# README for patch generator from ILD DB

Sylvain Kritter May 2nd 2016

# Tool used

1. RadianDICOM viewer to view dcm files.
2. Can also convert dicom images in jpg, bmp,
3. Total Image Converter to create jpeg, bpm and JPEG 2000 from dcm.

I have created in each ILD\_DB-textROIs dataset (35, 65,…) the directories

bmp,jp2k,jpg to store equivalent format from dcm files.

1. Spyder Python 3.5 for Python

# To run the patch generator

1. 3 python files needed:
2. final.py (this is the top, parameters are defined in it, no need to touch the 2 others)
3. fillshape.py
4. generatetabc.py
5. Preparation 2 python file needed
6. genima.py
7. renomsca.py

To run the patch generator, we need first to have the directory ILD\_DB-textROIs in the place where Python is launched.

In each dataset in directory ILD\_DB-textROIs, we need to add a directory named bmp or jpg to store bmp or jpg files generated from dcm

Bmp or jpg file are generated by TotalImageconverter, at top level keeping hierarchy.

Then genima.py is run to create the bmp or jpg directories and to put all the image files in bmp or jpeg directory.

For some scan, the scan are not following the number convention: scan number in the format <name>-<number in 4 digit>.dcm.

Renomscan.py is adding this number at the end of the scan file.

1. Customization in final.py
2. Patch image format: bmp or jpg
3. Dicom file size: 512 \* 512
4. Patch size: 32\*32
5. Threshold in % : 0.8 for patch area over ROI

# program generate files

1. From where it is launched:
2. ‘Jpeg’ directory , where jpeg images and text files of all ROI with patches is stored
3. ‘patch’ directory, where patches are store in sub directories, named after label and localization names in each scan, according to image format declared in top python (jpg or bmp). I think bmp is more accurate.
4. in each ILD\_DB-textROIs dataset (35, 65,…) , a directory named ‘patchfile’ where intermediate data is stored. Can be deleted afterward.

# Database Analysis

1. The Dcm files are identical in ILD\_DB\_txtROIs, ILD\_DB\_volumeROIs and ILD\_DB\_lungMasks
2. A ROI text file is in each data set In ILD\_DB\_txtROIs This Text file: corresponds to ROI in ILD\_DB\_volumeROIs roi\_mask of each dataset.
3. There is a set of ROI in roi\_mask in each dataset of ILD\_DB\_lungMasks: this is lung area