2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-2782-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/EMBC48229.2022.9871348

Transient Intervals of Significantly Different Whole Brain Connectivity Predict Recovery vs. Progression from Mild Cognitive Impairment: New Insights from Interpretable LSTM Classifiers

Yutong Gao, Vince D. Calhoun and Robyn L. Miller

Abstract— **The high dimensionality and complexity of time-varying measures of functional brain connectivity have created an environment in which a very rich transformation of the data remains difficult to map into disease states without some form of reduction (averaging, clustering, statistical blindness to the multivariate interactions between features that modulate their contributions). In this work, employing a recently developed architecture for long short-term memory classifiers that supports use of gradient-based model interpretability techniques, we predict progression or recovery from mild cognitive impairment (MCI) from an instantaneous (windowless) wavelet-based measure of dynamic functional network connectivity. This time-attention LSTM (TA-LSTM) model achieves 0.79 AUC on the task of predicting which MCI patients who will recover (RMCI) vs. those who will progress (PMCI) to AZD within a three-year timeframe. Using a common gradient-based model interpretation technique, saliency analysis, on this TA-LSTM points to potentially important predictive dynamic biomarkers, including the duration of the highly salient time intervals and the average connectivity patterns within these highly salient intervals.**

Index Term—**Brain Dynamics, Dynamic Functional Network Connectivity, Long Short-term Memory, Explainable Deep Learning, Wavelet Transform, rs-fMRI, Mild Cognitive Impairment, Alzheimer's Disease**

I. INTRODUCTION

Alzheimer's Disease (AZD) is a neurodegenerative disease that has risen to become the fifth leading cause of death in adults over 65 in the United States [1]. Mild cognitive impairment (MCI) is a precursor-stage of AZD that has a high chance of progressing to AZD in five years, but can still be reversed to cognitive health (CN) [2]. Therefore, prediction progression to AZD or recovery to CN is vital that offer early medical intervention.

Blood oxygenation level dependent (BOLD) functional magnetic resonance imaging (fMRI) measures changes in blood flow as a proxy for the dynamically varying activation levels in localized neuronal populations. In the resting state paradigm (rs-fMRI), subjects are scanned under task-free conditions, with no explicit experimental guidance. This paradigm has been used to study evolving configurations of connectivity in the brain [3], particularly associated with mental disorders [4]. Dynamic functional network connectivity (dFNC) is a transformed measure from rs-fMRI in which time-varying functional integration, typically computed as a set of network-pair correlations assessed on successive sliding windows through the scan (SWCdFNC) [5]. Windowless wavelet-based dynamic functional network connectivity (WWdFNC) captures connectivity at each timepoint using time-varying frequency domain information from the continuous wavelet transform [6]. WWdFNC functions as a low-pass filter on the dynamics and blurs the base of information that a recurrent neural network classifier can exploit for its task.

Deep learning has made significant advances in the neuroimaging domain for classification and prediction tasks by allowing the model to learn class discriminative patterns on various levels of abstraction from the data representation. Understanding and interpreting the model through post-hoc analysis will enable us to uncover the group-wise biomarkers. The saliency map approach [7] is a gradient-based "black-box decoding" approach that allows visualization of each input feature's value's contribution to the final prediction in the well-trained model. In general, gradient-based feature importance techniques applied to recurrent models are biased toward proximal timepoints due to the "vanishing gradient" phenomenon, making it difficult to leverage their powerful predictive performance on dynamic inputs for model interpretation. The time-attention LSTM overcomes this limitation by inserting an attention layer between the deepest LSTM layer and the final layer [8] (Figure 1).

This work introduces a model interpretation pipeline for identifying potential predictive time-varying group-wise connectivity biomarkers in the WWdFNC. We trained and tested the TA-LSTM model to predict patients with MCI who will progress to AZD vs. recover to HC in the next three years after rs-fMRI scanning. This is the first study of WWdFNC for feature learning on the deep learning model, and the results indicate WWdFNC as a feature representation that has more discriminative power than SWCdFNC. We then constructed a per-input saliency map to represent the contributing level to the final prediction. We performed several statistical analyses to find the predictive dynamic group-wise biomarkers, including the duration of the highly salient time intervals and connectivity attributes in the highlighted intervals.

II. MATERIALS AND METHODS

A. Participants and Data

We used data from the latest release in the Open Access Series of Imaging Studies (OASIS-3). OASIS-3 is a longitudinal dataset that included over one thousand participants

Research supported by NSF 2112455 and NIH R01MH123610

Y. Gao, V. D. Calhoun and R. L. Miller are with the Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS): Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, Georgia ygao11@student.gsu.edu, vcalhoun@gsu.edu, robyn.l.miller@gmail.com

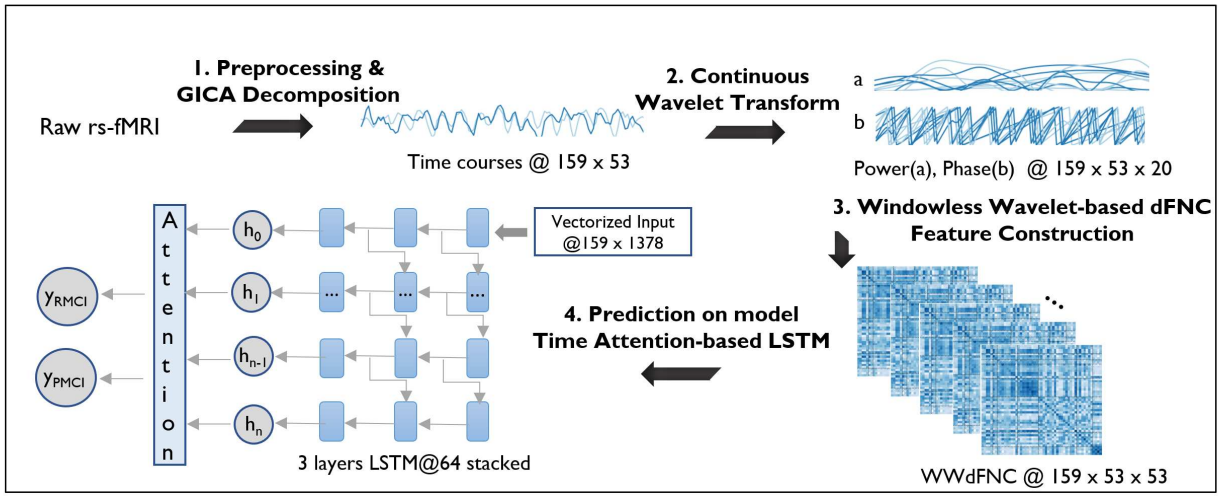

Fig. 1. The framework of Windowless Wavelet-based dFNC feature construction and Time-Attention LSTM model Architecture

at various stages of cognitive decline related to Alzheimer's Disease [9][1]. We used the clinical dementia rating (CDR) scale score at and three years after the scan together to identify the progression directions of MCI patients. The RMCI subjects had a CDR of 0.5 at the scan time and returned to 0 in three years. The PMCI subjects had a CDR of 0.5 at the scan time and progressed to a CDR of greater than 0.5 in the next three-year timeframe. We used one rs-fMRI scan per participant in our final sample dataset, and the final dataset consists of 94 rs-fMRI scans (50 RMCI, 44 PMCI) with age and gender balanced. The demographic information is summarized in Table I.

TABLE I
DEMOGRAPHIC AND CLINICAL INFORMATION

Mean ± SD	Recovery MCI	Progressive MCI	P value
Number	50	44	-
Age	73.81 ± 6.77	75.23 ± 7.15	0.33[a]
Gender(M/F)	27/23	27/17	0.47[a]
CDR_{Y0}	0.5	0.5	-
CDR_{Y3}	0 ± 0	1.13 ± 0.38	-

SD, Standard Deviation; MCI, Mild Cognitive Impairment; CDR_{Yn}, n-year after MR scan session; [a]Two sampled T-test

B. Data Preprocessing and Decomposition

We preprocessed the rs-fMRI using statistical parametric mapping (SPM12, `http://www.fil.ion.ucl.ac.uk/spm/`) by removing first five time points and performing the rigid body motion correction and slice timing correction. We used an echo-planar imaging (EPI) template to fit the rs-fMRI data into standard Montreal Neurological Institute (MNI) space and resampled to $3\times3\times3mm^3$ voxels. The data were smoothed using a Gaussian kernel (FWHM = 5mm), and were normalized to finalize the preprocessing.

Next, we decomposed the preprocessed rs-fMRI with group independent component analysis (GICA) to the independent components (ICs) and the corresponding time-courses (TCs) by adopting the NeuroMark pipeline [10]. Fifty-three pairs of ICs and TCs were selected and arranged into seven functional domains based on the spatial location, and seven domains include subcortical (SC), auditory (AU), sensorimotor (SM), visual (VI), cognitive control (CC), default mode (DM), and cerebellar (CB). In this work, we use z-scored TCs in the analysis.

[1]The OASIS-3 is a public use dataset and is not individually identifiable, the original experimental procedures involving human subjects were approved by the Institutional Review Board.

C. Windowless Wavelet-based dFNC (WWdFNC)

We used a windowless wavelet-based functional network connectivity measure as detailed in [6] to investigate time-varying connectivity between brain networks. The coupling status of networks at each timepoint $t \in\{1,2,\ldots,T\}$ is represented using WWdFNC, a wavelet-based measure. The WWdFNC starts by performing a continuous wavelet transform of each univariate network timeseries $s_k(t)$ using the complex Morlet wavelet at $J = 20$ evenly spaced frequencies. For each univariate network timeseries, this results in a complex-valued multivariate time-frequency domain timeseries (mTFTs), $S_k(t) \in \mathbb{C}^J$.

Assuming we have N samples and k TCs for each sample as input $S = [s_1, s_2, ..., s_k]$, we decompose k–th network's time-courses s_k into $P_k \in \mathbb{C}^{J\times T}$ and let $P_k^{j,t} \in \mathbb{C}$ denote the wavelet coefficient that represents the power and phase at frequency j in network k at time t. The network connectivity was then calculated by taking both power and phase synchrony into account. Power-weighted phase synchrony is used to compute the WWdFNC between k–th and l–th at t:

$$Conn_{k,l}^t = \sum_{j=1}^{20} \frac{p_k^t + p_l^t}{2} cos(\theta_k^t - \theta_l^t) \quad (1)$$

where p_k^t and θ_k^t are the power and phase coefficient for network k at time t, respectively.

D. Time Attention-based LSTM (TA-LSTM)

The attention mechanism was proposed and proved faster and more accurate than other state-of-art models [11], and also was proved efficiently to overcome the vanishing gradient issue in the gradient-based interpretation approach [8].

Meanwhile, we aim to build attention on the high-level time dimension feature after sequence learning instead of the general pooling computation after the attention layer. The time attention layer consists of scaled dot-product attention [11] with half dropout followed by global pooling over the time dimension. The time attention-weighted output c_t from the attention layer is computed by:

$$c_t = \frac{1}{\|e_j\|}\sum_{j=1}^{j}\alpha_{tt} \times e_{tj} \text{ and } \alpha_{tt} = \sigma(\frac{q_{tj} \times e_{tj}^{\top}}{\sqrt{\|e_j\|}}) \quad (2)$$

σ is softmax function, e_{tj} is hidden neural status of the last LSTM layer, q_{tj} is the query matrix produced by a half random dropout from e_{tj}. The TA-LSTM model consists of three LSTM layers with 64 hidden cells, the time-summarized attention layer, and one softmax layer to output the class probability. The architecture is shown in Figure 1.

III. Result

In this section, we report model performance and the interpretation of how the model learned on both temporal pattern and network connectivity level.

A. Model Performance on learning dFNC

We evaluated WWdFNC and SWCdFNC feature representations by training with TA-LSTM model in the ten-fold cross-validation manner. SWCdFNC was computed from NeuroMark preprocessed rs-fMRI in a sliding-window approach with the window size of 20 TR (44s) in steps of 1TR. In each fold, all 94 data were divided into 90% of training and 10% of testing. For every ten-fold cross-validation, each scan was tested once. We repeated ten-fold cross-validation five times with different shuffle parameters, resulting in 50 trials. We used the mean of 50 trials' AUC (Area Under the Curve), accuracy, sensitivity, specificity as the evaluation metrics. The performance is shown in Figure 2. The WWdFNC achieved 0.789 of AUC and 79.3% on the accuracy, increased an average of 0.06 on AUC metric, and 3.1% on accuracy compared to SWCdFNC.

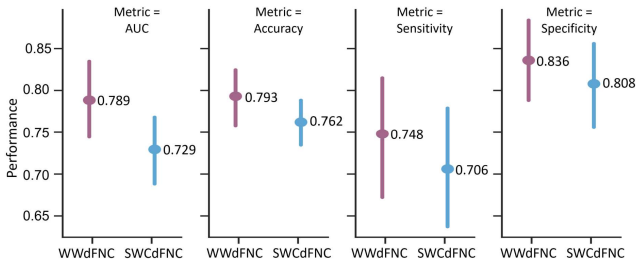

Fig. 2. Mean Model Performance of Prediction of RMCI vs. PMCI

B. Visualization of Most Salient Time Intervals

In the preceding section, we showed that our model trained on WWdFNC features could perform well in predicting the progression/ recovery of MCI patients in the next three years. In this section, we report the interpretation on the group-distinct brain function connectivity biomarkers in temporal patterns level. We constructed the saliency maps using the approach proposed in [7]. It applies the backpropagation algorithm to compute the gradient of the predicted class score y_c (before the Softmax layer) with respect to each input to construct the saliency maps with the same dimensionality as the trained features.

To analyze the temporal patterns from the saliency maps, we first averaged of all the connectivity at each t. We accessed the "strongly-contributing" time points (refer as $\vec{C}$) by thresholding with the upper 90^{th} percentile. The temporally highlighted patterns for both groups are shown in Figure 3(a). In the saliency maps, we observed the continuous highlighted intervals $\vec{C}$ in RMCI tend to last longer than in PMCI. We performed the independent 2-sample t-test between the length of $C_{\vec{RMCI}}$ and $C_{\vec{PMCI}}$, and there is the significant difference between two groups ($C_{\vec{RMCI}}$ = 3.78, $C_{\vec{PMCI}}$ = 2.99, p-value = $7.7e^{-3}$). The kernel density estimates shown in Figure 3(b), show a clearer pattern that $\vec{C}$ greater than or equal to three is the separator point: RMCI has a larger proportion of length of $\vec{C}$ greater or equal to three and verse vice for PMCI. We evaluated three more thresholds that equally range between 0.8 to 0.9 to broaden the range of "strongly-contributing" cutoffs. The results support the same conclusion that the highlighted time intervals in RMCI tend to last longer than in PMCI (Figure 3(c)).

C. Group-wise Connectivity Patterns within Intervals

We performed the statistical analysis on the cellwise properties on the "strongly-contributing" time intervals. Based on the observations reported in the preceding section, we selected all the $\vec{C}$ greater than or equal to 3. The independent samples t-test with multiple comparison correction results shown in middle plot of Figure 4. We also performed two additional statistical analyses for validation tests by applying no thresholding and 10% bottom gradient thresholding, and results are shown in the left and right plots of Figure 4. The global (no thresholding) cellwise plot shows few levels of significant difference; and no significant cells after FDR correction. The 10% lower thresholding that is evaluated for representing the non-contributing cellwise attributes behave in the opposite direction to the "strongly-contributing" and show less significance. The validation tests elucidate that the model is not strengthening the global-wise attributes of group difference for feature learning and show the discriminative temporal patterns that extracted in the saliency maps. With excluding the global cellwise connectivity patterns, we observed that the RMCI group, compared with the PMCI group, shows significant high connectivity in CC-DM, CC-VI, CC-CB, VI-SM, and VI-SC, and low connectivity in VI-VI, VI-CB, CC-CC.

IV. Conclusion and Discussion

Here our approach introduces a feasible pipeline to capture transiently realized connectivity biomarkers for advancing cognitive decline by exploiting the time-dependent feature learning of the TA-LSTM model. The highlighted "highly contributing" time intervals show few overlaps with the global average map, and the global average map has no significant elements after FDR correction, indicating the

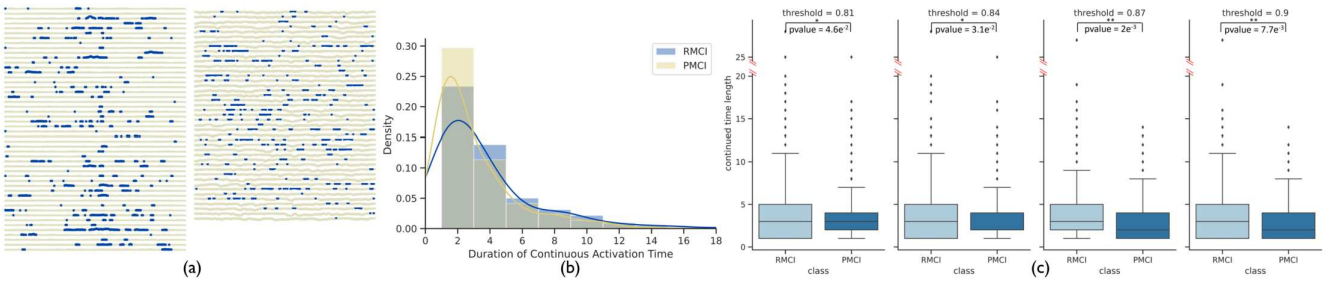

Fig. 3. The saliency maps were constructed on the trial with the highest AUC. (a) The sample-wise saliency maps are first averaged along the connectivity dimension at each time point, and the 10% highest gradient time points are highlighted. The left section shows the RMCI group and right section is for PMCI group. (b) The distribution of all supra-threshold time interval durations. We tested four equally spaced thresholds between $[0.81, 0.9]$ and performed the t-test on the length of supra-threshold intervals. The distribution and p value shown in (c).

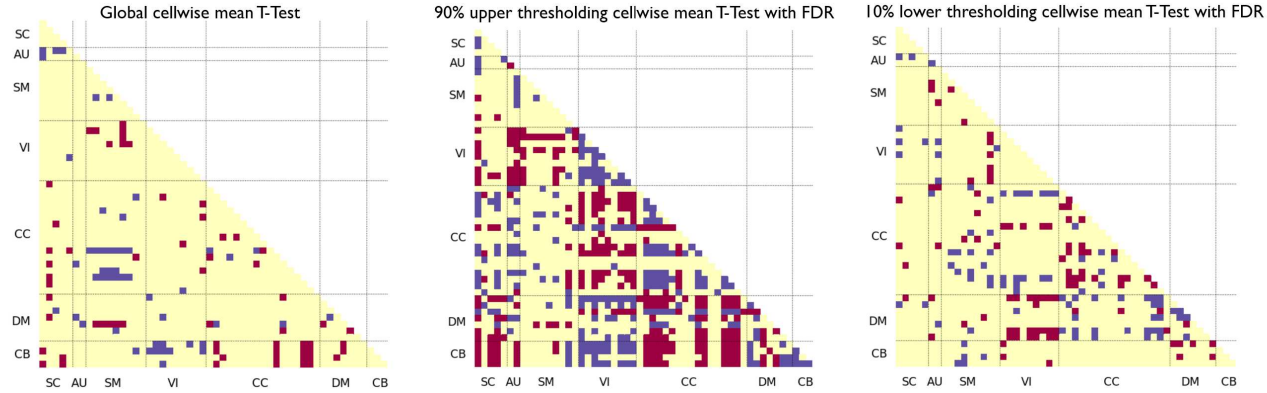

Fig. 4. T-test for differences between RMCI and PMCI in mean cellwise WWdFNC connectivity: all samples (left); within intervals exceeding the 90% upper thresholding for saliency (middle); within intervals under the 10% lower saliency thresholding (right). For the leftmost panel, we applied the no thresholding, for the middle we used intervals of length at least 3 exceeding the 90% upper saliency threshold, and for the left we used intervals of length at least 3 with saliency under the 10% bottom saliency threshold computed from all WWdFNC. We averaged the connectivity features within time intervals and performed the 2-sample T-test with multiple comparison correction (False Discovery Rate Correction $q = 0.05$). Red means the group average of RMCI is significantly greater than PMCI ($p < 0.05(FDR)$), blue means the group average of PMCI is significantly greater than RMCI ($p < 0.05(FDR)$)

group-wise time-varying patterns do not consistently exist in whole scan, and the patterns would be faded out in the average map. The low-vs-high saliency comparison validates time intervals contain different information with few relationships to the final prediction scores. Therefore, capturing the "highly contributing" time intervals over the full scan and investigating the corresponding connectivity patterns expands our knowledge of dynamic biomarkers for the future recovery or cognitive decline of impaired patients. We believe that the power to localize connectivity patterns whose transient appearance on intervals retained through sequential processing with memory in long short-term memory networks warrants continued focus on interpretable LSTMs. Furthermore, the additional accuracy achieved by using "instantaneous" WWdFNCs in this model suggests that greater temporal resolution of the input data can be productively exploited by LSTMs compared with SWCdFNC in which with the corase time resolution, highlighting the importance of continuing to refine our measures of time-varying connectivity.

REFERENCES

[1] A. s. Association, "2016 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 12, no. 4, pp. 459-509, 2016.

[2] S. Gauthier et al., "Mild cognitive impairment," (in English), Lancet, vol. 367, no. 9518, pp. 1262-1270, Apr 15 2006, doi: Doi 10.1016/S0140-6736(06)68542-5.

[3] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı, "The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery," Neuron, vol. 84, no. 2, pp. 262-274, 2014.

[4] R. L. Miller, V. M. Vergara, D. B. Keator, and V. D. Calhoun, "A method for intertemporal functional-domain connectivity analysis: application to schizophrenia reveals distorted directional information flow," IEEE Transactions on Biomedical Engineering, vol. 63, no. 12, pp. 2525-2539, 2016.

[5] Z. Fu et al., "Dynamic state with covarying brain activity-connectivity: on the pathophysiology of schizophrenia," Neuroimage, vol. 224, p. 117385, 2021.

[6] R. L. Miller and V. D. Calhoun, "Transient Spectral Peak Analysis Reveals Distinct Temporal Activation Profiles for Different Functional Brain Networks," in 2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), 2020: IEEE, pp. 108-111.

[7] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," in In Workshop at International Conference on Learning Representations, 2014: Citeseer.

[8] N. Lewis et al., "Can recurrent models know more than we do?," in 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI), 2021: IEEE, pp. 243-247.

[9] P. J. LaMontagne et al., "OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease," medRxiv, p. 2019.12.13.19014902, 2019, doi: 10.1101/2019.12.13.19014902.

[10] Y. Du et al., "NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders," Neuroimage Clin, vol. 28, p. 102375, 2020, doi: 10.1016/j.nicl.2020.102375.

[11] A. Vaswani et al., "Attention is all you need," Advances in neural information processing systems, vol. 30, 2017.