
		AUC	93.17	91.41	89.34
		F1-score	94.00	92.00	90.00
		Precision	94.00	92.00	90.00
		Recall	94.00	92.00	90.00

the AdaBoost classifier was tuned to achieve better accuracy (range 0.1 to 1.0; step size = 0.05). Empirically, we found a learning rate of 1.0 that resulted in the best group segregation.

We submitted the PSD attributes and corresponding class labels to the three classifiers individually. We separately analyzed group classification using the whole-brain data, hemisphere-specific data (e.g., LH and RH hemisphere), and per frequency band wise (e.g., θ, α, β, γ). The three classifiers were trained with the same training data set and were used to predict group membership with the same unseen test data set. The classifiers' performance metrics were calculated from the predicted class (obtained from unseen test data) and true class labels. Classifier performance in different classification scenarios (e.g., whole-brain, LH, and RH) are presented in TABLE I, II, and illustrated the accuracies in Fig. 4, 5, 6, and 7.



(a) Clear speech detection.



(b) Noise-degraded speech detection.

Fig. 4: Classifier accuracies for distinguishing NH and HI using PSD features from speech-evoked EEGs using whole-brain vs. hemispheres-specific (LH and RH) data.

IV. Results

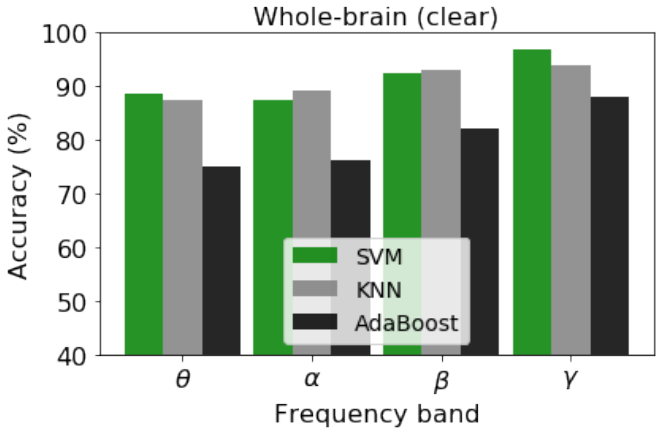
A. Performance Evaluation for Combining all Spectral Features

The accuracy of three classifiers obtained using combined all spectral features is presented in bar charts in Fig. 4. Accuracy and other performance metrics of three classifiers are reported in TABLE I.

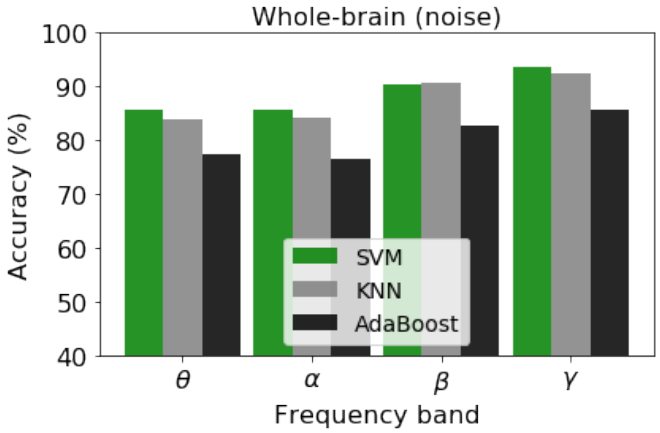
For clear speech, the SVM classifier yielded a maximum classification accuracy of 94.90%, with AUC 94.75%, F-1 score, precision, and recall 95.00% by using the whole-brain data. However, KNN (accuracy 93.56%, AUC 93.18%, F1-score, precision, and recall 94.00%), and AdaBoost (accuracy 90.08%, AUC 89.52%, F1-score, precision, and recall 90.00%) showed almost similar but a slightly less than SVM.

For noise-degraded speech perception, the SVM classifier revealed a maximum accuracy of 92.52%, AUC 91.12%, F1-score, precision, and recall 93.00% by using the whole-brain data. The KNN showed almost similar performance (accuracy 92.16%, AUC 90.89, F1-score, precision, and recall 92.00%) to the SVM, but AdaBoost accuracy was decreased by 8% (accuracy 84.48%, AUC 84.28%, F1-score, precision 85.00%, and recall 84.00%).

We also separately examined hemispheres-specific data (LH and RH alone) in distinguishing clear and noise-degraded speech. For clear speech perception, the LH data yielded the

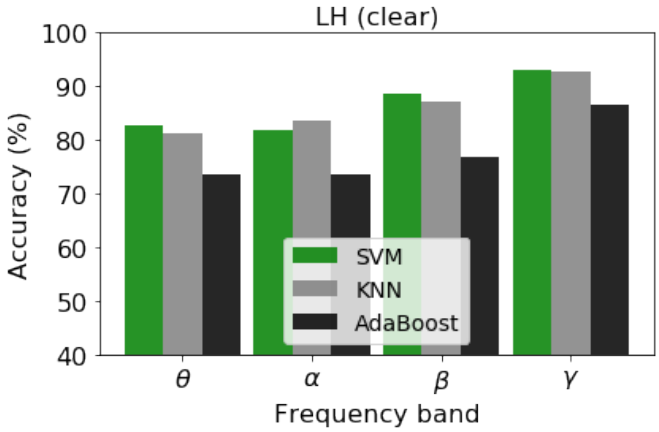


(a) Clear speech detection.

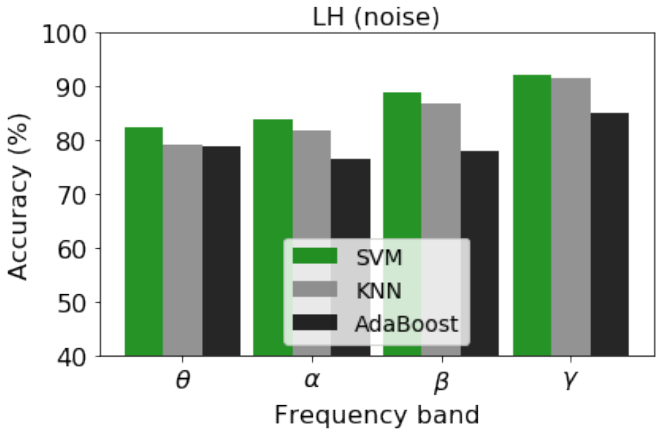


(b) Noise-degraded speech detection.

Fig. 5: Classifier accuracies per frequency band for distinguishing NH and HI using PSD from whole- brain data.



(a) Clear speech detection.



(b) Noise-degraded speech detection.

Fig. 6: Classifier accuracies per frequency band for distinguishing NH and HI using PSD from LH data.

classification accuracy of 93.83%, AUC 93.41%, F1-score, precision, and recall 94.00% by using the SVM, and the KNN showed (accuracy 90.34%, AUC 90.26%, F1-score, precision, and recall 90.00%) close to SVM. However, the AdaBoost showed slightly less accurate (accuracy 90.34%, AUC 90.26%, F1-score, precision, and recall 90.00%) than KNN.

For noise-degraded speech perception, classification was less accurate as compared to clear speech perception. By using the LH data alone, SVM showed accuracy 91.96%, AUC 90.86%, F1-score, precision, and recall 92.00%; whereas KNN, and AdaBoost were less accurate compared to the SVM. Results are reported in TABLE I. The RH data showed the lowest accuracy among all analysis scenarios with all classifiers (reported in TABLE I).

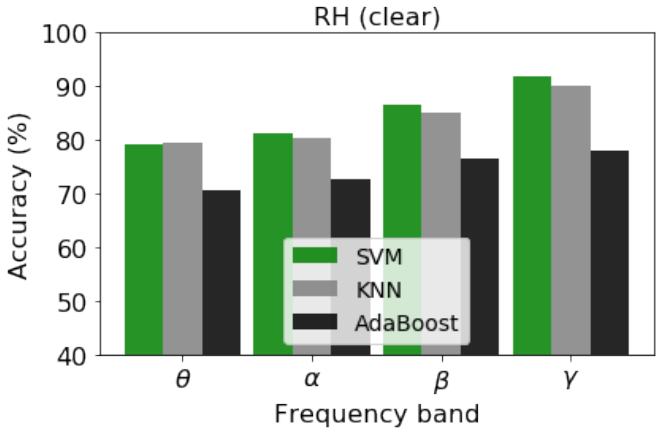
For both clear and noise-degraded speech perception, it was noticeable that all classifiers showed maximum accuracy and AUC by using whole-brain data. The LH features showed slightly lower accuracy and AUC score (reported in TABLE I) than the whole-brain. The RH is less robust than LH but still well above chance level (i.e., 50%). Notably, the SVM provided the best classification performance among other classifiers (KNN and AdaBoost) for all classification scenarios (whole-brain, LH, and RH).

B. Performance Evaluation for Different Frequency Bands

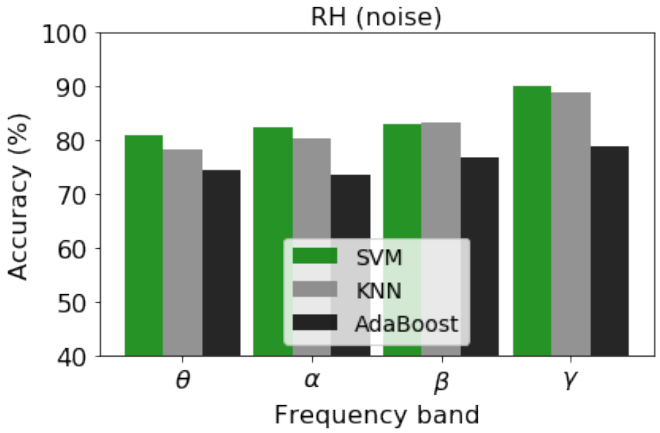
We also examined the power of individual frequency bands to segregate NH vs. HI. We used the individual frequency band spectral features (e.g., PSD) as input to the classifiers with the corresponding class labels (e.g., NH and HI). Group classification was performed again via three different classifiers (SVM, KNN, and AdaBoost). The accuracies of the classifiers in different analysis scenarios and conditions are presented in Fig. 5, 6, and 7. Fig. 5 represents the accuracy for the whole-brain, Fig. 6 for the LH, and Fig. 7 for the RH.

For clear-speech perception, when we used the PSD features of the θ frequency band from the whole-brain data; SVM, KNN, and AdaBoost classifiers showed the accuracies of 88.73%, 87.39%, and 75.06%, respectively. The classifiers were conducted on the LH data that yielded accuracies of 82.57% (SVM), 81.26 % (KNN), and 73.45% (AdaBoost). These results showed that the LH data showed slightly less accurate than the whole-brain data. RH data yielded the lowest accuracies for all classifiers. The RH data showed classification accuracies of 79.08%, 79.35%, and 70.50%, respectively, for SVM, KNN, and AdaBoost.

For noise-degraded speech perception, with the same frequency band feature (e.g., PSD of θ frequency band), classifiers' performance was less accurate than clear speech perception. Whole-brain data revealed accuracies of 85.59%, 83.93%, and 77.28% via SVM, KNN, and AdaBoost classifiers, re-



(a) Clear speech detection.



(b) Noise-degraded speech detection.

Fig. 7: Classifier accuracies per frequency band for distinguishing NH and HI using PSD from RH data.

spectively. The group classification was less accurate than whole-brain data while using LH data alone, accuracy of 82.54% (SVM), 79.22% (KNN), and 78.94% (AdaBoost). RH data showed the least accuracies (Fig.7b) among the three classification scenarios.

The spectral features of α and β frequency band yielded a similar group classification accuracy (maximum 90 and minimum 76 %) as θ band for clear and noise-degraded speech perception. Accuracies of the different classifiers are illustrated in Fig. 5, 6, and 7 corresponding to the individual frequency bands.

Interestingly, group segregation was improved by using the γ frequency band. Whole-brain data yielded the accuracies of 96.78%, 93.83%, and 87.93%, respectively for SVM, KNN, and AdaBoost for clear speech perception; whereas in noise-degraded speech perception, it was less robust accuracies of 93.62% (SVM), 92.35% (KNN) and 85.59% (AdaBoost). The group classification was slightly less accurate than the whole-brain while using the LH data alone. LH data yielded accuracies of 93.02% (SVM), 92.76% (KNN), and 86.59% (AdaBoost) for clear speech perception; whereas in noise-degraded speech perception, it was 92.24% (SVM), 91.68% (KNN), and 85.04% (AdaBoost). However, RH data showed lower accuracy than the LH data but still an accuracy above 85.00%.

Among these three classifiers, the SVM classifier showed the best group segregation across the board. From the individual frequency analysis, the γ frequency band provided the best group segregation via SVM. Thus, the other important performance metrics of the SVM classifier is reported corresponding to the best group segregation accuracy in TABLE II. Remarkably, all classifiers (SVM, KNN, and AdaBoost) revealed that classification was most robust while using the spectral feature of the γ frequency band for clear speech as well as for noise-degraded speech perception.

Our findings show whole-brain data yielded maximum group segregation while using combined all features as well as the individual γ frequency band spectral feature. However, classification using LH data alone was slightly less than the whole-brain features but better than the RH. Our results show an improvement of the classification accuracy by 11% over previous studies, [18], [20]. Moreover, our findings support the previous neuroimaging studies that higher frequency (e.g., γ) band are associated with auditory perception [42] and perceptual accuracy [43].

V. Conclusion

We developed a robust and efficient computational framework to differentiate NH vs. HI from PSD features. Our results demonstrate that higher frequency (γ) bands of the EEG are the most robust spectral features for segregating NH from HI listeners, especially using whole-brain data. Classification is also more robust when using LH as compared to RH features. Our frequency band analysis has the potential to be used in clinical settings for early detection of hearing loss, and building models for attentional and listening assistive devices to amplify particular speech sounds for hearing impaired listeners.In future work, we will to investigate what are the brain regions are engaged in categorical perception of speech.

Acknowledgements

This work was supported by National Institutes of Health (NIH/NIDCD R01DC016267), Natural Sciences and Engineering Research Council of Canada (NSERC 194536), and Department of Electrical and Computer Engineering at the University of Memphis. We thank the IIS for covering the conference registration fees.

References

[1] T. Vos, R. M. Barber, B. Bell, A. Bertozzi-Villa, S. Biryukov, I. Bolliger, F. Charlson, A. Davis, L. Degenhardt, D. Dicker *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013," *The Lancet*, vol. 386, no. 9995, pp. 743–800, 2015.

[2] D. L. Blackwell, J. W. Lucas, and T. C. Clarke, "Summary health statistics for us adults: national health interview survey, 2012." *Vital and health statistics. Series 10, Data from the National Health Survey*, no. 260, pp. 1–161, 2014.

[3] M. C. Liberman, "Noise-induced and age-related hearing loss: new perspectives and potential therapies," *F1000Research*, vol. 6, 2017.

[4] C. J. Murray, J. Abraham, M. K. Ali, M. Alvarado, C. Atkinson, L. M. Baddour, D. H. Bartels, E. J. Benjamin, K. Bhalla, G. Birbeck *et al.*, "The state of us health, 1990-2010: burden of diseases, injuries, and risk factors," *Jama*, vol. 310, no. 6, pp. 591–606, 2013.

[5] F. R. Lin, K. Yaffe, J. Xia, Q.-L. Xue, T. B. Harris, E. Purchase-Helzner, S. Satterfield, H. N. Ayonayon, L. Ferrucci, E. M. Simonsick *et al.*, "Hearing loss and cognitive decline in older adults," *JAMA internal medicine*, vol. 173, no. 4, pp. 293–299, 2013.

[6] G. A. Gates and J. H. Mills, "Presbycusis," *The lancet*, vol. 366, no. 9491, pp. 1111–1120, 2005.

[7] M. C. Killion and P. A. Niquette, "What can the pure-tone audiogram tell us about a patient's snr loss," *Hear J*, vol. 53, no. 3, pp. 46–53, 2000.

[8] C. Avcı and A. Akbaş, "Sleep apnea classification based on respiration signals by using ensemble methods," *Bio-Medical materials and engineering*, vol. 26, no. s1, pp. S1703–S1710, 2015.

[9] A. Saidatul, M. P. Paulraj, S. Yaacob, and M. A. Yusnita, "Analysis of eeg signals during relaxation and mental stress condition using ar modeling techniques," in *2011 IEEE International Conference on Control System, Computing and Engineering*. IEEE, 2011, pp. 477–481.

[10] F. Roux, M. Wibral, H. M. Mohr, W. Singer, and P. J. Uhlhaas, "Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory," *Journal of Neuroscience*, vol. 32, no. 36, pp. 12 411–12 420, 2012.

[11] M. Gärtner, L. Rohde-Liebenau, S. Grimm, and M. Bajbouj, "Working memory-related frontal theta activity is decreased under acute stress," *Psychoneuroendocrinology*, vol. 43, pp. 105–113, 2014.

[12] J. D. Beaver, C. J. Long, D. M. Cole, M. J. Durcan, L. C. Bannon, R. G. Mishra, and P. M. Matthews, "The effects of nicotine replacement on cognitive brain activity during smoking withdrawal studied with simultaneous fmri/eeg," *Neuropsychopharmacology*, vol. 36, no. 9, p. 1792, 2011.

[13] H. K. Maganti and M. Matassoni, "A perceptual masking approach for noise robust speech recognition," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2012, no. 1, p. 29, 2012.

[14] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of eeg using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[15] R. S. S. Kumari and J. P. Jose, "Seizure detection in eeg using time frequency analysis and svm," in *2011 international conference on emerging trends in electrical and computer technology*. IEEE, 2011, pp. 626–630.

[16] R. Wang, J. Wang, H. Yu, X. Wei, C. Yang, and B. Deng, "Power spectral density and coherence analysis of alzheimer's eeg," *Cognitive neurodynamics*, vol. 9, no. 3, pp. 291–304, 2015.

[17] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis," *Cancer informatics*, vol. 2, p. 117693510600200030, 2006.

[18] M. S. Mahmud, F. Ahmed, R. Al-Fahad, K. A. Moinuddin, M. Yeasin, C. Alain, and G. Bidelman, "Decoding age-related changes in the spatiotemporal neural processing of speech using machine learning," *bioRxiv*, p. 786566, 2019.

[19] C. Alain, "Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex," *Frontiers in systems neuroscience*, vol. 8, p. 8, 2014.

[20] G. M. Bidelman, M. S. Mahmud, M. Yeasin, D. Shen, S. R. Arnott, and C. Alain, "Age-related hearing loss increases full-brain connectivity while reversing directed signaling within the dorsal–ventral pathway for speech," *Brain Structure and Function*, vol. 224, no. 8, pp. 2661–2676, 2019.

[21] S. Sur and V. Sinha, "Event-related potential: An overview," *Industrial psychiatry journal*, vol. 18, no. 1, p. 70, 2009.

[22] F. Vecchio and S. Määttä, "The use of auditory event-related potentials in alzheimer's disease diagnosis," *International journal of Alzheimer's disease*, vol. 2011, 2011.

[23] Y. Liu, W. Zhou, Q. Yuan, and S. Chen, "Automatic seizure detection using wavelet transform and svm in long-term intracranial eeg," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 20, no. 6, pp. 749–755, 2012.

[24] A. X. Stewart, A. Nuthmann, and G. Sanguinetti, "Single-trial classification of eeg in a visual object task using ica and machine learning," *Journal of neuroscience methods*, vol. 228, pp. 1–14, 2014.

[25] X.-W. Wang, D. Nie, and B.-L. Lu, "Emotional state classification from eeg data using machine learning approach," *Neurocomputing*, vol. 129, pp. 94–106, 2014.

[26] G. M. Bidelman, C. N. Price, D. Shen, S. Arnott, and C. Alain, "Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults," *bioRxiv*, p. 568840, 2019.

[27] F. Maclennan-Smith, D. W. Swanepoel, and J. W. Hall III, "Validity of diagnostic pure-tone audiometry without a sound-treated environment in older adults," *International journal of audiology*, vol. 52, no. 2, pp. 66–73, 2013.

[28] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: a user-friendly application for meg/eeg analysis," *Computational intelligence and neuroscience*, vol. 2011, p. 8, 2011.

[29] C. M. Michel, M. M. Murray, G. Lantz, S. Gonzalez, L. Spinelli, and R. G. de Peralta, "Eeg source imaging," *Clinical neurophysiology*, vol. 115, no. 10, pp. 2195–2222, 2004.

[30] J. Song, C. Davey, C. Poulsen, P. Luu, S. Turovets, E. Anderson, K. Li, and D. Tucker, "Eeg source localization: sensor density and head surface coverage," *Journal of neuroscience methods*, vol. 256, pp. 9–21, 2015.

[31] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman *et al.*, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[32] D. Cruse, S. Chennu, D. Fernández-Espejo, W. L. Payne, G. B. Young, and A. M. Owen, "Detecting awareness in the vegetative state: electroencephalographic evidence for attempted movements to command," *PLoS One*, vol. 7, no. 11, p. e49933, 2012.

[33] G. M. Bidelman, "Induced neural beta oscillations predict categorical speech perception abilities," *Brain and language*, vol. 141, pp. 62–69, 2015.

[34] S. Casale, A. Russo, G. Scebba, and S. Serrano, "Speech emotion classification using machine learning algorithms," in *2008 IEEE international conference on semantic computing*. IEEE, 2008, pp. 158–165.

[35] M. S. Mahmud, M. Yeasin, D. Shen, S. R. Arnott, C. Alain, and G. M. Bidelman, "What brain connectivity patterns from eeg tell us about hearing loss: A graph theoretic approach," in *2018 10th International Conference on Electrical and Computer Engineering (ICECE)*. IEEE, 2018, pp. 205–208.

[36] M. S. Mahmud, F. Ahmed, M. Yeasin, and G. M. Bidelman, "Decoding categorical speech perception from evoked brain responses," in *IEEE Region 10 Symposium (TENSYMP) 2020*. IEEE, 2020.

[37] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets," *PloS one*, vol. 10, no. 3, p. e0118432, 2015.

[38] C.-W. Hsu, C.-C. Chang, C.-J. Lin *et al.*, "A practical guide to support vector classification," 2003.

[39] M. Bhasin and G. Raghava, "Svm based method for predicting hla-drb1* 0401 binding peptides in an antigen sequence," *Bioinformatics*, vol. 20, no. 3, pp. 421–423, 2004.

[40] P. Thanh Noi and M. Kappas, "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery," *Sensors*, vol. 18, no. 1, p. 18, 2018.

[41] Y. Qian, W. Zhou, J. Yan, W. Li, and L. Han, "Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery," *Remote Sensing*, vol. 7, no. 1, pp. 153–168, 2015.

[42] A. Knief, M. Schulte, O. Bertrand, and C. Pantev, "The perception of coherent and non-coherent auditory objects: a signature in gamma frequency band," *Hearing research*, vol. 145, no. 1-2, pp. 161–168, 2000.

[43] A. Yellamsetty and G. M. Bidelman, "Low-and high-frequency cortical brain oscillations reflect dissociable mechanisms of concurrent speech segregation in noise," *Hearing research*, vol. 361, pp. 92–102, 2018.