Deep Learning Based Large-scale Histological Tau Protein Mapping for Neuroimaging Biomarker Validation in Alzheimer's Disease

Maryana Alegro[1,2,3], Yuheng Chen[1], Dulce Ovando[1], Helmut Heinsen[6,7], Rana Eser[1], Duygu Tosun[4,5], Lea T. Grinberg[1,7]

[1]Department of Neurology, University of California San Francisco, San Francisco, CA, USA
[2]Berkeley Institute for Data Science, University of California Berkeley, CA, USA
[3]Bakar Institute for Computational Health Sciences, University of California San Francisco, CA, USA
[4]Department of Radiology, University of California San Francisco, San Francisco, CA, USA
[5]Veterans Affairs San Francisco, CA, USA
[6]Julius-Maximilians University Würzburg, Würzburg, Germany
[7]University of Sao Paulo Medical School, Sao Paulo, Brazil

Correspondence: lea.grinberg@ucsf.edu

Abstract

Deposits of abnormal tau protein inclusions in the brain are one of the pathological hallmarks of Alzheimer's disease (AD) and are the best predictor of neuronal loss and clinical decline. As such, imaging-based biomarkers to detect tau deposits in-vivo could leverage AD diagnosis and monitoring from earlier disease stages. Although several PET-based tau tracers are available for research studies, validation of such tracers against direct detection of tau deposits in brain tissue remain unresolved. Large-scale voxel-to-voxel correlation has been challenging because of the size of the human brain, deformation caused by tissue processing that precludes registration and the amount of tau inclusion. In this study, we created a semi-automated computer vision pipeline for segmenting tau inclusions in billion-pixel digital pathology images of whole human brains, aiming at generating quantitative, tridimensional tau density maps that can be used deciphering the distribution of tau inclusions along AD progression and validate PET tau tracers. Our pipeline comprises several pre and post-processing steps developed to handle the high complexity of these brain digital pathology images. SlideNet, a convolutional neural network designed to work with our large-datasets to locate and segment tau inclusions is the heart of the pipeline. Using our novel method, we have successfully processed >500 slides from two whole human brains, immunostained for two phospho-tau antibodies (AT100 and AT8) spanning several Gigabytes of images. Our network obtained strong tau inclusion segmentation results with ROC AUC of 0.89 and 0.85 for AT100 and AT8, respectively. Introspection studies further assessed the ability of our trained model to lean tau-related features. As the final results, our pipeline successfully created 3d tau density maps that were co-registered to the histology 3d maps.

Keywords

Machine learning, deep learning, convolutional neural networks, Alzheimer's disease, histopathology, digital pathology, big data, imaging

Introduction

Alzheimer's disease was the only leading cause of death to have increased in frequency – nearly a 90% increase between 2000 and 2014 [1]. Currently, more than 5 million people over the age of 65 in the US alone have a clinical diagnosis of Alzheimer's disease and effective disease-modifying treatments remain elusive [2]. Alzheimer's disease features a long silent period that can last more than a decade before symptoms onset [3]. It is undisputed that AD treatment is likely to be more effective if started during its asymptomatic phase. Thus, it is essential to have reliable and accurate biomarkers available to enhance detection of AD in its earliest stages, test the effectiveness of new therapies, and monitor disease progression.

AD main neuropathological hallmarks are the extracellular accumulation of β-amyloid in the form of neuritic plaques, and intracellular accumulation of misfolded tau protein (AD-tau) – in the form of neurofibrillary tangles (NFTs), neuropil threads and dystrophic neurites[4–7]. Its progression follows a characteristic, non-random anatomical, neuropathologically-confirmed pattern[8], on which current staging schemes are based. The amount and spreading of tau tangles directly correlate with clinical decline, making it a potential biomarker for early-stage AD detection.

Molecular neuroimaging tracers for PET scans is the modality with the best potential to detect tau tangles in-vivo while providing spatial information, key for staging the disease. As tau tracers are being developed, questions about their sensibility, specificity, and accuracy are arising, and correlations between histological and neuroimaging datasets are the gold-standard approach validating PET tracers. Although attempts to validate PET tau tracers have been published, most of these studies have been limited to correlation of few brain areas and compare average signals within a relatively large area rather than taking a voxel-to-voxel approach, making it challenging to truly understand if the PET signal represents the molecule of interest, the nature of off-target signal and whether sensitivity and specificity of the signal is similar across the different brain areas.

Two significant limitations preclude higher resolution histology to imaging correlation studies, including registration hurdles caused by the substantial geometric differences and between neuroimaging and histology image sets and dataset size. In previous studies, we demonstrated that it is possible to achieve excellent co-registration between histological and neuroimaging 3D reconstructions of human brains by combining sophisticated computer tools with improvements in histological process to minimize deformation[9,10].

Here, we present a semi-automated computational pipeline for creating high-resolution Tau density maps from whole human brain digital pathology images. Our pipeline uses a wide range of computer vision techniques together with modern deep learning algorithms and high-performance computing (HPC) capabilities, being able to process large-scale histological datasets, leveraging a point-to-point correlation to neuroimaging modalities. Using the proposed method, we processed more than 500 whole-brain slides coming from brains at intermediate and late AD stages, spanning several Gigabytes of data. We successfully mapped tau protein across the brains volumes and created 3d Tau density maps.

In the following sections, we provide a detailed description of the methods and quality control steps of this pipeline and the software we created. We will illustrate the utility of our pipeline for providing robust anatomical priors for validating neuroimaging methods in 2 cases, by showing unprecedented quantitative 3D whole human brain maps of tau protein detected by different antibodies at early stages.

Results

Computational Pipeline for mapping tau tangles

We engineered a comprehensive computational pipeline (Figure 1) for locating and measuring the density of tau tangles in large-scale whole human brain digital pathology datasets. Our pipeline comprises

independent modules (Figure 1, 1-6) that can be executed in parallel and are responsible for processing the different image modalities involved in the project. As part of the *imaging module* (Figure 1.1), we scanned our histological sections using a custom-built whole slide scanner and stitched the raw tiles using specialized software. Scanning time and amount of generated data were dependent on the size of the selected region of interest (ROI). On average, scanning a whole-brain section took 90 minutes, yielding approximately 1500 tiles (20Gbs of data). The stitching step took, on average, two days to complete on an HPC system, and generated a high-resolution histological image of approximately 80000x50000 pixels (~20Gbs). Figure 2a show an example of stitched image of a histological slide. The stitching step also generated a downsampled - 10% resolution version of the histological image. The blockface images created during block sectioning (see Histological processing, in Methods) were preprocessed in the *blockface preprocessing module* (Figure 1.2). First, each blockface image background was automatically removed, leaving just the brain areas. Next, all images were sorted and grouped. Since the block or camera may have moved during sectioning, slices were manually aligned to the middle slice using affine registration. Finally, slices were stacked together using in-house built software to form a 3d image volume, which was in turn used in the low-resolution histology pre-processing module.
In the *low-resolution histology pre-processing module* (Figure 1.3), the 10% resolution histological images had their background automatically segmented, leaving just the brain area. White matter (WM) masks were manually created and subtracted from the brain mask to create gray matter (GM) masks. The segmented low-resolution histological images were aligned to their corresponding pre-processed blockface images using a combination of registration algorithms and registration maps were stored for later use.
The core of our pipeline is the *high-resolution histology segmentation* module (Figure 1.4), responsible for segmenting tau inclusions and generating the tau density maps, here called heatmaps. In the heart of this module lies a convolutional neural network (CNN) we designed for this study, responsible for segmenting the tau tangle (see Methods for a detailed description). Figure 3 shows our CNN architecture. First, datasets for CNN training were created by sampling patches of GM, extracted from random locations distributed throughout the entire 3d brain volume. A different dataset was created for AT8 and AT100 markers. These patches were randomly split into 80% for training, 10% for testing and 10% for validation, yielding a total of 12 datasets. All patches were manually labeled with the help of a supervised learning segmentation algorithm. SlideNet was trained using a validation hold-out scheme, where training and testing set were used during training, but the validation set was just used for computing the network statistics. In parallel, high-resolution histological images were tiled into smaller images (Figure 1.4d).

During this step, the low-resolution GM masks were upsampled to match the histological image resolution and, were, in turn, also tiled. Histology tiles had their background segmented using their respective mask tiles, leaving just the GM regions. This allowed us to disregard tiles that did not contain GM, thus, avoiding unnecessary computations. Each tile was segmented using our trained CNN model, resulting in a probability map that was thresholded to create a binary mask. Heatmaps were computed for each tile (Figure 1.4f) and combined to form a high-resolution heatmap (Figure 1.4g) that had the same dimension of the original histology. A 10% resolution image was created for each high-resolution heatmap and registered to their respective blockface images using the registration maps pre-computed in the *low-resolution histology pre-processing* module (Figure 1.4h). We stacked all registered heatmaps together to create a 3d image volume. In the *MRI pre-processing* stage, MRI datasets went through a standard pre-processing pipeline, where they are corrected for field inhomogeneity, had their background and skull removed and are automatically registered to their respective blockface volumes. The final step of the pipeline, *heatmap to MRI 3d registration*, is responsible for having the heatmaps on MRI space. In other to do so we just overlaid our registered heatmaps on the registered MRI volumes, what guaranteed that we could perform point-to-point correlations between heatmap and MRI signal. We transported the heatmap to MRI space by inverting the 3d registration maps. The resulting datasets were visualized using Freeview for 2d visualization and Amira for 3d visualizations and reconstructions.
We have used this pipeline to successfully processed 524 slides from two brains with early and a late-stage Alzheimer's disease, respectively. As a result, we created 3d datasets of the tau tangle heatmap for both specimens, which can be used for clear visualization or quantitative correlations with MRI signal.

Figure 4 shows examples of early and late-stage heatmaps overlaid on their blockface images. Here, the hotter colors indicate higher density of tau.

Neural network is effective in locating tau inclusions

We developed a CNN responsible for locating tau tangle in the pre-processed histological images. We trained our models to classify pixels inside 204x204 patches of images as either tau or background. Each image tile generated by the pipeline *tiling* step (Figure 1.4d) was split into 2550 patches with 90 pixels of overlap and classified by our CNN. Unlike most common CNNs that work by outputting a single class score, our design is based on fully connected last layer, that outputs bidimensional probability maps of the tau locations. In order to quantitatively assess the performance of our CNN, we computed the receiver operating characteristic (ROC) curves for the AT8 and AT100 markers. Moreover, we computed the ROC for both the testing and validation datasets. The AT8 model obtained the area under the curve (AUC) of 0.85 for in both datasets, while the AT100 model performed a little better on the validation dataset – 0.87 on training dataset and 0.89 on the validation dataset. Figure 5 shows the ROC graphs. Figure 6 shows an example of histology tile (left) and the probability map created by the CNN (middle) and the thresholded binary mask (right).

We set the tau segmentation threshold value to 0.7, meaning only pixels with a confidence of being tau equal or larger than that value would be classified as such. In Figure 7 we can see where this threshold – indicated by the red stars on the graphs - falls on the Precision-Recall curve, for AT8 and AT100 testing and validation datasets. This value was chosen heuristically to guarantee high precision, thus reducing the number of false positives and guaranteeing that most pixels classified as tau really belong to that class.

Introspection shows network correctly identifies tau inclusions

Besides standard precision-recall and ROC curve evaluation, we also used neural network inspection techniques to assess if the network was learning tau tangle features. We first used gradient guided class activation maps (Grad-CAM)[11] to evaluate what features, on randomly selected images, were most important to drive the network to a particular decision. We also used perturbations for testing if the network was segmenting tau based solely on spatial localization or if it was using more relevant features.

Standard Grad-CAM works by computing the gradient between a user-defined class neuron in the last network layer and an intermediary target layer and then multiplying its mean value with the target class activations. This technique was designed to work with networks that have fully connect or a global pooling layer[12], which are responsible for mapping the information spread across the convolutional layers to a single neuron in the last. This is not the case for SlideNet, which outputs a 2d probability maps and lacks the property of fragmented mapping information to a single neuron. We adapted Grad-CAM for our design by selecting multiple output neurons inside a Tau tangle, on a randomly selected test image – meaning neurons that should be activated for the class Tau, with the help of a binary mask, and computing the mean of the Grad-CAM maps generated for each selected neuron. The result can be seen in Figure 8, where 8a. show the original test image, 8b the binary mask and 8c the Grad-CAM results, with brighter values indicating stronger influence in the network decision. We can see the features that activate the selected pixels inside mostly localized around both Tau tangles. We repeated the same experiment to determine if the network was also correctly discriminating background, by using a mask to select background pixels as can be seen in Figure 8d. The resulting Grad-CAM can be seen in Figure 8e, showing good localization of the background as well. Finally, we wanted to test if the network was learning to detect tau based just on pixel spatial localization or was using more relevant information. We created a perturbed image by partially covering the Tau tangle with a patch of background and repeating the background Grad-CAM experiment. In Figure 8f we have the perturbed image (red arrow shows the artificial patch), Figure 8g shows the binary mask and Figure 8i shows the Grad-CAM result. We can see that the network correctly recognized the patch as background.

Tridimensional tau maps

The final product of our pipeline was four 3d tau distribution datasets, showing the distribution of tau stained for AT8 and AT100, on two entire whole human brains. Figure 4 shows examples of tau map volumetric rendering overlaid on blockface images. The figure on the right shows AT100 marker distribution on the

early-stage AD specimen, and the figure on the left shows the same marker distribution in the late-stage case. Both images were rendered in the hippocampus area, a brain region traditionally affected by AD, to show the contrast between an early and late-stage case.

Discussion

Here we reported a scalable, parallelizable framework for mapping tau protein in whole human brains histological slides. Our framework was used to create 3d tau distribution maps that can easily be compared point-to-point to other neuroimaging modalities. Figure 10 shows a volumetric render of such maps. During this research we had to overcome three significant challenges: (1) dealing with large, unpredictable image deformation that precludes registration; (2) performing large-scale image segmentation on several hundreds of images; and (3) dealing with a large amount of data.

Materials and Methods

Specimen Procurement, Histological Processing and brain slabbing

The pipeline was developed and tested using two whole human brains. The first brain was classified as Braak stage 4 and neuropathological analysis[13]. The second brain belonged to a female patient with advanced dementia and severe Alzheimer pathology. Both individuals underwent AV1451 PET scan and structural MRI T1 weighted weeks before death. Upon death, the brains were scanned again in a 3t MRI and procedure as a whole. Both specimens were fixed by immersion in 4% buffered paraformaldehyde for 3 weeks. To avoid deformation, the brains were stored upside down, hanging by the circle of Willis for 3 days and after that, they were mounted in a plastic skull base molded after the subject's computerized tomography obtained as part of the in-vivo imaging and fixed for another 18 days inside a bucket filled with buffered 4% paraformaldehyde. After fixation, the specimens were embedded in celloidin and the blocks allowed to congeal. The blocks were sectioned using a sliding microtome (AO 880, American Optical, USA) equipped with a 14" long C-knife. The brains were sectioned in serial sets of coronal slides, each set containing five 160 µm-thick sections to obtain an accurately aligned stack of unstained serial section photographs, digital photographs were acquired directly from the blockface following each stroke During block sectioning using a high definition DSLR camera (EOS 5D Mark II, Canon, Tokyo, Japan) mounted on a copy stand arm (Kaiser Fototechnik, Germany) and linked to a computer. PET scan voxel size varies from 3 to 5 mm. During process, brain tissue shrinks about 30 to 40%. Thus, we chose to have 5 slides in each set, so each PET scan voxel would contain at least 2 full histological sets in the Z-axis, opening the opornutiy to probe multiple markers per voxel. Histological sections were stored in 70% alcohol for further processing. Details for tissue processing celloidin embedding and cutting have been debrided by us elsewhere[1,14–16]

Immunohistochemical labeling of abnormal tau protein inclusions

In Alzheimer's disease, tau protein becomes hyperphosphorylated and aggregated into inclusions. It is unclear which of these abnormal tau forms the tracers are binding to. We designed our protocol to accommodate different probing forms of protein accumulations. Immunostaining remains an attractive validating method to probe proteinaceous inclusion while preserving tissue integrity and providing spatial information. For this study, we chose two widely used antibodies against phospho-tau, At100, and AT8. As a rule, all 1st sections from each set were immunostained to detect tau protein phosphorylated at epitopes pSer202+Thr205 (AT8, Thermo). All 2nd sections from even-numbered sets were immunostained to detect tau protein phosphorylated at epitopes pThr212+Ser214 (AT100, Thermo). Immunohistochemistry reactions were done in batches of 50 to ensure homogeneity. Each batch contained positive and negative control for comparison. The controls were serial sections from a single subject not belonging to the study for quality control purposes.

For the quality verification process, LTG visually reviewed the positive control slide of each batch and compared the staining process to other tissue stained with the same antibody to assess the adequacy of the staining process. Staining was deemed unacceptable if no staining was present, very weak staining was observed as compared to another stained positive control slides, and/or very strong background,

similar to the level of staining, was observed. Batches whose positive control was deemed unacceptable were excluded and the staining process was repeated on the 2nd set of contiguous tissue sections, which in turn went through the same quality control process.

Whole Slide Imaging

Since our histological slides (4 x 6") were too big for commertialy available scanners, we developed our own cost-effective whole slide scanner (Figure 9). The hardware is comprised of a high-precision, 6" travel range, industrial XY stage (Griffin Motion), an Olympus manual focusing box, color CCD camera (Qimaging Micro publisher 6) and 5.5X machine vision objective (Navitar Zoom 6000) mounted directly on the camera. A lightbox with diffuser, mounted on top of the XY stage, is used for illumination. Slides are loaded directly to a 3d printed slide mount, on top of the lightbox. The scanner is controlled by a software developed in-house using Macro Manager 2.0.

Stitching: tiles are stitched together to form a high-resolution histological image using TeraStitcher[17], which is capable of working with several Gigabytes of data while maintaining a small memory footprint. During the stitching process a 10% resolution version of the image is also created and is used during the histology pre-processing and registration steps and also for visual inspection. This setup yielded a 1.22um pixel resolution.

Dataset Creation

In order to create our training, testing and validation datasets, we drew 1024x1024 pixel patches from randomly selected gray matter locations throughout the entire full-resolution brain image dataset. A set of 100 patches was created for each marker (*AT100* and *AT8*), from each brain respectively, totaling 4 datasets. The patch extraction routine was written in Python and completely automated, running on UCSFs' Wynton cluster in an embarrassingly parallel way.

Each patch was manually masked for background and tau inclusion with help of Fiji's Trainable Weka Segmentation plugin. Here, the user manually selected sample pixels belonging to *tau* and background classes that were, in turn, used to compute several features such as Gaussian filters, Hessian, membrane projections, mean, maximum, anisotropic diffusion, Lipschitz, Gabor, Laplacian, entropy, Sobel, difference of Gaussians, variance, minimum, median, bilateral filter, Kuwahara, derivatives, structure, and neighbor values. A linear SVM classifier (LibLinear) was then used to generate an initial tau segmentation. The user would retrain and refine the initial segmentation until they obtained satisfactory results. Masks where manually fine-tuned using an image editor (Gimp). All final masks went thought quality control by a highly experienced pathologist.

Finally, Patches for each maker were combined on bigger datasets, totaling 2 datasets of 200 patches each and were randomly split in 80% of patches for training, 10% for testing and the remaining for validation.

Histology Pipeline

Pre-processing blockface images

Blockface images have their background segmented using a semi-automated graph-based algorithm. Briefly, the user selects brain and background sample pixels using a graphical user interface (GUI). Images are then converted to LAB color space, and the algorithm computes mean color difference maps (ΔE). ΔE is defined as the distance in LAB space:

$$\Delta E = \sqrt{L_{diff}^2 + A_{diff}^2 + B_{diff}^2}$$

With L_{diff}, A_{diff}, B_{diff} being the difference values computed as:

$$L_{diff} = L_i - \mu(L_r)$$
$$A_{diff} = A_i - \mu(A_r)$$
$$B_{diff} = B_i - \mu(B_r)$$

Where L_i, A_i and B_i are image LAB channels, L_r, A_r, B_r are LAB channels of reference pixels selected by the user, and μ(.) is the mean. Pixels in ΔE, whose color is similar to the reference values, appear dark in (smaller distance) while cell pixels are brighter (larger distance). We compute brain and background ΔE

maps using the manually selected pixels as reference values and perform a histogram threshold using the Otsu's method to obtain binary masks. We, in turn, combine both masks to obtain the brain segmentation. It is common that several undesired objects linger after the initial segmentation. It is usually caused by reflections from liquid on top of the celloidin block. The segmentation is further refined using a graph-based method to remove the undesired objects. In this method, image objects and their relationship are modelled as a weighted graph, where connected structures are considered the vertices. Edge weights are computed using a similarity function computed from color and distance values. The graph is partitioned using NCuts, leaving just the brain area.

Commonly, the camera or brain have to be repositioned several times during sectioning to adjust for changes in block size, causing the blockface images to be misaligned in relation to each other. We use semi-automatic registration to correct this artifact. Having the middle slice as the reference image, we use Matlabs' registration GUI to select landmarks for computing affine registrations. Finally, the aligned blockface images are stacked together to form the blockface 3d volume using ITK. This 3d volume is later used as an intermediate space for creating the 3d tau-protein maps.

Low-res histology background segmentation

The 10% resolution histological image is converted to LAB color space. In that space, background pixels are consistently darker than brain pixels and segmentation is performed through histogram thresholding using the triangle algorithm. The resulting binary masks, here denominated brain masks, are used to erase all background pixels. Moreover, brain masks are combined with white matter mask to create gray matter masks that guide the entire segmentation process and are also used during patch extraction for generating our training and validation datasets.

2d registration to blockface image

After background segmentation, the 10% resolution histological images are aligned to their respective blockface images using a combination of manual and automatic registration. Due to presence of an excessive number of artifacts caused by histological processing, such as tearing, shrinking and shearing, we decided for a manual initialization of the registration using MIPAV spline-based registration. Here, the user manually selects landmarks on both the histology and blockface images. MIPAV then generates a warped image and a registration warp file. After the initial registration, the image goes through a diffeomorphic registration using the 2d SyN algorithm, which is based on the large deformation diffeomorphic deformation model metric mapping (LDDMM) method. Both the registered images and registration mappings are saved for use later in the pipeline.

Image and Mask Tiling

Here, each full-resolution histological image is first tiled to reduce memory footprint during image segmentation, with tile size corresponding to approximately to $5mm^2$ of tissue. Tile coordinates and dimensions are saved as XML metadata files. The histological images' respective 10% resolution gray matter masks are then rescaled to match their full-resolution dimension and tiles following the same procedure used for histology. Finally, histological tiles are masked using their respective gray matter tiles, leaving only the regions of interest on the image. Tile that are left empty or having less than 5% of tissue pixels are ignored during segmentation. This way, we reduce the overall computational time and guarantee we are only processing relevant information. Image and mask tiling routines were developed in Python and ran on UCSFs' Wynton cluster in an embarrassingly parallel way, having one pipeline instance for each histological image.

Deep-learning based segmentation

A convolutional neural network is the core of our segmentation pipeline. We designed SlideNet, a UNet based neural network capable of working with color information and outputting tau presence confidence maps that are later thresholded to create binary maps.

Model development and training

Our model has an 204x204x3 pixels input layer to accommodate RGB images – we chose this particular size to match $1mm^2$ of tissue in our WSI scanner resolution. In our model, the image is pushed through three contractions blocks, the bottleneck, and upsampled by three expansion blocks. Figure 3 shows the SlideNet architecture together with each layer tensor size.

Each contraction block is composed of two convolution layers that use 3x3 kernels, stride of 1 and ReLu activation, followed by a 2x2 max pooling layer and 0.1 rate dropout. The bottleneck is composed of two convolutional layers that use 1x1 kernels, stride of 1 and ReLu activation. Expansion blocks are composed of a 2x2 upsampling layer followed by a 0.1 rate dropout and two convolutional layers that use 3x3 kernels, stride of 1 and ReLu, with the exception of the last expansion block that use 3x3 upsampling. The last layer reshapes the data to a 20000x2 tensor and softmax activation.
Training is performed using standard backpropagation, with cross entropy as the loss function and Adam for optimization. The learning rate is estimated using the cyclical learning rate[11] method.
The network was developed in Keras (https://keras.io) on top of TensorFlow (https://www.tensorflow.org). We trained the network using minibatches of 32 images during 100 epochs, or until the loss curve plateaued. We also used massive data augmentation, performing real-time random rotations, shear, horizontal and vertical flips. All training and inference were performed on a NVIDIA Titan X GPU with 12Gb of RAM.
SlideNet outputs confidence maps of the existence of tau, which are thresholded to generate binary masks. Threshold value was set to 0.7 for both AT100 and AT8 datasets and pixel with confidence higher than this value were considered tau.

Heatmap computation
The binary tiles are transferred back to our cluster, where the heatmaps are computed. On each tile, we compute the mean amount of tau, indicated by the mean number of pixels belonging to foreground, inside a 1 μm^2 of tissue, roughly an 8x8 pixels block. The heatmaps are generated as tiles having the same dimension of binary tiles where each 1 μm^2 block is filed out with the mean number of pixels belonging to tau. Tiles are stitched together to create a heatmap having the same resolution as the original histological image and is latter resized to 10% of its resolution.

Heatmap normalization
Heatmaps are normalized to a reference slide. Our normalization procedure is inspired by grayscale histogram matching algorithms. Our method works by adjusting the histogram of the target images to that of a reference image. Since our images are continuous, we begin by binning them in 1000 bins and computing their normalized histograms. We then compute the cumulative distribution function (CDF) for both histograms. For each non-background pixel in the target image, we use the CDF to map its value to a bin and get its edge values. We then find the closest match to the target pixel CDF value in the reference image CDF and use the later to retrieve its respective bin edges. We finally obtain the normalized value by computing a linear mapping of the target pixel value from [c,d] to [a,c]. Our normalization step was developed in Python and run on Wynton in an embarrassingly parallel task.
Heatmap to blockface alignment and stacking
The 2d registration maps computed during the 2d registration to blockface image step are applied to the 10% resolution normalized heatmaps, yielding heatmap registered to their respective blockface images. These images are then stacked together to for a 3d volume.

MRI pipeline
Field inhomogeneity correction
Since field intensity inhomogeneity may reduce the overall processing quality, we begin our processing with a correction step for this artifact, using the previously validated N3 algorithm.
Brain segmentation
This step aims to eliminate all background noise and unnecessary information (such as bones and meninges) from the images that could cause registration errors and slow down the computational processing. We perform the segmentation by using the FSL BET tool.
Spatial Resampling
Histological images have higher resolution than their MRI counterparts. While digitized histology voxels have sub-millimetric volume, a standard clinical MRI acquisition has a voxel size of around 1x1x1 mm. Direct registration of these images will cause the histology to be resampled to a lower resolution space, introducing loss of contrast and anatomical information of interest. We instead upsample the MRI volume

to a higher resolution space (0.33x0.33x0.33 mm voxel size) by employing the cubic splines interpolation. We found this resolution to be computationally tractable, while still allowing visualization of fine details in histological images, such as deep brain structures – i. e. hippocampi, substantia nigra, and caudate nucleus. Figure 3f shows the same slices before (first row) and after (second row) skull stripping.

Histology to MRI 3D registration

Three-dimensional histology to MRI registration is performed using the 3D SyN algorithm. SyN is based on the large deformation diffeomorphic deformation model metric mapping (LDDMM) method whose mathematical properties are especially appealing for our problem. Diffeomorphic maps are smooth and invertible functions guaranteeing that no folds or vanishing tissue will occur during the registration process. These models also enforce a 1-to-1 mapping between the movable and reference voxels and are suitable to handle large deformations. SyN is a mature method that has been reported to outperform several popular registration methods. Here, SyN is used with Mattes mutual information for measuring similarity, given that it is the most robust option for handling multi-modality registration problems. All 3D registration maps computed in this step are stored for future use. Figure 3h is an example of 3D histology to MRI registration using the described method. The first row shows sagittal, axial, and coronal checkerboard representations of a histology slice, overlaid on its respective MRI. The right-hand side shows a 3D rendering of one registered volume overlaid on its MRI.

3d Reconstruction and Visualization

The 3d registration maps generated during the **Histology to MRI 3D registration** are inverted and used to warp the MRI to blockface space, allowing direct comparison between MRI signal and tau density. Freeview is used for visualization of 2d slices. Amira (ThermoFisher Scientific) is used for 3d reconstruction and visualization.

Declarations

Acknowledgements

We would like to thank the support of NVIDIA Corporation with the donation of a Titan V GPU.

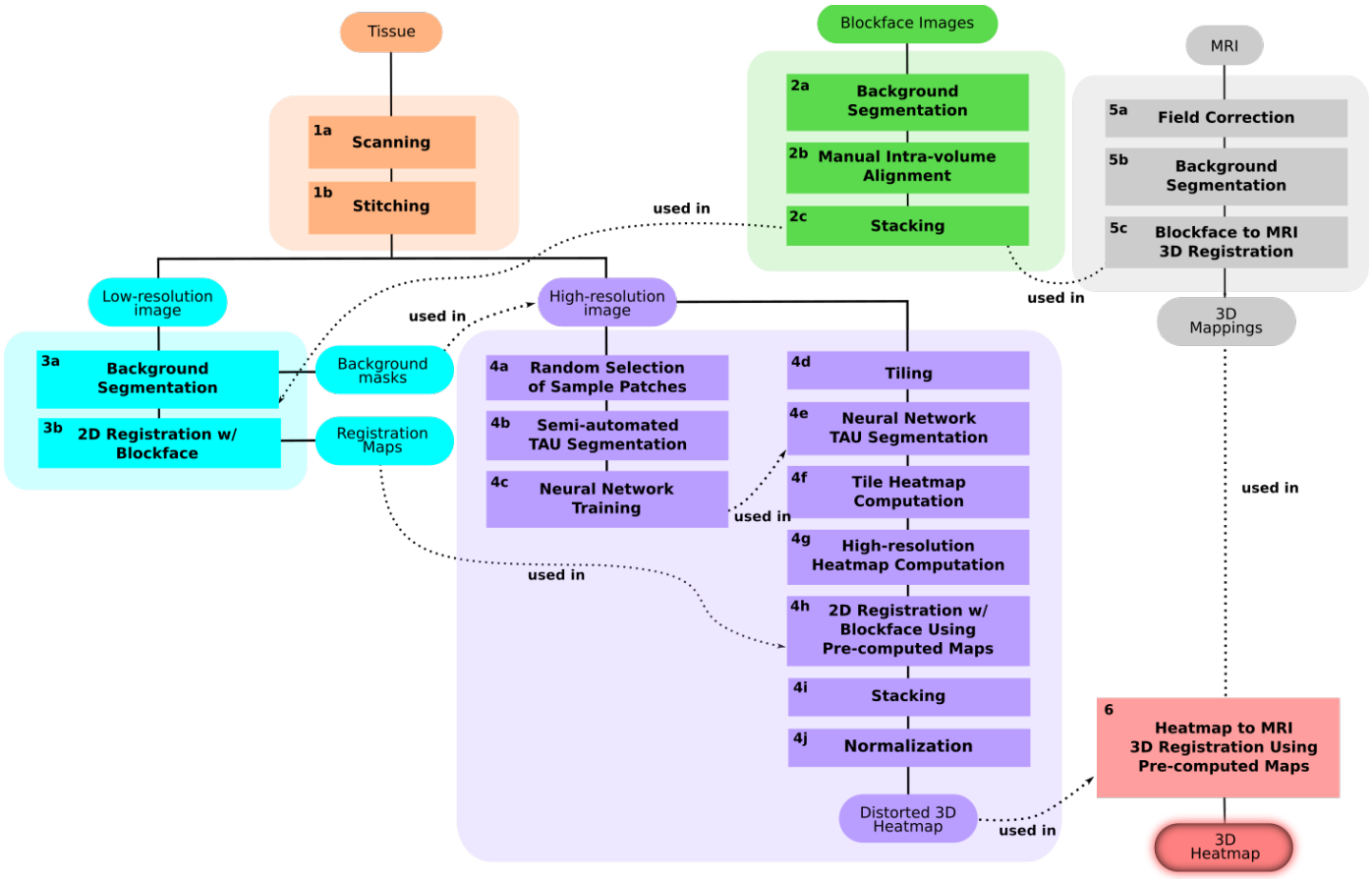

Figure 1: Tau tangle mapping pipeline. 1. Imaging module; 2. Blockface pre-processing module; 3. Low-resolution histology pre-processing module; 4. High-resolution histology segmentation module; 5. MRI pre-processing module; 6. 3d heatmap to MRI registration.

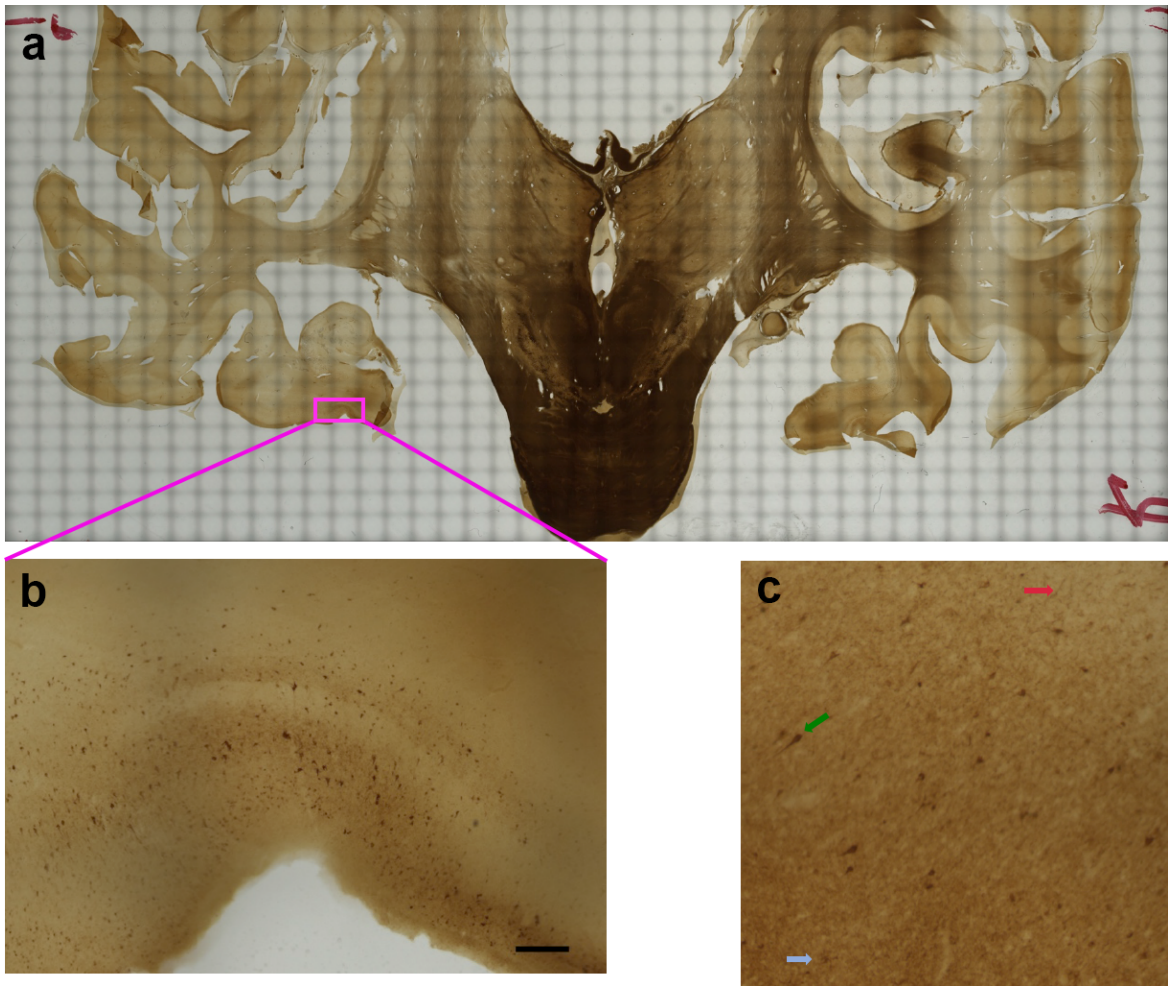

Figure 2: a) example of stitched image; b) 100% zoom of a hippocampus area (scalebar is 0.3mm); c) example of tau inclusions our system is capable of detecting. Green arrow is shows an NFT, red arrow shows a thread and blue arrow show as plaque.

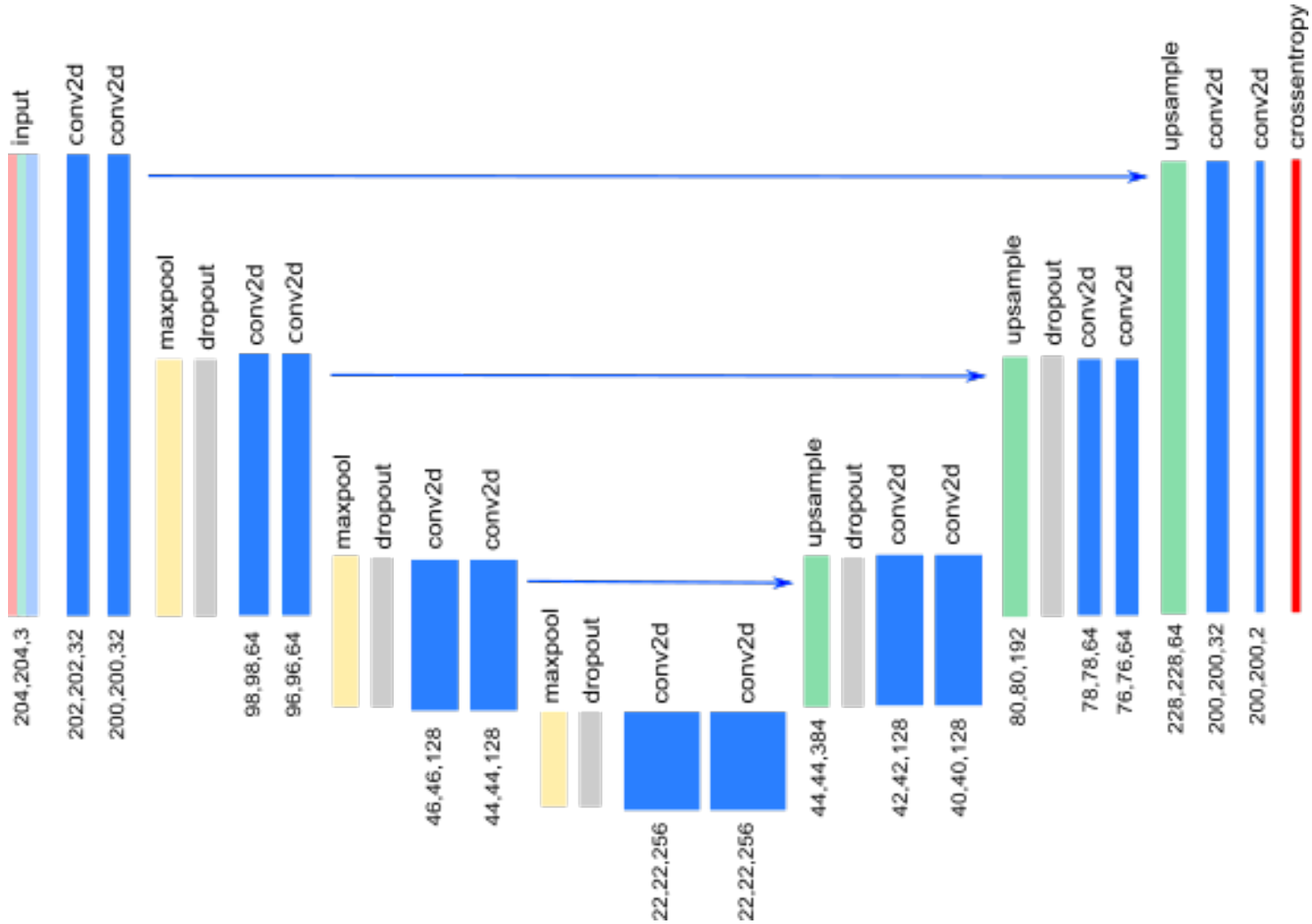

Figure 3: SlideNet architecture. The arrows indicate skip connections that are followed by cropped when the right-hand side layer is smaller than the left side layer.

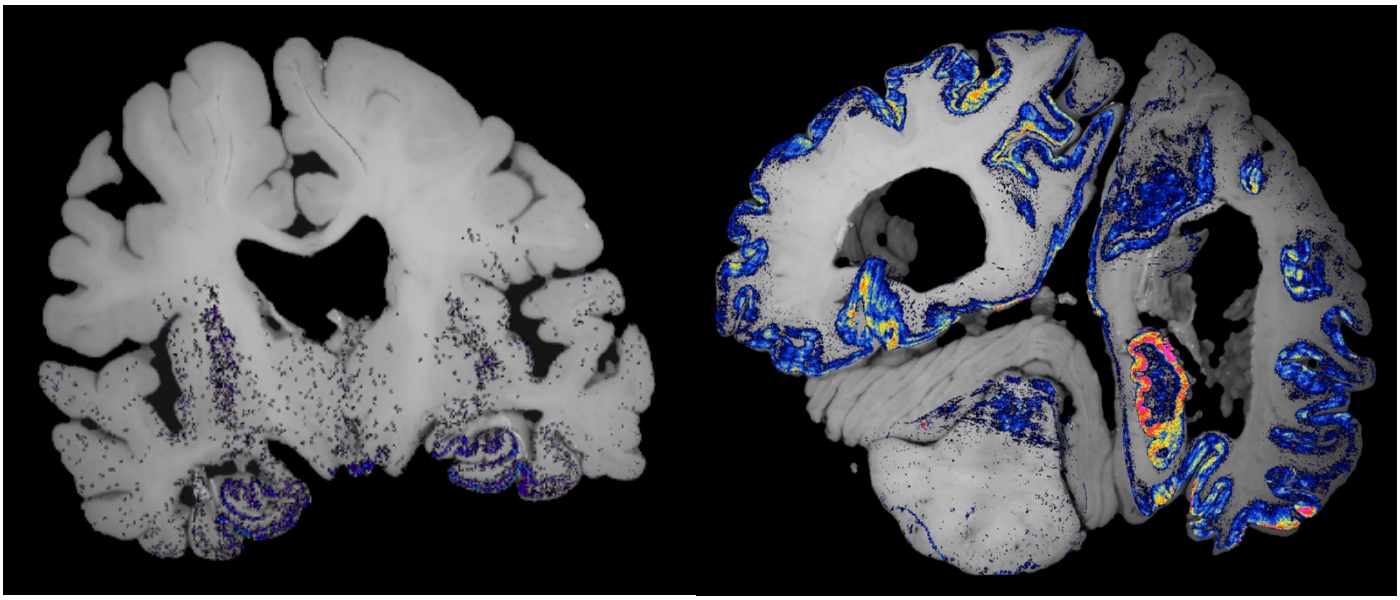

Figure 4: Examples of a heatmap from the early (left) and late (right) stage specimens. Here, warmer colors indicate higher density of tau tangles.

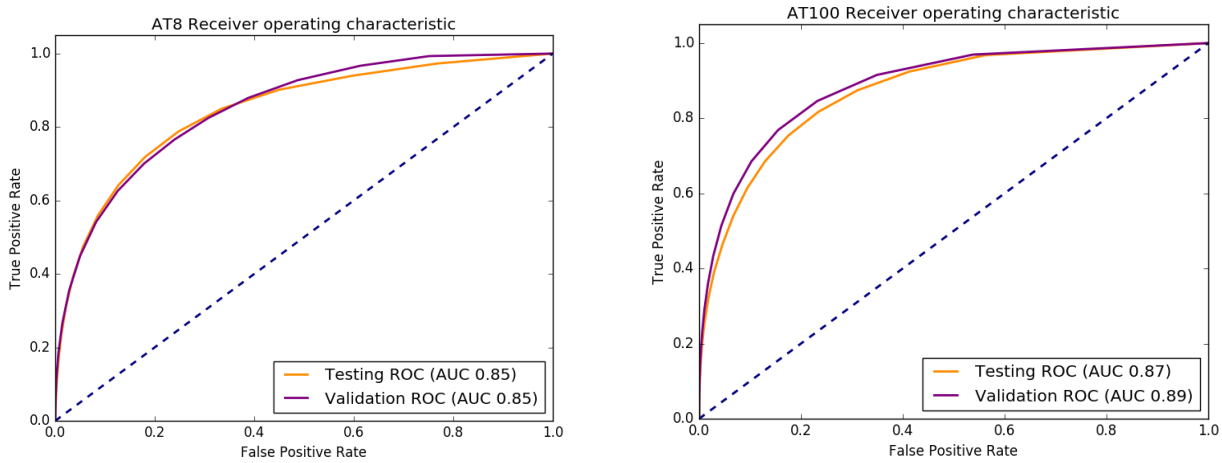

Figure 5: ROC curves

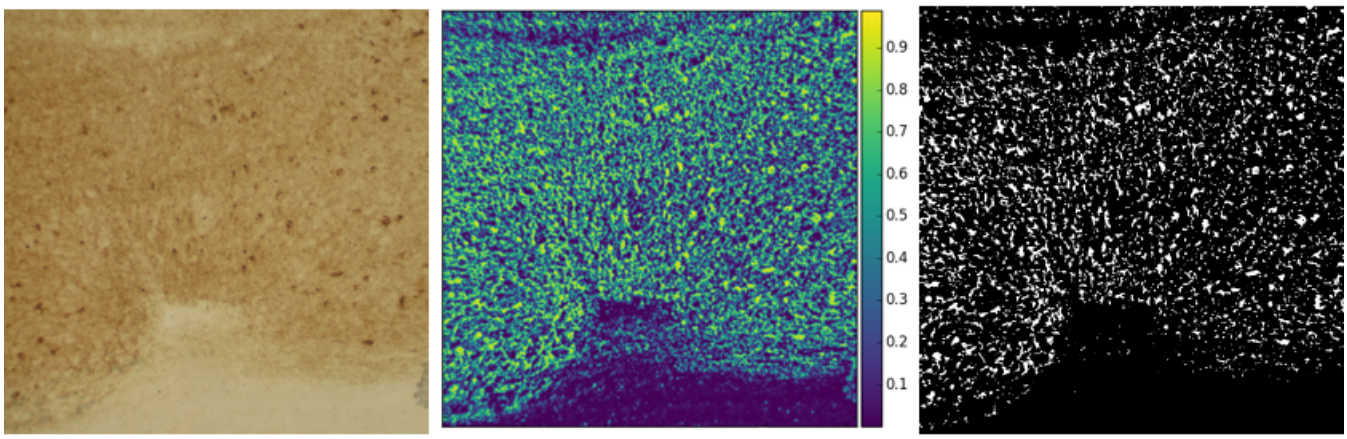

Figure 6: AT8 stained image patch (left) and its respective probability map (middle) and binary segmentation after thresholding at 0.7 (right)

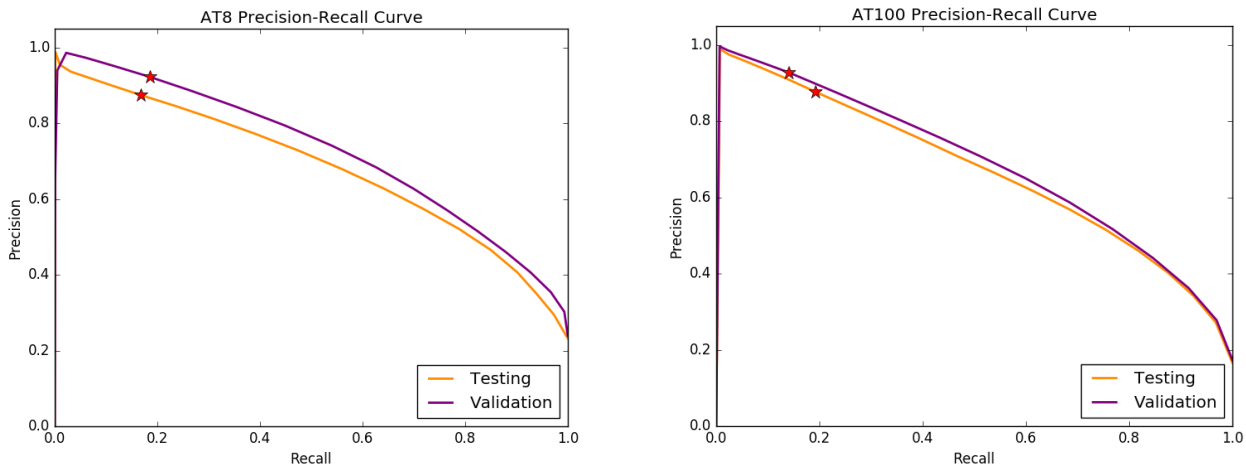

Figure 7: Precision-Recall curves for the AT8 and AT100 models. The red star indicates where is the 0.7 probability, which is used for thresholding the probability maps

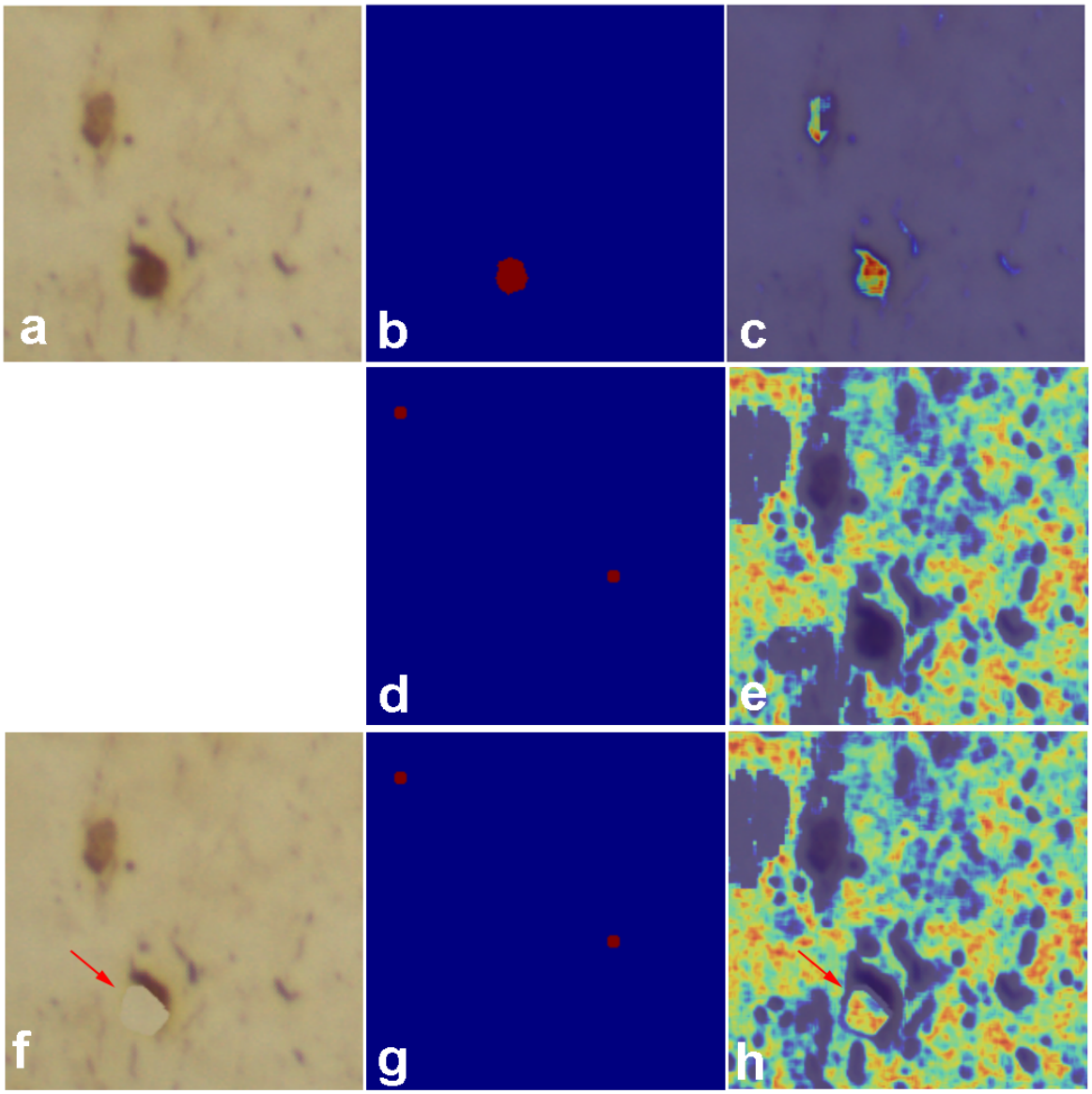

Figure 8: Network interpretability using Grad-CAM and perturbation techniques. a. Example of original tau tangle with some threads. b. Mask used for selecting pixels from (a). c. Grad-CAM result using pixels from mask in (b) as reference. d. mask used to locate background pixels. e. Grad-CAM result using background pixels as reference. f. Perturbed image, where a background image patch (indicated by the red arrow) was artificially placed to partially cover the tau inclusion. g. mask used for locating background pixel on perturbed image. h. Grad-CAM result using the perturbed image. The red arrow indicates the perturbed region that was correctly considered background.

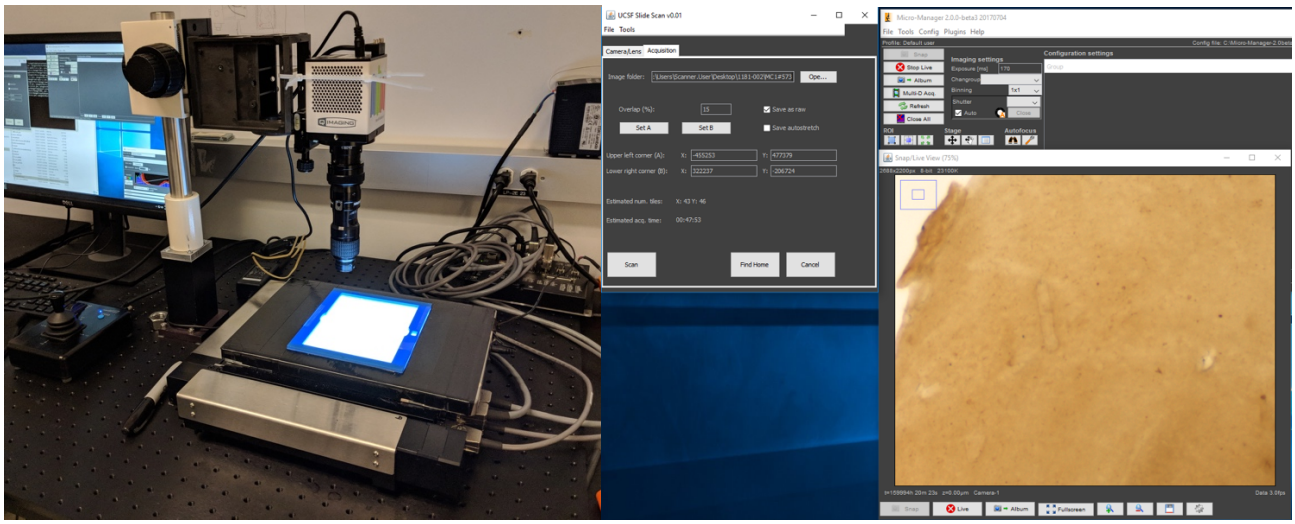

Figure 9: Left: our customized whole slide scanner. Right: User interface of the controller software developed using Micro Manager 2.0

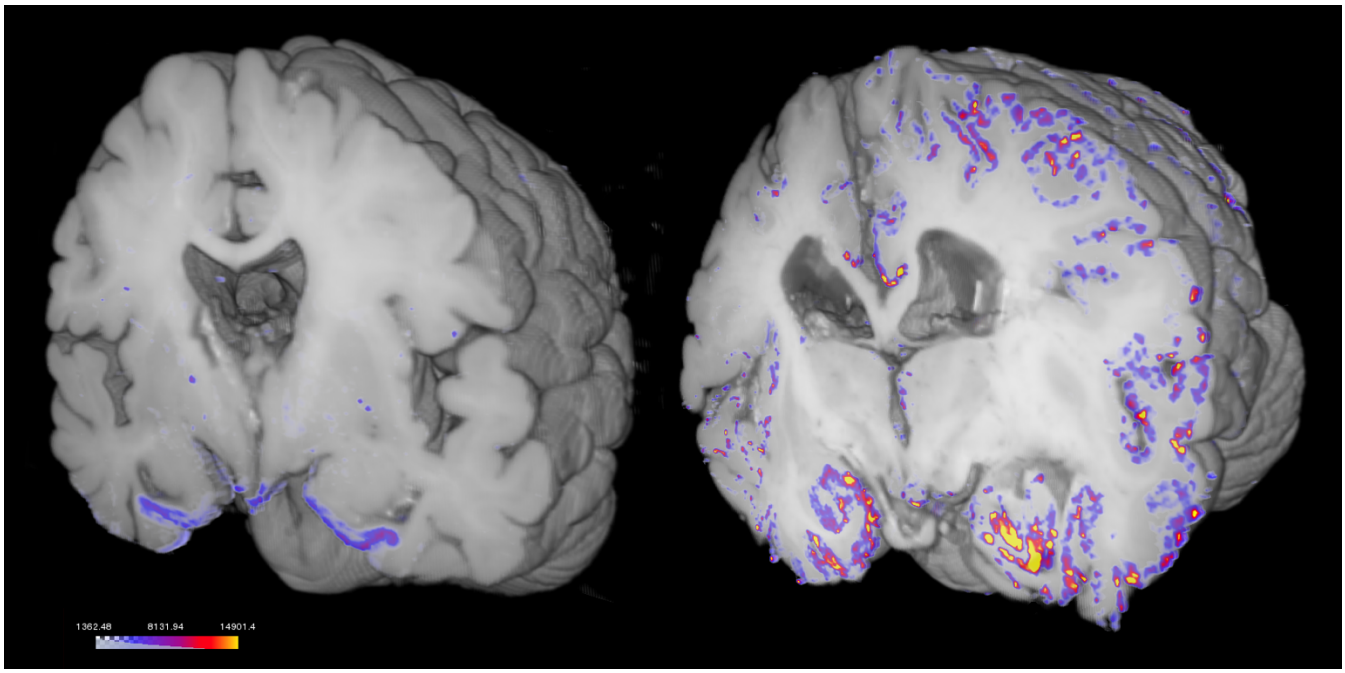

Figure 10: Volumetric reconstruction of the early (left) and late stage (right) AT100 tau heatmaps around the hippocampus area, overlaid on their respective blockface volumes

References

1. Montine, T. J. *et al.* National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. *Acta Neuropathol. (Berl.)* **123**, 1–11 (2012).
2. Association, Alzheimer's. Alzheimer's Disease Facts and Figures. (2018).
3. Jack, C. R. Alzheimer disease: new concepts on its neurobiology and the clinical role imaging will play. *Radiology* **263**, 344–361 (2012).
4. Grinberg, L. T. & Heinsen, H. Light at the beginning of the tunnel? Investigating early mechanistic changes in Alzheimer's disease. *Brain J. Neurol.* **140**, 2770–2773 (2017).
5. Simic, G. *et al.* Does Alzheimer's disease begin in the brainstem? *Neuropathol. Appl. Neurobiol.* **35**, 532–554 (2009).
6. Theofilas, P., Dunlop, S., Heinsen, H. & Grinberg, L. T. Turning on the Light Within: Subcortical Nuclei of the Isodentritic Core and their Role in Alzheimer's Disease Pathogenesis. *J. Alzheimers Dis. JAD* **46**, 17–34 (2015).
7. Theofilas, P. *et al.* Locus coeruleus volume and cell population changes during Alzheimer's disease progression: A stereological study in human postmortem brains with potential implication for early-stage biomarker discovery. *Alzheimers Dement. J. Alzheimers Assoc.* **13**, 236–246 (2017).
8. Theofilas, P. *et al.* Probing the correlation of neuronal loss, neurofibrillary tangles, and cell death markers across the Alzheimer's disease Braak stages: a quantitative study in humans. *Neurobiol. Aging* **61**, 1–12 (2018).
9. Ehrenberg, A. J. *et al.* Quantifying the accretion of hyperphosphorylated tau in the locus coeruleus and dorsal raphe nucleus: the pathological building blocks of early Alzheimer's disease. *Neuropathol. Appl. Neurobiol.* **43**, 393–408 (2017).
10. Bartzokis, G. Alzheimer's disease as homeostatic responses to age-related myelin breakdown. *Neurobiol. Aging* **32**, 1341–1371 (2011).

11. Smith, L. N. Cyclical Learning Rates for Training Neural Networks. *ArXiv150601186 Cs* (2015).
12. Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M. & Schmidhuber, J. Flexible, High Performance Convolutional Neural Networks for Image Classification. in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two* 1237–1242 (AAAI Press, 2011). doi:10.5591/978-1-57735-516-8/IJCAI11-210
13. Braak, H., Thal, D. R., Ghebremedhin, E. & Del Tredici, K. Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. *J. Neuropathol. Exp. Neurol.* **70**, 960–969 (2011).
14. Alegro, M. *et al.* Multimodal Whole Brain Registration: MRI and High Resolution Histology. in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* 634–642 (2016). doi:10.1109/CVPRW.2016.85
15. Theofilas, P. *et al.* A novel approach for integrative studies on neurodegenerative diseases in human brains. *J. Neurosci. Methods* **226**, 171–183 (2014).
16. Grinberg, L. T. & Heinsen, H. Computer-assisted 3D reconstruction of the human basal forebrain complex. *Dement. Neuropsychol.* **1**, 140–146 (2007).
17. Bria, A. & Iannello, G. TeraStitcher - a tool for fast automatic 3D-stitching of teravoxel-sized microscopy images. *BMC Bioinformatics* **13**, 316 (2012).