Medical Image Computing

Mini-Project

Age regression from brain MRI

Predicting the age from a brain MRI scan is believed to have diagnostic value in the context of a number of pathologies that cause structural changes and damage to the brain. Assuming an accurate predictor of brain age can be trained based on a set of healthy subjects, the idea is then to compare the predicted age obtained on a new patient scan with the real age of that patient. Discrepancy between predicted and real age might indicate the presence of pathology and abnormal changes to the brain. The objective for the coursework is to implement two different supervised learning approaches for age regression from brain MRI data. Data from 652 subjects will be provided. Each approach will require a processing pipeline with different components that you will need to implement using methods that were discussed in the lectures and tutorials. There are dedicated sections in the Jupyter notebook for each approach which contain some detailed instructions and some hints and notes.

Part A: Volume-based regression using brain structure segmentation

The first approach aims to regress the age of a subject from the volumes of brain tissues, including grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF). It is known that with increasing age the ventricles enlarge (filled with CSF), while it is assumed that grey and white matter volume might decrease. However, as overall brain volume varies across individuals, taking the absolute volumes of tissues might not be predictive. Instead, relative volumes need to be computed as the ratios between each tissue volume and overall brain volume. To this end, a three-class brain tissue segmentation needs to be implemented and applied to the provided 652 brain scans. Brain masks are provided which have been generated with a state-of-the-art brain extraction tool from the FSL toolkit.

Different regression techniques should be explored, and it might be beneficial to investigate what the best set of features is for this task. Are all volume features equally useful, or is it even better to combine some of them and create new features. How does a simple linear regression perform compared to a hypothesis with higher order polynomials? How about other regression methods such as support vector regression or regression trees? The accuracy of different methods should be evaluated using cross-validation, and average prediction accuracy should be reported.

Part B: Image-based regression using grey matter maps

The second approach will make use of grey matter maps that have been already extracted from the set of MRI scans and aligned to a common reference space to obtain spatially normalised maps. For this, we have used an advanced, state-of-the-art neuroimaging toolkit, called SPM12. The reference space corresponds to the MNI atlas. Because the grey matter maps are spatially normalised, voxel locations across images from different subjects roughly correspond to the same anatomical locations. This means that each voxel location in the grey matter maps can be treated as an individual feature. Because those maps are quite large, there would be a very large number of features to deal with. A dimensionality reduction using PCA needs to be performed before training a suitable regressor on the low-dimensional feature representation obtained with PCA. It might also be beneficial to apply some pre-processing before running PCA, which should be explored. The implemented pipeline should be evaluated using cross-validation using the same data splits as in part A, so the two approaches can be directly compared.

DELIVERABLES

As part of the mini-project, you are asked to produce two deliverables:

1) Implementations of the tasks specified in the given Jupyter notebook. For the mini-project, you are given a Jupyter notebook MIC-Mini-Project.ipynb that contains descriptions and some skeleton code for the tasks outlined above. You will need to add your implementations at the specified locations directly within the notebook.

The data for the mini-project can be downloaded here.

2) A paper-style report of up to four pages about your method, experiments, and results
Prepare a report with four sections: abstract, methods, results, conclusion. The report should discuss the pipelines of your two approaches, and present the results ideally with some figures and graphs. We recommend using a LateX online editor (e.g., Overleaf, Authorea).

PLEASE WRITE YOUR OWN CODE!!! PLUGIARISM FROM YOUR FRIENDS AND INTERNET WILL BE PENALIZED. GROUP WORK IS NOT ALLOWED