**Overall review of files and folders**

Files

* The notebook “Create Dataset.ipynb” - responsible for creating datasets for the networks.
* The following notebooks are responsible for training networks:

1. “Train 2D Classifier model.ipynb” - for 2D classification
2. “Train 3D Classifier model.ipynb” - for 3D classification
3. “Train Segmentation model” - for 2D segmentation

* The notebook “Feature Extraction.ipynb” - responsible for extracting features
* The notebook “Clustering and Exporting Data.ipynb” - responsible for clustering
* The notebook “Patch Visualisation.ipynb” - responsible for marking patches according to clusters

Folders

* The folder “datasets” contains both the mice data (images and masks) and the datasets for the networks (patches divided to train and test).
* The folder “results” is divided into 4 subfolders:

1. The folder “models” - contains models of trained networks
2. The folder “features” - contains features extracted from those models
3. The folder “clustering” - contains clustering’s results
4. The folder “patches\_mark” - contains images with marked patches according to the clustering

**A brief guide of the complete process**

1. Add new data scan(s) – create a new folder with the name of the scan, and in it create two folders "images" and "masks" – the first containing the slices and the second containing the masks. You would probably want to move this folder into the "miceData" folder in the "datasets" folder – since we saved the existing mice data in the same folder.

- "miceData" contains images in DICOM format. we have another folder "newMiceData" that contains PNG images, but it will be easier for you to work with the first one, without needing to convert by yourself.

1. Create a dataset for your network: open "Create Dataset.ipynb". You can find explanations about the required arguments and examples of 2D and 3D datasets creations. The dataset will be saved in a folder for you to name and after that, using "rearrange" – it will be copied and split into subfolders of "train" and "test".
2. Train a network - run the notebook that matches the network you want to train. The network will be trained and the best models will be saved.
3. Extract features - run “Feature Extraction.ipynb” to extract features from a saved model. Note that you need to load the model saved in the previous notebook.
4. Divide the test dataset into clusters based on the features: use “Clustering and Exporting Data.ipynb” for clustering and for saving full results. This notebook contains three parameters you can set (see “parameters” comment): Task (“classification\_2d”, “classification\_3d” and “segmentation”), epsilon and min\_samples. See paper for details.
5. Visualization: open "Patch Visualisation.ipynb". You can find explanations about the required arguments and how to mark the patches automatically (and manually if needed for some reason). By loading the file “clusters.csv” which was saved in the previous step, and using the main function, the slices (with marked patches) will be saved in a PNG format. The existing functions allow marking using 6 colors: "red" for cluster 0, “blue” for cluster 1, “yellow” for cluster 2, “pink” for cluster 3, “green” for cluster 4 and “white” for outliers (-1).