Installation Guide

System Installation

Requirements

The system requires the following conditions to be available on the machine:

- Internet connection
- Python 3 with `pip` package manager
- `conda` package manager

The `conda` command can be installed with either Anaconda or miniconda.

For Windows

The installation requires bash scripts being able to executed. Command Prompt or Powershell are not able to run bash script so we recommend using other solutions for Windows. We recommend:

- Git bash
- Windows Subsystem for Linux

Installation

Assume we are at the project's top-level folder, run the following command

```
$ ./SALIENCY_DISAGREEMENT/SETUP/setup_env.sh
```

The script `setup_env.sh` sets up the following:

- A new conda environment name `xai_disagreement`
- Install all required packages for training, generating explanations and visualization
- Download training data

After installation we should see a new directory `SALIENCY_DISAGREEMENT/SOURCE/pneumothorax-chest-xray-dataset/siim-acr-pneumothorax` that contains the training and test data.

Usage

Make sure to use the system with the `xai_disagreement` environment activated, using:

```
$ conda activate xai_disagreement
```

The system is run in three steps:

1. Training
2. Generating explanations
3. Visualization

Step 1: Training

To train black boxes, run:

```
$ ./SALIENCY_DISAGREEMENT/SOURCE/train.sh
```

After training the output black box checkpoints will be stored in the directory `pretrained_weights`. These checkpoints are required for step 2.

Step 2: Generating explanations

Launch Jupyter notebook using:

```
$ jupyter notebook
```

Run all cells within the notebook `generate_explanation.ipynb` to compute the explanations for the black boxes on the test set data. Computed explanations are stored in the directory `explanations` and are required for visualization in step 3.

Step 3: Visualization

Run all cells within the notebook `visualization.ipynb` to compute the disagreement heatmaps. The figures of the heatmaps are stored in `figures`. These figures are the results of our thesis.