# Unsupervised Domain Adaptation for Medical Image Segmentation Using Gradient Reversal Layer

Satoshi Kondo[1][0000-0002-4941-4920]

[1] Muroran Institute of Technology, Hokkaido 050-8585, Japan
kondo@mmm.muroran-it.ac.jp

**Abstract.** This is a write-up of our method submitted to Cross-Modality Domain Adaptation Challenge held in MICCAI2021 conference.

**Keywords:** Domain Adaptation, Segmentation, Gradient Reversal Layer.

## 1 Introduction

Unsupervised domain adaptation (UDA) is a type of domain adaptation and exploits labeled data from the source domain and unlabeled data from the target one [1, 2]. In the Cross-Modality Domain Adaptation for Medical Image Segmentation (crossMoDA) challenge held in MICCAI2021 conference, a large and multi-class dataset for unsupervised domain adaptation is introduced [3, 4]. In this challenge, contrast enhanced T1 MRI volumes for brain are provided as the source domain data, and high-resolution T2 MRI volumes are provided as the target domain data. The segmentation targets are vestibular schwannoma (VS) and the cochlea.

## 2 Proposed Method

Our method is based on a standard 3D segmentation deep neural network. We first train the segmentation network with supervised learning using the source domain dataset. In the domain adaptation phase, the feature maps obtained by the encoder part of the segmentation network are fed into Gradient Reversal Layer (GRL) [5] in addition to the decoder part of the segmentation network. The output of the GRL is fed into three fully connected layers which is the domain classifier.

The loss function in the supervised learning phase $\mathcal{L}_S$ is weighted summation of cross entropy loss $\mathcal{L}_c$ and Dice loss $\mathcal{L}_d$ as follows:

$$\mathcal{L}_S = \mathcal{L}_c + w_d \mathcal{L}_d \tag{1}$$

The loss function in the domain adaptation phase $\mathcal{L}_D$ is summation of the supervised loss as in Eq. (1) for source domain data and the domain classification loss $\mathcal{L}_b$, which is the binary cross entropy loss, for both of source and target data, as follows:

$$\mathcal{L}_D = \mathcal{L}_S + \mathcal{L}_b \tag{2}$$

We train two separate networks for VS and cochlea segmentation. The input data to each network is cubic patch extracted from the whole volumes.

## 3 Experimental Conditions

We employ 3D version of ENet [6] as the segmentation network. The segmentation network is first trained with samples from the source domain as mentioned in Section 2. Ninety-five samples (case ID 1 - 95) in the source domain are used as the training data and ten samples (case ID 96 - 105) are used as the validation data. We then train the whole network, i.e. the segmentation network, GRL and domain classification network, with samples from the source and target domains. One hundred and five samples (case ID 106 - 210) in the target domain are used as the training data. Each batch in the domain adaptation phase is composed the same number of the samples from the source and target domains.

As pre-processing, an input volume is resampled at 0.5 mm sampling pitch for all dimensions and normalized with mean and standard deviation values of each volume.

The number of epochs is 60 for both the supervised learning and domain adaptation phases. The initial learning rate are 1.0e-3 and 6.7e-5 for the supervised learning and domain adaptation phases, respectively. The optimizer is Adam [7] and the learning rate changes with cosine annealing. The weight $w_d$ in Eq. (1) is 3.4. The learning rates and the weight are optimized with hyper-parameter tuning using Optuna library [8]. The search range of the hyper-parameters are from 1e-5 to 0.1, from 1e-10 to 0.1 and from 0.1 to 4.0 for the initial learning rate in the supervised learning phase, the initial learning rate in the domain adaptation phase and the weight $w_d$, respectively. The batch sizes are 4 and 16 for VS and cochlea segmentation, respectively. The sizes of the cubic patch are 128 and 64 for VS and cochlea segmentation, respectively.

We employ the model showing the highest Dice value for the validation data in the source samples for the adaptation phase.

## 4 Results

The results for the validation dataset are shown in Table 1.

**Table 1.** Table captions should be placed above the tables.

| Structure | Dice score | ASSD |
|---|---|---|
| VS | $0.26 \pm 0.31$ | $11.4 \pm 11.2$ |
| Cochlea | $0.25 \pm 0.25$ | $11.7 \pm 11.3$ |

## References


1. Toldo, M., Maracani, A., Michieli, U., Zanuttigh, P.: Unsupervised Domain Adaptation in Semantic Segmentation: A Review. Technologies 8, 35 (2020).
2. Kouw, W. M., Loog, M.: A Review of Domain Adaptation without Target Labels. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(3), 766-785 (2021).
3. Shapey, J., et al: An artificial intelligence framework for automatic segmentation and volumetry of Vestibular Schwannomas from contrast-enhanced t1-weighted and high-resolution t2-weighted MRI. In: Journal of Neurosurgery JNS. (2020).
4. Shapey, J., Kujawa, A., Dorent, R., Wang, G., Bisdas, S., Dimitriadis, A., Grishchuck, D., Paddick, I., Kitchen, N., Bradford, R., Saeed, S., Ourselin, S., & Vercauteren, T.: Segmentation of vestibular schwannoma from MRI − An open annotated dataset and baseline algorithm, Scientific Data (2021).
5. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Lempitsky, V.: Domain-adversarial training of neural networks. The journal of machine learning research, 17(1), 2096-2030 (2016).
6. Paszke, A., Chaurasia, A., Kim, S., & Culurciello, E.: Enet: A deep neural network architecture for real-time semantic segmentation. arXiv:1606.02147 (2016).
7. Kingma, D. P., Ba, J.: Adam: A method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings (2015).
8. Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M.: Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2623-263 (2019).