Domain-Agnostic Segmentation
of Thalamic Nuclei from Joint Structural
and Diffusion MRI

Henry F. J. Tregidgo[1(✉)], Sonja Soskic[1], Mark D. Olchanyi[2,3],
Juri Althonayan[1], Benjamin Billot[1,4], Chiara Maffei[5], Polina Golland[4],
Anastasia Yendiki[5], Daniel C. Alexander[1], Martina Bocchetta[6,7],
Jonathan D. Rohrer[7], and Juan Eugenio Iglesias[1,4,5]

[1] Centre for Medical Image Computing, UCL, London, UK
`h.tregidgo@ucl.ac.uk`
[2] Neurostatistics Research Laboratory, MIT, Cambridge, USA
[3] Neurotechnology and Neurorecovery, MGH and Harvard Medical School,
Boston, USA
[4] Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, USA
[5] Martinos Center for Biomedical Imaging, MGH and Harvard Medical School,
Boston, USA
[6] Centre for Cognitive and Clinical Neuroscience, Brunel University, London, UK
[7] Dementia Research Centre, UCL, London, UK

Abstract. The human thalamus is a subcortical brain structure that
comprises dozens of nuclei with different function and connectivity, which
are affected differently by disease. For this reason, there is growing inter-
est in studying the thalamic nuclei *in vivo* with MRI. Tools are available
to segment the thalamus from 1 mm T1 scans, but the image contrast is
too faint to produce reliable segmentations. Some tools have attempted
to refine these boundaries using diffusion MRI information, but do not
generalise well across diffusion MRI acquisitions. Here we present the first
CNN that can segment thalamic nuclei from T1 and diffusion data of any
resolution without retraining or fine tuning. Our method builds on our
histological atlas of the thalamic nuclei and silver standard segmentations
on high-quality diffusion data obtained with our recent Bayesian adaptive
segmentation tool. We combine these with an approximate degradation
model for fast domain randomisation during training. Our CNN produces
a segmentation at 0.7 mm isotropic resolution, irrespective of the resolu-
tion of the input. Moreover, it uses a parsimonious model of the diffusion
signal (fractional anisotropy and principal eigenvector) that is compatible
with virtually any set of directions and b-values, including huge amounts
of legacy data. We show results of our proposed method on three hetero-
geneous datasets acquired on dozens of different scanners. The method
is publicly available at freesurfer.net/fswiki/ThalamicNucleiDTI.

1 Introduction

The human thalamus is a brain region with connections to the whole
cortex [6,30]. It comprises dozens of nuclei that are involved in diverse

functions like cognition, memory, sensory, motor, consciousness, language, and others [14,30,32]. Crucially, these nuclei are differently affected by diseases such as Parkinson's [17], Alzheimer's [9,10], or frontotemporal dementia [40]. Such differentiation has sparked interest in studying the thalamic nuclei *in vivo* with MRI. This requires automated segmentation methods at the subregion level, as opposed to the whole thalamus provided by neuromaging packages like FreeSurfer [15] or FSL [27], or by convolutional neural networks (CNNs) like DeepNAT [41] or SynthSeg [7].

Different approaches have been used to segment thalamic nuclei. Some methods have attempted to register manually labelled histology to MRI [20,23,29], but accuracy is limited by the difficulty in registering two modalities with such different contrasts, resolutions, and artifacts. Diffusion MRI (dMRI) has been used to spatially cluster voxels into subregions, based on similarity in diffusion signal [5,25,31] or connectivity to cortical regions [6,21]. Clustering based on functional MRI connectivity has also been explored [42]. However, such clusters are not guaranteed to correspond to anatomically defined nuclei.

Other methods have relied on specialised MRI sequences to highlight the anatomical boundaries of the thalamus, typically at 7T [24,36] or with advanced dMRI acquisitions [4]. A popular method within this category is "THOMAS", a labelled dataset of 7T white-matter-nulled scans that has been used to segment the thalamic nuclei with multi-atlas segmentation [35] and CNNs [38]. Its disadvantage is requiring such advanced acquisitions at test time, which precludes its application to legacy data or at sites without the required expertise or resources.

One approach that supports training and test data of different modalities is Bayesian segmentation, which combines a probabilistic atlas (derived from one modality) with a likelihood model to compute adaptive segmentations on other modalities using Bayesian inference [3]. A probabilistic atlas of thalamic nuclei built from 3D reconstructed histology is available on FreeSurfer, along with a companion Bayesian segmentation method to segment the nuclei from structural scans [18]. We have recently released an improved version of this method that incorporates dMRI into the likelihood model [37], but it inherits the well-known problems of Bayesian segmentation with partial voluming (PV) [39]. While this tool works well with high-resolution dMRI data (like the Human Connectome Project, or HCP [33]), the lack of PV modelling is detrimental at resolutions much lower than ∼1 mm isotropic. This is the case of virtually every legacy dataset, and many modern datasets that use lower resolutions to keep acquisition time short or for consistency with older timepoints (e.g., ADNI [19]).

Finally, there are also supervised discriminative methods that label the thalamus from dMRI. An early approach by Stough et al. [34] used a random forest to segment the thalamus into six groups of nuclei. As features, they used local measures like fractional anisotropy (*FA*) and the principal eigenvector (*V1*), and connectivity with cortical regions. A recent approach [13] segmented the whole thalamus using the six unique elements of the diffusion tensor image (*DTI*) as inputs. While these supervised approaches can provide excellent performance on the training domain, they falter on datasets with different resolution.
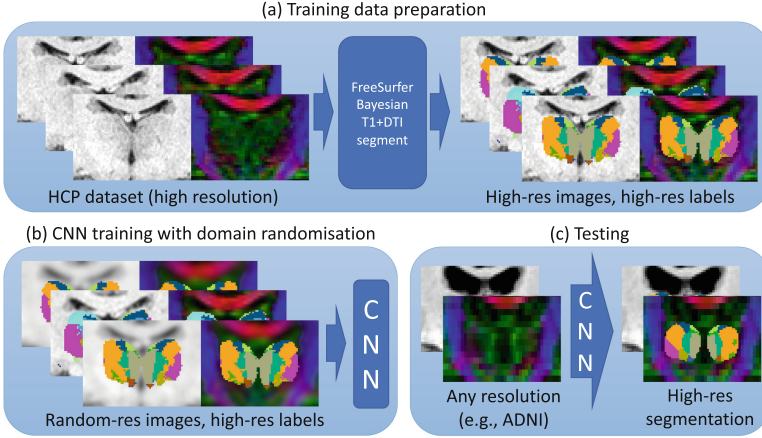
Fig. 1. Overview of the proposed method. (a) Generation of labelled training data. (b) Training with domain randomisation. (c) Testing. Images are in coronal view.

Here, we present the first CNN for joint dMRI and T1-weighted (*T1w*) images that can segment the thalamic nuclei without retraining or fine-tuning. We use domain randomisation to model resolution during training, which enables the CNN to produce super-resolved 0.7 mm isotropic segmentations, independently of the resolution of the input images. Aggressive data augmentation is used to ensure robustness against variations in contrast, shape and artifacts. Finally, our CNN uses a parsimonious representation of the dMRI data (FA+V1), which makes our publicly available tool compatible with virtually every dMRI dataset.

2   Methods

A summary of our method is shown in Fig. 1. We use our joint T1/DTI Bayesian method in FreeSurfer to segment the thalamic nuclei from a large number of modern, high-quality scans. These segmentations are used as silver standard to train a CNN, thus circumventing the need for manual segmentations. Our approach uses a hybrid domain randomisation and augmentation strategy that enables the network to generalise to virtually any diffusion dataset.

2.1   Training Dataset, Preprocessing, and Data Representation

To make the CNN compatible with legacy datasets, we choose a simple representation based on the FA and V1 of the DTI fit at each voxel. DTI only requires 7 measurements and is thus compatible with even the oldest datasets. As in many DTI visualisation tools, we combine the FA and V1 into a single 3×1 red-green-blue vector at every voxel. This RGB vector has brightness proportional to the FA, and its colour encodes the direction of V1 as shown in Fig. 1a.

To obtain accurate training segmentations from the Bayesian method in FreeSurfer [37] we require a high-resolution dataset with reduced PV artifacts. We choose the HCP, which includes 0.7 mm isotropic T1 and 1.25 mm isotropic dMRI with 90 directions and three b-values (1000, 2000, and 3000 s/mm^2). We use the HCP to generate our targets and then generate training images at a wide spectrum of resolutions by increasing the voxel size with a degradation model. We consider two RGB images per subject, derived from DTI fits of the b = 1000 and b = 2000 shells, respectively. For each of the two DTI fits, the Bayesian method yields three different sets of segmentations (corresponding to three available likelihood models). All six segmentations are defined on the .7 mm grid (Fig. 1a, right), and comprise 23 thalamic nuclei per hemisphere (46 total) [37]. We exclude the Pc, Pt and VM nuclei provided by the Bayesian segmentation as they are not labelled in every training example due to their small volumes (2–3 mm^3).

2.2   Domain Randomisation and Data Augmentation

We employ domain randomisation and aggressive data augmentation for both our T1 and diffusion data in order to model: *(i)* the degradation in quality from HCP to more standard acquisition protocols, and *(ii)* the variability in appearance due to differences in acquisitions and scanners at test time.

Domain Randomisation for Resolution: at the crux of our method is the domain randomisation of input resolutions. At every iteration, we randomly sample the voxel dimensions for the T1 and DTI (independently) in two steps. First, we sample a "coarse" scalar voxel size from a uniform distribution between 1 and 3 mm. Then, we sample the voxel side length in each direction from a normal distribution centred on this "coarse" mean with $\sigma = 0.2$ mm.

Next, we resample the T1 and RGB channels to the sampled resolution. For the T1, we use a publicly available PV-aware degradation model [8], which accounts for variability in slice thickness and slice spacing. For the RGB, one should theoretically downsample the original diffusion-weighted images, and recompute the DTI at the target resolution. However, the exact characteristics of the blurring depend on the set of directions and b-values, which will not be the same for the training and test datasets. Moreover, recomputing the DTI is too slow for on-the-fly augmentation. Instead, we apply the degradation model to the RGB image directly, which can be done very efficiently. This is only an approximation to the actual degradation, but in practice, the domain randomisation strategy minimises the effects of the domain gap created by the approximation.

Data Augmentations: we also apply a number of geometric and intensity augmentations, some standard, and some specific to our dMRI representation.

– *Global geometric augmentation:* we use random uniform scaling (between 0.85 and 1.15) and rotations about each axis (between −15 and 15°). Rotations are applied to the images and also used to reorient the V1 vectors.

- *Local geometric augmentation:* we deform the scans with a piecewise linear deformation field, obtained by linear interpolation on a $5 \times 5 \times 5 \times 3$ grid. V1 vectors are reoriented with the PPD method ("preservation of principle direction" [1]).
- *Local orientation augmentation:* we generate a smooth grid of random rotations between $[-15°, 15°]$ around each axis using piecewise linear interpolation on a $5 \times 5 \times 5 \times 3$ grid. These are applied to V1 to simulate noise in principle direction
- *DTI "speckles":* To account for infeasible FA and V1 voxels generated by potentially unconstrained DTI fitting, we select random voxels in the low resolution images (with probability $p = 1 \times 10^{-4}$), randomise their RGB values, and renormalise them so that their effective FA is between 0.5 and 1.
- *Noise, brightness, contrast, and gamma:* we apply random Gaussian noise to both the T1 and FA; randomly stretch the contrast and modify the brightness of the T1; and apply a random gamma transform to the T1 and FA volumes.

These augmentations are applied to the downsampled images. After that, the augmented images are upscaled back to 0.7 mm isotropic resolution. This ensures that all the channels (including the target segmentations) are defined on the same grid, independently of the intrinsic resolution of the inputs (Fig. 1b). At test time, this enables us to produce .7 mm segmentations for scans of any resolution (prior upscaling to the .7 mm grid).

2.3   Loss

We build on the soft dice loss [26]. Since some labels in the atlas are very small, we implemented a grouped soft dice by combining nuclei into 10 larger groupings [37]. We then combined this Dice with the average Dice of the individual nuclei and the Dice of the whole thalamus into the following composite loss:

$$\mathcal{L} = -\sum_l SDC(X_l, Y_l) - \sum_g SDC\left(\{X_l\}_{G_g}, \{Y_l\}_{G_g}\right) - SDC\left(\{X_l\}_{l \neq 0}, \{Y_l\}_{l \neq 0}\right),$$

(1)

where, $X_l = \{x_i^l\}$ and $Y_l = \{y_i^l\}$ are the predicted and ground truth probability maps for label $l \in [0, \ldots, L]$; G_g is the set of label indices in nuclear group $g \in [1, \ldots, 10]$, label $l = 0$ corresponds to the background and SDC is the soft Dice coefficient: $SDC(X, Y) = (2 \times \sum_i x_i y_i)/(\sum_i x_i^2 + \sum_i y_i^2)$.

2.4   Architecture and Implementation Details

Our CNN is a 3D U-net [11,28] with 5 levels (2 layers each), $3 \times 3 \times 3$ kernels and ELU activations [12]. The first level has 24 features, and every level has twice as many features as the previous one. The last layer has a softmax activation. The loss in Eq. 1 was optimised for 200,000 iterations with Adam [22]. A random crop of size $128 \times 128 \times 128$ voxels (guaranteed to contain the thalami) was used at every iteration. The T1 scans are normalised by scaling the median white matter
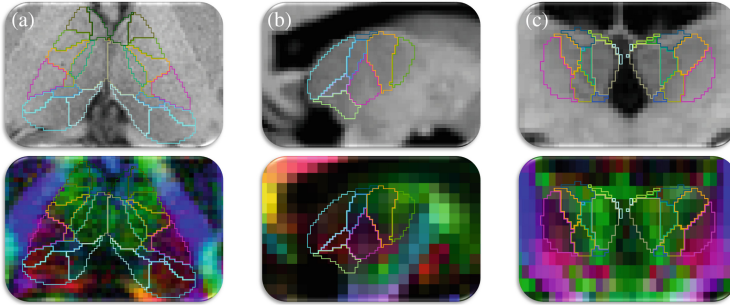
Fig. 2. Example segmentations from: (a) HCP (axial); (b) LOCAL (sagittal); and (c) ADNI (coronal). Top: T1. Bottom: RGB encoding. The CNN produces accurate segmentations at high isotropic resolution despite the heterogeneous acquisitions.

intensity to 0.75 and clipping at [0, 1]. The DTI volumes are upsampled to the 0.7 mm space of the T1s (using the log domain [2]) prior to the RGB computation. To generate a training target we combine all six segmentation candidates (three likelihood models times two shells) in a two step process: *(i)* averaging the one-hot encodings of each segmentation, and *(ii)* coarsely segmenting into 10 label groups and renormalising the soft target.

For validation purposes, we used the Bayesian segmentations of 50 withheld HCP subjects and 14 withheld ADNI subjects. Even though the Bayesian segmentation of ADNI data is not reliable enough to be used as ground truth for evaluation (due to the PV problems described in the Introduction), it is still informative for validation purposes – particularly when combined with HCP data. The final model for each network is chosen based on the validation loss averaged between the HCP and ADNI validation sets.

3   Experiments and Results

3.1   MRI Data

We trained on 600 HCP subjects (see Sect. 2.1). For evaluation, we used:

HCP: 10 randomly selected subjects (not overlapping with the training data), with manual segmentations of 10 groups of labels (the same as in Sect. 2.3).

LOCAL: 21 healthy subjects (9 males, ages 53–80), each with a 1.1 mm isotropic T1 and a test-retest pair of 2.5 mm isotropic dMRI (64 directions, b=1,000).

ADNI: 90 subjects from the ADNI, 45 with with Alzheimer's disease (AD) and 45 healthy controls (73.8±7.7 years; 44 females), each with a T1w ($1.2\times1\times1$ mm, sagittal) and dMRI ($1.35\times1.35\times2.7$ mm, axial, 41 directions, b=1,000) scan.

3.2   Competing Methods and Ablations

To the best of our knowledge, the only available tool that can segment T1/dMRI of any resolution is Freesurfer [37]. We therefore compare our method with this

Table 1. Mean Dice for ground truth comparison (left columns) and test re-test (right columns). Dice is shown for labels grouped into: histological labels ("hist", 23 labels), manual protocol ("manual", 10 labels), nuclear groups [37] ("nuclear", 5 labels), and whole thalamus. CNNs are sorted in descending order of average Dice across columns.

Model	Dice with manual tracing:			Test re-test Dice score:			
	Manual	Nuclear	Whole	Hist	Manual	Nuclear	Whole
Bayesian	0.547	0.639	0.885	0.716	0.826	0.880	0.946
CNN No defm.	0.585	0.678	0.895	0.824	0.885	0.917	0.963
CNN single target	0.582	0.676	0.905	0.824	0.883	0.911	0.962
CNN full model	0.576	0.674	0.903	0.819	0.876	0.910	0.961
CNN majority vote	0.579	0.676	0.897	0.816	0.875	0.909	0.961
CNN no rotation	0.580	0.672	0.902	0.808	0.873	0.909	0.961
CNN no speckle	0.583	0.680	0.903	0.799	0.869	0.907	0.958
CNN average target	0.575	0.669	0.898	0.799	0.867	0.906	0.959
CNN Dice Loss	0.544	0.618	0.806	0.824	0.867	0.899	0.963

algorithm, along with seven ablations of model options: using only the Dice loss; three variations on the way of merging the candidate Bayesian segmentations into a training target (average one-hot; majority voting; and selecting a segmentation at random); and three ablations on the augmentation (omitting the "speckle" DTI voxels, the random V1 rotations, and the piecewise linear deformation).

3.3   Results

Qualitative results are shown in Fig. 2. Our CNN successfully segments all scans at 0.7 mm resolution, despite the different input voxel sizes. Quantitative results are presented below for three experimental setups, one with each dataset.

Direct Evaluation with Manual Ground truth Using HCP: We first evaluated all competing methods and ablations on the 10 manually labelled subjects. Table 1 (left columns) shows mean Dice scores at different levels of granularity. Thanks to the ground truth aggregation, domain randomisation and aggressive augmentation, most of the CNNs produce higher accuracy than the Bayesian method at every level of detail – despite having been trained on automated segmentations from the Bayesian tool. The only ablation with noticeable lower performance is the one using the Dice of only the fine histological labels (i.e., no Dice of groupings), which highlights the importance of our composite Dice loss.

Test-Retest Using LOCAL: Table 1 (right columns) shows Dice scores between the segmentations of the two available sets of images for the LOCAL dataset, for different levels of granularity. All the networks are more stable than the Bayesian method, with considerably higher test-retest dice scores.

Best-Performing CNN: Considering Table 1 as a whole, the CNN with the highest mean Dice is the one without local geometric augmentation. We hypothesise that this is because the benefit of this augmentation is negligible due to
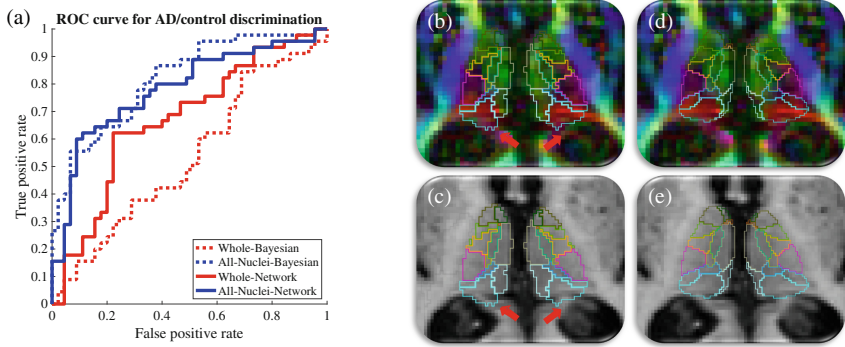
Fig. 3. Comparison of Bayesian and CNN segmentation on ADNI. (a) ROC curves for AD discrimination. (b) Axial view of RGB encoding of a subject with Bayesian segmentation overlaid. (c) Corresponding view of T1w scan. (d–e) Corresponding plane with CNN segmentation. Red arrows point at oversegmentation by the Bayesian method. (Color figure online)

Table 2. Area under the curve and accuracy at elbow for AD discrimination.

Method	Whole thal AUC	Whole thal acc. at elbow	Nuclei AUC	Nuclei acc. at elbow
Bayesian	53.3	57.8	81.9	74.4
CNN	67.1	70.0	78.3	75.6

the large number of training cases, and thus does not compensate for the loss of performance due to the approximations that are required to augment on the fly.

Group Study Using ADNI: We segmented the ADNI subjects with the best-performing CNN, and computed volumes of the thalamic nuclei normalised by the intracranial volume (estimated with FreeSurfer). We then computed ROC curves for AD discrimination using a threshold on: *(i)* the whole thalamic volume; and *(ii)* the likelihood ratio given by a linear discriminant analysis (LDA, [16]) on the volumes of the 23 nuclei (left-right averaged). The ROC curves are shown in Fig. 3(a). The area under the curve (AUC) and accuracy at the elbow are shown in Table 2. The LDA from the CNN and Bayesian methods show no significant difference in overall discriminative power. However, the atrophy detected by the CNN shows greater significance with the AV and VA reaching significance after correction for multiple comparisons ($p < 0.002$ Wilcoxon rank-sum). Additionally, there is a significant increase in the discriminative power of the whole thalamus from the CNN compared to the Bayesian method ($p < 0.005$ paired DeLong test). This indicates the external boundary of our method may be more useful than that provided by the Bayesian method and often corresponds to a reduction of oversegmentation into the pulvinar (Figs. 3b–e).

4 Discussion and Conclusion

We have presented the first method that can segment the thalamic nuclei from T1 and dMRI data obtained with virtually any acquisition, solving the problems posed by PV to Bayesian methods. Using Bayesian segmentations generated from multiple diffusion models while applying hybrid domain randomisation and augmentation methods, we remarkably improve upon both the accuracy and reliability of our source segmentations. Our tool is robust against misregistration from geometric distortion, which is generally more problematic in frontal and occipital regions. Nuclei volumes resulting from the tool show similar discriminative power to those provided by the Bayesian tool, while improving the utility of whole thalamus measurements and increasing segmentation resolution. Crucially, our use of the FA and V1 representation of dMRI data as input means that our tool is compatible with virtually every dMRI dataset. Publicly sharing this ready-to-use tool as part of FreeSurfer 7.4 will enable neuroimaging studies of the thalamic nuclei without requiring any expertise in neuroanatomy or machine learning, and without any specialised computational resources.

References

1. Alexander, D., Pierpaoli, C., Basser, P., Gee, J.: Spatial transformations of diffusion tensor magnetic resonance images. IEEE Trans. Med. Imaging **20**(11), 1131–1139 (2001)
2. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magn. Reson. Med. **56**(2), 411–421 (2006)
3. Ashburner, J., Friston, K.: Unified segmentation. Neuroimage **26**(3), 839–51 (2005)
4. Basile, G., Bertino, S., Bramanti, A., Ciurleo, R., et al.: In vivo super-resolution track-density imaging for thalamic nuclei identification. Cereb Cortex **31**, 5613–36 (2021)
5. Battistella, G., Najdenovska, E., Maeder, P., Ghazaleh, N., et al.: Robust thalamic nuclei segmentation method based on local diffusion magnetic resonance properties. Brain Struct. Funct. **222**(5), 2203–16 (2017)
6. Behrens, T.E., Johansen-Berg, H., Woolrich, M., Smith, S., et al.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. **6**(7), 750–57 (2003)
7. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., et al.: SynthSeg: segmentation of brain MRI scans of any contrast and resolution without retraining. Med Image Anal. **86**, 102789 (2023)
8. Billot, B., Robinson, E., Dalca, A.V., Iglesias, J.E.: Partial volume segmentation of brain MRI scans of any resolution and contrast. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12267, pp. 177–187. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59728-3_18

9. Braak, H., Braak, E.: Alzheimer's disease affects limbic nuclei of the thalamus. Acta Neuropathol. **81**(3), 261–268 (1991)

10. Braak, H., Braak, E.: Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. **82**(4), 239–259 (1991)

11. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49

12. Clevert, D.A., Unterthiner, T., Hochreiter, S.: Fast and accurate deep network learning by exponential linear units. arXiv preprint arXiv:1511.07289 (2015)

13. Ewert, C., Kügler, D., Yendiki, A., Reuter, M.: Learning anatomical segmentations for tractography from diffusion MRI. In: Computing dMRI Workshop 2020, pp. 81–93

14. Fama, R., Sullivan, E.V.: Thalamic structures and associated cognitive functions: relations with age and aging. Neurosci. Biobehav. R **54**, 29–37 (2015)

15. Fischl, B., Salat, D.H., Busa, E., Albert, M., et al.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron **33**(3), 341–355 (2002)

16. Fisher, R.A.: The use of multiple measurements in taxonomic problems. Ann. Eugenic **7**(2), 179–88 (1936)

17. Henderson, J.M., Carpenter, K., Cartwright, H., Halliday, G.M.: Loss of thalamic intralaminar nuclei in progressive supranuclear palsy and Parkinson's disease: clinical and therapeutic implications. Brain **123**(7), 1410–21 (2000)

18. Iglesias, J.E., Insausti, R., Lerma-Usabiaga, G., Bocchetta, M., et al.: A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. Neuroimage **183**, 314–26 (2018)

19. Jack, C.R., Jr., Bernstein, M.A., Fox, N.C., Thompson, P., et al.: The Alzheimer's disease neuroimaging initiative: MRI methods. J. Magn. Reson. Imaging **27**, 685–91 (2008)

20. Jakab, A., Blanc, R., Berényi, E.L., Székely, G.: Generation of individualized thalamus target maps by using statistical shape models and thalamocortical tractography. Am. J. Neuroradiol. **33**(11), 2110 (2012)

21. Johansen-Berg, H., Behrens, T., Sillery, E., Ciccarelli, O., et al.: Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. Cereb. Cortex **15**(1), 31–39 (2005)

22. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. arXiv:1412.6980 (2014)

23. Krauth, A., Blanc, R., Poveda, A., Jeanmonod, D., et al.: A mean three-dimensional atlas of the human thalamus: generation from multiple histological data. Neuroimage **49**(3), 2053–62 (2010)

24. Liu, Y., D'Haese, P.F., Newton, A.T., Dawant, B.M.: Generation of human thalamus atlases from 7 T data and application to intrathalamic nuclei segmentation in clinical 3 T T1-weighted images. Magn. Reson. Med. **65**, 114–128 (2020)

25. Mang, S., Busza, A., Reiterer, S., Grodd, W.: Klose: Thalamus segmentation based on the local diffusion direction: a group study. Magn. Reson. Med. **67**, 118–26 (2012)

26. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 3DV Conference 2016, pp. 565–571 (2016)