MICCAI paper outline

1 Problem outline

Quickly reintroduce the problem and objective of lesion-behaviour mapping : finding which brain regions play a role in certain types of behavior through measured scores from patients with lesions. Quickly explain topographical bias, explain that it causes univariate methods to fail

2 Current misconceptions

Cite the original paper which concludes that multivariate methods (i.e. ML-based) fail due to topographical bias, and that only causal methods can overcome this bias. "Original paper" refers to https://www.biorxiv.org/content/10.1101/2019.12.17.878355v1

3 Causal structure of the problem

Give a simple graph outlining the structure of the problem (see figure 1 for a sample) , explain that there is no confounding bias, because there is no backdoor path between the causes and the outcome, and so causal methods are not necessary. Mention that this is only true with this amount of observed variables, and that additional observations (age, etc) could be confounders that
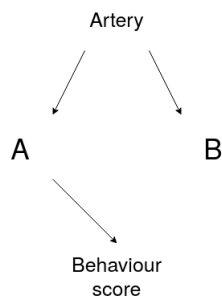

FIGURE 1 – sample figure

would warrant the use of causal inference. It's also worth mentioning that there IS a backdoor path when two or more regions influence the behaviour scores.

Q : Not sure if I should mention causal inference methods for hidden confounders anyways. They seem to still be pretty experimental and controversial (like Blei's paper). Perhaps I could make a passing mention to instrumental variables.

4 Experiments

4.1 Original paper experiments

Quickly describe the experiments of the original paper. Say that we believe the SVM model (as opposed to multivariate methods in general) to be susceptible to topographical bias

4.2 Our experiments

Describe the dataset used : Lesymap, the same one used in the original paper.
Describe the way we generate behavioral scores : single ROI, association of two ROIs with OR / AND / + associations, cite https://europepmc.org/article/med/9712015 to show that the choice of simulated outcomes is grounded in the lesion literature.

Show there is topographical bias involved with the ROIs picked : A glass brain plot would be most elegant, but I'm not sure how to make one that looks good.

Describe the models we use : BART (as our "causal model" representative), Desparsified Lasso, SVR (because it was in the original paper), and Random Forest + feature importance (to see if the good performance of BART on AND/OR scenarios can be credited to the tree structure of its underlying model).

Describe the way we assert model performance in ROI recovery : we compute Z-scores for the weights / feature importance / ATEs of each model (so each brain region gets a Z-score), and then use these Z-scores to compute precision/recall curves. We then take the AUC of these curves as the final indicator of model performance.

Show our results (see figure 2 for a sample)

5 Explanation in the case of a linear model

(Thanks to Julie for the idea) Take the simple case of a linear causal model, and show that thanks to strong ignorability, this causal model amounts to doing regression conditional on the regions not considered as treatment (ex : investigating the effect of ROI n, conditioning of the brain regions that are not n.)

Though I see a few unclear areas in this reasoning, and I will probably need to discuss it further with Julie when writing that part.

6   Conclusion

Conclude that our experiments illustrate the fact that topographical bias does not require causal inference methods to be overcome.

Explain the possibility of going causal by measuring added variables which could act as confounders (age, etc), but are unfortunately not found in public brain lesion datasets, or by using algorithms that model latent unobserved confounders.
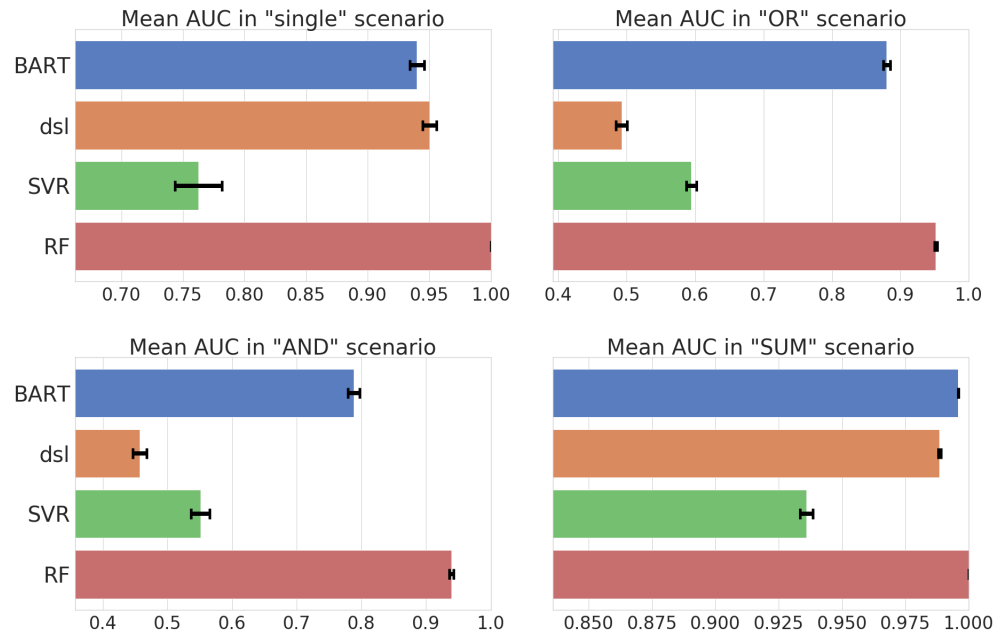

FIGURE 2 – Sample figure yet again, will need to adjust plot width and add a bar for Shapley values