Enhanced Brain Tumor Segmentation Using CoAtNet with Multi-Modal Fusion

Abstract

This research aims to enhance brain tumor segmentation by employing the CoAtNet model, which integrates multiple types of medical images, such as MRI, PET, and CT scans. Traditional methods often rely on a single imaging modality, which can limit accuracy due to each scan's specific strengths and weaknesses. By combining these different imaging modalities into a unified approach, CoAtNet seeks to improve the precision of tumor detection and segmentation. This project will explore how this multi-modal fusion technique can provide a more comprehensive view of brain tumors, potentially leading to better diagnostic and treatment outcomes.

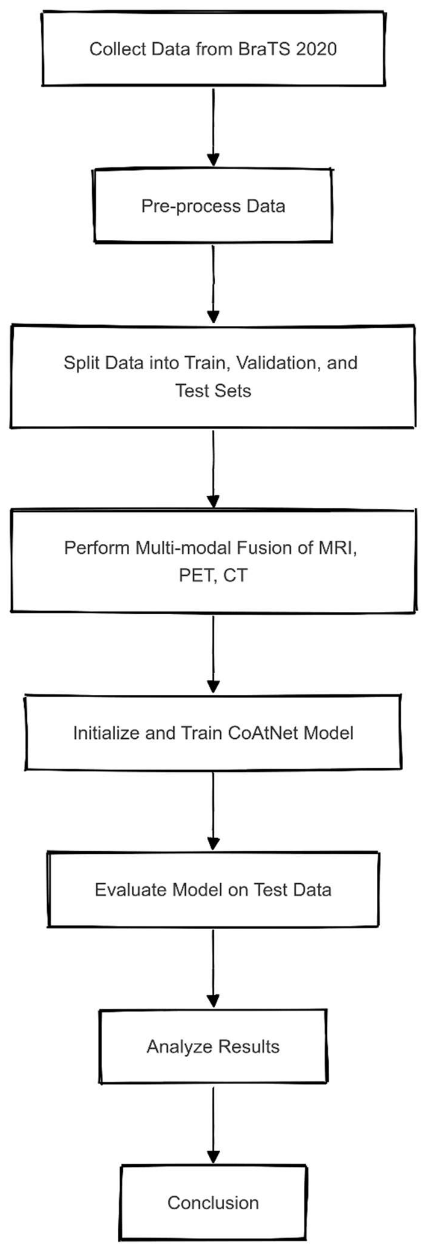
Introduction

Brain tumor segmentation is a crucial task in medical imaging, as accurate identification of tumor regions significantly affects diagnosis, treatment planning, and patient outcomes. Imaging modalities such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and Computed Tomography (CT) each offer distinct advantages—MRI provides structural detail, PET reveals metabolic activity, and CT offers rapid imaging with high contrast—but relying on a single modality can result in incomplete segmentation due to the limitations of each type. Recent advances in deep learning and multi-modal fusion techniques have shown promise in improving segmentation accuracy by integrating complementary information from multiple imaging sources. CoAtNet, a model designed to fuse MRI, PET, and CT data, leverages a combination of convolutional neural networks (CNNs) and transformers to capture both local and global features, making it a strong candidate for brain tumor segmentation tasks. Building on previous work with hybrid models like TranSiam and U-Net-based architectures, this research explores how CoAtNet's multi-modal fusion approach can enhance diagnostic precision and lead to improved treatment outcomes.

Objective

The objective of this research is to enhance brain tumor segmentation by employing the CoAtNet model, which integrates multiple medical imaging modalities such as MRI, PET, and CT scans. Traditional methods that rely on a single imaging modality can limit accuracy, as each scan has its own strengths and weaknesses. By combining these different imaging modalities, CoAtNet aims to improve the precision of tumor detection and segmentation, ultimately providing a more comprehensive view of brain tumors. This research will explore how multi-modal fusion can lead to better diagnostic accuracy and treatment outcomes for brain tumors.

Flowchart



Literature Review

1) TranSiam, a dual-path network that combines CNN and transformers for improved medical image segmentation. The model extracts detailed information using convolution and captures global features with an Improved Convolutional Transformer (ICMT) block. A novel TMM block is introduced to fuse multimodal data using cross- and self-attention mechanisms. Evaluated on the BraTS 2019 and 2020 datasets, TranSiam outperforms classic methods, achieving a high Dice score of 89.34%, with improved sensitivity and specificity. It effectively balances accuracy and computational efficiency.
2) A brain tumor segmentation framework using multi-modal MRI data, enhancing accuracy by integrating information from different modalities. It uses a hybrid fusion strategy with self-supervised learning to capture cross-modal dependencies. The model, built on a U-Net design, achieves higher segmentation accuracy on the BraTS 2019 dataset, preventing overfitting.
3) Brain tumor segmentation using multi-modal MRI data using a self-supervised learning strategy. The model uses a hybrid fusion of modality-specific features, a fully convolutional neural network, ResNet-50, Atrous Spatial Pyramid Pooling, and a Hybrid Attentional Fusion Block. Experimental results show superior segmentation performance, high average Dice score, improved sensitivity and specificity, and better edge detection of tumor regions.
4) Brain tumor segmentation issues with incomplete MRI data. It proposes a Region-Aware Fusion Network (RFNet) that uses a Region-aware Fusion Module to fuse available modalities based on their sensitivity to specific tumor regions. The model, built on a 3D U-Net architecture, improves segmentation accuracy across multiple datasets, even with missing data.
5) A robust brain tumor segmentation method using multi-modal MRI data. It proposes a deep neural network consisting of a feature-enhanced generator, a correlation constraint block, and a segmentation network. The generator synthesizes missing MRI modalities, while the block ensures coherence between modalities. The segmentation network, a multi-encoder U-Net, performs the final segmentation task. The method outperforms state-of-the-art by 3.5%, 17%, and 18.2%.
6) A deep neural network has been developed to improve brain tumor segmentation from multi-modal MRI data. The model uses a multi-encoder 3D U-Net architecture, integrating information from MRI modalities like FLAIR, T2, T1, and T1c. Key components include Modality-level Cross-Connection (MCC) and Attentional Feature Fusion Module (AFFM). Experiments on the BraTS 2018 dataset show significant improvements in segmentation accuracy, especially for challenging tumor sub-regions.
7) A deep learning architecture has been developed to improve brain tumor segmentation using multimodal MRI images. The model uses separate encoders for each modality and fuses them using a Bi-Directional Feature Pyramid Network (Bi-FPN). Experiments on the MICCAI BraTS2018 and BraTS2020 datasets show that the model outperforms existing methods with superior Dice coefficients, making it a strong candidate for automatic brain tumor segmentation tasks. The model's robustness and accuracy make it a strong candidate for automatic brain tumor segmentation tasks

8) A deep learning model for brain tumor segmentation uses multimodal MRI data to enhance accuracy. The model uses multiple modality-specific encoders to extract features from different MRI modalities and an attention mechanism to emphasize informative features. Tested on the BraTS 2017 dataset, this multi-encoder and attention-based approach outperforms single-encoder methods and achieves superior performance in segmenting brain tumors, particularly in the enhancing tumor region.
9) Brain tumor segmentation using multi-modal MRI images, targeting different tumor regions. It uses an ensemble of three UNet models, including 3D MRI patches, brain parcellation information, and a combination of 3D and 2D inputs. Tested on the BraTS2018 and BraTS2020 datasets, the method achieved superior segmentation performance with Dice scores of 91.03% for whole tumor, 86.44% for tumor core, and 80.58% for enhancing tumor.
10) Brain tumor segmentation using multi-modality MRI data using a deep convolutional neural network. The model uses a U-shaped fully convolutional network and deep residual learning to fuse features from multiple MRI modalities. Experiments on the BraTS2021 dataset showed improved segmentation accuracy with Dice scores of 83.3% for enhancing tumors, 89.07% for tumor cores, and 91.44% for whole tumors.
11) Brain tumor segmentation by generating synthetic multi-modal medical image pairs using a Generative Adversarial Network (GAN). TumorGAN synthesizes new image pairs by combining semantic labels and brain contours from real data, improving data augmentation. The model was validated on the BraTS 2017 dataset, showing improved segmentation performance and increased Dice scores compared to traditional methods, demonstrating its effectiveness in synthetic data augmentation.
12) A three-stage network is developed to segment brain tumors using multi-sequence MRI data. The model uses a 3D U-Net for initial segmentation, a multi-encoder network for fusion, and a final stage for refinement. A new loss function is introduced for multi-class segmentation. The method was evaluated on the BraTS 2017 dataset, achieving improved segmentation performance with Dice scores of 89.4 for the whole tumor, 81.6 for the tumor core, and 73.0 for the enhancing tumor.
13) Brain tumor segmentation using multimodal MRI by combining deep semantic features and edge information. The model consists of three modules: semantic segmentation using the Swin Transformer, edge detection using CNNs, and feature fusion with graph convolution. Tested on the BraTS benchmarks, it showed superior performance and improved tumor boundary detection accuracy.
14) For predicting brain tumor recurrence locations using multi-modal fusion and nonlinear correlation learning. The model uses transfer learning, multi-scale multi-channel feature fusion, and a nonlinear correlation learning module to improve tumor prediction from limited datasets. It also uses a Kullback-Leibler divergence-based loss function to enhance correlation between different modalities. The model outperforms baseline models with a 63.5% Dice Similarity Coefficient and 69.2% sensitivity.
15) A deep learning model for semantic segmentation of brain tumors from multimodal MRI images using a modified Link-Net architecture and a Squeeze and Excitation ResNet152 model. The model achieved a 99.2% accuracy, outperforming traditional methods like U-Net and FCN, and was validated by neurosurgeons. The model's robustness in tumor segmentation is demonstrated through its integration of multi-modal imaging data.

Reference

Li, X., Ma, S., Tang, J., & Guo, F. (2022). TranSiam: Fusing multimodal visual features using transformer for medical image segmentation. *arXiv preprint arXiv:2204.12185*.

Fang, F., Yao, Y., Zhou, T., Xie, G., & Lu, J. (2021). Self-supervised multi-modal hybrid fusion network for brain tumor segmentation. *IEEE Journal of Biomedical and Health Informatics*, *26*(11), 5310-5320.

Yang, H., Zhou, T., Zhou, Y., Zhang, Y., & Fu, H. (2023). Flexible fusion network for multi-modal brain tumor segmentation. *IEEE Journal of Biomedical and Health Informatics*, *27*(7), 3349-3359.

Ding, Y., Yu, X., & Yang, Y. (2021). RFNet: Region-aware fusion network for incomplete multi-modal brain tumor segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 3975-3984).

Zhou, T., Canu, S., Vera, P., & Ruan, S. (2021). Feature-enhanced generation and multi-modality fusion based deep neural network for brain tumor segmentation with missing MR modalities. *Neurocomputing*, *466*, 102-112.

Zhou, T. (2023). Modality-level cross-connection and attentional feature fusion based deep neural network for multi-modal brain tumor segmentation. *Biomedical Signal Processing and Control*, *81*, 104524.

Syazwany, N. S., Nam, J. H., & Lee, S. C. (2021). MM-BiFPN: multi-modality fusion network with Bi-FPN for MRI brain tumor segmentation. *IEEE Access*, *9*, 160708-160720.

Zhou, T., Ruan, S., Guo, Y., & Canu, S. (2020, April). A multi-modality fusion network based on attention mechanism for brain tumor segmentation. In *2020 IEEE 17th international symposium on biomedical imaging (ISBI)* (pp. 377-380). IEEE.

Zhang, Y., Zhong, P., Jie, D., Wu, J., Zeng, S., Chu, J., ... & Tang, X. (2021). Brain tumor segmentation from multi-modal MR images via ensembling UNets. *Frontiers in radiology*, *1*, 704888.

Zhang, G., Zhou, J., He, G., & Zhu, H. (2023). Deep fusion of multi-modal features for brain tumor image segmentation. *Heliyon*, *9*(8).

Li, Q., Yu, Z., Wang, Y., & Zheng, H. (2020). TumorGAN: A multi-modal data augmentation framework for brain tumor segmentation. *Sensors*, *20*(15), 4203.

Zhou, T., Canu, S., & Ruan, S. (2020). Fusion based on attention mechanism and context constraint for multi-modal brain tumor segmentation. *Computerized Medical Imaging and Graphics*, *86*, 101811.

Zhu, Z., He, X., Qi, G., Li, Y., Cong, B., & Liu, Y. (2023). Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI. *Information Fusion*, *91*, 376-387.

Zhou, T., Noeuveglise, A., Modzelewski, R., Ghazouani, F., Thureau, S., Fontanilles, M., & Ruan, S. (2023). Prediction of brain tumor recurrence location based on multi-modal fusion and nonlinear correlation learning. *Computerized Medical Imaging and Graphics*, *106*, 102218.

Ramasamy, G., Singh, T., & Yuan, X. (2023). Multi-modal semantic segmentation model using encoder based Link-Net architecture for BraTS 2020 challenge. *Procedia Computer Science*, *218*, 732-740.