Chapter 14
Gray and White Matters Segmentation in Brain CT Images Using Multi-task Learning from Paired CT and MR Images

Taohai Han and Hongkai Wang

Abstract Computer tomography (CT) has been routinely used for decades in clinical neuroimaging. Compared to Magnetic resonance imaging (MR), CT is more readily available and more cost-effective, but the soft-tissue contrast is much lower. In brain CT images, the unclear soft-tissue boundaries and the high noise level hamper accurate segmentation of the gray and white matter, leaving obstacles for the subsequent geometrical quantification of brain structures. To address this challenge, this paper specifically acquires same-patient CT and MR image pairs and proposes a multi-task learning model for simultaneous tissue segmentation and modality transfer. We aim to use the modality transfer task to learn corresponding MR and CT features which assists the segmentation task to achieve more accurate results than single modality learning. Moreover, we add a Shannon entropy loss function to the training loss to further combat the high noise influence and reduce the fragmentation problem of the segmentation results. Experimental results proved that our multi-task framework achieves more accurate segmentation than the single segmentation task, and the Shannon entropy loss results in much fewer broken brain regions than the state-of-the-art (SOTA) U-net method. Our study provides a useful tool for clinical neural CT image analysis.

T. Han
School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Dalian, Liaoning, China

H. Wang (✉)
Liaoning Key Laboratory of Integrated Circuit and Biomedical Electronic System, Dalian University of Technology, Dalian, Liaoning, China
e-mail: wang.hongkai@dlut.edu.cn

R. Kountchev et al. (eds.), *New Approaches for Multidimensional Signal Processing*,
Smart Innovation, Systems and Technologies 385,
https://doi.org/10.1007/978-981-97-0109-4_14

14.1 Introduction

With the aging of the population and the growing number of age-related mental diseases, craniocerebral examinations have become more and more important for early disease prevention and diagnosis. Brain CT, especially high-resolution brain CT, is a commonly used tool for brain disease screening. Compared with cranial MR images, brain CT has the advantage of easy access and a large inventory [1], but the soft-tissue contrast and signal-to-noise ratio are relatively lower. The popularity of brain CT results in a large number of high-resolution brain CT images for clinical examination and diagnosis, which increases the burden of the human reviewer [2]. To perform computer-aided diagnosis based on the brain CT images, automated segmentation of brain regions is a prerequisite step. Since CT doesn't have good soft-tissue contrast as MR, it is almost infeasible to segment sub-cortical brain regions from CT images. Therefore, the segmentation of entire gray and white matters becomes a practical clinical need.

Despite the rapid development of computer-assisted medical image segmentation technology, especially deep learning technology in the last decade, there is still no ideal algorithm for gray and white matter segmentation from CT images. The reason is that the contrast between gray and white matter in brain CT images is particularly low thus the boundary is very blurred [3, 4], which we can see in Fig. 14.1. Besides, the high noise level makes it even more difficult to achieve accurate segmentation results. Due to these problems, even experienced physicians have difficulty in manually labeling white and gray matter from CT images. Therefore, the labeling required for deep learning training is difficult to produce, hampering the implementation of deep learning methods in this task.

Unlike CT-based segmentation, gray and white matter segmentation from MR images is already a well-solved problem, there are many existing solutions for MR-based brain region segmentation, such as FreeSurfer [5], SPM [6], BrainSuite [7], etc. Nevertheless, there are still a few studies based on CT images. Li et al. concluded that the continuous improvement of CT scanning quality leads to increasing feasibility for soft-tissue segmentation from CT [8]. Qian et al. [9] used the active contour model for cerebrospinal fluid segmentation of brain CT images. And Manniesing [10] proposed

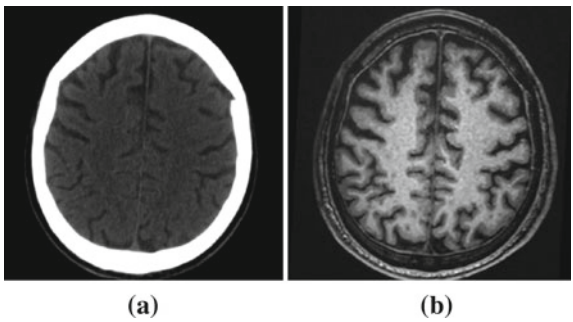

Fig. 14.1 Comparison of brain CT image and brain MR image. **a** The brain CT image. **b** The brain MR image

a CT-based segmentation method with a post-correction process. Yahiaoui et al. [11] targeted the segmentation of ischemic stroke regions on brain CT, and Li et al. [12] studied the segmentation of hemorrhagic lesions. In summary, most of the existing studies on brain region segmentation algorithms are based on MR images, while most of the segmentation algorithms for brain CT focus on specific diseases, rather than gray and white matter regions.

Considering the large amounts of daily acquired brain CT images and the clinical need for white and gray matter segmentation, it is highly demanding to develop automated algorithms for CT-based gray and white matter segmentation. Although the low soft-tissue contrast and high noise level impose difficulties for accurate segmentation, the recent progress in multi-task deep learning shed light on novel solutions to this problem. Multi-task learning puts several tasks requiring similar data features together for simultaneous deep network training so that different tasks help each to achieve better performance than any single task alone. For medical image segmentation, a series of studies have shown that multi-task learning effectively promotes inter-task cooperation and improves overall segmentation accuracy [13–15].

In this study, we take advantage of the multi-task learning strategy to overcome the low soft-tissue contrast and high noise interference and improve the segmentation accuracy of gray and white matter from brain CT images. Through simultaneous learning of CT segmentation and CT-to-MR modality transfer, the network learns the corresponding latent features between the two modalities, therefore exceeding the performance of learning from a single modality (CT). Our study has the following main contributions:

(i) Unlike many existing unpaired domain adaption methods, our study especially acquired 27 pairs of same-patient CT and MR images. In this way, pixel-level strong inter-modality feature correspondence is established, so that the network learns the corresponding high-contrast MR features to promote the segmentation accuracy of CT images.
(ii) The acquisition of paired MR and CT images solves the labeling difficulty of CT images. We used the well-recognized FreeSurfer software to segment the MR images and map the tissue labels into CT image space via inter-modality registration. Therefore, the training labels for the CT segmentation network are obtained.
(iii) We overcome the high noise interference in the CT images by adding the Shannon entropy loss which encourages the high confidence network. As a result, the low confidence segmentation caused by the noise is removed and the segmentation result becomes more compact.

14.2 Method

14.2.1 Image Dataset

The dataset for this study was specially acquired pairs of brain CT and MR images of 27 normal subjects, i.e., each of the samples have corresponding CT and MR images of the same patient. First, we used FreeSurfer to segment the brain regions of the MR images to obtain the white and gray matter labels. Then, using rigid registration, the MR images are aligned to the corresponding CT images, and the segmentation labels are mapped into CT image space via the rigid transformation resulting from the registration. This study focuses on 2D image segmentation thus 2260 2D slices are exacted from the 27 image series, each containing one brain CT slice, one label map, and one MR slice aligned to the CT space. The size of each slice image is 512 $\times$ 512. We use the CT image as the input of the multi-task network, the white matter, and gray matter labels to supervise the segmentation task, and the aligned MR image to supervise the modality transfer task. Figure 14.2 shows the flow chart for making brain CT image segmentation labels.

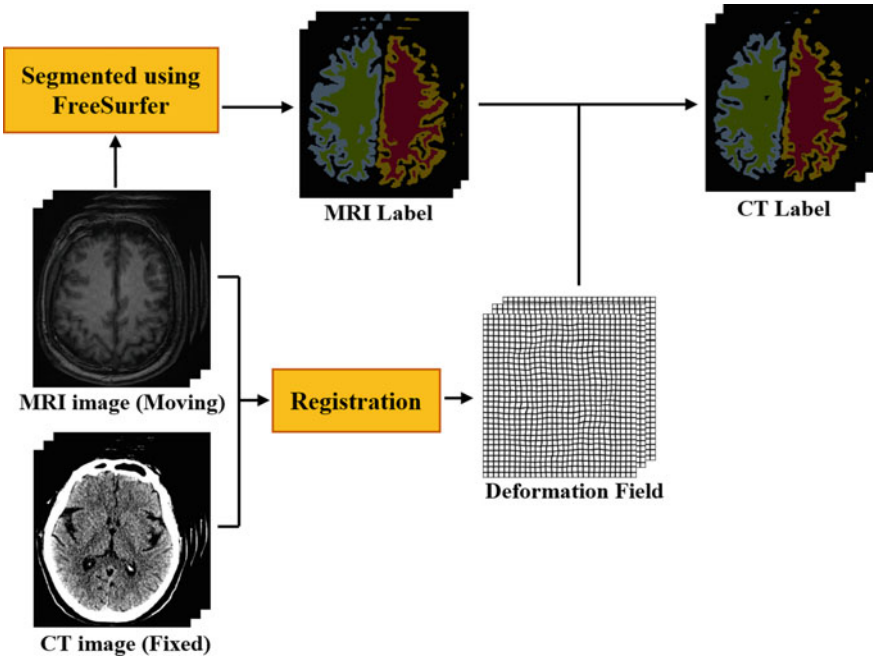

Fig. 14.2 Workflow of the proposed method for making brain CT labels

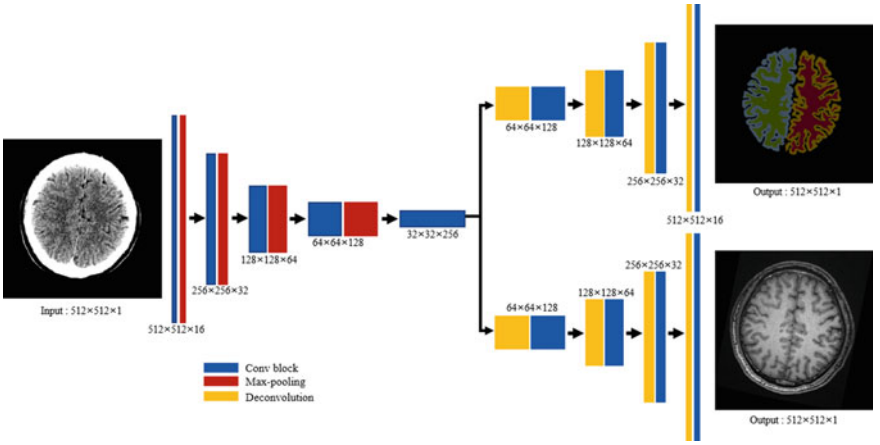

Fig. 14.3 The architecture of the proposed multi-task learning network model

14.2.2 Network Architecture

Our end-to-end multi-task network architecture is shown in Fig. 14.3. The network contains one encoder part and two decoder branches for the segmentation task and the modality conversion task, respectively. The two tasks share the same encoder containing four down-sampling layers, such a structure fully acquires multi-scale information for feature extraction of the brain CT. The two tasks have their independent decoders, and both contain four up-sampling layers. The image segmentation and modality transfer tasks, in turn, drive the encoder to extract the common features required for both tasks, thus improving the performance of each task by learning relative features from the other task.

14.2.3 Loss Function

For the segmentation task, we use a hybrid loss function, including a cross-entropy loss term [16] and a Shannon entropy loss [17] term, respectively.

The cross-entropy loss function is a commonly used loss function for image segmentation, which calculates the cross-entropy between the network prediction probability and the ground truth label:

$$L_{ce} = -\sum_{c=1}^{M} y_c \log(p_c) \tag{14.1}$$

where M represents the number of segmentation classes, y_c represents the one-hot encoding of the corresponding pixel labels, and p_c represents the prediction probability of the neural network that the pixel belongs to that species. For the whole image, the cross-entropy loss of all the pixels is averaged.

In addition to the conventional cross-entropy loss, this paper introduces the Shannon entropy loss function to enhance the prediction confidence of the segmentation task:

$$L_{se} = -\sum_{c=1}^{M} p_c \log(p_c) \tag{14.2}$$

where p_c represents the prediction probability of the neural network for the pixel belonging to each class. For the whole image, the Shannon entropy loss of all the pixels is averaged. We adopt the Shannon entropy loss to encourage networks to make more confident predictions, and to combat the low prediction confidence caused by the high CT noise.

The overall segmentation loss function is the weighted sum of the cross-entropy loss and the Shannon entropy loss:

$$L_{seg} = \alpha L_{ce} + (1 - \alpha) L_{se} \tag{14.3}$$

where α is the factor that adjusts the ratio of the two loss functions.

For the modality transfer task, the SSIM (Structural Similarity) loss function [18] is used in this paper. The SSIM index is often used to measure the structural similarity between two images. The SSIM index is sensitive to the perception of local details:

$$SSIM\,(p, g) = \frac{\left(2\mu_p\mu_g + c_1\right)\left(2\sigma_{pg} + c_2\right)}{\left(\mu_p^2 + \mu_g^2 + c_1\right)\left(\sigma_p^2 + \sigma_g^2 + c_2\right)} \tag{14.4}$$

where μ_p and μ_g represent the pixel means of the predicted and labeled images, respectively. σ_p and σ_g represent the standard deviations of the predicted and labeled images, respectively. σ_{pg} represents the covariance of the predicted and labeled images. c_1 and c_2 are constant terms preventing the divisor from being zero.

The corresponding SSIM loss function is then defined as

$$L_{trans} = 1 - SSIM \tag{14.5}$$

The overall loss function of the model is defined by the following equation:

$$L_{mix} = \beta L_{seg} + (1 - \beta) L_{trans} \tag{14.6}$$

where β is the factor that adjusts the ratio of the two loss functions.

14.2.4 Model Training

In this study, the brain CT images input to the network are intercepted with grayscale between [20,60] and the MR images aligned with the CT images are intercepted with grayscale between [100,500] and normalized respectively. We take 0.5 for both α and β in the loss function. We randomly select 2000 2D slices from the dataset as the training set and the remaining 260 slices as the test set. The model training process uses Adam optimizer to optimize the parameters of the model, and the learning rate is set to 1×10^{-4}, and the learning rate decay with multiple steps is also set. The model was trained on two 3090Ti graphics cards manufactured by NVIDIA and built using PyTorch 1.9.

14.3 Result and Analysis

For evaluation, we compared the proposed method with the state-of-the-art (SOTA) U-Net model [19] based on the 260 test slices. To verify the contribution of the modality transfer task and the Shannon entropy loss function, we also conducted ablation studies to compare our method with the multi-task learning without Shannon entropy loss, and the single task segmentation with the Shannon entropy.

We chose the Dice coefficient [20] to evaluate the segmentation accuracy, which is defined as

$$Dice = 2\frac{|R_r \cap R_s|}{|R_r| + |R_s|} \tag{14.7}$$

where R_r is the brain area obtained by network prediction and R_s is the label obtained by mapping the FreeSurfer segmentation of MR image into the CT space. A dice coefficient close to 1 indicates an accurate segmentation and 0 indicates completely inaccurate segmentation.

Table 14.1 reveals that our method obtains the overall highest Dice in comparison with the other three methods. Compared with the U-Net network, we achieved a Dice improvement of 0.03 and 0.02 for white matter and gray matter, respectively. Ablation studies removing either the modality transfer task or the Shannon entropy loss resulted in a drop in the Dice coefficient. It is interesting to see that the Shannon entropy loss contributes slightly more than the modality transfer task, but this should be further validated with more test data in the future study.

Figure 14.4 shows the segmentation results of the four methods at different anatomical positions. We demonstrate the registered MR and the FreeSurfer labels (the ground truth) along with the input CT images. It can be seen that our method produces visually closer results to the ground truth. Compared to the original U-Net result, the addition of Shannon entropy loss results in fewer broken regions, while the multi-task learning also helps to produce more accurate segmentation, especially for the complex-shaped white matter region.

Table 14.1 Comparison table of dice coefficients

Method	White matter	Gray matter
U-Net	0.71 ± 0.05	0.65 ± 0.04
U-Net + Shannon entropy loss	0.74 ± 0.04	0.67 ± 0.04
U-Net + modality transfer	0.73 ± 0.04	0.66 ± 0.04
Our method	$\mathbf{0.74 \pm 0.03}$	$\mathbf{0.67 \pm 0.04}$

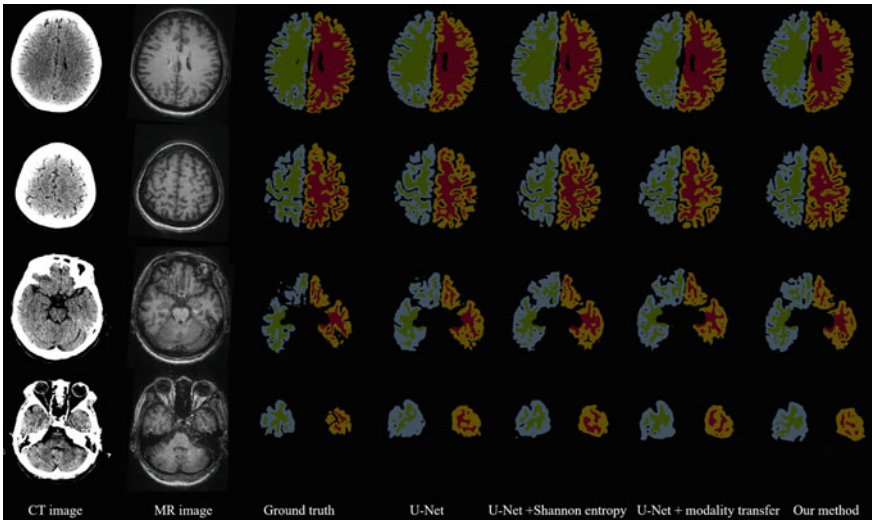

Fig. 14.4 Comparison of our method with the other three methods

14.4 Conclusion

This study addresses the difficult CT-based gray and white matter segmentation problem using multi-task learning based on specially acquired CT-MR image pairs of the same patients. Using the CT-to-MR modality transfer as an auxiliary task, the corresponding MR image feature is learned by the network to promote the segmentation of the CT image. We also added the Shannon entropy to the loss function to overcome the high noise interference. This work has potential value for improving the data analysis ability for clinical brain CT images.

Acknowledgements This work was supported in part by the National Key Research and Development Program No. 2020YFB1711500, 2020YFB1711501 and 2020YFB1711503, the general program of National Natural Science Fund of China (No. 81971693, 61971445), the funding of Dalian Engineering Research Center for Artificial Intelligence in Medical Imaging, Hainan Province Key Research and Development Plan ZDYF2021SHFZ244, the Fundamental Research Funds for the Central Universities (No. DUT22YG229), the funding of Liaoning Key Lab of IC & BME System and Dalian Engineering Research Center for Artificial Intelligence in Medical Imaging.

References

1. Fawzi, A., Achuthan, A., Belaton, B.: Brain image segmentation in recent years: a narrative review. Brain Sci. **11**(8), 1055 (2021). https://doi.org/10.3390/brainsci11081055
2. Irimia, A., Maher, A.S., Rostowsky, K.A., Chowdhury, N.F., Hwang, D.H., Law, E.M.: Brain segmentation from computed tomography of healthy aging and geriatric concussion at variable spatial resolutions. Front. Neuroinform. **13**, 9 (2019). https://doi.org/10.3389/fninf.2019.00009.
3. Gupta, V., et al.: Automatic segmentation of cerebrospinal fluid, white and gray matter in unenhanced computed tomography images. Acad. Radiol. **17**(11), 1350–1358 (2010). https://doi.org/10.1016/j.acra.2010.06.005
4. Zhao, C., Carass, A., Lee, J., He, Y., Prince, J.L.: Whole brain segmentation and labeling from CT using synthetic MR images. In: Machine Learning in Medical Imaging, pp. 291–298. Springer International Publishing (2017). https://doi.org/10.1007/978-3-319-67389-9_34.
5. Greve, D.N., et al.: Different partial volume correction methods lead to different conclusions: an 18F-FDG-PET study of aging. Neuroimage **132**, 334–343 (2016). https://doi.org/10.1016/j.neuroimage.2016.02.042
6. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage **26**(3), 839–851 (2005). https://doi.org/10.1016/j.neuroimage.2005.02.018.
7. Shattuck, D.W., Leahy, R.M.: BrainSuite: an automated cortical surface identification tool. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000, pp. 50–61. Springer Berlin Heidelberg (2000). https://doi.org/10.1007/978-3-540-40899-4_6.
8. Li, Z., et al.: Improvement of image quality and radiation dose of CT perfusion of the brain by means of low-tube voltage (70 KV). Eur Radiol. **24**(8), 1906–1913 (2014). https://doi.org/10.1007/s00330-014-3247-1.
9. Qian, X., Wang, J., Guo, S., Li, Q.: An active contour model for medical image segmentation with application to brain CT image: an active contour model for medical image segmentation. Med. Phys. **40**(2), 021911 (2013). https://doi.org/10.1118/1.4774359.
10. Manniesing, R., et al.: White matter and gray matter segmentation in 4D computed tomography. Sci Rep. **7**(1), 119 (2017). https://doi.org/10.1038/s41598-017-00239-z
11. Yahiaoui, A.F.Z., Bessaid, A.: Segmentation of ischemic stroke area from CT brain images. In: 2016 International Symposium on Signal, Image, Video and Communications (ISIVC), pp. 13–17. Tunis, Tunisia (2016). https://doi.org/10.1109/ISIVC.2016.7893954.
12. Li, L., et al.: Deep Learning for Hemorrhagic Lesion Detection and Segmentation on Brain CT Images. IEEE J. Biomed. Health Inform. **25**(5), 1646–1659 (2021). https://doi.org/10.1109/JBHI.2020.3028243.
13. Murugesan, B., Sarveswaran, K., Shankaranarayana, S.M., Ram, K., Joseph, J., Sivaprakasam, M.: Conv-MCD: a plug-and-play multi-task module for medical image segmentation. In: Machine Learning in Medical Imaging, pp. 292–300. Springer International Publishing (2019). https://doi.org/10.1007/978-3-030-32692-0_34.
14. Tan, C., Zhao, L., Yan, Z., Li, K., Metaxas, D., Zhan, Y.: Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 1221–1224. Washington, DC (2018). https://doi.org/10.1109/ISBI.2018.8363791.
15. Murugesan, B., Sarveswaran, K., Shankaranarayana, S.M., Ram, K., Joseph, J., Sivaprakasam, M.: Psi-Net: shape and boundary aware joint multi-task deep network for medical image segmentation. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 7223–7226. Berlin, Germany (2019). https://doi.org/10.1109/EMBC.2019.8857339.
16. Zhang, Z., Sabuncu, M.R.: Generalized cross entropy loss for training deep neural networks with noisy labels. arXiv, Nov. 29, 2018. Accessed 07 Jan 2023. [Online]. Available: http://arxiv.org/abs/1805.07836

17. Liu, B., Desrosiers, C., Ben Ayed, I., Dolz, J.: Segmentation with mixed supervision: confidence maximization helps knowledge distillation. Med. Image Anal. **83**, 102670 (2023). https://doi.org/10.1016/j.media.2022.102670.
18. Zhao, H., Gallo, O., Frosio, I., Kautz, J.: Loss functions for image restoration with neural networks. IEEE Trans. Comput. Imaging. **3**(1), 47–57 (2017). https://doi.org/10.1109/TCI.2016.2644865.
19. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, pp. 234–241. Springer International Publishing (2015). https://doi.org/10.1007/978-3-319-24574-4_28.
20. Zou, K. H., et al.: Statistical validation of image segmentation quality based on a spatial overlap index1. Acad. Radiol. **11**(2), 78–189 (2004). https://doi.org/10.1016/S1076-6332(03)00671-8.