ISSN: 2664-9640

Brain Glioma Segmentation Model based on Multi-scale Information Fusion

Hongwei Zhang, Junshan Chen*

School of Chong Qing, Normal University, Chongqing, China

*Corresponding Author

Abstract

Predicting histological grading and genetic biomarkers of gliomas using medical imaging and deep learning is a challenging task that is important for personalised treatment and survival prediction. Previous studies require physicians to manually outline the lesion area, which consumes a lot of human resources. Therefore, in this paper, we propose a brain glioma segmentation model based on multi-scale information fusion. Our model is based on the U-Net architecture, which incorporates the use of a multi-scale information fusion module. Compared to the original network model, our approach shows a higher lesion focus and shows strong potential for practical applications.

Keywords

Medical Image Segmentation; Adult Diffuse Glioma; MRI; CNN.

1. Introduction

Gliomas are tumours that arise from glial cells and account for 40-45% of all brain tumours, with a high degree of malignancy and the highest incidence of primary brain tumours. Gliomas have become the primary tumours with the highest incidence in the adult central nervous system (accounting for 40-45% of all intracranial tumours), and their aggressive biological behaviour and poor prognostic features have attracted much attention[1]. Clinical practice has shown that the median survival of patients with pathologically confirmed malignant gliomas is generally maintained at 12-24 months, and that low-grade diffuse gliomas have the potential to progress to higher grades, making complete cure difficult to achieve. Post-operative neurological and cognitive deficits and the financial burden of ongoing treatment present a multidimensional challenge that affects the quality of patient survival .

According to the WHO 2021 edition of the CNS tumour classification guidelines, molecular pathological features have been formally incorporated into the glioma grading and diagnostic system [2]. Current research focuses on seven key molecular markers: 1p/19q co-deletion [3], IDH1/2 mutations, EGFR amplification, TERT promoter variants , PTEN dysfunction, P53 pathway abnormalities, ATRX gene inactivation and CDKN2A/B deletion[4]. These markers have demonstrated significant clinical value in prognostic assessment and targeted therapy.

The current gold standard for genotyping is based on surgical/biopsy tissue samples combined with immunohistochemistry (IHC) for protein expression profiling[5]. However, invasive sampling is limited by factors such as patient tolerance, anatomical location of the lesion and risk of postoperative complications[6]. Although techniques such as RT-qPCR, FISH and NGS can provide information at the molecular level, their clinical application faces several limitations, including high sensitivity to sample quality, limited detection throughput, strict requirements for operational standardisation and high detection costs. In particular, second-generation sequencing (NGS) typically requires 5-7 working days to complete data analysis, which may affect the timeliness of clinical decision making[7]. Therefore, the establishment of a new typing technology system that combines economy, convenience and high accuracy has

ISSN: 2664-9640

become an urgent scientific challenge. And the task of accurate segmentation before typing is an important task to achieve glioma genotyping.

Yogananda et al [8] proposed a deep learning based 3D-Dense-UNet model for 1p/19q co-deletion status classification and tumour segmentation of gliomas. The study used T2w MRI and migration learning based on data from 368 patients acquired by TCIA and TCGA.The 3-fold cross-validation results showed that the Dice score for tumour segmentation was 0.8452.The study demonstrated that multi-task learning combined with migration learning and the use of only one MRI modality not only simplifies the research process, but also improves the efficiency of genotyping through shared image features and accuracy.

Wang et al [9]proposed SGPNet, a 3D multitasking residual framework for glioma segmentation. The model uses T1, T2, T1Gd and T2-FLAIR MRI modalities to extract hierarchical features from residual units. Experimental results show that SGPNet performs well on the BraTS 2020 and TCGA datasets with a segmentation Dice coefficient of 0.935. It was found that the lesion region information extracted in the segmentation task has a significant impact on the predictive effect of genotypes, further emphasising the critical role of the segmentation task.

Cheng et al [10] proposed MTTU-Net, a multi-task learning framework based on multimodal MRI, combining CNN and Transformer for brain glioma segmentation and tumour segmentation, with a Dice score of 0.8984 in the BraTS 2020 dataset through shared features and uncertainty pseudo-label selection (UPS). The study demonstrated that multi-task learning combined with uncertainty weight balancing can better perform the two tasks of brain glioma segmentation and genotyping simultaneously, which significantly improves the performance of the model and provides a more reliable clinically-assisted genotyping application for technical support.

In this study, the U-Net network model was used as the backbone network, and the application of a multi-scale information fusion module (MSF) was proposed as a way to achieve accurate segmentation of glioma tumour regions for subsequent glioma genotyping tasks.

2. Data and Methods

2.1. Data

The data in this article comes from TCIA[11], the First Affiliated Hospital of Chongqing Medical University, and TCGA[12]. The data selection mainly relies on manually selected lesion information by doctors. Through t1c, t2f, and segmented images of lesion information selected by doctors, slice images with lesions are selected channel by channel and band by band. When segmented images show the presence of lesions, their t1c and t2f data slices are saved, and normalization and Resiz operations are performed simultaneously.

2.2. Methods

The U-Net model proposed by Ronneberger et al.[13] is one of the first algorithms to use fully convolutional networks for semantic segmentation. The model is called U-Net because its network structure is a symmetric U. The pooling layer is replaced by an upsampling operator and contains a large number of feature channels, making it suitable for multi-scale image segmentation tasks.

The U-net structure is divided into two parts, the encoder and the decoder, with the encoder on the left and the decoder on the right. In the encoder part, features are extracted by four rounds of downsampling to extract the high-level semantic features of the image, which can effectively capture the spatial information and morphological features of the brain tumour; then the low-resolution feature maps obtained by the encoder are successively up-sampled by the four groups of decoder blocks, which operate in the opposite way to the encoder to gradually recover the spatial resolution of the image. U-Net has been used in the field of medical image

ISSN: 2664-9640

segmentation, especially the The research and application of brain tumour segmentation has been widely used and recognised.The U-Net model diagram is shown in Figure 1.

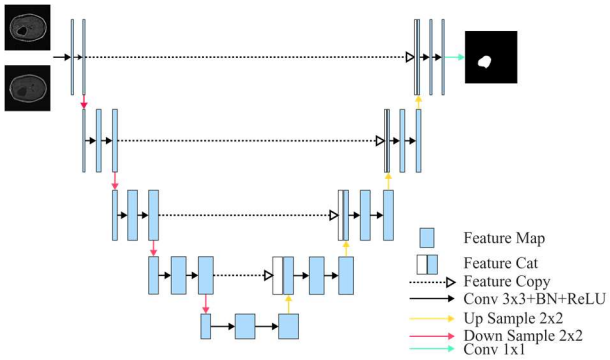


Figure 1. U-net network model

In the standard U-Net architecture, the decoder part passes and fuses the high-level semantic information from the encoder and the progressively recovered detail information in the decoder by connecting high-resolution feature maps with low-resolution feature maps (i.e., splicing operation). However, simple splicing operations often fail to fully utilize the complex relationships between these feature maps, which may result in redundant or insufficient information. To solve this problem, we propose a multi-scale information fusion module (Mut-Scale Fusion, MSF), and the structure of MSF is schematically shown in Figuere 2.

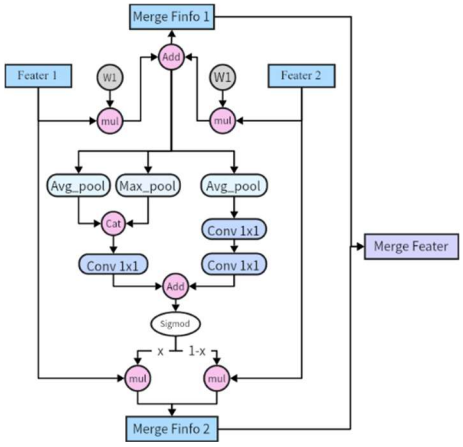


Figure 2. Mut-Scale Fusion

MSF combines Channel Attention (CA) and Spatial Attention (SA) mechanisms for optimizing the fusion process of low semantic information with high semantic information in U-Net decoder. The module receives as input two feature maps at different scales in the U-Net decoder, denoted as x (low semantic information) and y (high semantic information). These two feature maps contain lower and higher level semantic information respectively.

$$Merge\_Info = par1 \cdot x + par2 \cdot y + par3 \quad (1)$$

par1, par2 and par3 as weights and paranoia of high and low semantic information, which are all learnable parameters in the code.

The channel attention mechanism (CA) learns the importance of each channel by performing global average pooling on the fused feature maps and focuses on the more critical channel

ISSN: 2664-9640

information during the fusion process to improve the representational ability of the network, while the spatial attention mechanism (SA) focuses on the important spatial locations in the image. It generates a spatial attention map to optimise the region focusing ability of the image through average pooling and maximum pooling operations on the input feature map.

The results of channel attention and spatial attention are merged by a sigmoid activation function to obtain a merged attention weight. This weight is used to weight the input feature maps x and y. The weighting formula is given in Eq 2:

$$Merge_{Feater} = Merge_{Info} + \sigma(ca + sa) \cdot x + \big(1 - \sigma(ca + sa)\big) \cdot y \tag{2}$$

The MSF module combines the channel attention and spatial attention mechanisms to weight the high semantic and low semantic information in the fused decoder in a more fine-grained way, which enables U-Net to capture the detailed features and global contextual information more accurately in image segmentation tasks, the schematic of the MSF module is shown in Fig. The schematic of the MU-Net is shown in Figure 3.

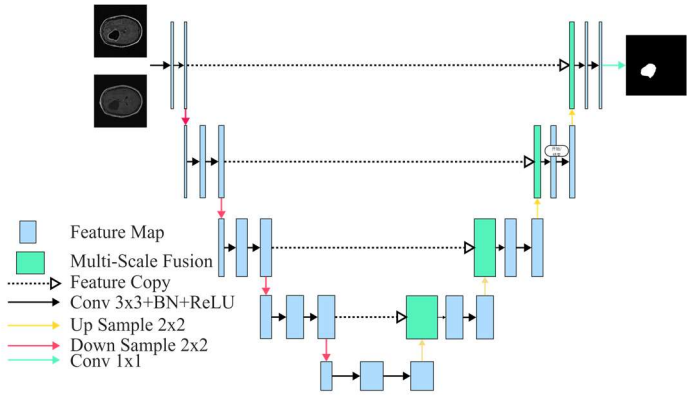
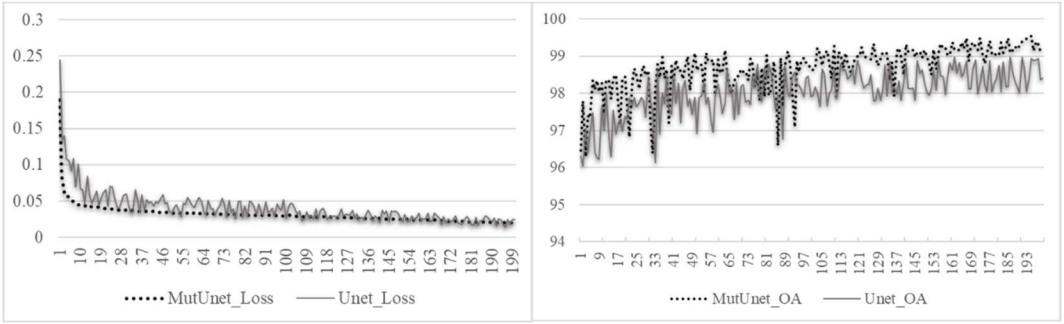


Figure 3. Mu-Net

3. Expermengts and Analysis

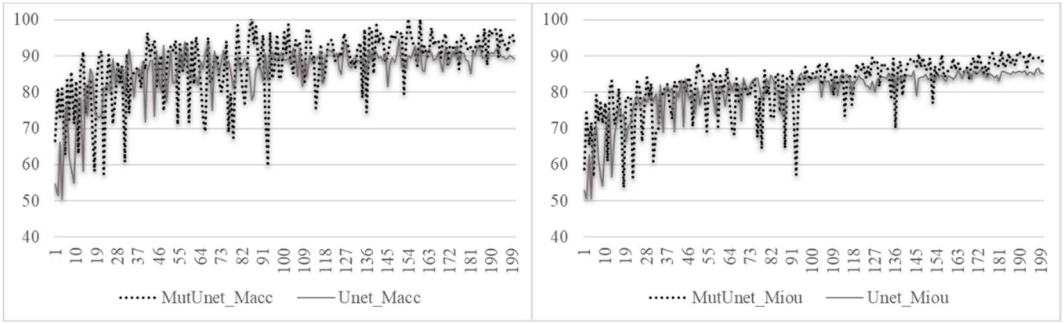
3.1. Dataset and Experimental Setup

The experiment is based on the execution of 200 iterations, and the widely used Adam optimiser is used to ensure efficient parameter updates and a stable training process. The hardware equipment used for the experiments is configured as follows 12th generation Intel® Core™ i7-12700F processor (2.10 GHz), 32.0 GB of RAM (31.8 GB of actual available RAM), and NVIDIA RTX 3070 graphics card.These configurations provide powerful computational support for the experiments, allowing the model training to proceed efficiently and stably. The results are shown in Figure X. Figure 4 shows the training results for the data.



(a) (b)

ISSN: 2664-9640



(c) (d)

Figure 4. Model Indicator Chart

Figure 4 shows how the MU-Net model performs against the standard U-Net architecture on several key metrics, including loss (, overall accuracy, accuracy, and average intersection and merger ratio. It is clear from the results that MU-Net outperforms U-Net on all metrics, as amply demonstrated by the trends in the graphs.

The training loss curves of the two models are illustrated in Fig. 4 (a). MU-Net significantly outperforms the standard U-Net in both the speed and magnitude of the loss decline (continuous folded line), while the loss decline curve of MU-Net is smoother and continues to decline, eventually reaching a lower loss value. As can be seen in Fig. 4(b), MU-Net shows more consistent and higher accuracy throughout the training process, while U-Net has more obvious fluctuations, with accuracy decreasing from time to time.The fluctuations of U-Net indicate that it may be more sensitive to different training phases during the training process, while MU-Net is more stable.

Figure 4(c) compares the accuracy (mAcc) of the two models. the accuracy of MU-Net is consistently higher than that of U-Net, and the change of its accuracy is smoother. The accuracy of MU-Net fluctuates less than that of U-Net, indicating that it is more stable during the training process and is able to maintain a high prediction accuracy.U-Net shows more fluctuations during the training process with occasional accuracy degradation, which may affect its performance in real-world applications.

Figure 4(d) evaluates the average intersection and merger ratio (mIoU) of the two models. mIoU growth of MU-Net is smoother compared to U-Net, while U-Net shows more fluctuation and occasional stagnation.

3.2. Experiments and Analysis

In Figure 5, the inference results of two models, U-Net and MU-Net, are shown, where the green contour indicates the area of manually outlined lesions, the blue area represents the inference results of the U-Net model, and the red colour is the inference results of the MU-Net model. As can be seen from Figure 5, the MU-Net model shows a significant advantage in dealing with lesions of different sizes. The main source of this advantage is the MSF (Multi-Scale Fusion) module introduced in MU-Net, which enables MU-Net to effectively handle lesions of different sizes and maintain high segmentation accuracy. In contrast, the U-Net model is prone to leakage when dealing with larger lesions, especially when dealing with the larger lesions in Figure X-c, the inference result of U-Net does not completely cover the lesion area and shows more significant leakage.

In addition, the performance of U-Net also has some shortcomings when dealing with smaller lesions. Although U-Net is more stable in its overall segmentation effect, U-Net often has missing or unclear segmentation boundaries when segmenting smaller lesions. This means that in small lesion scenarios requiring fine segmentation, the performance of U-Net is suboptimal and prone to blurred or inaccurate boundaries. In contrast, MU-Net is able to accurately capture the details

ISSN: 2664-9640

of small lesions while ensuring the accuracy of large lesions through the multi-scale feature of the MSF module, which greatly improves the segmentation accuracy of small lesions and reduces the phenomenon of missed segmentation. Meanwhile, the U-Net model is prone to segmentation leakage or large differences in segmentation boundaries when faced with smaller lesions.

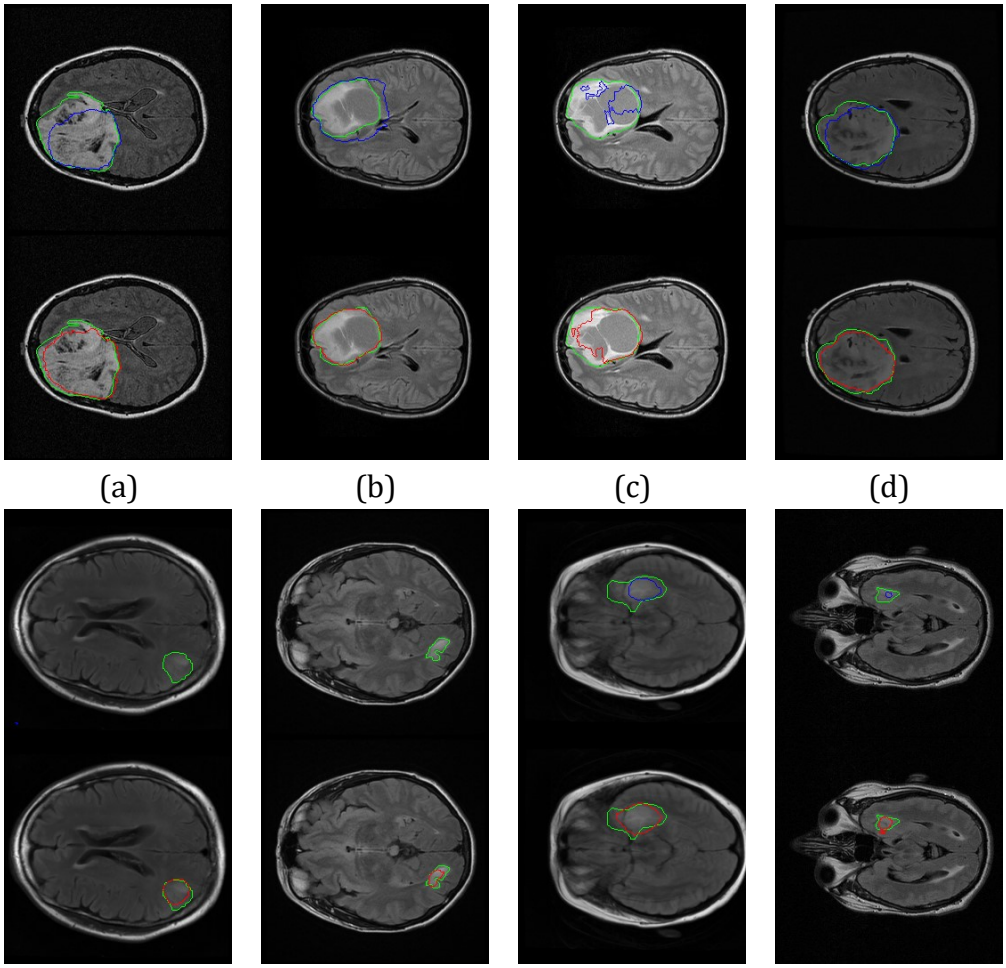


Figure 5. Model heatmap

4. Conclusion

In this study, we propose a new multi-scale information fusion module to improve the U-Net network, which ultimately achieves accurate segmentation of glioma lesion regions. This innovative approach realizes the enhancement of the performance of the common U-Net network, which greatly facilitates the task of lesion segmentation and subsequent genotyping of gliomas. It has certain practical application value. Although the proposed method achieved good segmentation results, it still needs to be tested and improved using more data in the future, and the patient's age, gender, and pathological characteristics need to be subsequently taken into account for greater help in clinical decision-making.

References

[1] Weller M, Wick W, Aldape K, et al. Glioma[J]. Nature reviews Disease primers, 2015, 1(1): 1-18.

[2] Sahm F, Brandner S, Bertero L, et al. Molecular diagnostic tools for the World Health Organization (WHO) 2021 classification of gliomas, glioneuronal and neuronal tumors; an EANO guideline[J]. Neuro-oncology, 2023, 25(10): 1731-1749.

[3] Zhang S, Yin L, Ma L, et al. Artificial Intelligence Applications in Glioma With 1p/19q Co-Deletion: A Systematic Review[J]. Journal of Magnetic Resonance Imaging, 2023, 58(5): 1338-1352.

[4] Nicholson A G, Tsao M S, Beasley M B, et al. The 2021 WHO classification of lung tumors: impact of advances since 2015[J]. Journal of Thoracic Oncology, 2022, 17(3): 362-387.

[5] Lu V M, O'Connor K P, Shah A H, et al. The prognostic significance of CDKN2A homozygous deletion in IDH-mutant lower-grade glioma and glioblastoma: a systematic review of the contemporary literature[J]. Journal of neuro-oncology, 2020, 148: 221-229.

[6] Di Bonaventura R, Montano N, Giordano M, et al. Reassessing the role of brain tumor biopsy in the era of advanced surgical, molecular, and imaging techniques—a single-center experience with long-term follow-up[J]. Journal of Personalized Medicine, 2021, 11(9): 909.

[7] Gisina A, Kholodenko I, Kim Y, et al. Glioma stem cells: novel data obtained by single-cell sequencing[J]. International journal of molecular sciences, 2022, 23(22): 14224.

[8] Yogananda C G B, Shah B R, Yu F F, et al. A novel fully automated MRI-based deep-learning method for classification of 1p/19q co-deletion status in brain gliomas[J]. Neuro-oncology advances, 2020, 2(Supplement_4): iv42-iv48.

[9] Wang Y, Wang Y, Guo C, et al. SGPNet: A Three-Dimensional Multitask Residual Framework for Segmentation and IDH Genotype Prediction of Gliomas[J]. Computational Intelligence and Neuroscience, 2021, 2021(1): 5520281.

[10] Cheng J, Liu J, Kuang H, et al. A fully automated multimodal MRI-based multi-task learning for glioma segmentation and IDH genotyping[J]. IEEE Transactions on Medical Imaging, 2022, 41(6): 1520-1532.

[11] Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository[J]. Journal of digital imaging, 2013, 26: 1045-1057.

[12] Ceccarelli M, Barthel F P, Malta T M, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma[J]. Cell, 2016, 164(3): 550-563.

[13] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, 2015: 234-241.