Automatic Segmentation of Infant Brain Ventricles with Hydrocephalus in MRI Based on 2.5D U-Net and Transfer Learning

Kenji Ono, Yutaro Iwamoto, and Yen-Wei Chen
Information Science and Engineering, Ritsumeikan University, Shiga, Japan
Email: is0252ek@ed.ritsumei.ac.jp, yiwamoto@fc.ritsumei.ac.jp, chen@is.ritsumei.ac.jp

Masahiro Nonaka
Kansai Medical University, Osaka, Japan
Email: nonakamasa65@gmail.com

***Abstract*—The goal of our study is to segment and quantify brain ventricles in infants with hydrocephalus. The Hydrocephalus is a brain disease in which cerebrospinal fluid accumulates in the ventricles, which expand abnormally. The ventricles then press on other brain tissues, leading to the risk of multiple functional and developmental disorders. Segmenting brain ventricles is necessary for early detection and surgical follow-up. Unfortunately, there are few studies on patients with hydrocephalus and infant ventricles are complex and diverse with limited data. Moreover, using conventional automatic segmentation by atlas and machine learning with handcrafted features is difficult to segment the infant brain ventricles with hydrocephalus because of the above data-specific issues. Here, we propose a deep automatic method based on 2.5D U-Net and transfer learning to segment the infant brain ventricles with hydrocephalus. We apply a network architecture that combines low-level features with high-level features to improve learning efficiency, and to maintain the correlation in the slice direction. The input images of the network are multi-slice images (the target slice image and its neighbor slices). Furthermore, we apply transfer learning using adult datasets to deal with limited data and fine-tuning in the hydrocephalus infant datasets. In our experiments, our proposed method outperforms conventional methods and improves the DICE from 58% to 72%.**

***Index Terms*—deep learning, 2.5D, U-net, transfer learning, hydrocephalus infant ventricular, MRI**

I. Introduction

Our study focuses on infant patients with hydrocephalus. The hydrocephalus is a brain disease in which cerebrospinal fluid accumulates in the ventricles, which is expanded abnormally, leading to the risk of multiple functional and developmental disorders [1]. The difference between normal, hydrocephalus adult and infant are shown in Fig. 1. Shunt surgery is an effective treatment, and it is necessary for surgical follow-up to quantitatively evaluate the ventricular volume before and after surgery and for surgical follow-up. An automatic segmentation method is needed because manual segmentation requires more burden and highly experienced labor resources.

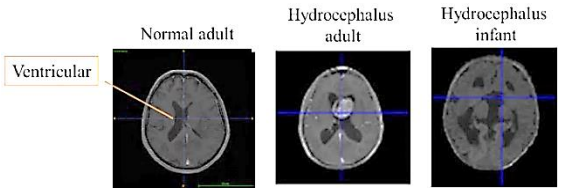


Figure 1. Axial slice of normal, hydrocephalus adult and infant.

Conventional methods based on basic ventricular features [2] or using a brain atlas [3] are difficult to be applied to segmentation of infant brain ventricles with hydrocephalus because of their complicated and diverse shapes. Moreover, machine-learning methods with handcrafted features [4], [5] have problems with the mis-segmentation of similar regions (e.g., neighboring tumors) and face overfitting problems due to the limited data.

Therefore, we propose a deep automatic segmentation model called 2.5D U-Net for hydrocephalus-affected infant brain ventricles. Our proposed method and contributions are as follows.

- 2.5D U-Net: multi-slice inputs including the target slice image and its neighbored slices enable us to segment the 3D data effectively because of considering the correlation in the slice direction.
- Transfer Learning (TL) with an adult dataset: we show that TL with an adult dataset could improve the segmentation accuracy of hydrocephalus-affected infant brain ventricles. This is beneficial because more adult data are available and hydrocephalus-affected adult brain ventricles are easy to segment (automatically or manually). Therefore, we have more data for training.
- The proposed method outperforms conventional methods: the DICE is improved from 58% to 72%.

The proposed method improves the accuracy of the conventional method and enables robust segmentation even for limited data.

Manuscript received January 13, 2020; revised May 8, 2020.

doi: 10.18178/joig.8.2.42-46

The structure of the paper is as follows. The related work is summarized in Section II. The proposed method is described in Section III. The experimental results are presented in Section IV, and finally a summary is given in Section V.

II. Related Work

Recently, Fully Convolutional Networks (FCNs) [6] are widely used for semantic segmentation. U-Net is one of FCNs and is often used for medical image segmentation [7]. The architecture of U-Net is shown in Fig. 2. It is composed of encoder for extracting features and decoder for reconstruction of images. In order to combine low level features and high level features and to achieve accurate localization, skip connections (white arrows in Fig. 2) are used to connect the encoders and decoders. Since U-Net is proposed for 2D image segmentation, the segmentation of 3D structures such as medical volume images is realized by repeating the segmentation slice by slice. This approach do not include the context information along the z axis, the consistency among slices is lost.

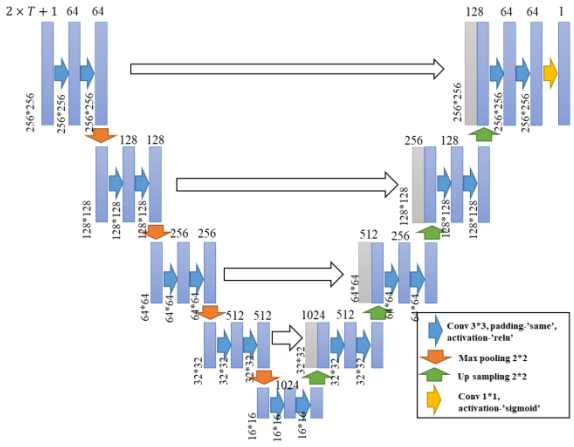


Figure 2. U-net architecture.

In order to do 3D image segmentation, 3D U-Net [8], [9] is also proposed as an extension of U-Net. In 3D U-Net, each layer of U-Net is replaced with 3D convolution and 3D max pooling respectively. Though the 3D U-Net can include the context information along the z axis, it has many parameters and leads to overfitting if the training dataset is small.

Moreover, the transfer learning is attracting attention, and is considered as an important issue in medical imaging because of the limited training data [10]-[14]. ImageNet or other non-medical image datasets are widely used for pre-training.

III. Proposed Methods

A. 2.5D U-Net

In our study, though we also use a U-Net-based network for remaining location information of the ventricle, which is in the central part of the brain, we propose a 2.5D U-Net that uses multi-slice features as input to keep the spatial information in the slice direction. Since the proposed method is based on the 2D model, the method is possible to effectively segment 3D tissues without overfitting.

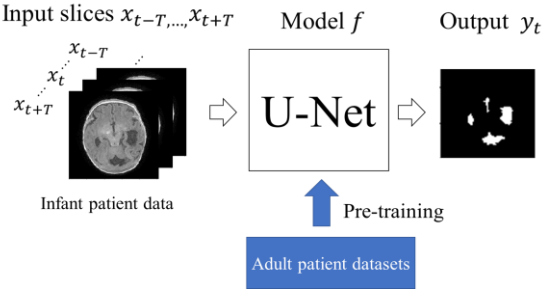


Figure 3. Overview of 2.5D U-net with TL.

The overview of the proposed 2.5D U-Net is indicated in Fig. 3. It has 2×T+1 input channels, which corresponds to the target slice and its neighboring slices (-T ~T slices). The input-output function of the proposed 2.5D network is shown in (1):

$$y_t = f(x_{t-T}, \ldots, x_{t+T} | \boldsymbol{\theta}) \tag{1}$$

where y_t is the output of the target slice t, x_t is the input target slice, T is the number of neighboring slices, which is an arbitrary constant and determines how many neighbored slices are included in the input to the network. $\boldsymbol{\theta}$ is the model parameters.

We apply focal loss, which enables efficient learning with fewer data and prevents overfitting with complex and diverse datasets [15], while binary cross-entropy, in which each pixel uses the same weight, leads to a poorly trained network with insufficient data. Here, more efficient learning is required owing to the complex and diverse datasets. Therefore, we apply a focal loss that reduces the weights of the high-probability regions and places the focus on learning in difficult regions, such as the complex boundaries of hydrocephalus infant brain ventricles; this can be expressed as (2),

$$\begin{aligned} \text{Focal Loss} = &-\sum_{i \in P} \alpha (1-p_i)^{\gamma} \log(p_i) \\ &-\sum_{i \in N} (1-\alpha) p_i^{\gamma} \log(1-p_i) \end{aligned} \tag{2}$$

here, α and γ are constants, p_i is the probability at pixel i, and P and N are label sets (positive and negative, respectively). By changing α, the weight of each label is adjusted, whereas γ adjusts the focus degree applied to the lower-probability area.

B. Transfer Learning with Adult Datasets

Medical imaging often deals with abnormal cases, unlike natural images. There are complex and diverse data, and situations with limited data often occur. This makes it difficult to obtain features for machine learning and may lead to overfitting. In order to overcome this problem, Transfer Learning (TL) using ImageNet or other non-medical image dataset is widely used to pre-train the network. After the pre-training, the limited medical training data are used for fine-tuning. In this study, we propose to use adult datasets for pre-training and fine-

tuning with target data from infant patients with hydrocephalus. This is the first experiment to demonstrate the effectiveness of applying TL to hydrocephalus infant brain ventricular segmentation. Compared with conventional ImageNet or other non-medical image dataset, the adult brain datasets have similar structure with the target infant brain and is more suitable for pre-training. When discussing transfer learning, we should consider for updating or frozen the weight of each layer of the network, and which dataset is effective with transfer learning for the task. However, the discussion will remain in the future, or refer to the [10]-[14], which indicate transfer learning is more accurate than the scratch. Therefore, all encoder decoder layer weights are fine-tuned in our experiments.

IV. Experiments and Results

In order to demonstrate the effectiveness of the proposed method, we conducted several experiments to compare the proposed 2.5D with the conventional K-mean [4], and U-Net [7]. We also compared cases with and without transfer learning (TL).

A. Experimental Settings

Table I shows the experimental settings. The dataset was provided by Kansai Medical University. We applied leave-one-out cross-validation on six infants patients with hydrocephalus and took the average value of segmentation accuracy for quantitative performance evaluation. The training and validation inputs had a volume size of 64×256×256. We used an Adam optimizer [16] for the optimization. The learning rate was 10^{-4}, the batch size was 4, and the epoch was 100. Batch normalization is applied after two convolution and dropout is applied after activation in the decoder. The dropout probability is 0.5 during training and 0.0 during testing.

TABLE I. Experimental Settings

Batch size	4
Epochs	100
Volume size	64 ×256 ×256
Pretraining dataset	30 adults
Target dataset	6 infants
Optimizer	Adam [16]
α (focal loss)	0.25
γ (focal loss)	2

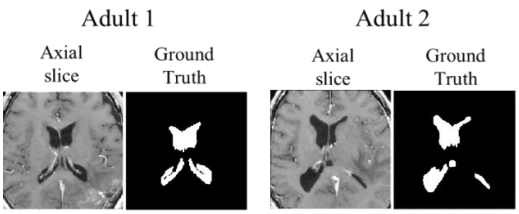


Figure 4. Axial slice of adult data for transfer learning

The setting of transfer learning using adult patient data is as same as fine-tuning with hydrocephalus infant patients. As shown in Fig. 4, they are adult patient's brain. During Pre-training, total or 30 volumes are used, each volume is used by cropping 64×256×256, and the number of epochs, optimizer, batch size, loss and parameters are the same as fine tuning of infant hydrocephalus patients.

B. Evaluation Metrics

We used two quantitative measures for the evaluation: DICE and the Volume Ratio (VR). DICE is a measure of the similarity of the ground truth to the prediction, and VR is the ratio of the ground truth and the predicted volume used to quantitatively evaluate the ventricular volume change before and after surgery. The equations for these metrics are as follows (3) and (4):

$$\mathrm{DICE} = \frac{2\sum_i P_i \cdot G_i}{\sum_i P_i^2 + \sum_i G_i^2} \tag{3}$$

$$\mathrm{VR} = \frac{V_P}{V_G} \tag{4}$$

where P_i is the prediction result and G_i is the ground truth, each being a binary vector, and V_P and V_G represent the volumes of the prediction and ground truth, respectively.

C. Results

Table II shows the accuracy of each model. The proposed 2.5D U-Net achieved higher DICE than U-Net. In particular, T=2, 3, which increase neighbored slices is higher DICE than T=1, while DICE of T=4 is slightly worse than T=3. From these results, it can be said while a model that gives correlation by considering neighbored slices is effective, increasing the number of input slices is not enough to cause that overfitting may occur due to the number of model parameters increases. Therefore, it is necessary to adjust the effective number of slices according to the degree of correlation in the slice direction for each target. As shown in Fig. 5, the 2.5D U-Net has less noise than the U-Net, and the segmented region is more clearly. As shown in Fig. 6, while U-Net has a lot of noise segmented and structural defects, 2.5D U-Net is continuous in the slice direction and can be segmented as a lump.

TABLE II. DICE and VR for Each Model

Model	DICE	VR
K-means++ [4]	0.53±0.18	3.30±2.61
U-Net [8]	0.58±0.11	1.14±0.27
2.5D U-Net ($T = 1$)	0.61±0.19	1.21±0.46
2.5D U-Net ($T = 2$)	0.62±0.21	1.14±0.49
2.5D U-Net ($T = 3$)	0.66±0.17	1.17±0.42
2.5D U-Net ($T = 4$)	0.64±0.16	1.44±0.62
2.5D U-Net ($T = 1$) + TL	**0.72**±0.11	**0.99**±0.09
2.5D U-Net ($T = 2$) + TL	0.68±0.15	1.16±0.23
2.5D U-Net ($T = 3$) + TL	0.69±0.12	1.11±0.24
2.5D U-Net ($T = 4$) + TL	0.66±0.18	1.09±0.22

In case that effectiveness of transfer learning with the adult dataset, as shown in Table II, the DICE of each model with transfer learning is higher than that without transfer learning. In addition, the DICE of 2.5D U-Net (T=1) with TL is higher than those of other U-Net + TL (T=2, 3, 4). Therefore, it can be said that the transfer learning using adult patients is effective for segmentation

of infant patients because the network learned the general features of the ventricle as prior knowledge via the pre-training. The transfer learning enables to perform stable learning and improve the segmentation accuracy. The reason why T=1 is more accurate than T=2 , T=3 and T=4 is the difference in ventricular characteristics between adult patients and infant hydrocephalus patients. In the case of T=2, T=3 and T=4 the feature of adults is captured strongly by using more parameters in the pre-training. However, since ventricular features differ between adults and infant hydrocephalus. Then T=1 , which can effectively perform fine-tuning with the appropriate number of parameters than the others. In addition, as shown in Fig. 5, the segmentation of 2.5D U-Net using transfer learning is more stable and clearer. As shown in Fig. 6, the 3D visualization also indicates that 2.5D U-Net with transfer learning is closest to the ground truth and can be segmentation stably.

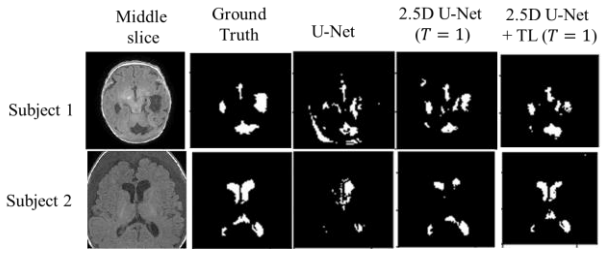


Figure 5. Qualitative results for the middle axial slices of the selected models.

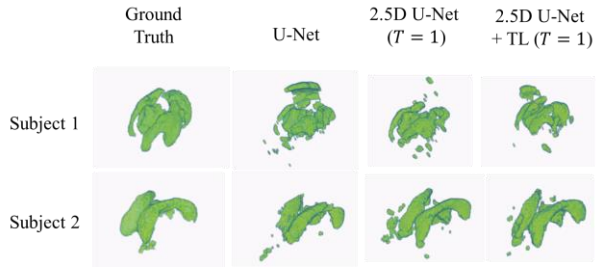


Figure 6. 3D visualization of the selected models.

V. Conclusions

In this paper, we proposed a 2.5D U-Net model and TL method with adult datasets for segmentation of hydrocephalus infant ventricular. The proposed method can include the context information along the z axis with a 2D network architecture resulting accurate 3D segmentation even with limited training samples. In our experiments, our proposed method achieved better results (DICE was improved from 58% to 72%) compared to conventional segmentation methods.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Kenji Ono designed and conducted the study, wrote the initial draft of the manuscript and analysis the data; Yutaro Iwamoto and Yen-Wei Chen assisted in the research, and the preparation of the manuscript; Masahiro Nonaka have contributed to data collection and interpretation; All authors approved the final version of the manuscript, and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgement

This work is supported in part by the Grant-in Aid for Scientific Research from the Japanese Ministry for Education, Science, Culture and Sports (MEXT) under the Grant Nos. 18H03267, 18K18078.

References

[1] J. G. Mandell, *et al.*, "Volumetric brain analysis in neurosurgery: Part 2. Brain and CSF volumes discriminate neurocognitive outcomes in hydrocephalus," *Journal of Neurosurgery: Pediatrics*, vol. 15, no. 2, pp. 125-132, 2015.

[2] B. B. O'hayon, *et al.*, "Frontal and occipital horn ratio: A linear estimate of ventricular size for multiple imaging modalities in pediatric hydrocephalus," *Pediatric Neurosurgery*, vol. 29, no. 5, pp. 245-249, 1998.

[3] J. Ashburner and K. J. Friston, "Unified segmentation," *Neuroimage*, vol. 26, no. 3, pp. 839-851, 2005.

[4] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proc. the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2007.

[5] F. Yepes-Calderon, M. D. Nelson, and J. G. McComb, "Automatically measuring brain ventricular volume within PACS using artificial intelligence," *PloS One*, vol. 13, no. 3, p. e0193152, 2018.

[6] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431-3440.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015.

[8] Ö. Çiçek, *et al.*, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2016.

[9] P. Mlynarski, *et al.*, "3D convolutional neural networks for tumor segmentation using long-range 2D context," *Computerized Medical Imaging and Graphics*, vol. 73, pp. 60-72, 2019.

[10] V. Cheplygina, "Cats or CAT scans: Transfer learning from natural or medical image source datasets?" *Current Opinion in Biomedical Engineering*, 2019.

[11] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299-1312, 2016.

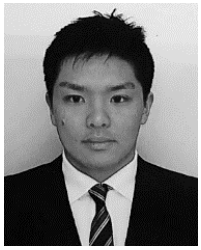
[12] S. Chen, K. Ma, and Y. Zheng, "Med3D: Transfer learning for 3D medical image analysis," arXiv preprint arXiv:1904.00625, 2019.

[13] K. Hara, H. Kataoka, and Y. Satoh, "Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet?" in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

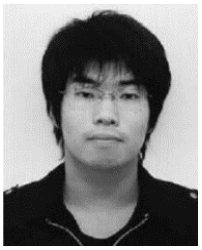
[14] H. Ravishankar, P. Sudhakar, R. Venkataramani, S. Thiruvenkadam, P. Annangi, N. Babu, and V. Vaidya, "Understanding the mechanisms of deep transfer learning for medical images," in *Deep Learning and Data Labeling for Medical Applications*, Springer, 2016, pp. 188-196.

[15] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE International Conference on Computer Vision*, 2017.

[16] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015.



Kenji Ono is a master student in information science and engineering at the Ritsmeikan University. He received his B.E. from Ritsumeikan University in 2018. His research interests are image processing, machine learning and medical image processing.



Yutaro Iwamoto is assistant professor in information science and engineering at the Ritsmeikan University. He received his B.E., M.E. and D.E. from Ritsumeikan University in 2011, 2013 and 2017, respectively. His research interests are image processing, machine learning and medical image processing.

Yen-Wei Chen is professor in information science and engineering at the Ritsmeikan University. He received his B.E., M.E. and D.E. from Kobe University and Osaka University in 1985, 1987 and 1990, respectively. His research interests are pattern recognition, image processing, machine learning and medical image processing.

Masahiro Nonaka is professor in neurosurgery (pediatric neurosurgery) at the Kansai Medical University. He received his M.D. and Ph.D. from Osaka University in 1992 and 2001, respectively. His research interests are congenital malformation, brain tumor and image guided neurosurgery.