The missing link: Predicting connectomes from noisy and partially observed tract tracing data

Our understanding of the wiring map of the brain, known as the connectome, has increased greatly in the last decade, mostly due to technological advancements in neuroimaging techniques and improvements in computational tools to interpret the vast amount of available data. Despite this, with the exception of the C. elegans roundworm, no definitive connectome has been established for any species. In order to obtain this, tracer studies are particularly appealing, as these have proven highly reliable. The downside of tract tracing is that it is costly to perform, and can only be applied ex vivo. In this paper, we suggest that instead of probing all possible connections, hitherto unknown connections may be predicted from the data that is already available. Our approach uses a 'latent space model' that embeds the connectivity in an abstract physical space. Regions that are close in the latent space have a high chance of being connected, while regions far apart are most likely disconnected in the connectome. After learning the latent embedding from the connections that we did observe, the latent space allows us to predict connections that have not been probed previously. We apply the methodology to two connectivity data sets of the macaque, where we demonstrate that the latent space model is successful in predicting unobserved connectivity, outperforming two baselines and an alternative model in nearly all cases. Furthermore, we show how the latent spatial embedding may be used to integrate multimodal observations (i.e. anterograde and retrograde tracers) for the mouse neocortex. Finally, our probabilistic approach enables us to make explicit which connections are easy to predict and which prove difficult, allowing for informed follow-up studies.