Pipeline comparisons of convolutional neural networks for structural connectomes: predicting sex across 3,152 participants

Pipeline comparisons of convolutional neural networks for structural connectomes: predicting sex across 3,152 participants

Hon Wah Yeung[1], Saturnino Luz[2], Simon R. Cox[3],
Colin R. Buchanan[3], Heather C. Whalley[1] and Keith M. Smith[2,4]

***Abstract*—With several initiatives well underway towards amassing large and high-quality population-based neuroimaging datasets, deep learning is set to push the boundaries of what is possible in classification and prediction in neuroimaging studies. This includes those that derive increasingly popular structural connectomes, which map out the connections (and their relative strengths) between brain regions. Here, we test different Convolutional Neural Network (CNN) models in a benchmark sex prediction task in a large sample of N=3,152 structural connectomes acquired from the UK Biobank, and compare results across different connectome processing choices. The best results (76.5% test accuracy) were achieved using Fractional Anisotropy (FA) weighted connectomes, without sparsification, and with a simple weight normalisation through division by the maximum FA value. We also confirm that for structural connectomes, a Graph CNN approach, the recently proposed BrainNetCNN, outperforms an image-based CNN.**

I. Introduction

There is increasing interest in data science for developing new machine learning methods, with deep learning being a rapidly emerging field. With recent advances in neuroimaging, scans with high resolution are now widely produced in increasing quantities, providing ever greater potential for machine learning to make strides in aiding classification of brain disorders and diseases from medical images. It is thus imperative to work towards the most robust and powerful methodological pipelines for such classifications.

A Convolutional Neural Network (CNN) is a particular class of deep neural network which employs a mathematical operation called convolution. It consists of an input layer, an output layer and several hidden layers in between [1]. This kind of neural network is suitable for images, including brain imaging data. Previous studies have performed CNN on T1-Weighted MRI scans and achieved successful results for sex and age prediction [2], [3]. Researchers are also now looking at the possibilities of applying deep learning techniques to brain imaging data in the context of mental health and neurological research. For example, several research groups have worked on classification of Alzheimer's disease [4], [5].

Apart from imaging techniques that capture the morphometry of the brain, the brain's structural connectivity map, the connectome, can also be extracted from diffusion MRI (dMRI), with connections encoded in adjacency matrices. It is believed that healthy people and those with mental illness do not only differ in brain morphometric measures but also in brain connectivity patterns. Therefore, incorporating connectivity measures should provide more information to the learning algorithm and, thereby, provide better results for classification. Conventional CNN works perfectly on 3D T1-weighted MRI scans as it can capture the spatial locality of the brain image. However, feeding in adjacency matrices into the CNN could lead to misrepresentations of the structural connectome. This is because the topological locality of a graph is different from looking at the spatial locality, which a conventional convolution layer is designed for, of the adjacency matrix. To this end, Kawahara et al. [6] recently proposed a deep neural network model called BrainNetCNN. This network is composed of convolutional filters (edge-to-edge, edge-to-node and node-to-graph convolutional filters) that are able to capture the topological locality of structural brain networks. These filters are derived from basic properties of an adjacency matrix [6].

In this study, we explore two different deep neural network constructions for the connectomes: i) The conventional ImageCNN and ii) the BrainNetCNN proposed in [6], based on their ability in the benchmark task of predicting sex. More importantly, we explore the effect of several key considerations of connectome definitions on performance including i) weight definition: comparing Mean Diffusivity (MD) and Fractional Anisotropy (FA); ii) sparsity and binarisation: comparing the full weighted network against 5% density weighted and binary networks; iii) weight normalisation: comparing original weights against widespread normalisations. Moreover, this study used a dataset with 3152 participants, which was significantly larger than that in [6] (115 infants), and therefore could better test the generalisability of this novel method.

II. Materials

The network data used in this study have been published previously [7] and are outlined below.

A. Participants

Participants were recruited from the UK Biobank. A subset of participants underwent brain MRI at the UKB imaging centre in Cheadle, Manchester, UK. The initial release of diffusion MRI (dMRI) data included 5,455 participants. In total, 567 were excluded from the current study at the stage of scanning, due to incompatible dMRI acquisition. By following the dMRI quality control procedures as suggested in UKB Brain Imaging Documentation, a further 1,314 participants were removed prior to release.

[1]Division of Psychiatry, University of Edinburgh, Kennedy Tower h.w.yeung@sms.ed.ac.uk
[2]Usher Institute, University of Edinburgh, Edinburgh, UK
[3]Department of Psychology, University of Edinburgh, Edinburgh, UK
[4]Health Data Research UK, London, UK

B. MRI Acquisition

All imaging data were acquired using a single Siemens Skyra 3T scanner. 3D T1-weighted volumes were acquired using a MP RAGE sequence at $1 \times 1 \times 1$ mm resolution with $208 \times 256 \times 256$ field of view. The dMRI data were acquired using a spin-echo EPI sequence (50 $b = 1000s/mm^2$, 50 $b = 2000s/mm^2$ and 10 $b = 0s/mm^2$) resulting in 100 distinct diffusion-encoding directions, FOV $= 104 \times 104$ mm, imaging matrix $= 52 \times 52$, 72 slices, slice thickness $= 2$ mm. Water diffusion parameters were estimated for FA, which measures the degree of anisotropic water molecule diffusion, and MD, which measures the magnitude of diffusion. Details of the MRI protocol and processing can be found in [8], [9].

C. Network Construction

Each T1-weighted image was segmented into 85 distinct neuroanatomical Regions-Of-Interest (ROI) using volumetric segmentation and cortical reconstruction (FreeSurfer v5.3.0), 34 cortical structures per hemisphere were identified using the Desikan-Killany atlas [10]. Brain stem, accumbens area, amygdala, caudate nucleus, hippocampus, pallidum, putamen, thalamus and ventral diencephalon were also extracted with FreeSurfer. A cross-modal nonlinear registration method was used to align ROIs from T1-weighted volume to diffusion space (skull stripping [11], initial alignment by affine transformation with 12 degrees of freedom (FLIRT; [12]) followed by a nonlinear deformation method (FNIRT; [13])).

Networks were constructed by identifying connections between all ROI pairs. The endpoint of a streamline was recorded as the first ROI encountered when tracking from the seed location. Successful connections were recorded in an 85×85 adjacency matrix. Two network weightings, FA and MD, were computed. For each weighting, an adjacency matrix was computed with element, a_{ij}, recording the mean value of the diffusion parameter in voxels identified along all interconnecting streamlines between nodes i and j. All matrices were made symmetric since afferent and efferent connections are indistinguishable for tractography. Self-connections were removed, setting diagonal entries to zero.

In total, 3,152 participants (44.6–77.1 years of age, 1,495 male) remained after participants were excluded at quality control or due to failure in processing. On average, 6.01 million streamlines were seeded per subject of which 1.49 million (24.9%) were found to successfully connect between nodes following the tracking procedure and removal of self-connections. The FA and MD networks were produced from the same set of streamlines, where the range of values of MD is 0 - 0.003 $\times 10^{-3} mm^2/s$ and for FA is 0 - 0.9. Before any thresholding was introduced, the mean value of network density (percentage of non-zero entries in a adjacency matrix) across subjects was 68.4% (SD = 3.2). Proportional-thresholding was used to keep only connections present in at least 2/3 of subjects, which result in connection density of $\sim$60% after thresholding. Both of the dMRI-based weightings followed approximately normal distributions.

III. Methods

In this analysis, two different CNN models were used: a conventional ImageCNN and the newer BrainNetCNN. Further density thresholding and binary graphs were also considered, four different ways of adjacency matrix transformations were employed. This gives twelve different representations of the original adjacency matrices for both FA and MD.

A. ImageCNN

The ImageCNN model in this study was composed of some basic layers of CNN. More details of the layers can be found in [1], with architecture in Figure 1a. The 85×85 adjacency matrices were fed into a 3×3 2D convolution, with zero padding, stride 2 and 16 filters, and then followed by a ReLU layer, batch normalization layer as well as a 2×2 Max Pooling Layer with stride 2. The layers from the 2D convolution to batch normalization layer was repeated again, but with 32 filters instead. There were then the fully connected layers with 256, 64 and 2 nodes respectively, having a ReLU layer and a dropout layer with $prob = 0.5$ in between the fully connected layer. Finally, there was the classification layer as well as the softmax layer.

B. BrainNetCNN

On top of the basic layers, some special layers were designed for the BrainNetCNN [6]:

1) Edge-to-Edge (E2E) layer: This computed a weighted sum of edge weights between shared nodes for each of the edges. This worked like a convolution layer with a cross-shaped filter, which is, the (i, j) entry output is given by weighted sum of the i-th row and weighted sum of the j-th column. This can be written as:

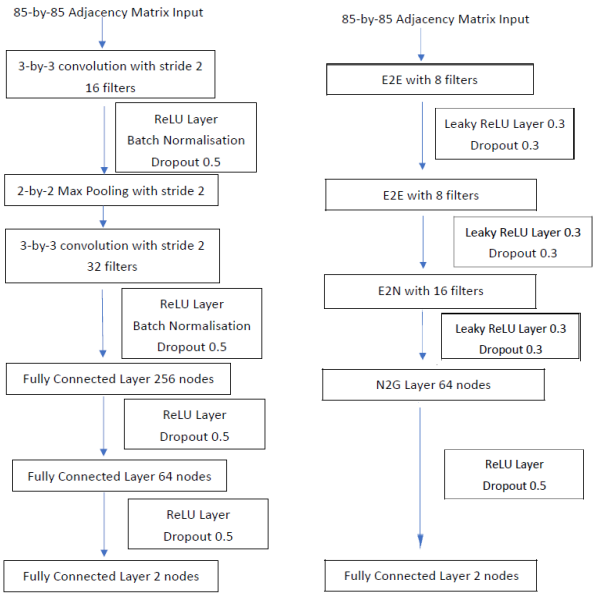
$$(A^{k,n})_{ij} = \sum_{f=1}^{F_{k-1}} A_{i,:}^{k-1} \cdot w_r^{k-1,f,n} + A_{:,j}^{k-1} \cdot w_c^{k-1,f,n}$$

where $A^{k,n}$ is the filtered adjacency, $A_{i,:}$ is the i-th row of A and $A_{:,j}$ is the j-th column of A, and $w_r^{k-1,f,n}$ and $w_c^{k-1,f,n}$ are the learnt row and column weights for the n-th filter at the $k-1$-th layer,respectively. F_{k-1} is the number of feature maps at the $k-1$-th layer.
2) Edge-to-Node (E2N) layer: In the original paper, due to inconsistencies arising from combining the row and column 1D convolutions, they only took the row 1D convolution to represent the node responses. We replicated this here.
3) Node-to-graph (N2G) layer: This was basically a fully connected layer connected to an E2N layer.

Choosing a larger number of filters did not improve the results so we chose to use 8 filters in E2E layer. The Architecture of the BrainNetCNN is shown in Figure 1b.

C. Graph Pruning and Graph Adjacency transformation

The adjacency matrices have densities of $\sim$60%, which some consider as over-connected and containing false connections [14]. Therefore, we limit the density to 5% by



(a) ImageCNN Architecture (b) BrainNetCNN Architecture

Fig. 1: Architectures of both neural network models, both were connected to classification layer and softmax layer at the end.

retaining the strongest weighted connections. The binarized version of the 5%-density graphs were also considered.

Different representations of the adjacency matrices may also increase the effectiveness and efficiency of model training. Hence, three adjacency matrix transformations were considered:

1) Symmetric Normalization (SymNorm):

$$A_{sym} = D^{-1/2} A D^{-1/2}$$

2) Random Walk Normalization (RWNorm):

$$A_{rw} = D^{-1} A$$

3) Maximum Value Normalization (MaxNorm):

$$A_{max} = \frac{A}{\max\limits_{1 \le i,j \le 85} A_{ij}}$$

where A is the adjacency matrix and D the degree matrix. In total, we had three types of adjacency matrices (original, 5% density, binarized 5% density) together with four types of transformations (original weights, SymNorm, RWNorm and MaxNorm), which gave twelve different adjacency matrix form for each of MD and FA weights.

D. Hyperparameters for the models

For both models we used stochastic gradient descent and the same set of hyperparameters for training, with momentum $= 0.9$, learning rate $= 0.01$, mini batch size $= 128$, L_2-norm regularization $= 0.0005$, validation frequency $= 30$. Training was stopped when no better validation accuracy was found in the next six subsequent validations.

E. Experimental set-up

Fifteen shuffle splits with 70% /15% /15% for training\validation\test were chosen and same set of splits were used for the two different deep learning models on each adjacency matrix form. The mean and standard deviation of the accuracies for each part of the split were then recorded.

IV. Results

For original weights, both models trained when FA matrices were used but failed to train on the MD matrices, as shown in Table I. Moreover, they also failed to train on the 5% density graph of the original MD matrices, though training was succesful for the binarized 5% density graph.

Weight	Mean Diffusion		Fractal Anisotropy	
CNN	Image	BrainNet	Image	BrainNet
Validation	52.3(2.38)	55.5(6.74)	72.4(2.51)	74.8(3.00)
Training	52.0(5.08)	58.0(7.63)	76.4(6.00)	83.0(4.74)
Test	52.5(2.29)	55.8(5.89)	72.2(2.75)	75.3(2.46)

TABLE I: The performances (accuracy mean percentage $\pm$ standard deviation) on sex prediction for ImageCNN and BrainNetCNN using original mean diffusion and fractional anistropy matrices.

The binarized 5% density graph on FA performed more consistently over different adjacency transformations for both models. Tables II and III show the results for ImageCNN and BrainNetCNN, respectively. Less overfitting (i.e. more consistency between training and test) was seen in the BrainNetCNN.

Normalisation	Original	SymNorm	RWNorm	MaxNorm
Validation	71.0(2.01)	68.3(1.62)	68.8(1.56)	70.4(1.74)
Training	97.0(1.98)	98.0(1.72)	96.4(1.79)	97.6(2.10)
Test	71.3(2.12)	69.2(2.13)	69.2(2.23)	71.2(1.58)

TABLE II: Sex prediction performances (accuracy mean percentage $\pm$ standard deviation) with ImageCNN for different adjacency transformations over the binarized 5% density graph of FA.

Normalisation	Original	SymNorm	RWNorm	MaxNorm
Validation	71.7(1.93)	68.7(2.27)	70.8(2.27)	72.2(1.89)
Training	86.6(2.53)	93.4(1.85)	85.3(3.50)	85.7(3.05)
Test	72.0(2.28)	68.7(2.02)	71.3(2.26)	71.9(1.65)

TABLE III: Sex prediction performances (accuracy mean percentage $\pm$ standard deviation) with BrainNetCNN for different adjacency transformations over the binarized 5% density graph on FA.

However, the best performance for both models was seen with MaxNorm on original weighted FA matrices, as shown in Table IV.

CNN	Image	BrainNet
Validation	73.3(3.87)	76.7(1.74)
Training	81.8(5.91)	84.2(3.06)
Test	73.6(2.86)	76.5(2.15)

TABLE IV: Performances (accuracy mean percentage $\pm$ standard deviation) on sex prediction for ImageCNN and BrainNetCNN using MaxNorm on original weighted FA matrices.

For each of the twelve transformations for the adjacencies of MD and FA described above, BrainNetCNN

(71.0%(4.37)) performed consistently better than the ImageCNN (68.6%(4.83)) in terms of mean accuracies (though differences were within 3%) and generally provided more consistent results across the shuffle splits.

V. Discussion

Our results show that BrainNetCNN outperforms ImageCNN in sex prediction on structural connectomes. This demonstrates that incorporating the non-local (in terms of matrix position) connectivity information from the adjacency matrix is more powerful than naively treating connectome adjacency matrices as images in a regular CNN framework.

In terms of connectivity measures, FA measures provided better and more consistent performances over all adjacency transformations, in terms of test accuracies, for both ImageCNN (70.2%(1.90)) and BrainNetCNN (72.3%(2.02)) than using the MD measures (ImageCNN: 67.0%(6.15), BrainNetCNN: 70.0%(5.54)). This agrees with previous analysis of sex in UKB connectomes [15] and supports the view that FA may be more appropriate as a connectivity measure as it accounts for the directionality of the diffusion measured across the tracts [16].

The best results are achieved on unthresholded, weighted connectomes with normalisation by maximum weight. On the other hand, the performances on binarized 5% density adjacency matrices over the 4 different transformations are more consistent than original weighted and 5% density weighted connectomes. There is much debate over the density required to perform robust connectome analysis. Some view weak connections as spurious and to be excluded [14], others view more connections as more information, allowing for some degree of noise in the data [17]. This may explain why the fully weighted connectomes show generally the best performance while sparse binary representations have greater consistency. Furthermore, while weights can cause complications in network analysis [17] this does not appear to extend to hypothesis-free classifications using GCNNs.

As far as the transformation of the adjacency matrices is concerned, both models perform the best using MaxNorm for FA measures. This indicates that linear transformations preserve meaningful information in the weights which appear to be obscured by non-linear transformations.

VI. Conclusion

In conclusion, the BrainNetCNN outperforms an ImageCNN in sex classification. This possibly implies that the BrainNetCNN is better at capturing important features and properties of adjacency matrices. Suitable alterations on the adjacency matrices may yield better results. Further analysis is needed for investigating the effect of alterations on adjacency matrices and optimal graph density, visualising the activations of certain neural network layers and identifying which regional connections have higher predictive power.

Acknowledgements

This study was supported and funded by the Wellcome Trust Strategic Award "Stratifying Resilience and Depression Longitudinally" (STRADL) (Reference 104036/Z/14/Z), and was also supported by National Institutes of Health (NIH) research grant R01AG054628. The research was conducted using the UK Biobank resource, with approved project number 10279. Structural brain imaging data from UK Biobank was processed at the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), which is a part of the cross-council Lifelong Health and Wellbeing Initiative (MR/K026992/1). CCACE received funding from Biotechnology and Biological Sciences Research Council (BBSRC), Medical Research Council (MRC), and was also supported by Age UK as part of The Disconnected Mind project. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF). KMS was supported by Health Data Research UK, an initiative funded by UK Research and Innovation Councils, NIH Research (England) and the UK devolved administrations, and leading medical research charities. SRC was also supported by the Medical Research Council (MR/R024065/1).

References

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[2] J. H. Cole *et al.*, "Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker," *NeuroImage*, vol. 163, pp. 115–124, 2017.

[3] L. Yuan, X. Wei, H. Shen, L.-L. Zeng, and D. Hu, "Multi-center brain imaging classification using a novel 3D CNN approach," *IEEE Access*, vol. 6, pp. 49 925–49 934, 2018.

[4] S. E. Spasov, L. Passamonti, A. Duggento, P. Liò, and N. Toschi, "A multi-modal convolutional neural network framework for the prediction of alzheimer's disease," in *Proceedings of the EMBC 2018*, 2018, pp. 1271–1274.

[5] M. Liu, J. Zhang, E. Adeli, and D. Shen, "Landmark-based deep multi-instance learning for brain disease diagnosis," *Medical image analysis*, vol. 43, pp. 157–168, 2018.

[6] J. Kawahara *et al.*, "BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.

[7] C. R. Buchanan *et al.*, "The effect of network thresholding and weighting on structural brain networks in the UK Biobank," *NeuroImage*, p. 116443, 2020.

[8] F. Alfaro-Almagro *et al.*, "Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank," *Neuroimage*, vol. 166, pp. 400–424, 2018.

[9] K. L. Miller *et al.*, "Multimodal population brain imaging in the uk biobank prospective epidemiological study," *Nature neuroscience*, vol. 19, no. 11, p. 1523, 2016.

[10] R. S. Desikan *et al.*, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[11] S. M. Smith, "Fast robust automated brain extraction," *Human brain mapping*, vol. 17, no. 3, pp. 143–155, 2002.

[12] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[13] J. L. Andersson *et al.*, "Non-linear registration aka spatial normalisation fmrib technial report tr07ja2," *FMRIB Analysis Group of the University of Oxford*, 2007.

[14] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *PLOS Computational Biology*, vol. 3, 02 2007.

[15] S. J. Ritchie *et al.*, "Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants," *Cerebral Cortex*, vol. 28, no. 8, pp. 2959–2975, 05 2018.

[16] J. Soares, P. Marques, V. Alves, and N. Sousa, "A hitchhiker's guide to diffusion tensor imaging," *Frontiers in neuroscience*, vol. 7, 2013.

[17] K. Smith, D. Abásolo, and J. Escudero, "Accounting for the complex hierarchical topology of EEG phase-based connectivity in network binarisation," *PLOS ONE*, vol. 12, no. 10, 2017.