HPI Digital Health Center:

Research Group: Machine Learning, headed by Prof. Dr. Christoph Lippert

Medical Imaging Model Zoo

Motivation

Deep learning models have been demonstrated as a useful method to detect abnormal characteristics in medical images and could therefore support the work of medical practitioners. Hospitals have accumulated large databases of medical images from routine radiologic diagnostics, such as the Imaging Research Warehouse (IRW) at Mount Sinai. With millions of images, IRW represents one of the world's largest databases, covering a broad range of body parts, medical indications and image modalities (CT, MRI, etc.) that promises to become a unique resource to foster the uptake of deep learning in clinical imaging. Turning this data into a public resource for open science is a monumental undertaking.

In this project, you can contribute to this effort by developing a model zoo that allows to store and retrieve versioned deep-learning models that have been trained on medical images for a variety of tasks. We will start training deep-learning models that help organizing the images, i.e. by different body parts or views and detect objects such as medical devices. We will then create a model store to keep track of these models and add new ones. A basic scenario to make use of the model store is displayed below.
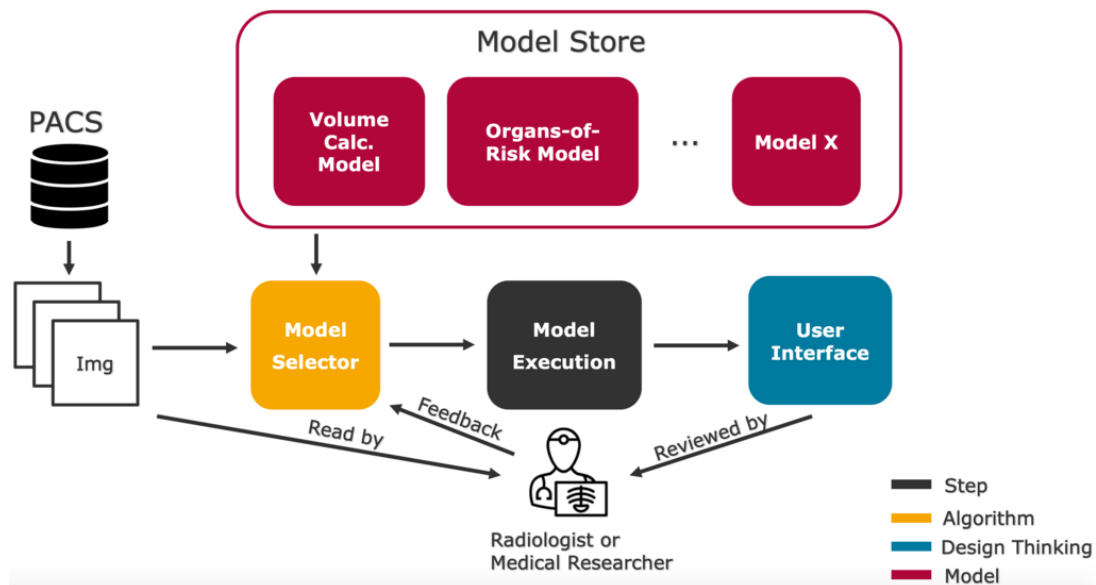
Medical images are stored in and can be accessed from a Picture Archiving and Communication System (PACS). Extracted images are input to a model selector. The model selector can be seen as the brain of the system, which similar to a radiologist, decides based on the image

content or meta data which models apply. The whole available knowledge is located in a model store. After a model has been selected, it is executed, and the output is displayed to a graphical user interface that the corresponding medical expert can review.

The outcome of this project will be a model-store that lists trained models and their evaluation and where new models can be uploaded. This store could potentially be accessed by a model selector that seeks to find a suitable model for a certain image-set and task.

Work Packages

The work packages comprise the following steps:

1. Familiarizing with different public data sets (CheXpert, MURA, …) through exploratory data analysis

2. Building different classification and segmentation models with these datasets that aim at identifying objects or characterizing or segmenting image

3. Evaluate model performance with different metrics (i.e. AUC, confusion matrix)

4. Building the backend and frontend of model store.

5. Develop a user interface to apply models on input data

Requirements

Since the development of the model store is the main part of this project, the team should have a strong software engineering background including knowledge of modern web frameworks. Team members interested in the development of deep learning models for various datasets should have attended at least one machine learning course and be able to extract and preprocess datasets. The team should be familiar with Python and its ecosystem (e.g. jupyter, numpy, pandas, deep learning frameworks).

Contact

If you have any open questions, do not hesitate to get in touch. We are located on the 1st floor of the Digital Health Center.

M.Sc. Benjamin Bergner
E: benjamin.bergner@hpi.de

M.Sc. Jana Fehr
E: jana.fehr@hpi.de

Prof. Dr. Christoph Lippert
E: christoph.lippert@hpi.de