ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/27-36
E. Sri Varshini *et. al.*, / International Journal of Engineering & Science Research

Brain Tumor Detection using UNET and Classification using Transformers

E. Sri Varshini[1], J. Manoghnaa[2], S. Spurthi[3], R. Sirisha[4]
Department of Artificial Intelligence and Data Science [1,2,3,4]
Stanley College Of Engineering and Technology for Women, Hyderabad[1,2,3,4]

ABSTRACT

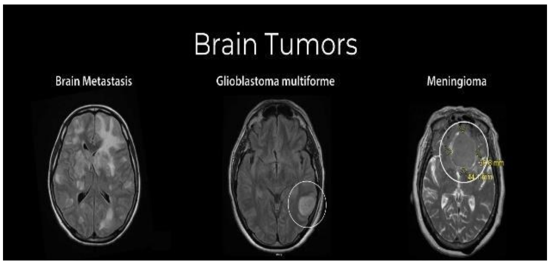
Patients can be recovered easily if the any tumor is at its earlier stage. In this study brain tumor is classified using Magnetic Resonance Imaging (MRI) scan of the brain. Previously there is need of huge medical staff as well as more time for brain tumor diagnosis. Recently there is advancement in the medical diagnosis which makes process automatic diagnosis and it is able to classify MRI scan images. Many machine learning techniques has been implemented for brain tumor diagnosis but there are some limitations such as huge time required for training and lower accuracy. So, to overcome the limitations of existing techniques in this project 3D UNET is used for identification of location of tumor. Transformer based deep learning model is used for classification of tumor into 4 different categories. Performance of proposed algorithm with existing algorithm such as VGG16 is compared to show the superiority of proposed transformer-based model. Proposed model based on transformers is more efficient than state of art techniques which is measured using accuracy, precision, recall and F-score. To show the performance comparison between two algorithms bar graph is plotted which shows the superior performance of proposed transformer-based model for brain tumor classification.

Keywords: Magnetic Resonance Imaging (MRI) , Brain tumor identification and segmentation, UNET segmentation, Transformers.

I. INTRODUCTION

The accurate segmentation of brain tumors and early detection stands essential for successful therapeutic planning because these conditions become lethal medical conditions. Brain tumor detection requires MRI because it produces exceptional tissue visualization and contrast that helps identify brain cancer. The manual procedure for separating brain tumors from MRI scans proves lengthy and error-prone to human evaluators while needing advanced medical expertise. Deep learning-based automated techniques serve as effective instruments to boost the accuracy of brain tumor segmentation alongside their performance in classification tasks.

FCNNs particularly UNET have established themselves as top performers for medical image segmentation tasks. The UNET architecture with its encoder-decoder framework supported by skip connections maintains efficient performance by processing both precise spatial information and contextual overview details thus being powerful for segmenting tumors with different shapes and dimensions. The limited receptive field of UNET leads to challenges in modeling long-range dependencies because this limitation affects the quality of segmented complex tumor structures.



Fg.1.1 Different types of Brain Tumor

Basically, in the above image the brain tumor is classified into three different categories based on the presence of tumor n particular region.

Medical diagnostic procedures and treatment planning depend heavily on the segmentation process of brain tumors using MRI scans. UNET and similar traditional CNN models achieve widespread usage since they excel at feature extraction. The issue with such networks is their inability to detect extensive dependencies between features so they produce substandard tumor segmentations for complex anatomical structures. To overcome this problem new studies have merged Transformers with existing UNET frameworks. A transformer-enhanced residual UNET operates within TransResUNET to perform glioma segmentation through self-attention mechanisms that enhance feature expression [1]. The combination of transformer-based architecture with CNN in BiTr-UNET makes MRI brain tumor segmentation better than using individual CNN models while also reaching enhanced performance levels [2].

Different combination methods have enabled better segmentation performance. The medical imaging technology ETUNET merges efficient Transformer blocks with UNET to optimize 3D tumor segmentation operations [4]. Swin UNETR uses Swin Transformers to execute brain tumor semantic segmentation and extracts hierarchical features for contextual multi-scale information detection [5]. The HUT architecture combines parallel execution of UNET and Transformer-based pipelines in order to enhance MRI dataset segmentation performance [7]. Brain tumor analysis benefits from segmentation precision improvements when UNET operates together with Transformers.

The attention of researchers has shifted toward Transformer-based models because their self-attention capability allows them to detect distant relationships between model inputs. The network performs effective tumor differentiation and enhanced diagnostic precision by implementing Transformers in classification operations following UNET-based segmentation algorithms. The joint use of UNET for segmentation with Transformers for classification establishes a strong system which detects tumors while properly determining benign versus malignant clusters to support medical decision processes.

The proposed framework combines UNET for brain tumor segmentation with Transformers for determining accurate tumor type classifications. The proposed framework integrates convolutional network spatial feature extraction with Transformer-based long-range feature dependency learning features. The proposed model demonstrates superior performance than conventional CNN-based algorithms according to experimental benchmark testing which shows its practical medical application viability.

II. LITEARTURE SURVEY

The research suggests using a deep-learning method to perform accurate MRI scan brain tumor segmentation through a hybrid ResNet U-Net structure with Transformer blocks. A model merges hierarchically extracted features through ResNets combined with the attention capabilities of Transformers to handle long-distance connections. Spatial consistency during decoding becomes improved through the adoption of transposed convolutions together with skip connections. The BraTS2019 dataset evaluation showed the model produced exceptional results such as dice scores measuring 0.91, 0.89, and 0.84 for whole, core, and enhancing tumor regions together with a total accuracy of 98%. [1]

The research work presents BiTr-UNET as a combined CNN-Transformer framework dedicated to multi-modal MRI brain tumor segmentation tasks. Through the combination of CNNs for extracting local features with transformers managing distance dependencies the model achieves better segmentation results. BiTr-UNET achieved top performance on BraTS2021 by delivering median Dice scores of 0.9335 and 0.9304 and 0.8899 with minimal Hausdorff distances which demonstrates its ability to segment tumors precisely. [2]

An integrated UNET and transformer-based modules system serves as the solution for brain tumor segmentation in this research. The proposed approach improves segmentation accuracy because it combines local features of CNNs with a transformer-based attention mechanism which captures global patterns. The model obtained substantial dice scores of 0.94, 0.921, 0.83, and 0.94 during evaluation on BraTS datasets from 2015 to 2021 thus proving its capability for automatic brain tumor segmentation.[3]

The research brings value to brain tumor segmentation methods by integrating Transformer modules into the UNET structure to rectify the long-range dependency limitations of CNNs. A new CNN-Transformer encoder together with a spatial-channel attention layer in the bottleneck system and cross-attention-based skip connections has been proposed to enhance both feature utilization and boundary localization capabilities. The model reached excellent performance results in both BraTS2018 (0.854 DSC and 6.688 HD95) and BraTS2020 (0.862 DSC and 5.455 HD95) by achieving high DSC values with low HD95 rates.[4]

The research establishes Swin UNETR as a combined model which integrates Swin Transformer features with U-Net capabilities for executing 3D brain tumor segmentation tasks. Through the integrated shifted window self-attention and multi-resolution feature extraction system the model successfully connects distant dependencies together with maintaining spatial information. Swin Transformer operates as the hierarchical encoder within the network and the decoder features an FCNN structure with skip connections for better segmentation results. During BraTS 2021 evaluation Swin UNETR achieved validation results that placed it among the leading performers for medical image segmentation.[5]

In this work you develop a new 3D UNET-Transformer hybrid model for brain tumor segmentation that merges Contextual Transformer (CoT) and Double Attention (DA) attention systems. Both CoT components allow better extraction of contextual features and DA elements enhance the handling of long-range dependencies during the process. The conducted experiments demonstrate higher segmentation accuracy than current methods do. Through this method medical image processing becomes faster which benefits healthcare professionals during diagnostic and therapeutic planning work processes. Please let me know whether you require more details and refining work.[6]

Your Hybrid UNET Transformer (HUT) model produces better results for single-modality lesion segmentation and multi-modality brain tumor segmentation through dual implementation of UNET and Transformer pipelines simultaneously. The Transformer pipeline implements global attention capabilities as well as long-range dependencies through the utilization of UNET decoder outputs. The self-supervised training method enhances segmentation precision through its implementation. HUT provides superior performance to SPiN on ATLAS data through improved Dice scores by 4.84% and Hausdorff Distance by 41% yet also leads nnUNET to enhanced results on BraTS20 by achieving 0.96% increased Dice score and a 4.1% reduction in Hausdorff Distance. Additional input is available if you need more adjustments to this text.[7]

III.PROPOSED METHOD

Proposed model uses UNET for segmentation while transformer models for classification of brain MRI scan.

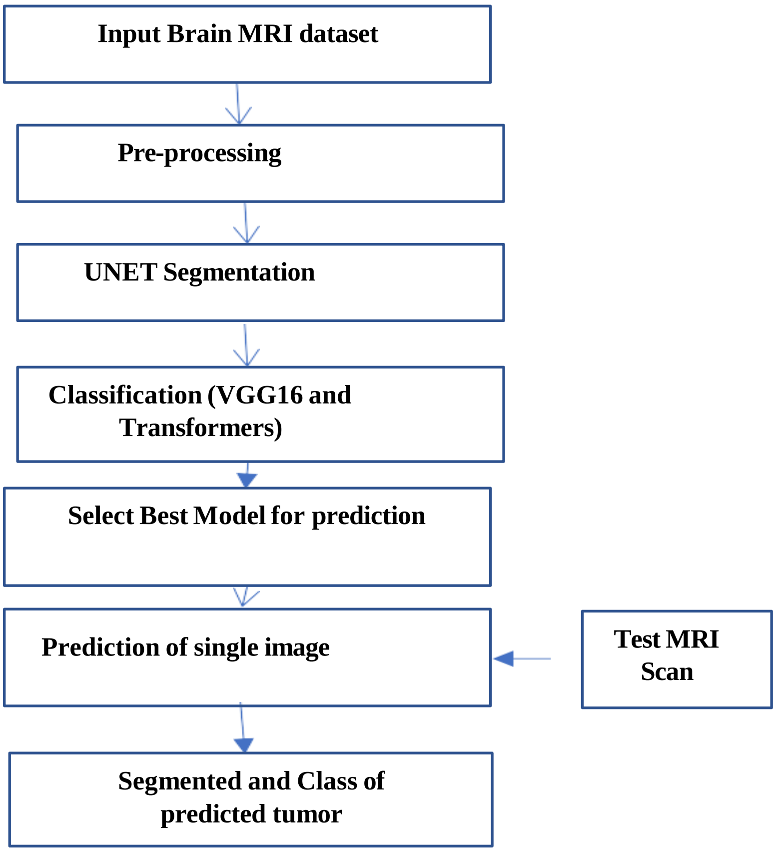


Fig. 3.1 Block Diagram of Proposed Model

The brain tumor detection system uses U- Net segmentation technology coupled with Transformer deep learning models for classification needs. The workflow includes data preparation followed by segmentation of lesions and then the extraction of features until the last stage where classification occurs.

The preprocessing of MRI images includes three sequential operations consisting of noise reduction along with intensity normalization and contrast enhancement to enhance segmentation quality. Medical image segmentation depends on U-Net which functions as a recognized deep learning architecture for identifying tumor regions precisely. The model operates successfully by recognizing global and local image features which results in ideal tumor boundary recognition.

The obtained essential tumor characteristics like shape, size and texture emerge from the subsequent feature extraction process. The

classification model based on Transformers receives extracted features before utilizing self-attention to discover sophisticated spatial relations which enhance accuracy levels. The approach produces effective separation between tumors that are benign or malignant.

This method intends to boost tumor detection precision through its combination of state-of-the-art segmentation with classification components. Precise medical image segmentation depends on U-Net's capabilities while the robust classification functions are possible with Transformers' excellent feature extraction abilities. Future experiments predict that the proposed combined model will achieve superior performance levels in accuracy and computational speed when compared to traditional CNN-based systems.

Proposed model steps are explained below,

a) MRI scan Dataset

In this step dataset of brain tumor is selected which has different types of brain tumor scans based on their cell of presence.

b) Preprocessing on images

The input scan is given to preprocessing which uses normalization and shuffling. Normalization helps to reduce the complexity in processing. Shuffling helps understanding the images n better way.

c) UNET based Brain tumor segmentation

UNET based segmentation helps to get the exact region of tumor present in the brain MRI scan. UNET is deep learning-based algorithm which helps to understand the image in better way.

d) Existing VGG16 and Proposed Transformer model

VGG16 is the pretrained model in our application VGG16 is considered as existing state of art model. This model has lower accuracy.

e) Prediction on test data using Proposed model

Proposed model uses UNET segmentation and Transformer deep learning model is used for classification. Proposed model weights are saved and those weights are used for further prediction.

IV. RESULT ANALYSIS

The proposed detection method consisting of both transformer components and U-Net segments effectively identifies brain tumors within MRI scans. The U-Net architecture properly locates tumor regions through segmentation before transformers use their spatial dependency capabilities in image classification. The transformer model achieved superior performance through key metrics including accuracy and precision as well as recall and F1-score since it surpassed traditional CNN-based approach results.

The experimental results demonstrate the model successfully classifies tumors with solid performance and excellent discrimination between tumor types. Transformers installed alongside other systems produce advanced feature extraction capability which allows better universal applicability to various datasets. Performance evaluation demonstrates that this U-Net transformer combination is an ideal approach to medical image processing and enables automatic brain tumor detection.

Fig.4.1 Overall GUI for Proposed Model

Using tkinter library from python GUI is implemented. GUI is for easy understanding of users or patients.

Dataset Link is :

https://www.kaggle.com/datasets/masoudnickparva r/brain-tumor-mri-dataset

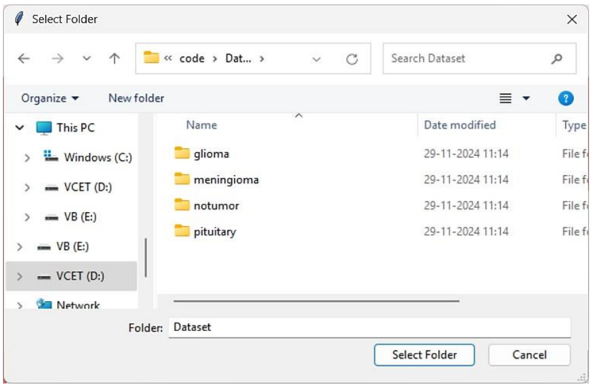


Fig. 4.2 Different Types of Brain Tumors

Proposed application has four different types of brain tumors. These types are based on tumor present in area of brain region.

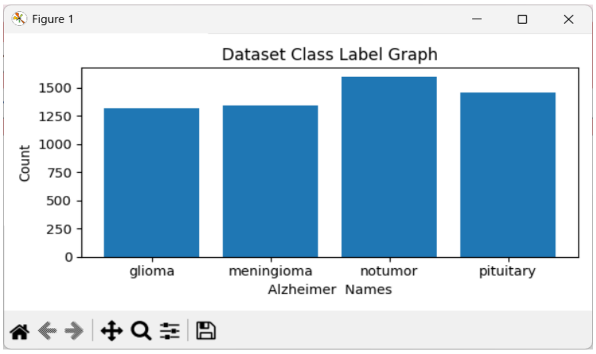


Fig. 4.3 Count of each type of Tumor Class

In the dataset each class count is calculated and shown using bar graph. Each bar represents the count of the respective class.

Fig.4.4 Total Images Present in the Dataset is Shown

Total images present n the dataset are 5712 images. These images are from 4 different categories.

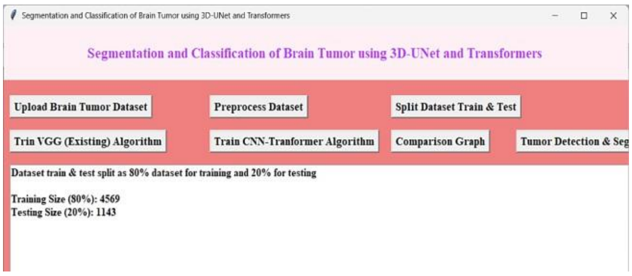


Fig. 4.5 Data Splitting (Training and Testing Division)

Our dataset is divided into two parts as training and testing part. Dataset splitting is used for model training and model testing.

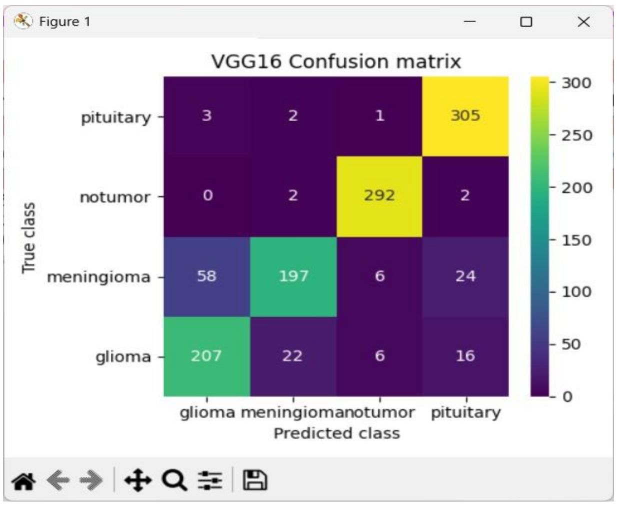


Fig. 4.6 Confusion Matrix for Existing Method

In this application VGG16 is designed as state of art technique as it is giving lower accuracy compared to proposed transformer model.

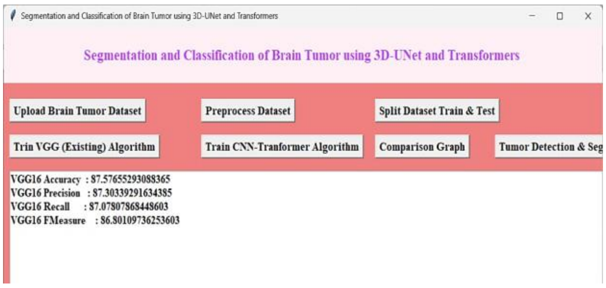


Fig. 4.7 Performance of the existing technique

For existing technique VGG16, different metrics such as accuracy, precision, recall and f-score are calculated.

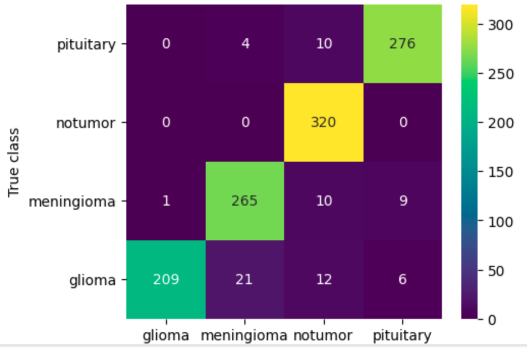


Fig. 4.8 Confusion matrix of proposed transformer-based model

Proposed transformer-based model for prediction gives better accuracy compared to state of art technique.

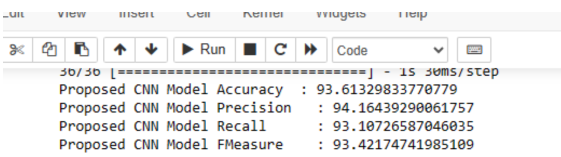


Fig. 4.9 Performance of the Proposed technique

Proposed transformer-based technique is implemented and tested on 20% test data to calculate the performance.

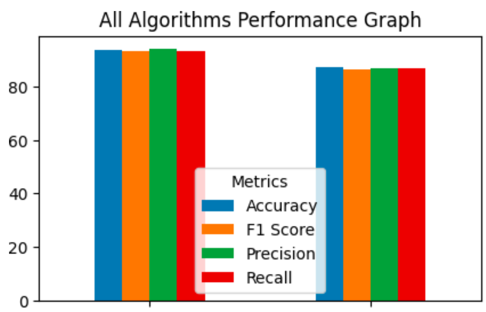


Fig. 4.10 Performance Comparison between a) Proposed model b) Existing VGG16 model

Based on above bar graph we can say that proposed transformer model works far better than state of art techniques.

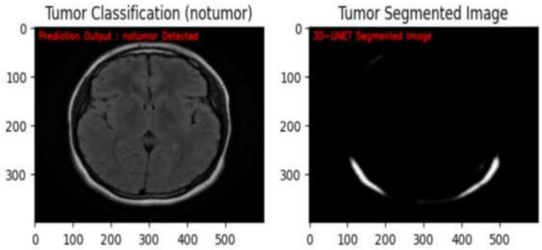


Fig. 4.11 a) Tumor Classification results b) Segmentation Results

It is observed that predicted results by the proposed transformer-based model is 'No Tumor'.

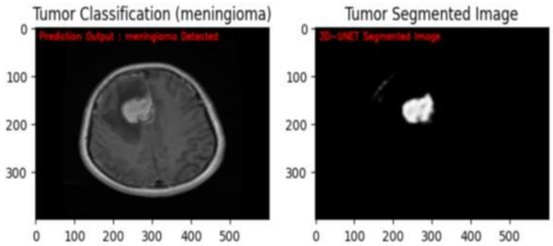


Fig. 4.11 Proposed model prediction results a) Tumor Classification results b) Segmentation Results

In a) image is classified as meningioma tumor and respective segmentation is shown in figure b).

V. CONCLUSION

Proposed model is successfully designed using python and graphical user interface for demonstration. UNET is the deep learning structure which is used for image segmentation and marking the region on image is also used which is advancement in segmentation. Transformer models can classify MRI brain scan into different categories as per the training. For demonstrate the proposed model we used GUI (Graphical User Interest). Proposed model is a type of computer aided diagnostic technique which makes application to make automatic classification without manual intervention. Comparative analysis is performed between existing VGG16 and proposed UNET based transformer model for classification. It is observed from bar graph visualization that existing state of art techniques has degraded performance over proposed UNET based transformer model.

REFERENCES

1. Rasool, Novsheena, Javaid Iqbal Bhat, Niyaz Ahmad Wani, Naveed Ahmad, and Mohammed Alshara. "TransResUNET: revolutionizing glioma brain tumor segmentation through transformer-enhanced residual UNET." *IEEE Access* (2024).
2. Jia, Qiran, and Hai Shu. "Bitr-UNET: a cnn-transformer combined network for mri brain tumor segmentation." In *International MICCAI brainlesion workshop*, pp. 3-14. Cham: Springer International Publishing, 2021.
3. Shedbalkar, Jayashree, K. Prabhushetty, and R. H. Havaldar. "UNET and Transformer-Based Model for Multi-Modality Brain Tumor Segmentation." *SSRG Int. J. Electron. Commun. Eng* 10 (2023): 22-35.
4. Zhang, Wang, Shanxiong Chen, Yuqi Ma, Yu Liu,

and Xu Cao. "ETUNET: Exploring efficient transformer enhanced UNET for 3D brain tumor segmentation." *Computers in Biology and Medicine* 171 (2024): 108005.

5. Hatamizadeh, Ali, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R. Roth, and Daguang Xu. "Swin UNETr: Swin transformers for semantic segmentation of brain tumors in mri images." In *International MICCAI brainlesion workshop*, pp. 272-284. Cham: Springer International Publishing, 2021.
6. Nguyen-Tat, Thien B., Thien-Qua T. Nguyen, Hieu- Nghia Nguyen, and Vuong M. Ngo. "Enhancing brain tumor segmentation in MRI images: A hybrid approach using UNET, attention mechanisms, and transformers." *Egyptian Informatics Journal* 27 (2024): 100528.
7. Soh, Wei Kwek, Hing Yee Yuen, and Jagath C. Rajapakse. "HUT: Hybrid UNET transformer for brain lesion and tumour segmentation." *Heliyon* 9, no. 12 (2023).
8. Liang, Junjie, Cihui Yang, and Lingguo Zeng. "3D PSwinBTS: An efficient transformer-based UNET using 3D parallel shifted windows for brain tumor segmentation." *Digital Signal Processing* 131 (2022): 103784.
9. Tiwary, Pradeep Kumar, Prashant Johri, Alok Katiyar, and Mayur Kumar Chhipa. "Deep Learning-Based MRI Brain Tumor Segmentation with EfficientNet- Enhanced UNET." *IEEE Access* (2025).
10. Lyu, Yu, and Xiaolin Tian. "MWG-UNET++: Hybrid Transformer U-Net Model for Brain Tumor Segmentation in MRI Scans." *Bioengineering* 12, no. 2 (2025): 140.