Research Project Proposal

Feature Modeling for Anomaly Detection in Neuroimaging

Denis Kovačević
denis.kovacevic@student.uni-tuebingen.de

Advisors: **Thomas Wolfers, Alexander Frotscher**

University of Tübingen
MSc in Machine Learning

1 Objective

The objective of this research is to explore methods for anomaly detection using feature modeling within neuroimaging. The project aims to determine how pre-trained deep learning models can enhance these specific tasks in the neuroimaging domain. The research will evaluate the performance of different feature extraction and modeling techniques in identifying and localizing anomalies in MRI data.

2 Rationale

In neuroimaging, detecting anomalies like tumors or lesions is critical for diagnosing various brain conditions. Supervised approaches, like segmentation networks, require extensive labeled data, which is scarce in medical imaging and fails when labels can not be obtained.

Our approach focuses on feature modeling for anomaly detection in an unsupervised manner. By leveraging deep feature representations from pre-trained models, we can capture patterns of normal brain structure and detect deviations without requiring labeled anomalies. This allows us to effectively model the neuroimaging data and identify potential abnormalities, addressing the data limitations in this domain.

3 Approach

1. **Data Collection:** We will use neuroimaging datasets (e.g., MRI) suitable for anomaly detection, which have already been gathered and prepared for analysis.
2. **Self-Supervised Feature Learning:** We will explore self-supervised learning techniques to extract meaningful features from neuroimaging data. These methods allow us to learn robust representations from the data itself without requiring labeled anomalies.
3. **Feature Extraction:** We will implement a feature extraction pipeline designed to capture features that can effectively model the presence of anomalies. This involves extracting features from the pre-trained networks at specific layers and creating patches of them.
4. **Modeling Technique:** We will model the distribution of normal data using a Multivariate Gaussian distribution, which will allow us to detect outliers (anomalies) by identifying points that deviate from this distribution using the Mahalanobis distance as an anomaly score.
5. **Evaluation:** We will evaluate the performance of the feature extraction and modeling techniques using metrics such as Dice score for segmentation accuracy and AUPRC (Area Under Precision-Recall Curve) for detecting anomalies.

4 Possible Issues

One of the main challenges in this project is finding the right algorithms for pre-training models when adapting them from natural image datasets to medical images like MRI scans. This is particularly critical for fine-grained tasks such as anomaly localization. While the project aims to mitigate these issues by leveraging self-supervised techniques and transfer learning, optimizing these approaches specifically for neuroimaging remains a significant challenge.