

Description of Source Data

Anton Zhitomirsky

December 29, 2023

1 Structure of source Data

Results are structured in the file:

/vol/biomedic3/bglocker/nnUNet

-rwxr-xr-x	1	bglocker	biomedia	236	Sep	24	15:16	exports
drwxr-sr-x	9	bglocker	biomedia	9	Nov	25	10:55	nnUNet_preprocessed
drwxr-sr-x	9	bglocker	biomedia	10	Nov	25	10:50	nnUNet_raw
drwxr-sr-x	9	bglocker	biomedia	9	Nov	25	12:20	nnUNet_results
drwxr-sr-x	11	bglocker	biomedia	11	Dec	16	09:10	nnUNet_testing
-rw-r--r--	1	bglocker	biomedia	644	Oct	20	07:20	run_nnunet_0.sh
-rw-r--r--	1	bglocker	biomedia	644	Oct	20	07:20	run_nnunet_1.sh
-rw-r--r--	1	bglocker	biomedia	644	Oct	20	07:20	run_nnunet_2.sh
-rw-r--r--	1	bglocker	biomedia	644	Oct	20	07:21	run_nnunet_3.sh
-rw-r--r--	1	bglocker	biomedia	644	Oct	20	07:21	run_nnunet_4.sh

2 nnUNet_raw

nnUNet_raw has the original (training) images with manual annotations. Each Dataset below is treated as a binary segmentation problem. See Section5

drwxr-sr-x	4	bglocker	biomedia	5	Sep	17	13:47	Dataset001_Anorectum
drwxr-sr-x	3	bglocker	biomedia	5	Sep	17	20:24	Dataset002_Bladder
drwxr-sr-x	3	bglocker	biomedia	5	Sep	17	20:27	Dataset003_CTVn
drwxr-sr-x	3	bglocker	biomedia	5	Sep	17	20:28	Dataset004_CTVp
drwxr-sr-x	3	bglocker	biomedia	5	Sep	17	20:29	Dataset005_Parametrium
-rw-r--r--	1	bglocker	biomedia	135	Nov	25	10:50	note

3 nnUNet_results

The raw files from Section2 are used to train an nnUNet model. Which does a 5-fold cross validation, resulting in five models, each trained on 80 subjects and tested on 20 (there is a total of 100 subjects with manual annotations).

3.1 Script

Slurm script

Each of the scripts are presented as bash scripts in `/vol/biomedic3/bglocker/nnUNet/run-nnunet_*.sh`. These are used to schedule the python program into the Slurm scheduler for running in the cloud.

```
#!/bin/bash

# Example of running python script in a batch mode
#SBATCH -c 4                # Number of CPU Cores
#SBATCH -p gpus             # Partition (queue)
#SBATCH --gres gpu:1        # gpu:n, where n = number of GPUs
#SBATCH --mem 12288         # memory pool for all cores
#SBATCH --nodelist lory     # SLURM node
#SBATCH --output=slurm.%N.%j.log # Standard output and error log

# Source virtual environment (pip)
source /vol/biomedic3/bglocker/coding/envs/nnunet/bin/activate

# Set env variables
source /vol/biomedic3/bglocker/nnUNet/exports

# Run python script
nnUNetv2_train 1 3d_fullres 4 # where 4 refers to which dataset we're
                             # training
```

- Source virtual environment
This is activating the already created python environment.
- Set env variables
contains paths for raw, results and preprocessed directories
- Run python scripts
runs the script at `/vol/biomedic3/bglocker/coding/envs/nnunet`

Python script

No access

4 nnUNet_testing

The models are then tested against 10 hold out manual segmentations with no manual segmentations.

5 Viewing the Data

The viewing tool used is ItkSnap, which was developed as an open source tool for viewing medical imaging scans. The view (Figure1) shows how you would see input data.

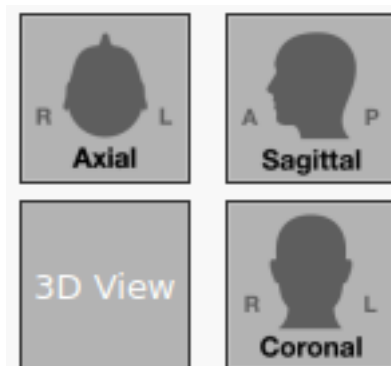


Figure 1: view of all input data

With that we can use this tool to view input data. Here, the R and L stand for right and left respectively, and the A and P stand for Anterior and Posterior. We can provide a few other examples of viewing data displayed below in Figure2. We are further provided with manual annotation of the substructures. Figure3 shows an example of the annotation of the Anorectum. You can enable the 3D visual model through Edit > 3D Panel > Toggle 3D view.

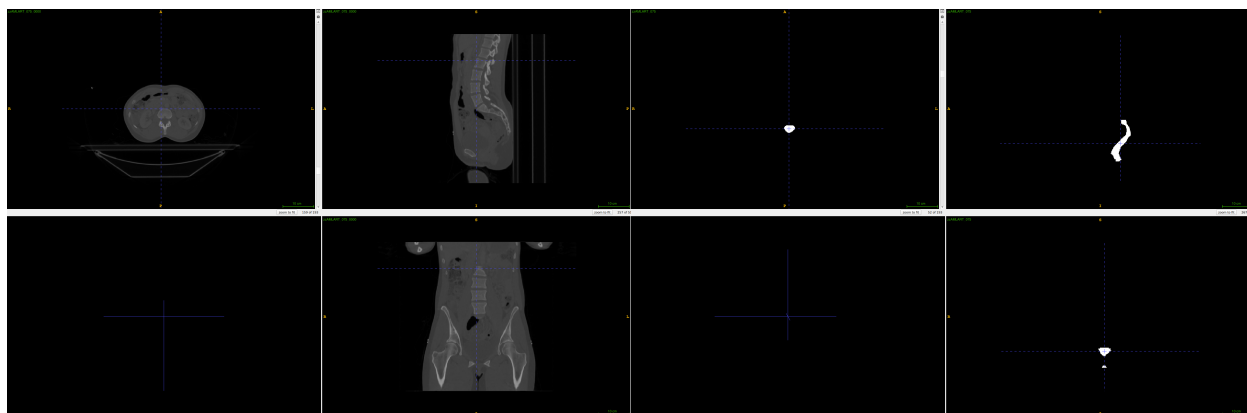


Figure 2: ItkSnap view of the Anorectum Raw Image

Figure 3: ItkSnap view of the Anorectum Raw Image