Lesion-symptom mapping using deep learning and explainable AI in small vessel disease

Cerebral small vessel disease (SVD) is a common cerebrovascular disorder that typically occurs during aging and can result in stroke, cognitive impairment, as well as behavioural or functional problems. It often co-occurs with vascular cognitive impairment, causing long-term disabilities and worsening the quality of life of the patients. The diagnosis can be challenging and, it is often unclear which factors cause these brain lesions and how they relate to the clinical symptoms. Lesion-symptom mapping (LSM) aims to identify strategic brain lesion locations that have the greatest impact on cognition. This allows a better understanding of the relationship between brain structure and function, and the cognitive impact of vascular lesions.

Several methods have been developed to assess LSM over the years; however, they have limitations. For instance, they do not consider the entire lesion image, nor do they identify multiple lesion types and cognitive domains simultaneously. Deep learning (DL) models overcome these limitations due to their complex architectures, which allow them to capture the correlation between multiple inputs and outputs simultaneously. However, one drawback of DL models is their lack of interpretability. It is not always clear how the model arrived at its predictions. To tackle this issue, explainable artificial intelligence (XAI) methods are used to understand the decision-making process of the model through attribution maps. These maps are images that highlight the areas of the input image that are important to the model.

The aim of this research is to develop a DL-LSM approach that predicts multiple cognitive outputs and identifies the lesion locations responsible for these predictions using XAI. The approach presented is validated in a simulation study that replicates brain-behaviour relationships using WMH segmentations of 821 memory clinic patients. This allows to evaluate the DL-LSM ability to locate predefined regions of interest (ROIs) that are associated to artificial cognitive scores. These simulations involved three experiments, including generating artificial cognitive scores based on the lesion volume within three ROIs, studying the impact of adding noise into the scores on the DL and XAI methods, and exploring whether intercorrelations between different ROIs in the artificial cognitive scores could be detected. For each experiment, convolutional neural networks (CNN) were developed to predict the artificial cognitive scores, and XAI was used to identify the locations that influenced these predictions. The methods were evaluated by quantifying the model's predictive performance, identifying the ROIs in the XAI's attribution maps, and quantifying the intercorrelation of the detected ROIs.

The study demonstrates that DL algorithms and XAI techniques can predict multiple cognitive scores and identify lesion regions related to these scores in 3D WMH lesion maps of SVD patients. Additionally, the study shows that DL models remain robust to artificial cognitive scores with up to 20% of added noise and can identify intercorrelations between different lesioned areas associated with artificial cognitive scores.

In conclusion, the findings suggest that a DL-LSM approach can help understand the underlying brain mechanisms that lead to the cognitive impact of vascular lesions. This approach also provides new insights into improving LSM techniques that are already used in clinical practice. Overall, it shows promise in enhancing the understanding and management of SVD-related cognitive impairment.