Functional connectome fingerprinting using shallow feedforward neural networks

Gokce Sarar[a,b] , Bhaskar Rao[b], and Thomas Liu[a,c,d,1]

[a]Center for Functional MRI, University of California San Diego, La Jolla, CA 92093; [b]Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093; [c]Department of Radiology, University of California San Diego, La Jolla, CA 92093; and [d]Department of Psychiatry, University of California San Diego, La Jolla, CA 92093

Although individual subjects can be identified with high accuracy using correlation matrices computed from resting-state functional MRI (rsfMRI) data, the performance significantly degrades as the scan duration is decreased. Recurrent neural networks can achieve high accuracy with short-duration (72 s) data segments but are designed to use temporal features not present in the correlation matrices. Here we show that shallow feedforward neural networks that rely solely on the information in rsfMRI correlation matrices can achieve state-of-the-art identification accuracies ($\geq$ 99.5%) with data segments as short as 20 s and across a range of input data size combinations when the total number of data points (number of regions $\times$ number of time points) is on the order of 10,000.

fMRI | connectome | neural network | fingerprinting

Functional connectome fingerprinting based on the similarity of correlation coefficient matrices computed from resting-state functional MRI (rsfMRI) data can identify individuals with high accuracy ($> 98\%$) using long-duration (> 12 min) scans, but considerably lower accuracy ($\sim 68\%$) is obtained when the data duration is decreased to 72 s (1). Recurrent neural networks (RNN) can achieve high accuracy (98.5%) with short duration (72 s) data, presumably reflecting their ability to capture both spatial and temporal features (2, 3). However, it has been shown that high RNN performance can be achieved even when the temporal order of the fMRI data is permuted (4), suggesting that the temporal features are not critical for identification. Here we introduce two shallow feedforward neural networks that can achieve high identification accuracy without the need for recurrent connections. Furthermore, we use these networks to estimate the minimum size of the data needed to robustly identify subjects with high mean accuracy ($\geq 99.5\%$) from short segments of rsfMRI data. Since identification accuracy reflects the ability to effectively extract information from functional connectomes, additional insight into the methods and minimum data sizes that achieve high performance can guide the development of extended approaches to detect other differences in functional connectivity, such as disease-related changes.

The two networks considered are shown in Fig. 1 A and B. The input to the correlation neural network (corrNN) consists of the upper triangular elements of the correlation coefficient matrix C estimated from a data matrix X consisting of z-normalized time series (of length N) from M regions of interest (ROIs). For identification of L subjects, the network structure consists of a fully connected classification layer with L units, a batch normalization layer, and a softmax layer. The norm-based neural network (normNN) uses the z-normalized data X as the input. The first stage is a fully connected layer that projects the data onto K hidden units using the $M \times K$ weight matrix W to form the $N \times K$ intermediate matrix $Y = XW$. In the second stage, the L_2 norm across the time dimension (i.e., across each column of Y) is computed for each hidden unit to form a summary measure of similarity over the collection of N time points. The resulting vector $F = \sqrt{diag(Y^T Y)} = \sqrt{diag(W^T CW)}$ comprises K

features extracted from the correlation matrix C. The kth feature is proportional to the variance in the direction of the kth column vector of W. If these vectors are randomly oriented and constrained to be unit norm, then the features represent a random sampling of the "peanut"-shaped surface of directional variances (5). The subsequent stages in the network are a batch normalization layer, a fully connected classification layer with L hidden units, a second batch normalization layer, and a softmax layer.

Results

We assessed the performance of the two networks using data from the Human Connectome Project (HCP) (6). Two rsfMRI scans acquired on day 1 were used for training, while the two scans from day 2 were used for validation and testing.

For $M = 379$ ROIs, $N = 100$ time points (72-s duration) per segment, and $K = 256$ hidden units, the mean classification accuracies of the corrNN and normNN models were 99.8% and 99.6%, respectively, for an initial set of 100 subjects, and 100.0% and 99.7% for a second independent set of 100 subjects. These accuracies are higher than those reported (94.3 to 98.5%) for RNN models (2, 3). For comparison, the mean classification accuracy using the similarity of the correlation coefficients was 79.4% for 100 time points per segment, which is higher than the 68% mean accuracy reported in ref. 1 using data from a different dataset.

We used a greedy search algorithm to assess the relative importance of the ROIs with respect to model accuracy. Importance maps are shown in the top rows of Fig. 1 C and D for corrNN and normNN, respectively, with the subsequent rows thresholded to highlight the top 15 to 60 ROIs. When considering the top 60 ROIs, the highest numbers of ROIs are found in region 22 (dorsolateral prefrontal cortex) followed by regions 17 (inferior parietal cortex), 14 (lateral temporal cortex), 16 (superior parietal cortex; for CorrNN), 21 (inferior frontal cortex), and 3 (dorsal stream visual cortex), where brain regions are as defined in ref. 7.

We used the top ROIs to evaluate CorrNN and NormNN performance with 15 to 60 ROIs and 5 to 1,000 time points, as shown in Fig. 1 E and F, respectively. As the number of ROIs decreases, the number of time points needed to achieve higher accuracy increases. Defining 99.5% as the threshold for high mean accuracy, we observed that this threshold is surpassed with as few as $M = 60$ ROIs and $N = 100$ time points for CorrNN

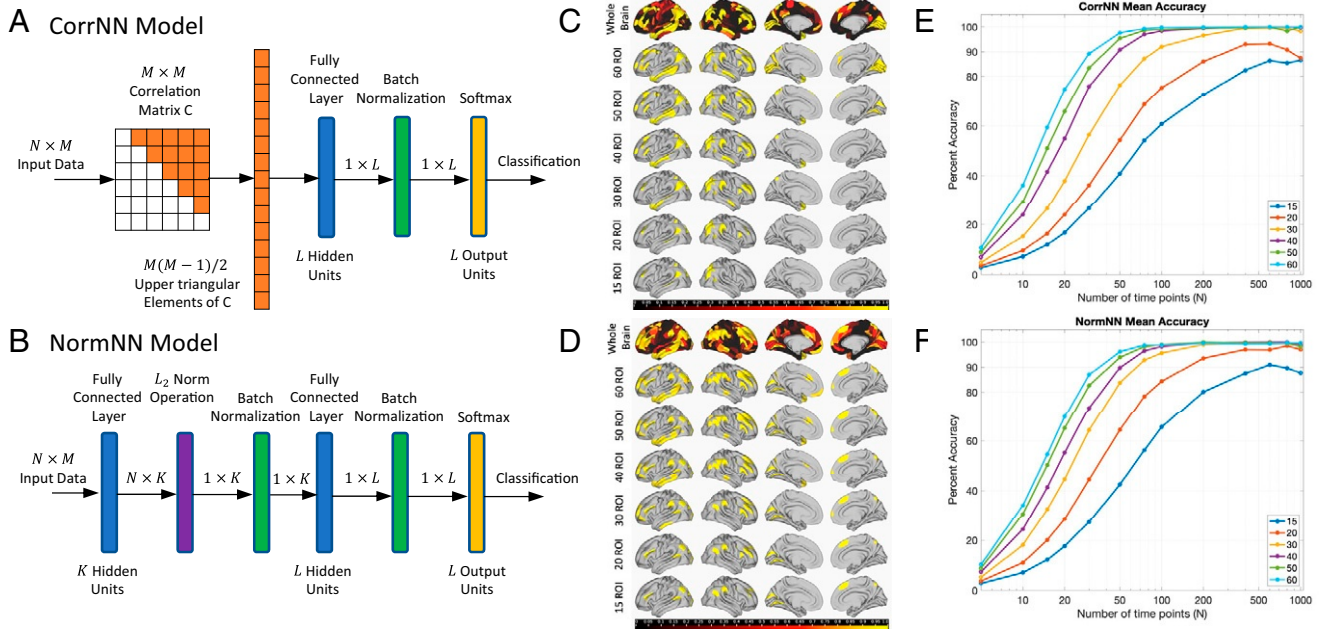[1]To whom correspondence may be addressed. Email: ttliu@ucsd.edu.

Fig. 1. (*A* and *B*) CorrNN and NormNN model structures. (*C* and *D*) Top rows are maps showing the relative importance of the ROIs for identification accuracy, with maximum importance of 1.0 indicated in yellow. The remaining rows are thresholded to show the locations of the top 15 to 60 ROIs. (*E* and *F*) Mean identification accuracies as a function of the number of time points and ROIs.

and 40 ROIs and 200 time points for NormNN, corresponding to $M \times N = 6,000$ or 8,000 total data points, respectively.

To further explore the dependence on the number of ROIs and time points, we considered combinations (M, N) where the total number of data points was constrained to be equal to or close to either 6,000 or 10,000 (see Fig. 2 legend). For CorrNN, high mean accuracies are obtained for two of the combinations (dark red squares) with 6,000 data points and for all five of the combinations (dark red diamonds) with 10,000 points.

For NormNN, the number of parameters exhibits a linear dependence on the number of ROIs (M) as compared to the quadratic dependence for CorrNN (see Fig. 2 legend). To better compare the models, we increased K by powers of 2 up to the value $K_{eq} = 0.5L\left(M^2 - M\right)/(M + L + 3)$ for which the numbers of NormNN and CorrNN parameters were equivalent, while also including K_{eq} as one of the possible options. In Fig. 2*B*, we show NormNN accuracies obtained for either 1) the minimum value of $K \geq 256$ that surpassed the 99.5% threshold or 2) the value $K \leq K_{eq}$ that achieved the highest accuracy when the threshold was not met. High mean accuracies were obtained for two and four of the combinations with 6,000 and 10,000 data points, respectively.

As shown by the histograms, the high mean CorrNN and NormNN accuracies correspond to robust identification performance, with the majority of the trials demonstrating 100% prediction accuracy. These accuracies were obtained with global signal regression (GSR), and were significantly greater than those obtained without GSR for both CorrNN ($\bar{\Delta} = 0.60$; $t_{10} = 5.00$; $p = 0.0005$) and NormNN ($\bar{\Delta} = 0.78$; $t_{10} = 3.93$; $p = 0.0028$) where $\bar{\Delta}$ denotes the mean difference in accuracy. Without GSR, only two of the CorrNN combinations and two of the NormNN combinations exhibited accuracies greater than 99.5%.

Using the ROIs determined from the first 100 subjects, we evaluated performance on the second set of 100 subjects for the combinations denoted in Fig. 2. High mean CorrNN accuracies ($\geq 99.5\%$) were maintained for both of the previously identified high-performance combinations with 6,000 points and for four of

the combinations with 10,000 points, with the remaining combination $(379, 27)$ exhibiting slightly lower accuracy (99.28%) for the second dataset. Thus, the same set of ROIs can offer comparable and high levels of performance across independent datasets.

For both sets of subjects, the mean number of CorrNN prediction errors was not significantly correlated (across subjects) with the mean framewise displacement (FD) measure of subject motion ($|r| < .05$; $p > 0.64$). Correlations were higher but did not reach significance when using a filtered version of the FD measure (*SI Appendix, Extended Methods*), with values of $r = 0.19$ ($p = 0.06$) and $r = 0.13$ ($p = 0.19$) for the first and second subject groups, respectively. When viewed within the context of the high CorrNN accuracies that can be achieved, these results suggest that any effects of subject motion on performance are fairly weak.

For NormNN, we find that the first layer trained weights are randomly distributed so that the features after the L_2 norm operation represent an approximately uniform sampling of the directional variance surface of C. Indeed, high performance can also be achieved by replacing the first layer with a set of random Gaussian weights. The generalizability of the features across datasets exhibits a dependence on the number of units K. For example, when using first layer weights trained using the first set of subjects, performance for the combination $(100, 100)$ with $K = 256$ drops from 99.82% for the first 100 subjects to 98.79% for the second 100 subjects. Increasing to 1,024 units with weights trained using the first set yields accuracies of 99.91% and 99.63% for the first and second sets, respectively. Comparable accuracy levels (99.81% and 99.65%) are obtained when using random weights for the first layer. Thus, generalizability of the NormNN features increases when there is a higher number of features to characterize the directional variance.

Discussion

We have shown that shallow feedforward models can identify subjects based solely on information in rsfMRI correlation matrices, robustly achieving high accuracies ($\geq 99.5\%$) with 6,000 to
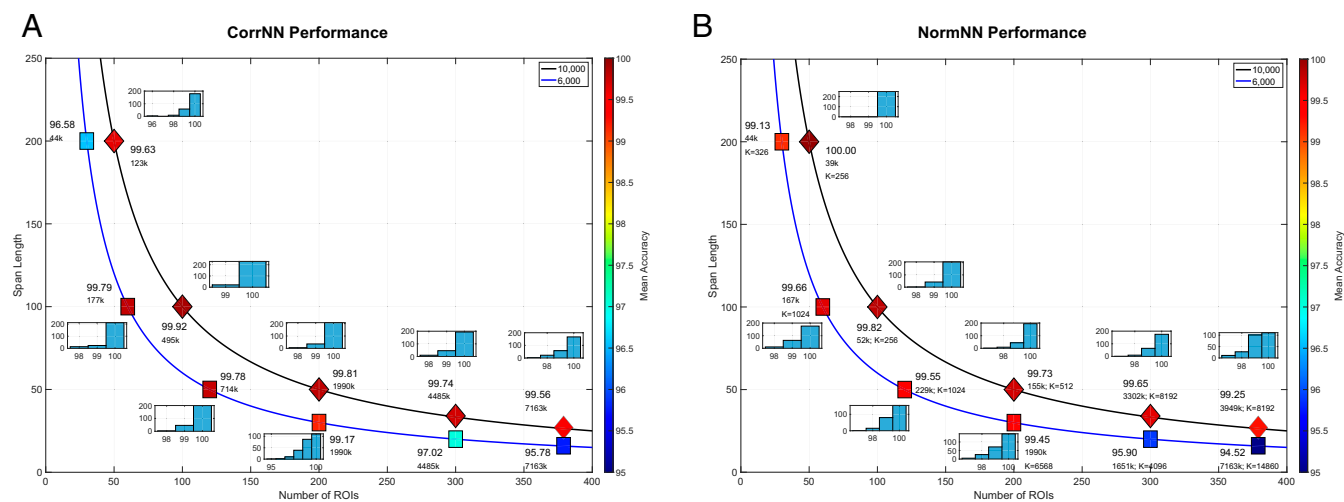
Fig. 2. (*A*) CorrNN and (*B*) NormNN identification accuracies for combinations (*M*, *N*) of numbers of ROIs (*M*) and span lengths (*N*) that are constrained to have either 6,000 (blue curves) or 10,000 (black curves) data points, with the exception of the combinations (379, 16), (300, 34), and (379, 27), which have 6,064, 10,200, and 10,233 data points, respectively. Mean accuracies are indicated by labels and color scale. The numbers of model parameters (in thousands) for CorrNN $\left(0.5L\left(M^2 - M + 6\right)\right)$ and NormNN ($K(M + L + 3) + 3L$) are also listed (with $L = 100$), as are the numbers of hidden units (*K*) for NormNN combinations. For combinations where CorrNN mean accuracy is greater than 99%, autoscaled histograms show the distribution of identification accuracies obtained over 250 test trials per combination.

10,000 data points. For comparison, the convolutional RNN presented in ref. 3 achieved 98.5% accuracy with 23,600 data points. In comparing the two feedforward models, NormNN can attain high accuracy with fewer model parameters, while CorrNN may serve as a better foundation for future work, as it uses correlation coefficient features that are more directly interpretable than the NormNN directional variance features.

Consistent with prior observations (1), high performance can be achieved when using a subset of the ROIs, including those located in frontoparietal and lateral temporal regions. The same set of ROIs can be used to achieve high performance across independent datasets, suggesting that the predictive value of intersubject variability in the functional boundaries and connectivity of these regions generalizes across datasets.

While combinations with span lengths as short as 27 points (19.5 s; CorrNN (379, 27)) can offer high performance, they require a large number of model parameters. In contrast, combinations with fewer ROIs but increased span lengths (e.g., (100, 100)) achieve high performance with one to two orders of magnitude fewer parameters. For NormNN, the number of trainable parameters can be further decreased through the use of random weights in the first layer.

As in prior studies (1–3), the current study utilized the HCP dataset, in which the data were acquired on two consecutive days (6). Although substantial variations in functional connectivity can occur on short time scales (i.e., minutes to hours) due to factors such as temporal fluctuations in vigilance (8), our results indicate that high performance can be obtained over a 1-d interval even in the presence of these factors. Future large-scale studies will be needed to assess whether high identification accuracy can be obtained over longer intervals (i.e., weeks to years).

The effectiveness of the feedforward networks for distinguishing individuals with relatively little data suggests that similar future approaches may have the potential to more fully utilize the information contained in rsfMRI data to better identify disease-related differences.

Materials and Methods

HCP preprocessing of the data included motion correction, detrending, denoising, and registration (7). The 379 ROIs were defined using 360 cortical ROIs from ref. 7 and 19 subcortical ROIs from ref. 6. Data were averaged within each ROI, and GSR was applied. Training, testing, and validation of the models were performed with Keras and TensorFlow. Further details are provided in *SI Appendix, Extended Methods*.

1. E. S. Finn *et al.*, Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
2. S. Chen, X. Hu, Individual identification using the functional brain fingerprint detected by the recurrent neural network. *Brain Connect.* **8**, 197–204 (2018).
3. L. Wang, K. Li, X. Chen, X. P. Hu, Application of convolutional recurrent neural network for individual recognition based on resting state fMRI data. *Front. Neurosci.* **13**, 434 (2019).
4. G. Sarar, S. Wang, J. Ren, T. T. Liu, "Functional connectome fingerprinting using recurrent neural networks does not depend on the temporal structure of the data" in *Proceedings of the 27th Annual Meeting of the ISMRM* (International Society for Magnetic Resonance in Medicine, 2019), p. 3856.
5. D. Bartz, K. Hatrick, C. W. Hesse, K.-R. Müller, S. Lemm, Directional variance adjustment: Bias reduction in covariance matrices based on factor analysis with an application to portfolio optimization. *PloS One* **8**, e67503 (2013).
6. D. C. Van Essen *et al.*, The Wu-Minn Human Connectome Project: An overview. *Neuroimage* **80**, 62–79 (2013).
7. M. F. Glasser *et al.*, A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
8. T. T. Liu, M. Falahpour, Vigilance effects in resting-state fMRI. *Front. Neurosci.* **14**, 321 (2020).
9. T. T. Liu, FC_Fingerprint_NN. Bitbucket. https://bitbucket.org/ttliu07/fc_fingerprint_nn/src/master/. Deposited 24 January 2021.
10. T. T. Liu, FC_Fingerprint_NN. Open Science Framework. https://osf.io/uj2n7/. Deposited 22 March 2021.

Sarar et al.
Functional connectome fingerprinting using shallow feedforward neural networks

PNAS | 3 of 3
https://doi.org/10.1073/pnas.2021852118