MSC Information Studies
Data Systems Project 2021-22

Project Title

Explainability in medical image analysis

Project Stakeholder (name and or organization)

Quin

Stakeholder contact details

Majd Zreik
Lead computer vision
m.zreik@quin.md

Brief project description

AI is still seen as a magic black box, it is non-intuitive, and difficult for people to understand and trust. In the medical image analysis field, this problem is amplified and leads to a limited adoption of such AI systems. The proposed projects aim to adapt the exciting explainability tools for the medical domain to demystify and better explain medical AI decisions.
Topics to be covered:

- Implement XAI: In literature, there are numerous methods to explain the predictions of an AI system. However, there is no central platform that can run several implementations simultaneously and intuitively on the same input image to provide the user an easy way to examine the different results. In this project, we aim to build an intuitive graphical interface with a plug-and-play nature for implementations of several explainable methods running on an input image.
- Evaluate XAI: In literature, there are numerous methods to explain the predictions of an AI system. However, evaluating the performance of such methods is hard. This project aims to research for the optimal evaluation metrics and to build a graphical interface to visualize such metrics for a given input image. Moreover, we aim to define and build a feedback interface for the user to interact and evaluate the explainable outcome on top of the input image.
- Visualize XAI: In literature, the common way to visualize the explainable result is by overlaying a colorful heatmap on top of the input image. For medical image analysis this is not optimal as it might interfere with the specialist reading the input image. This project aims to research and define new ways of visualizing such results for the medical image analysis field.

Quin will provide a link to a public dataset of X-ray images of knees suspected with osteoarthritis (OA). The students would use the dataset to train a model and use the trained model to perform the different projects.

Key challenge/problem or message to be communicated within the project domain

MSC Information Studies
Data Systems Project 2021-22

See project description

Exemplar projects or additional reference materials

Dataset: https://data.mendeley.com/datasets/56rmx5bjcr/1
Similar paper: https://pubs.rsna.org/doi/pdf/10.1148/ryai.2020190065
Similar repo: https://github.com/MIPT-Oulu/DeepKnee
Review papers about XAI in medical image analysis:
https://arxiv.org/abs/2005.13799 and https://arxiv.org/abs/2004.14545

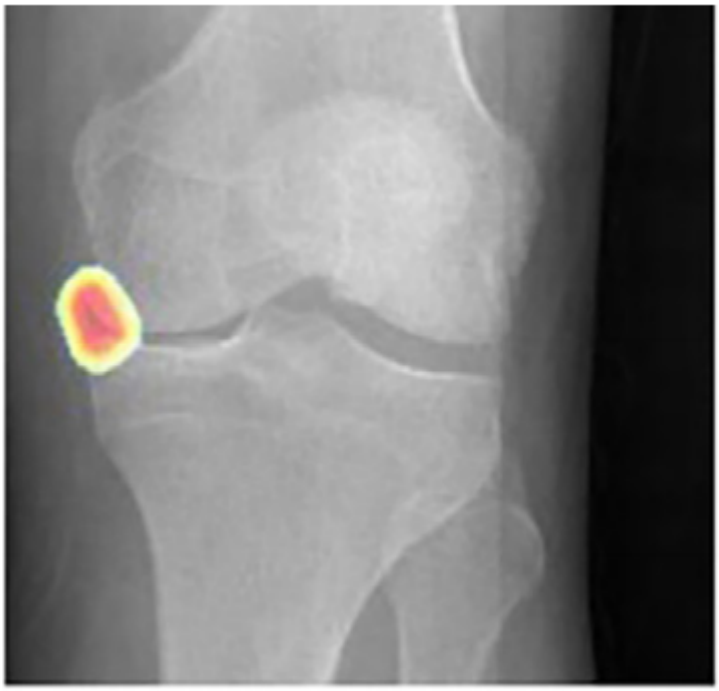

Figure 1: from https://pubs.rsna.org/doi/pdf/10.1148/ryai.2020190065

Goals or key criteria for an explorative project
Find the optimal way to implement, evaluate and visualize XAI for a medical image analysis platform that take a human-in-the-loop into account.

Suggested key requirements or success factors for an implementation

- Implement XAI:
  - a plug-and-play functional graphical interface for running several explainable methods on an input image.
- Evaluate XAI:
  - optimal evaluation metrics (that take the clinicians opinion into account)
  - a functional graphical interface to visualize such metrics for a given input image
  - a feedback interface for the clinician to interact and evaluate the explainable outcome on top of the input image.
- Visualize XAI:

- Interviews with clinicians
- define optimal ways of visualizing examinable results for the medical image analysis field.
- Visualize these results on a given input image

Challenges or constraints envisioned (if any)
Data is from a public dataset. We don't see any issues here.

Any specific technical or content requirements
The dataset is quite large (7GB). Only the *ClsKLData* (data for classification) should be used. For simplicity, instead of dealing with a 5 class classification problem (kl-0,1,2,3,4), students might binarize it (kl 0, 1 and kl 2,3,4) or even choose only two classes (kl-0 and kl-3).
If students want to avoid training the model, several implementations are available online that include the trained models on the same dataset

Any additional comments or thoughts

We would like to have meta-groups of 15 students, consisting of 3 groups working on the three different subprojects decomposing the main problem.