Funded by NIH National Institute of Mental Health (MH098786, MH060877, MH064168)

Measurement-oriented deep-learning workflow for improved segmentation of myelin and axons in high-resolution images of human cerebral white matter

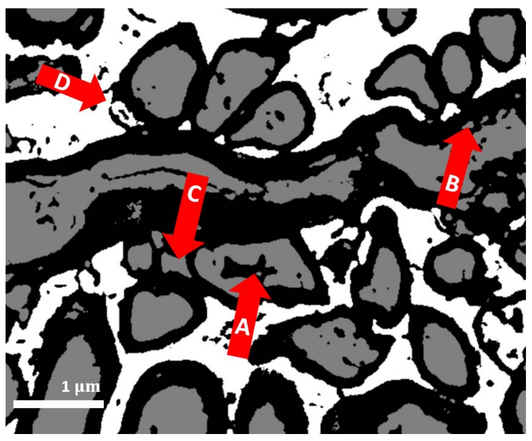
New York State Psychiatric Institute

P. Janjic[1*], K. Petrovski[1], B. Dolgoski[1], J. Smiley[2], P. Zdravkovski[3], G. Pavlovski[3], Z. Jakjovski[3], N. Davceva[3], V. Poposka[3], A. Stankov[3], G. Rosoklija[1,4,5], G. Petrushevska[3], L. Kocarev[1], A.J. Dwork[1,4,5,6]

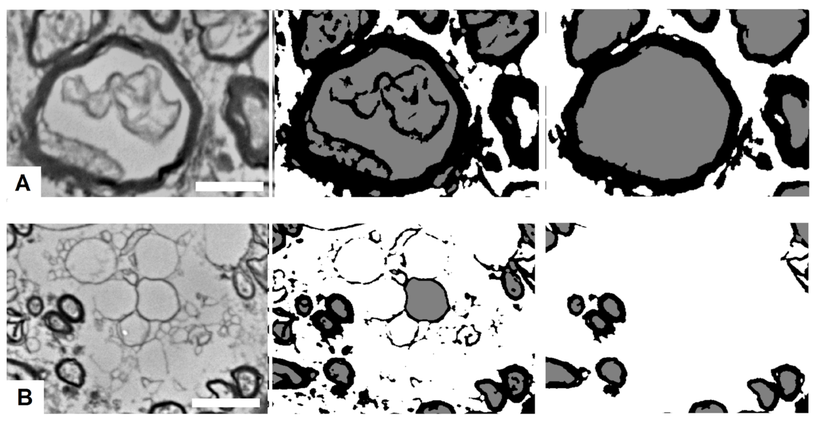
[1] Macedonian Academy of Sciences and Arts, Skopje, North Macedonia, [*]E-mail: predrag.a.janjic@gmail.com, [2] Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York, USA, [3] School of Medicine, Ss. Cyril and Methodius University, Skopje, North Macedonia, [4] Division of Molecular Imaging and Neuropathology, New York State Psychiatric Institute, New York, USA, [5]Dept. of Psychiatry, and [6] Dept. of Pathology and Cell Biology, Columbia University, New York

Segmentation of myelinated axons in high-contrast EM images poses challenges

Standard segmentation of high-contrast electron micrographs (EM) of human white matter accurately identifies myelin, but does not translate easily into measurements of individual axons and their myelin.



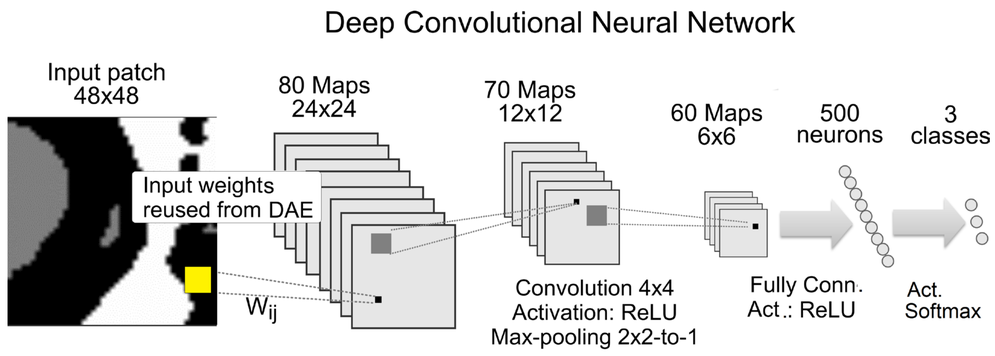
Typical misclassifications in 3-class segmentation (*Visiomorph)*, **(A)** segmenting nuclei or dark debris as myelin, **(B)** segmenting artificial gaps in myelin as axons, **(C)** segmenting regions surrounded by a ring of myelinated axons as axons, and **(D)** segmentation of axons as background when there is a defect in the surrounding myelin. Pixel classes: myelin (black), axon (gray), background (white)



Automated segmentation of pre-frontal white matter using *unsupervised* learning methods, can not overcome errors due to high contrast non-fiber structures and spurious objects, which they qualify as regular fiber ROIs.

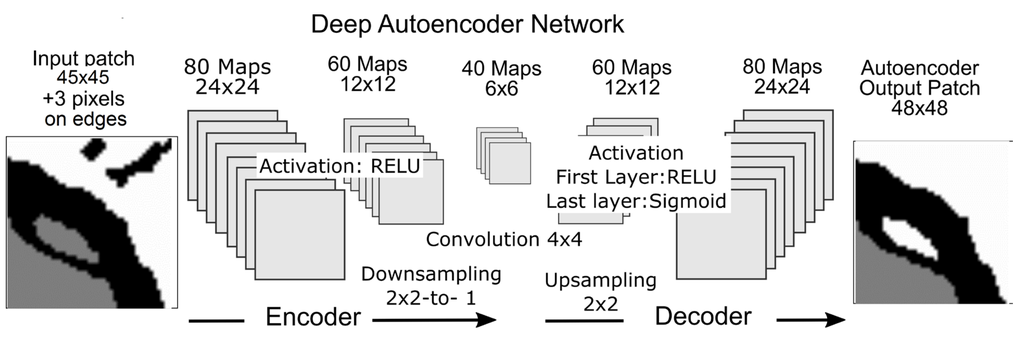
Most popular segmentation toolboxes like *Weka Segmenter* (*ImageJ*), *Visiomorph VIS,* or various add-ins for *Matlab,* based on classical ML feature extractors, saturate at 55% - 65% pixel accuracy on our images of autopsy samples of human frontal white matter. Scale bar = 1 micron.

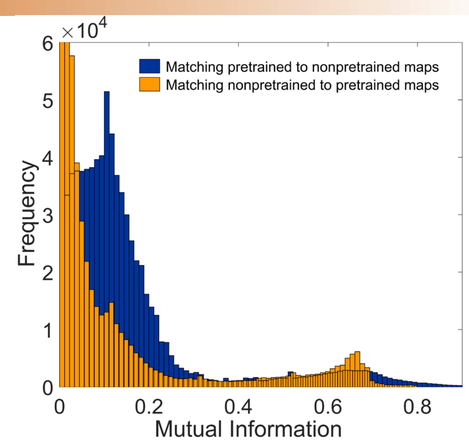
Convolutional networks learn to segment the central pixel of small input fragments



Deer Convolutional Neural Network (DNN) takes 6 to 8 million small input fragments (45 X 45 pixels, +3 pixels padding) from a set of 30 pre-segmented 2,048 x 2,048 pixel images to learn to classify the central pixel. The trained network is then used as a free classifier to revise the preliminary segmentation of images it has not seen.

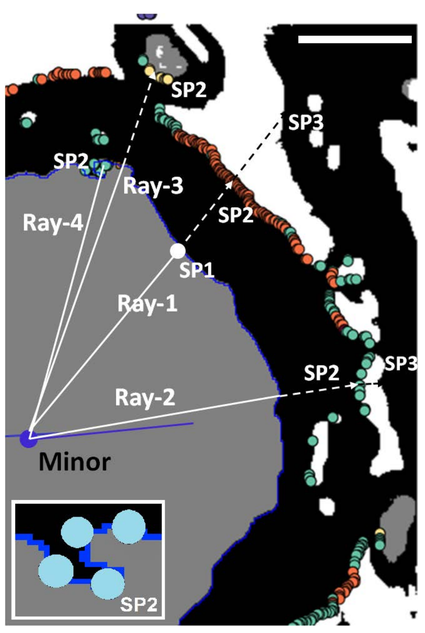
A separate DNN in *De-noising Auto-Encoder (DAE)* architecture (BELOW) is suggested for pre-training of the weights to the input layer of the DNN, instead of their random initialization. This stage improves the specificity of the feature-maps / layer, and hence their discrimination capacity.



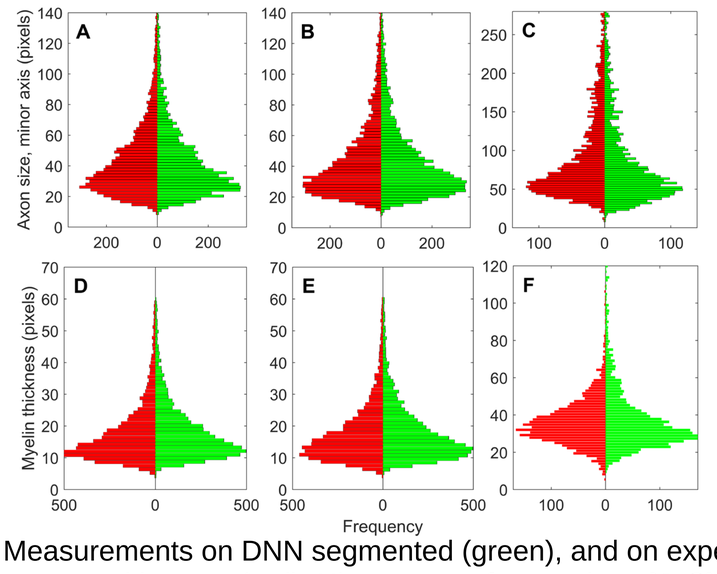


Histograms of maximal mutual information (M) between an activated map from either set to all maps in the other set, (Pre-trained, RND) or (RND, Pre-trained)

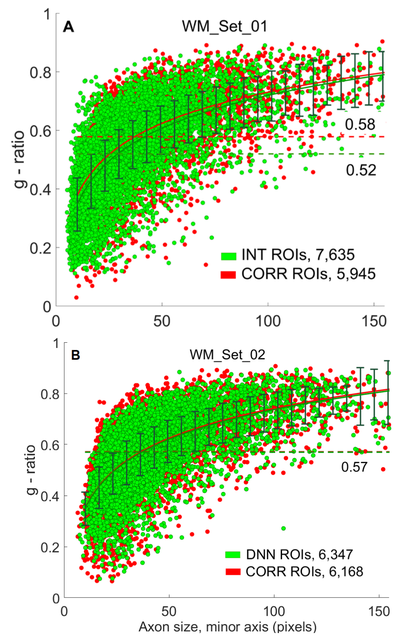
The measurements on DNN segmented and annotated images are interchangeable



Automated ray measurement tool (**RMT**) measures average axon diameter and myelin thickness along a ray (white line) from the center of the fiber ROI in DNN segmented image. Several stopping points (SP's) are used to verify the given ray meets *regular boundary point*, or to disqualify it. Together with additional conditions this step discards most spurious fiber ROIs prior to final averaging. The ROI boundary (blue line) is obtained from the *ROI Manager Tool* in *ImageJ*. Scale bar = 1 micron.



Measurements on DNN segmented (green), and on expert-annotated images (red). Histograms of axon sizes **(A-C)**, and myelin thickness **(D-F),** in two datasets from human white matter and one from the optic nerve. DNN was trained only by a subset of the first dataset.



Axonal g-ratio estimated from DNN**,** segmented (green) and, CORR, annotated images (red). **(A,B)** Scatter-plots of average g-ratio for each ROI. **(A)** Measurements on pre-segmented image set (standard ML tools) with ~30% spurious ROIs, and **(B)** on another dataset after DNN improved segmentation resulting in less than 3% mismatch in number of fiber ROIs. Dashed lines represent the mean over all ROIs.

Improved segmentation enables closing to a single pixel measurement precision

Dependent var. in different image sets		ICC(2,1) DNN vs. CORR	ICC(2,1) INT vs. CORR
WM_Set_01	axon	0.98	0.53
	myelin	0.92	0.90
	minor	0.98	0.45
	g-ratio	0.70	0.07
WM_Set_02	axon	0.97	0.54
	myelin	0.98	0.97
	minor	0.98	0.45
	g-ratio	0.96	0.30
ON_Set_01	axon	0.59	0.12
	myelin	0.96	0.88
	minor	0.47	0.32
	g-ratio	0.91	0.23

Intraclass correlation coefficients (ICC) between measured variables in classically segmented and DNN segmented datasets, with those measured on annotated datasets. For ICC's defined see Shrout PE, Fleiss, JL., *Psychological Bull* 1979, 86(2):420-8.

Statistical similarity of the histograms (above) was checked by the chi-squared distance $d_{xy} = \sum_{i=1}^{n}(x_i - y_i)^2/(x_i + y_i)$ where x_i, y_i represent the frequency counts of ROIs. For the corresponding three datasets it gave normalized values:
WM_Set_01: $d_{axon} = 0.85\%$, $d_{myelin} = 4.6\%$,
WM_Set_02: $d_{axon} = 0.56\%$, $d_{myelin} = 4.0\%$,
ON_Set_01: $d_{axon} = 1.78\%$, $d_{myelin} = 7.3\%$.

ROI number / Image set	White matter Set_01	White matter Set_02	Opt. nerve Set_01
ROIs at **INTERIM** (pre-segmented)	11,277	10,977	3,932
ROIs in **DNN segmented**	7,897	9,379	6,762
ROIs discarded in **RMT** by ray measurements	522	1,661	2,122
ROIs discarded in **RMT** by additional conditions	1,421	1,370	2,017
Final ROI number after RMT	5,954 (7,635)*	6,347 (7,851)*	2,623 (3,000)*
ROI number **CORR**	5,945	6,168	2,775