TRANSFORMER NETWORK FOR BRAIN GLIOMA SEGMENTATION IN MRI IMAGES

by Jana Publication & Research

Submission date: 14-Jun-2025 11:57AM (UTC+0700)
Submission ID: 2690340364
File name: IJAR-52220.docx (318.88K)
Word count: 4167
Character count: 25230

TRANSFORMER NETWORK FOR BRAIN GLIOMA SEGMENTATION IN MRI IMAGES

1. ……

Manuscript Info ……

Manuscript History

Received: xxxxxxxxxxxxxxx
Final Accepted: xxxxxxxxxxx
Published: xxxxxxxxxxxxxxx

Key words:-
xxxxxxxxx

Abstract ……

Glioma is a type of tumor that originates in the neuroglial cells of the brain or spinal cord, forming a mass that can press on surrounding tissue and cause symptoms. To diagnose glioma and to assess the tumor volume, manual segmentation of gliomas in MRI images is normally performed. However, it is time-consuming and prone to errors due to diagnostic variability among experts. This study proposes a deep learning approach using a Transformer Network to enhance segmentation accuracy and improve diagnostic efficiency. The research utilizes the BraTS 2021 dataset, consisting of 374 MRI scans with ground truth labels, to train and evaluate the Transformer Network model. The model incorporates an EfficientNet-B1 backbone for computational efficiency and is trained with optimal parameters: a learning rate of 0.0001, batch size of 5, and 200 epochs. Results indicate that the Transformer Network achieved a Dice coefficient of 0.921, significantly outperforming the baseline deep learning segmentation method, which is U-Net model (0.827), demonstrating superior segmentation accuracy. In conclusion, the Transformer Network proves more effective and accurate than traditional methods for brain glioma segmentation. Future research should focus on expanding datasets and computational resources to further enhance model performance. This study is expected to contribute to an improved glioma diagnosis and treatment planning.

……

Introduction :-

Introduction :-

Glioma is a type of brain tumor that originates in the brain or spinal cord and arises from neuroglial cells, which support neuronal function by regulating impulses and supplying nutrients (Verkhratsky et al., 2023). Representing approximately 33% of all brain tumors, gliomas are among the most common and serious forms of central nervous system malignancies (Molnar et al., 2015). The World Health Organization (WHO) classifies brain tumors into four grades based on cellular characteristics. Low-grade tumors (Grades I and II) are typically benign and slow-growing, while high-grade tumors (Grades III and IV) are more aggressive and malignant, accounting for roughly 80% of all brain tumor cases.

Magnetic Resonance Imaging (MRI) is a widely utilized tool in the diagnosis and treatment planning of neurological disorders, including brain tumors (Fuzari et al., 2020). MRI provides high-resolution images across multiple

anatomical planes—axial, sagittal, and coronal—allowing detailed visualization of brain structures, vascular systems, and pathological regions (Afshar, Mohammadi & Plataniotis, 2018). Although the use MRI is critically important in clinical settings, the manual interpretation of these images is often labor-intensive, subject to variability among radiologists, and prone to diagnostic inaccuracies (Despotović et al., 2015). As a result, automated techniques are increasingly adopted to support and enhance diagnostic precision.

Deep learning, a domain within artificial intelligence, has gained prominence in medical imaging for to automatically extract meaningful features from raw data. In the context of brain tumor analysis, deep learning models assist clinicians in diagnosing and delineating tumors from healthy tissues, ultimately improving treatment planning and patient outcomes (Ahamed et al., 2023). Automated segmentation of brain tumors in MRI scans remains a complex challenge due to factors such as tumor heterogeneity, variations in image quality, and limited availability of annotated data. Reliable segmentation methods are crucial to accurately differentiate tumor boundaries, particularly for treatment planning and monitoring. Traditional models like U-Net have shown strong performance in brain tumor segmentation and are often considered benchmarks, outperforming or matching other architectures such as CNNs and Capsule Networks in various studies.

At present, Transformer-based architectures are rapidly becoming the new standard in medical image segmentation that achieve better performance than the previous state-of-the-art models. For instance, Liu et al., (2023) demonstrated superior performance in cardiac MRI segmentation using a Swin Transformer Network, achieving a Dice coefficient of 92.28%. Transformers, initially developed for natural language processing, have shown significant potential in vision tasks due to their ability to model long-range dependencies and global contextual relationships, which becomes the limitation of conventional CNN-based models. The growing success of Transformers in visual domains has encouraged their adoption in medical imaging, especially given their adaptability to varying input dimensions and robustness in handling complex structures. In the case of glioma segmentation, this adaptability and ability to capture spatial context may result in more precise and reliable segmentation outcomes.

This study explores the effectiveness of a Transformer-based deep learning model, specifically the Segtran architecture, for glioma segmentation using MRI data. By combining the strengths of attention mechanisms and neural network-based feature extraction, the study aims to enhance segmentation accuracy, reduce diagnostic variability, and support clinical decision-making. The performance of the Transformer model is benchmarked against the well-established U-Net architecture to evaluate its potential advantages in brain tumor segmentation tasks.

Methodology :-

The transformer architecture was first introduced by Vaswani et al. in 2017. The architecture is based on self-attention mechanisms, which replaced traditional recurrent neural networks (RNNs) and convolutional neural networks (CNNs) in sequence-to-sequence tasks. The model can handle long-range dependencies and parallel processing (Sajun et al., 2023) and has shown good performance especially in natural language processing (NLP) tasks. Then, these networks are explored in image classification problem which resulted in the use of Vision Transformer (ViT) in 2020 (Takahashi et al., 2024).

In terms of medical image segmentation problem that requires precise delineation of anatomical structures, the Transformer network is integrated with CNN based models to achieve superior performance in various image

medical tasks such as TransUNet and Swin-UNet. The integration has allowed long-range dependencies information to be captured, while maintaining the local feature extraction capabilities of CNNs (Chen et al., 2024) (Pu et al., 2024). The methodology used in this study is based on Transformer network.

The Transformer architecture is composed of two main components: the encoder and the decoder. Each part comprises of similar components such as layer normalization, masked multi-head attention modules, position-wise feed-forward networks, and multi-head attention modules. The decoder's masked multi-head attention mechanism restricts attention to earlier positions, ensuring that predictions depend only on previously generated outputs. This structure enables the Transformer to perform sequence-to-sequence tasks effectively, such as in machine translation and speech recognition.

Shaohua Li et al., (2021) proposed Segtran, a novel framework for medical image segmentation based on Squeeze-and-Expansion Transformers. In contrast to conventional models like U-Net, Segtran overcomes the constraint of limited effective receptive fields by incorporating a squeezed attention block for regularization and an expansion block to capture diverse feature representations. Additionally, it introduces a novel positional encoding strategy to improve spatial continuity in image processing by using a continuity inductive bias where spatial relationships are important. This paper utilized Segtran as proposed by Shaohua Li et al., (2021) to investigate its effectiveness in segmenting the brain glioma in MRI images.

In this mode, a pretrained CNN backbone was used to extract rich visual features from input MRI slices. For 2D inputs, EfficientNet-B4 was selected due to its balance between performance and computational efficiency. To retain higher spatial resolution, the initial convolutional stride was reduced. The feature maps obtained from multiple stages of the backbone provided multi-scale representations for downstream processing. To encode spatial information crucial for image understanding, a learnable sinusoidal positional encoding was added to the CNN-extracted features. This encoding combined sine and cosine functions with learnable parameters, allowing the model to adaptively represent spatial continuity and locality, which are vital for segmenting coherent anatomical structures.

The fused visual and positional features were flattened and passed through a stack of Squeeze-and-Expansion Transformer layers, the core innovation of the Segtran framework. In this layer, there are two main components namely Squeezed Attention Block (SAB) and Expanded Attention Block (EAB). The SAB block is used to reduce computational overhead and mitigate overfitting risks associated with standard attention mechanisms, attention computations were compressed using a set of learned codebook vectors (inducing points). This created a more compact attention map that still preserved global contextual relationships. On the other hand, in EAB block, multiple attention modes were introduced using a mixture-of-experts approach. Each mode processed the features independently, and the outputs were aggregated using dynamic mode attention. This allowed the model to learn diverse contextual representations and increased its modeling capacity.

Since high-level features from the CNN backbone were low in spatial resolution, a dual Feature Pyramid Network (FPN) structure was applied. The input FPN upsampled coarse CNN features before feeding them into the transformer layers. The output FPN further upsampled the transformer output to enhance spatial granularity before segmentation. This FPN configuration was bottom-up, preserving semantic richness during upsampling. The final feature map, now at a higher resolution, was processed by a simple 1×1 convolutional segmentation head, producing pixel-wise class confidence scores for each segmentation label, including whole tumor (WT) and the background.

i. Dataset

The study uses the BraTS 2021 Glioma brain tumor dataset from the Perelman School of Medicine at the University of Pennsylvania, updated annually (*http://www.braintumorsegmentation.org/*). BraTS is part of a segmentation

challenge aligned with MICCAI (https://miccai.org/) and focuses on advanced methods for brain tumor segmentation in multimodal MRI. The BraTS 2021 dataset includes 374 clinically obtained multiparametric MRI scans with confirmed diagnoses, comprising T1, T2, FLAIR, and T1Ce images. Enhanced contrast in T1ce images is crucial for distinguishing meningiomas from gliomas. T1-weighted images show high-fat content tissues as bright and CSF as dark, while T2-weighted images show water-filled areas as bright and high-fat content tissues as dark. FLAIR images highlight extended T2 tissue areas as bright and suppress CSF signals, appearing dark. Figure 1 shows a sample of the dataset of different types of MRI subimages.

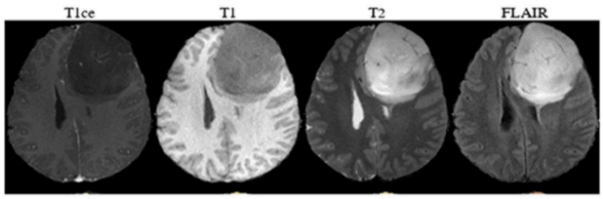

Figure 1. Four Sub-Files Formed (T1ce, T1, T2, and FLAIR)

ii. Training and Testing

During the training phase, the process was conducted over three epochs with a batch size of five, using the standard AdamW optimization method. AdamW is an optimization algorithm used for training deep learning models, which helps improve generalization and prevent overfitting by decoupling weight decay (Loshchilov & Hutter, 2019). The parameter refinement phase is conducted based on the input data model, selecting the best hyperparameters. Early stopping based on the validation loss was used to prevent overfitting.

iii. Model Evaluation

The Dice Similarity Coefficient (DSC) is a commonly used evaluation method in brain tumor segmentation problems (Kao et al., 2020). The Dice coefficient also serves as a general evaluation method in all BraTS challenges, with many researchers using it in their work (Isensee et al., 2017; Havaei et al., 2017; Kamnitsas et al., 2017; Myronenko et al., 2018). Therefore, the Dice coefficient will be used as the evaluation method in this study to measure the similarity between the predicted tumor segmentation and the actual tumor segmentation.

The formula for the Dice coefficient is:

$$DSC = \frac{2*|A \cap B|}{|A|+|B|} \quad (1)$$

In this formula, A and B are two different sets to be compared. The Dice score ranges from 0 to 1, with a value of 1 indicating a perfect match between the predicted and ground truth labels, and 0 indicating no overlap. If there is no similarity between the two sets, the Dice coefficient will be 0. When the actual ground truth segmentation closely matches the predicted one, the Dice coefficient will be equal to 1. The Dice score is calculated for each tumor region (whole tumor, tumor core, and enhancing core) and averaged to obtain the mean Dice score.

According to Zhou et al. (2019), the binary cross-entropy loss function is commonly used in image segmentation. The binary cross-entropy (BCE) is calculated by comparing the predicted pixel vector to the actual pixel vector, where the pixel vector y_i represents the ground truth segmentation mask and $\hat{y}$ represents the predicted segmentation mask.

$$LOSS_{BCE} = -\sum_{i=1}^{N} y_i log\hat{y} \quad (2)$$

However, imbalanced datasets in terms of class representation can lead to incorrect behavior of the loss function. To address class imbalance, the weighted binary cross-entropy (WBCE) loss function is used. This is particularly useful in cases where the ground truth segmentation mask has many background pixels and very few tumor regions (Long, Shelhamer & Darrell, 2015). The $LOSS_{WBCE}$ is defined as follows, where w_i represents the weights assigned to each class.

$$LOSS_{WBCE} = -\sum_{i=1}^{N} w_i y_i log\hat{y} \quad (3)$$

Based on Dong et al. (2017), the basic U-Net model for brain tumor segmentation employs the Dice Loss function, defined as the inverse of the Dice Coefficient. The Dice Loss is shown in Equation 4 where G and P are two different sets to be compared.

$$LOSS_{DSC} = 1 - \frac{2*|G \cap P|}{|G|+|P|} \quad (4)$$

This study utilizes Visual Studio Code with Python for brain glioma segmentation in MRI images. The following parameters are used:

- *math for basic mathematical functions.*
- *numpy for scientific computing.*
- *torch for training neural networks.*
- *torch.nn for neural network layers.*
- *torch.nn.Parameter for learnable parameters.*
- *torch.nn.functional for mathematical operations.*
- *networks.segtran_shared for Segtran model configuration.*
- *train_util.batch_norm for batch normalization.*
- *efficientnet.model for efficient models.*
- *argparse.Namespace for command-line arguments.*

Results and Discussion :-

The study employed the BraTS 2021 dataset, which comprises 374 MRI scans with confirmed brain tumor annotations, to evaluate the performance of a Transformer-based neural network for glioma segmentation. The dataset was divided into 85% for training and 15% for testing. EfficientNet-B1 was selected as the backbone for its

computational efficiency and strong feature extraction capabilities. The model architecture consisted of three convolutional layers followed by three Transformer layers, enabling both local and global feature representation.

Training was conducted using the AdamW optimizer to mitigate overfitting, with an initial learning rate of (29)01, a batch size of 5, and a total of 200 epochs. Training performance was monitored using the training loss and the Dice Similarity Coefficient (DSC) as the primary evaluation metric. Although the model required substantial memory due to its architecture, it achieved excellent segmentation results, outperforming models trained with fewer epochs. The configuration demonstrated robust learning and generalization across the dataset. We experimented with different parameter settings to determine the optimal hyperparameters based on the dataset size, batch size, number of iterations, and learning rate.

i. Dataset size

Table 1 shows the Dice coefficient values for different dataset sizes used to train and test Segtran Transformer network and U-Net models. The models were trained with various dataset sizes while keeping parameters like epochs, learning rate, and batch size constant (150 epochs, learning rate of 0.0001, batch size of 1).

Table 1 Comparison of Dice coefficient values for different dataset sizes for Transformer and U-Net networks

Dataset	Testing Data	Transformer	U-Net
100	10	0.904	0.767
200	20	0.864	0.787
300	20	0.823	0.801
320	54	0.829	0.817

Results indicate that larger datasets yield higher Dice coefficients. The dataset was maximized to 374 samples, split into 320 for training and 54 for testing. Larger datasets improve model accuracy by providing more data for interpreting variations and patterns, reducing overfitting, and addressing class imbalance. Larger datasets are used to enhance model accuracy by (Loshchilov & Hutter, 2019).

ii. Iteration

The training process consists of several iterations, known as epochs, during which the model learns from the training data and adjusts parameters to improve its performance (Wang et al. 2019). Table 2 represents the Dice coefficient values for different numbers of epochs used to train Transformer network and U-Net models.

Table 2 Comparison of Dice coefficient values for different iterations for Transformer and U-Net networks

Dataset	Training Data	Iteration	Transformer	U-Net
320	54	100	0.733	0.643
320	54	150	0.829	0.817
320	54	200	0.856	0.821
320	54	250	0.849	0.758
320	54	300	0.818	0.675

With the dataset size set to maximum, and both the learning rate of 0.0001 and batch size of 1 are held constant, the study found that training for 200 epochs yielded the highest Dice coefficient—0.856 for the Transformer network and 0.821 for U-Net. Training for 200 epochs enabled the model to learn complex patterns and extract essential features necessary for accurate tumor segmentation. However, exceeding this number of epochs may lead to overfitting, where the model performs well on the training data but poorly on unseen data. Therefore, 200 epochs were identified as the optimal training duration.

iii. Learning Rate

Table 3 presents the Dice coefficient values obtained for different numbers of training epochs using the Transformer network and U-Net models. To ensure a fair comparison, other parameters—namely the dataset size (set to its maximum), learning rate (0.0001), and batch size (1)—were kept constant. The training process involves multiple epochs during which the model iteratively learns and updates its parameters (Wang et al., 2019).

Table 3 Comparison of Dice coefficient values for different learning rates for Transformer and U-Net networks

Dataset	Training Data	Learning Rate	Transformer	U-Net
320	54	0.0001	0.856	0.821
320	54	0.001	0.657	0.616
320	54	0.01	0.007	0.003
320	54	0.1	0.001	0.000

The study found that 200 epochs yielded optimal performance, achieving a Dice coefficient of 0.856 for the Transformer network and 0.821 for U-Net. Training beyond 200 epochs may result in overfitting, where the model becomes overly specialized to the training data and performs poorly on unseen data (Li et al., 2021). Thus, 200 epochs were identified as the optimal training duration.

iv. Batch Size

Batch size refers to the number of data samples processed at once before updating model parameters in an iteration (Wang et al. 2019). Table 4 presents the Dice coefficient values for different batch sizes used with the Transformer network and U-Net models. To assess model performance across varying batch sizes, other parameters—dataset size (maximum), number of epochs (200), and learning rate (0.0001)—were kept constant. Batch size refers to the number of data samples processed simultaneously before the model updates its parameters in each iteration (Wang et al., 2019). The study found that a batch size of 5 yielded the best performance, with the Transformer network achieving a Dice coefficient of 0.921, outperforming U-Net, which achieved a Dice coefficient of 0.827.

Table 4. Comparison of Dice coefficient values batch size for Transformer and U-Net networks

Dataset	Training Data	BS	Transformer	U-Net
320	54	1	0.856	0.821
320	54	2	0.878	0.822
320	54	4	0.880	0.825
320	54	5	0.921	0.827

v. Segmentation Performance using Optimum Parameters

Under identical training conditions, the Transformer model outperformed U-Net in brain tumor segmentation, achieving a Dice coefficient of 0.921 compared to 0.827 for U-Net (Table 5). Both models utilized EfficientNet-b1 as the backbone, were trained for 200 epochs with a batch size of 5, and employed the AdamW optimization algorithm with a learning rate of 0.0001. The superior performance of the Transformer model can be attributed to its self-attention mechanism, which enables the extraction of global contextual features and long-range dependencies across the input MRI scans. This capability is particularly beneficial in medical image segmentation, where tumors may show complex shapes. While U-Net relies heavily on local convolutional operations, the Transformer architecture enhances feature representation by capturing information of different regions of the image. This leads to more accurate segmentation and better delineation of tumor boundaries. These findings suggest that the Transformer network is more effective at capturing critical spatial and contextual information, making it a more suitable architecture for brain tumor segmentation tasks.

Table 5. Comparison of Transformer and U-Net networks on testing dataset

Parameters	Transformer	U-Net
Backbone	EfficientNet-b1	EfficientNet-b1
Layers	3	3
Opitimization	AdamW	AdamW
Batch Size	5	5
Learning Rate	0.0001	0.0001
No of Iteration	200	200
Dice Coefficient	0.921	0.827

Figure 2 illustrates the training loss curve during the training process of the Transformer model. Training loss represents the error between the model's predictions and the actual values in the training dataset. Ideally, the loss should decrease progressively with each iteration or epoch, indicating that the model is improving its predictive accuracy. A consistent downward trend in the graph shows effective learning and model optimization. As shown in the figure, the training process achieves stable convergence, suggesting that the model is learning efficiently and avoiding issue of overfitting.

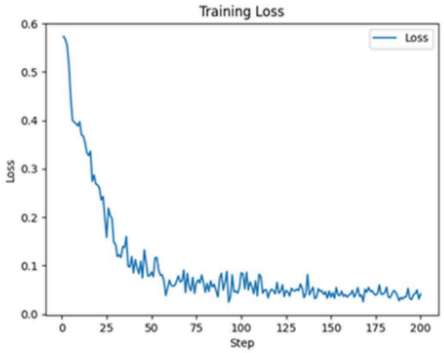

Figure 2 Training Loss

Figure 3 shows the cross-entropy loss graph. Cross-entropy loss measures the difference between the predicted distribution by the model and the actual distribution in the data. In brain tumor segmentation, it evaluates how well the model predicts the correct pixels for each class (e.g., tumor vs. non-tumor). The graph displays a decrease in cross-entropy loss with each training iteration. A consistent decrease indicates that the model is improving in recognizing patterns and making accurate predictions.

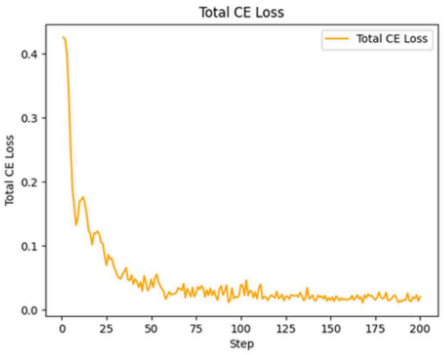

Figure 3 Cross Entropy Loss

Finally, Figure 4 shows the Dice loss graph. This graph displays changes in loss based on the Dice coefficient during the training process. The Dice coefficient measures similarity between two data sets, in this case, between the model's predicted segmentation and the actual segmentation. Dice loss is the inverse of the Dice coefficient, meaning lower loss indicates more accurate segmentation. A consistent decrease in Dice loss shows that the model is improving in tumor segmentation.

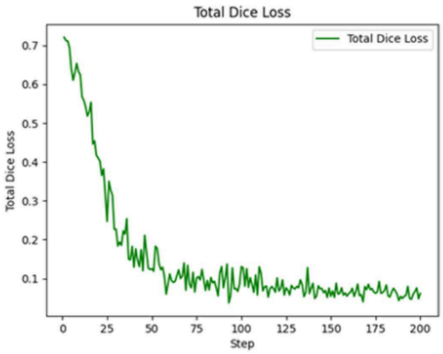

Figure 4 Total Dice Loss

Conclusion :-

In conclusion, this study evaluated the effectiveness of Segtran, a transformer-based neural network for brain glioma segmentation using the BraTS MRI dataset and benchmarked its performance against the widely used U-Net architecture. The model demonstrated superior segmentation performance under optimal conditions, including an EfficientNet-B1 backbone, three transformer layers, the AdamW optimizer, a batch size of 5, a learning rate of 0.0001, and 200 training epochs. The Transformer model achieved a Dice coefficient of 0.921, significantly outperforming U-Net, which achieved 0.827. These results highlight the Transformer Network's capability to deliver more accurate and efficient tumor segmentation, supporting its potential as a valuable tool in medical image analysis.

Acknowledgement:-

The authors would like to acknowledge the Ministry of Higher Education Malaysia (MOHE) for the Fundamental Research Grant Scheme (FRGS) Project Code: FRGS/1/2024/TK07/UKM/02/6 and the Universiti Kebangsaan Malaysia.

References :-

1. Afshar, P., Mohammadi, A., & Plataniotis, K.N. 2018. Brain Tumor Type Classification via Capsule Networks. Proceedings - International Conference on Image Processing, ICIP, pp. 3129-3133. IEEE Computer Society.
2. Ahamed, M. F., Hossain, M. M., Nahiduzzaman, M., Islam, M. R., Islam, M. R., Ahsan, M., & Haider, J. 2023. A review on brain tumor segmentation based on deep learning methods with federated learning techniques. Computerized Medical Imaging and Graphics, 110, 102313. https://doi.org/10.1016/j.compmedimag.2023.102313

3. Chen, J., Mei, J., Li, X., Lu, Y., Yu, Q., Wei, Q., Luo, X., Xie, Y., Adeli, E., Wang, Y., Lungren, M. P., Zhang, S., Xing, L., Lu, L., Yuille, A., & Zhou, Y. 2024. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. Medical image analysis, 97, 103280. https://doi.org/10.1016/j.media.2024.103280
4. Despotović, I., Goossens, B., & Philips, W. 2015. MRI Segmentation of the Human Brain: Challenges, Methods, and Applications. Computational and Mathematical Methods in Medicine, 2015(1), 450341. https://doi.org/10.1155/2015/450341
5. Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. 2017. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. Annual Conference on Medical Image Understanding and Analysis, pp. 506-517.
6. Fuzari, H.K.B., Dornelas de Andrade, A., Vilar, C.F., Sayão, L.B., Diniz, P.R.B., Souza, F.H., & de Oliveira, D.A. 2018. Diagnostic accuracy of magnetic resonance imaging in post-traumatic brachial plexus injuries: A systematic review. Clinical Neurology and Neurosurgery, 164, pp. 5-10.
7. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., & Larochelle, H. 2017. Brain tumor segmentation with Deep Neural Networks. Medical Image Analysis, 35, pp. 18-31.
8. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., & Maier-Hein, K.H. 2018. Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. International MICCAI Brainlesion Workshop, pp. 287-297.
9. Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis, 36, 61-78. https://doi.org/10.1016/j.media.2016.10.004
10. Kao, PY., Ngo, T., Zhang, A., Chen, J.W. & Manjunath, B.S. 2019. Brain Tumor Segmentation and Tractographic Feature Extraction from Structural MR Images for Overall Survival Prediction. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Lecture Notes in Computer Science(), vol 11384, pp 128-14. https://doi.org/10.1007/978-3-030-11726-9_12
11. Li, S., Sui, X., Luo, X., Xu, X., Liu, Y., & Goh, R. (2021). Medical Image Segmentation Using Squeeze-and-Expansion Transformers. ArXiv. https://arxiv.org/abs/2105.09511
12. Liu, K., & Liang, Y. (2023). A fusion-attention swin transformer for cardiac MRI image segmentation. Iet Image Processing. https://doi.org/10.1049/ipr2.12936
13. Long, J., Shelhamer, E., & Darrell, T. 2015. Fully Convolutional Networks for Semantic Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431-3440.
14. Loshchilov, I., & Hutter, F. 2019. Decoupled Weight Decay Regularization. Proceedings of the International Conference on Learning Representations (ICLR).
15. Myronenko, A. 2018. 3D MRI brain tumor segmentation using autoencoder regularization. ArXiv. https://arxiv.org/abs/1810.11654
16. Pu, Q., Xi, Z., Yin, S., Zhao, Z. & Zhao. L. 2024. Advantages of transformer and its application for medical image segmentation: a survey. BioMedical Engineering OnLine, 23 (14). https://doi.org/10.1186/s12938-024-01212-4
17. Sajun, A. R., Zualkernan, I., & Sankalpa, D. 2023. A Historical Survey of Advances in Transformer Architectures. Applied Sciences, 14(10), 4316. https://doi.org/10.3390/app14104316
18. Takahashi, S., Sakaguchi, Y., Kouno, N., Takasawa, K., Ishizu, K., Akagi, Y., Aoyama, R., Teraya, N., Bolatkan, A., Shinkai, N., Machino, H., Kobayashi, K., Asada, K., Komatsu, M., Kaneko, S., Sugiyama, M., & Hamamoto, R. 2024. Comparison of Vision Transformers and Convolutional Neural Networks in Medical Image Analysis: A Systematic Review. Journal of medical systems, 48(1), 84. https://doi.org/10.1007/s10916-024-02105-8
19. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. 2017. Attention Is All You Need. 31st Conference on Neural Information Processing Systems (NIPS 2017) ArXiv. https://arxiv.org/abs/1706.03762
20. Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., & Vercauteren, T. 2019. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. Neurocomputing, 335, 34. https://doi.org/10.1016/j.neucom.2019.01.103
21. Zhou, Y., Wang, X., Zhang, M., Zhu, J., Zheng, R., & Wu, Q. MPCE: A Maximum Probability Based Cross Entropy Loss Function for Neural Network Classification in IEEE Access, vol. 7, pp. 146331-146341, 2019, doi: 10.1109/ACCESS.2019.2946264.
22. Verkhratsky, A., & Butt, A. 2023. Neuroglia: function and pathology. Elsevier.

TRANSFORMER NETWORK FOR BRAIN GLIOMA SEGMENTATION IN MRI IMAGES

ORIGINALITY REPORT

31%	23%	20%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.journalijar.com Internet Source	5%
2	repository.stiki.ac.id Internet Source	3%
3	arno.uvt.nl Internet Source	2%
4	www.mdpi.com Internet Source	2%
5	eitca.org Internet Source	1%
6	www.doria.fi Internet Source	1%
7	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Springer Science and Business Media LLC, 2020 Publication	1%
8	Submitted to Asia Pacific University College of Technology and Innovation (UCTI) Student Paper	1%
9	www.nature.com Internet Source	1%
10	"Information Processing in Medical Imaging", Springer Science and Business Media LLC, 2019 Publication	1%
11	Submitted to Universiti Kebangsaan Malaysia Student Paper	

		1 %
12	Submitted to Virginia Polytechnic Institute and State University Student Paper	1 %
13	Submitted to Swinburne University of Technology Student Paper	1 %
14	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2019", Springer Science and Business Media LLC, 2019 Publication	1 %
15	iris.poliba.it Internet Source	1 %
16	arxiv.org Internet Source	1 %
17	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Springer Nature, 2019 Publication	<1 %
18	Submitted to Brunel University Student Paper	<1 %
19	dokumen.pub Internet Source	<1 %
20	ebin.pub Internet Source	<1 %
21	www.scirp.org Internet Source	<1 %
22	Chao Li, Xinggang Wang, Wenyu Liu, Longin Jan Latecki. "DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks", Medical Image Analysis, 2018	<1 %

	Publication	
23	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2024", Springer Science and Business Media LLC, 2024 Publication	<1 %
24	Submitted to The University of Manchester Student Paper	<1 %
25	openprairie.sdstate.edu Internet Source	<1 %
26	Naseeb Singh, V.K. Tewari, P.K. Biswas. "Vision transformers for cotton boll segmentation: Hyperparameters optimization and comparison with convolutional neural networks", Industrial Crops and Products, 2025 Publication	<1 %
27	export.arxiv.org Internet Source	<1 %
28	www.arxiv-vanity.com Internet Source	<1 %
29	Shoffan Saifullah, Rafał Dreżewski, Anton Yudhana, Maciej Wielgosz, Wahyu Caesarendra. "Modified U-Net with attention gate for enhanced automated brain tumor segmentation", Neural Computing and Applications, 2025 Publication	<1 %
30	unidel.edu.ng Internet Source	<1 %
31	www.researchgate.net Internet Source	<1 %
32	"Ophthalmic Medical Image Analysis", Springer Science and Business Media LLC,	<1 %

	2022 Publication	
33	Submitted to University of Westminster Student Paper	<1%
34	associatedgroups.com Internet Source	<1%
35	www.biorxiv.org Internet Source	<1%
36	Lalit Mohan Goyal, Tanzila Saba, Amjad Rehman, Souad Larabi-Marie-Sainte. "Artificial Intelligence and Internet of Things - Applications in Smart Healthcare", CRC Press, 2021 Publication	<1%
37	Piyush Bhushan Singh, Pawan Singh, Harsh Dev, Devanshu Batra, Brijesh Kumar Chaurasia. "HViTML: Hybrid vision transformer with machine learning-based classification model for glaucomatous eye", Multimedia Tools and Applications, 2025 Publication	<1%
38	Reza Azad, Amirhossein Kazerouni, Moein Heidari, Ehsan Khodapanah Aghdam et al. "Advances in medical image analysis with vision Transformers: A comprehensive review", Medical Image Analysis, 2023 Publication	<1%
39	link.springer.com Internet Source	<1%
40	www.frontiersin.org Internet Source	<1%
41	www.jianshu.com Internet Source	<1%

42	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2020", Springer Science and Business Media LLC, 2020 Publication	<1 %
43	Asim Zaman, Mazen M. Yassin, Irfan Mehmud, Anbo Cao, Jiaxi Lu, Haseeb Hassan, Yan Kang. "Challenges, optimization strategies, and future horizons of advanced deep learning approaches for brain lesion segmentation", Methods, 2025 Publication	<1 %
44	Mrutyunjaya Panda, Ajith Abraham, Biju Gopi, Reuel Ajith. "Computational Intelligence for Oncology and Neurological Disorders - Current Practices and Future Directions", CRC Press, 2024 Publication	<1 %
45	Mushtaq Mahyoob Saleh, Musab Elkheir Salih, Mohamed A. A. Ahmed, Altahir Mohamed Hussein. "From Traditional Methods to 3D U-Net: A Comprehensive Review of Brain Tumour Segmentation Techniques", Journal of Biomedical Science and Engineering, 2025 Publication	<1 %
46	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical and Computer Technologies", CRC Press, 2025 Publication	<1 %
47	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2021", Springer Science and Business Media LLC, 2021 Publication	<1 %
48	Hung P. Do, Yi Guo, Andrew J. Yoon, Krishna S. Nayak. "Accuracy, uncertainty, and	<1 %

adaptability of automatic myocardial ASL segmentation using deep CNN", Magnetic Resonance in Medicine, 2019
Publication

49	Mehdi Ghayoumi. "Generative Adversarial Networks in Practice", CRC Press, 2023 Publication	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		