



Learning Ontology-Based Hierarchical Structural Relationship for Whole Brain Segmentation

Junyan Lyu^{1,2}, Pengxiao Xu¹, Fatima Nasrallah², and Xiaoying Tang¹(✉)

¹ Department of Electronic and Electrical Engineering, Southern University of Science and Technology, Shenzhen, China
tangxy@sustech.edu.cn

² Queensland Brain Institute, The University of Queensland, Brisbane, Australia

Abstract. Whole brain segmentation is vital for a variety of anatomical investigations in brain development, aging, and degradation. It is nevertheless challenging to accurately segment fine-grained brain structures due to the low soft-tissue contrast. In this work, we propose and validate a novel method for whole brain segmentation. By learning ontology-based hierarchical structural knowledge with a triplet loss enhanced by graph-based dynamic violate margin, our method can mimic experts' hierarchical perception of the brain anatomy and capture the relationship across different structures. We evaluate the whole brain segmentation performance of our method on two publicly-accessible datasets, namely JHU Adult Atlas and CANDI, respectively possessing fine-grained (282) and coarse-grained (32) manual labels. Our method achieves mean Dice similarity coefficients of 83.67% and 88.23% on the two datasets. Quantitative and qualitative results identify the superiority of the proposed method over representative state-of-the-art whole brain segmentation approaches. The code is available at <https://github.com/CRazorback/OHSR>.

Keywords: Whole brain segmentation · Hierarchical ontology · U-Net · Triplet loss · Structural MRI

1 Introduction

Quantitative analysis of structural MRI of the human brain is essential in various anatomical investigations in brain development, aging, and degradation. Accurate segmentation of brain structures is a prerequisite for quantitative and particularly morphometric analysis [5, 8, 23]. However, whole brain segmentation is challenging due to the low soft-tissue contrast, the high anatomical variability, and the limited labeled data, especially for fine-grained structures.

During the past decade, MRI based whole brain segmentation approaches have been explored. Multi-atlas based methods [19, 24] are shown to be simple

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43901-8_37.

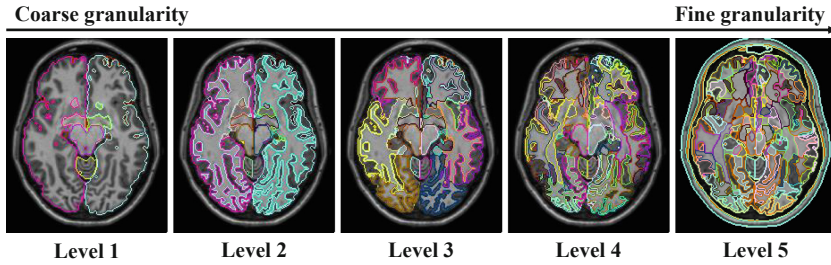


Fig. 1. Illustration of the five-level ontology-based hierarchical structural relationship. The labels are created in a coarse-to-fine manner.

yet effective and serve as the de facto standard whole brain segmentation methods. With a limited number of labeled data, multi-atlas based methods propagate labels from the atlas images to the target images through registration (mostly non-linear or even diffeomorphic registration), joint label fusion, and possibly corrective learning. With diffeomorphic registration, multi-atlas based methods are empowered of enhanced segmentation accuracy and topology-preserving capabilities that well accommodate the potential need of incorporating shape prior [4]. Nevertheless, multi-atlas based methods suffer from low computational efficiency and fail to leverage sophisticated contextual information when the tissue contrast in the structural MRI data is low.

Recently, deep learning has demonstrated state-of-the-art (SOTA) performance on various medical image segmentation tasks [6, 12, 16], also serving as an alternative solution to the whole brain segmentation task. QuickNAT [21] trains three 2D U-Nets respectively on the axial, coronal and sagittal views, and then aggregates them to infer the final segmentation. Pre-training with auxiliary labels derived from Freesurfer is conducted to alleviate QuickNAT's reliance on manually annotated data. SLANT [11] applies multiple independent 3D U-Nets to segment brain structures in overlapped MNI subspace, followed by label fusion. Wu et al. [26] propose a multi-atlas and diffeomorphism based encoding block to determine the most similar atlas patches to a target patch and propagate them into 3D fully convolutional networks. These deep learning methods improve over conventional multi-atlas based methods by considerable margins in terms of both computational efficiency and segmentation accuracy.

However, existing methods tend to ignore the ontology-based hierarchical structural relationship (OHSR) of the human brain's anatomy. Most of them assume all brain structures are disjoint and use multiple U-Nets to separately perform voxel-wise predictions for each of multiple structures of interest. It has been suggested that neuroanatomical experts recognize and delineate brain anatomy in a coarse-to-fine manner (Fig. 1) [17, 18]. Concretely, at the highest level of the most coarse granularity (Level 1 in Fig. 1), the brain can be simply decomposed into telencephalon, diencephalon, mesencephalon, metencephalon, and myelencephalon. At a lower level, the cerebral cortex, cerebral nuclei and white matter can be identified within the telencephalon. At the lowest level of the

most fine granularity, namely Level 5 in Fig. 1, the most fine-grained structures such as the hippocampus and the amygdala in the limbic system are characterized. It shall be most desirable if a neural network can learn brain anatomy in a similar fashion. Inspired by OHSR of brain anatomy, in the manifold space of a neural network, the distance between the feature vectors of the hippocampus and the amygdala should be smaller than that between the feature vectors of the hippocampus and any other structure that does not also belong to the limbic system. In other words, embeddings of fine-grained structures labeled as a same class at a higher level are supposed to be more similar than those of fine-grained structures labeled as different classes at the same higher level. Such prior knowledge on brain anatomy has been ignored or only implicitly learned in both multi-atlas based and deep learning based whole brain segmentation methods. Moreover, image contrast is not the only anatomical clue to discriminate structure boundary, especially for fine-grained structures. For instance, the anterior limb and posterior limb of the internal capsule are both part of white matter, which cannot be separated based on intensities and contrasts but can be differentiated by the sharp bend and feature-rich neighboring gray matter anatomy. This further suggests the importance of exploring and capturing OHSR.

To mimic experts' hierarchical perception of the brain anatomy, we here propose a novel approach to learn brain's hierarchy based on ontology, for a purpose of whole brain segmentation. Specifically, we encode the multi-level ontology knowledge into a voxel-wise embedding space. Deep metric learning is conducted to cluster contextually similar voxels and separate contextually dissimilar ones using a triplet loss with dynamic violate margin. By formatting the brain hierarchy into a directed acyclic graph, the violate margin can be easily induced by the height of the tree rooted at triplet's *least common subsumer*. As a result, the network is able to exploit the hierarchical relationship across different brain structures. The feature prototypes in the latent space are hierarchically organized following the brain hierarchy. To the best of our knowledge, this is the first work to incorporate ontology-based brain hierarchy into deep learning segmentation models. We evaluate our method on two whole brain segmentation datasets with different granularity and successfully establish SOTA performance.

2 Methodology

The proposed approach builds upon standard 3D U-Net and makes use of multi-level ontology knowledge from the brain hierarchy to enhance the whole brain segmentation performance. In subsequent subsections, we first revisit the standard triplet loss [9, 14] and its dilemma in learning brain hierarchy. Then we describe how we construct the brain hierarchy graph and how we measure the semantic similarity between brain structures based on the constructed graph. After that, we introduce dynamic violate margin to the triplet loss.

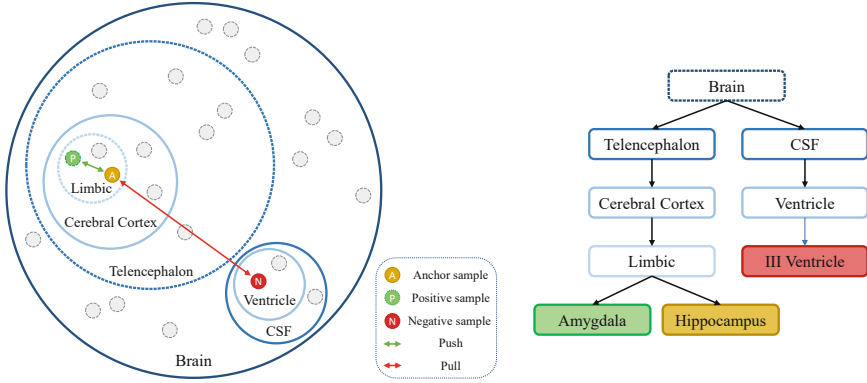


Fig. 2. An intuitive illustration of hierarchy-based triplet loss and its corresponding toy example graph of brain hierarchy. The loss tends to group structures labeled as a same class at a higher ontology level.

2.1 Triplet Loss

The goal of the triplet loss is to learn a feature embedding space wherein distances between features correspond to semantic dissimilarities between objects. Given a triplet $\mathcal{T}_i = \{f_i, f_i^+, f_i^-\}$ comprising an anchor voxel-wise feature vector f_i , a positive voxel-wise feature vector f_i^+ which is semantically similar to the anchor vector, and a negative voxel-wise feature vector f_i^- which is semantically dissimilar to the anchor vector, the triplet loss is formulated as

$$l_{\text{triplet}}(\mathcal{T}_i) = [\langle f_i, f_i^+ \rangle - \langle f_i, f_i^- \rangle + \mathcal{M}]_+, \quad (1)$$

where $\langle \cdot, \cdot \rangle$ is a distance function to evaluate the semantic dissimilarity between two feature vectors. The violate margin \mathcal{M} is a hyperparameter that defines the minimum distance between positive and negative samples. It forces the gap between $\langle f, f^+ \rangle$ and $\langle f, f^- \rangle$ to be larger than \mathcal{M} and ensures the model does not learn trivial solutions. $[\cdot]_+ = \max\{0, \cdot\}$ is the hinge loss, which prevents the model from being updated when the triplet is already fulfilled. During training, the overall objective of the voxel-wise triplet loss is to minimize the sum of the loss over all triplets in a mini-batch, namely

$$\mathcal{L}_{\text{triplet}} = \frac{1}{N} \sum_{\mathcal{T}_i \in \mathcal{T}} l_{\text{triplet}}(\mathcal{T}_i), \quad (2)$$

where N is the total number of triplets in a mini-batch. Note that a triplet is allowed to consist of voxel-wise feature vectors from different subjects when the batch size is larger than 1. This strategy of sampling an anchor's neighbor enables the model to learn the global context in the brain instead of the local context in a subspace of the brain, since it is infeasible to train a 3D U-Net with whole brain MRI data.

However, it is challenging to apply the standard triplet loss to learn brain hierarchy: positive or negative is ill-defined with a fixed violate margin. For instance, the violate margin between $\langle f_{\text{hippo}}, f_{\text{amyg}} \rangle$ and $\langle f_{\text{hippo}}, f_{\text{fimb}} \rangle$ is certainly different from that between $\langle f_{\text{hippo}}, f_{\text{amyg}} \rangle$ and $\langle f_{\text{hippo}}, f_{\text{IIIvent}} \rangle$: the hippocampus, the amygdala, and the fimbria all belong to the limbic system while the third ventricle belongs to cerebrospinal fluid (CSF). As such, a distance function $d_{\mathcal{G}}(\cdot, \cdot)$ is required to measure the semantic dissimilarity between two brain structures, and can be then used to determine the corresponding violate margin.

2.2 Measuring Semantic Dissimilarity

Let $\mathcal{G} = (V, E)$ be a directed acyclic graph with vertices V and edges $E \subseteq V^2$. It specifies the hyponymy relationship between structures at different ontology levels. An edge $(u, v) \in E$ indicates u is an ancestor vertex of v . Specifically, v belongs to u at a higher ontology level. The structures of interest $S = \{s_1, \dots, s_n\} \subseteq V$ are of the lowest ontology level. An example is shown in Fig. 2.

A common measure for the dissimilarity $d_{\mathcal{G}} : S^2 \rightarrow \mathbb{R}$ between two structures is the height of the tree rooted at the *least common subsumer* (LCS) divided by the height of the whole brain hierarchy tree, namely

$$d_{\mathcal{G}}(u, v) = \frac{h(\text{lcs}(u, v))}{h(\mathcal{G})}, \quad (3)$$

where the height of a tree $h(\cdot) = \max_{v \in V} \psi(\cdot, v)$ is defined as the length of the longest path from the root to a leaf. $\psi(\cdot, \cdot)$ is defined as the number of edges in the shortest path between two vertices. $\text{lcs}(\cdot, \cdot)$ refers to the ancestor shared by two vertices that do not have any child also being an ancestor of the same two vertices. With respect to the example hierarchy in Fig. 2, the LCS of the hippocampus and the amygdala is the limbic and the LCS of the hippocampus and the third ventricle is the brain. Given the height of the example hierarchy is 4, we can easily derive that $d_{\mathcal{G}}(\text{hippo}, \text{amyg}) = \frac{1}{4}$ and $d_{\mathcal{G}}(\text{hippo}, \text{IIIvent}) = 1$. Non-negativity, symmetry, identity of indiscernibles, and triangle inequality always hold for $d_{\mathcal{G}}(\cdot, \cdot)$ since the brain hierarchy is a tree and all structures of interest are leaf vertices, thus being a proper metric [2].

2.3 Dynamic Violate Margin

With $d_{\mathcal{G}}(\cdot, \cdot)$, we can define positive and negative samples and their violate margin in the triplet loss. We sample triplet $\mathcal{T}_i = \{f_i, f_i^+, f_i^-\}$ satisfying $d_{\mathcal{G}}(v_i, v_i^-) > d_{\mathcal{G}}(v_i, v_i^+)$, where f, f_i^+, f_i^- are the feature vectors of the voxels respectively labeled as v, v_i^+, v_i^- . Then the violate margin $\tilde{\mathcal{M}}$ can be determined dynamically

$$\begin{aligned} \tilde{\mathcal{M}} &= 0.5(\mathcal{M}_{\tau} + \mathcal{M}_{\epsilon}), \\ \mathcal{M}_{\tau} &= d_{\mathcal{G}}(v_i, v_i^-) - d_{\mathcal{G}}(v_i, v_i^+), \\ \mathcal{M}_{\epsilon} &= \frac{1}{N_v(N_v - 1)} \sum_{a, b \in v} \langle f_a, f_b \rangle. \end{aligned} \quad (4)$$

$\mathcal{M}_\tau \in (0, 1]$ is the hierarchy-induced margin required between negative pairs and positive pairs in terms of $d_{\mathcal{G}}(\cdot, \cdot)$. \mathcal{M}_ϵ is the tolerance of the intra-class variance, which is computed as the average distance between samples in v . In this work, we adopt the cosine distance as our distance function in latent space: $\langle f_a, f_b \rangle = \frac{1}{2}(1 - \frac{f_a \cdot f_b}{\|f_a\| \|f_b\|}) \in (0, 1]$. The triplet loss can thus be reformulated as

$$l_{\text{triplet}}(\mathcal{T}_i) = [\langle f_i, f_i^+ \rangle - \langle f_i, f_i^- \rangle + \tilde{\mathcal{M}}]_+. \quad (5)$$

Collectively, the overall training objective to learn OHSR is

$$\mathcal{L} = \mathcal{L}_{\text{seg}} + \lambda \mathcal{L}_{\text{triplet}}, \quad (6)$$

where λ is a hyperparameter.

3 Experiment

3.1 Datasets and Implementation

We evaluate our method on two public-accessible datasets with manually labeled fine-grained or coarse-grained brain structures. The first one is JHU Adult Atlas [25], containing 18 subjects with an age range of 27–55 years. T1-weighted MRI images are acquired in the MPRAGE sequence at Johns Hopkins University using 3T Philips scanners. All images are normalized to the MNI152 1mm space. The images are initially segmented into 289 structures using a single-subject atlas followed by substantial manual corrections. The optic tract, skull and bone marrow are excluded from our experiment, ending up with 282 structures. Two types of five-level OHSR are provided. Type I is based on classical definitions of the brain ontology and Type II is more commonly used in clinical descriptions [1, 7, 10]. The second one is Child and Adolescent Neuro Development Initiative (CANDI) [13], consisting of 103 1.5T T1-weighted MRI scans with 32 labeled structures. The subjects are aged 4–17 and come from both healthy and neurological disorder groups. A two-level hierarchical relationship is built by grouping the 32 structures into white matter, gray matter or CSF.

For JHU Adult Atlas, 8, 2 and 8 images are randomly selected and used for training, validation and testing. For CANDI, we split the dataset into training (60%), validation (10%), and testing (30%) sets following a previous study [15]. Foreground image patches of size $96 \times 112 \times 96$ are randomly cropped as the input of our model. To enlarge the training set, random scaling, gamma correction and rotation are applied. Image intensities are normalized using Z-score normalization. AdamW with a batch size of 4 is used to optimize the training objective, wherein $\mathcal{L}_{\text{seg}} = \mathcal{L}_{\text{dice}} + \mathcal{L}_{\text{ce}}$. $\mathcal{L}_{\text{dice}}$ and \mathcal{L}_{ce} are respectively the Dice loss and the cross entropy loss. The hyperparameter $\lambda \in [0, 0.5]$ is scheduled following a cosine annealing policy. The initial learning rate is 5×10^{-4} and decays following a polynomial function. The model is trained for 30,000 steps, with the best model saved based on the validation Dice. All experiments are conducted using PyTorch 1.13.1 with NVIDIA Tesla V100 GPUs. More dataset and implementation details are provided in the supplementary material.

Table 1. Comparisons with SOTA for segmenting 282 fine-grained brain structures on JHU Adult Atlas, in terms of DSC. “S” denotes small structures with a size smaller than 1000 mm³. “M” denotes medium structures with a size between 1000 mm³ and 5000 mm³. “L” denotes large structures with a size larger than 5000 mm³. “A” indicates all structures. The best results are highlighted in **bold**. Please note all compared results are based on self-reimplementation.

Method	DSC(%)			
	S	M	L	A
U-Net [3]	66.56 ± 15.31	73.37 ± 16.22	78.58 ± 6.79	73.82 ± 14.07
QuickNAT [21]	69.30 ± 19.55	72.09 ± 20.76	78.97 ± 9.15	74.25 ± 17.25
nnU-Net [12]	80.04 ± 14.60	82.85 ± 13.06	84.79 ± 5.84	82.97 ± 11.56
Proposed	80.76 ± 13.69	83.51 ± 13.13	85.52 ± 5.38	83.67 ± 11.26

Table 2. Comparisons with SOTA for segmenting 32 coarse-grained brain structures on CANDI, in terms of DSC. The best result is highlighted in **bold**.

Method	DSC(%)
U-Net [20]	85.70 ± 9.70
QuickNAT V2 [22]	86.20 ± 9.50
nnU-Net [12]	87.97 ± 8.46
ACEnet [15]	88.10 ± 7.40
Proposed	88.23 ± 6.32

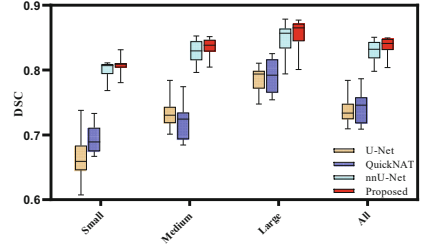


Fig. 3. DSC boxplots of small, medium, large and all structures from different methods on JHU Adult Atlas.

3.2 Evaluation Results

We now report quantitative and qualitative evaluation results. The Dice similarity coefficient (DSC) is used to quantitatively measure the segmentation accuracy. We compare our proposed method with SOTA methods and conduct several ablation studies to demonstrate the effectiveness of our method.

Comparisons with SOTA. To fairly compare with SOTA methods, we use the identical 3D U-Net and data augmentation hyperparameters as the “3d_fullres nnUNetTrainerV2_noMirroring” configuration in nnU-Net. As summarized in Table 1, our method obtains the highest overall mean DSC of 83.67% on JHU Adult Atlas, with an improvement of 0.70% over previous best method, i.e., nnU-Net. The improvements are consistent over small, medium and large brain structures, demonstrating the robustness of our method. The boxplot comparison results are illustrated in Fig. 3. Detailed improvement of each structure is presented in the supplementary material. Table 2 shows the results on CANDI. Please note all compared results except for nnU-Net in that table are directly copied from ACEnet paper [15] since CANDI has been explored more than JHU

Table 3. Analysis of the effect of OHSR on JHU Adult Atlas.

Method	DSC(%)
U-Net	73.82 \pm 14.07
U-Net + OHSR (Type I)	77.34 \pm 14.52 (+3.52)
U-Net + OHSR (Type II)	77.99 \pm 12.93 (+4.17)
nnU-Net	82.97 \pm 11.56
nnU-Net + OHSR (Type I)	83.49 \pm 11.21 (+0.52)
nnU-Net + OHSR (Type II)	83.67 \pm 11.26 (+0.70)

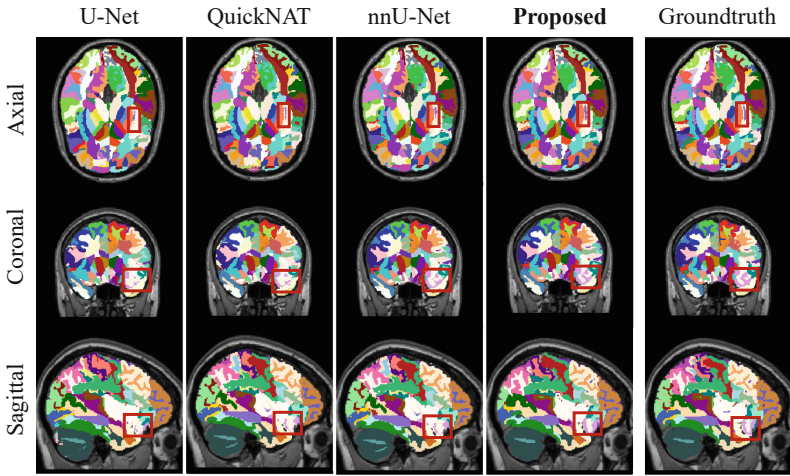


Fig. 4. Qualitative comparisons between our proposed method and SOTA. The regions of interest are highlighted in bounding boxes.

Adult Atlas. Our method achieves an average DSC of 88.23%, performing not only better than the hyperparameter-tuning method nnU-Net, but also better than ACENet encoding anatomical context via an attention mechanism. These results also suggest that our method can leverage both simple (two-level) and sophisticated (five-level) hierarchies to enhance whole brain segmentation.

Ablation Studies. We first evaluate the effectiveness of incorporating OHSR into U-Net. The experiments are conducted on JHU Adult Atlas. As shown in Table 3, learning OHSR using a triplet loss with graph-based dynamic margin improves U-Net by 3.52% in DSC. This indicates our method can empower a relatively small network to segment fine-grained brain structures with considerable accuracy. Our method even outperforms nnU-Net by 0.70% in DSC, clearly demonstrating its superiority. We further compare the performance between two types of OHSR. As tabulated in Table 3, Type II ontology achieves bet-

ter performance, indicating mimicking clinicians to understand brain hierarchy is of greater help.

Qualitative Results. Qualitative comparisons are demonstrated in Fig. 4. From the axial view, we can clearly see the external capsule is well-segmented by our proposed method, while other methods can hardly differentiate its boundary. From the coronal and sagittal views, we observe that our method can better capture and preserve the overall shape of the lateral frontal-orbital gyrus.

4 Conclusion

In this paper, we propose a novel approach to learn brain hierarchy based on ontology for whole brain segmentation. By introducing graph-based dynamic violate margin into the triplet loss, we encode multi-level ontology knowledge into a voxel-wise embedding space and mimic experts' hierarchical perception of the brain anatomy. We successfully demonstrate that our proposed method outperforms SOTA methods both quantitatively and qualitatively. We consider introducing hierarchical information into the output space as part of our future efforts.

Acknowledgement. This study was supported by the National Natural Science Foundation of China (62071210); the Shenzhen Science and Technology Program (RCYX20210609103056042); the Shenzhen Science and Technology Innovation Committee (KCXFZ2020122117340001); the Shenzhen Basic Research Program (JCYJ20200925153847004, JCYJ20190809120205578).

References

1. Ackerman, S.: Discovering the brain (1992)
2. Barz, B., Denzler, J.: Hierarchy-based image embeddings for semantic image retrieval. In: IEEE Winter Conference on Applications of Computer Vision, pp. 638–647 (2019)
3. Cardoso, M.J., et al.: MONAI: an open-source framework for deep learning in healthcare. arXiv preprint [arXiv:2211.02701](https://arxiv.org/abs/2211.02701) (2022)
4. Ceritoglu, C., Tang, X., Chow, M., et al.: Computational analysis of LDDMM for brain mapping. *Front. Neurosci.* **7**, 151 (2013)
5. Chupin, M., G  rardin, E., Cuingnet, R., et al.: Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus* **19**(6), 579–587 (2009)
6. Coup  , P., et al.: AssemblyNet: a large ensemble of CNNs for 3D whole brain MRI segmentation. *Neuroimage* **219**, 117026 (2020)
7. Djamanakova, A., Tang, X., Li, X., et al.: Tools for multiple granularity analysis of brain MRI data for individualized image analysis. *Neuroimage* **101**, 168–176 (2014)
8. Erickson, K.I., Voss, M.W., Prakash, R.S., et al.: Exercise training increases size of hippocampus and improves memory. *Proc. Natl. Acad. Sci.* **108**(7), 3017–3022 (2011)

9. Ge, W.: Deep metric learning with hierarchical triplet loss. In: Proceedings of the European Conference on Computer Vision, pp. 269–285 (2018)
10. Gong, Y., Wu, H., Li, J., Wang, N., Liu, H., Tang, X.: Multi-granularity whole-brain segmentation based functional network analysis using resting-state fMRI. *Front. Neurosci.* **12**, 942 (2018)
11. Huo, Y., et al.: 3D whole brain segmentation using spatially localized atlas network tiles. *Neuroimage* **194**, 105–119 (2019)
12. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
13. Kennedy, D.N., Haselgrove, C., Hodge, S.M., Rane, P.S., Makris, N., Frazier, J.A.: CANDIShare: a resource for pediatric neuroimaging data. *Neuroinformatics* **10**, 319–322 (2012)
14. Li, L., Zhou, T., Wang, W., Li, J., Yang, Y.: Deep hierarchical semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1246–1257 (2022)
15. Li, Y., Li, H., Fan, Y.: ACENet: anatomical context-encoding network for neuroanatomy segmentation. *Med. Image Anal.* **70**, 101991 (2021)
16. Lyu, J., Zhang, Y., Huang, Y., Lin, L., Cheng, P., Tang, X.: AADG: automatic augmentation for domain generalization on retinal image segmentation. *IEEE Trans. Med. Imaging* **41**(12), 3699–3711 (2022)
17. Mai, J.K., Majtanik, M., Paxinos, G.: Atlas of the Human Brain. Academic Press, Cambridge (2015)
18. Puelles, L., Harrison, M., Paxinos, G., Watson, C.: A developmental ontology for the mammalian brain based on the prosomeric model. *Trends Neurosci.* **36**(10), 570–578 (2013)
19. Tang, X., et al.: Bayesian parameter estimation and segmentation in the multi-atlas random orbit model. *PLoS ONE* **8**(6), e65591 (2013)
20. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
21. Roy, A.G., Conjeti, S., Navab, N., Wachinger, C.: Alzheimer’s Disease Neuroimaging Initiative: QuickNAT: a fully convolutional network for quick and accurate segmentation of neuroanatomy. *NeuroImage* **186**, 713–727 (2019)
22. Roy, A.G., Navab, N., Wachinger, C.: Recalibrating fully convolutional networks with spatial and channel “squeeze and excitation” blocks. *IEEE Trans. Med. Imaging* **38**(2), 540–549 (2018)
23. Tang, X., Qin, Y., Wu, J., et al.: Shape and diffusion tensor imaging based integrative analysis of the hippocampus and the amygdala in Alzheimer’s disease. *Mag. Reson. Imaging* **34**(8), 1087–1099 (2016)
24. Wang, H., Yushkevich, P.A.: Multi-atlas segmentation with joint label fusion and corrective learning—an open source implementation. *Front. Neuroinf.* **7**, 27 (2013)
25. Wu, D., et al.: Resource atlases for multi-atlas brain segmentations with multiple ontology levels based on T1-weighted MRI. *Neuroimage* **125**, 120–130 (2016)
26. Wu, J., Tang, X.: Brain segmentation based on multi-atlas and diffeomorphism guided 3D fully convolutional network ensembles. *Pattern Recognit.* **115**, 107904 (2021)