Multimodal Brain Tumor Segmentation Using Contrastive Learning Based Feature Comparison with Monomodal Normal Brain Images

Huabing Liu[1], Dong Nie[2], Dinggang Shen[3,4], Jinda Wang[5], and Zhenyu Tang[1(✉)]

[1] School of Computer Science and Engineering, Beihang University,
Beijing 100191, China
tangzhenyu@buaa.edu.cn

[2] Alibaba Inc., Hangzhou, China

[3] School of Biomedical Engineering, ShanghaiTech University,
Shanghai 201210, China

[4] Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200230, China

[5] Sixth Medical Center, Chinese PLA General Hospital, Beijing 100853, China

Abstract. Many deep learning (DL) based methods for brain tumor segmentation have been proposed. Most of them put emphasis on elaborating deep network's internal structure to enhance the capacity of learning tumor-related features, while other valuable related information, such as normal brain appearance, is often ignored. Inspired by the fact that radiologists are often trained to compare with normal tissues when identifying tumor regions, in this paper, we propose a novel brain tumor segmentation framework by adopting normal brain images as reference to compare with tumor brain images in the learned feature space. In this way, tumor-related features can be highlighted and enhanced for accurate tumor segmentation. Considering that the routine tumor brain images are multimodal while the normal brain images are often monomodal, a new contrastive learning based feature comparison module is proposed to solve incomparable issue between features learned from multimodal and monomodal images. In the experiments, both in-house and public (BraTS2019) multimodal tumor brain image datasets are used to evaluate our proposed framework, demonstrating better performance compared to the state-of-the-art methods in terms of Dice score, sensitivity, and Hausdorff distance. Code: https://github.com/hbliu98/CLFC-Brain-Tumor-Segmentation.

Keywords: Brain tumor segmentation · Normal brain images · Feature comparison · Contrastive learning

1 Introduction

Brain tumor segmentation using multimodal magnetic resonance (MR) images is an essential task for subsequent diagnosis and treatment. In the past few years,

L. Wang et al. (Eds.): MICCAI 2022, LNCS 13435, pp. 118–127, 2022.
https://doi.org/10.1007/978-3-031-16443-9_12

many deep learning (DL) based segmentation methods have been proposed and achieved great success [6,11,13,20]. For example, Dong et al. [6] proposed a 2D U-Net [15] for end-to-end brain tumor segmentation, where the soft Dice loss is introduced to handle unbalanced training samples. To utilize volumetric information and multi-scale contextual information, Kamnitsas et al. [11] proposed a 3D convolutional neural network (CNN) with dual-pathway architecture called DeepMedic, to extract features from tumor brain images at multiple scales simultaneously. Wang et al. [20] further integrated Transformers [17] into 3D U-Net to build long-range dependency and learn global semantic features, based on which the segmentation performance can be enhanced.

Although existing DL-based methods have shown promising results, most of them focused on improving the learning capacity of tumor-related features by elaborating deep network's internal structure, while other valuable related information, such as normal brain appearance, is often ignored [18,19]. It is known that radiologists are often trained to compare with normal tissues when identifying tumor regions. Following this observation, the anomaly detection has been introduced for tumor segmentation [1,3]. These methods compare pathological images with their respective normal appearance images reconstructed by the autoencoder [16]. In this way, pathological regions can be highlighted and easily segmented. Note that existing anomaly detection based methods usually work with monomodality [2]. However, in the context of brain tumor segmentation, routine tumor brain images are multimodal, e.g., T1, T1 contrast-enhanced (T1c), T2, and FLAIR MR images, while normal brain images, which are used to train the autoencoder, are often monomodal, e.g., T1 MR images. Comparing multimodal images with monomodal images is difficult, therefore anomaly detection based methods cannot be directly applied to multimodal brain tumor segmentation.

In this paper, we propose a novel multimodal brain tumor segmentation framework, where monomodal normal brain images are adopted as reference and compared with multimodal tumor brain images in the feature space to impel the segmentation performance. To solve the incomparable issue between features learned from different modalities, a new Contrastive Learning based Feature Comparison (CLFC) module is proposed to align the features learned from monomodal normal brain images (normal brain features) to the features learned from multimodal tumor brain images (tumor brain features) at normal brain regions, i.e., non-tumor regions. In this way, tumor regions in tumor brain features, i.e., tumor-related features, can be effectively highlighted and enhanced using normal brain features as reference. Our proposed framework is evaluated using in-house and public (BraTS2019) tumor brain image datasets. Experimental results show that our proposed framework outperforms the state-of-the-art methods in terms of Dice score, sensitivity, and Hausdorff distance in both datasets. The contributions of our work can be summarized as follows:

- We propose a novel deep framework for multimodal brain tumor segmentation, where external information, i.e., monomodal normal brain images, is utilized as reference to impel the segmentation performance.

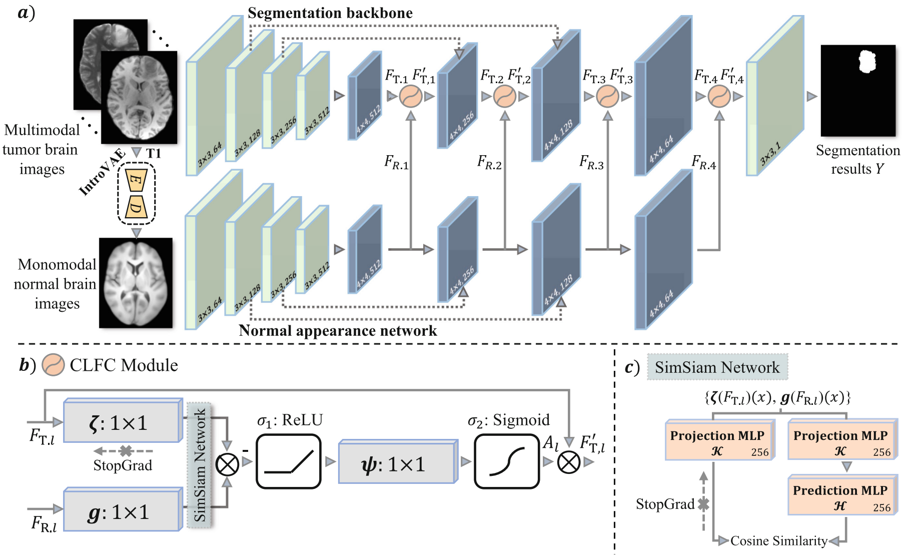

Fig. 1. a) The proposed framework, composed of a segmentation backbone and a normal appearance network; b) Structure of CLFC module; and c) Structure of SimSiam. Kernel sizes and output feature channels are marked at the bottom of each layer.

- A contrastive learning based feature comparison (CLFC) module is designed to solve the incomparable issue between features learned from multimodal tumor brain images and monomodal normal brain images.

2 Method

Our proposed framework is composed of two sub-networks: 1) the segmentation backbone and 2) the normal appearance network, as shown in Fig. 1a. The input of the segmentation backbone is multimodal tumor brain images, while the normal appearance network takes monomodal (T1) normal brain images as the input. The monomodal normal brain images are produced by introspective variational autoencoder (IntroVAE) [8] from the T1 modality contained in the input multimodal tumor brain images. The segmentation backbone and the normal appearance network are of encoder and decoder structures, where each encoder is composed of four convolution layers followed by batch normalization, ReLU and maxpooling layer, and each decoder has four transposed convolution layers to perform upsampling. At the end of the segmentation backbone, an extra convolution layer is designed to produce final segmentation results. In both sub-networks, at each level l of both the decoders, the tumor brain features $F_{\mathrm{T},l}$ learned by the segmentation backbone and the normal brain features $F_{\mathrm{R},l}$ learned by the normal appearance network are aligned and compared by the CLFC modules to produce $F'_{\mathrm{T},l}$, which contains enhanced tumor-related features, making the final segmentation to a high accuracy level. Considering the high computational efficiency of 2D convolution and also the advantage of 3D spatial information, 2.5D slices containing the slice to be segmented and its K adjacent slices (i.e., $2K+1$ slices) are adopted as the input of our framework.

2.1 The Normal Appearance Network

As aforementioned, the input monomodal normal brain images is set to T1 MR images, which is a commonly adopted modality for normal brain imaging in the context of clinical routine. Instead of using normal brain images from healthy subjects, the monomodal normal brain images are reconstructed from the T1 modality contained in the input multimodal tumor brain images using IntroVAE. In this way, except tumor regions, the reconstructed normal brain images have similar anatomical structure as the original tumor brain images [12,21].

The IntroVAE. The IntroVAE extends traditional VAE to an adversarial learning framework, by which the reconstructed images are of detailed structures. It is composed of an encoder for projecting input brain images to a latent distribution and a decoder for producing reconstructed brain images based on the learned distribution. In the training phase, T1 normal brain MR images from healthy subjects are used to train the IntroVAE, by which low-dimensional manifold representing normal brain appearance can be obtained. In the inference phase, the trained model projects the input tumor brain image to a certain point on the manifold, which represents the closest normal brain appearance. Finally, the corresponding normal brain image can be reconstructed from the point. For more details of the IntroVAE, please refer to the original paper [8].

Based on the reconstructed monomodal normal brain images, normal brain features can be learned in the normal appearance network. At each level of the decoder in the normal appearance network, the learned normal brain features $F_{\mathrm{R},l}$, $l = 1, ..., 4$ are sent to the segmentation backbone as reference to enhance the segmentation performance (discussed below).

2.2 The Segmentation Backbone

The segmentation backbone takes concatenated multimodal tumor brain images as input, and outputs tumor segmentation results. To improve segmentation results, at each decoding level of the segmentation backbone, the learned tumor brain features $F_{\mathrm{T},l}$ are compared with the learned normal brain features $F_{\mathrm{R},l}$, by which tumor regions in $F_{\mathrm{T},l}$, i.e., those tumor-related features, can be highlighted and enhanced. Ideally, for feature vectors $F_{\mathrm{T},l}(x)$ at position x, which are at normal/tumor regions, $F_{\mathrm{R},l}(x)$ should have consistent/inconsistent features at the corresponding positions. In this way, tumor regions in $F_{\mathrm{T},l}$ can then be effectively highlighted according to the feature consistency. Unfortunately, in the context of multimodal brain tumor segmentation, $F_{\mathrm{T},l}$ and $F_{\mathrm{R},l}$ are incomparable (multimodal vs. monomodal).

To tackle the above issue, we propose a new contrastive learning based feature comparison (CLFC) module as shown in Fig. 1b. The CLFC module is composed of two main steps. The *first* step is feature alignment, where two 1×1 convolution layers $\boldsymbol{\zeta}$ and $\boldsymbol{g}$ are adopted to align $F_{\mathrm{T},l}$ and $F_{\mathrm{R},l}$ at normal regions. As a result, feature vectors at normal regions are consistent in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ and $\boldsymbol{g}(F_{\mathrm{R},l})$. To achieve effective feature alignment, a contrastive learning method called the simple Siamese network (SimSiam) [4] is adopted (see Fig. 1c). The SimSiam network is composed of an encoder $\mathcal{F}$, a projection MLP $\mathcal{K}$, and a prediction MLP $\mathcal{H}$. It takes each positive sample pair $\{I_1, I_2\}$ (samples of the same class) as input and maximizes the cosine similarity of output vectors p_1 and v_2:

$$D(I_1, I_2) = -\frac{p_1}{||p_1||_2} \cdot \frac{v_2}{||v_2||_2}, \tag{1}$$

where $p_1 = \mathcal{H}(\mathcal{K}(\mathcal{F}(I_1)))$, $v_2 = \mathcal{K}(\mathcal{F}(I_2))$ and $||\cdot||_2$ is the L2-norm. In our framework, feature vector pairs $\{F_{\mathrm{T},l}(x), F_{\mathrm{R},l}(x)\}$, $x \in \Omega_{\mathrm{NR}}$, where Ω_{NR} denotes normal regions, are defined as positive samples. Note that Ω_{NR} is determined according to the manually labeled tumor mask in the training dataset. In the CLFC module, $\boldsymbol{\zeta}$ and $\boldsymbol{g}$ are regarded as the encoder $\mathcal{F}$ of the SimSiam network, so $\{\boldsymbol{\zeta}(F_{\mathrm{T},l})(x), \boldsymbol{g}(F_{\mathrm{R},l})(x)\}$ are sent to the projection and prediction MLPs to compute the cosine similarity. The loss of the SimSiam network is defined as:

$$\mathcal{L}_{\mathrm{Sim}}(F_{\mathrm{T},l}, F_{\mathrm{R},l}) = \frac{1}{2} \sum_{x \in \Omega_{\mathrm{NR}}} D(F_{\mathrm{T},l}(x), F_{\mathrm{R},l}(x)) + D(F_{\mathrm{R},l}(x), F_{\mathrm{T},l}(x)). \tag{2}$$

It is worth noting that the SimSiam network is used in the training phase but removed in the inference phase. Moreover, to make the segmentation backbone focus on tumor segmentation rather than feature alignment, a stop gradient operation is also applied before $\boldsymbol{\zeta}$ as shown in Fig. 1b.

After the feature alignment step, the *second* step is feature comparison. Since feature vectors in the aligned $F_{\mathrm{T},l}$ and $F_{\mathrm{R},l}$, i.e., $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ and $\boldsymbol{g}(F_{\mathrm{R},l})$, are consistent at normal regions, tumor regions in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ can be easily identified using $\boldsymbol{g}(F_{\mathrm{R},l})$ as reference. Specifically, we measure the consistency between $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ and $\boldsymbol{g}(F_{\mathrm{R},l})$ and identify the regions of low feature consistency, i.e., $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ at tumor regions vs. $\boldsymbol{g}(F_{\mathrm{R},l})$ at normal regions, to produce the attention map A_l. By using A_l as the mask, tumor-related features in $F_{\mathrm{T},l}$ can be enhanced. The whole procedure of the CLFC module can be summarized as:

$$\begin{aligned} A_l &= \sigma_2(\boldsymbol{\psi}(\sigma_1(-\boldsymbol{\zeta}(F_{\mathrm{T},l}) \otimes \boldsymbol{g}(F_{\mathrm{R},l})))), \\ F'_{\mathrm{T},l} &= A_l \cdot F_{\mathrm{T},l}, \end{aligned} \tag{3}$$

where σ_1 is ReLU, σ_2 is sigmoid activation function, $\boldsymbol{\psi}$ is the 1×1 convolution layer, and $F'_{\mathrm{T},l}$ is the output of the CLFC module containing enhanced tumor-related features. The final segmentation results Y are produced from $F'_{\mathrm{T},4}$ by the last convolution layer in the segmentation backbone. The Dice loss [5] is adopted for Y, and the final loss function of our framework is $\mathcal{L} = \mathcal{L}_{\mathrm{Dice}} + \mathcal{L}_{\mathrm{Sim}}$.

3 Experiments

Both in-house and public datasets are used to evaluate our proposed framework. Specifically, the in-house dataset contains multimodal tumor brain MR images of T1c, B0, mean diffusivity (MD), and fractional anisotropy (FA) modalities from 104 glioblastoma patients. The public dataset is BraTS2019, which includes tumor brain T1, T1c, T2 and FLAIR brain MR images of 335 glioma patients. Manually labeled tumor masks are available for each patient in both datasets. All images are aligned with MNI152 [7] using affine transformation and normalized by histogram matching. Besides our framework, the state-of-the-art segmentation method nnU-Net [9] is also evaluated. Moreover, ablation experiments of our framework are conducted. Specifically, the Baseline-1 uses only the segmentation backbone of our framework (no normal appearance network), and the Baseline-2 is similar to our framework, but it has no contrastive learning (no SimSiam network) in the CLFC module during training. For both datasets, five-fold patient-wise cross validation is adopted to evaluate each method. The input of all methods under evaluation is 2.5D slices with $K = 2$, i.e., five slices. The batch size is set to four, and the maximal number of epochs is 300. The accuracy of tumor segmentation results is quantified using patient-wise Dice score, sensitivity, and 95% Hausdorff distance (HD95). All methods are implemented using PyTorch and trained with RTX 3090 GPU.

It is worth noting that the IntroVAE is separately trained with public dataset IXI [10] containing 581 normal brain T1 MR images. Specifically, the encoder and decoder of IntroVAE are trained iteratively with the learning rates of $1\ \times 10^{-4}$ and $5\ \times 10^{-3}$, respectively. The batch size and number of epochs are set to 120 and 200. Other hyper-parameters, like weight terms in the loss function, keep the same as the original paper. After integration into our framework, the parameters of the IntroVAE is fixed during the training of our framework. Moreover, for in-house dataset, which has no T1 modality, we adopt tumor brain T1c MR images as input of IntroVAE to get the reconstructed normal brain T1 MR images, as T1c is T1 with contrast agent in the vessel and both modalities exhibit similar appearance of gray matter and white matter at normal regions [14].

3.1 Evaluation of Segmentation Results

Figure 2 shows some examples of the tumor segmentation results using each method under evaluation. It is clear that as benefiting from feature comparison using normal brain images as reference, our framework can detect subtle normal and tumor regions, which are hard to distinguish using tumor brain images alone, especially in the regions marked in red circles in Fig. 2.

Details of the evaluation results are shown in Table 1, and our framework outperforms all the other methods under evaluation in terms of Dice score, sensitivity, and HD95 in both datasets. The patient-wise Wilcoxon signed rank test is adopted to compare the Dice scores from different methods. For the in-house/public dataset, the p values are 0.0185/0.0076 (Ours vs. nnU-Net), $1.0320 \times 10^{-4}/2.7976 \times 10^{-7}$ (Ours vs. Baseline-1), and $7.7959 \times 10^{-4}/4.9657 \times 10^{-4}$

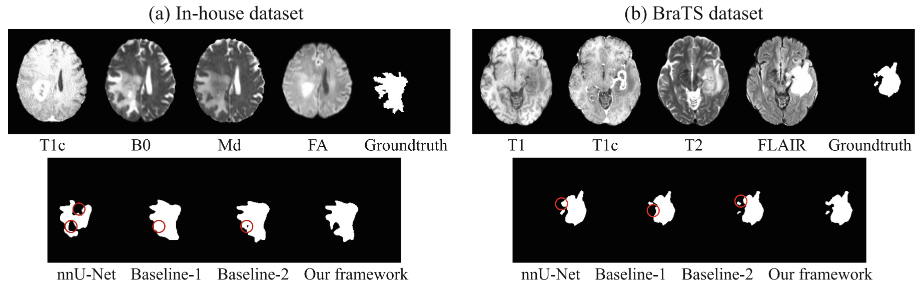

Fig. 2. Examples of the segmentation results using all methods under evaluation. Our framework can detect subtle normal and tumor regions, which cannot be recognized by the other methods (as marked in red circles). (Color figure online)

(Ours vs. Baseline-2), respectively. Baseline-2 achieves better segmentation accuracy than Baseline-1, which shows that the normal appearance network plays a positive role in tumor segmentation. But without the contrastive learning, $F_{\mathrm{T},l}$ and $F_{\mathrm{R},l}$ are difficult to be compared with each other, and the improvement using Baseline-2 is limited. Our framework adopts the contrastive learning for feature alignment in the CLFC module, by which $F_{\mathrm{T},l}$ and $F_{\mathrm{R},l}$ can be well aligned and the segmentation performance is significantly enhanced.

Table 1. Evaluation results of tumor segmentation.

Method	In-house dataset			BraTS dataset		
	Dice (%)	Sensitivity (%)	HD95 (mm)	Dice (%)	Sensitivity (%)	HD95 (%)
nnU-Net	87.44 ± 7.35	85.99 ± 12.28	6.07 ± 3.53	90.53 ± 7.24	89.51 ± 9.82	8.05 ± 17.22
Baseline-1	87.90 ± 3.27	88.86 ± 5.36	6.47 ± 5.36	90.54 ± 7.71	90.00 ± 11.56	4.63 ± 4.96
Baseline-2	88.66 ± 3.61	87.60 ± 6.28	4.74 ± 2.03	91.18 ± 4.90	90.79 ± 8.95	4.23 ± 4.21
Ours	**89.70 ± 3.68**	**90.85 ± 5.50**	**4.29 ± 1.81**	**91.86 ± 4.63**	**91.90 ± 8.02**	**3.84 ± 3.04**

3.2 Evaluation of Contrastive Learning Based Feature Comparison

In the CLFC module, the contrastive learning plays an important role. To give an intuitive visualization of the effect using contrastive learning, distributions of feature vectors in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ and $\boldsymbol{g}(F_{\mathrm{R},l})$ are shown in Fig. 3.

Specifically, feature vectors are divided into three types: tumor regions in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$, normal regions in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$, and normal regions in $\boldsymbol{g}(F_{\mathrm{R},l})$. All feature vectors are projected onto a 2D plane using the PCA based dimension reduction. Clearly, as compared with Baseline-2 (without contrastive learning), in our framework (with contrastive learning), feature vectors at normal regions in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$ and $\boldsymbol{g}(F_{\mathrm{R},l})$ can be more effectively aligned and are more different from the feature vectors at tumor regions in $\boldsymbol{\zeta}(F_{\mathrm{T},l})$. As a result, tumor regions can be easily identified through the feature comparison. Figure 4 shows some examples of the attention maps A_l produced at each decoding level of Baseline-2 and

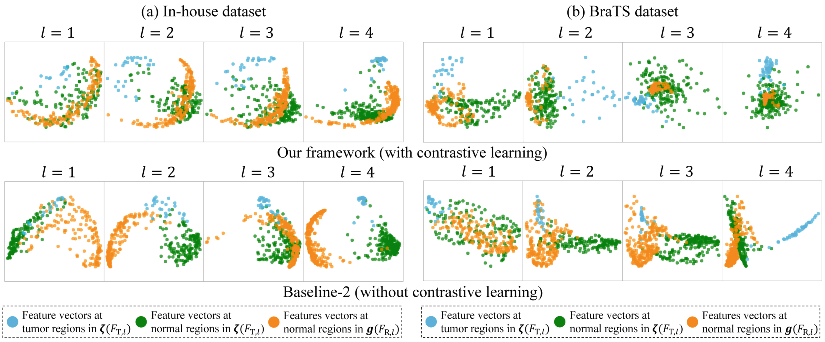

Fig. 3. Distributions of feature vectors in our framework and Baseline-2.

our framework. It is clear that the attention maps produced in our framework has more concentrated tumor regions, and the segmentation results are more consistent with the ground truth than Baseline-2.

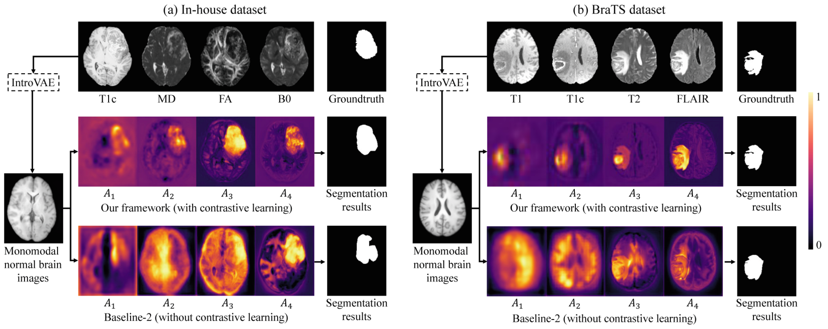

Fig. 4. Examples of the attention maps A_l in our framework and Baseline-2.

4 Conclusion

We proposed a novel multimodal brain tumor segmentation framework, where external information of normal brain appearance was used as reference to highlight and enhance the tumor-related features. Moreover, a contrastive learning based feature comparison (CLFC) module was proposed to address the incomparable issue between features learned from multimodal tumor brain images and monomodal normal brain images, based on which high-quality attention maps A_l as well as tumor-related features $F'_{T,l}$ can be produced for better segmentation results. Both in-house and public BraTS2019 datasets were used to evaluate our framework. The experimental results showed that our framework outperforms the state-of-the-art methods with statistical significance, and the proposed normal appearance network and the proposed CLFC module both play effective role in the segmentation. Since the CLFC module is effective for binary segmentation, our framework currently works with whole tumor segmentation. In the future work, we will extend our framework to segmentation of tumor sub-regions, e.g., edema, enhancing tumor, necrotic and non-enhancing tumor core.

References

1. Astaraki, M., Toma-Dasu, I., Smedby, Ö., Wang, C.: Normal appearance autoencoder for lung cancer detection and segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11769, pp. 249–256. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32226-7_28
2. Baur, C., Denner, S., Wiestler, B., Navab, N., Albarqouni, S.: Autoencoders for unsupervised anomaly segmentation in brain MR images: a comparative study. Med. Image Anal., 101952 (2021)
3. Baur, C., Wiestler, B., Albarqouni, S., Navab, N.: Deep autoencoding models for unsupervised anomaly segmentation in brain MR images. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds.) BrainLes 2018. LNCS, vol. 11383, pp. 161–169. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11723-8_16
4. Chen, X., He, K.: Exploring simple siamese representation learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 15750–15758 (2021)
5. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)
6. Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y.: Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In: Valdés Hernández, M., González-Castro, V. (eds.) MIUA 2017. CCIS, vol. 723, pp. 506–517. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60964-5_44
7. Evans, A., Collins, D., Milner, B.: An MRI-based stereotactic brain atlas from 300 young normal subjects, 408, Anaheim. In: Proceedings of the 22nd Symposium of the Society for Neuroscience (1992)
8. Huang, H., He, R., Sun, Z., Tan, T., et al.: Introvae: introspective variational autoencoders for photographic image synthesis. In: Advances in Neural Information Processing Systems 31 (2018)
9. 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)
10. IXI: Information extraction from images. www.brain-development.org
11. Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017)
12. Luo, Y., et al.: Adaptive rectification based adversarial network with spectrum constraint for high-quality pet image synthesis. Med. Image Anal. **77**, 102335 (2022)
13. Nie, D., Wang, L., Adeli, E., Lao, C., Lin, W., Shen, D.: 3-D fully convolutional networks for multimodal isointense infant brain image segmentation. IEEE Trans. Cybern. **49**(3), 1123–1136 (2019)
14. Radue, E.W., Weigel, M., Wiest, R., Urbach, H.: Introduction to magnetic resonance imaging for neurologists. Continuum Lifelong Learn. Neurol. **22**(5), 1379–1398 (2016)

15. 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
16. Siddiquee, M.M.R., et al.: Learning fixed points in generative adversarial networks: from image-to-image translation to disease detection and localization. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 191–200 (2019)
17. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems 30 (2017)
18. Wang, K., et al.: Semi-supervised medical image segmentation via a tripled-uncertainty guided mean teacher model with contrastive learning. Med. Image Anal. **79**, 102447 (2022)
19. Wang, L., Shi, F., Lin, W., Gilmore, J.H., Shen, D.: Automatic segmentation of neonatal images using convex optimization and coupled level sets. NeuroImage **58**(3), 805–817 (2011)
20. Wang, W., Chen, C., Ding, M., Yu, H., Zha, S., Li, J.: TransBTS: multimodal brain tumor segmentation using transformer. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 109–119. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87193-2_11
21. Xiang, L., et al.: Deep embedding convolutional neural network for synthesizing CT image from T1-weighted MR image. Med. Image Anal. **47**, 31–44 (2018)