CENTER FOR BIOMEDICAL IMAGING IN ONCOLOGY
Lurie Family Imaging Center

Dana-Farber
Cancer Institute
HARVARD MEDICAL SCHOOL
TEACHING HOSPITAL

DATA PREPARATION: Instruction for machine learning data preparation

STEP 1
- Run ClinicalVolumes tumor segmentation. Save segmentation of the tumor into a "seg" folder. one "seg" folder for each animal

STEP 2
- **RGB image creation** (Python)
  1) Open "Dicom_to_tiff/1_RGBcreator/Python" folder

  2) Inside the "data_dir" folder, create a new directory (Note: create one directory for each animal)

     Name each directory similarly to the following: LFIC_MR_0001_20200101

     The name must have the same number of letters and underscore as the suggested. The name

     will be used to name the RGB images

  3) Move or copy the dicom files into the corresponding folder inside the "data_dir" directory

  4) Open "RGB_creator_multidir" in a Python editor

     Select image options as requested

  5) Run "RGB_creator.py" from the terminal with the command:

     python RGB_creator_multidir.py

 - The program will convert all the dicom files inserted in the different folders in .tif 256x256x3

STEP 3
- **8bits mask image creation** (MATLAB)
  1) Copy   a   single   "seg"   folder   (folder   with   the   segmentation   dicom   files)   into

     "Dicom_to_tiff/1_RGBcreator/Matlab" folder

  2) Open Gray_scale_and_8uint_image.m file in MATLAB and follow the instruction inside the file.

     (no need to clean the folders from old data)

  3) Segmented masks will be created and found in the "Images_GrayS_8bits" folder

- The program must be run for one animal (one dicom dataset) at the time

Note: If needed to **rename**, **invert the order** or **rotate** the .tif image files:
1. open the "Dicom_to_tiff/ 3_IMAGE_editing" folder

2. move the tif images into the corresponding folder (The name of the folder will be used to name or rename the images)

3. Open "IMAGE_editing_multidir.py" in a Python editor

   Select image options as requested

4. Run "IMAGE_editing_multidir.py" from the terminal with the command: python IMAGE_editing_multidir.py

STEP 4
- Move the tif image files and mask files in a folder XX. One folder per animal

Instruction for machine learning data preparation (Skull_stripping)

STEP 1

- Run ClinicalVolumes tumor segmentation. Save segmentation of the tumor into a "seg" folder.

- Run ClinicalVolumes full brain segmentation. Save segmentation of the brain into a "seg_skull_stripping" folder.

STEP 2

- **RGB image creation _ Skull_stripping_mask creation** (MATLAB)

  1) Move the "seg_skull_stripping" folder into "Dicom_to_tiff/2_RGBcreator_Skull_stripping/Matlab" folder

  2) Open Skull_stripping_mat file in MATLAB and follow the instruction reported in the file. (no need to clean the folders from old data)

  3) Segmented brains .tif files will be found in the "Skull_stripping_mask" folder

- The program must be run for one animal (one dicom dataset) at the time

- **RGB image creation _ Skull_stripping_mask creation** (Python)

  4) Open "Dicom_to_tiff/2_RGBcreator_Skull_stripping /Python" folder

  5) Inside the "data_dir" folder create a new directory (Note: create one directory for each animal)

     Name each directory similarly to the following: LFIC_MR_0001_20200101

     The name must have the same number of letters and underscore as the suggested. The name will be used to name the RGB images

  6) Inside the "Skull_stripping_mask" folder create a directory of the same name as in (5)

  7) Move or copy the dicom files into the corresponding "data_dir" folder

  8) Move or copy the .tif skull stripping masks files (created in STEP 2) into the corresponding "Skull_stripping_mask" folder

9) Open "RGB_creator_multidir_stripping.py" in a Python editor

   Select image options as requested

10) Run "RGB_creator_multidir_stripping.py" from the terminal with the command:

   python RGB_creator_multidir_stripping.py

- The program will convert all the dicom files inserted in the different folders in .tif 256x256x3

STEP 3
- **8bits mask image creation** (MATLAB)
   4) Copy the "seg" folder into "Dicom_to_tiff/1_RGBcreator/Matlab" folder

   5) Open Gray_scale_and_8uint_image.m file in MATLAB and follow the instruction inside the file.

   (no need to clean the folders from old data)

   6) Images will be created in the "Images_GrayS_8bits" folder

- The program must be run for one animal (one dicom dataset) at the time

Note: If needed to **rename**, **invert the order** or **rotate** the tif image files:
   5. open the "Dicom_to_tiff/ 3_IMAGE_editing" folder

   6. move the tif images into the corresponding folder (The name of the folder will be used to name

   or rename the images)

   7. Open "IMAGE_editing_multidir.py" in a Python editor

   Insert the directory paths as requested

   Select image options as requested

   8. Run "IMAGE_editing_multidir.py" from the terminal with the command: python

   IMAGE_editing_multidir.py

STEP 4
- Move the .tif image files and .tif mask files in a folder XX. One folder per animal