



Deep Learning on Brain Images in Autism: What Do Large Samples Reveal of Its Complexity?

Matthew Leming^(✉)  and John Suckling 

Department of Psychiatry, Cambridge University, Cambridge CB20SZ, UK
m1784@cam.ac.uk

Abstract. Deep learning models for image classification face two recurring problems: they are typically limited by low sample size and are abstracted by their own complexity (the “black box problem”). We address these problems with the largest functional MRI connectome dataset ever compiled, classifying it across gender and Task vs rest (no task) to ascertain its performance, and then apply the model to a cross-sectional comparison of autism vs typically developing (TD) controls that has proved difficult to characterise with inferential statistics. Employing class-balancing to build a training set, a convolutional neural network was classified fMRI connectivity with overall accuracies of 76.35% (AUROC 0.8401), 90.71% (AUROC 0.9573), and 67.65% (AUROC 0.7162) for gender, task vs rest, and autism vs TD, respectively. Saliency maps demonstrated that the deep learning model is capable of distinguishing complex patterns across either wide networks or localized areas of the brain, and, by analyzing maximal activations of the hidden layers, that the deep learning model partitions data at an early stage in its classification.

Keywords: Autism · Big data · Functional connectivity · Deep learning

1 Introduction

Motivated by reports of increased head circumference in children diagnosed with autism, the first measurements with MRI reported increased total brain, total tissue, and total lateral ventricle volumes in autistic adults [31]. Many similar studies followed, leading to a general consensus that brain volume was increased in autism. Moreover, in a highly cited article [8], increases in brain volume were suggested to occur in the first few years of life when diagnostic symptoms - social communication challenges, restricted and repetitive behaviours - also emerge. Since then, as further evidence has accumulated, the period of early brain overgrowth has been restricted to the first year of life [35], although large-scale longitudinal studies have failed to reproduce these meta-analytic findings [2, 13]. Localising putative changes to brain structure has proved to be an even greater

difficultly, with discrepancies between meta-analyses even though there is a significant overlap in the primary literature [6, 10, 41].

To address variations in data acquisition and processing that make between study comparisons less powerful, publicly available large-sample datasets are now a key aspect of imaging research. The ABIDE (http://fcon_1000.projects.nitrc.org/indi/abide/) multi-centre initiative has made available over 2000 images in two releases, but cross-sectional analyses of structural MRI have failed to observe significant differences [17, 42]. The majority of these studies have used the established voxel brain morphometry technique (VBM), to estimate voxelwise tissue occupancies, and mass-univariate statistical testing. Other morphological properties of the cortex may yield greater sensitivity [25].

The measurement of correlation, or ‘functional connectivity’, between time-series of blood oxygenation level dependent (BOLD) endogenous contrast estimated from brain regions whilst in resting wakefulness has been demonstrated as a reproducible measurement on an individual basis [16]. Functional connectivity estimates represented as undirected graphs (connectomes) of nodes (brain regions) and edges (connectivity strengths) show promise in localising differences in resting activity to specific large-scale brain networks [40], and although there is cautionary evidence using the ABIDE dataset and others [32], it would appear that statistically significant differences in connectivity are generally observable, although variable in their presentation.

Computing power and access to large datasets have led to a resurgence in the popularity of neural networks (NNs) as a tool for data classification. In parallel, because of their wide applicability in representing complex data such as proteins and social networks, connectors have undergone significant development in terms of global and local characteristics. Some recent work has used NNs for processing connectomes, including whole-graph classification, clustering into sub-graphs, and node-wise classification [4, 9, 19, 21, 26, 29].

In this article, we leverage publicly available datasets to amass and automatically pre-process a total of 39,461 functional MRIs from nine different multi-centre collections. We first classify them based on gender and task vs rest (no task) as a test of the validity of the application NNs to imaging data due to the known connectivity differences identified using inferential statistics [1, 37]. We then classify autistic individuals from typically developing (TD) controls. All classifications were undertaken using a convolutional neural network (CNN) that uniquely encodes multi-layered connectivity matrices, an extension of the deep learning architecture previously described in [23]. To incentivise the model to classify based on phenotypic differences rather than site differences, class balancing techniques were used when building the training and test sets and compared against the fully-inclusive samples. Key outputs of the CNN were salience maps [23, 38] that highlighted areas of the connectome the model preferentially focused on when performing its classification, and activation maximization [15] of a hidden layer inspected to visualize how the model partitioned the dataset following classification.

Table 1. Populations present in each dataset. Note that this represents the data that was successfully preprocessed and does not include data that failed this stage.

Collection	Subjs	Conns	Rest	Task	Age		Sex				Disorders		
					Min	Max	Mean	Stddev	F	M	Depr	ASD	Alzh
1000 FC	833	833	833	0	7.88	85	26.47	11.22	475	358	0	0	0
ABCD	1424	11789	5142	6647	0.42	11.08	10.05	0.64	5553	6236	1557	124	0
ABIDE	658	658	658	0	6	64	17.18	7.94	84	574	0	307	0
ABIDE II	646	674	674	0	5.32	55	14.48	7.63	155	464	0	293	0
ADNI	158	309	309	0	56	95	73.38	7.22	171	138	0	0	157
BioBank	11275	11275	11275	0	40	70	54.93	7.50	6172	5103	1486	0	0
ICBM	83	353	0	353	19	74	43.53	14.58	169	184	0	0	0
OPEN FMRI	1409	6548	997	5551	5.89	78	27.11	10.47	2774	3037	182	81	0
NDAR	1284	7022	5080	1942	0.25	53.42	20.51	8.05	3434	3588	0	404	0
All	17770	39461	24968	14493	0.25	95	28.97	19.88	18987	19682	3225	1209	157

In attempting to classify components of this accumulated dataset, we sought to address the following questions: (1) How effective is a machine learning paradigm at classifying fMRI connectomes? (2) Which areas or networks of the brain do models focus on when undertaking classifications? (3) How does the model partition large datasets during classification? (4) Can the model effectively classify functional connectivities taken from multiple sources without relying explicitly on site differences to do so? (5) What is the best current evidence for cross-sectional differences in functional connectivity that characterise autism?

2 Methods

2.1 Datasets and Preprocessing

Datasets were acquired from OpenFMRI [33,34]; the Alzheimer’s Disease Neuroimaging Initiative (ADNI); ABIDE [12]; ABIDE II [11]; the Adolescent Brain Cognitive Development (ABCD) Study [5]; the NIMH Data Archive, including the Research Domain Criteria Database (RDoCdb), the National Database for Clinical Trials (NDCT), and, most predominantly, the National Database for Autism Research (NDAR) [18]; the 1000 Functional Connectomes Project [14]; the International Consortium for Brain Mapping database (ICBM); and the UK Biobank; we refer to each of these sets as *collections*. OpenFMRI, NDAR, ICBM, and the 1000 Functional Connectomes Project are collections that comprise different datasets submitted from unrelated research groups; ADNI, ABIDE, ABIDE II, ABCD, and the UK Biobank are collections that were acquired as part of a larger research initiative.

Data were preprocessed using the fMRI Signal Processing Toolbox and the Brain Wavelet Toolbox [30] and parcellated with the 116-area Automated Anatomical Labelling (AAL) atlas that defined the nodes of the connectome, with the edges weighted by the correlation of the wavelet coefficients from the decomposition of the pre-processed BOLD time-series in each of four temporal scales: 0.1–0.2 Hz, 0.05–0.1 Hz, 0.03–0.05 Hz, and 0.01–0.03 Hz.

Datasets with regional dropout or which otherwise failed the parcellation stage were omitted from the analysis. Redundant datasets across collections were discarded. Multiple instances of connectivity matrices from the same individuals were used, though contributions from the same individuals were not shared between the training, validation, and test sets. The numbers of participants, total numbers of datasets used as well as phenotypic distributions, are shown in Table 1.

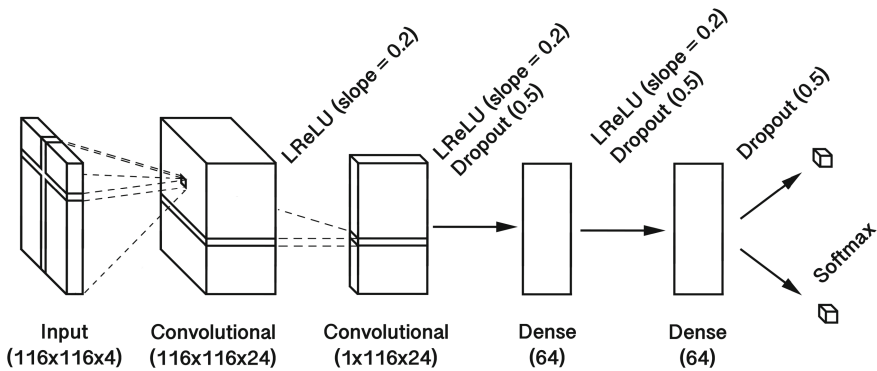


Fig. 1. The structure of the neural network, based on BrainNetCNN.

2.2 Neural Network Model and Training

The data used for training and testing the CNN were $4 \times 116 \times 116$ (4 wavelet scales and 116 nodes) symmetric connectivity (wavelet coefficient correlation) matrices with values scaled on $[0, 1]$.

To classify data, we employed a CNN with cross-shaped filters described in [23]; to allow the network to train on connectivity matrices (Fig. 1). We re-implemented the architecture of [23] using Keras [7], a popular machine learning library, leveraging the advantages of other software libraries that support Keras. Additionally, this re-implementation extended the model to include multiple channels in the inputs, as opposed to single connectivity matrices.

The CNN was constructed with: 24 edge-to-edge filters; 24 edge-to-node filters; 2 fully-connected layers, each with 64 nodes; and a final softmax layer. Three leaky rectified linear unit (ReLU) layers, with a slope of 0.2, and three dropout layers, with a dropout rate of 0.5, were also used in the network. Specifications are shown in Fig. 1. The model was trained using an Adam optimizer with batch sizes of 64, otherwise Keras defaults were used. Models were trained for 250 epochs, and the epoch with the highest validation accuracy was selected.

2.3 Set Division

Datasets were partitioned into three sets: the training set, comprised of two-thirds of the data and used to train the model; the validation set, comprised of

one-sixth of the data and used to select the epoch at which training stopped; and the test set, used to assess the trained classifier performance.

For all classifications, balancing was used such that each class comprised approximately half of the datasets. To account for gender, age, and possible scanning site differences between datasets, we report the inter-dataset classification accuracy as well as the global accuracy. Two different class-balancing approaches were used when building the sets: one selected two age-matched cases for gender and task vs rest classification, and the other selected a case and corresponding control from the same collection for the autism vs TD classification.

2.4 Test Set Evaluation

Inter-data Classification. Following the training of the models, the accuracy and the area under the receiver operating characteristic curve (AUROC) were calculated as measures of machine learning performance on the test set. This was to determine if one group in the classification outperformed the other in training leading to a biasing of the overall accuracy.

Activation Maximization. Activation maximization [15] is a technique to determine the maximally activated hidden units in response to the test set of the CNN layers following training. Activation maximization was applied to the 116×24 second layer of our network (Fig. 1) as this two-dimensional convolutional layer acts as a bottleneck, and is thus easier to interpret and visualize. This layer is naturally stratified by 24 *filters*, each with 116 nodes that correspond to parcellated brain areas. To offset the influence of spurious maximizations, we opted to record the 10 datasets that maximally activated each hidden unit, displaying their mode, collection, gender, age group, and task/rest; for example, if 6 connectomes that maximally activated a unit were from Biobank and four from Open fMRI, Biobank would be displayed as maximally activating that unit.

Saliency Maps. We deployed saliency maps [23, 28, 38] using a previous Keras implementation [27] to display the parts of the connectivity matrix the CNN emphasised in its classification of the test set. Class saliency extractions operate by taking the derivative of the CNN classification function (approximated as a first-order Taylor expansion, estimated via back-propagation) with respect to an input matrix, with the output being the same dimensions as the input [38]. Saliency extractions are particularly advantageous when applied to connectivity matrices, because unlike typical 2D images these matrices are spatially static (i.e. each part of the matrix represents the same connection in the brain, across all datasets), and thus global tendencies of the model can be visualized. Saliency maps for each adjacency matrix were averaged and displayed to demonstrate on which aspects of the connectome the CNN was most focused when performing the classification.

Table 2. Populations present in the training set for each round of classification, with and without class balancing

	With class balancing						Without class balancing					
	Autism		Rest v Task		Gender		Autism		Rest v Task		Gender	
	non-ASD	ASD	Rest	non-Rest	Female	Male	non-ASD	ASD	Rest	non-Rest	Female	Male
1000 FC	0	0	507	0	320	271	0	0	625	0	358	272
ABCD	61	82	3088	3634	3370	3275	7301	82	3091	4292	3416	3991
ABIDE	232	229	461	0	62	375	263	230	493	0	60	434
ABIDE II	221	218	408	0	113	303	284	221	505	0	117	346
ADNI	0	0	4	0	106	88	0	0	230	0	135	95
BIOBANK	3	3	597	0	3967	3704	0	0	8458	0	4633	3825
ICBM	0	0	0	185	99	105	0	0	0	217	121	94
NDAR	122	130	2164	909	2194	2020	4307	138	3215	1230	2180	2240
OPEN FMRI	61	62	372	2792	1977	2059	0	0	883	4030	2059	2299
TOTAL	700	724	7601	7520	12208	12200	12155	671	17500	9769	13079	13596

2.5 Experiments

We performed the classification on class-balanced datasets that then classified based on gender, task vs rest, and autism vs TD controls. For gender and task vs rest classifications, we also balanced classes by age; that is, the distribution of ages for each group was the same. For autism vs TD controls, we balanced across collections to minimise site differences, and also as a proxy for age whilst maximising the sample size.

Additionally, we trained models without the use of class balancing, only excluding collections that entirely lacked a particular class. The number of connectomes used in each experiment’s training set, with and without class balancing, are given in Table 2. Where training was successful, we report the overall classification accuracy and AUROC for the balanced and unbalanced test sets.

3 Results

The results displayed a tendency of the model to use particular filters to sequester data by different variables, especially if it were attempting to classify by that variable, although the model divided data across certain filters independent of the classification variable. While gender, task vs rest, and autism vs TD controls each have a small proportion of their filters wholly activated by the datapoints of a single collection (which may be easy to distinguish based on differences between MRI scanners), the majority of filters were activated by a variety of different collections, indicating the effective synthesis of data from different sources. Those comparisons that saw the highest classification accuracy tended to activate individual filters in most nodes, indicating the network’s tendency to group data early in the architecture, prior to the fully-connected layers.

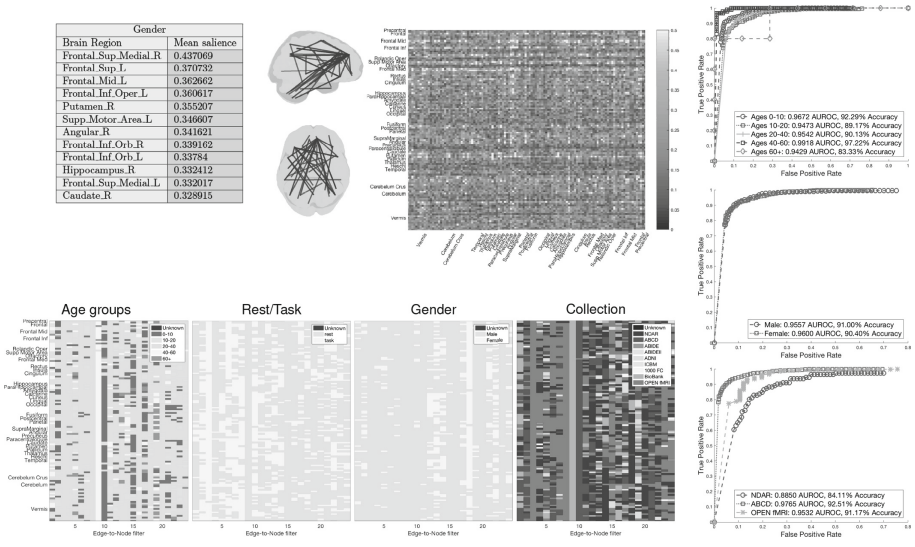


Fig. 2. Activation maximization, salience map, and ROC-curve results on gender classification

3.1 Gender Comparison

With class balancing, classification accuracy on the test group was 76.35% (AUROC 0.8401). Classification was most successful on the UK Biobank collection (87.31% accuracy). When stratifying by age, the CNN was able to obtain higher performance distinguishing gender in older age groups than younger age groups, and better able to achieve classification of gender in resting than task-based fMRI (78.96% versus 71.44%).

In activation maximization Fig. 2, filters 3/8, 9, and 10 were almost entirely dedicated to the classification of OPEN fMRI, Biobank, and ABCD collections, respectively. Filter 18 was activated by females, whilst most other filters are activated by males. Filter 3 was activated by task-based fMRI from OPEN fMRI, while filter 15 as activated by resting-state fMRI from no particular collection. Salience maps indicated that gender classification utilized a wider spread of areas, focusing on networks in the frontal lobe.

Without class balancing, the results on gender were a comparable 0.8406 AUROC and 76.88% accuracy.

3.2 Task vs Rest (No Task)

With balanced classes, task vs rest fMRI classification was successful with 90.71% (AUROC 0.9573) of the test set correctly assigned. The training set, whilst balanced by age, had a high imbalance between collections. The AUROC of those collections that contributed substantial amounts both resting-state and task participants - i.e., NDAR, ABCD, and Open fMRI - had comparable AUROCs to

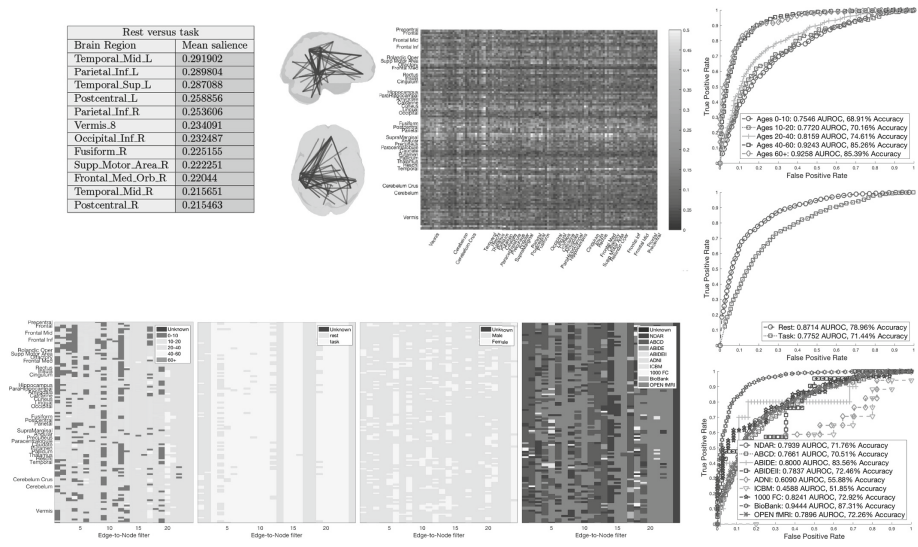


Fig. 3. Activation maximization, salience map, and ROC-curve results on resting-state-versus-task classification

that obtained overall. Furthermore, the salience map focused on the default mode network in the left hemisphere and its connection to the right frontal medial orbital area. Together, this suggests that the main influence in classification was not site differences.

In activation maximization six filters were dedicated to the resting-state class, fourteen to task-based fMRI, and four were mixed Fig. 3. This is likely indicative of the deep learning model using a simpler characterization of resting-state fMRI than task-based, which used more of its internal memory to capture the distinguishing patterns.

Without class balancing, the model achieved a higher 0.9792 AUROC, presumably displaying the effects of using more data on these models, as the age-balancing technique applied effectively discarded nearly half the training set data. However, a classification based partially on age groups is also possible.

3.3 Autism vs TD Controls

With class balancing, the overall performance on the test set was 67.65% (AUROC 0.7162). Autism classifications were highly dependent on the collection used, though the final accuracies were above chance for all collections. Class balancing was necessary, as data from autistic individuals comprised a relatively low percentage; collections with data from autistic individuals - Open fMRI, ABIDE I and II, NDAR, and ABCD - had <10%. Without class balancing the model failed to converge, simply classifying every datapoint as a TD control.

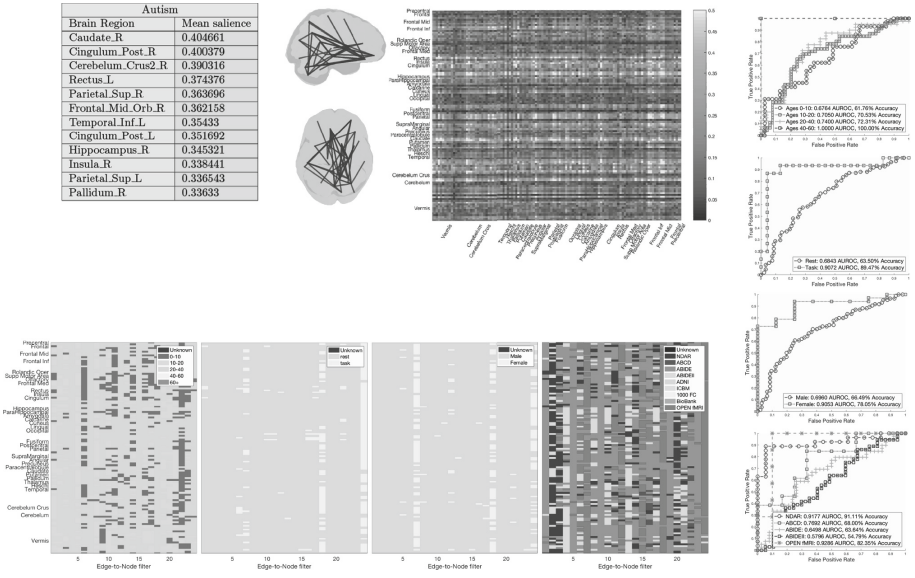


Fig. 4. Activation maximization, salience map, and ROC-curve results on autism classification

In activation maximization of the second layer, autism classification used filters 1, 5, 9, and 23 for the ABIDE collection, and filters 7, 18, and 24 were mostly used for females from the Open fMRI collection Fig. 4. The majority of nodes were maximally activated by data from the ABIDE I and II collections, although a disproportionately high number were used to classify Open fMRI autism data, which comprised < 10% of the total. A surprisingly low proportion of data from the NDAR and ABCD collections maximally activated the nodes, even though its classification was relatively successful and comprised a more substantial portion of the dataset. Most other nodes were maximally activated by the male resting-state data, which reflects the autism dataset as a whole. The salience maps indicate autism was classified using specific, localised regions of the brain; notably, bilateral posterior cingulum and the right caudate nucleus.

4 Discussion

This work describes how large and diverse imaging data might be analyzed by deep learning models, encouraging the aggregation of publicly available collections. Data were partitioned based on clear and logical features of the images, and that, even with imperfect classification accuracies, deep learning models are capable of recognizing highly complex patterns in large datasets representing large-scale brain networks and localized structures.

The neuroscientific objective of this study was to use the available imaging data with deep learning to describe the pattern of functional brain changes that

distinguishes individuals with autism from TD controls. With the absence of any gold standard in the cross-sectional comparison, we first undertook other classifications that have more secure, robust findings in the extant literature to confirm the veracity of the developed methods.

When classifying gender, the model was influenced by diffuse areas connected to the frontal lobe (Fig. 2). This is consistent with previous findings in gender comparisons of functional imaging, which did not find differences in brain activity between specific areas, but rather differences in local functional connectivity over large areas of the cortex [39]. Gender classifications were most successful with larger collections with more consistent image quality (e.g. ABCD and BioBank), rather than smaller collections of very high-quality images.

Deep learning models are prone to sorting data by different variables relatively early on in the classification process (Figs. 2, 3, and 4), reserving different filters for different classes of data.

Task vs rest functional connectivity classifications, as expected, identified the major components of the well-known default mode network (Fig. 3), a set of bilateral and symmetric regions that is suppressed during exogenous stimulation. More filters maximally activated by task fMRI than resting-state, indicating the greater variation that characterizes task fMRI, which is related to cognitive performance [20]. The high classification accuracies and detected patterns gives credibility to the use CNN with neuroimages.

Bilateral areas with some correspondence to the default mode network, particularly parietal, temporal and frontal medial regions, were identified as salient to the comparison of the autism vs TD controls: Figs. 3 and 4. Notably, autistic individuals were additionally classified by connections to the cerebellum and deep structures (caudate and hippocampus). Prior cross-sectional studies of functional connectivity in autistic individuals have primarily thresholded connectivity estimates (i.e. correlations), whereas here all existing connections were included, both positive and negative. A comparison of connectors using a highly matched sub-set of the ABIDE II collection [24] found global differences, and reduced network segregation within the default mode network and primary auditory and somatosensory cortical regions, and between these regions and other large networks.

Model accuracy was lower compared to the highest rates reported in literature [3, 22, 28], although this result should be viewed with several caveats. The dataset used in this analysis was larger and more complex than any other previously analyzed, consisting of data from many collections. Direct comparisons of machine learning classification methods is difficult as there are no universally accepted methods to divide collections into training and test sets (unlike standardized competitions in other fields, such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [36]). Furthermore, our exclusion criteria differed, and, because we opted to use multiple scanning sessions from single subjects during training, we also used data in ABIDE not employed in previous studies. Class balancing may also have significantly affected the classification accuracy. Nevertheless, this was necessary to avoid spuriously large accuracies

due to the highly skewed ratio of autism:TD, where high rates of classification the larger groups lead to biases to the overall rate. Lastly, preprocessing methods and exclusion criteria are not typically shared across studies, and thus differences due to the input data cannot be discounted.

More generally, our deep learning model employed multichannel input. Although this has long been the standard in 2D image classification (for instance, RGB images), it has not been utilized before in the classification of connectomes. Theoretically, this provides an advantage, since it encodes more information about the underlying timeseries. In practice, multichannel inputs generally increased the accuracy of our model by 2–3% over the single-channel models tested.

We used salience maps [38] to identify connections and areas that the model incorporated in its classification; this method has previously been used in deep learning on functional connectivity [23, 28] as is an effective method of dissecting neural networks. However, a caveat to this is that salience maps are imperfect indicators of areas of importance in the data that may not give a complete depiction of the distinguishing features.

One of the key methods we used to interrogate the results from our deep learning model was activation maximization. Previously, activation maximization has been used for intuiting the internal configuration of neural networks rather than for interpretation purposes [15]. In this study, while some filters were solely activated by data from single collections, the majority by mixed data from different collections suggesting an ability to account for site differences during classification. Deployment of activation maximisation here led to specific observations: variation of task-based fMRI is far greater than during rest (six filters maximally activated by rsfMRI and 14 by tasks); dataset sequestering happens even without successful classification; the number of filters activated maximally by a particular dataset is not necessarily proportional to the classification accuracy of that dataset.

5 Conclusion

With careful class-balancing, deep learning models are capable of good quality classifications across mixed collections detecting differences in brain networks, and functions of localized structures, or functional connections over large areas. Saliency maps highlighted key spatial elements of the classification and activation maximisation gave insights into the types of features on which the CNN based its classification. This deep learning model is an example of the apparatus to leverage publicly accessible large volumes of data for discovery science.

Acknowledgements. This study used publicly available datasets, each with their own acknowledgements. For brevity, we have not included the full text, but recognise the contributions of the Alzheimer’s Disease Neuroimaging Initiative, International Consortium for Brain Mapping, National Database for Autism Research, NIH Pediatric MRI Data Repository, National Database for Clinical Trials, Research Domain Criteria Database, Adolescent Brain Cognitive Development Study, UK Biobank Resource, 1000

Functional Connectomes Project, ABIDE I and II, and Open fMRI. This research was co-funded by the NIHR Cambridge Biomedical Research Centre and Marmaduke Sheild. ML is supported by a Gates Cambridge Scholarship from the University of Cambridge.

References

1. Arbabshirani, M., Havlicek, M., Kiehl, K., Pearlson, G., Calhoun, V.: Functional network connectivity during rest and task conditions: a comparative study. *Hum. Brain Mapp.* **34**, 2959–2971 (2012). <https://doi.org/10.1002/hbm.22118>
2. Blanken, L., et al.: A prospective study of fetal head growth, autistic traits and autism spectrum disorder. *Autism Res.* **11**, 602–612 (2018). <https://doi.org/10.1002/aur.1921>
3. Brown, C., Kawahara, J., Hamarneh, G.: Connectome priors in deep neural networks to predict autism. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (2018). <https://doi.org/10.1109/ISBI.2018.8363534>
4. Bruna, J., Zaremba, W., Szlam, A., LeCun, Y.: Spectral networks and locally connected networks on graphs. In: ICLR (2014)
5. Casey, B., Dale, A.: The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018). <https://doi.org/10.1016/j.dcn.2018.03.001>
6. Cauda, F., et al.: Grey matter abnormality in autism spectrum disorder: an activation likelihood estimation meta-analysis study. *J. Neurol. Neurosurg. Psychiatry* **82**, 1304–1313 (2011). <https://doi.org/10.1136/jnnp.2010.239111>
7. Chollet, F.: Keras (2015). <https://github.com/fchollet/keras>
8. Courchesne, E., Carper, R., Akshoomoff, N.: Evidence of brain overgrowth in the first year of life in autism. *JAMA* **290**, 337–344 (2003). <https://doi.org/10.1001/jama.290.3.337>
9. Defferrard, M., Bresson, P., Vandergheynst, X.: Convolutional neural networks on graphs with fast localized spectral filtering. In: NIPS, pp. 3844–3852 (2016)
10. DeRamus, T., Kana, R.: Anatomical likelihood estimation meta-analysis of grey and white matter anomalies in autism spectrum disorders author links open overlay panel. *NeuroImage: Clin.* **7**, 525–536 (2015). <https://doi.org/10.1016/j.nicl.2014.11.004>
11. Di Martino, A., et al.: Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci. Data* **4**, 170010 (2017). <https://doi.org/10.1038/sdata.2017.10>
12. Di Martino, A., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* **19**, 659–67 (2014). <https://doi.org/10.1038/mp.2013.78>
13. Dinstein, I., Haar, S., Atsmon, S., Schtaerman, H.: No evidence of early head circumference enlargements in children later diagnosed with autism in Israel. *Mol. Autism* **8** (2018). <https://doi.org/10.1186/s13229-017-0129-9>
14. Dolgin, E.: This is your brain online: the functional connectomes project. *Nat. Med.* **16**, 351 (2010). <https://doi.org/10.1038/nm0410-351b>
15. Erhan, D., Bengio, Y., Courville, A., Vincent, P.: Visualizing higher-layer features of a deep network. Technical report 1341, University of Montreal (2009)
16. Finn, E., et al.: Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci.* **18**, 1664–1671 (2015). <https://doi.org/10.1038/nn.4135>

17. Haar, S., Berman, S., Behrmann, M., Dinstein, I.: Anatomical abnormalities in autism? *Cereb. Cortex* **26**, 1440–1452 (2016). <https://doi.org/10.1093/cercor/bhu242>
18. Hall, D., Huerta, M., McAuliffe, M., Farber, G.: Sharing heterogeneous data: the national database for autism research. *Neuroinformatics* **10**, 331–339 (2012). <https://doi.org/10.1007/s12021-012-9151-4>
19. Hamilton, W., Ying, R., Leskovec, J.: Representation learning on graphs: methods and applications. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* (2017)
20. Hasson, U., Nusbaum, H., Small, S.: Task-dependent organization of brain regions active during rest. *PNAS* **106**, 10841–10846 (2009). <https://doi.org/10.1073/pnas.0903253106>
21. Hechtlinger, Y., Chakravarti, P., Qin, J.: A generalization of convolutional neural networks to graph-structured data. *arXiv* (2017)
22. Heinsfeld, A., Franco, A., Craddock, R., Buchweitz, A., Meneguzzia, F.: Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: Clin.* **17**, 16–23 (2018). <https://doi.org/10.1016/j.nicl.2017.08.017>
23. Kawahara, J., et al.: BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage* **146**, 1038–1049 (2017). <https://doi.org/10.1016/j.neuroimage.2016.09.046>
24. Keown, C., Datko, M., Chen, C., Maximo, J., Jahedi, A., Müller, R.: Network organization is globally atypical in autism: a graph theory study of intrinsic functional connectivity. *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging* **2**, 66–75 (2017). <https://doi.org/10.1016/j.bpsc.2016.07.008>
25. Khundrakpam, B., Lewis, J., Kostopoulos, P., Carbonell, F., Evans, A.: Cortical thickness abnormalities in autism spectrum disorders through late childhood, adolescence, and adulthood: a large-scale MRI study. *Cereb. Cortex* **27**, 1721–1731 (2017). <https://doi.org/10.1093/cercor/bhx038>
26. Kipf, T., Welling, M.: Semi-supervised classification with graph convolutional neural networks. In: *ICLR 2017* (2017)
27. Kotikalapudi, R., Contributors: keras-vis (2017). <https://github.com/raghakot/keras-vis>
28. Khosla, M., Jamison, K., Kuceyeski, A., Sabuncu, M.R.: 3D convolutional neural networks for classification of functional connectomes. In: Stoyanov, D., et al. (eds.) *DLMIA/ML-CDS-2018*. LNCS, vol. 11045, pp. 137–145. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_16
29. Nikolentzos, G., Meladianos, P., Tixier, A.J.-P., Skianis, K., Vazirgiannis, M.: Kernel graph convolutional neural networks. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds.) *ICANN 2018*. LNCS, vol. 11139, pp. 22–32. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01418-6_3
30. Patel, A., Bullmore, E.: A wavelet-based estimator of the degrees of freedom in denoised fmri time series for probabilistic testing of functional connectivity and brain graphs. *NeuroImage* **142**, 14–26 (2016). <https://doi.org/10.1016/j.neuroimage.2015.04.052>
31. Piven, J., Arndt, S., Bailey, J., Haverkamp, S., Andreasen, N., Palmer, P.: An MRI study of brain size in autism. *Am. J. Psychiatry* **152**, 1145–1149 (1995). <https://doi.org/10.1176/ajp.152.8.1145>
32. Plitt, M., Barnes, K., Martin, A.: Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *NeuroImage Clin.* **7**, 359–66 (2015). <https://doi.org/10.1016/j.nicl.2014.12.013>

33. Poldrack, R., et al.: Toward open sharing of task-based fMRI data: the OpenfMRI project. *Front. Neuroinform.* **7** (2013). <https://doi.org/10.3389/fninf.2013.00012>
34. Poldrack, R., Gorgolewski, K.: OpenfMRI: open sharing of task fMRI data. *NeuroImage* **144**, 259–261 (2017). <https://doi.org/10.1016/j.neuroimage.2015.05.073>
35. Redcay, E., Courchesne, E.: *Biol. Psychiatry* **58**, 1–9 (2005). <https://doi.org/10.1016/j.biopsych.2005.03.026>
36. Russakovsky, O., et al.: ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015). <https://doi.org/10.1007/s11263-015-0816-y>
37. Satterthwaite, T., et al.: Linked sex differences in cognition and functional connectivity in youth. *Cereb. Cortex* **25**, 2383–2394 (2015). <https://doi.org/10.1093/cercor/bhu036>
38. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep inside convolutional networks: visualising image classification models and saliency maps. In: *Workshop at International Conference on Learning Representations* (2014)
39. Tomasi, D., Volkow, N.: Gender differences in brain functional connectivity density. *Hum. Brain Mapp.* **33**, 849–860 (2013). <https://doi.org/10.1002/hbm.21252>
40. Wang, W., et al.: Altered resting-state functional activity in patients with autism spectrum disorder: a quantitative meta-analysis. *Front. Neurol.* **9**, 556 (2018). <https://doi.org/10.3389/fneur.2018.00556>
41. Yang, J., Hofmann, J.: Action observation and imitation in autism spectrum disorders: an ALE meta-analysis of fMRI studies. *Brain Imaging Behav.* **10**, 960–969 (2016). <https://doi.org/10.1007/s11682-015-9456-7>
42. Zhang, W., Groen, W., Mennes, M., Grevén, C., Buitelaar, J., Rommelse, N.: Revisiting subcortical brain volume correlates of autism in the abide dataset: effects of age and sex. *Psychol. Med.* **48**, 654–668 (2018). <https://doi.org/10.1017/S003329171700201X>