Deep multivariate autoencoder for capturing complexity in Brain Structure and Behaviour Relationships

No Author Given

No Institute Given

Abstract. Diffusion MRI is a powerful tool that serves as a bridge between brain microstructure and cognition. Recent advancements in cognitive neuroscience have highlighted the persistent challenge of understanding how individual differences in brain structure influence behavior, especially in healthy people. While traditional linear models like Canonical Correlation Analysis (CCA) and Partial Least Squares (PLS) have been fundamental in this analysis, they face limitations, particularly with high-dimensional data analysis outside the training sample. To address these issues, we introduce a novel approach using deep learning—a multivariate autoencoder model—to explore the complex non-linear relationships between brain microstructure and cognitive functions.

Keywords: diffusionMRI · Cognitive decoding · Multivariate Learning.

1 Introduction

In cognitive neuroscience, a significant gap remains in understanding how interindividual differences in brain structure affect behavior [3]. Diffusion magnetic resonance imaging (dMRI) provides insights into tissue microstructure [1] enabling a more comprehensive understanding of the relationship between brain architecture and behavior.

Recent advancements in the field have shifted from one-to-one mappings between brain regions and cognition, derived from focal brain lesion studies [2], to a regional multivariate perspective [3]. Despite these advancements, challenges persist, particularly in research on healthy individuals. Most current knowledge is biased towards pathological conditions, not reflecting the complexity of brain-behavior relationships in the general population. Menon et al. [5] are among the few studies that relate GM microstructure and cognition in healthy individuals.

The Brain Structure and Behavior (BSB) community has recently emphasized two critical aspects on dMRI: techniques and methodological models [3]. From the technical perspective, the focus has primarily been on tractography and white matter, leaving grey matter and microstructure underexplored [3]. Additionally, methodologically, multivariate linear models used in BSB for dMRI data, like Canonical Correlation Analysis (CCA) and Partial Least Squares (PLS), face challenges in generalizability [3, 4]. These limitations highlight the need for more sophisticated models [5, 3, 10].

2 Methodology

We developed a multivariate Encoder-Decoder model to predict cognitive processes from diffusion MRI (dMRI) signals. Data preprocessing involved filtering subjects with missing information and outliers. We used the 3T dMRI and cognitive data from 779 subjects in the HCP database, focusing on the insula for its role in encoding cognitive-control functions, as noted by Menon et al. and others [5, 11, 12]. We standardized each cognitive feature at batch level and restored them to their original values post-prediction for consistent evaluation. To define the brain structure data, and based on previous works [7, 8, 9], we hypothesize that the diffusion signal is modulated by microstructure and that microstructure has a role in modulating cognition and performed a multishell analysis with 3 b-values.

The proposed model, illustrated in fig. 1, features two encoder modules—one for brain structure data ($\phi(X) = z$) and another for cognitive data($\theta(Y) = z'$)—each compressing their respective inputs into latent-dimensional embeddings $zandz'(n = 64)$. These embeddings are then used by a shared decoder (ψ) to reconstruct cognitive data, capturing complex, non-linear relationships between brain structure and behavior. The model's training involved sequentially optimizing the encoders and the decoder using mean squared error loss functions, ensuring consistent and accurate reconstructions of cognitive features from the embeddings:

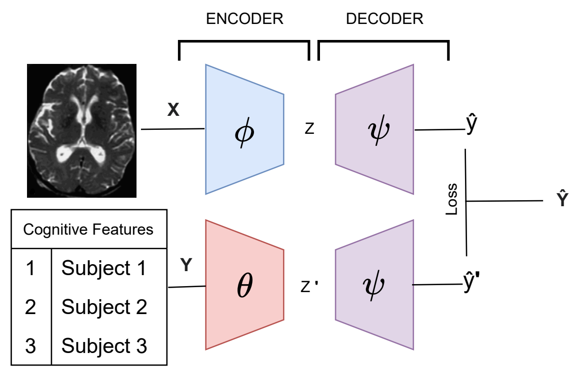

Fig. 1: The model architecture features in two encoders ϕ and θ that transform X (dMRI data) and Y (cognitive data), respectively, into lower-dimmensional embedding spaces (z and z'). These embeddings are then processed by a shared decoder ϕ which reconstructs the cognitive data to produce $\hat{y}$ and $\hat{y}'$, facilitating the prediction of cognitive functions. The design allows for the exploration of complex multivariate patterns between brain structure and cognitive features.

$$\mathcal{L}_{\text{encoders}} = \mathcal{L}(z, z')$$

$$\mathcal{L}_{\text{decoder}} = \alpha\mathcal{L}(z, z') + \beta\mathcal{L}(y, \hat{y}) + \gamma\mathcal{L}(y, \hat{y'})$$

Where α, β, and γ are weight hyperparameters that can be adjusted to balance the importance of each term. In our analysis, all the losses had the same importance $\alpha = \beta = \gamma = 1$.

3 Results and Discussion

The performance of our multivariate Encoder-Decoder model was compared with traditional methods (CCA and PLS) using $k = 5$ k-fold cross-validation to ensure robust evaluation. We used Spearman correlation values, to measure predictive accuracy, and the results are shown in table 1 alongside our model's best results obtained for one sample. These values indicate that our model outperforms CCA and PLS in predicting multivariate cognitive processes during the validation phase, demonstrating superior out-of-sample performance. Despite the noticeable gap between training and validation results, that suggests potential overfitting, in the context of complex phenomena such as BSB, our primary focus is on achieving good out-of-sample performance rather than minimizing the gap between training and validation set performance. These correlations results align with significant findings in the field, such as those by Menon et al., where correlations around 0.2 are considered high, affirming a meaningful relationship between dMRI attenuations and behavior. Therefore, the table reflects the peak performance of our model from one sample, highlighting its potential despite variability in generalization.

Feature	CCA		PLS		Ours	
	Training	**Validation**	**Training**	**Validation**	**Training**	**Validation**
Age	**1.00**	**0.17**	0.22	0.13	0.37	0.03
Working Memory Acc.	0.45	0.14	0.47	0.16	**0.57**	**0.20**
Working Memory RT	0.24	0.14	0.32	0.06	**0.49**	**0.26**
Relational Task Acc.	0.34	**0.16**	0.45	0.01	**0.37**	0.039
Relational Task RT	0.25	**0.15**	0.15	0.11	**0.32**	0.054
Gambling Task	0.30	0.15	0.09	0.09	**0.47**	**0.20**
Gambling Task RT	0.11	0.03	0.25	0.00	**0.52**	**0.17**
List Sorting	0.17	-0.05	0.24	0.16	**0.57**	**0.29**
Flanker	0.03	0.06	0.35	0.07	**0.54**	**0.07**
Card Sorting	0.37	0.12	**0.47**	0.15	0.41	**0.21**
Picture Sequence	0.19	0.04	0.16	**0.19**	**0.55**	-0.01
Processing Speed	0.20	0.04	0.32	0.06	**0.47**	**0.21**

Table 1: Model comparison. Spearman's correlation factor for each cognitive feature in training and validation using CCA, PLS, and our model, based in multimodal learning. The highest values per row, for both training and validation, are in bold to indicate the best model performance.

4 Conclusions

This study demonstrates that our novel neural network-based model improves the prediction of cognitive functions from diffusion MRI (dMRI) data compared to traditional linear models such as Canonical Correlation Analysis (CCA) and Partial Least Squares (PLS). Our approach, which integrates distinct encoder modules for brain structure and cognitive data with a shared decoder module, offers a more understanding of the relationship between diffusion MRI attenuations and cognitive functions. The model's better performance, as indicated by superior validation Spearman coefficients, highlights its ability to capture complex, non-linear relationships that traditional models struggle to uncover.

Despite the promising results, the training-validation gap observed in our model suggests potential overfitting. This gap, larger than that seen in the PLS model, underscores the need for careful interpretation. It also indicates that cognitive variables might not be fully independent, as evidenced by the varied performance across cognitive features. These observations align with recent findings by Menon [5], which suggest that latent space models could provide further insights by uncovering underlying patterns not captured by existing methods.

Our study acknowledges that while the neural network model offers a significant advancement over classical approaches, it is not yet optimized for interpretability. Nonetheless, the model presents a flexible and powerful starting point for exploring brain-behavior relationships, particularly in handling the high-dimensional and variable nature of neuroimaging data.

Future research should focus on integrating additional data dimensions, such as structural MRI data, and exploring latent space models. These directions could reveal deeper patterns and relationships of the cognitive processes and their neural bases which will help in advancing both theoretical and practical applications in cognitive neuroscience.

Bibliography

[1] Le Bihan, D., Iima, M. (2015). Diffusion magnetic resonance imaging: what water tells us about biological tissues. *PLoS Biology*, **13**(7),

[2] Genon, S., Reid, A., Langner, R., Amunts, K., Eickhoff, S. B. (2018). How to characterize the function of a brain region. *Trends in Cognitive Sciences*, **22**(4), 350-364.

[3] Genon, S., Eickhoff, S. B., Kharabian, S. (2022). Linking interindividual variability in brain structure to behaviour. *Nature Reviews Neuroscience*, **23**(5), 307-318.

[4] Helmer, M., Warrington, S., Mohammadi-Nejad, A. R., Ji, J. L., Howell, A., Rosand, B., ... Murray, J. D. (2024). On the stability of canonical correlation analysis and partial least squares with application to brain-behavior associations. *Communications Biology*, **7**(1), 217.

[5] Menon, V., Gallardo, G., Pinsk, M. A., Nguyen, V. D., Li, J. R., Cai, W., Wassermann, D. (2020). Microstructural organization of human insula is linked to its macrofunctional circuitry and predicts cognitive control. *eLife*, **9**, e53470.

[6] Kang, N., Chung, S., Lee, S. H., Bang, M. (2024). Cerebro-cerebellar grey matter abnormalities associated with cognitive impairment in patients with recent-onset and chronic schizophrenia. *Schizophrenia*, **10**(1), 11.

[7] Afzali, M., Pieciak, T., Newman, S., Garyfallidis, E., Özarslan, E., Cheng, H., Jones, D. K. (2021). The sensitivity of diffusion MRI to microstructural properties and experimental factors. Journal of Neuroscience Methods, 347, 108951.

[8] Avram, A. V., Sarlls, J. E., Barnett, A. S., Özarslan, E., Thomas, C., Irfanoglu, M. O., ... Basser, P. J. (2016). Clinical feasibility of using mean apparent propagator (MAP) MRI to characterize brain tissue microstructure. Neuroimage, 127, 422-434.

[9] Fick, R. H., Wassermann, D., Caruyer, E., Deriche, R. (2016). MAPL: Tissue microstructure estimation using Laplacian-regularized MAP-MRI and its application to HCP data. NeuroImage, 134, 365-385.

[10] Mihalik, A., Chapman, J., Adams, R. A., Winter, N. R., Ferreira, F. S., Shawe-Taylor, J., ... Alzheimer's Disease Neuroimaging Initiative. (2022). Canonical correlation analysis and partial least squares for identifying brain–behavior associations: A tutorial and a comparative study. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 7(11), 1055-1067.

[11] Ham, T., Leff, A., de Boissezon, X., Joffe, A., Sharp, D. J. (2013). Cognitive control and the salience network: an investigation of error processing and effective connectivity. Journal of Neuroscience, 33(16), 7091-7098.

[12] Molnar-Szakacs, I., Uddin, L. Q. (2022). Anterior insula as a gatekeeper of executive control. Neuroscience Biobehavioral Reviews, 139, 104736.