Longitudinal Speech Biomarkers for Automated Alzheimer's Detection

Jordi Laguarta
Massachusetts Institute of Technology School of Engineering https://orcid.org/0000-0001-9010-6412
Ferran Hueto
Massachusetts Institute of Technology School of Engineering
Prithvhi Rajasekaran
Massachusetts Institute of Technology School of Engineering
Sanjay Sarma
Massachusetts Institute of Technology School of Engineering
Brian Subirana (subirana@mit.edu)
MIT AutoID Laboratory, Cambridge, MA, USA 2Harvard University, Cambridge, MA, USA
https://orcid.org/0000-0002-3069-5275

Research

Keywords: multimodal deep learning, transfer learning, explainable speech recognition, brain model, AI diagnostics

Posted Date: August 18th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-56078/v1

RESEARCH

Longitudinal Speech Biomarkers for Automated Alzheimer's Detection

Jordi Laguarta[1], Ferran Hueto[1,2], Prithvi Rajasekaran[1], Sanjay Sarma[1] and Brian Subirana[1,2*]

*Correspondence: subirana@mit.edu
[1]MIT AutoID Laboratory, Cambridge, MA, USA
[2]Harvard University, Cambridge, MA, USA
Full list of author information is available at the end of the article

Abstract

Background: We introduce a novel speech processing framework, the MIT CBMM Open Voice Brain Model (OVBM), combining implementations of the 4 modules of intelligence: The brain OS chunks and overlaps audio samples and transfers CNN features from the sensory stream and cognitive core creating a multi-modal graph neural network of symbolic compositional models for the target task.

Methods: Our approach consists of pre-training models to extract acoustic features from selected biomarkers and then leverage transfer learning to combine the biomarker feature extractors into a graph neural network to provide an explainable diagnsotic for Alzheimer's Dementia (AD) using speech recordings.

Results: We apply OVBM to the automated diagnostic of Alzheimer's Dementia patients, achieving above state-of-the-art accuracy of 93.8% using only raw audio, while extracting a personalized subject saliency map to track relative disease progression of 16 explainable biomarkers.

Conclusion: By using independent biomarker models, OVBM lets health experts explore biomarker features and whether there are common biomarkers features between AD and other diseases like COVID-19. We present a novel lungs and respiratory tract biomarker created using 200.000+ cough samples to pre-train a model discriminating English from Catalan coughs. Transfer Learning is subsequently used to transfer features from this model with various other biomarker OVBM models. This strategy yielded consistent improvements in AD detection, no matter the combination used. This cough dataset sets a new benchmark as largest audio health dataset with 30.000+ subjects participating in April 2020, demonstrating for the first time cough cultural bias.

Keywords: multimodal deep learning; transfer learning; explainable speech recognition; brain model; AI diagnostics

1 Introduction

We introduce a novel speech processing framework, the MIT CBMM Open Voice Brain Model (OVBM), combining recent advances in our understanding of the four modules of the human brain as researched at MIT's Center for Brain Minds and Machines (CBMM) [1]. Our multimodal framework combines elements of the four CBMM modules of intelligence: the sensory stream and the cognitive core modules are both based on biomarkers trained with specialized CNNs; the brain OS module chunks and overlaps audio samples and transfers features from the sensory stream and cognitive core modules to create a graph neural network for the target task; the fourth module, symbolic compositional models, aggregates the multi-modal predictions from the brain OS input chunks and uses several mechanisms to establish

Table 1 A review of other AD diagnostic algorithms on the same dataset from [10]. Our top performing model only uses audios while Yancheva et al. used both audios and transcripts. Orimaye et al. only used 35 patients hence risking high variance. Karlekar et al. only used transcripts.

Author	Date	Accuracy(%)
Yancheva et al [11]	2016	80.0
Orimaye et al.[12]	2016	87.5
Karlekar et al.[13]	2018	91.1
Laguarta et al.	**2020**	**93.8**

final predictions and explainable visualizations. Previous work, instead, has very little inspiration on the different stages of human intelligence or focuses solely on modelling a small part of the brain[2, 3].

The OVBM framework can be used for various tasks such as speech segmentation and transcription - in this paper we demonstrate it in the individualized and explainable diagnostic of Alzheimer's Dementia (AD) patients, where, as shown in Table 1 we achieve above state-of-the-art accuracy of 93.8%[4] and using only raw audio as input, while extracting, for each subject a saliency map with the relative disease progression of 16 biomarkers. Even with expensive CT scans, to date experts can not create consistent biomarkers [5, 6, 7] even when including emotional biomarkers, unlike our approach which automatically develops them from free speech. Experts point at this lack of biomarkers as the reason why no new drug has been introduced in the last 16 years despite AD[8] being the sixth leading cause of death in the United States [9], and one of the leading unavoidable causes for loss of healthy life. Our emphasis on detecting relevant biomarkers corresponding to the different stages of disease onset, led us to build ten sub-models using four datasets.

By using independent biomarkers, our framework lets us explore whether there may be common biomarkers between AD and other diseases such as COVID-19. We found that cough features, in particular, are very useful biomarker enablers as shown in several experiments reported in this paper. To do so, over 200.000 cough samples were used to pre-train a model discriminating English from Catalan coughs, and then transfer learning was leveraged to transfer its features and integrate it into OVBM brain model, always showing improvements in AD detection, no matter what combination was used. This COVID-19 cough dataset we created approved by MIT's IRB 2004000133 sets a new benchmark as the largest audio health dataset, with over 30.000 subjects participating in less than four weeks in April 2020.

In Section 2 we present the different components of the CBMM Open Voice Brain Model AD detector, from Sections 3 to 6 we introduce the 16 biomarkers with results and a novel personalized AD biomarker comparative saliency map. We conclude in Section 7 with a brief summary and implications for future research.

2 Methods

2.1 Overview of the OVBM Framework

The OVBM in Figure 1 frames a four-module system to test biomarker combinations and provides an explainable diagnostic framework for a target task of AD discrimination.

The Sensory Stream is responsible for pre-training models on large speech datasets to extract features of individual physical biomarkers. The Brain OS splits audio into overlapping chunks and leverages transfer learning strategies to fine-tune the

biomarker models to the smaller target dataset. For longitudinal diagnosis, it includes a round-robin five stage graph neural network that marks salient events in continuous speech. The Cognitive Core incorporates medical knowledge specific to the target task to train cognitive biomarker feature extractors. The Symbolic Compositional Model combines fine-tuned biomarker models into a graph neural network and its predictions on individual audio chunks are aggregated and fed into an aggregating engine to reach a final diagnostic plus a patient saliency map. To enable doctors to gain insight into the specific condition of a given patient, one of the novelties of our approach is that the outputs at each unique module are extracted to create a visualization in the form of a health diagnostic saliency map showing the impact of the selected biomarkers. This saliency map could be used as a form to longitudinally track and visualize disease progression.

2.1.1 OVBM Applied to AD detection

Next, we review each of the four OVBM modules in the context of AD, introducing 16 biomarkers and gradually explaining the partial GNN architecture shown in Figure 1. To be able to compare models, our baselines and 8 of the biomarkers are based on the ResNet50 CNN due to its state-of-the-art performance on medical speech recognition tasks such as [14]. All audio samples are processed with the MFCC package published by [15], and padded accordingly. We operate on Mel Frequency Cepstral Coefficients (MFCC), instead of spectrograms[16] because of its resemblance to how the human cochlea captures sound[17]. All audio data uses the same MFCC parameters (Window Length: 20ms, Window Step: 10ms, Cepstrum Dimension: 200, Number of Filters: 200, FFT Size: 2048, Sample rate: 16000). All datasets follow a 70/30 train-test split and models are trained with an Adam optimizer[18].

The dataset from DementiaBank, ADrESS [19], is used for training the OVBM framework and fine-tuning all biomarker models on AD detection. It consists of subject recordings in full enhanced audio and short normalised sub-chunks, along with the recording transcriptions and metadata from 78 AD and 78 non-ad patients. For the approach of this study focusing purely on audio processing we only use the full enhanced audio and patient metadata.

Figure 1 OVBM GNN Architecture at a given Brain OS time.

2.2 OVBM AD Sensory Stream Biomarkers

We have selected four biomarkers previously inspired by medical community choices [20, 21, 22, 23, 24], as reviewed next:

2.2.1 ***Biomarker 1 (Muscular Degradation)***

One of the main early-stage biomarkers is memory loss[24], which occurs both at a conceptual level as well as at a muscular level[21]. We follow memory decay models from [25, 26] to capture this muscular metric by degrading input signals for all train and test sets with the **Poisson** mask in Equation 1, a commonly occurring

distribution in nature[27]. We use as a Possion function a mask with input MFCC image = I_x, output mask = $M(I_X)$, $\lambda = 1$, and k = each value in I_x:

$$M(I_x) = Pr(\lambda)I_x \tag{1}$$

$$Pr(X = k) = \lambda^k e^{-k} k! \tag{2}$$

2.2.2 ***Biomarker 2 (Vocal cords)***

Significant research in AD such as [20] has proven that the disease impacts motor neurons. In other diseases, like Parkinson's, where motor neurons are affected, vocal cords have proven to be one of the first muscles affected [28]. Therefore, we test vocal cords as a biomarker for this experiment. We train a Wake Word (WW) model to learn vocal cord features on LibriSpeech - an audiobook dataset with ≈1,000 hours of speech from [29] by creating a balanced sample set of 2 second audio chunks, half containing the word 'Them' and half without. A ResNet50[30] is trained for binary classification of **'Them'** on 3s audio chunks(lr:0.001, val_acc: 89%).

2.2.3 ***Biomarker 3 (Sentiment)***

Clinical evidence supports the importance of sentiments in AD early-diagnosis [22, 31], and different clinical settings emphasize different sentiments, such as doubt[32] or frustration[32]. We train a Sentiment Speech classifier model to learn **intonation** features on the RAVDESS - an emotional speech dataset by [33] of actors speaking in 8 different emotional states. A ResNet50[30] is trained on categorical classification of 8 corresponding intonations such as calm, happy or disgust(lr: 0.0001, val_acc on 8 classes: 71%).

2.2.4 ***Biomarker 4 (Lungs and Respiratory Tract)***

The human cough is already used to diagnose several diseases using audio recognition [34, 35] as it provides information corresponding biomarkers in the lungs and respiratory tract[36]. People with chronic lung disorders are more than twice as likely to have AD[23], therefore we hypothesise features extracted from a cough classifier could be valuable for AD diagnosis. We use the **cough** dataset collected through MIT Open Voice for COVID-19 detection [37], strip all but the spoken language of the person coughing (English, Catalan), and split audios into 6s chunks. A ResNet50[30] is trained on binary classification (Input: MFCC 6s Audio Chunks (1 cough) - Output: English/Catalan, lr: 0.0001, val_acc: 86%).

2.3 OVBM Brain OS Biomarkers

The Brain OS is responsible for transferring learned features from the individual biomarker models in the Sensory Stream and Cognitive Core, integrating them into the OVBM model, and training it for a target task, in this case AD detection.

To make the most out of the short patient recordings, we split each patient recording into overlapping audio chunks, allowing chunks to overlap into intervals (0s-4s, 2s-6s, 4s-8s). With all the pre-trained biomarker models in the sensory stream and

cognitive modules, we select the best pre-trained biomarkers, concatenate them together and pass their outputs into a 1024 neuron deeply connected neural network layer and ReLU activation. We also incorporate at this point metadata such as gender. We test three Brain OS transfer learning strategies: (1) CNNs are used as fixed feature extractors without any fine-tuning (2) CNNs are fine-tuned by training all layers (3) Only the final layers of the CNN are fine-tuned.

2.4 OVBM Cognitive Core Biomarkers

Neuropsychological tests are a common screening tool for AD[32]. These tests, among others, evaluate a patient's ability to remember uncommon words, contextualize, infer actions and detect saliency[32, 31]. In the case of this AD dataset, all patients are asked to describe the Cookie Theft picture created by [38], where a set of words such as 'kitchen' (**context**), 'tipping' (**unique**), 'jar' (**inferred**) and 'overflow' (**salient**), are used to capture four cognitive biomarkers. To keep the richness of speech, we train four wake word models from LibriSpeech[29] with ResNet50s following the same approach as Biomarker 2. The four chosen cognitive biomarkers detect patients' ability on: context, uniqueness, inference and saliency.

2.5 OVBM Symbolic Compositional Models Biomarkers

This module fine-tunes previous modules outputs into a graph neural network. Predictions on individual audio chunks for one subject are aggregated and fed into competing models to reach a final diagnostic. In the AD implementation, given we had at most 39 overlapping chunks, three simple aggregation metrics are compared: averaging, linear positive (more weight given to later chunks) and linear negative (more weight given to earlier chunks).

3 Results

3.1 Sensory Stream

3.1.1 Biomarker 1 (Muscular Degradation)

As shown in Table 2, the Poisson biomarker brings a unique improvement to each model except for Cough, consistent with both inherently capturing features that contain muscular degradation.

Table 2 Impact of Poisson mask on AD performance. Baseline is a ResNet50 trained on the AD task without transfer learning.

Model	W/o Poisson(%)	With Poisson(%)
Baseline	65.6	68.8
Cough	**75.0**	**75.0**
Intonation	68.8	75.0
Wake-Word 'Them'	75.0	78.1
Multi-Modal	90.6	93.8
Avg Improvement(%)		3.1

3.1.2 Biomarker 2 (Vocal cords)

Illustrated in Table 3 and Figure 3, this vocal cords model proves it's contribution of unique features, which without fine-tuning to the AD task performs as well as the baseline ResNet50 fully trained on AD, and significantly beats it when fully fine-tuned.

3.1.3 Biomarker 3 (Sentiment)

As illustrated by Table 3 and Figure 2, this biomarker captures unique features for AD detection, and when only fine-tuning its final five layers outperforms a fully trained ResNet50 on AD detection.

3.1.4 Biomarker 4 (Lungs and Respiratory Tract)

Figure 3 and Table 3, justify the features extracted by this cough model as valuable for the task of AD detection by capturing a unique set of samples and improving performance. Further, Figure 2 validates it's impact on the top performing multi-modal, justifying the relevance of this novel biomarker.

Figure 2 Impact of sensory stream biomarkers on OVBM performance through leave-one-out strategy - removing transferred knowledge one at a time. Top dotted sections of bars indicates there is always performance gain from the cough biomarker. Baselines are the OVBM trained on AD without any transfer learning. In the other bars, a biomarker model is removed and replaced with an AD pre-trained ResNet50, hence removing the transferred feature knowledge but conserving computational power. This proves the complementarity of the biomarker models since all are needed for maximum results.

Figure 3 Sensory Stream Saliency Bar Chart: To illustrate the potential of our approach we show the strength of the simplest transfer models we tried. The most surprising, perhaps, is that the simple wakeword model to find the word 'Them' is as powerful as the baseline. If we let the model fine-tune the last few (0-5-10) layers then it goes well beyond it. Our novel Cough database, inspired in the effect of AD in the respiratory tract also shows surprising results, even without any adaptation at all. If we let fine-tuning of the whole model, it's validation accuracy improves $\approx 10\%$ with respect to the baseline. Baseline is the same OVBM architecture trained on AD without any transfer learning of features.

Table 3 To illustrate the complementary nature of the biomarkers we show the unique AD patients detected by the simplest transfer models we tried. Each transfer model detects unique patients reinforcing the potential of combining biomarkers.

Biomarker	**Model Name**	**Unique(%)**
Respiratory Tract	Cough	9.38
Sentiment	Intonation	18.75
Vocal cords	WW 'THEM'	15.63
R. Tract & Sentiment	Cough & Tone.	0
R. Tract & Vocal cords	Cough & WW	6.25
Sentiment & Vocal cords	Tone. & WW	3.13
In All 3		40.63
In Neither of the 3		6.25

3.2 Brain OS

From Figure 4, it is evident AD detection improves as chunk length increases since attention-marking has more per-chunk information to formulate a better AD story. From this attention-marking index (quantity of information required in a chunk for a confident diagnosis) we select chunk sizes **2s**, **8s**, **14s**, and **20s**, shown in Figure 6, as the Brain OS biomarkers, establishing individual AD progression. In terms of transfer learning strategies, Figure 3 shows that fine-tuning all layers always leads to top results, however for most models almost no fine-tuning is required to beat the baseline.

Figure 4 The two top lines illustrate the full OVBM performance, with it's biomarker feature models, as a function of chunk size. PT refers to individually fine-tuning each biomarker model for AD before re-training the whole OVBM. The middle line shows the OVBM without the cognitive core, illustrating how it boosts performance by about 10% at all time chunk sizes. Baseline PT is the OVBM architecture with each ResNet50 inside individually trained on AD before retraining them together in the OVBM architecture.

3.3 Cognitive Core

We could show the same saliency bar chart in Figure 3 and a uniqueness table such as Table 3 to illustrate the impact of each cognitive biomarker, but due to sake of repetition we opt not to. Instead in Figure 4 we show the impact of removing the cognitive core on the top OVBM performance which drops ≈10%, validating the relevance of the cognitive core biomarkers.

3.4 Symbolic Compositional Models

In Figure 5: averaging proves to be the most effective method for aggregating chunk predictions. Positive linear outperforming the negative linear indicates the later audio chunks are more informative than the front ones. Figure 6 includes 4 biomarkers derived from combining chunk predictions from biomarker models of the three other modules. With more data and longitudinal recordings, the OVBM GNN may add other biomarkers.

Figure 5 Relation between chunk size and AD discrimination error, showing increased importance of the later chunks.

Figure 6 **A:** Saliency map to study the explainable AD evolution for all the patients in the study based on the predictions of individual biomarker models. BrainOS (2,8,14,20) show the model prediction at different chunk sizes. This map could be used to longitudinally monitor subjects where a lower score on the biomarkers could indicate a more progressed AD subject. **B:** Saliency map comparing AD+ subject S092 and AD- subject S019.

4 Conclusion

We conclude by providing a few insights further supporting our brain-inspired model for audio health diagnostics as presented above. We have proven the success of the OVBM framework, setting the new benchmark for state-of-the-art accuracy of AD classification on the largest AD dataset available [19]d, despite only incorporating audio signals - the first to do so using GNNs[39]. We have shown the additive nature of biomarker feature exctractor models for the biomarkers chosen in this study, where the OVBM sets a platform for AI speech diagnostic expects to hypothesize and test relevant biomarkers for the audio diagnostic of a specific disease. Future work may improve this benchmark by also incorporating into OVBM longitudinal GNN's natural language biomarkers using NLP classifiers or multi-modal graph neural networks incorporating non-audio diagnostic tools[40].

5 Discussion

One of the most surprising insights of all is the discovery of cough as a new biomarker (Figure 2), one that improves any of the intermediate models tested. Further, our leave-one-out approach to evaluate biomarker models validates the OVBM as a framework on which medical experts can hypothesize and test out the impact of existing and novel biomarkers for the diagnostic of AD using speech. We are the first to report that cough biomarkers have information related to gender and culture, and are also the first to demonstrate how they improve AD classification as illustrated in the saliency charts (Figure 3).

A promising finding is the model's explainability, introducing the biomarker AD saliency map, offering a novel method to evaluate patients longitudinally on a set of physical and neuropsychological biomarkers as shown on Fig. 6. In future research, longitudinal data may be collected to properly test the onset potential of OVBM GNN discrimination in continuous speech. We hope this approach brings the AI health diagnostic experts closer to the medical community and helps monitor patients as well as accelerate research for treatments by providing longitudinal and explainable tracking metrics that can help succeed adaptive clinical trials of urgently needed innovative interventions.

6 Declarations

6.1 Competing interests

The authors declare that they have no competing interests.

6.2 Ethical Approval and Consent to participate

The COVID-19 cough dataset was collected through an open website with MIT's IRB 2004000133.

6.3 Consent for publication

Consent for publication has been granted by all authors.

6.4 Availability of data and materials

Not applicable.

6.5 Competing interests

Not applicable.

6.6 Funding

Not applicable.

6.7 Authors' contributions

Not applicable.

6.8 Acknowledgements

Not applicable.

6.9 Authors' information

Please contact Brian Subirana at subirana@mit.edu for any questions, clarifications, or feedback.

Author details

[1]MIT AutoID Laboratory, Cambridge, MA, USA. [2]Harvard University, Cambridge, MA, USA.

References

1. Modules | The Center for Brains, Minds & Machines. https://cbmm.mit.edu/research/modules Accessed 2020-05-14
2. Nassif, A.B., Shahin, I., Attili, I., Azzeh, M., Shaalan, K.: Speech recognition using deep neural networks: A systematic review. IEEE Access **7**, 19143–19165 (2019)
3. Reddy, D.R.: Speech recognition by machine: A review. Proceedings of the IEEE **64**(4), 501–531 (1976)
4. Pulido, M.L.B., Hernández, J.B.A., Ballester, M.Á.F., González, C.M.T., Mekyska, J., Smékal, Z.: Alzheimer's disease and automatic speech analysis: a review. Expert Systems with Applications, 113213 (2020)
5. James, H.J., Van Houtven, C.H., Lippmann, S., Burke, J.R., Shepherd-Banigan, M., Belanger, E., Wetle, T.F., Plassman, B.L.: How accurately do patients and their care partners report results of amyloid-β pet scans for alzheimer's disease assessment? Journal of Alzheimer's Disease (Preprint), 1–12 (2020)
6. Henriksen, K., O'Bryant, S., Hampel, H., Trojanowski, J., Montine, T., Jeromin, A., Blennow, K., Lönneborg, A., Wyss-Coray, T., Soares, H., Bazenet, C., Sjögren, M., Hu, W., Lovestone, S., Karsdal, M., Weiner, M., Laske, C.: The future of blood-based biomarkers for alzheimer's disease. Alzheimer's Dementia **10**, 115–131 (2014). doi:10.1016/j.jalz.2013.01.013
7. Morsy, A., Trippier, P.: Current and emerging pharmacological targets for the treatment of alzheimer's disease. Journal of Alzheimer's Disease **72**, 1–33 (2019). doi:10.3233/JAD-190744
8. Zetterberg, H.: Blood-based biomarkers for alzheimer's disease—an update. Journal of neuroscience methods **319**, 2–6 (2019)
9. Association, A.: 2019 alzheimer's disease facts and figures. Alzheimer's & Dementia **15**(3), 321–387 (2019)
10. Lyu, G.: A review of alzheimer's disease classification using neuropsychological data and machine learning. In: 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp. 1–5 (2018)
11. Yancheva, M., Rudzicz, F.: Vector-space topic models for detecting alzheimer's disease. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 2337–2346 (2016)
12. Orimaye, S.O., Wong, J.S.-M., Golden, K.J.: Learning predictive linguistic features for alzheimer's disease and related dementias using verbal utterances. In: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 78–87 (2014)
13. Karlekar, S., Niu, T., Bansal, M.: Detecting linguistic characteristics of alzheimer's dementia by interpreting neural models. arXiv preprint arXiv:1804.06440 (2018)
14. Ghoniem, R.M.: Deep genetic algorithm-based voice pathology diagnostic system. In: Métais, E., Meziane, F., Vadera, S., Sugumaran, V., Saraee, M. (eds.) Natural Language Processing and Information Systems, pp. 220–233. Springer, Cham (2019)
15. Lyons, J., Wang, D.Y.-B., Gianluca, Shteingart, H., Mavrinac, E., Gaurkar, Y., Watcharawisetkul, W., Birch, S., Zhihe, L., Hölzl, J., Lesinskis, J., Almér, H., Lord, C., Stark, A.: Jameslyons/python_speech_features: Release V0.6.1. doi:10.5281/zenodo.3607820. https://doi.org/10.5281/zenodo.3607820
16. Lee, H., Pham, P., Largman, Y., Ng, A.Y.: Unsupervised feature learning for audio classification using convolutional deep belief networks. In: Bengio, Y., Schuurmans, D., Lafferty, J.D., Williams, C.K.I., Culotta, A. (eds.) Advances in Neural Information Processing Systems 22, pp. 1096–1104. Curran Associates, Inc., ??? (2009)
17. Krijnders, J., t Holt, G.: Tone-fit and mfcc scene classification compared to human recognition. Energy [dB] **400**(450), 500 (2017)
18. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
19. Luz, S., Haider, F., de la Fuente, S., Fromm, D., MacWhinney, B.: Alzheimer's dementia recognition through spontaneous speech: The ADReSS Challenge. In: Proceedings of INTERSPEECH 2020, Shanghai, China (2020). https://arxiv.org/abs/2004.06833
20. Heckman, P.R., Blokland, A., Prickaerts, J.: From age-related cognitive decline to alzheimer's disease: a translational overview of the potential role for phosphodiesterases. In: Phosphodiesterases: CNS Functions and Diseases, pp. 135–168. Springer, ??? (2017)
21. Wirths, O., Bayer, T.: Motor impairment in alzheimer's disease and transgenic alzheimer's disease mouse models. Genes, Brain and Behavior **7**, 1–5 (2008)
22. Galvin, J., Tariot, P., Parker, M.W., Jicha, G.: Screen and intervene: The importance of early detection and treatment of alzheimer's disease. The Medical Roundtable General Medicine Edition (2020)
23. Dodd, J.W.: Lung disease as a determinant of cognitive decline and dementia. Alzheimer's research & therapy **7**(1), 32 (2015)
24. Chertkow, H., Bub, D.: Semantic memory loss in dementia of alzheimer's type: What do various measures measure? Brain **113**(2), 397–417 (1990)
25. Subirana, B., Bagiati, A., Sarma, S.: On the forgetting of college academice: at" ebbinghaus speed"? Technical report, Center for Brains, Minds and Machines (CBMM) (2017)
26. Cano-Cordoba, F., Sarma, S., Subirana, B.: Theory of intelligence with forgetting: Mathematical theorems explaining human universal forgetting using "forgetting neural networks". Technical report, Center for Brains, Minds and Machines (CBMM) (2017)

27. Reed, W.J., Hughes, B.D.: From gene families and genera to incomes and internet file sizes: Why power laws are so common in nature. Physical Review E **66**(6), 067103 (2002)
28. J. Holmes, R., M. Oates, J., J. Phyland, D., J. Hughes, A.: Voice characteristics in the progression of parkinson's disease. International Journal of Language & Communication Disorders **35**(3), 407–418 (2000)
29. Panayotov, V., Chen, G., Povey, D., Khudanpur, S.: Librispeech: an asr corpus based on public domain audio books. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5206–5210 (2015). IEEE
30. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
31. Costa, A., Bak, T., Caffarra, P., Caltagirone, C., Ceccaldi, M., Collette, F., Crutch, S., Della Sala, S., Démonet, J.F., Dubois, B., *et al.*: The need for harmonisation and innovation of neuropsychological assessment in neurodegenerative dementias in europe: consensus document of the joint program for neurodegenerative diseases working group. Alzheimer's research & therapy **9**(1), 27 (2017)
32. Baldwin, S., Farias, S.T.: Unit 10.3: Assessment of cognitive impairments in the diagnosis of alzheimer's disease". Current protocols in neuroscience/editorial board, Jacqueline N. Crawley...[et al.], 10–3 (2009)
33. Livingstone, S.R., Russo, F.A.: The ryerson audio-visual database of emotional speech and song (ravdess): A dynamic, multimodal set of facial and vocal expressions in north american english. PloS one **13**(5) (2018)
34. Abeyratne, U.R., Swarnkar, V., Setyati, A., Triasih, R.: Cough sound analysis can rapidly diagnose childhood pneumonia. Annals of biomedical engineering **41**(11), 2448–2462 (2013)
35. Pramono, R.X.A., Imtiaz, S.A., Rodriguez-Villegas, E.: A cough-based algorithm for automatic diagnosis of pertussis. PloS one **11**(9) (2016)
36. Bennett, W.D., Daviskas, E., Hasani, A., Mortensen, J., Fleming, J., Scheuch, G.: Mucociliary and cough clearance as a biomarker for therapeutic development. Journal of aerosol medicine and pulmonary drug delivery **23**(5), 261–272 (2010)
37. Subirana, B., Hueto, F., Rajasekaran, P., Laguarta, J., Puig, S., Malvehy, J., Mitja, O., Trilla, A., Moreno, C.I., Valle, J.F.M., et al.: Hi sigma, do i have the coronavirus?: Call for a new artificial intelligence approach to support health care professionals dealing with the covid-19 pandemic. arXiv preprint arXiv:2004.06510 (2020)
38. Goodglass, H., Kaplan, E., Barressi, B.: Cookie theft picture. Boston diagnostic aphasia examination. Philadelphia, PA: Lea & Febiger (1983)
39. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S.Y.: A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems (2020)
40. Parisot, S., Ktena, S.I., Ferrante, E., Lee, M., Guerrero, R., Glocker, B., Rueckert, D.: Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease. Medical image analysis **48**, 117–130 (2018)

Figures

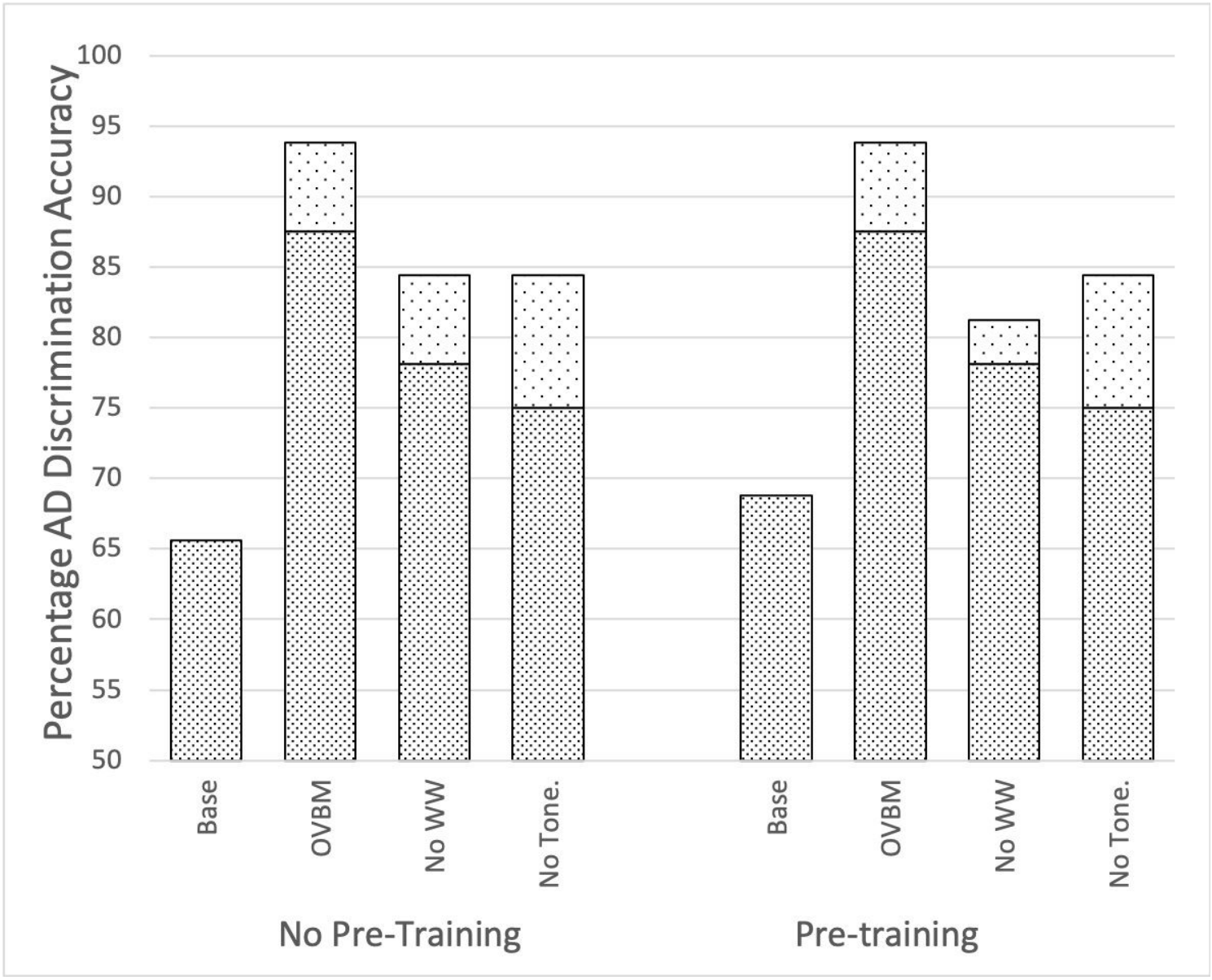

Figure 1

OVBM GNN Architecture at a given Brain OS time.

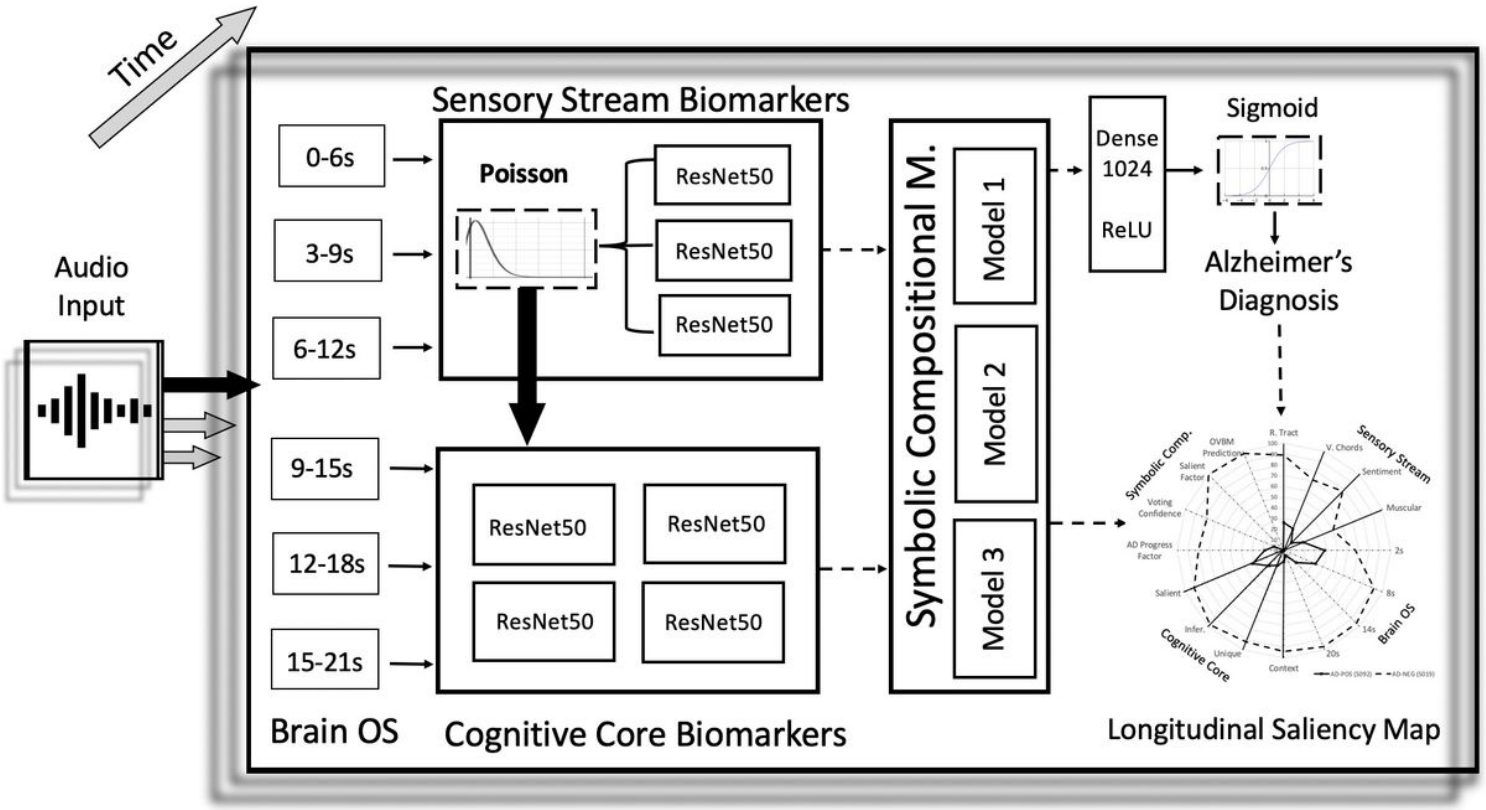

Figure 2

Impact of sensory stream biomarkers on OVBM performance through leave-one-out strategy - removing transferred knowledge one at a time. Top dotted sections of bars indicates there is always performance gain from the cough biomarker. Baselines are the OVBM trained on AD without any transfer learning. In the other bars, a biomarker model is removed and replaced with an AD pre-trained ResNet50, hence removing the transferred feature knowledge but conserving computational power. This proves the complementarity of the biomarker models since all are needed for maximum results.

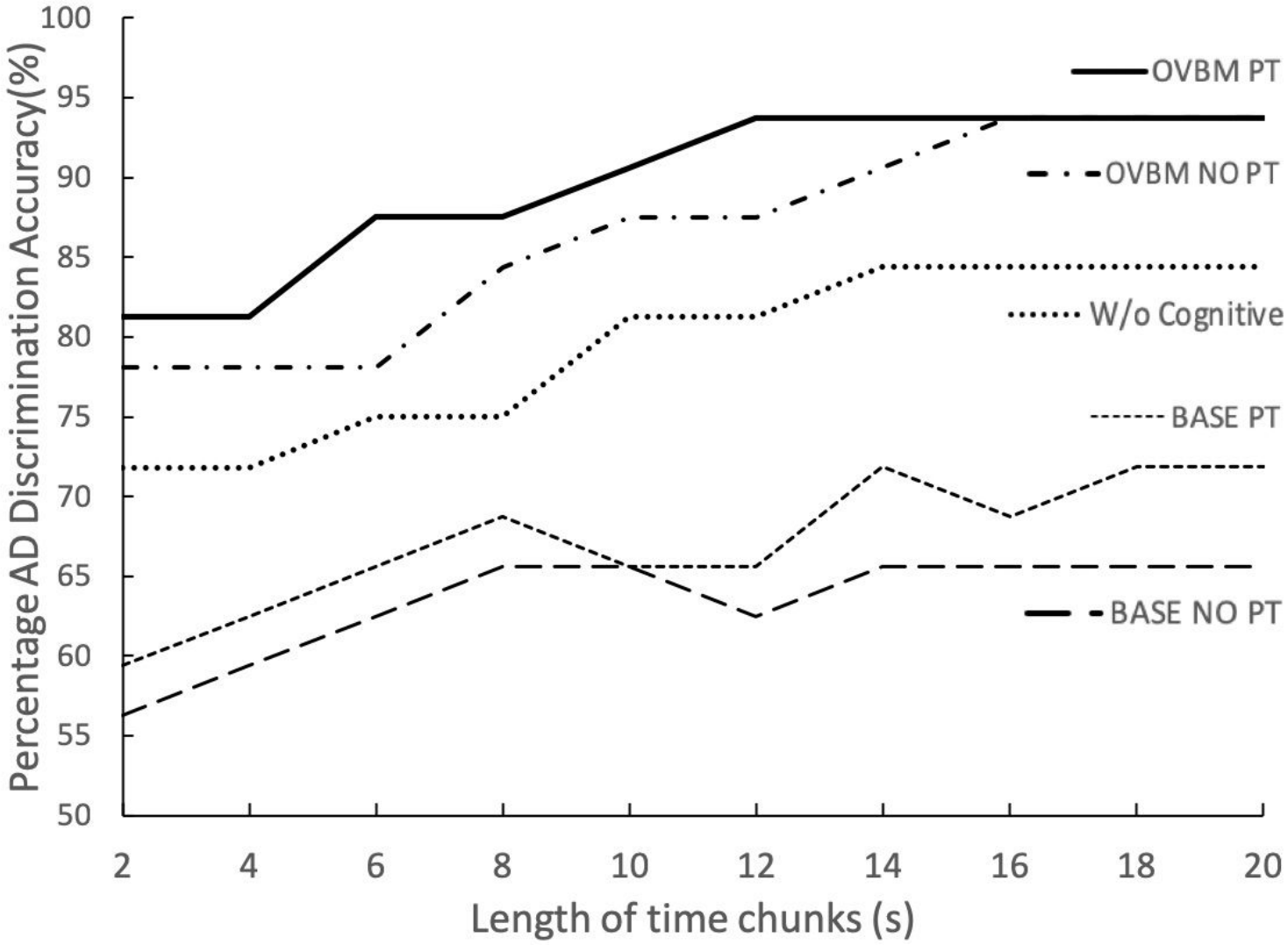

Figure 3

Sensory Stream Saliency Bar Chart: To illustrate the potential of our approach we show the strength of the simplest transfer models we tried. The most surprising, perhaps, is that the simple wakeword model to nd the word 'Them' is as powerful as the baseline. If we let the model ne-tune the last few (0-5-10) layers then it goes well beyond it. Our novel Cough database, inspired in the e□ect of AD in the respiratory tract also shows surprising results, even without any adaptation at all. If we let ne-tuning of the whole model, it's validation accuracy improves □ 10% with respect to the baseline. Baseline is the same OVBM architecture trained on AD without any transfer learning of features.

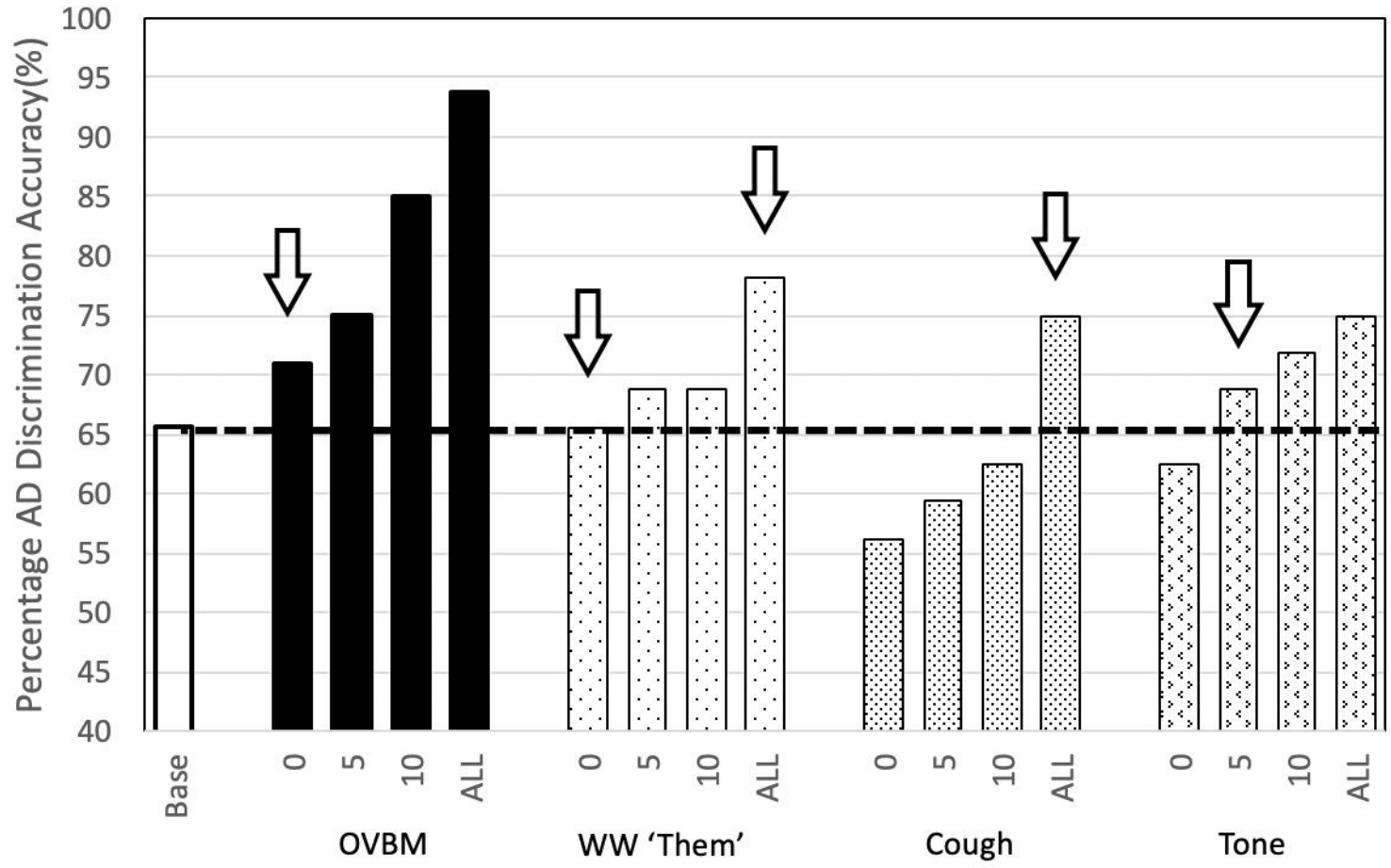

Figure 4

The two top lines illustrate the full OVBM performance, with it's biomarker feature models, as a function of chunk size. PT refers to individually ne-tuning each biomarker model for AD before re-training the whole OVBM. The middle line shows the OVBM without the cognitive core, illustrating how it boosts performance by about 10% at all time chunk sizes. Baseline PT is the OVBM architecture with each ResNet50 inside individually trained on AD before retraining them together in the OVBM architecture.

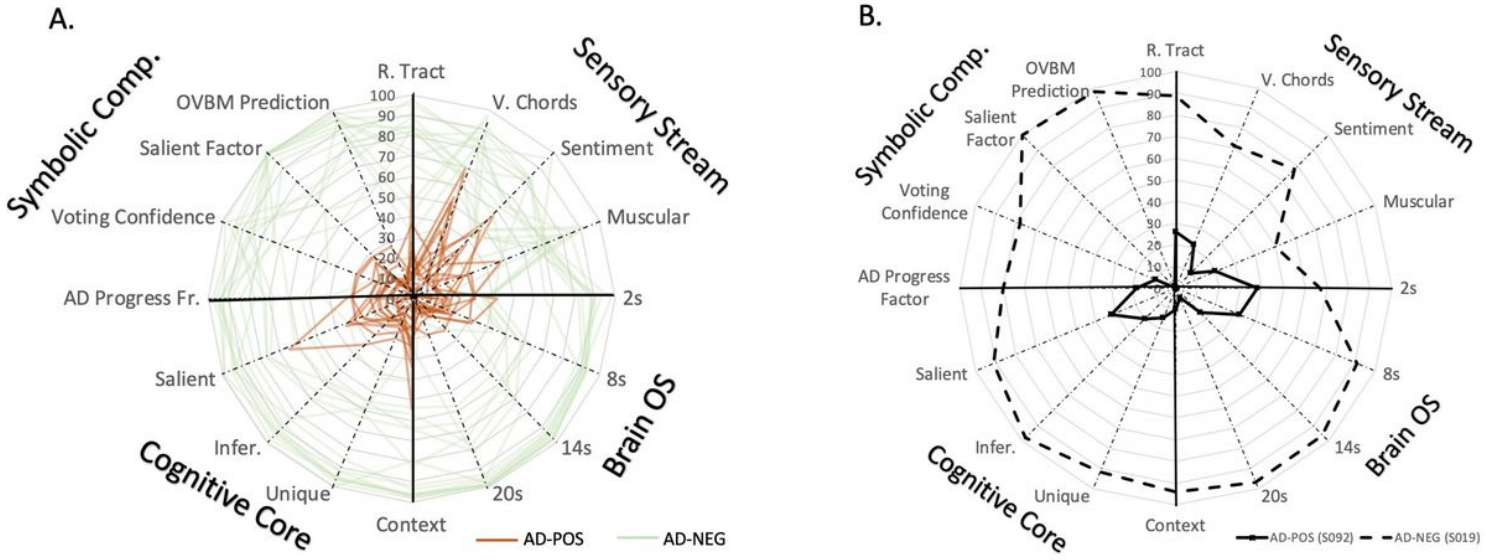

Figure 5

Relation between chunk size and AD discrimination error, showing increased importance of the later chunks.

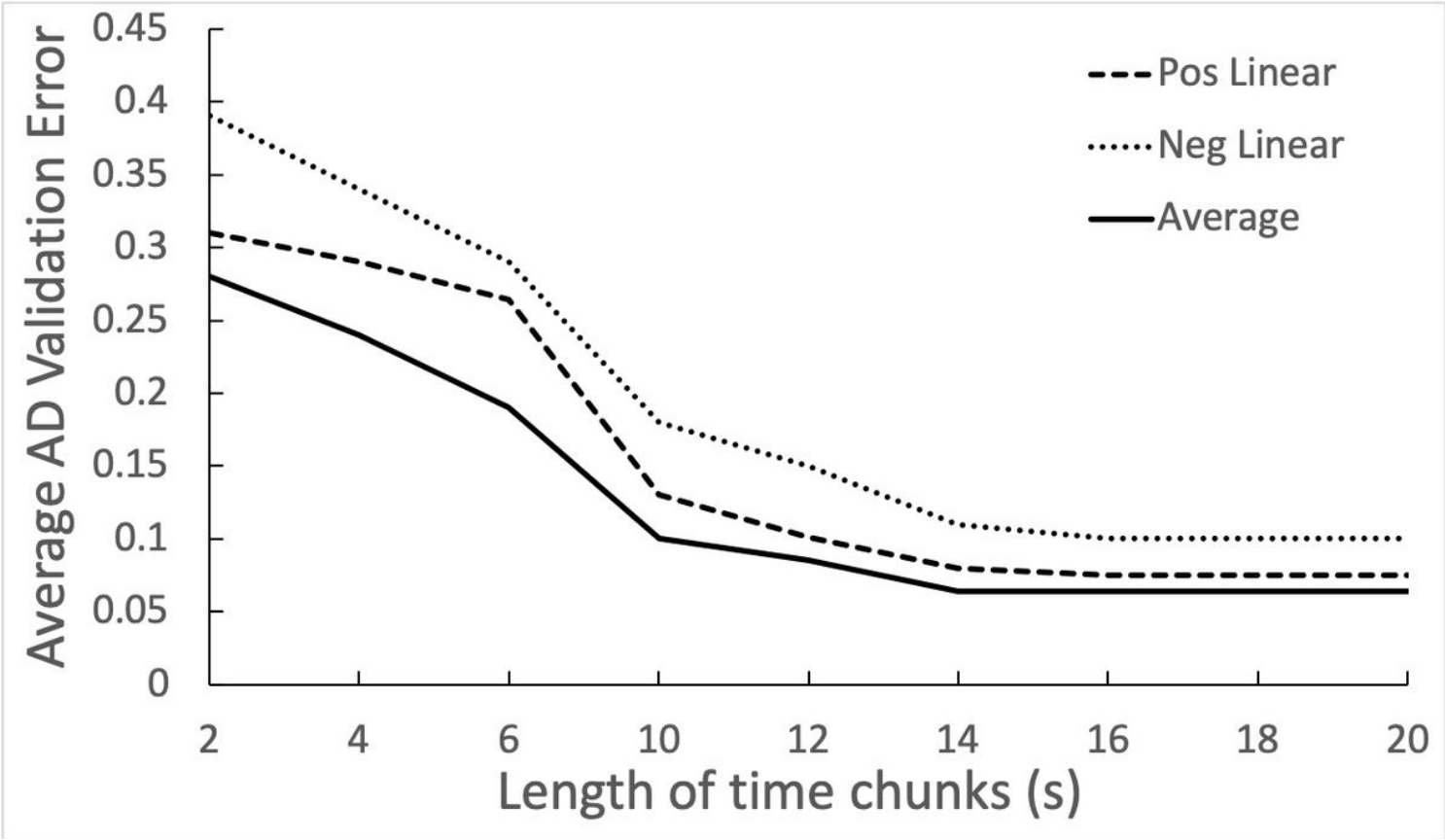

Figure 6

Saliency map to study the explainable AD evolution for all the patients in the study based on the predictions of individual biomarker models. BrainOS (2,8,14,20) show the model prediction at di erent chunk sizes. This map could be used to longitudinally monitor subjects where a lower score on the biomarkers could indicate a more progressed AD subject. B: Saliency map comparing AD+ subject S092 and AD- subject S019.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- readme.html
- bmcarticle.tex
- bmcartbiblio.sty
- bmcart.cls
- spbasic.bst

- bmcmathphys.bst
- bmcarticle.pdf
- bmcarticle.bib
- bmcarticle2col.pdf
- vancouver.bst