

CSC/MAT-338 Final Project: MedMNIST3D

Profs. McKinney and Yoon

Project Overview

For this project, you and your partner will work with one of the six MedMNIST3D datasets. Some are classification oriented and some are detection oriented (e.g. is an anomaly present). The MedMNIST3D datasets comprise collections of 28 by 28 by 28 images gathered by medical imaging techniques. Your goal is to design a TensorFlow neural network that works with your dataset. You will need to understand what the data is and what the task is. You will also need to consider ethical aspects of the project.

Starter Notebook

We provide an initial Jupyter notebook that shows a few ways to visualize the data, using matplotlib. There are other ways to potentially visualize the data. The first cell is what is important: it contains the starter code to properly import the dataset and convert it to a format appropriate for TensorFlow.

Project Milestones

For this project, you will have several milestones in addition to the final product.

Milestone 1: Due to Canvas 2359 Friday 5 December. 1 point.

This milestone's goal is for you to have downloaded your assigned dataset and started to understand what your neural network needs to do. To do this, you will need to read the MedMNIST website and its associated papers (<https://medmnist.com/>). What is your data? What type of medical imaging produced it, and what is the medical question at issue with your dataset? **Deliverable: Jupyter notebook with initial code and markdown cells containing descriptions of your data.**

Milestone 2: Due to Canvas 2359 Tuesday 9 December. 1 point.

At this point you should also have an initial neural network architecture with appropriate loss function and should have run it on at least one or two epochs to confirm that the architecture is starting to learn. We assume that these epochs are just test epochs and will run on your local hardware rather than our shared GPUs.

Deliverable: Jupyter notebook.

Milestone 3: Due to Canvas 2359 Friday 12 December. 1 point.

For this milestone, you should have a complete neural network architecture that has run on the shared GPU system, giving you results from the validation dataset to compare against the MedMNIST benchmarks. Details about how to use the GPU system will be posted to Canvas over break.

Deliverable: Jupyter Notebook.

Milestone 4: Due to Canvas 0800 Wednesday 17 December.

You will upload two things to Canvas: your presentation slides and the final Jupyter notebook (even if your notebook has not changed since the last milestone. These will be two separate assignments on Canvas to keep the files organized). We will load them onto one computer to minimize transition times for the presentations. **Failure to upload will result of a grade of 0/12 for the presentation.**

Deliverable: Presentation (PPT or PDF) and Jupyter notebook.

Final Submissions: Due to Canvas 1200 Wednesday 17 December. 2 points.

You will submit a peer evaluation sheet (on paper), where you will give feedback to each presenting group. You will fill this out as the presentations are delivered and turn the paper in when all presentations are complete. Failure to take this seriously will result in loss of points.

You will complete a partner evaluation form on Canvas after we finish the presentations. You will each indicate how you and your partner divided labor and collaborated. Failure to take this seriously will result in loss of points.

Presentation Requirements

Each presentation should be between 7 and 9 minutes. During this, you should:

- Briefly describe the dataset and task. What medical question is trying to be answered?
- Briefly describe your architecture and how it compared to the MedMNIST benchmarks. If you used network elements that we haven't discussed in class (e.g. LSTM), give a very brief description of them.
- Briefly describe your view of the ethical considerations of **one** aspect of the project. For example, you could consider the ethics of how the data was collected, or how your neural network could be used. In doing so, you should connect your chosen ethical aspect to a **specific part** of one of our ethics readings. (We realize that there are many ethical considerations worth discussing. Focus on one for the sake of time.)

Grading:

Each milestone has a point value associated with it as listed above. The presentation will be graded out of 12 points: 4 points for dataset/task, 4 points for architecture, and 4 points for ethics. These will correspond to letter grades, e.g. 4 = A, 3 = B. The final work product will also be graded out of 4 points using the same method.