Interactive Design of Convolutional Neural Networks for Medical Image Analysis

Alexandre Xavier Falcão

Institute of Computing, University of Campinas, Brazil
`afalcao@ic.unicamp.br`

Convolutional neural networks (CNNs) have played a role in image analysis with several well-succeeded applications involving object detection, segmentation, and identification. The design of a CNN model traditionally relies on the pre-annotation of a large dataset, the choice of the model's architecture, and the tunning of the training hyperparameters. These models are sought as "black-boxes", implying that one cannot explain their decisions. Explainable artificial intelligence (XAI) has appeared to address the problem and avoid the wrong interpretation of the results. However, the importance of user and designer participation in the machine learning loop has called little attention yet.

In medical image computing, data annotation is costly, often scarce, and depends on an expert in the application domain (the user). The choice of the model's architecture and the training hyperparameter tunning rely on the network designer (an expert in AI). The user absence in the machine learning loop leaves essential questions with no answer (e.g., what are the most relevant samples for annotation?), while the lack of interactive methodologies to learn filters and model's architecture limits the designer to the interpretation of the model. The user and designer should then actively participate in the data annotation and training processes, both assisted by the machine, to increase human understanding and control, reduce human effort, and improve interpretation of the results.

This lecture addresses part of the above problems by presenting an interactive methodology for the design of CNN filters from markers in medical images, and a semi-automatic data annotation method guided by feature projection. The user starts the training process by selecting a few images per class and drawing strokes (markers) in regions that discriminate the classes. The designer defines an initial network architecture, and the filters of the CNN are automatically computed with no need for backpropagation. The user and designer may decide about the most suitable filters based on data visualization. The image features extracted by the CNN are projected in 2D for semi-automatic data annotation. The user analyzes the 2D projection, annotates the most challenging samples, while a semi-supervised classifier propagates the labels to the remaining ones. The annotated dataset can then be used to revisit the design of the CNN model, as illustrated for applications of medical image computing.