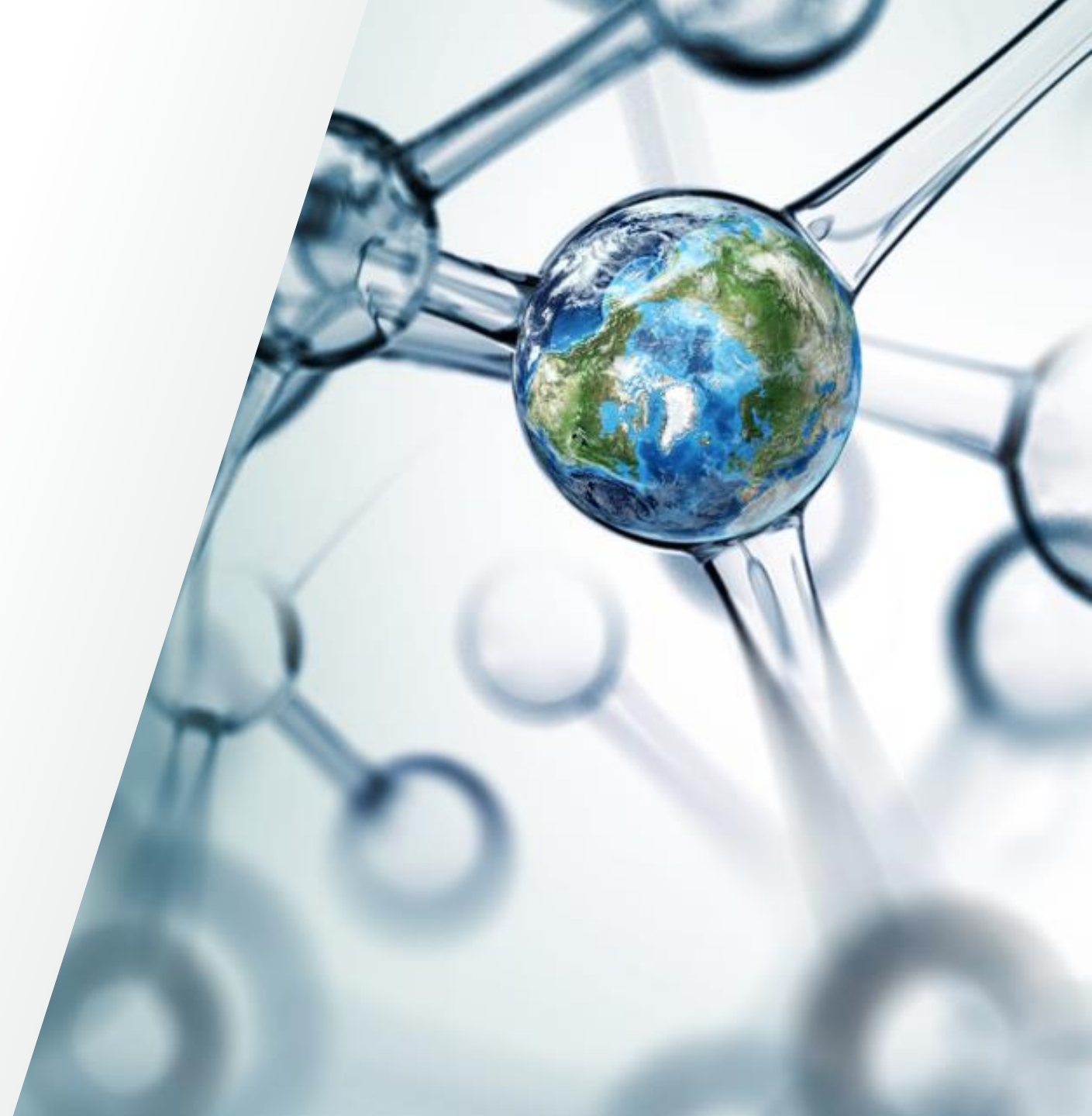


Extraction of Patches for Deep Learning

February 2021

 The world leader in serving science

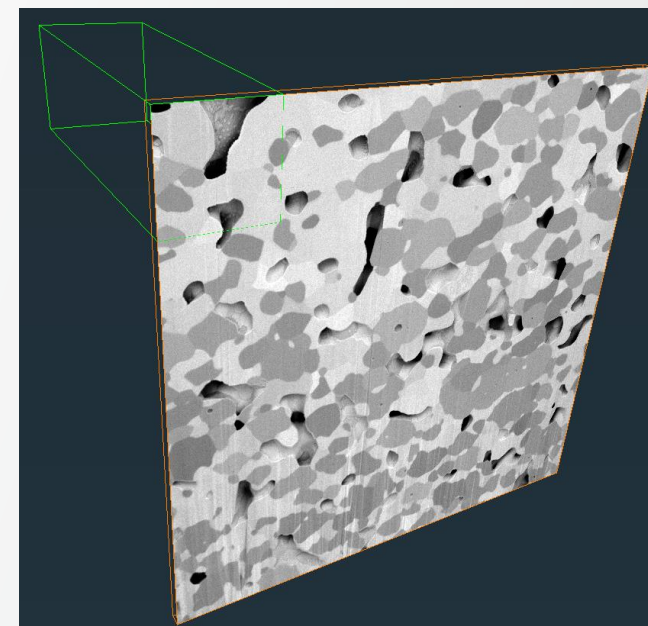
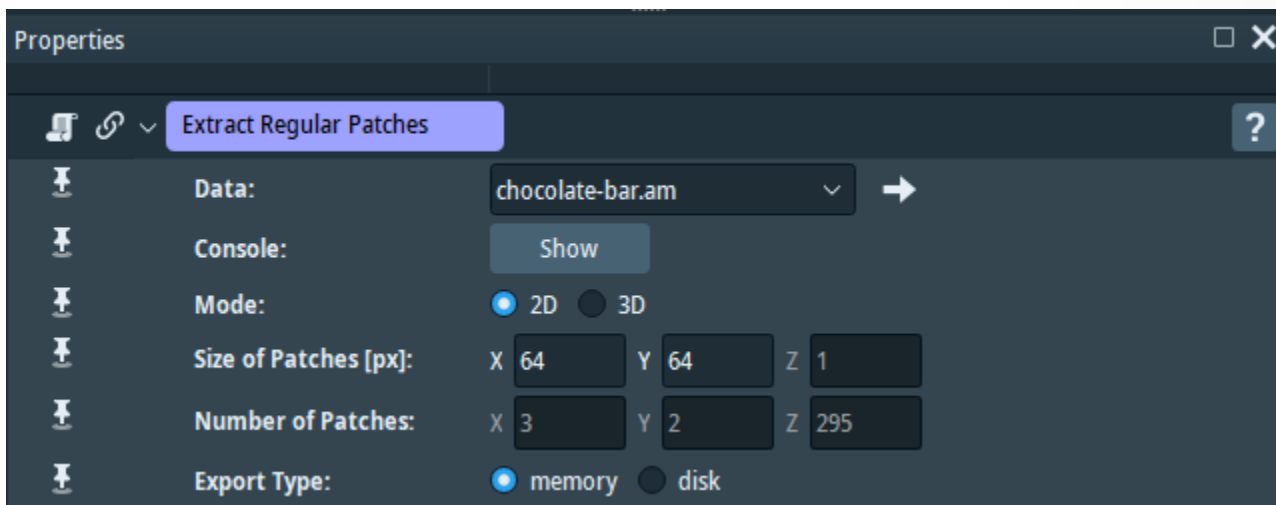


Why extracting patches?

- Properties of the Deep Learning Training module
 - Training and validation data are split from input volumes
 - Training batches are generated from randomly extracted patches
- Potential limitations
 - You need more control on where the patches come from, to balance the training and validation data
 - You have some classes which are under-represented, and want to emphasize more on these during the training
 - You need to manually segment your ground truth, but cannot afford to process the full data, and it is not appropriate to work on a subvolume
- Features proposed
 - All modules of this package allows for a reproducible extraction
 - They can, and should, be applied to both the images and the associated segmentation to ensure the consistency
 - All modules will only extract patches of the requested size. Patches location that would be too close to the image border are skipped.
 - Extraction modules allow extracting in memory (as a data module), or as 2D TIF files on the disk.
 - *Note for PerGeos users: all proposed modules can be used from the Explore workroom. However, some convenience features or manipulations are only available with Amira or Avizo.*

Extracting regular, adjacent patches

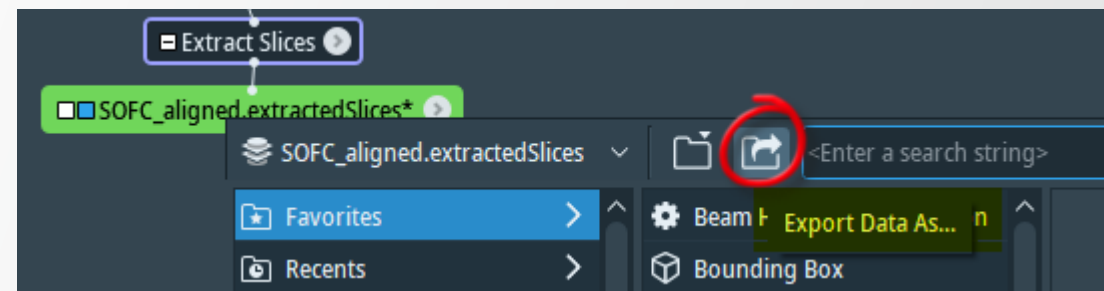
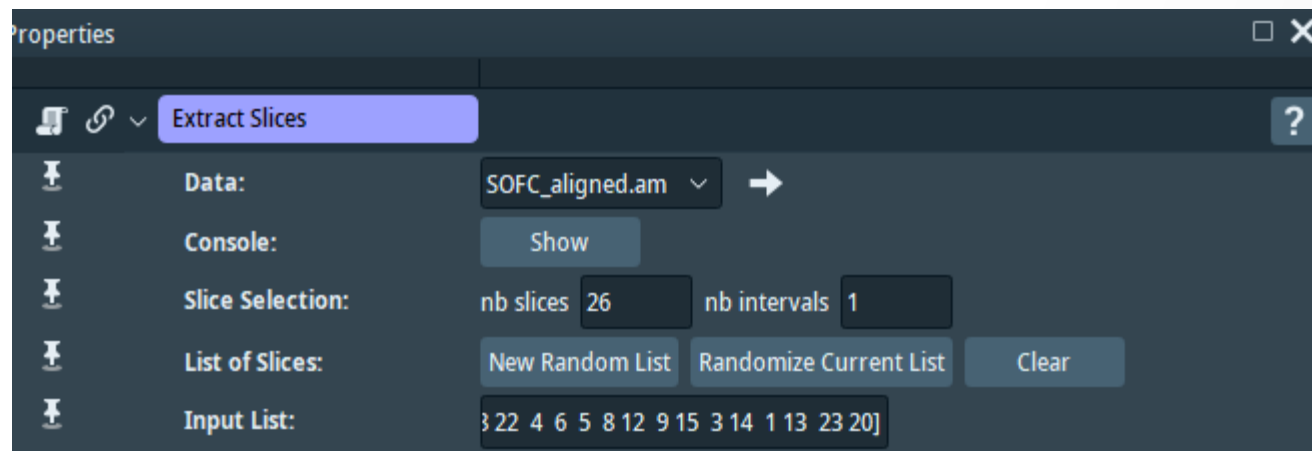
- The module « **Extract Regular Patches** » splits the data into adjacent patches
 - Each patches will then be considered as a slice
 - The patches are extracted from left to right, top to bottom, for all slices.
 - Only complete patches are extracted, so pixels on the right and bottom part of the original data will likely not be extracted.
- If you have already segmented your data, you can use the same module first on grayscale data, and then on the ground truth segmentation dataset to ensure consistency in your training data.
- You may want to randomize the patches afterwards, using « Extract Slices » module (see next slide)



(orange) Original data, and (green) data reformatted with *Extract Regular Patches*

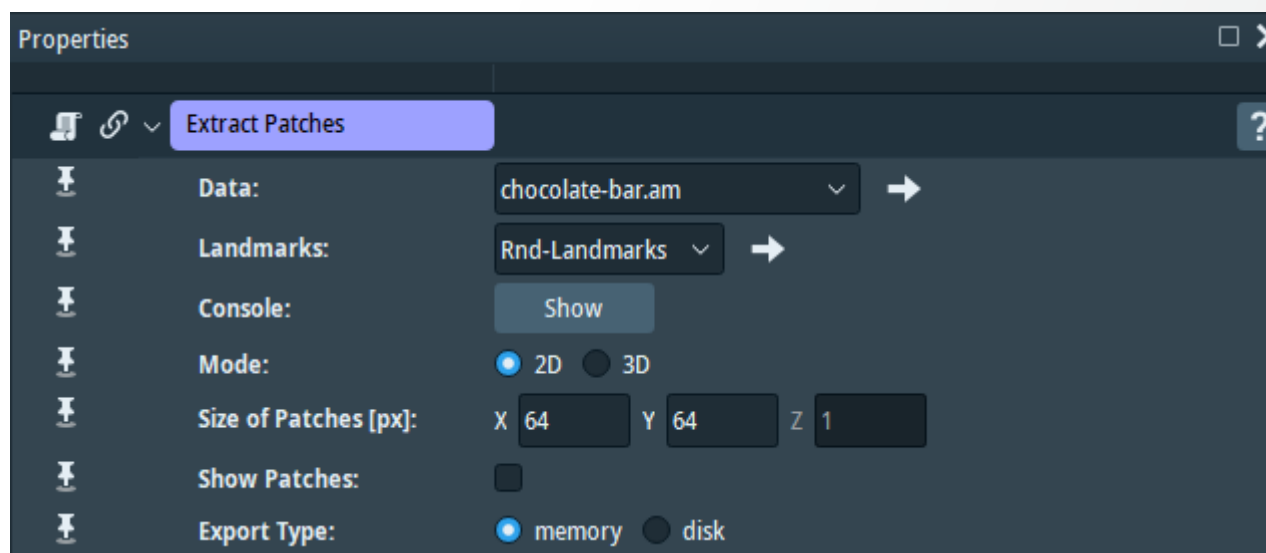
Extracting random slices

- The module « **Extract Slices** » simply extracts slices from the input volume
- The '**Input List**' vector indicates the slice indices to be extracted
 - It can be manually defined, or edited.
- **New Random List** allows randomly selecting 'nb slices' to be extracted.
 - These slices are determined by a random shuffling of all slice indices.
- Since the slice indices to be extracted are fixed, you can use the same module first on grayscale data, and then on the ground truth segmentation dataset to ensure consistency in your training data.
- If you need to export the randomized slices to the disk: right-click on the output dataset, and export as 2D TIFF.



Extracting patches at specific locations

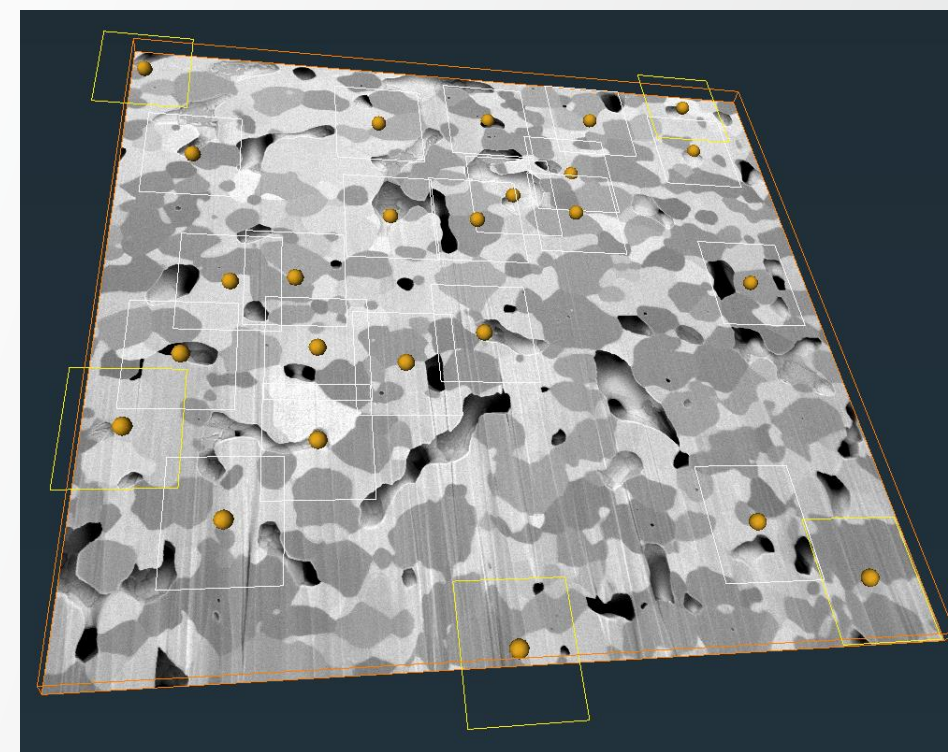
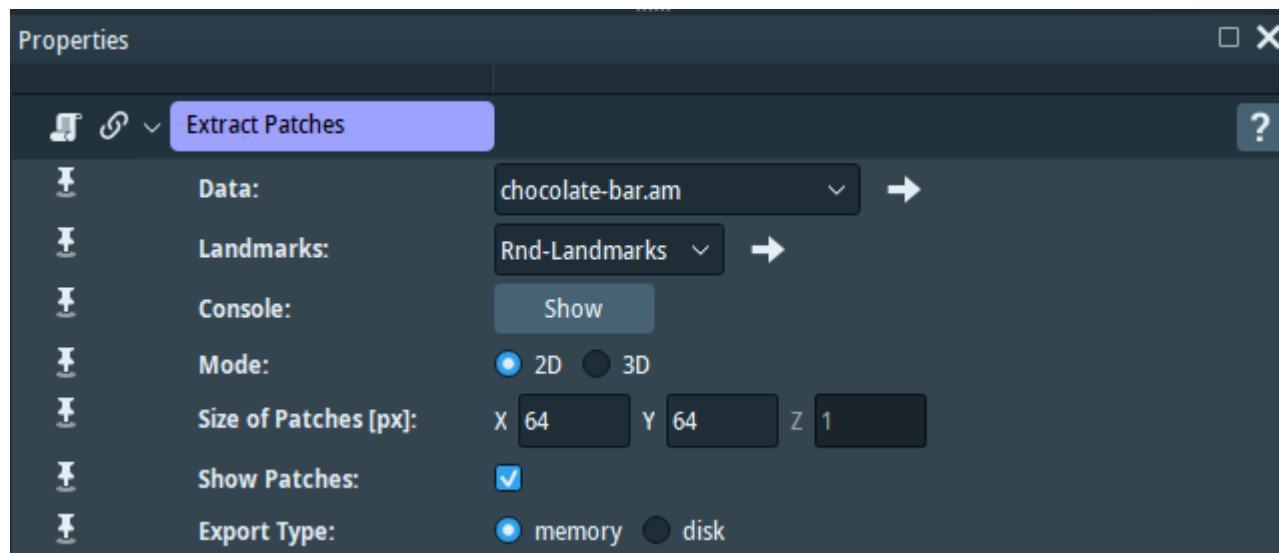
- The module « **Extract Patches** » extracts patches at locations defined by a 'Landmark' dataset
- The patches are extracted in an order corresponding to the indices of the Landmarks in the set.
- Only entire patches are extracted: if a landmark is too close to the image border, it will be skipped.
- See 'Random Landmarks' for randomly shuffling the landmark indices.



- **Note for PerGeos users:** Landmarks cannot be visualized, and cannot be manually edited. However the option Show Patches will still display the bounding boxes of the patches

Extracting patches at specific locations

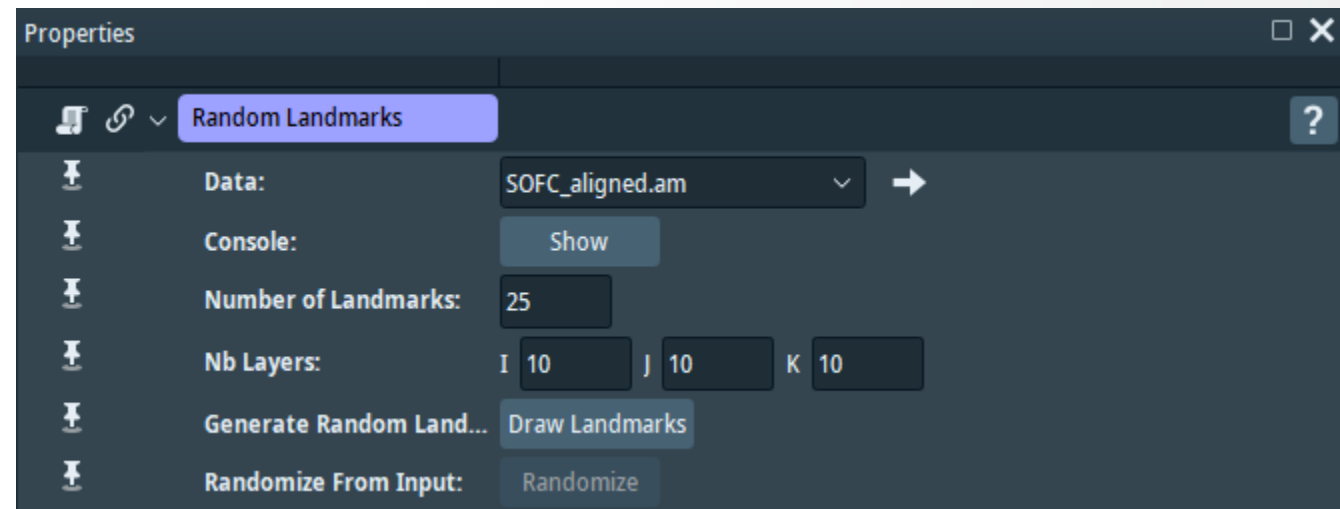
- The module « **Extract Patches** » extracts patches at locations defined by a 'Landmark' dataset
- The patches are extracted in an order corresponding to the indices of the Landmarks in the set.
- Only entire patches are extracted: if a landmark is too close to the image border, it will be skipped.
- See 'Random Landmarks' for randomly shuffling the landmark indices.



The option 'Show Patches' shows with a yellow lineset the patches that will not be extracted, because they are not fully inside the volume

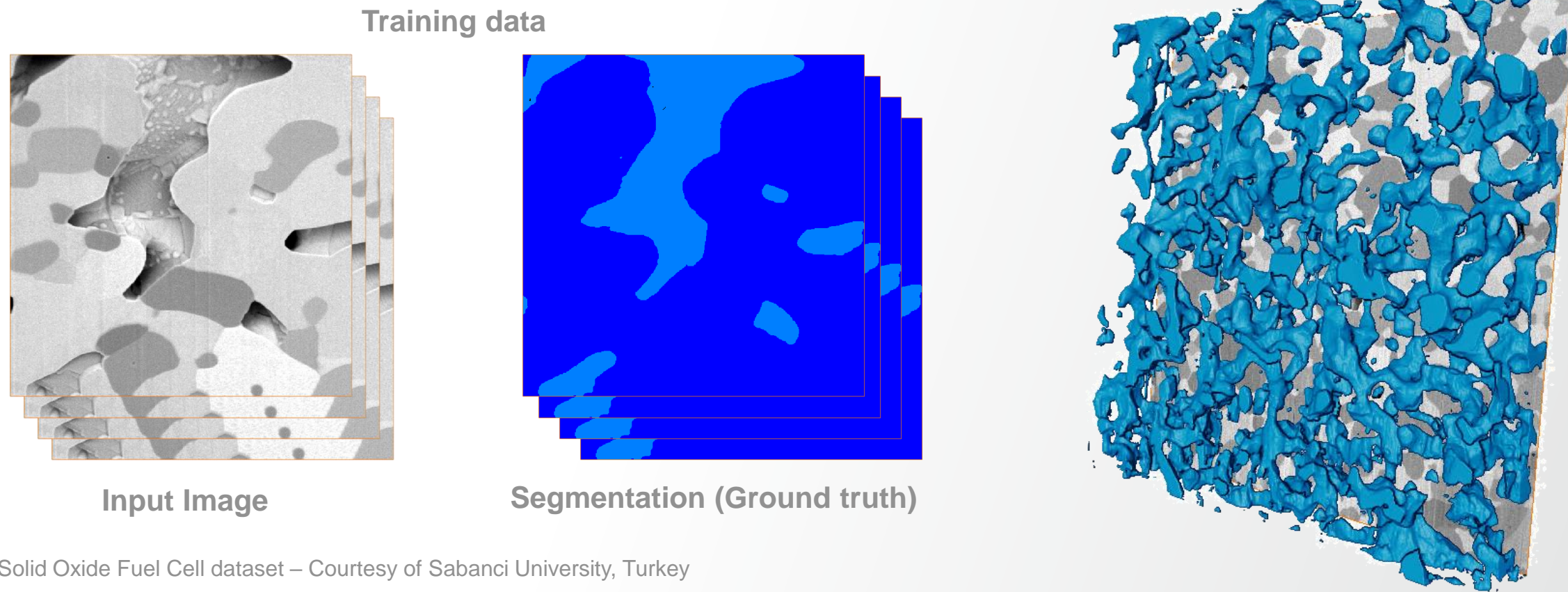
Defining Landmarks

- Existing modules *(Amira and Avizo only, not for PerGeos)*:
 - You can create an 'Landmarks' object by right-clicking in an empty area of the Project View
 - Create Object... Landmarks
 - The Landmark object has a 'Landmark Editor' allowing manual adding, deletion, or displacement of landmarks.
- The module 'Merge Landmark' with option 'Unify Sets' allows fusing two landmark data objects into a single one
- Xtra module « **Random Landmark** »
 - Right-click in an empty area of the Project View, Create Object... Random Landmarks
 - You can connect it to
 - An image, or a ROI Box module: in this case, random landmarks will be generated in the bounding box of the connected Data.
 - A Landmark Set: in this case, you can Randomize the indices of the landmarks.
This option can be useful to shuffle landmarks that were manually generated.



Manual Segmentation of patches in context

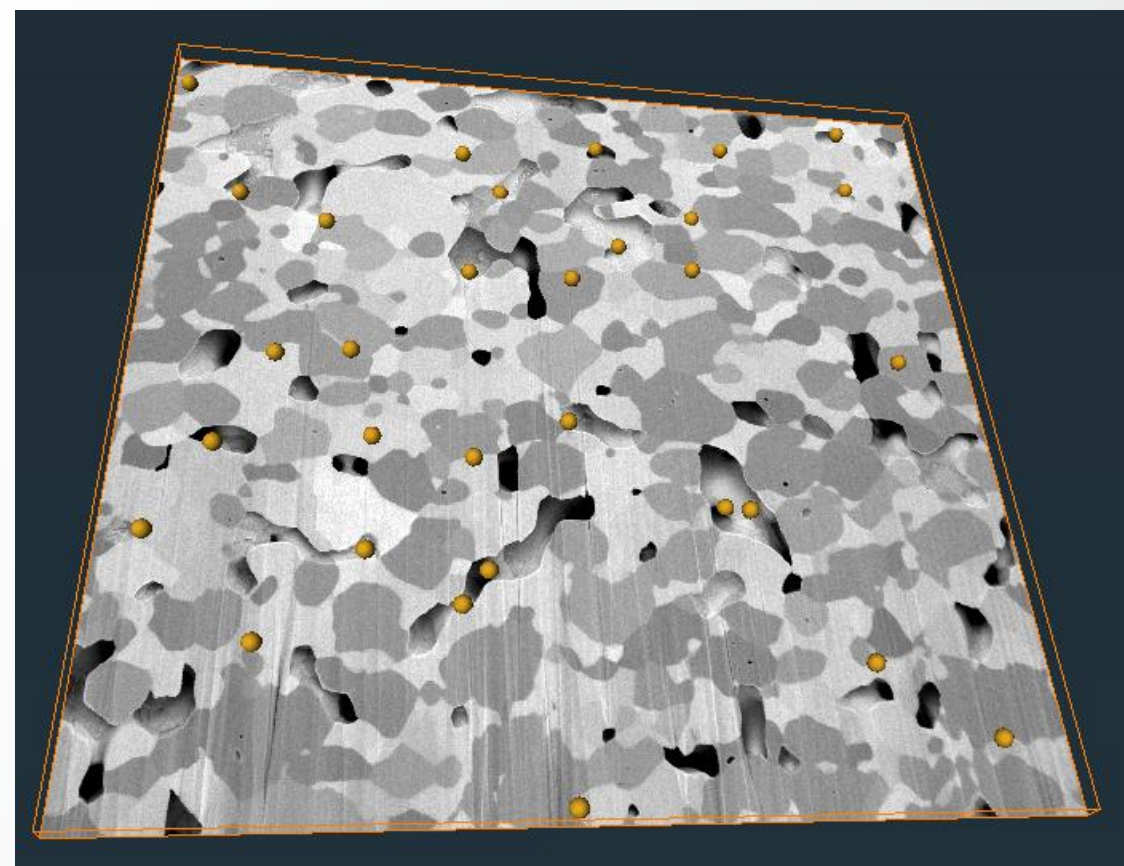
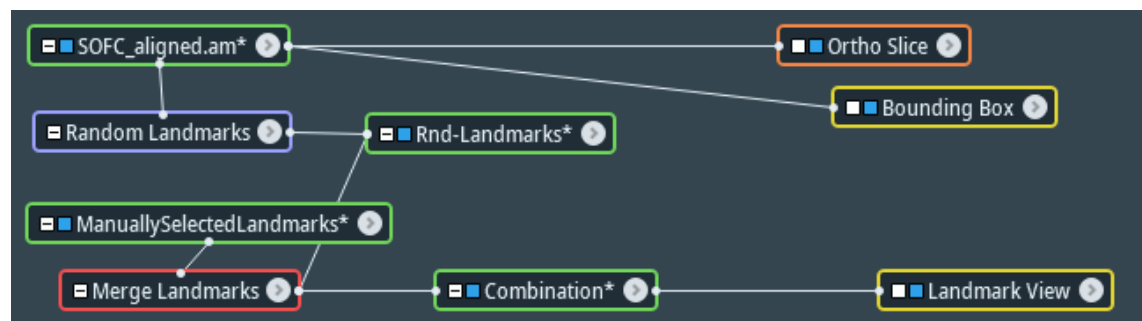
- Deep Learning is often used when only manual segmentation is truly reliable.
- Challenging segmentation tasks usually require that the data is observed in its context
- The next few slides propose a workflow for doing so, illustrated with the segmentation of pores in an Solid Oxide Fuel Cell data sample(*) imaged with FIB-SEM, and presenting pore-back artefacts.



(*) Solid Oxide Fuel Cell dataset – Courtesy of Sabanci University, Turkey

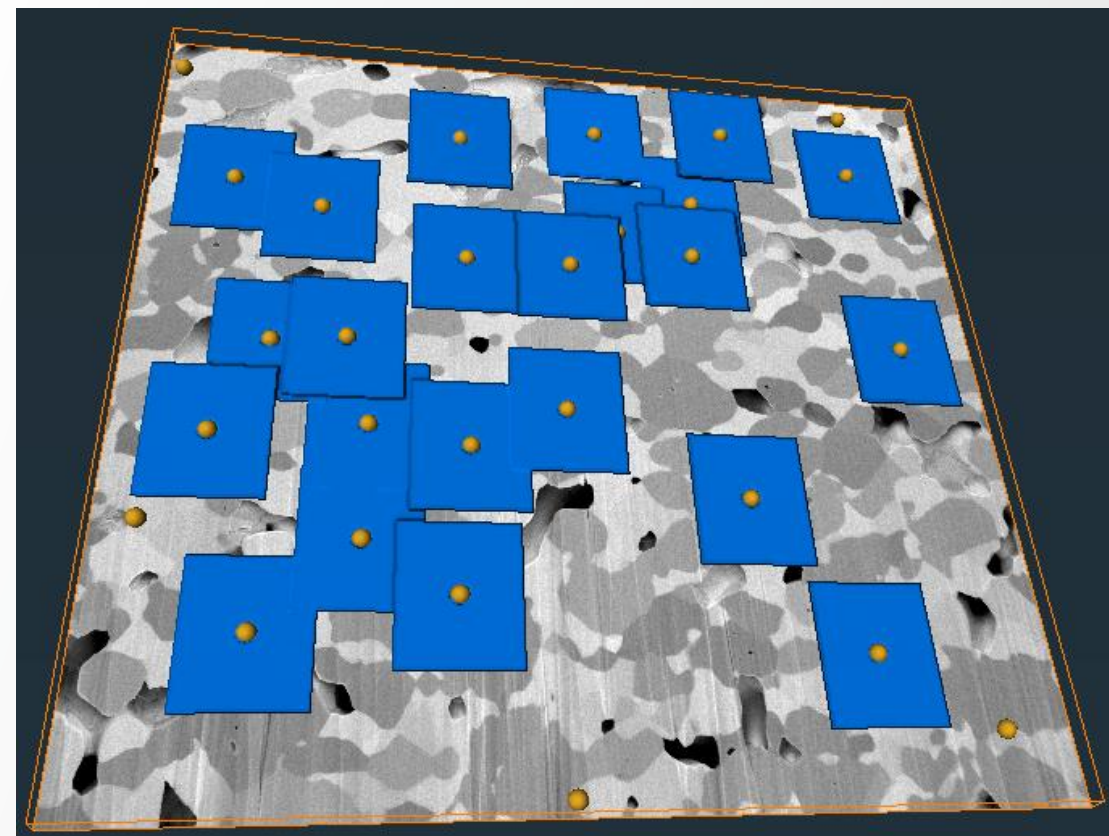
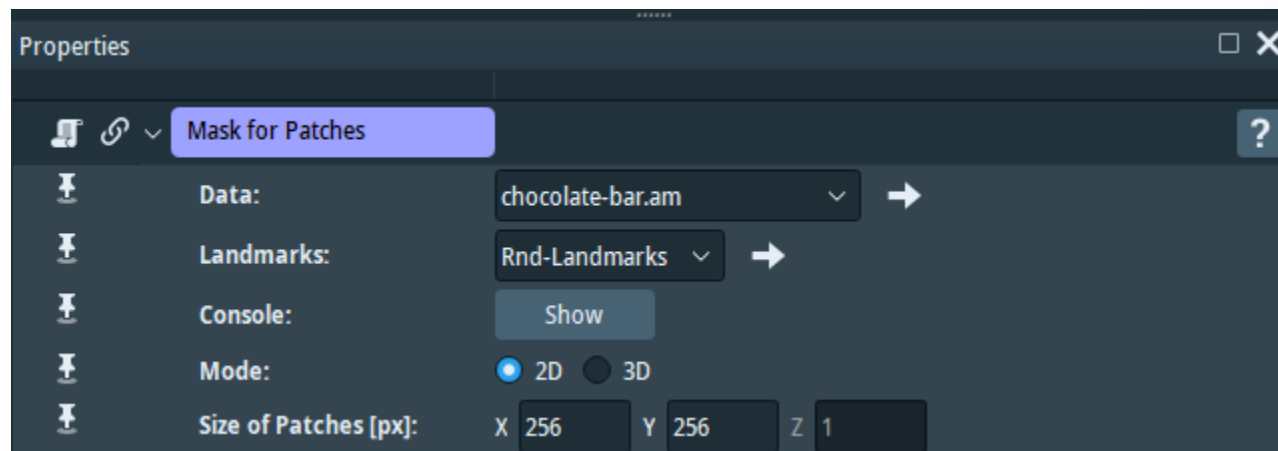
Step 1: create landmarks

- Create an appropriate Landmark set
 - Either manually or randomly positioned landmarks, or mixing both techniques
 - If necessary, use « Random Landmarks » to shuffle to order of landmarks



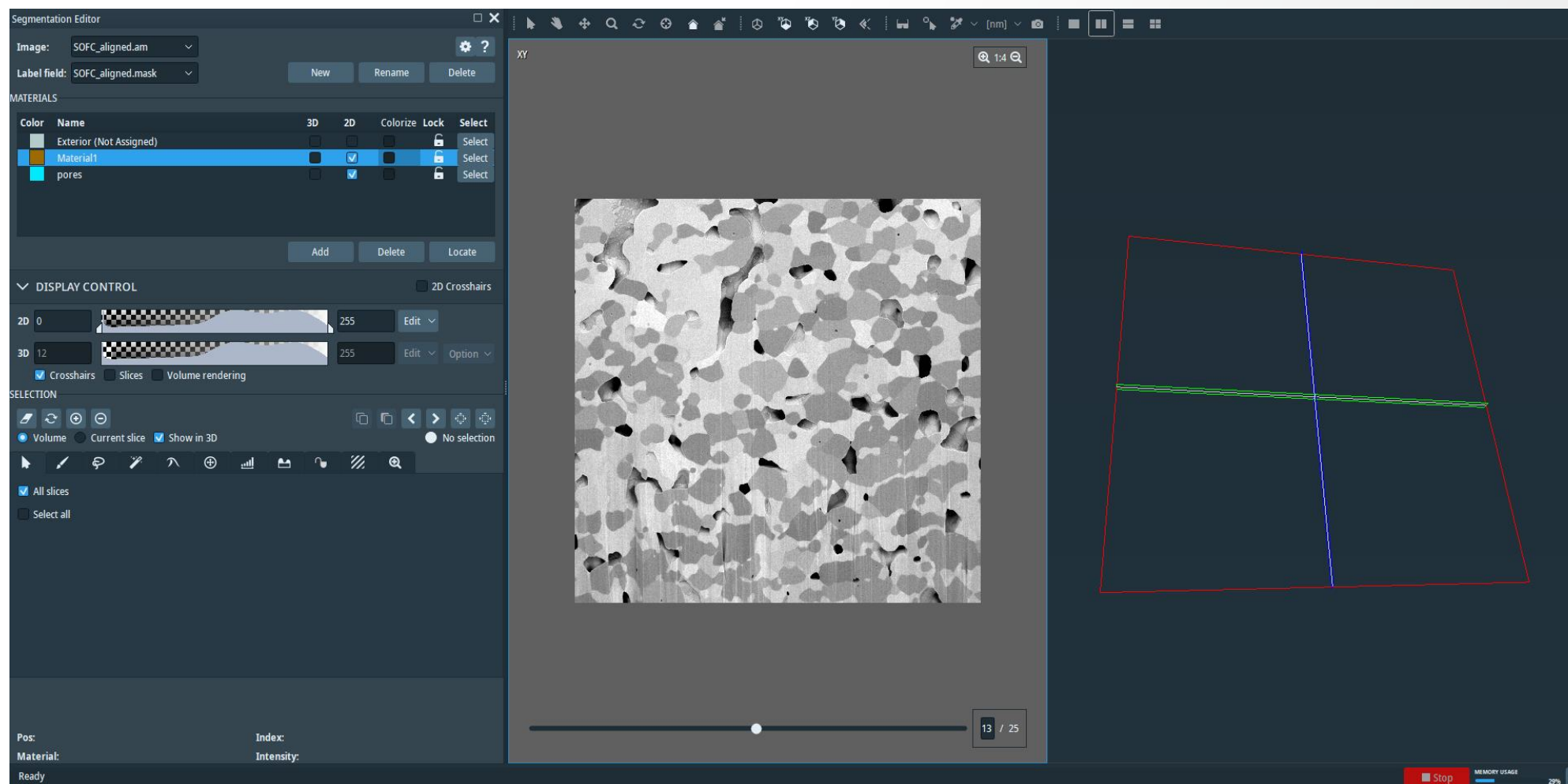
Step 2: Binary Mask for Patches

- Use « Mask for Patches » to create a binary mask of the location of the patches



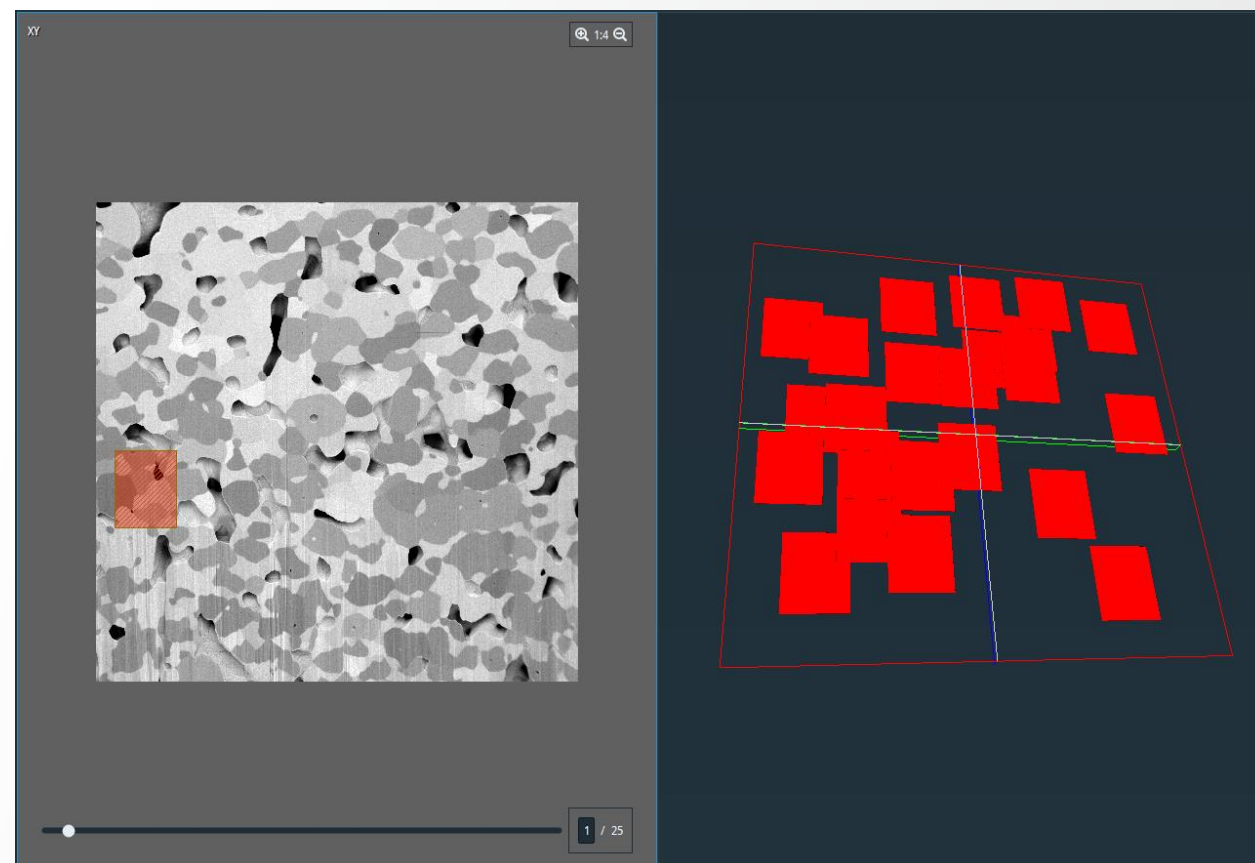
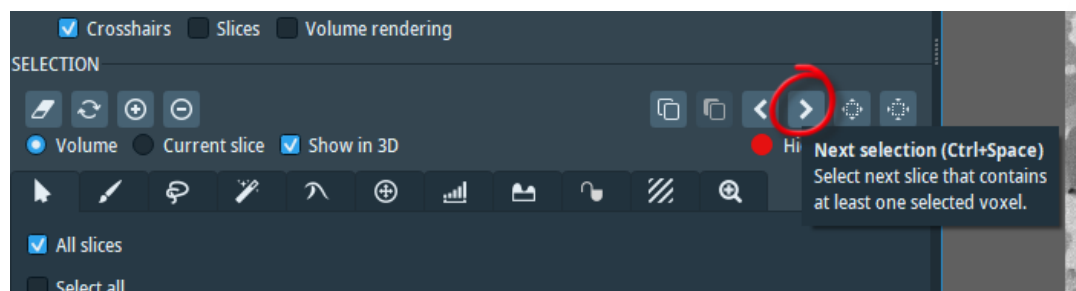
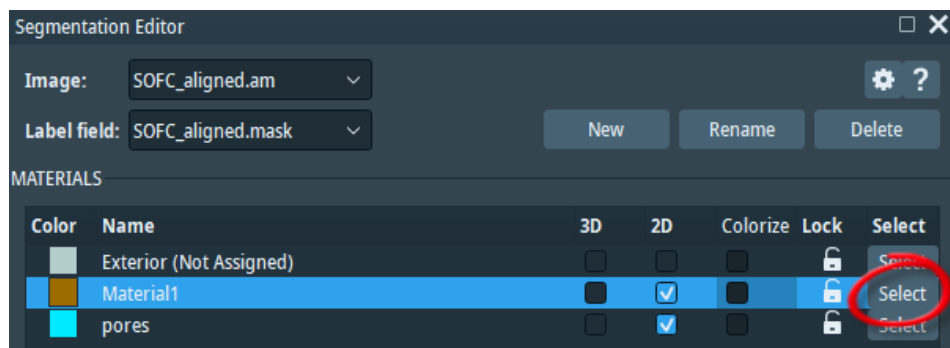
Step 3: Enter the Segmentation Editor

- Enter the Segmentation Editor.
- Configure it to 2 viewers mode
- See the 'material' corresponding to patches location is 'Material1'
- Create a new material for the manual segmentation



Step 4a: Browse through patches

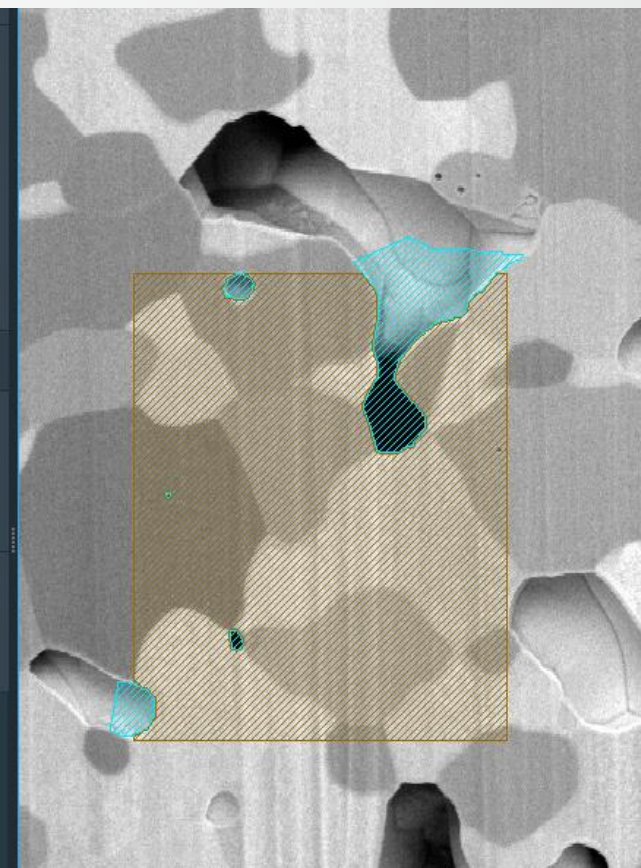
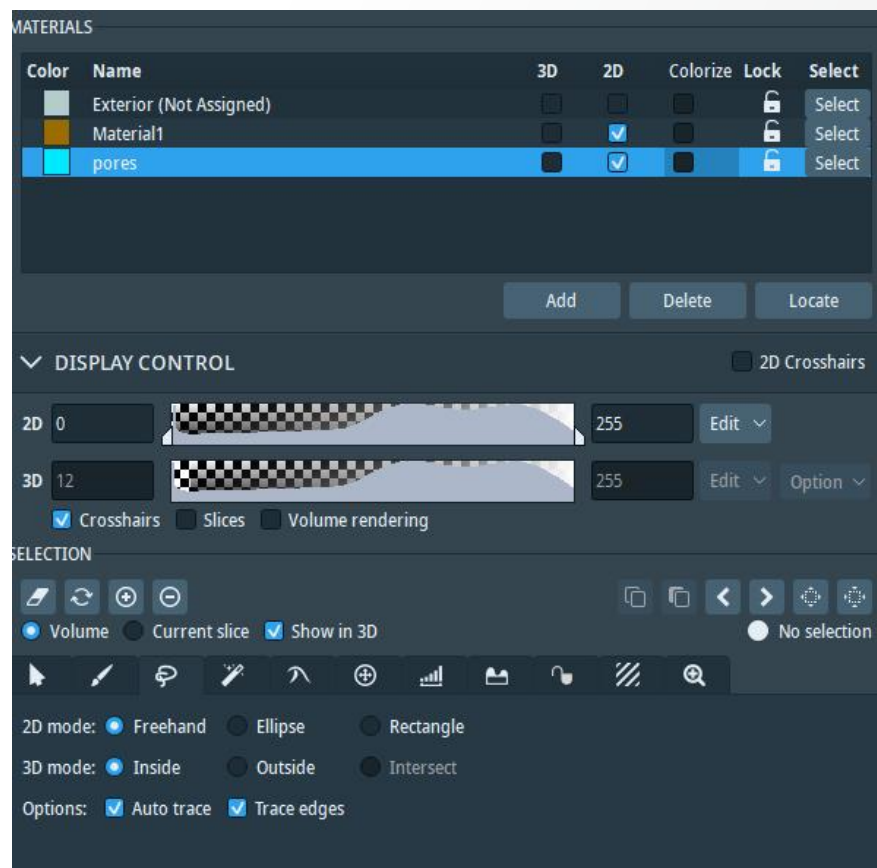
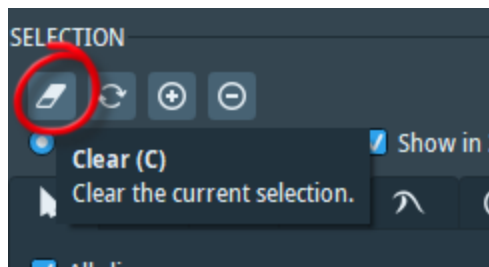
- Browse to slice 0
- Press « Select » on the first material to highlight all the patches areas
- Press 'Next Selection' to navigate to the first slice containing a patch (in our case, slice 0 had none)



- *Note for PerGeos users: the features 'Next selection' and 'Previous selection' are not proposed in the PerGeos Segmentation workroom, the proposed workflow will require manually browsing through the slides.*

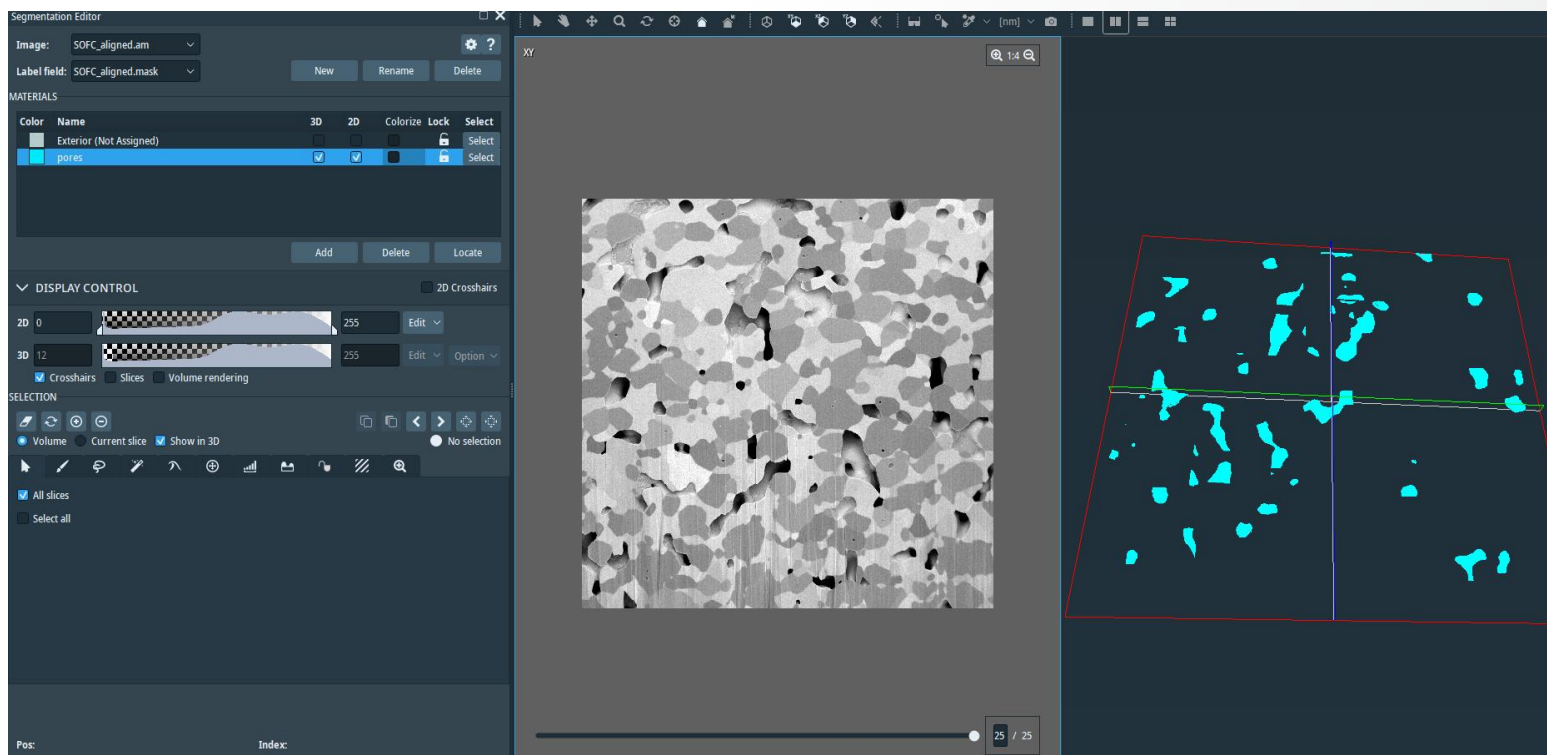
Step 4b: Browse through patches

- Clear the selection
- Manually segment the pores visible in the patch using any available tool
 - You can zoom in, and browse to the previous/next slides to get a better understanding of the data locally
 - It is not mandatory to stick to the strict limits of the patch
- Assign this selection to the 'pores' material
- Repeat for all patches in the slice



Step 4c: Browse through patches

- Repeat:
 - Select Material1 again, and browse to the next slice containing a selection as in Step 4a
 - Repeat segmentation of the patches as in Step 4b
- Until all slices are processed
- Finally, Delete 'Material1'
- Go Back to the Project Room



Step 5: Browse through patches

- Extract patches for the image
 - Create a « Extract Patches » module on the grayscale image
 - Connect to the appropriate Landmark
 - Make sure to use the same patch size as in Mask For Patches (Step 2)
 - Apply
- Repeat the same operations for the segmentation dataset
- Remember to save the project, or at least the landmark object to keep track of where your training data comes from. If you need more data for training, you may want to come back to the project to pick relevant locations

