Ensemble learning with 3D convolutional neural networks for functional connectome-based prediction

Meenakshi Khosla[a], **Keith Jamison**[b], **Amy Kuceyeski**[b,c], **Mert R. Sabuncu**[a,d]

[a]School of Electrical and Computer Engineering, Cornell University

[b]Radiology, Weill Cornell Medical College

[c]Brain and Mind Research Institute, Weill Cornell Medical College

[d]Nancy E. & Peter C. Meinig School of Biomedical Engineering, Cornell University

Abstract

The specificty and sensitivity of resting state functional MRI (rs-fMRI) measurements depend on preprocessing choices, such as the parcellation scheme used to define regions of interest (ROIs). In this study, we critically evaluate the effect of brain parcellations on machine learning models applied to rs-fMRI data. Our experiments reveal an intriguing trend: On average, models with stochastic parcellations consistently perform as well as models with widely used atlases at the same spatial scale. We thus propose an ensemble learning strategy to combine the predictions from models trained on connectivity data extracted using different (e.g., stochastic) parcellations. We further present an implementation of our ensemble learning strategy with a novel 3D Convolutional Neural Network (CNN) approach. The proposed CNN approach takes advantage of the full-resolution 3D spatial structure of rs-fMRI data and fits non-linear predictive models. Our ensemble CNN framework overcomes the limitations of traditional machine learning models for connectomes that often rely on region-based summary statistics and/or linear models. We showcase our approach on a classification (autism patients versus healthy controls) and a regression problem (prediction of subject's age), and report promising results.

Graphical Abstract
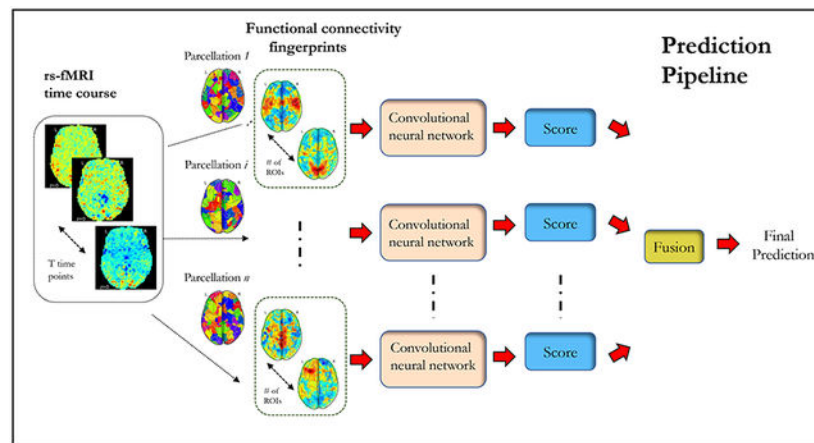
Keywords

Functional connectivity; fMRI; Convolutional Neural Networks; Autism Spectrum Disorder; ABIDE

1. Introduction

Functional connectivity, as often captured by correlations in resting state functional MRI (rs-fMRI) data, has produced novel insights linking differences in brain organization to individual or group-level characteristics. Recently, machine learning models are being increasingly applied to study and exploit individual variation in functional connectivity data [1, 2, 3]. These models often employ hand-engineered features, such as pairwise correlations between regions of interest (ROIs) and network topological measures of clustering, modularity, small-worldness, integration, or segregation [4, 5, 6]. The ROIs are usually computed based on a pre-defined atlas or a parcellation scheme. The choice of the ROIs can have a significant impact on downstream analyses [7, 8, 9].

Brain ROIs can be defined based on macro-anatomical features, cytoarchitecture, functional activations, and/or connectivity patterns [10, 11, 12, 13]. A common approach is to derive the ROIs either based on input from experts and/or using a data-driven strategy on a small number of subjects. Expert-defined ROIs are challenging to standardize across studies [14] and often rely on arbitrary decisions. Data-driven ROIs, on the other hand, can be biased by the selection of the subjects, especially for regions that exhibit large variability across the population. Popular data-driven techniques include clustering, dictionary learning and Independent Component Analysis (ICA) [15, 16, 9]. Such methods can be sensitive to confounds such as motion, while initialization, optimization, and other algorithmic choices can also significantly influence the results [17]. A parcellation scheme not only defines the boundaries of ROIs, but also restricts the analysis to a certain spatial scale. Abraham et al. [18] showed that among various preprocessing decisions, the choice of region definition has the greatest impact on predictive accuracy with data-driven extraction based on dictionary learning outperforming ICA/clustering and other reference atlases.

Given the arbitrary nature of a chosen parcellation scheme and its impact on predictive models, we hypothesized that machine learning models can benefit markedly from an ensemble strategy that integrates across different scales and ROI definitions. Figure 1 shows a general schematic of our proposed framework. In this work, we conducted a thorough empirical evaluation of different choices for brain parcellations.

Another important factor in connectome-based machine learning pertains to the choice of the classification algorithm. A large body of related work in the literature has focused on simple linear predictive models using vectorized connectivity data. A relatively recent trend is to exploit neural networks for graph-structured data, such as Graph Convolution Networks or BrainNet-CNN, to make individual-level predictions on connectomes. Ktena et al. [19] applied spectral graph convolutions in a distance-metric learning framework to train a k-nearest neighbor classifier on connectivity data. In a similar vein, Kawahara et al. [20] proposed the BrainNetCNN architecture that extends convolutional neural networks (CNNs) to handle graph-structured data. CNNs are motivated via the translation-invariance property of image-based classification problems and can exploit voxel/pixel resolution data. On the other hand, BrainNetCNN works directly with an adjacency matrix derived from the connectome data, while disregarding spatial information. The model parameter count would scale according to the number of ROIs, making the utilization of voxel-level connectivity infeasible with this approach. As we discuss below, we propose an alternative representation of connectivity data, which allows us to leverage modern deep learning architectures, like CNNs, to build a prediction model that exploits the full-resolution 3D spatial structure of rs-fMRI without having to learn too many model parameters.

In this work, we consider two applications: discrimination of autism patients and healthy controls; and regression of age. The first problem is a particularly challenging one. Several previous studies have reported altered functional connectivity patterns in Autism Spectrum Disorder (ASD) patients [21, 22, 23, 24]. While studies using small samples have reported classification accuracies over 75% [25], application of similar models on large heterogeneous datasets, such as ABIDE [26], have shown more modest performance levels over a wide range of connectome preprocessing schemes (accuracies that range 60–67%) [18].

Our main contributions in this paper are:

- An extensive evaluation of the influence of brain parcellations on functional connectome-based machine learning models

- An ensemble learning strategy for combining predictions from multiple classifiers corresponding to different brain parcellations

- An easy-to-implement 3D CNN framework for connectome-based classification

2. Materials and Methods

2.1. Dataset

The Autism Brain Imaging Data Exchange (ABIDE) is a multi-site consortium aggregating and openly sharing anatomical, functional MRI and phenotypic datasets of individuals

diagnosed with ASD, as well as healthy controls (HC) [26]. The first phase of ABIDE (ABIDE-I) collected data from 1,112 individuals, comprising 539 individuals diagnosed with ASD and 573 typical controls across 17 sites. The second phase (ABIDE-II) aggregated 1,114 additional datasets, comprising 521 individuals with ASD and 593 healthy controls across 19 sites.

2.2. Preprocessing of fMRI Data

The Preprocessed Connectomes Project (PCP) released preprocessed versions of ABIDE-I using several pipelines [27]. We used the data processed through the Configurable Pipeline for the Analysis of Connectomes (CPAC). This pipeline performs motion correction, global mean intensity normalization and standardization of functional data to MNI space ($3\times3\times3$ mm resolution) before the extraction of ROI time series. Among the different strategies in the release, our analysis used data de-noised by regression of nuisance signals including motion parameters, CompCor WM+CSF components, and global signal, followed by band-pass filtering (0.01–0.1Hz). We note that we have experimented with alternate preprocessing strategies that include/exclude the global signal regression and CompCor steps. These results are presented in the Supplementary Section 7.6.

We preprocessed the ABIDE-II dataset following the same sequence of steps listed for ABIDE-I in CPAC (using the version v1.0.2a). Since manual quality control (QC) was not yet available for ABIDE-II, we performed an automatic QC by selecting those subjects that retained at least 100 frames or 4 minutes of fMRI scans after motion scrubbing [28]. Motion scrubbing was performed based on Framewise Displacement (FD), discarding one volume before and two volumes after the frame with FD exceeding 0.5mm [29].

2.3. Cohort selection

In our experiments, we used ABIDE-I subject data that passed manual QC by all the functional raters. This yielded a final sample size of 774 ABIDE-I subjects, comprising 379 subjects with ASD and 395 typical controls. As an independent test dataset, we employed ABIDE-II subjects from sites that participated in ABIDE-I and used the same MRI sequence parameters for data collection. After automatic QC, we ended up with a final ABIDE-II sample size of 163 individuals with ASD and 230 healthy controls. For age prediction, we only considered healthy controls. Furthermore, subjects whose age were more than 3.5 standard deviations away from the median were excluded from the task of age prediction. Table 1 summarizes the dataset characteristics for the two prediction tasks considered in this study.

2.4. Extracting ROI time series from atlases

In our experiments, we considered all atlases that were used for ROI time series extraction in PCP. These include the following seven atlases: Talaraich and Tournoux (TT, R=97), Harvard-Oxford (HO, R=111), Automated Anatomical Labelling (AAL, R=116), Eickho-Zilles (EZ, R=116), Dosenbach 160 (DOS160, R=161), Craddock 200 (CC200, R=200), and Craddock 400 (CC400, R=392), where R is the number of ROIs [30, 31, 32, 33, 34, 35, 36, 37, 38].

For our 3D CNN model, described below, the parcellated regions were used as target ROIs to derive the input connectivity features at the voxel level. For the non-CNN benchmark models, also described below, each atlas was used to define a corresponding connectivity matrix which was fed as input to each model after collapsing into a vector. We report results for ensemble learning strategies as well, where we combined the predictions of models corresponding to individual atlases.

2.5. Creating stochastic parcellations

Stochastic parcellations were created by Poisson Disk Sampling using the method described in [39]. Given a number of ROIs, this approach divides the gray matter voxels (as defined by a given mask) into roughly equal-sized parcels while ensuring that the parcels do not cross hemisphere boundaries. Stochasticity is introduced in the ROI center locations, and all the remaining voxels are assigned to the closest region center. These centers are kept a minimum distance apart based on the desired number of regions in the parcellation. Further details about the sampling approach are provided in Supplementary Section 7.2. All parcellations were created in the MNI152 template at a 3mm resolution, same as the resolution of the preprocessed functional data. For creating these parcellations, we relied on a whole brain gray matter mask including sub-cortical structures. To create the mask, we took the union of the gray matter tissue prior provided in the standard MNI152 template and the cortical mantle mask used in [16]. Some example stochastic parcellations are shown in Figure 2 against atlases at similar resolutions.

2.6. 3D Convolutional Neural Network Approach

Here, we present our novel strategy to adopt a 3D CNN architecture for use with connectomic data.

Loosely reminiscent of the biological visual system, CNNs use spatially localized filters to detect local image features. Unlike fully connected layers where every unit is connected to all other units of the previous layer, convolutional layers employ a structured arrangement where each unit is connected to only a small subset of spatially connected units in the input image channels. Further, the weights of these connections are shared between the units of the convolutional layer so that the same feature can be detected regardless of its spatial location. Mathematically, a convolutional layer of the form $Y=O_W(X)$ operates on an M-dimensional input $X(v)=(X_1(v),\ldots,X_M(v))$ by applying a set of filters $\{W=\{w_{m,n}\}, m=1,\ldots,M; n=1,\ldots,N\}$. Here, $\mathbf{v}$ used to index the pixel or voxel is (in case of 3D convolution). After applying an elementwise non-linearity ϕ (such as a logistic function), this produces an N-dimensional output $Y(v)=(Y_1(v),\ldots,Y_1(v))$. Each element $Y_n(v)$, known as a feature map, is thus given as,

$$Y_n(v) = \phi\left(\sum_{m=1}^{M}\left(X_m * w_{m,n}\right)(v)\right), \quad (1)$$

where * denotes the standard spatial convolution operation. The convolutional layers in CNNs are often interspersed with pooling layers that reduce the size of feature maps and

offer translation invariance. Max-pooling is the most popular pooling operation. It down-samples each input feature map (commonly referred to as a channel) separately by selecting the maximum feature response in pre-fixed local neighborhoods. A max-pooling $Y_i = P(X_i)$ operation on channel i is thus defined as, $Y_i(v) = \text{Max}\left(X_i(\bar{v}): \bar{v} \text{ in neighborhood of v}\right)$. In 3D, for example, the neighborhood can be a $3 \times 3 \times 3$ cube around each voxel. The convolutional and max-pooling layers form the backbone of a CNN. A CNN architecture is constructed by combining multiple layers that successively learn more complex features from the input images. For example, with L layers the output can be mathematically expressed as $(O_{W(L)}, \dots P \bigcirc O_{W(1)})(X)$. Since we are considering an image classification problem, we add fully connected layers to the flattened output at the end of a CNN.

Research in visual recognition has shown that fully connected feedfor-ward architectures dont scale well to full images. Instead, neural network architectures with local connectivity, such as CNNs, are much more suitable when dealing with high-dimensional images. The shared weights of the CNN architecture facilitate learning with fewer parameters. 3D Convolutional layers thus transform an input 4D (3D multi-channel) volume to an output 4D volume. Each layer learns a set of spatial filters that activate in response to distinct visual patterns. Replicating or convolving each filter across the volume allows the corresponding pattern to be detected irrespective of its spatial location. Finally, the outputs from all filters are stacked along the 4th dimension to create a 4D feature map. Multiple convolutional layers coupled with pooling operations create global representations from local patterns. Stacking fully connected layers at the end after convolutional and down-sampling operations dramatically reduces the model parameter count for classification.

In our proposed approach, the input to the CNN is formed by concatenating voxel-level maps of "connectivity fingerprints", which are represented as a multi-channel 3D volume. Each channel is a connectivity feature, such as the Pearson correlation between each voxel's time series and the average signal within a target ROI. In our implementation, we use both atlas-based and stochastic brain parcellation schemes to define target ROIs. The total number of input channels thus represents the number of ROIs used for creating voxel-level fingerprints. For each parcellation scheme (atlas-based or stochastic), we trained a separate model.

In our experiments, we employed a simple CNN architecture, illustrated in Fig. 3. Our architecture has several convolutional layers, interspersed with max-pooling based down-sampling layers, followed by a couple of densely connected layers. The models were trained with a mini-batch size of 64, until convergence of validation loss. For classification, we used binary cross-entropy, whereas for regression we adopted mean squared di erence as the loss function. The neural network weights were optimized via stochastic gradient descent (SGD) for classification and Adam for regression. The learning rate and momentum for SGD were set to 0.001 and 0.9 respectively. Learning rate of Adam was set to 0.0005. For age regression, we employ a stochastic weight averaging strategy where we average the neural network weights over last 20 epochs. The same architecture and settings were used for all atlases and stochastic parcellations. We note that each atlas is defined on a unique gray matter mask. To ensure that all prediction models (benchmark and proposed) relied on

information from the same voxels, the atlas-specific gray matter mask was applied to the voxel-level connectivity fingerprint data before feeding into the proposed convolutional architecture. For stochastic parcellations, the custom gray matter mask as described above was used for masking the fingerprints. The code and stochastic parcellations have been made available at: https://github.com/mk2299/Ensemble3DCNN_connectomes.

2.7. Benchmark Methods

In our experiments, we implemented following benchmark methods.

2.7.1. Ridge Regression—A linear regression model was trained with squared loss and a times the squared norm of the weight vector (See Appendix). For classification, the ground truth labels were encoded as ±1 for the two output categories. We tested 10 linearly spaced values for thehyper-parameter a in the range[0.1,10] and report for the value with the highest cross-validation accuracy.

2.7.2. Support Vector Machine—We implemented a standard SVM as a benchmark (See Appendix). We found that a radial basis function (RBF) kernel performed better than a linear model. Thus we report results for the RBF-kernel SVM. The two hyper-parameters (RBF kernel width γ and and misclassification cost weight C) were fine-tuned by maximizing cross-validation accuracy via a grid search. For regression, we implemented the standard SVR scheme with an ϵ-insensitive loss function, optimizing for the ϵ-tube and penalty parameter of the error term via grid search.

2.7.3. Fully Connected Architecture—The fully-connected neural network (FCN) architecture takes as input functional connectivity estimates between pairs of ROIs, which is vectorized and processed by a feed-forward network. We implemented following architecture, which performed best on ABIDE-I cross-validation: 4 fully connected hidden layers, with 800, 500, 100 and 20 numbers of features and each linear layer followed by an elementwise Exponential Linear Unit (ELU) activation. Dropout regularization parameter was set to 0.2 and applied to each layer during training. For classification, the output node was a sigmoid, and cross-entropy loss was used. For age prediction, the sigmoidal output was replaced with a linear activation and mean squared di erence was used as the loss function. The models were trained with a mini-batch size of 64, until convergence of validation loss.

SGD was used as the optimizer with learning rate and momentum set to 0.01 and 0.9 respectively for classification. For age prediction, a smaller learning rate of 0.001 was used.

2.7.4. BrainNet Convolutional Neural Networks—BrainNet CNN, originally proposed in [20], utilizes specialized kernels to handle connectomic data. Their work described novel edge-to-edge, edge-to-node and node-to-graph convolutional layers that can potentially capture topological relationships between network edges. For BrainNet CNN, we implemented the following architecture that worked best on ABIDE-I cross-validation: 1 edge-to-node layer with 256 filters, followed by a node-to-graph layer with 128 output nodes and finally a dense layer with single output. A leaky ReLU non-linearity with alpha equal to 0.33 was applied to the output of each layer except the last layer. The activation of the last

layer was set to linear and sigmoid for the regression and classification tasks, respectively. Dropout regularization with rate 0.2 was used for the edge-to-node layer. Similar to [20], Euclidean loss was minimized for age regression, whereas cross-entropy loss was used to optimize the classification models. The models were trained for 1000 iterations using SGD with momentum equal to0.9. The learning rate was set to 0.0005 for age prediction and 0.008 for ASD/Healthy classification. The training curves were monitored for atlases to ensure convergence.

2.8. Ensemble Learning

In our experiments, we explored two ensemble learning strategies. The first one is what we call multi-atlas ensemble (or MA-Ensemble). MA-Ensemble averages the predictions of the models of a specific method (e.g., BrainNet CNN) computed using each one of the seven atlases. For classification, the final prediction is computed as the majority vote of the individual binary class predictions. For regression, the ensemble prediction is simply the mean. The second ensemble strategy (SP-Ensemble) averages across the models of a specific method computed using stochastic parcellations. In our experiments, unless stated otherwise, we used 30 stochastic parcellations at each of the following four spatial scales: 110, 160, 200 and 400 ROIs. These scales were chosen in accordance with existing atlases. Thus the SP-Ensemble's prediction was computed based on fusing 120 (30 $\times$ 4 scales) models. We also implemented single-scale SP-Ensemble models, which averaged over the 30 parcellations at the same spatial scale.

2.9. Visualizing the CNN model

In order to understand the connectivity features captured by the CNN model, we employed the saliency map approach of [41]. This visualization technique computes the gradient of the output prediction with respect to the input image voxel values, i.e., the 3D volume, using a single backward pass through the trained neural network. We then computed voxel-level saliency as the maximum absolute gradient value across all input channels corresponding to different target ROIs. More formally, consider an input image I, representing the connectivity fingerprints of V voxels with R ROI signals. The saliency weights $w \in \mathbb{R}^{V \times R}$ are computed by taking the absolute value of the gradient of neural network output O with respect to the input image, i.e., $w = \left|\frac{\partial O}{\partial I}\right|$. In order to obtain the saliency at the voxel level $S \in \mathbb{R}^V$, we take the maximum across all the ROIs, i.e., $S_i = \max_{1 \; j \; R} w_{ij}$. Finally, to visualize an ensemble model, we averaged the individual saliency maps that made up the ensemble.

3. Results

3.1. Experiments

In our experiments, we considered two tasks: i) binary classification of autism vs healthy, and ii) age prediction. For each task, we implemented two evaluation schemes. First, we conducted 10-fold cross-validation on the ABIDE-I dataset, so that we could present results that were comparable to previously reported classification results such as [1, 18]. Second, we trained each model on the entire ABIDE-I dataset and computed test performance on the independent ABIDE-II set. We report classification accuracy and the receiver operating

curves (ROC), along with corresponding area under the curves (AUC) for each of these scenarios under various combinations of parcellation schemes and prediction algorithms. For age prediction, we report the root mean squared error (RMSE).

3.2. Evaluation of Prediction Performance

Table 2 shows the independent test performance for different models on the classification problem. The proposed 3D CNN approach performs at least as good as, and often better than, the benchmark methods, including the fully-connected deep neural network (FCN) and BrainNetCNN. In particular, the 3D CNN approach performs favorably against other algorithms for all but two parcellation schemes, including the ensembles. Similarly, the SP-Ensemble achieves the best ABIDE-I cross-validation for most algorithms, including the 3D CNN. The ABIDE-I cross-validation results, reported in Table S.2, are in general compatible with the independent test results, where the 3D CNN and SP-Ensemble techniques mostly outperform the competition. Figure 4 shows the Receiver Operating Characteristic (ROC) curves for SP-Ensemble models for the different algorithms on the independent ABIDE-II test dataset. We observe that the 3D-CNN SP-Ensemble achieves an AUC of ~77% and an accuracy of ~72% on independent ABIDE-II data, slightly better than the state-of-the-art cross-validation on ABIDE-I for ASD/HC classification [24], with FCN and Brain-Net CNN ensembles yielding a similar performance. ROC Curves for individual atlases are shown in Figure S.4.

Table 3 lists independent test results for the age prediction task on ABIDE-II, and Table S.3 reports the 10-fold cross-validation error on ABIDE-I. The 3D CNN approach consistently shows superior performance, yielding the best results for all parcellation schemes. Similar to the classification scenario, SP-Ensemble or MA-Ensemble also yield the best cross-validation and independent test performance values for the majority of the algorithms, including 3D CNN. Overall, the best accuracy is achieved by SP-Ensemble 3D CNN, which yields a root mean squared error of **3.28** years on ABIDE-I cross-validation and 2.15 years on the independent ABIDE-II dataset. We also estimated mean absolute error (MAE) of all models on ABIDE-II and observed a similar trend, as reported in Table S.6.

3.3. Comparison of stochastic parcellations and atlases

Here, our objective is to conduct a detailed investigation of how the choice of ROIs affects prediction performance for different machine learning (ML) algorithms. For each ML algorithm and each parcellation we have a model trained on the ABIDE-I data, which we then used on the independent ABIDE-II data to quantify prediction accuracy. Figure 5 shows the distribution of accuracy values (estimated with a kernel density model) obtained using stochastic parcellations, while also illustrating the results for each of the atlases and the scale-specific SP-ensembles. The scale-specific SP-Ensemble strategy, as the name implies, averaged the models corresponding to the 30 stochastic parcellations in each scale. We observe that the atlas-based models performed no better than typical stochastic parcellation models, independent of scale and algorithm. This result offers an intriguing possibility: perhaps we do not need anatomically or functionally derived brain parcellations to train machine learning models since stochastic parcellations perform equally well or no worse in practice.

Our proposed SP-Ensemble CNN strategy yielded accuracy results that were about as good as the best scale-specific SP-Ensemble model. Finally, the ensemble models were almost always better than the atlas-based models and they compared favorably against the individual stochastic parcellation models. The same observations can be made for ABIDE-I cross-validation (see Supplementary Figure S.1).

In above analysis, one potential confound was the different gray matter masks of atlases and stochastic parcellations (SPs). In order to account for this confound, we conducted following analysis. For each of the atlases, we generated 100 SPs using the same gray matter mask as the atlas. We excluded DOS160 because it does not rely on a well-defined gray matter mask and places discontiguous 4.5 mm spherical regions over fixed coordinates in the brain (sampling only 5% of brain voxels). We then trained on each of these SPs using the same hyper-parameters that were found to be optimal for the corresponding atlas. Here, we show the results for ridge regression (the model that was fastest to train), but we obtained similar results for all other algorithms as well. As can be seen from Figure 6, for most atlases and corresponding gray matter masks, the model trained on the atlas ROIs performed no better than an average SP model. Furthermore, and importantly, the SP-Ensemble (computed by averaging across SPs on the atlas-specific mask) yielded better performance than the atlas models for all atlases.

3.4. Visualization

An important goal of machine-learning tools in neuroimaging is to generate novel insights linking imaging biomarkers with disease or phenotypic traits. Visualization techniques for CNNs can help reveal important features used by the model for discriminating between output classes. Figure 7 shows the saliency maps computed for the SP-Ensemble CNN ASD classification and age prediction models. As can be seen from these maps, the precuneus, often considered a core node of the default mode network [42], seems to play a significant role for both prediction problems. However, there are also salient regions that are unique to each problem. For example, the anterior cingulate/ventromedial prefrontal cotex, a region that has been linked to autism [43], was distinctly highlighted for the ASD classification problem. The left parietal cortex was also emphasized for ASD prediction, which is consistent with the laterilized activation observed in this region in Autism patients [44]. On the other hand, for age prediction, the left dorsolateral pre-frontal cortex (dlPFC) is a uniquely salient region. The dlPFC is associated with executive functions, such as working memory and abstract reasoning. For working memory, dlPFC's function seems to be age-associated and more lateralized in younger adults [45].

4. Discussion

In this study, we presented a detailed empirical analysis of how the choice of ROIs can impact the performance of machine learning models trained on functional connectomes. We considered several machine learning algorithms, together with a range of spatial scales and parcellation schemes, including the popular atlas-based techniques and a stochastic approach. Our analysis suggests that using a single atlas for summarizing the connectome data is often sub-optimal for training machine learning models, and significantly more

accurate predictions can be achieved with an ensemble approach that averages across models trained with different parcellation schemes. Furthermore, we demonstrated that averaging across stochastic parcellations can achieve very high accuracy values, often surpassing atlas-based models. Our findings resonate with several other studies that compare stochastic parcellations and atlases, although in different contexts. Craddock et al. [46] compared spatially constrained functional parcellations obtained from spectral clustering with anatomically constrained parcellations produced from random clustering. Random parcellations performed as well as functional parcellations and better than anatomical atlases on metrics of cluster homogeneity and representation accuracy. Based on this, the study reflected that su ciently small ROIs perform well for functional network analysis regardless of their spatial position. Fornito et al. [47] generated stochastic parcellations by randomly sub-dividing the AAL atlas and showed that functional organizational properties are independent of the parcellation template at the same network resolution, although significant variability is observed across scales. Studies on di usion-MRI based anatomical networks have similarly shown that topological attributes and network organizational parameters are consistent across different parcellation schemes, including random parcellations [48, 39].

Another main contribution of this study is a novel approach to employ a 3D CNN architecture on functional connectivity data. Convolutional neural networks achieve state-of-the-art performance on many image-based prediction tasks, as they take advantage of the full spatial resolution of the data and the translation invariance property of the problem. Our proposed approach treats voxel-level connectivity fingerprints as input channels to a conventional 3D CNN framework. Spatial convolutions can capture local structural or topographic patterns in the data, such as connectivity gradients. Successively stacking convolutional layers in our architecture would hierarchically yield higher-order features that can capture information relevant for classification. Studies have shown that individual-level network topography serves as a fingerprint of human behavior [49]. Our multi-channel input image comprising connectivity fingerprints, coupled with CNNs, provides a natural framework to capture individual-level differences in topography as they relate to behavior or disease. This strategy contrasts with current practice where the input to machine learning models are pairwise ROI functional correlations. This makes the model more susceptible to uncertainty caused by parcellation choice. This can be seen in our experiments where there is relatively larger variance in prediction performance across atlases for the fully connected neural network. Thus, CNNs with connectivity map inputs can o er a more robust alternative to classification approaches that only rely on ROI-level connectivity information, such as the BrainNet-CNN. Our results demonstrate that when tailored for connectomes, CNNs o er a promising opportunity to probe brain networks in disease.

Machine learning practitioners have to make a number of preprocessing choices in extracting connectomic features to analyze. While there is no onesize-fits-all solution across different tasks, in the context of machine learning models of functional connectivity, we present some interesting empirical observations below.

4.1. Ensemble learning

The motivation behind using multiple stochastic parcellations for prediction is grounded in the concept of ensemble learning. The core idea is to integrate out a latent variable (i.e., parcels or ROI definitions) from the learning problem [50]. This approach also makes the predictions more robust to the precise parcellation scheme. As shown above, the performance of atlas-based models can vary significantly (~ 5–10% for parcellations at the same scale). In such a scenario, ensemble learning over multiple stochastic parcellations can be a robust strategy that yields reliable predictions.

4.2. Network granularity

We explored the impact of network granularity on prediction performance of machine learning algorithms for connectomes. Our analysis suggests that better prediction performance can be expected with parcellations at higher granularity upto ~ 400 ROIs. To further investigate this trend on ROI-level models, we trained the fully-connected network (FCN), that is generally the best performing baseline algorithm, on both the prediction tasks for the 1024 node parcellation proposed in [48]. As can be seen from Table 4, an atlas with 1024 regions is comparable to the CC200 atlas for ASD/HC classification in ABIDE-II. However, the performance actually degrades significantly (in comparison to CC200 or CC400) for the age prediction task.

Our evaluations contradict with a previously reported result that a coarser network scale (~ 100–150 ROIs) is more suitable for autism classification [18]. In their paper, these conclusions were drawn by comparing the performances achieved with a few atlases. However, inferring trends from a small number of atlases can be misleading, since factors like the boundary definitions of structures (cortical/subcortical) or the particular gray matter mask used, will effect results. Stochastic parcellations can control for these confounds and depict unbiased trends across network scales.

4.3. Number of gray matter voxels

Our empirical study suggests that there is no direct correlation between the number of voxels in the gray matter mask and a model's prediction performance. However, we do observe that the choice of gray matter mask can impact results. For example, the DOS160 atlas with as few as ~ 3,039 voxels shows performance no worse than other atlases at the same resolution(HO, EZ, TT and AAL) with ~ 20× more voxels.

4.4. Visualization

Saliency maps provide a valuable visualization strategy to probe deep neural network models. We visualized the saliency maps from 3D CNN models trained on ROIs extracted using both atlases and stochastic parcellations. As shown in Figure 7 and Supplementary Figures S.2 and S.3, these maps are remarkably consistent.

These maps reveal that the precuneus, which is a hub of the default mode network and associated with ASD and age, plays an important role for both prediction problems. There were also uniquely highlighted regions, such as the anterior cingulate/ventromedial prefrontal cortex for ASD classification and the left dorsolateral prefrontal cortex (dlPFC)

for age prediction. Several studies have suggested the potential of DMN connectivity as a neuropheno-type of autism. Chen at el. [51] trained a random forest classifier that distinguished ASD subjects from healthy controls with high accuracy, and showed that default mode and somatosensory regions contribute significantly to diagnostic accuracy. Similarly, Abraham et al. [18] revealed discriminative connections in the DMN for ASD/HC classification within a larger heterogeneous cohort of the ABIDE dataset. Furthermore, it has been shown that the connectivity of posterior cingulate cortex (PCC) and aberrations in the medial prefrontal cortex node of the DMN can predict social deficits in children with ASD [52]. Our results corroborate the findings of these studies, and suggest a crucial involvement of DMN in autism.

4.5. Influence of motion

Several studies have shown differences in head motion parameters during fMRI between healthy controls and diseased populations, or between subjects from different age groups [53, 54]. This, in turn, can manifest as artifacts in the derived resting-state connectivity [55]. Although our independent test data was motion scrubbed, we performed additional analyses to rule out the confounding effect of motion in classifier decisions. We selected a cohort of 151 ASD subjects with motion-matched healthy controls from our independent dataset and analyzed the correlation of 4 motion parameters with classifier predictions. These include the root-mean-square framewise displacement, mean relative displacement, maximum absolute displacement and the number of micro-movements greater than 0.5mm. These summary statistics were chosen in accordance with previous reports of motion artifacts in rsfMRI[28]. As shown in Figure 8, no significant correlations were observed between motion variables and the predictions of SP-Ensemble (model average over all atlases). In this motion-matched cohort, classification accuracy of 71.8% was obtained using 3D-CNN.

For our regression task, there was no significant correlation between a subject's age and any of these motion parameters in our cohorts.

4.6. Recommendations

Based on our experiments, we make two claims in this study: (a) 3D-CNN performs favorably compared to alternative baseline algorithms, and (b) Ensemble models that average across parcellation schemes consistently perform better than individual atlas-based models and are thus a safer choice for supervised machine learning on connectomes. This is because individual atlases can show significant variability in classification/regression performance and finding the optimal atlas for a prediction task among the wide range of available atlases might not be feasible. Figure 9 shows the probability density estimates for the diﬀerence in performance between (a) 3D-CNN versus baseline algorithms as evaluated with the SP-Ensemble strategy, and (b) SP-Ensemble versus single atlas implemented with the 3D-CNN model. These estimates are presented for both our prediction tasks. For this experiment, we estimate the evaluation metrics (AUC-ROC for ASD/HC classification and RMSE for age regression) on 10,000 bootstrapped samples from ABIDE-II. These results demonstrate that the SP-Ensemble approach consistently achieves an accuracy as good as the best performing single-atlas model. Further, the 3D-CNN model consistently outperforms the baseline algorithms for the age prediction task, with more prominent

improvements for individual atlas models. This can be seen from Tables 2 and 3. We note that when using the ensemble strategy, the differences between models are marginal and might be irrelevant in some practical applications. For instance, the SP-Ensemble performance on ASD/HC classification task is comparable among 3D-CNN, FCN or BrainNet-CNN, with slight improvements over linear models. Thus, if time and/or computational resources impose constraints, it might be more suitable to prefer simpler models like FCN or SVM over 3D-CNN for example, especially with the ensemble approach.

4.7. Limitations and future work

Throughout our analysis, Pearson's correlation was chosen to measure functional connectivity strength between different brain regions. Several other correlation metrics, including tangent-based and partial correlation have been shown to yield superior classification performance in prior studies [9, 18]. While we do not expect this to affect the general conclusions and findings of our study, the choice of the correlation metric still remains an arbitrary decision in any machine learning pipeline for connectomes.

Due to the heavy computational burden required for training multiple deep learning models, we only considered one particular scheme for creating stochastic parcellations, i.e., Poisson Disk Sampling. Alternative strategies for creating random parcellations have also been proposed, for instance, through stochastic sub-division of anatomically derived ROIs into smaller parcels [56]. It is also possible to randomize several other more popular schemes for parcellating the brain, such as, using Ward's clustering on functional data from sub-samples of the population [50] or creating Geometric parcellations with different initializations [13].

While the proposed CNN approach achieves promising accuracy on autism detection and age prediction, there is room for further improvement. We have not yet conducted a comprehensive optimization of the convolutional architecture. Furthermore, there are likely more optimal choices than target ROI-based correlations that are used as input to the model. An interesting alternative would be select random gray matter vertices for connectivity profiling, as proposed in [16]. We envision an end-to-end learning strategy that can enable the optimization of these connectomic features.

Saliency maps provide an appealing visualization technique by mapping the neural network activations back to input voxel space. Several modifications to gradient-based back-propagation have been reported in literature that can potentially highlight more informative features learnt by the model [57, 58]. Further, the use of saliency maps need not be restricted to depicting group-averaged discriminative features. Unsupervised learning on saliency maps can provide novel insights into clinical subtypes of disease. It is also important to note that machine learning techniques do not unequivocally provide evidence for the salient features being directly associated with the disease or other target variables. However, when combined with detailed future investigations, they can spur clinical discoveries.

4.8. Conclusion

The results presented in our paper showcase the utility of ensemble learning for connectomes. Functional network based prediction models are impacted by several a priori

choices, the most pivotal of which is the ROI definition. We demonstrate that ensembles of stochastic parcellations yield predictions that are significantly more robust and accurate compared to single atlas-based approaches. Further, our experiments highlight the potential of convolutional neural network models for connectome-based classification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

6.   References

[1]. Plitt M, Barnes KA, Martin A, Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards, Neuroimage Clin 7 (2015) 359–366. [PubMed: 25685703]

[2]. Mennes M, Vega Potler N, Kelly C, Di Martino A, Castellanos FX, Milham MP, Resting state functional connectivity correlates of inhibitory control in children with attention-deficit/hyperactivity disorder, Front Psychiatry 2 (2011) 83. [PubMed: 22470352]

[3]. Varoquaux G, Baronnet F, Kleinschmidt A, Fillard P, Thirion B, Detection of brain functional-connectivity di erence in post-stroke patients using group-level covariance modeling, Med Image Comput Comput Assist Interv 13 (2010) 200–208. [PubMed: 20879232]

[4]. Brown CJ, Hamarneh G, Machine learning on human connectome data from MRI, CoRR 1611.08699 (2016).

[5]. Kaiser M, A Tutorial in Connectome Analysis: Topological and Spatial Features of Brain Networks, ArXiv e-prints (2011).

[6]. Alexander-Bloch A, Vértes PE, Stidd R, Lalonde F, Clasen L, Rapoport JL, Giedd JN, Bullmore ET, Gogtay N, The anatomical distance of functional connections predicts brain network topology in health and schizophrenia., Cerebral cortex 23 1 (2013) 127–38. [PubMed: 22275481]

[7]. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beck-mann CF, Nichols TE, Ramsey JD, Woolrich MW, Network modelling methods for FMRI (2011).

[8]. Yao Z, Hu B, Xie Y, Moore P, Zheng J, A review of structural and functional brain networks: small world and atlas, Brain Informatics 2 (2015) 45–52. [PubMed: 27747502]

[9]. Dadi K, Rahim M, Abraham A, Chyzhyk D, Milham M, Thirion B,Varoquaux G, Benchmarking functional connectome-based predictive models for resting-state fMRI, 2018 Working paper or preprint.

[10]. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C,Van Der Kouwe A, Killiany R, Kennedy D, Klaveness S, et al., Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain, Neuron 33 (2002) 341–355. [PubMed: 11832223]

[11]. Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J,Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, et al., A multi-modal parcellation of human cerebral cortex, Nature 536 (2016) 171–178. [PubMed: 27437579]

[12]. Eickho SB, Thirion B, Varoquaux G, Bzdok D, Connectivity-based parcellation: Critique and implications, Human brain mapping 36 (2015) 4771–4792. [PubMed: 26409749]

[13]. Arslan S, Ktena SI, Makropoulos A, Robinson EC, Rueckert D,Parisot S, Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex, NeuroImage 170 (2018) 5–30. Segmenting the Brain. [PubMed: 28412442]

[14]. Yushkevich PA, Amaral RS, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, Bocchetta M, Burggren AC, Carr VA,Chakravarty MM, et al., Quantitative comparison of 21 protocols for

labeling hippocampal subfields and parahippocampal subregions in in vivo mri: towards a harmonized segmentation protocol, Neuroimage 111 (2015) 526–541. [PubMed: 25596463]

[15]. Varoquaux G, Gramfort A, Pedregosa F, Michel V, Thirion B, Multi-subject dictionary learning to segment an atlas of brain spontaneous activity, in: Biennial International Conference on Information Processing in Medical Imaging, Springer, pp. 562–573.

[16]. Thomas Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR,Lashkari D, Hollinshead M, Ro man JL, Smoller JW, Zllei L,Polimeni JR, Fischl B, Liu H, Buckner RL, The organization of the human cerebral cortex estimated by intrinsic functional connectivity, Journal of Neurophysiology 106 (2011) 1125–1165. [PubMed: 21653723]

[17]. Thirion B, Varoquaux G, Dohmatob E, Poline J-B, Which fmri clustering gives good brain parcellations?, Frontiers in neuroscience 8 (2014) 167. [PubMed: 25071425]

[18]. Abraham A, Milham MP, Martino AD, Craddock RC, Samaras D, Thirion B, Varoquaux G, Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example, NeuroImage 147 (2017) 736–745. [PubMed: 27865923]

[19]. Ktena SI, Parisot S, Ferrante E, Rajchl M, Lee M, Glocker B,Rueckert D, Metric learning with spectral graph convolutions on brain connectivity networks, NeuroImage 169 (2018) 431–442. [PubMed: 29278772]

[20]. Kawahara J, Brown C, P Miller S, Booth B, Chau V, Grunau R,Zwicker J, Hamarneh G, Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment 146 (2016).

[21]. Cherkassky VL, Kana RK, Keller TA, Just MA, Functional connectivity in a baseline resting-state network in autism., Neuroreport 17 16 (2006) 1687–90. [PubMed: 17047454]

[22]. Assaf M, Jagannathan K, Calhoun V, Miller L, Stevens M, Sahl R,O'Boyle J, Schultz R, Pearlson G, Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients 53 (2010) 247–56.

[23]. Monk CS, Peltier SJ, Wiggins JL, Weng S-J, Carrasco M, Risi S,Lord C, Abnormalities of intrinsic functional connectivity in autism spectrum disorders, NeuroImage 47 (2009) 764–772. [PubMed: 19409498]

[24]. Heinsfeld AS, et al., Identification of autism spectrum disorder using deep learning and the abide dataset, in: NeuroImage: Clinical.

[25]. Yahata N, Morimoto J, Hashimoto R, Lisi G, Shibata K, K. Y, A small number of abnormal brain connections predicts adult autism spectrum disorder, Nature Communications 7 (2016).

[26]. Di Martino A, O'Connor D, Chen B, Alaerts K, Anderson JS,Assaf M, Balsters JH, Baxter L, Beggiato A, Bernaerts S, Blanken LME, Bookheimer SY, Braden BB, Byrge L, Castellanos FX, Dapretto M, Delorme R, Fair DA, Fishman I, Fitzgerald J,Gallagher L, Keehn RJJ, Kennedy DP, Lainhart JE, Luna B,Mostofsky SH, Mller R-A, Nebel MB, Nigg JT, O'Hearn K,Solomon M, Toro R, Vaidya CJ, Wenderoth N, White T, Craddock RC, Lord C, Leventhal B, Milham MP, Enhancing studies of the connectome in autism using the autism brain imaging data exchange ii, Scientific data 4 (2017) 170010. [PubMed: 28291247]

[27]. Craddock C, Benhajali Y, Chu C, Chouinard F, Evans A, Jakab A,Khundrakpam BS, Lewis JD, Li Q, Milham M, Yan C, Bellec P, The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivative, Frontiers in Neuroinformatics (2013).

[28]. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL,Petersen SE, Methods to detect, characterize, and remove motion artifact in resting state fMRI, NeuroImage 84 (2014) 320–341. [PubMed: 23994314]

[29]. Muschelli J, Nebel MB, Caffo BS, Barber A, Pekar JJ, Mostofsky SH, Reduction of motion-related artifacts in resting state fmri using acompcor 96 (2014).

[30]. Frazier JA, et al., Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder., The American journal of psychiatry 162 7 (2005).

[31]. Goldstein JM, Seidman LJ, Makris N, Ahern T, O'Brien LM,Caviness VS, Kennedy DN, Faraone SV, Tsuang MT, Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability, Biol. Psychiatry 61 (2007) 935–945. [PubMed: 17046727]

[32]. Makris N, Goldstein JM, Kennedy D, Hodge SM, Caviness VS,Faraone SV, Tsuang MT, Seidman LJ, Decreased volume of left and total anterior insular lobule in schizophrenia, Schizophr. Res 83 (2006) 155–171. [PubMed: 16448806]

[33]. Smyser CD, Dosenbach NU, Smyser TA, Snyder AZ, Rogers CE, Inder TE, Schlaggar BL, Neil JJ, Prediction of brain maturity in infants using machine-learning algorithms, Neuroimage 136 (2016) 1–9. [PubMed: 27179605]

[34]. Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC,Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT,Albert MS, Killiany RJ, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, Neuroimage 31 (2006) 968–980. [PubMed: 16530430]

[35]. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F,Etard O, Delcroix N, Mazoyer B, Joliot M, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, Neuroimage 15 (2002) 273–289. [PubMed: 11771995]

[36]. Cameron C, James G, Holtzheimer P, Hu X, HS M., A whole brain fmri atlas generated via spatially constrained spectral clustering, Human Brain Mapping 33 (2011) 1914–1928. [PubMed: 21769991]

[37]. Lancaster JL, Woldor MG, Parsons LM, Liotti M, Freitas CS,Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT, Automated talairach atlas labels for functional brain mapping, Human Brain Mapping 10 (2000) 120–131. [PubMed: 10912591]

[38]. Eickho SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR,Amunts K, Zilles K, A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data, Neuroimage 25 (2005) 1325–1335. [PubMed: 15850749]

[39]. Schirmer MD, Developing Brain Connectivity - Effects of Parcellation Scale on Network Analysis in Neonates (Doctoral dissertation, King's College London) (2015).

[40]. Ioffe S, Szegedy C, Batch normalization: Accelerating deep network training by reducing internal covariate shift, CoRR abs/1502.03167 (2015).

[41]. Simonyan K, Vedaldi A, Zisserman A, Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps, ArXiv e-prints (2013).

[42]. Utevsky AV, Smith DV, Huettel SA, Precuneus is a functional core of the default-mode network, Journal of Neuroscience 34 (2014) 932–940. [PubMed: 24431451]

[43]. Watanabe T, Yahata N, Abe O, Kuwabara H, Inoue H, Takano Y,Iwashiro N, Natsubori T, Aoki Y, Takao H, et al., Diminished me-dial prefrontal activity behind autistic social judgments of incongruent information, PloS one 7 (2012) e39561. [PubMed: 22745788]

[44]. Koshino H, Carpenter PA, Minshew NJ, Cherkassky VL, Keller TA, Just MA, Functional connectivity in an fmri working memory task in high-functioning autism, Neuroimage 24 (2005) 810–821. [PubMed: 15652316]

[45]. Reuter-Lorenz PA, Jonides J, Smith EE, Hartley A, Miller A,Marshuetz C, Koeppe RA, Age differences in the frontal lateralization of verbal and spatial working memory revealed by pet, Journal of cognitive neuroscience 12 (2000) 174–187. [PubMed: 10769314]

[46]. Craddock RC, James G, Holtzheimer PE, Hu XP, Mayberg HS, A whole brain fMRI atlas generated via spatially constrained spectral clustering, Human Brain Mapping 33 (2012) 1914–1928. [PubMed: 21769991]

[47]. Fornito A, Zalesky A, Bullmore ET, Network scaling effects in graph analytic studies of human resting-state FMRI data, Front Syst Neurosci 4 (2010) 22. [PubMed: 20592949]

[48]. Zalesky A, Fornito A, Harding IH, Cocchi L, Yucel M, Pantelis C,Bullmore ET, Whole-brain anatomical networks: does the choice of nodes matter?, Neuroimage 50 (2010) 970–983. [PubMed: 20035887]

[49]. Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, Sun N,Zuo XN, Holmes AJ, Eickho SB, Yeo BTT, Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion, Cereb. Cortex (2018).

[50]. Da Mota B, Fritsch V, Varoquaux G, Frouin V, Poline J, T. B, Enhancing the reproducibility of group analysis with randomized brain parcellations, in: Medical Image Computing and

Computer-Assisted Intervention - MICCAI 2013. Lecture Notes in Computer Science, vol 8150 Springer, Berlin, Heidelberg.

[51]. Chen CP, Keown CL, Jahedi A, Nair A, Pflieger ME, Bailey BA, Muller RA, Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism, Neuroimage Clin 8 (2015) 238–245. [PubMed: 26106547]

[52]. Menon V, Developmental pathways to functional brain networks: emerging principles, Trends Cogn. Sci. (Regul. Ed.) 17 (2013) 627–640. [PubMed: 24183779]

[53]. Satterthwaite TD, Wolf DH, Loughead J, Ruparel K, Elliott MA,Hakonarson H, Gur RC, Gur RE, Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth, NeuroImage 60 (2012) 623–632. [PubMed: 22233733]

[54]. Fair D, Nigg J, Iyer S, Bathula D, Mills K, Dosenbach N, Schlaggar B, Mennes M, Gutman D, Bangaru S, Buitelaar J, Dickstein D,Di Martino A, Kennedy D, Kelly C, Luna B, Schweitzer J, Velanova K, Wang Y-F, Mostofsky S, Castellanos F, Milham M, Distinct neural signatures detected for adhd subtypes after controlling for micro-movements in resting state functional connectivity mri data, Frontiers in Systems Neuroscience 6 (2013) 80. [PubMed: 23382713]

[55]. Van Dijk KR, Sabuncu MR, Buckner RL, The influence of head motion on intrinsic functional connectivity mri, Neuroimage 59 (2012) 431–438. [PubMed: 21810475]

[56]. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O, Mapping the structural core of human cerebral cortex, PLOS Biology 6 (2008) 1–15. [PubMed: 20076695]

[57]. Zintgraf LM, Cohen TS, Adel T, Welling M, Visualizing deep neural network decisions: Prediction di erence analysis, CoRR abs/1702.04595 (2017).

[58]. Selvaraju RR, Das A, Vedantam R, Cogswell M, Parikh D, Batra D, Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization, CoRR abs/ 1610.02391 (2016).

[59]. Bridson R, Fast poisson disk sampling in arbitrary dimensions, in: ACM SIGGRAPH 2007 Sketches, SIGGRAPH '07, ACM, New York, NY, USA, 2007.
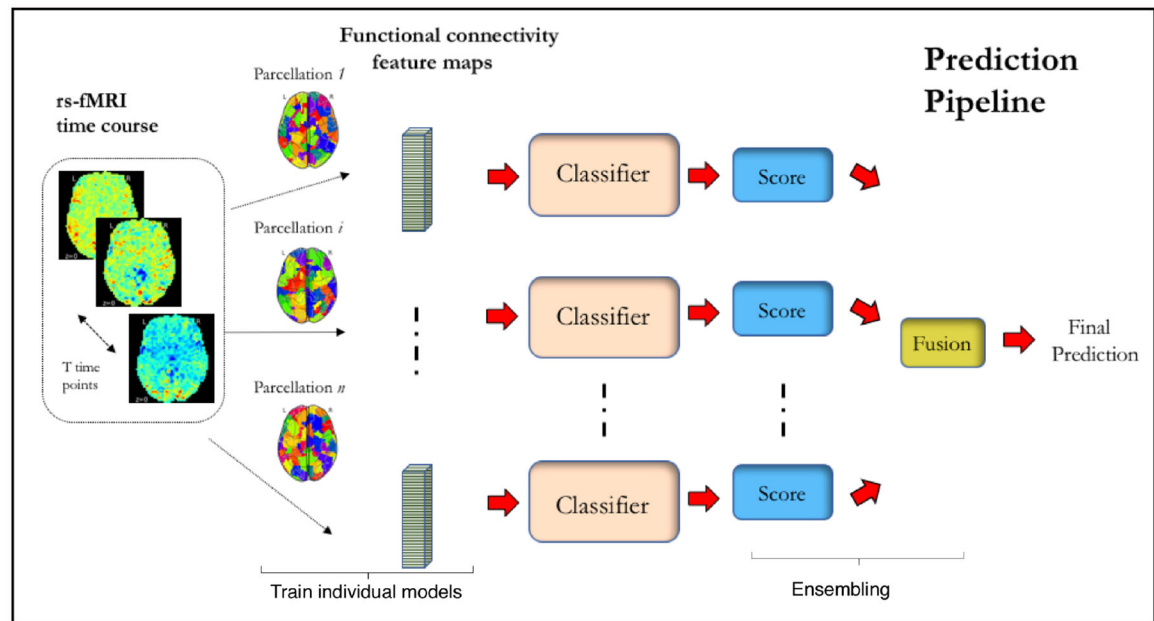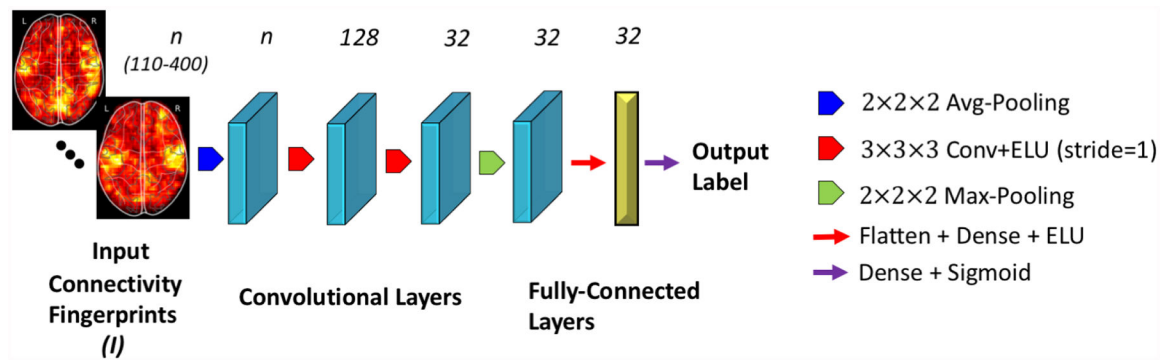
Figure 1:
A general illustration of the proposed approach
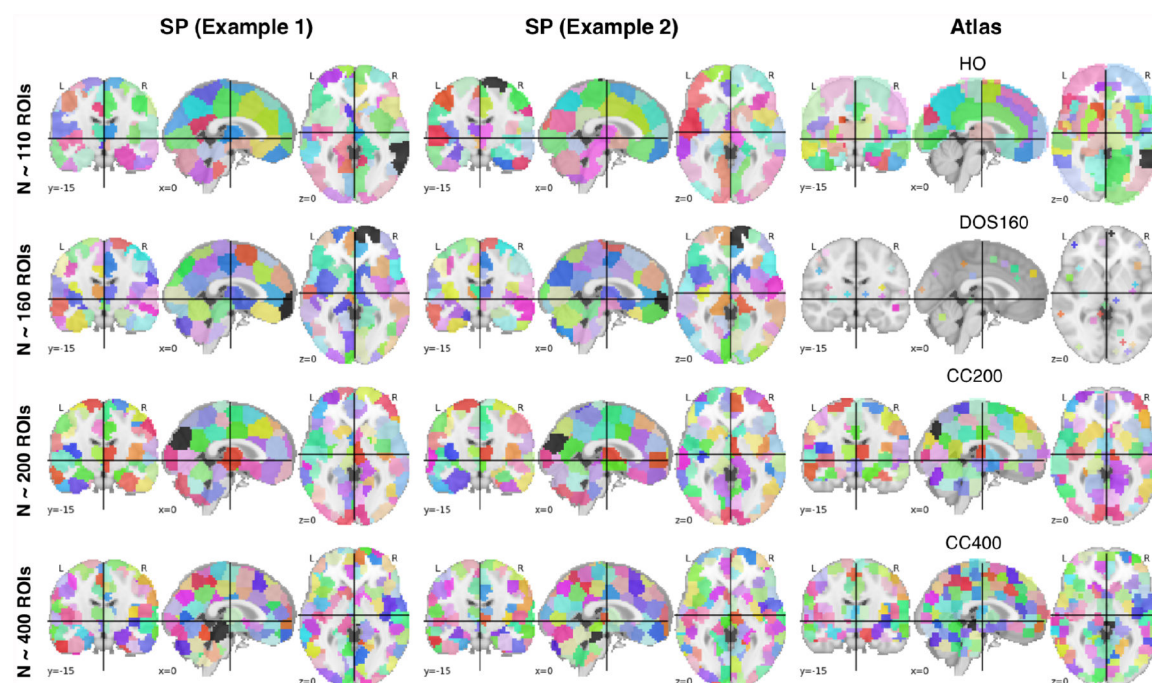
Figure 2:
ROI masks for example SPs and atlas at each of the four spatial scales considered in this study.
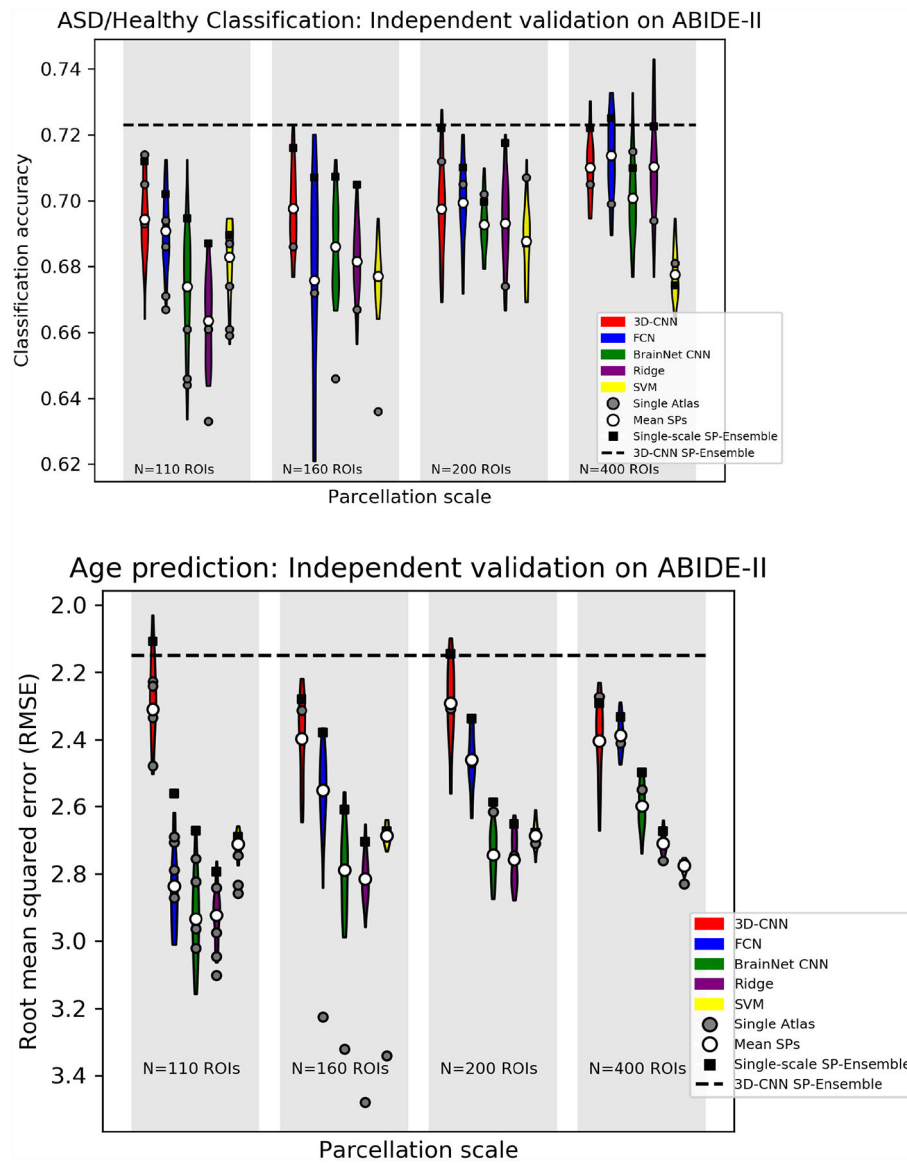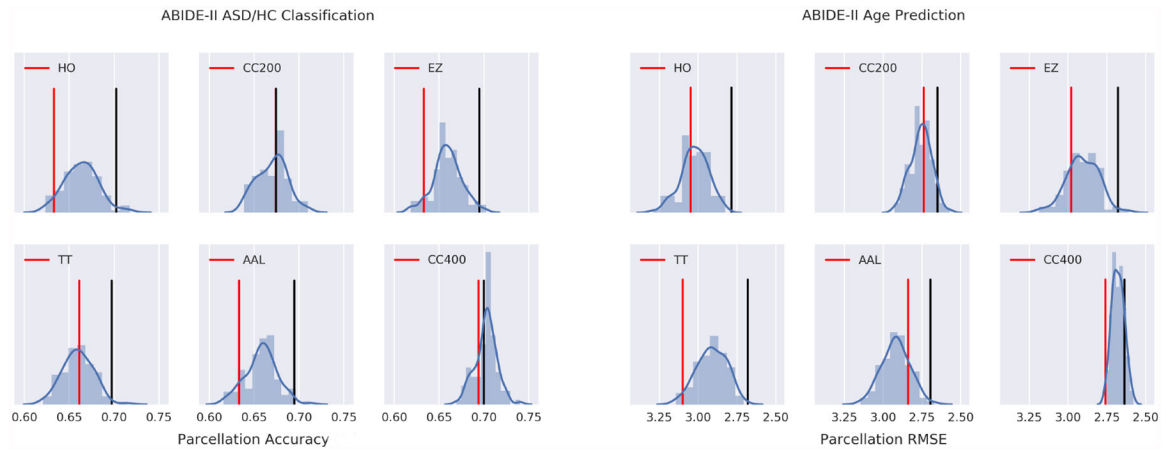
Figure 3:
Proposed CNN approach. All operations are in 3D volume. 2D correlation maps are shown for illustration only. For the age prediction task, an additional Max-Pooling and Batch-Normalization[40] operation followed the first and second convolutional layer.
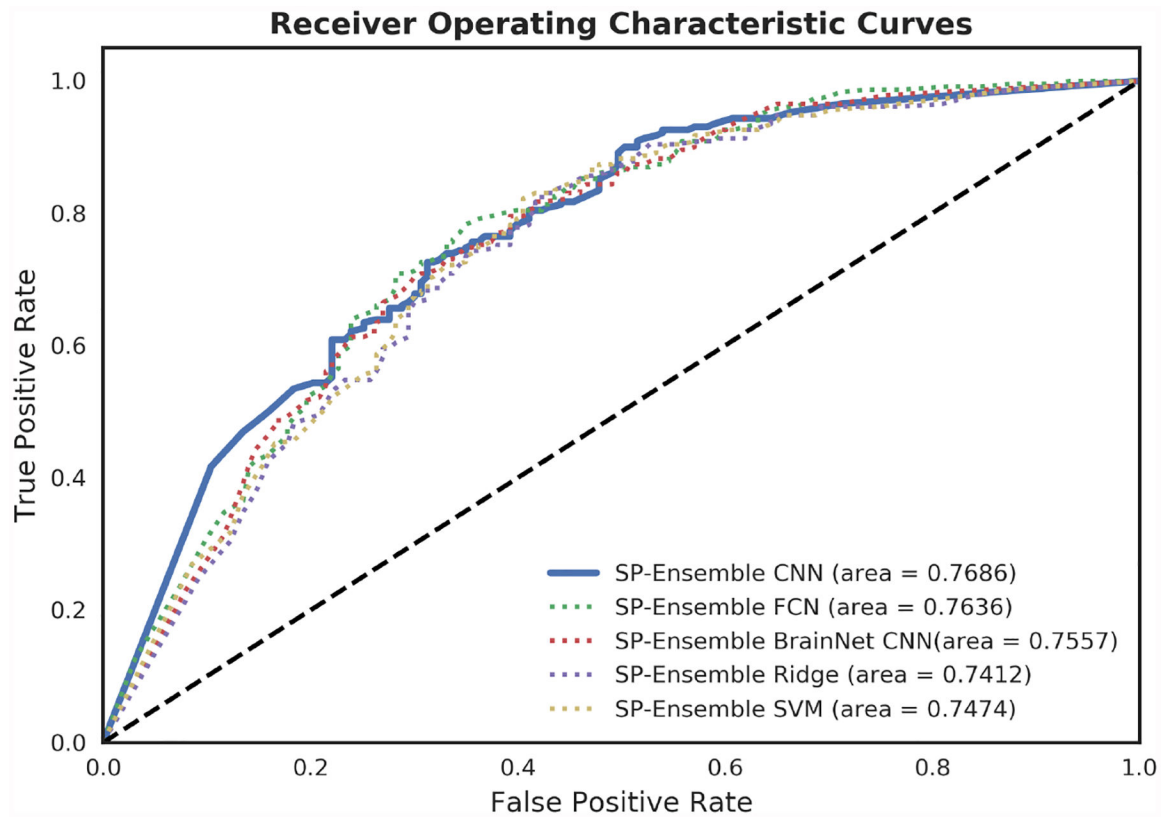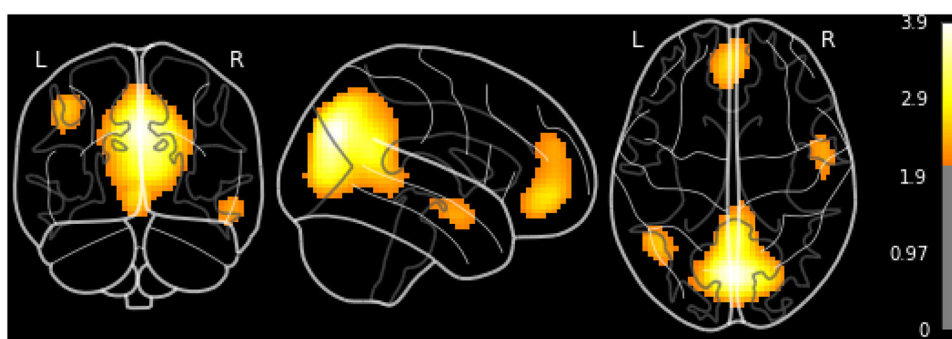
Figure 4:
ASD-HC Classification: Receiver Operating Curves for independent validation on ABIDE-2
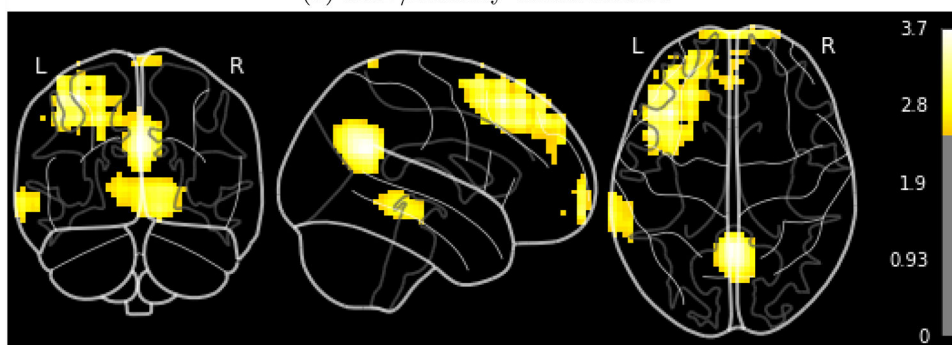
Figure 5:
Violin plots showing the spread of prediction accuracies/errors for stochastic parcellations at multiple network scales for different classification models. Mean accuracy/error of individual violins is denoted by 'Mean SPs'. Performance of individual atlases is compared with SPs with the closest15# of ROIs and is denoted as 'Single Atlas'. Results are computed by training models on entire ABIDE-1 cohort and testing on the independent ABIDE-2 cohort.

Figure 6:
Distribution of Ridge models' performance for stochastic parcellations created using the same gray-matter mask as the corresponding atlas. Red denotes the atlas model's accuracy and black indicates the SP-Ensemble accuracy.
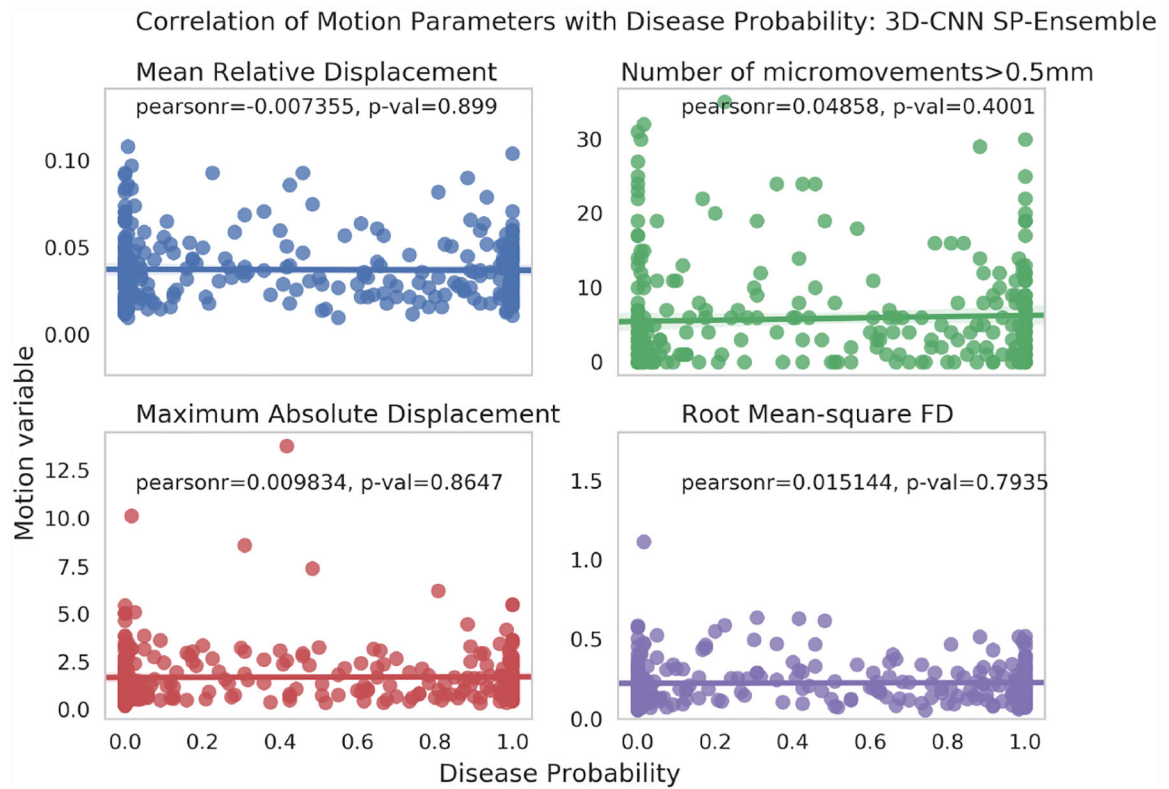
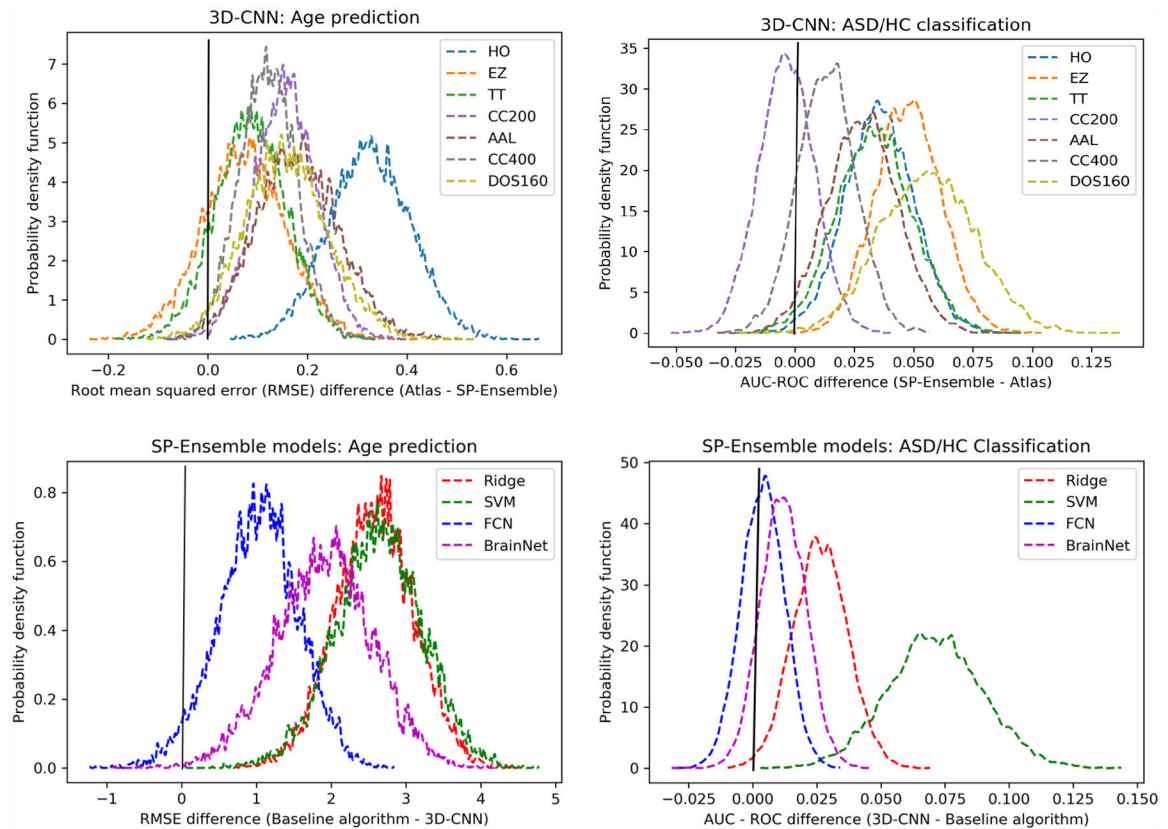(a) ASD/Healthy Classification

(b) Age prediction

Figure 7:
Mean saliency maps of trained 3D-CNN models for SP-Ensemble

Figure 8:
Motion correlations

Figure 9:
Kernel density estimates of the probability distributions for the performance di erence between models, computed based on 10000 bootstrap samples from ABIDE-II. Values to the left of the black vertical line indicate bootstrap samples where the proposed approach (3D CNN or SP-Ensemble) under-performed compared to the competing method.

Table 1:

Composition of Cohorts

Dataset	Prediction	Sample Size	Median Age (Range) in yrs
ABIDE-I	Age	387	13.8 (6.5–29.1)
ABIDE-I	ASD/HC	379/395	13.9 (6.5–56.2)
ABIDE-II	Age	213	10.6 (5.8–18.8)
ABIDE-II	ASD/HC	163/230	11.0 (5.2–38.9)

Table 2:

Classification accuracy for ASD vs. Control: Independent test on ABIDE-II of baseline models and proposed CNN approach. For each row, best results are bolded. For each column, best results are *italicized*. Green indicates better performance, whereas orange/red highlights worse performance.

ASD/HC Classification Accuracy (ABIDE-II)					
Parcellation	Ridge	SVM	FCN	BrainNet	3D-CNN
HO	63.3	**68.7**	67.7	66.1	67.7
CC200	67.4	70.7	*71.5*	70.2	***72.8***
EZ	63.3	66.1	63.8	64.4	**66.4**
TT	66.1	67.4	65.9	67.4	**70.0**
CC400	69.4	68.2	69.9	*71.5*	70.5
AAL	63.3	65.9	65.4	64.6	**69.5**
DOS160	66.7	63.6	66.1	64.6	**67.0**
MA-Ensemble	69.7	70.0	69.9	70.7	**71.7**
SP-Ensemble	*71.7*	*71.2*	71.2	70.5	**72.3**

Table 3:

Root mean squared error (RMSE in years) for age prediction: Independent test on ABIDE-II for benchmark models and proposed CNN approach. For each row, best results are **bolded**. For each column, best results are *italicized*.

Age RMSE (ABIDE-II)					
Parcellation	**Ridge**	**SVM**	**FCN**	**BrainNet**	**3D-CNN**
HO	3.05	2.86	2.79	2.82	**2.48**
CC200	2.74	2.71	2.47	2.62	**2.31**
EZ	2.98	2.72	2.71	2.96	**2.23**
TT	3.10	2.83	2.87	3.02	**2.24**
CC400	2.76	2.83	2.41	2.55	**2.27**
AAL	2.84	2.74	2.69	2.75	**2.33**
DOS160	3.48	3.34	3.22	3.32	**2.31**
MA-Ensemble	2.72	2.81	2.47	*2.55*	**2.15**
SP-Ensemble	*2.68*	*2.69*	*2.38*	*2.55*	***2.15***

Table 4:

Classification/regression performance of FCN with a high-resolution parcellation (~ 1024 ROIs) [48]

FCN Results using Zalesky's 1024 node parcellation		
	ABIDE-I (10-fold)	**ABIDE-II**
DX (% accuracy)	72.0	72.0
Age (RMSE)	3.54	2.89