Brain Tumor Segmentation

Subjects: Computer Science, Artificial Intelligence

Contributor: Mengxian Chi, Hong An, Xu Jin, Zhenguo Nie

Brain tumor segmentation plays a crucial role in the diagnosis, treatment planning, and monitoring of brain tumors. Accurate segmentation of brain tumor regions from multi-sequence magnetic resonance imaging (MRI) data is of paramount importance for precise tumor analysis and subsequent clinical decision making. The ability to delineate tumor boundaries in MRI scans enables radiologists and clinicians to assess tumor size, location, and heterogeneity, facilitating treatment planning and evaluating treatment response. Traditional manual segmentation methods are time-consuming, subjective, and prone to inter-observer variability. Therefore, the automatic segmentation algorithm has received widespread attention as an alternative solution. For instance, the selforganizing map (SOM) is an unsupervised exploratory data analysis tool that leverages principles of vector quantization and similarity measurement to automatically partition images into self-similar regions or clusters. Segmentation methods based on SOM have demonstrated the ability to distinguish high-level and low-level features of tumors, edema, necrosis, cerebrospinal fluid, and healthy tissue.

brain tumor segmentation CNNs feature pyramid lightweight model

1. Attention Mechanisms in Convolutional Neural Networks

The channel attention mechanism ^[1] focuses on enhancing important channel information while suppressing less relevant channels, achieved by computing attention weights along the channel dimension. It typically involves global average pooling and a multi-layer perceptron (MLP) to generate attention weights, which are then used to weight the original feature maps. On the other hand, the spatial attention mechanism 2 aims to highlight important spatial locations and suppress unimportant ones. It can be implemented through convolutional operations with different kernel sizes or using self-attention mechanisms to compute attention weights between spatial positions. These attention mechanisms, when combined, form the Convolutional Block Attention Module (CBAM) ^[3], which integrates both channel and spatial attention. Additionally, a recently proposed attention mechanism called coordinate attention [4] focuses on modeling the relationships between different positions in the feature maps by incorporating coordinate information. It utilizes an MLP to process the coordinate information and generates position weights, which are then multiplied with the original feature maps. The 3D structures of these commonly used attention mechanisms are illustrated in **Figure 1**. While these attention mechanisms enhance the modeling capability of CNNs by capturing important channel, spatial, and positional information, they also introduce additional computational overhead.



Figure 1. Attention mechanisms with (a) channel attention, (b) spatial attention, (c) CBAM, and (d) coordinate attention.

2. Single-Path and Multi-Path Convolutional Neural Networks for Brain Tumor Segmentation

CNNs have shown great promise in achieving accurate and efficient segmentation results, revolutionizing the way brain tumors are analyzed and diagnosed due to their ability to automatically learn discriminating features from input data ^[5]. Initially, single-path CNNs were employed, where a single data processing stream was utilized ^[6]. These networks, which take multi-modal brain tumor magnetic resonance imaging (MRI) scans as input, sequentially pass the data through a combination of convolutional layers, pooling layers, and non-linear activation layers, ultimately performing segmentation using a classifier at the end of the model. Single-path CNNs are characterized by their simplicity in structure and shallow hierarchy, but their segmentation performance may be sub-optimal. As brain tumor images are inherently complex and diverse, relying solely on a single processing path may limit the network's ability to capture and represent the intricate details present in different modalities.

To address this limitation, multi-path CNNs have been introduced, featuring multiple parallel convolutional pathways ^[Z]. This architecture allows for the processing of input information at multiple scales, providing a larger receptive field and the potential for enhanced segmentation accuracy. However, it is worth noting that multi-path CNNs tend to exhibit a higher level of complexity and require a larger model size to accommodate the increased number of pathways. Furthermore, an inherent challenge arises from the uneven distribution of tumor regions, where certain tumor areas may exhibit varying sizes and proportions compared to others.

To tackle this class imbalance issue, cascaded CNNs have been proposed as a potential solution ^[8]. By cascading multiple network models, each designed to segment a specific region of interest, cascaded CNNs enable the transformation of the multi-region tumor segmentation problem into a series of binary segmentation tasks. One of the key advantages of cascaded CNNs is their ability to consider the unique relationships between sub-regions when predicting subsequent segmentation tasks. This can be particularly beneficial in minimizing false positives, as each network operates on regions extracted from the output of the previous network. However, an important point

to consider is that cascaded CNNs, in contrast to single-path and multi-path CNNs, are not end-to-end and require additional time for training and testing due to the sequential nature of the cascaded segmentation process.

3. The U-Net and Its Variants for Brain Tumor Segmentation

The U-Net architecture consists of an encoder–decoder structure coupled with skip connections ^[9]. The encoder path incorporates a series of convolutional and pooling layers to progressively extract hierarchical features and reduce spatial resolution. The decoder path utilizes up-sampling and transposed convolutional layers to recover the spatial information and generate segmentation maps. Skip connections connect the corresponding encoder and decoder layers, allowing for the model to fuse low-level and high-level features. This design enables U-Net to capture both local and global contextual information, facilitating the accurate delineation of tumor boundaries.

Initially, the research focused on 2D segmentation networks operating within individual 2D image planes. U-Net [10] has demonstrated its efficacy in capturing tumor boundaries and distinguishing tumor regions from healthy brain tissue. U-Net++ [11] extends the U-Net architecture by incorporating nested and dense skip pathways, enabling the capture of multi-scale contextual information for precise brain tumor segmentation. SegResNet [12] combines U-Net architecture with the residual network (ResNet) to enhance feature representation and segmentation performance, effectively capturing both local and global contextual information. To further improve feature representation, DynU-Net ^[13] integrates a dynamic routing algorithm inspired by capsule networks into the U-Net architecture, enabling the capture of hierarchical relationships among different tumor regions. MS-Net ^[14] is a medical image segmentation technique based on a codec structure composed of a Multi-Scale Attention Module (MSAM) and a Stacked Feature Pyramid Module (SFPM). MSAM dynamically adjusts the receptive fields to capture different levels of context details, while SFPM adaptively increases the weight of the features of interest to focus the network's attention on the target region. Fusion factor ^[15] is introduced to control the amount of information transferred from deep to shallow layers in Feature Pyramid Networks (FPN) for tiny object detection. The paper explores how to estimate the effective value of the fusion factor for a specific dataset by statistical methods. However, these 2D networks may disregard the crucial depth information inherent in the MRI images, consequently impeding their ability to comprehensively utilize the rich local and global contextual information available.

Therefore, 3D U-Net ^[16] was developed to extend the U-Net framework for processing volumetric data, enabling the segmentation of brain tumors in 3D medical images. By considering spatial dependencies along the three dimensions, SCAR U-Net ^[17] improves the accuracy of tumor segmentation in volumetric scans. V-Net ^[18] is another extension of U-Net that incorporates a volumetric residual learning framework. It leverages 3D convolutional neural networks and residual connections to capture fine-grained details in volumetric data. The evolution of mainstream 2D segmentation networks into their 3D counterparts has resulted in significant improvements in brain tumor segmentation performance ^[19]. DSTGAN ^[20] presents a spatiotemporal generative adversarial learning approach for segmentation and quantification of myocardial infarction without contrast agents. The approach utilizes a generator and a discriminator module, which consist of three seamlessly connected networks to extract the morphological and motion abnormalities of the left ventricle, learn the complementarity between segmentation and quantification tasks, and leverage adversarial learning to enhance the accuracy of

estimation. However, it remains crucial to strike a balance between model complexity and computational feasibility, considering the practical constraints and available computational resources.

Simultaneously, the transformer ^[21] architecture has gained significant popularity in natural language processing (NLP) and has found applications in medical image analysis ^{[22][23]}. Initially developed for sequence modeling tasks, transformers have showcased their ability to capture long-range dependencies and capture contextual information effectively. Building upon this success, researchers have extended transformers to medical image analysis, leading to the emergence of models. UNETR [24] combines the transformer architecture with the U-Net framework, enabling the modeling of long-range dependencies and achieving state-of-the-art performance in brain tumor segmentation. Similarly, SwinUNETR ^[25] integrates the Swin Transformer, a hierarchical vision transformer, with the U-Net framework, effectively capturing global and local dependencies with reduced computational complexity. nnFormer ^[26] is a novel approach using a 3D transformer to segment medical images based on interleaved convolution and self-attention operations. It introduces local and global volume-based self-attention to learn volume representations and outperforms previous transformer-based methods on three public datasets. SeMask [27] proposes a semantically masked transformer network for semantic segmentation of images. The network leverages an additional semantic layer to incorporate semantic information about the image, which improves the performance of the pre-trained transformer backbone. However, these transformer-based U-Net models face challenges such as increased model size, longer training time, and higher computational requirements, which can limit their practicality in real-world applications.

References

- Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
- Zhu, X.; Cheng, D.; Zhang, Z.; Lin, S.; Dai, J. An empirical study of spatial attention mechanisms in deep networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6688–6697.
- Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
- Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.
- 5. Sarvamangala, D.; Kulkarni, R.V. Convolutional neural networks in medical image understanding: A survey. Evol. Intell. 2022, 15, 1–22.

- Dvořák, P.; Menze, B. Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation. In Proceedings of the Medical Computer Vision: Algorithms for Big Data: International Workshop, MCV 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, 9 October 2015; Revised Selected Papers. Springer: Cham, Switzerland, 2016; pp. 59– 71.
- Sedlar, S. Brain tumor segmentation using a multi-path CNN based method. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, 14 September 2017; Revised Selected Papers. Springer: Cham, Switzerland, 2018; pp. 403–422.
- Wang, G.; Li, W.; Ourselin, S.; Vercauteren, T. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, 14 September 2017; Revised Selected Papers. Springer: Cham, Switzerland, 2018; pp. 178–190.
- 9. Du, G.; Cao, X.; Liang, J.; Chen, X.; Zhan, Y. Medical image segmentation based on U-Net: A review. J. Imaging Sci. Technol. 2020, 64, 020508.
- Maurya, S.; Kumar Yadav, V.; Agarwal, S.; Singh, A. Brain Tumor Segmentation in mpMRI Scans (BraTS-2021) Using Models Based on U-Net Architecture. In Proceedings of the International MICCAI Brainlesion Workshop, Online, 27 September 2021; Springer: Cham, Switzerland, 2021; pp. 312–323.
- Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. Unet++: A nested U-Net architecture for medical image segmentation. In Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 20 September 2018; Proceedings. Springer: Cham, Switzerland, 2018; pp. 3–11.
- Myronenko, A. 3D MRI brain tumor segmentation using autoencoder regularization. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 16 September 2018; Revised Selected Papers, Part II. Springer: Cham, Switzerland, 2019; pp. 311–320.
- Futrega, M.; Milesi, A.; Marcinkiewicz, M.; Ribalta, P. Optimized U-Net for brain tumor segmentation. In Proceedings of the International MICCAI Brainlesion Workshop, Online, 27 September 2021; Springer: Cham, Switzerland, 2021; pp. 15–29.
- Zhang, B.; Wang, Y.; Ding, C.; Deng, Z.; Li, L.; Qin, Z.; Ding, Z.; Bian, L.; Yang, C. Multi-scale feature pyramid fusion network for medical image segmentation. Int. J. Comput. Assist. Radiol. Surg. 2023, 18, 353–365.

- Gong, Y.; Yu, X.; Ding, Y.; Peng, X.; Zhao, J.; Han, Z. Effective fusion factor in FPN for tiny object detection. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 1160–1168.
- Chen, M.; Wu, Y.; Wu, J. Aggregating multi-scale prediction based on 3D U-Net in brain tumor segmentation. In Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, 17 October 2019; Revised Selected Papers, Part I. Springer: Cham, Switzerland, 2020; pp. 142–152.
- Chi, M.; An, H.; Jin, X.; Wen, K.; Nie, Z. SCAR U-Net: A 3D Spatial-Channel Attention ResU-Net for Brain Tumor Segmentation. In Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences, Amsterdam, The Netherlands, 13–15 October 2022; pp. 497– 501.
- 18. Hua, R.; Huo, Q.; Gao, Y.; Sui, H.; Zhang, B.; Sun, Y.; Mo, Z.; Shi, F. Segmenting brain tumor using cascaded V-Nets in multimodal MR images. Front. Comput. Neurosci. 2020, 14, 9.
- 19. Avesta, A.; Hossain, S.; Lin, M.; Aboian, M.; Krumholz, H.M.; Aneja, S. Comparing 3D, 2.5 D, and 2D Approaches to Brain Image Auto-Segmentation. Bioengineering 2023, 10, 181.
- 20. Xu, C.; Howey, J.; Ohorodnyk, P.; Roth, M.; Zhang, H.; Li, S. Segmentation and quantification of infarction without contrast agents via spatiotemporal generative adversarial learning. Med. Image Anal. 2020, 59, 101568.
- 21. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. Adv. Neural Inf. Process. Syst. 2017, 30, 5998–6008.
- 22. He, K.; Gan, C.; Li, Z.; Rekik, I.; Yin, Z.; Ji, W.; Gao, Y.; Wang, Q.; Zhang, J.; Shen, D. Transformers in medical image analysis. Intell. Med. 2023, 3, 59–78.
- Wang, T.; Nie, Z.; Wang, R.; Xu, Q.; Huang, H.; Xu, H.; Xie, F.; Liu, X.J. PneuNet: Deep learning for COVID-19 pneumonia diagnosis on chest X-ray image analysis using Vision Transformer. Med. Biol. Eng. Comput. 2023, 61, 1395–1408.
- Hatamizadeh, A.; Tang, Y.; Nath, V.; Yang, D.; Myronenko, A.; Landman, B.; Roth, H.R.; Xu, D. UNETR: Transformers for 3d medical image segmentation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 4–8 January 2022; pp. 574–584.
- Hatamizadeh, A.; Nath, V.; Tang, Y.; Yang, D.; Roth, H.R.; Xu, D. Swin UNETR: Swin transformers for semantic segmentation of brain tumors in mri images. In Proceedings of the International MICCAI Brainlesion Workshop, Online, 27 September 2021; Springer: Cham, Switzerland, 2021; pp. 272–284.

- 26. Zhou, H.Y.; Guo, J.; Zhang, Y.; Han, X.; Yu, L.; Wang, L.; Yu, Y. nnFormer: Volumetric medical image segmentation via a 3D transformer. IEEE Trans. Image Process. 2023, 32, 4036–4045.
- Jain, J.; Singh, A.; Orlov, N.; Huang, Z.; Li, J.; Walton, S.; Shi, H. Semask: Semantically masked transformers for semantic segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, New Orleans, LA, USA, 18–24 June 2022; pp. 752–761.

Retrieved from https://encyclopedia.pub/entry/history/show/125609