Interpretable Alzheimer's Disease Classification with Clinical Features

Project Management and Software Development
for Medical Applications

General Info

Contact Person: Tom Nuno Wolf

Contact Email: tom_nuno.wolf@tum.de

Project Abstract

Recently, we have proposed an inherently interpretable model for Alzheimer's Disease classification on 3D PET images. However, interpretation of 3D images is fundamentally challenging. In clinical practise, clinicians rely on a set of clinical features, which are typically 2D images. To this end, we want train, evaluate and interpret neural networks on such features in leverage applicability to the clinic. Additionally, we will evaluate the explanations given by an inherent interpretable neural network for agreement with the medical literature.

Background and Motivation

The group ai-med constantly push the application of AI in the medical domain. As a group within the Institute of Diagnostic and Interventional Radiology of the Klinikum Rechts der Isar, we work closely with physicians and practitioners to form the future of AI-assisted radiological workflows. Although many promising applications of AI to medical data can be found in the literature, the application and deployment of such methods in the clinical environment is a different story. On the one hand, "real-world" data is manifold compared to training data typically available, which is why AI often fails in the clinical setting. On the other hand, practitioners do not trust an automatic system, unless they fully understand its reasoning process. To this end, inherently interpretable neural networks allow to move from black-box models to human-understandable decision making.

Student's Tasks Description

Prior work on classification of Alzheimer's Disease with SSP maps is limited. As a result, the student needs to perform a benchmark with a neural network on the different projections and permutations of projections to take into account. Once a competitive performance with 3D models is achieved, the student will adapt an existing inherently interpretable neural network to the data. Lastly, the explanations given by this model will be evaluated in terms of quality (compared to the medical literature).

Technical Prerequisites

Python, PyTorch/Tensorflow, Pandas, matplotlib

References

T.N. Wolf, S. Pölsterl, C. Wachinger: DAFT: A Universal Module to Interweave Tabular Data and 3D Images in CNNs, 2022

T.N Wolf, S. Pölsterl, C. Wachinger: Don't PANIC: Prototypical Additive Neural Network for Interpretable Classification of Alzheimer's Disease, 2023

Please send the completed proposal to ardit.ramadani@tum.de, lennart.bastian@tum.de and tianyu.song@tum.de. Please note that this proposal will be evaluated by the BMC coordinators and will be assigned to a student only in case of acceptance.