

Hierarchical Brain Tumor Detection And Classification Using U-Net-Dcwhdnn Intelligent Soft Computing Learning Architectures

Ganesh Ramchandra Patil^{*1}, Dr. Shailendrakumar Mahadeo Mukane²

^{*1}Ph.D Research Scholar SVPM's College Of Engineering, Malegaon Bk, Baramati, Maharashtra-413115, India. Savitribai Phule Pune University, Pune.

²Principal/Research Guide, SVPM's College of Engineering, Malegaon Bk, Baramati, Maharashtra-413115, India. Savitribai Phule Pune University, Pune.

Abstract

Brain tumor is one of the leading causes of mortality and it requires early diagnosis. The diverse characteristics of brain tumor makes the segmentation and detection process complicated. This paper presents a novel approach, deep convolutional wild horse deformable neural network model (DCWDNN) for accurate disease diagnosis. The deep learning approach, UNet transformer architecture with self-attention encoder and decoder (UNet-TSA) is involved for segmenting tumor region by capturing long range dependencies. Moreover, the self-attention module is added to extract the fine grained feature information with minimal computational complexity. The accuracy of proposed tumor detection approach UNet-TSA is enhanced with wild horse soft computing (WHSC) based hyper parameter optimization approach. It ensures robustness to resolve problem in complex, and high dimensional space. The proposed approach is evaluated with BraTS 2018 dataset in which the accuracy, precision, recall performance is improved to 0.998, 0.991, and 0.992 with the Mean square error (MSE) of 0.0020. The performance of the proposed approach is compared with state-of-the-art DL approaches to show the effectiveness of proposed approach.

Index terms: Brain tumor segmentation, UNet transformer, self-attention encoder decoder, tumor detection, deformable neural network, wild horse soft computing.

1. Introduction

The emergence of brain tumors stems from the unchecked and swift proliferation of cells, posing a serious threat to life if left untreated in its early stages [1]. It was estimated that approximately 0.7 million individuals in America were living with brain tumors. Of these, 0.86 million cases were diagnosed, with 60,800 patients classified as benign and 26,170 as malignant. The survival rate for malignant patients in the US is reported to be 35% [2]. In recent years, brain tumor segmentation and detection are accomplished with machine learning (ML) and deep learning (DL) approaches. The inherent advantages of these approaches contribute to more accurate results in the detection of brain tumors [3]. Magnetic Resonance Imaging (MRI) serves as a widely utilized non-invasive imaging technology for the detection, diagnosis, and treatment monitoring of brain diseases. By generating detailed three-dimensional images, MRI enables neurologists to accurately pinpoint abnormalities in brain images. This advanced imaging modality plays a crucial role in facilitating precise diagnosis and treatment planning for various neurological conditions [4-5]. Major tumor types which are categorized using the developed deep learning approaches are benign (noncancerous) and malignant (cancerous) [6]. Segmentation of tumors means isolating and delineating the tumor region from the surrounding healthy brain tissue in MRI scans. Some models initially segmented and then classified into respective classes for enhancing the performance [7]. Most of the models use the hand craft and deep learning feature still have a moderate results in the detection due to high computation [8].

Different models can be used for segmentation and classification of brain tumor that aims to provide better performance. Support vector machine (SVM) is one of the classifier used

for the brain tumors which possess large labeled data [9]. One of the most familiar deep learning models in classification is convolutional neural network (CNN). These CNN model is used for brain tumor detection may also possess less performance due to vanish gradient problems [10-11]. Some researches handle this issue by using variants of recurrent neural networks (RNN) like long short term memory (LSTM) etc. Even though these models provided the Computational expense which makes the detection difficult [12-13]. Some model used the auto encoders to provide segmented portions which give better results with high complexity [14-15]. The model that provide a good and effective way to detect brain tumor based on MRI to help the doctor in making decisions is still on rely.

Despite significant progress in brain tumor detection and classification, precise segmentation and classification remain challenging due to the diverse array of tumor characteristics. Tumors vary in location, shape, and size, complicating the diagnostic process. Existing models face hurdles such as limited generalizability, interpretability, and efficiency. Generalizability issues arise when models struggle to perform accurately on new data outside the training dataset, leading to unreliable predictions. Lack of interpretability hinders clinicians' understanding of the models' decision-making processes, crucial for informed decision-making in clinical settings. Additionally, the need for large and expensive datasets for training deep learning models poses a significant challenge. Acquiring annotated medical imaging data is time-consuming and labor-intensive.

To address these challenges, innovative approaches such as transfer learning, explainable AI, and data augmentation techniques are essential. Collaborative efforts between researchers, clinicians, and industry stakeholders are crucial for developing robust diagnostic tools that improve the accuracy, efficiency, and interpretability of brain tumor detection and classification. Hence, these challenges motivated as to present a novel deep learning model that provides solution for the existing issues.

The remaining contribution is listed as follows:

- To segment the tumor region with a novel Unet transformer based segmentation approach. It captures the long the long range dependencies with less computational overhead. The fine grained information is extracted during segmentation process using self-attention and skip connection.
- To enable accurate diagnosis of brain tumor using DCWHDNN that tunes the DL parameter to produce optimal performance. The deformable convolution layers are utilized to enhance the robustness and generalization procedure.
- To select optimal parameter by resolving slow convergence and local optima issue in complex high dimensional space. Also, the performance of the proposed approach is evaluated with standard metrics and compared with state-of-the-art approaches.

The paper organization is given as follows. Section 2 provides the analysis and comparison of existing researches. Section 3 provides detailed description of proposed methodology with architecture. The experimental details, results, and discussion are given in section 5. Finally, the conclusion and future directions are provided in section 5.

2. Related work

Some of the recent works on brain tumor detection using machine learning and deep learning approaches are described as follows.

Patil et al. [16] designed an ensemble model with shallow convolutional neural network (SCNN) and VGG16 with T1C modality MRI image. The collected features from both deep learning models were merged to increase the classification accuracy of three different

types of tumors, hence improves the model's performance in terms of accuracy and loss information. The ensemble deep convolutional neural network model (EDCNN) shown that the combination of deep learning models aims to increase multiclass classification problem accuracy. In the evaluation, the model achieved 97.77% accuracy in Brain Tumor Detection. Moreover, the model possesses high computational cost, less robustness and Interpretability challenges.

Ibrahim et al. [17] introduced a novel hybrid model that improves detection and classification capabilities by fusing CNNs with the Particle Swarm Optimization (PSO) technique. The PSO algorithm was used to find the CNN hyper-parameter configuration that performs efficiently. The CNN architectures were then fitted with these optimal parameters for classification. For Alzheimer's disease, two datasets were used in this work, a worldwide Kaggle dataset and the Alzheimer's disease Neuroimaging Initiative (ADNI). The experimental evaluation proved the effectiveness of the suggested model, obtaining previously unheard-of accuracy rates of 98.50%, 98.83%, and 97.12% for the aforementioned datasets, respectively. This model assures scalability and resolves the challenges related to generalization.

Archana et al. [18] developed a unique way for diagnosing brain malignancies using the Bagging Ensemble with K-Nearest Neighbor (BKNN) which improves the classification accuracy. Initially, a U-Net architecture is used for segmenting tumor region. Then, the k-NN prediction algorithm is executed based on bagging technique. Unet was utilized based on the objective of enhancing the accuracy and uniformity of parameter distribution in the layers. During classification, each decision tree was fitted on a slightly different training dataset. Since, each tree has little variations and produces slightly different predictions. Bagged decision trees are effective for classification and it enhances the accuracy up to 97.7% throughout the entire process. The model has less efficiency due to expensive and large datasets to calculate the distance among each new point and each old point.

Kanchanamala et al. [19] suggested an effective detection approach called Exponential deer hunting optimization-based Shepard CNN (ExpDHO-based ShCNN) for the identification and classification of brain tumors. Tumor identification employs ShCNN approach, which was trained with the developed optimization method known as ExpDHO. Finally, tumor categorization (DHOA) was carried out using the developed ExpDHO method, which combines the Exponential weighted moving average (EWMA) with the Deer Hunting Optimization method. In addition, the developed method produced positive outcomes when measured by standard performance metrics. The accuracy, sensitivity and specificity of brain tumor detection were reached to 0.929, 0.934, and 0.939 respectively. In case of brain tumor classification 0.917 accuracy, 0.918 sensitivity, and 0.919 specificity was attained with lack of reliability.

Woźniak et al. [20] provided a correlation learning approach (CLM) for deep learning model that relies on traditional neural network and CNNs. CNN was assisted as a base model for identifying the informative region with pooling and convolution layer. The primary neural classifier learns more quickly and becomes more efficient for tumor diagnosis. The CLM model can achieve approximately 96% accuracy, as well as approximately 95% precision and recall, according to the results. Processing large dataset involves generalizability and interpretability challenges. The overview of issues that arise in the existing system is depicted in Table 1.

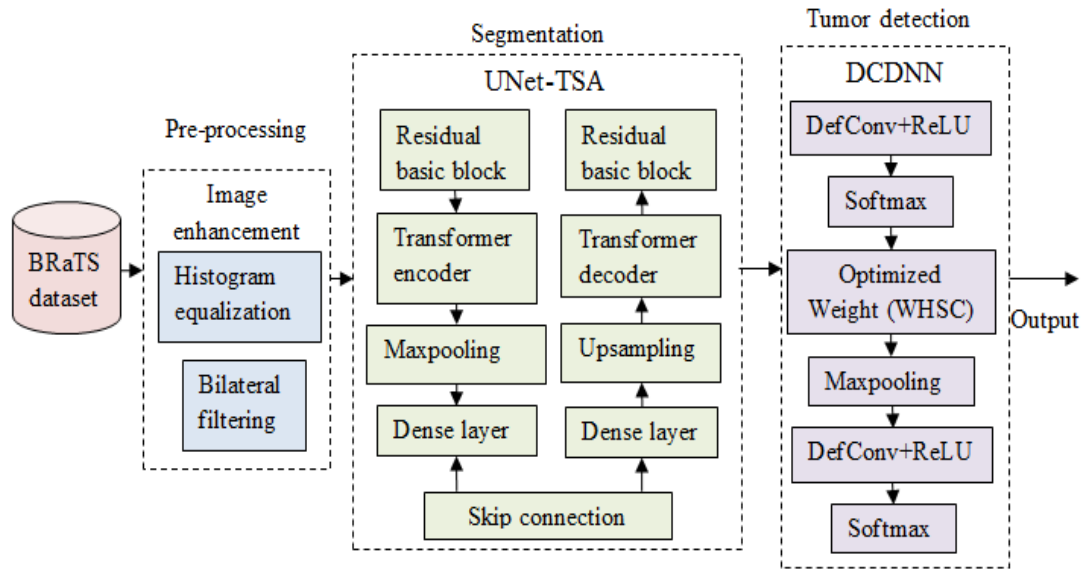
Table 1: Analysis of existing techniques with its disadvantages

Author name and Reference	Technique used	Performance	Disadvantage
Patil et al. [16]	SCNN-VGG16	Achieved 97.77% accuracy	High computational cost, less robustness and Interpretability challenges.
Ibrahim et al. [17]	CNN-PSO	Obtained accuracy rates of 98.50% precision, 98.83% sensitivity, and 97.12% specificity	Less scalability and generalizability challenges.
Archana et al. [18]	BKNN	Achieved 97.7% accuracy	Less efficiency due to very expensive large datasets to calculate the distance among each new point and each old point.
Kanchanamala et al. [19]	ExpDHO-based Deep CNN	Achieved 0.929 precision, 0.934 sensitivity, and 0.939 specificity.	The model is limited with lack of reliability.
Woźniak et al. [20]	CLM-CNN	Obtain 96% accuracy	Processing large dataset possess generalizability and interpretability challenges.

3. Proposed methodology

The brain tumor is one of the leading causes of death and it can be efficiently detected through segmentation process. The overall architecture of proposed approach is shown in figure 1. Initially, from the dataset images are collected and preprocessed with histogram equalization and bilateral filtering. The contrast of the image is enhanced with histogram equalization approach. It can be accomplished by distributing most frequent pixel values throughout the image. Also, the bilateral filter is applied to minimize the presence of noise from the image. This approach is based on replacing the intensity of pixel values with weighted average from neighboring pixel. Hence, it preserves edge information. After preprocessing, the image is fed to the segmentation process in which the tumor regions are separated from normal tissue region. The preprocessing is invoked with UNet-TSA and then the tumor is detected with DCDNN approach.

Figure 1: Overall architecture of proposed approach



3.1 UNet-TSA based Segmentation

The Unet architecture is integrated with transformer architecture through self-attention for improving the segmentation performance. UNet-TSA architecture is shown in figure 2. MHSA module of transformer encoder and decoder are shown in figure 3. The self-attention modules are applied in both encoder and decoder part for capturing the long range dependencies in varying scales with less overhead. The complexity of self-attention can be minimized with relative position encoding. The self-attention decoder extracts the fine grained information through skip connection in encoder. The proposed approach permits the transformer initialization in convolution network without pre-training. The convolution layer collects the details of neighboring pixel and captures the long range dependencies. The long range relationship on high dimensional feature map is attained with self-attention approach.

Figure 2: Structure of proposed UNet-TSA

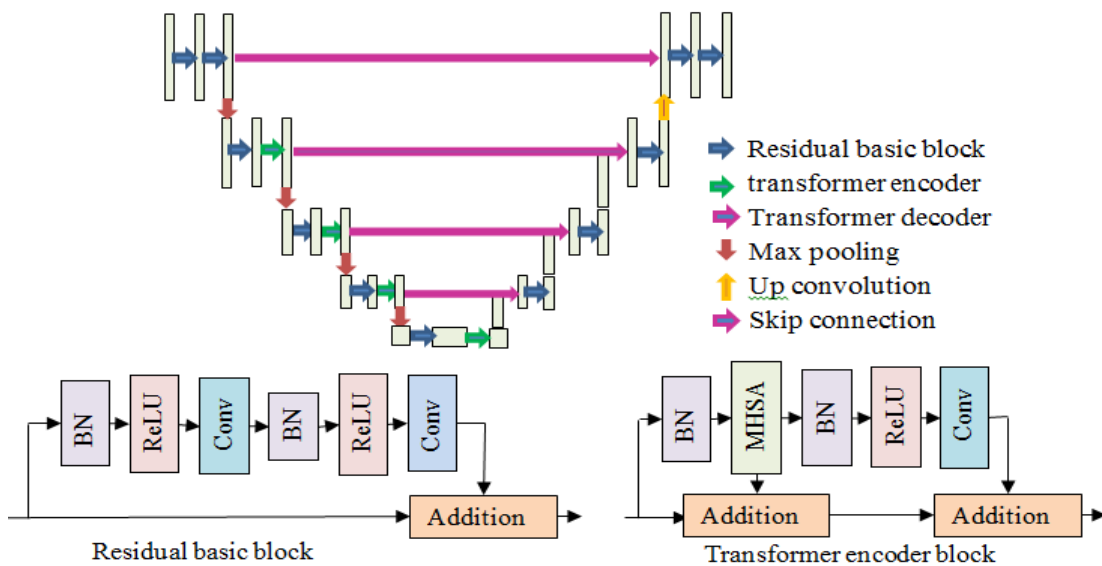
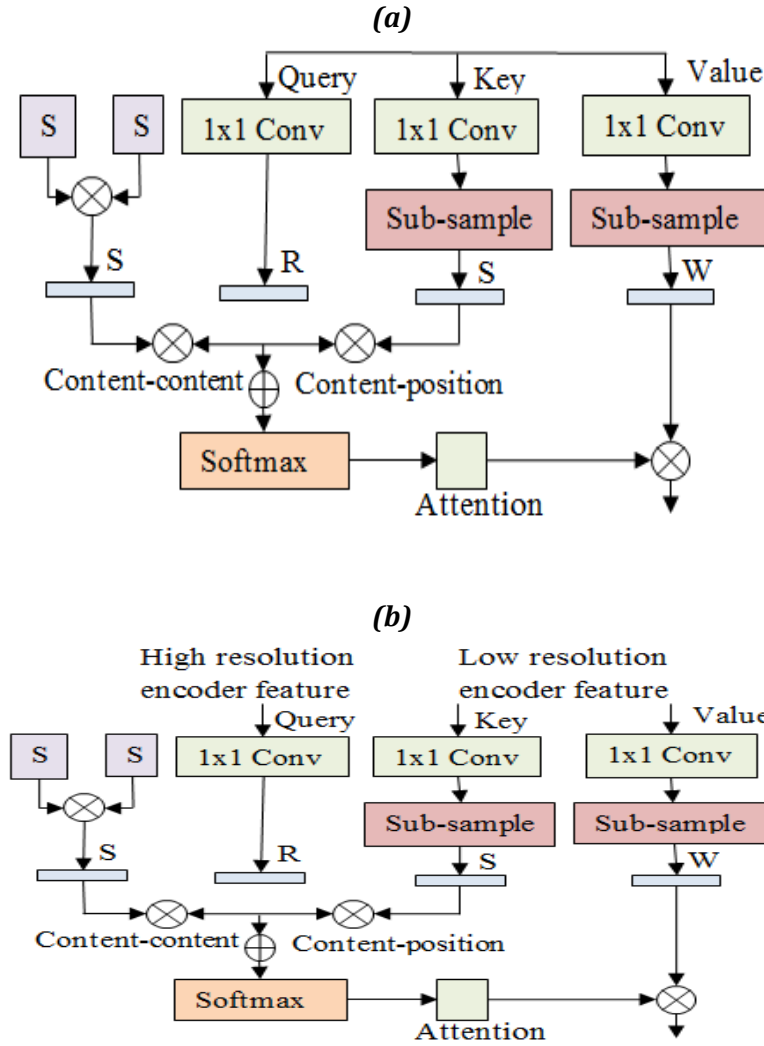


Figure 3: Proposed MHSA module (a) MHSA of transformer encoder (b) MHSA of transformer decoder



The transformer is made of multi-head self-attention which permits joint attention from various representation subspaces. The multi-head outputs are integrated and transformed to the feed forward neural network. Considering the input feature map $Y \in S^{D \times I \times X}$ where I represents the height, X is a width, and D indicates the amount of channels. The convolution layers are utilized for the projection of Y into query R , key L , and value W embedding such that $R, L, W \in S^{e \times I \times X}$. In each embedding process, the head dimension is represented with e , and the R, L, W values are flattened and transposed with the sequence of $o \times e$ where $o = IX$. The self-attention output can be formulated based on scaled dot product.

$$Atn(R, L, W) = sm \left(\frac{RL^T}{\sqrt{e}} \right) W \quad (1)$$

The projection matrix or similarity matrix is represented with $Q \in S^{o \times o}$. It calculates the normalized pairwise dot product between key elements. The contextual details from the values are estimated based on aggregating matrix to enable global receptive field.

The position information is eliminated with self-attention module. It makes the structured information ineffective for the segmentation process. Hence, relative position encoding is added based on width and height information. In between pixel j and pixel k , Pairwise attention is added with relative position encoding.

$$Pe_{j,k} = \frac{r_j^U}{\sqrt{e}} (l_k + s_{k_y-j_y}^X + s_{k_z-j_z}^I) \quad (2)$$

where r_j is the query vector for pixel j , $s_{k_y-j_y}^X, s_{k_z-j_z}^I$ are the learnable embedding of width $k_y - j_y$. Similarly, the width and height are calculated after low dimensional projection.

The relative position embedding based self-attention is given as,

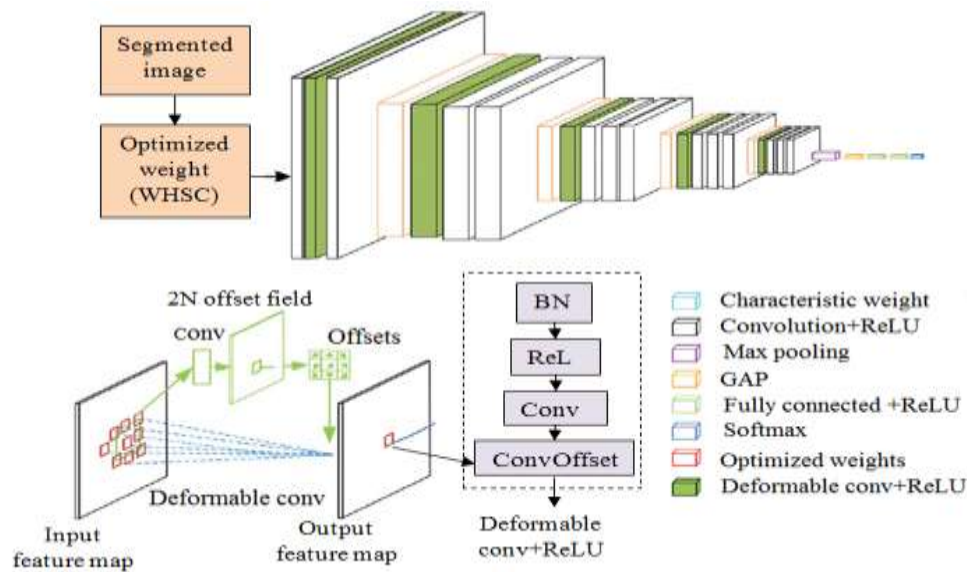
$$Attention(R, \bar{L}, \bar{W}) = sm \left(\frac{R\bar{L}^T + T_I^{re} + T_X^{re}}{\sqrt{e}} \right) \bar{W} \quad (3)$$

Where, $R\bar{L}^T, T_I^{re}$, and T_X^{re} are relative position metrics, and $\bar{W}$ represent the dimension.

3.2 Tumor detection with DCDNN

DCDNN approach improves the robustness, and generalization performance of tumor detection process using deformable convolution layers. The performance of DCDNN is further enhanced with WHSC based hyper parameter optimization algorithm. The DCDNN architecture is shown in figure 4. The deformable convolution layer has four layers including DC kernel, ReLU, convolution kernel, and batch normalization. The convolution and DC kernels are used for accurate morphological feature estimation. Feature mapping of DCDNN is passed through ReLU activation function. The batch normalization layer enhances the training speed while utilizing more number of parameters. Max pooling layers are used to select the most discriminative feature information by minimizing feature dimension. Then the average pooling layers are utilized for average mapping and preserve spatial details of the image. It minimizes the overfitting probability. The fully connected layers are formulated by computing weighted sum of various features with the interconnection of subsequent layers. Finally, softmax layer is considered for output probability estimation and it produces final detection outcome.

Figure 4: architecture of DCDNN



Consider the input feature map y and the convolution network H and the sampling value x , the normal convolution grid H can be formulated as:

$$H = \{(-1,-1), (-1,0), \dots, (0,1), (1,1)\} \quad (4)$$

Hence, the location n_0 from the output feature map z is given as:

$$z(n_0) = \sum_{n_j \in H} x(n_j) \cdot y(n_0 + n_j) \quad (5)$$

where y is the input feature map, x represents the sampled value weight and location mean of H . The normal grid of deformable convolution is improved with the offset Δn_j .

$$z(n_0) = \sum_{n_j \in H} x(n_j) \cdot y(n_0 + n_j + \Delta n_j) \quad (6)$$

where Δn_j is obtained from the additional convolution layer and was computed with shift values of sample data.

3.2.1 Wild Horse soft computing (WHSC) approach

WHSC approach resolves slow convergence, and local optimal problem in complex high dimensional space. The performance of wild horse optimization is enhanced with soft computing based initialization. It depends on the social behavior of wild horses which perform mate, pursue, command, graze, and dominate. The initial population is split into several groups where O indicates the number of population and H indicates the number of groups. All groups are associated with a leader called stallion and hence, the total population have H stallion. Similarly, the residual population of mare and foal were distributed among these groups. The grazing performance is described as follows.

$$Y_{j,H}^k = 2A \cos(2\pi SA) \times (Sn^k - Y_{j,H}^k) + Sn^k \quad (7)$$

where $Y_{j,H}^k$ represents the current position of foal/mare group members, A indicates the adaptive process, S is the stochastic number ranges from -2 to 2, and Sn^k represents the location of stallion. The adaptive process A is computed as below.

$$Q = \vec{S}_1 < T_{DR}; I_{DX} = (Q = 0); A = S_2 \Theta I_{DX} + \vec{S}_3 \Theta(\sim I_{DX}) \quad (8)$$

Where Q indicates the vector from 0 to 1, $\vec{S}_1, \vec{S}_3$ indicates the arbitrary value ranges from 0 to 1, S_2 refers the uniform arbitrary value from 0 to 1, T_{DR} is the adaptive parameter starts with 1 and it minimizes until reaches 0, and I_{DX} is the identity parameter. The execution process of T_{DR} is calculate as follows.

$$T_{DR} = 1 - It \times \left(\frac{1}{\max It} \right) \quad (9)$$

Where It indicates the previous iteration and $\max It$ represents the maximum number of iterations. For enabling mate behavior of horse the foal drives to temporary group. The cross over function or mate performance is given as

$$Y_{H,L}^q = Co(Y_{H,j}^r, Y_{H,k}^A) j \neq k \neq l, q = r = finish, Co = Mean \quad (10)$$

Where $Y_{H,L}^q$ represents the position of horse q at L th group, $Y_{H,j}^r$ is the position of r^{th} horse at group j , $Y_{H,k}^A$ represents the horse A at position k , and Co indicates the crossover process. The stallion used to lead the group to water hole. The water hole is competed by the group and dominates group use that. Then, the next group gets chance to utilize the water hole. This procedure is represented as follows.

$$\bar{S}n_{H_j} = \begin{cases} 2A \cos(2\pi SA) \times (W_H - Sn_{H_j}) + W_H, & \text{if } S_3 > 0.5 \\ 2A \cos(2\pi SA) \times (W_{VH} - Sn_{H_j}) - W_H, & \text{if } S_3 \leq 0.5 \end{cases} \quad (11)$$

where $\bar{S}n_{H_j}$ indicates the next position of leader, W_H represents the water hole position, and S_3 is a random value. In this strategy, the leader stallion is chosen based on fitness value. The position of leaders and corresponding members were updated based on the following formula.

$$\bar{S}n_{H_j} = \begin{cases} Y_{H,j}, & \text{if } \cos u(Y_{H,j}) < \cos u(Sn_{H_j}) \\ Sn_{H_j}, & \text{if } \cos u(Y_{H,j}) > \cos u(Sn_{H_j}) \end{cases} \quad (12)$$

Where $Y_{H,j}$ is a foal position, Sn_{H_j} is a stallion position, $u(Y_{H,j})$ is the fitness of foal and $u(Sn_{H_j})$ is the fitness of stallion. For enhancing the performance, the initialization process is accomplished based on following levy flight distribution.

$$Y_{j,H} = Y_{j,H} - \alpha(Y_{j,H} - Y_h) \otimes ly(\mu) = Y_{j,H} + \frac{0.01v}{|w|^{1/\mu}} (Y_{j,H} - Y_h) \quad (13)$$

where $Y_{j,H}$ indicates the j^{th} position of group member, γ is the step scaling size, Y_h is the global optimum solution, $\otimes$ indicates element-wise multiplication, $y(\mu)$ represents levy flight exponent, and the values of v and w are estimated as,

$$v \sim O(0, \sigma_v^2), w \sim O(0, \sigma_w^2) \quad (14)$$

Where O represents the dimension, and σ is the standard deviation. In this way, the new candidate solution is generated and processed. It improves the capacity for balancing exploration and exploitation strategies. The proposed optimization approach improves the stability and flexibility during exploration and exploitation strategies of optimization algorithm.

4. Experimental analysis

The proposed approach is simulated with BRaTS 2018 , Figshare Dataset, Brain Tumor Classification MRI Dataset Task 1 dataset in python software. Performance of proposed approach is evaluated through standard performance metrics and compared with existing approaches. The hardware and software requirement of the proposed work is declared in table 2.

Table 2: Hardware and Software requirement

Hardware requirement		
Sl. No	Hardware type	Specification
1	Processor	Intel(R) Core(TM)2 Duo CPU E8400 @ 3.00GHz 2.99 GHz processor
2	Storage	195 GB
3	RAM	8.00 GB
4	Display	SVGA monitor resolution 1024x768
Software requirement		

Sl. No	Software type	Specification
1	Operating System	Windows 10 Pro
2	Implementation software	PYTHON

4. 1 Experimental results and discussions

Brain tumors are pose a critical health challenge, demanding accurates the diagnosis and prognosis for effective treatment.Evaluation on multiple datasets shows strong performance in segmentation, classification (accuracy, precision, recall, F1-score, AUC) and survival analysis, confirming its reliability and clinical relevance. The proposed work was implemented on a system running Windows 11 Pro with Python 3.12.7 as the primary programming language. Development was carried out using Jupyter Notebook and Visual Studio Code, supported by libraries such as NumPy, Pandas, Matplotlib, and Seaborn. For DL, Tensor Flow and Keras frameworks were employed to design and train the models. The system was equipped with 32 GB RAM, ensuring efficient data processing and model execution.

Table 3: Sample output of BraTS 2018 dataset

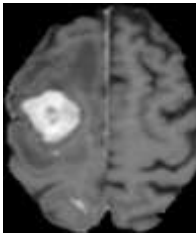
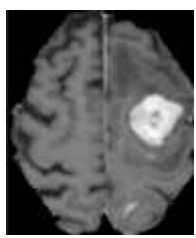
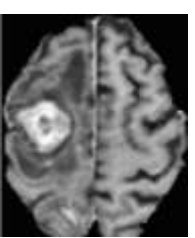
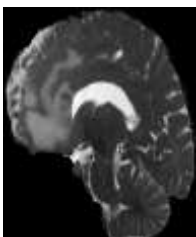
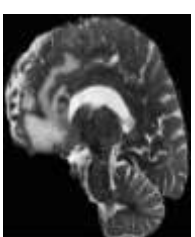
Original image	Input Image	Preprocessed Image	Feature Extracted Image	Segmented image
				
				

Table 4: Sample output of Figshare Dataset

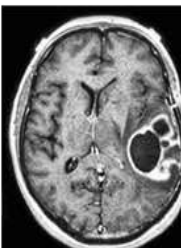
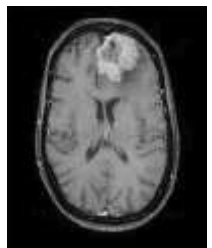
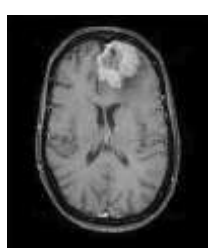
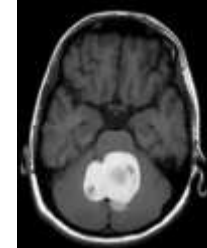
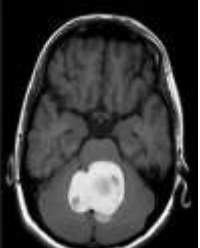
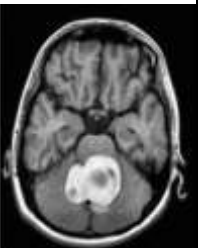
Original image	Input Image	Preprocessed Image	Segmented image	Feature Extracted Image
				
				

Table 5: Sample output of Brain Tumor Classification MRI Dataset

Original image	Input Image	Preprocessed Image	Segmented image	Feature Extracted Image
				
				

4.2 Dataset Description

Brats2018 Dataset: It is widely used in medical imaging research and provides multimodal 3D brain MRIs with expert-annotated tumor segmentations. Each case includes four MRI modalities (T1, T1c, T2, and FLAIR) and annotations of three tumor sub-regions combined into whole tumor, tumor core, and enhancing tumor regions. Data were collected from 19 institutions using diverse MRI scanners, supporting robust analysis of brain tumor detection and segmentation. Table 2 shows the sample output of BraTS2018 dataset.

Figshare Dataset: The Figshare brain tumor dataset consists of 3,064 T1-weighted contrast enhanced MRI images from 233 patients, covering three tumor types: meningioma (708 slices), glioma (1,426 slices), and pituitary tumor (930 slices). To manage the file size, the dataset is divided into four subsets, each containing 766 slices in separate. Table 3 shows the sample output of Figshare dataset.

Brain tumor classification MRI dataset : Brain tumors are aggressive diseases that can afflict both adults and children. It make up 85–90% of primary CNS tumors, and each year about 11,700 new cases are discovered. The 5-year survival rate for brain tumors is 34% for men and 36% for women. The classification of these tumors into benign, malignant, and pituitary kinds necessitates precise diagnosis and treatment in order to enhance patient outcomes. Table 4 shows the sample output of Brain tumor classification MRI dataset.

Figure 5: Confusion Matrix (a) BraTS2018 dataset, (b) Figshare dataset, and (c) Brain tumor classification MRI dataset

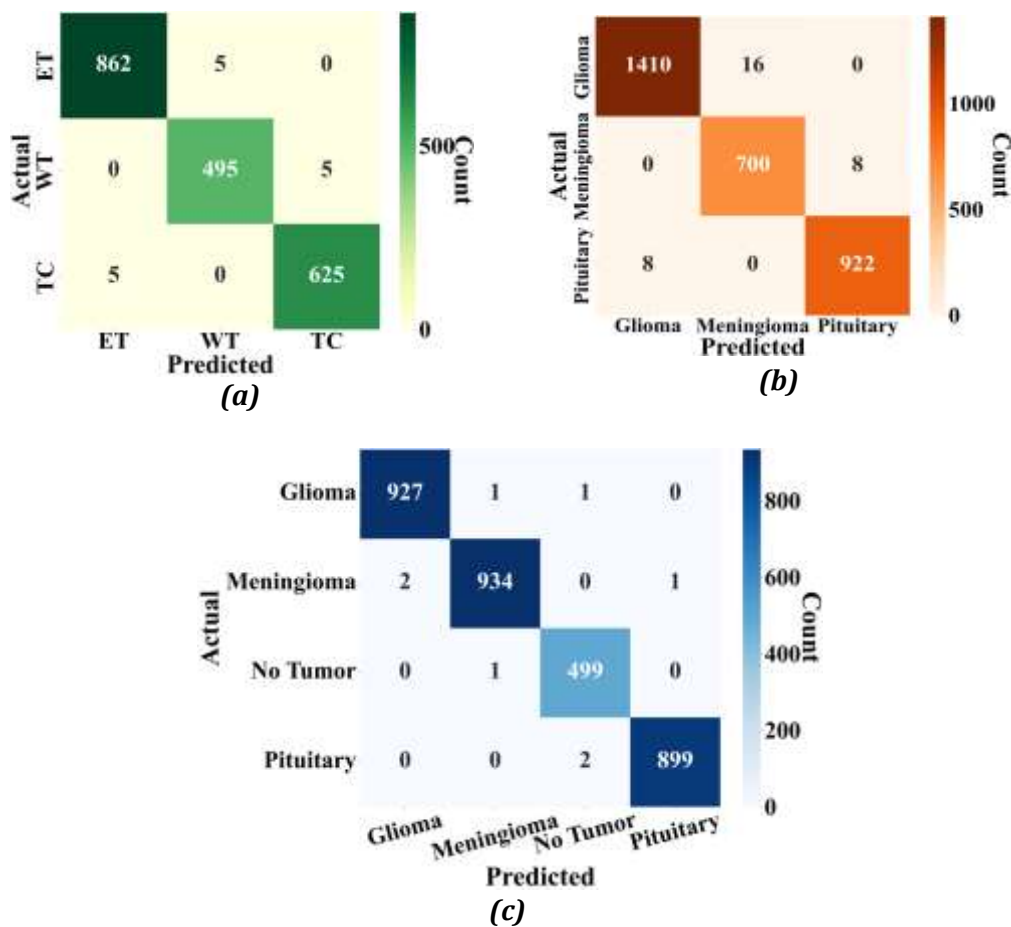


Figure 5 presents the confusion matrices for brain tumor classification. In Figure 8(a), the diagonal values represent correctly classified cases, while the off-diagonal values indicate misclassifications. Figure 8(b) shows strong diagonal dominance, suggesting that most predictions align with the actual labels. Figure 8(c) demonstrates that the models achieve high accuracy in distinguishing tumor subtypes (ET, WT, TC), tumor categories (Glioma, Meningioma, Pituitary), and tumor versus no-tumor cases.

Figure 6: Label distribution (a) BraTS2018 dataset, (b) Figshare dataset, and (c) Brain tumor classification MRI dataset

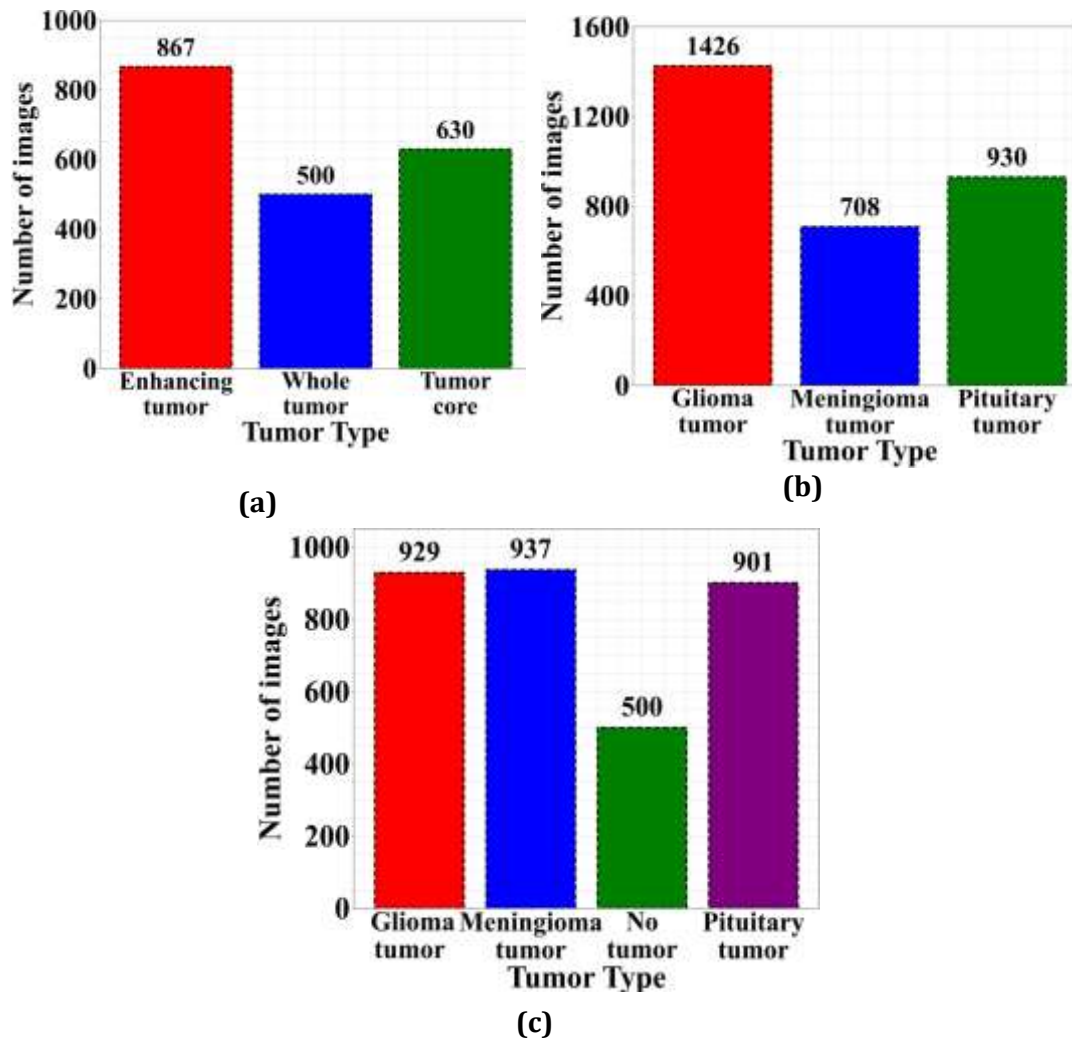


Figure 6 illustrates the label distribution of tumor images across different categories. In Figure 9(a), enhancing tumors (867) occur more frequently compared to whole tumors (500) and tumor core (630). Figures 9(b) and 9(c) show the class distribution of the datasets, with glioma tumors having the highest count (1426 and 929, respectively), followed by meningioma and pituitary tumors, and a smaller number of no-tumor cases (500).

Figure 7: ROC Curve (a) BraTS2018 dataset, (b) Figshare dataset, (c) Brain tumor classification MRI dataset

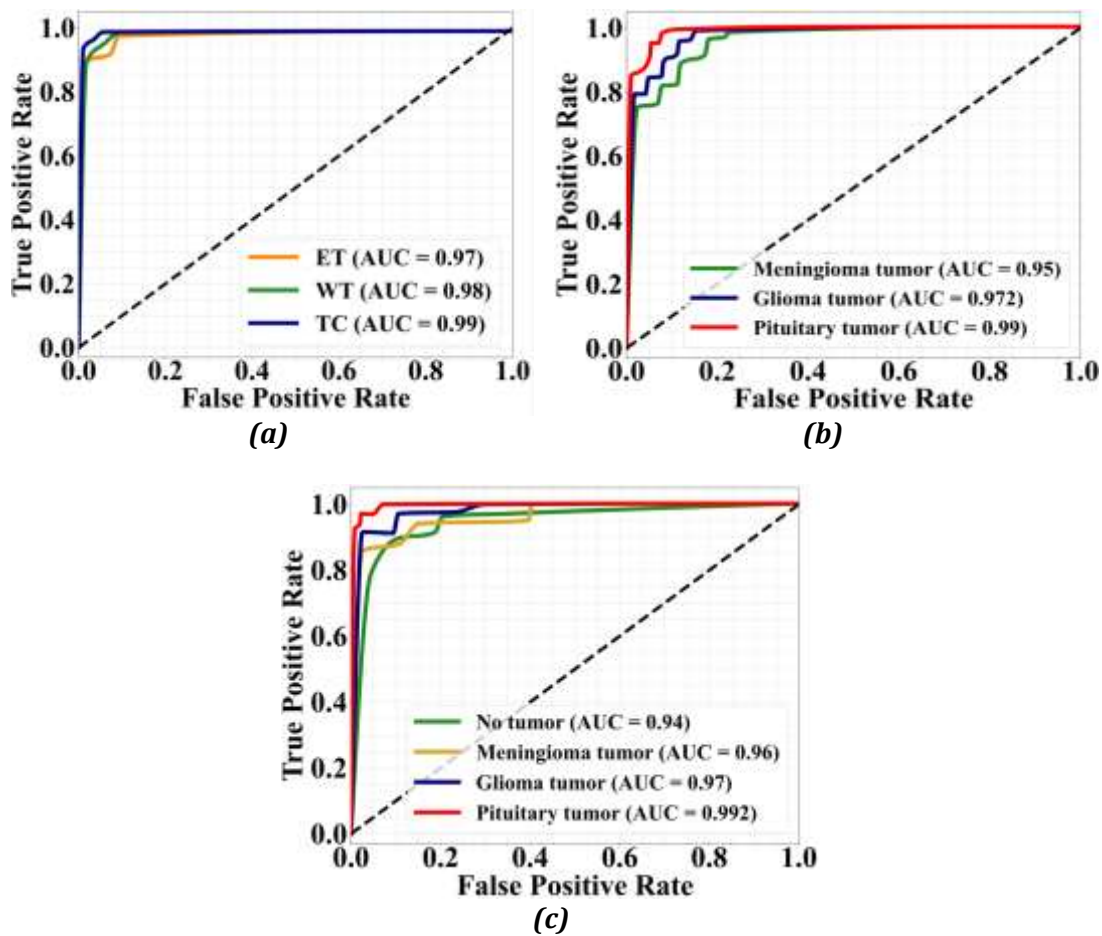


Figure 7 shows the Receiver Operating Characteristic (ROC) curves for different tumor classification models, illustrating the diagnostic performance. Figure 10(a) depicts the trade-off between the true positive rate and false positive rate for a specific tumor type. Figure 10(b) presents the Area under the Curve (AUC) values, which reflect model accuracy, with higher values indicating better performance. Figure 10(c) highlights that certain tumor types, such as TC (AUC = 0.99) and Pituitary tumor (AUC = 0.992), achieve the highest classification accuracy. Overall, the models demonstrate strong predictive capabilities, with AUC values ranging from 0.94 to 0.992.

Figure 8: Accuracy & Loss

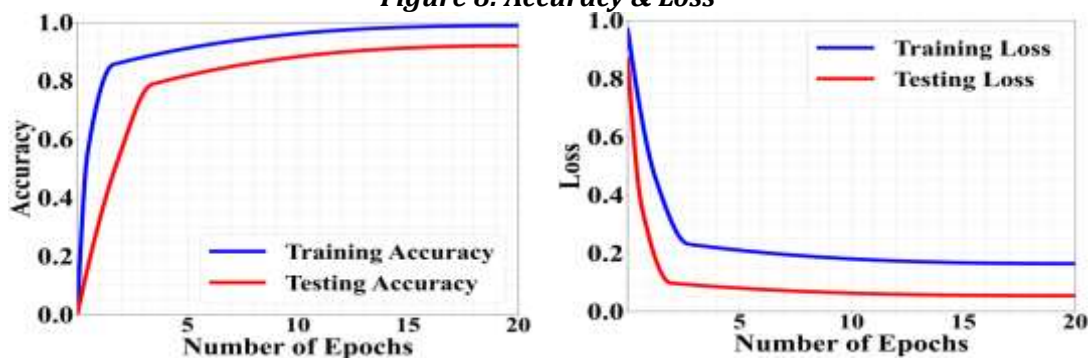


Figure 8 shows the training and testing performance of a model over 20 epochs. Accuracy steadily increases for both the training and testing, with training accuracy slightly higher, indicating good learning without major overfitting. Meanwhile, loss rapidly decreases and stabilizes for both sets, confirming effective model convergence.

The accuracy, precision, recall, f1-score, and specificity performance of the proposed approach is shown in figure 6. Performance of the proposed approach is compared with existing DL approaches including U-Net, GoogleNet, DenseNet201, ResNet50, and VGGNet. When compared with existing approaches, better performance was provided with the proposed approach. For all classification related metrics, the performance is above 0.98 that is 0.983 accuracy, 0.980 precision, 0.981 recall, 0.980 f1-score, and 0.986 specificity. Higher performance of the proposed approach indicates the accuracy and efficiency of classification process. It is due to high level contextual feature representation and accurate tumor region segmentation.

Figure 9: Comparison of classification performance

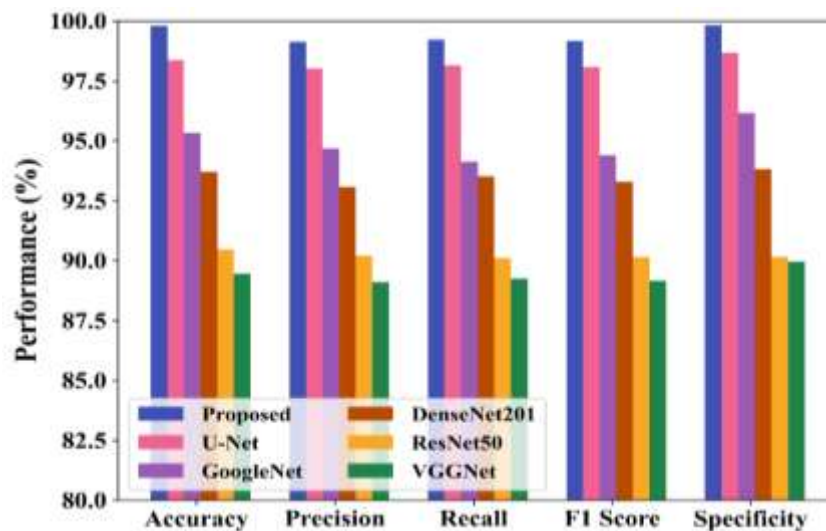
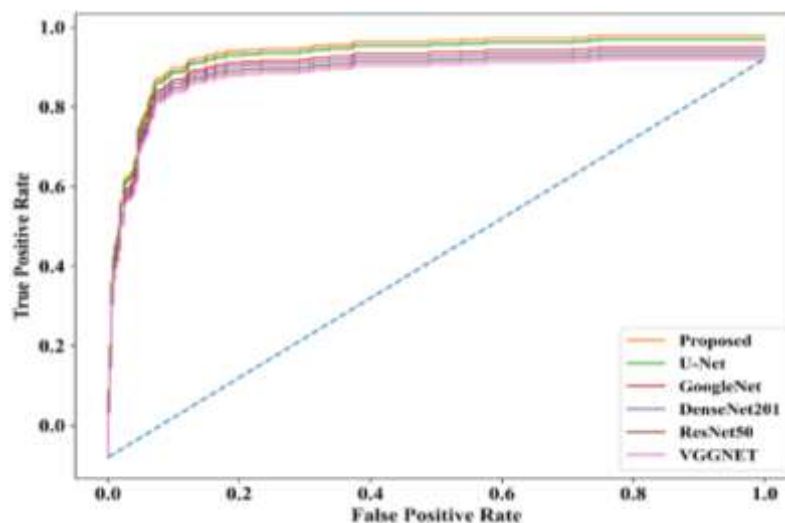


Figure 10: Performance comparison of AUC curve



The AUC performance comparison is shown in figure 10. The AUC curve is plotted based on true positive and false positive rates of the DL classification approaches. AUC values obtained are higher than existing AUC values. This indicates the higher level of true positives and lower level of false positives. Better AUC performance is due to correct classification of positive cases and lower number of incorrect classifications. The accuracy performance with varying number of epochs is shown in figure 8. When comparing with testing accuracy, training accuracy is higher for proposed approach. Both training and testing accuracy is lower for the existing approaches. Initially, the accuracy is lower than 0.6 and it is increased to 0.98 at epoch 50. After 50 epochs, the performance is not much deviated and nearly constant level is achieved for both training and testing procedure. The accuracy of DL approach is affected with the optimization procedure. Hence, the accuracy of proposed approach is improved with WHSC approach.

Figure 11: Accuracy performance comparison with varying number of epochs (a) Training accuracy (b) Testing accuracy

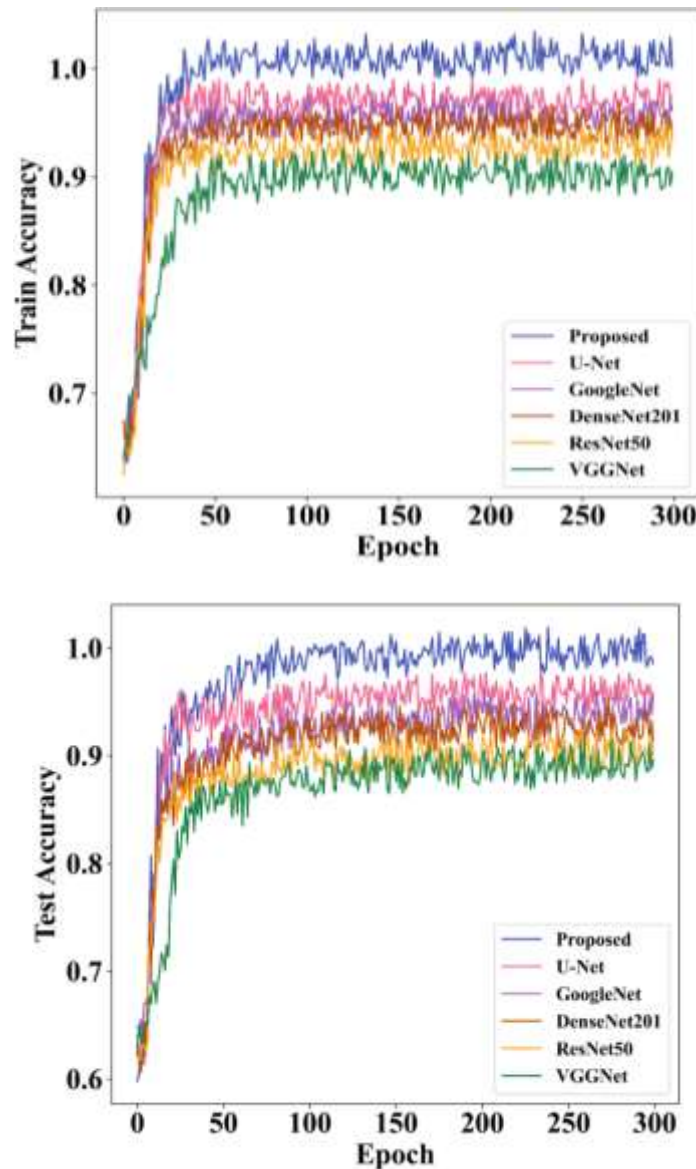
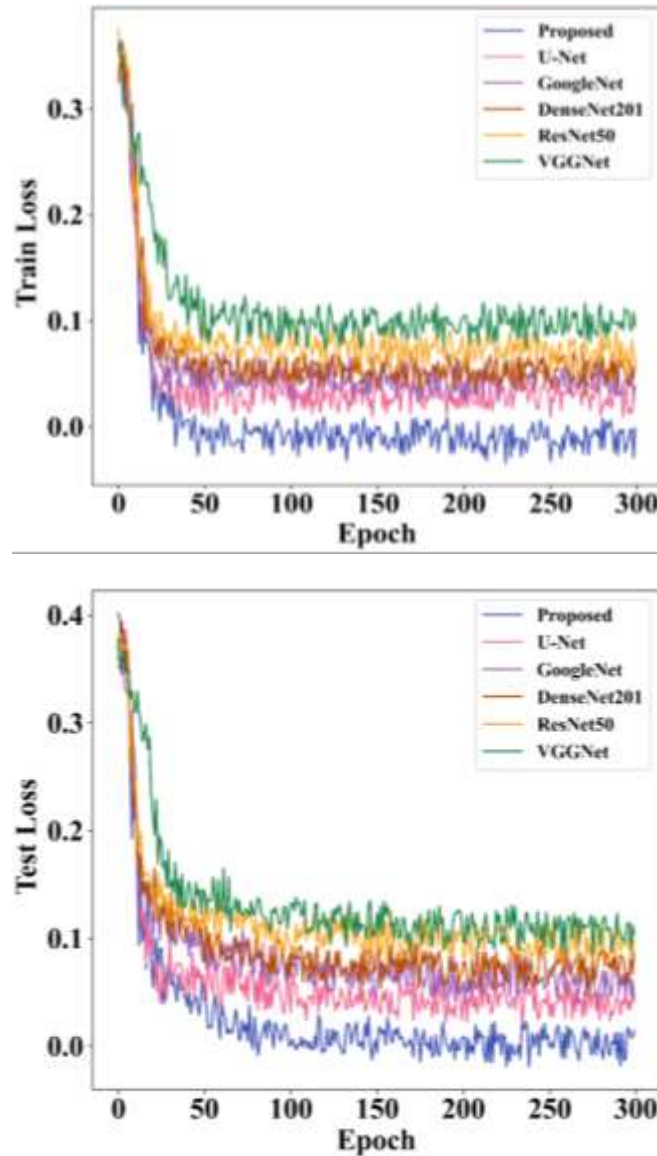


Figure 12: Loss performance comparison (a) Training loss (b) Testing loss

The loss performance comparison is given in figure 12. The training and testing loss of the proposed approach is compared with existing DL approaches. The training loss is lower than the testing loss for all approaches. When increasing the number of epochs, the loss is reduced. During the initial epochs, the loss is low. When the number of epochs is increased, the loss is reduced. The dice score, and IoU performance of proposed approach is shown in figure 10. The dice score performance is higher for the proposed approach and lower for the existing approaches. Dice score is 0.85 for the proposed approach. In case of existing approaches, the dice score performance is lower than 0.80. Higher dice score indicates the high similarity between predicted and ground truth results. Dice score is enhanced with better learning procedure and it indicates the segmentation accuracy. The IoU performance 0.996 for the proposed approach and lower than 0.98 for existing approaches. Higher IoU of proposed approach indicates the much-overlapped region between ground truth image and predicted image.

Figure 13: Performance comparison with existing approaches (a) Dice score (b) IoU

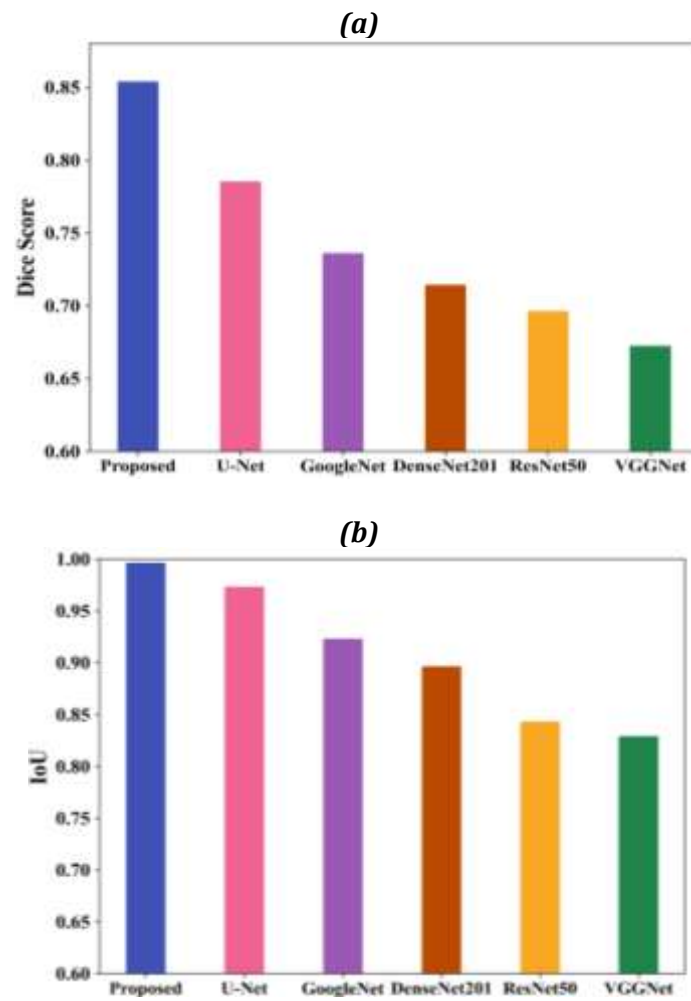
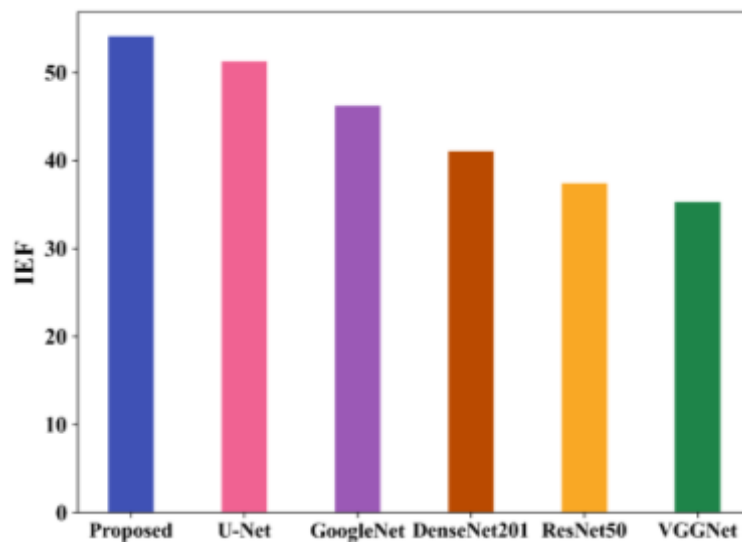
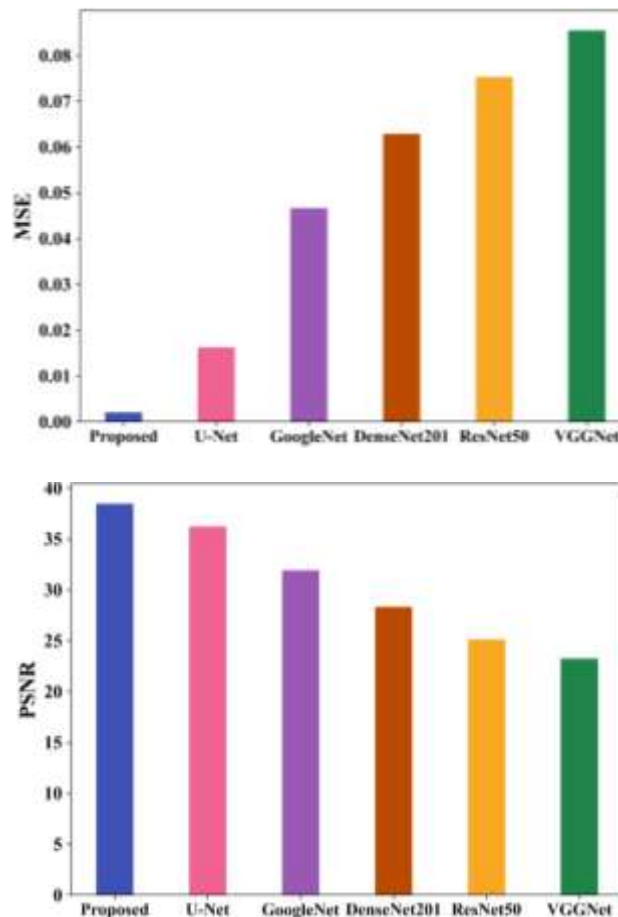


Figure 14: IEF Comparison with existing approaches



The IEF performance is compared with the existing approaches as shown in figure 14. Higher IEF in proposed approach indicates the better image quality for the task of tumor prediction. IEF performance is higher for the proposed approach and lower for the existing approaches. The IEF value obtained with U-Net, GoogleNet, DenseNet201, ResNet50, VGGNet, and proposed approach is 51.27, 42.20, 41.08, 37.41, 35.27, and 54.12. The proposed preprocessing approach using histogram equalization and bilateral filtering enhances the image quality for better segmentation and prediction.

Figure 15: Comparison of proposed segmentation quality (a) MSE (b) PSNR



The MSE, and PSNR performance comparison is given in figure 15. The MSE of the proposed approach is lower than the existing approach. Prediction error of the proposed approach is lower than the prediction error of other approaches. Lower error of the proposed approach is due to generalization, low convergence, and robustness. MSE is 0.0020 for proposed approach and higher than 0.01 for all other existing approaches. PSNR performance of proposed approach is achieved as 38.48. The higher PSNR indicates the image quality which is higher during tumor detection process. High PSNR based images promotes accurate detection process. For the existing approaches, the PSNR values are lower than 35. The lower values indicate the inefficient processing. The comparison of results with existing approaches is given in table 3. When compared with all performance metrics, the proposed approach has achieved the performance at high level.

Table 6: Results comparison with existing approaches

Approach es	U- Net	Google Net	DenseNet2 01	ResNet 50	VGG Net	Propos ed
Accuracy	0.983	0.953	0.937	0.904	0.895	0.998
Precision	0.980	0.946	0.930	0.902	0.890	0.991
Recall	0.981	0.941	0.935	0.901	0.892	0.992
F1-Score	0.980	0.944	0.933	0.901	0.891	0.991
Specificity	0.986	0.961	0.938	0.901	0.899	0.998
MSE	0.016	0.046	0.062	0.095	0.105	0.0020
Dice score	0.785	0.736	0.714	0.696	0.672	0.854
PSNR	36.23	31.93	28.34	25.13	23.24	38.48
IoU	0.973	0.923	0.896	0.843	0.828	0.996
IEF	51.27	42.20	41.08	37.41	35.27	54.12

5. Conclusion

In this paper, DCWDNN approach is developed for brain tumor segmentation. The long range dependencies from the tumor regions were captured with UNet-TSA based segmentation approach. Segmentation performance was enhanced with fine grained feature information. The complexity of the proposed approach is minimized with self-attention module. Robustness and better generalization was achieved with the proposed tumor detection process. The classification accuracy was improved with WHSC approach. The performance of the proposed approach is evaluated with BRaTS 2021 dataset in which the accuracy, precision, recall, f1-score, specificity, MSE, dice score, PSNR, IoU, and IEF performance is enhanced with 0.998, 0.991, 0.992, 0.991, 0.998, 0.0020, 0.854, 38.48, 0.996, and 54.12 respectively. When compared with existing approaches, the performance of the proposed approach is high and it indicates the superiority of the proposed approach. In future, the performance of the proposed approach can be further extended with hybrid approaches by evaluating large amount of data to ensure its effectiveness.

References

1. Sharif, M.I., Khan, M.A., Alhussein, M., Aurangzeb, K. and Raza, M., 2022. A decision support system for multimodal brain tumor classification using deep learning. *Complex & Intelligent Systems*, 8(4), pp.3007-3020.
2. Usha, M.P., Kannan, G. and Ramamoorthy, M., 2024. Multimodal brain tumor classification using convolutional Tumnet architecture. *Behavioural Neurology*, 2024(1), p.4678554.
3. Al Khalil, Y., Ayaz, A., Lorenz, C., Weese, J., Pluim, J. and Breeuwer, M., 2024. Multi-modal brain tumor segmentation via conditional synthesis with Fourier domain adaptation. *Computerized Medical Imaging and Graphics*, 112, p.102332.
4. Khalighi, S., Reddy, K., Midya, A., Pandav, K.B., Madabhushi, A. and Abedalthagafi, M., 2024. Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. *NPJ precision oncology*, 8(1), p.80.
5. Nehru, V. and Prabhu, V., 2024. Automated multimodal brain tumor segmentation and localization in MRI images using hybrid res2-unext. *Journal of Electrical Engineering & Technology*, 19(5), pp.3485-3497.

6. Yu, Z., Li, X., Li, J., Chen, W., Tang, Z. and Geng, D., 2024. HSA-net with a novel CAD pipeline boosts both clinical brain tumor MR image classification and segmentation. *Computers in Biology and Medicine*, 170, p.108039.
7. Umarani, C.M., Gollagi, S.G., Allagi, S., Sambrekar, K. and Ankali, S.B., 2024. Advancements in deep learning techniques for brain tumor segmentation: A survey. *Informatics in Medicine Unlocked*, 50, p.101576.
8. Sun, K., Zhang, Y., Liu, J., Yu, L., Zhou, Y., Xie, F., Guo, Q., Zhang, H., Wang, Q. and Shen, D., 2024. Achieving multi-modal brain disease diagnosis performance using only single-modal images through generative AI. *Communications Engineering*, 3(1), p.96.
9. Gong, Z., Xu, T., Peng, N., Cheng, X., Niu, C., Wiestler, B., Hong, F. and Li, H.B., 2024. A multi-center, multi-parametric MRI dataset of primary and secondary brain tumors. *Scientific Data*, 11(1), p.789.
10. Yan, Y., Yang, C., Chen, W., Jia, Z., Zhou, H., Di, Z. and Xu, L., 2025. Multimodal MRI and artificial intelligence: shaping the future of glioma. *Journal of Neurorestoratology*, 13(2), p.100175.
11. Rutoh, E.K., Guang, Q.Z., Bahadar, N., Raza, R. and Hanif, M.S., 2024. GAIR-U-Net: 3D guided attention inception residual u-net for brain tumor segmentation using multimodal MRI images. *Journal of King Saud University-Computer and Information Sciences*, 36(6), p.102086.
12. Abidin, Z.U., Naqvi, R.A., Haider, A., Kim, H.S., Jeong, D. and Lee, S.W., 2024. Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: A prospective survey. *Frontiers in Bioengineering and Biotechnology*, 12, p.1392807.
13. Wang, J., Lu, S.Y., Wang, S.H. and Zhang, Y.D., 2024. RanMerFormer: Randomized vision transformer with token merging for brain tumor classification. *Neurocomputing*, 573, p.127216.
14. Yurtsever, M.E., Atay, Y., Arslan, B. and Sagiroglu, S., 2024. Development of brain tumor radiogenomic classification using GAN-based augmentation of MRI slices in the newly released gazi brains dataset. *BMC Medical Informatics and Decision Making*, 24(1), p.285.
15. Ahmed, M.M., Hossain, M.M., Islam, M.R., Ali, M.S., Nafi, A.A.N., Ahmed, M.F., Ahmed, K.M., Miah, M.S., Rahman, M.M., Niu, M. and Islam, M.K., 2024. Brain tumor detection and classification in MRI using hybrid ViT and GRU model with explainable AI in Southern Bangladesh. *Scientific reports*, 14(1), p.22797.
16. Kar, S. and Singh, P.K., 2025. MBTC-Net: Multimodal brain tumor classification from CT and MRI scans using deep neural network with multi-head attention mechanism. *Medicine in Novel Technology and Devices*, p.100382.
17. Ketabi, S., Wagner, M.W., Hawkins, C., Tabori, U., Ertl-Wagner, B.B. and Khalvati, F., 2025. Multimodal contrastive learning for enhanced explainability in pediatric brain tumor molecular diagnosis. *Scientific Reports*, 15(1), p.10943.
18. Liu, F., Zhang, Y., Lu, T., Wang, J. and Wang, L., 2025. Hierarchical in-out fusion for incomplete multimodal brain tumor segmentation. *Scientific Reports*, 15(1), p.23017.
19. Wang, Y., Xu, J., Guan, Y., Ahmad, F., Mahmood, T. and Rehman, A., 2025. MSegNet: A multi-view coupled cross-modal attention model for enhanced MRI brain tumor segmentation. *International Journal of Computational Intelligence Systems*, 18(1), p.63.
20. Zarenia, E., Far, A.A. and Rezaee, K., 2025. Automated multi-class MRI brain tumor classification and segmentation using deformable attention and saliency mapping. *Scientific Reports*, 15(1), p.8114.
21. Chandrasekaran, S., Aarathi, S., Alqhatani, A., Khan, S.B., Quasim, M.T. and Basheer, S., 2025. Improving healthcare sustainability using advanced brain simulations using a

- multi-modal deep learning strategy with VGG19 and bidirectional LSTM. *Frontiers in Medicine*, 12, p.1574428.
22. Batool, A. and Byun, Y.C., 2025. A lightweight multi-path convolutional neural network architecture using optimal features selection for multiclass classification of brain tumor using magnetic resonance images. *Results in Engineering*, 25, p.104327.
 23. Dorfner, F.J., Patel, J.B., Kalpathy-Cramer, J., Gerstner, E.R. and Bridge, C.P., 2025. A review of deep learning for brain tumor analysis in MRI. *NPJ Precision Oncology*, 9(1), p.2.
 24. <https://ioft-data.engin.umich.edu> › brats-2018-dataset
 25. <https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset>
 26. <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>
 27. Liu, X. and Nguyen, T.D., 2024. Medical images enhancement by integrating CLAHE with wavelet transform and non-local means denoising. *Academic Journal of Computing & Information Science*, 7(1), pp.52-58.
 28. Kanchanamala, P., Kuppusamy, V. and Ganesan, G., 2025. QDCNN-DMN: A hybrid deep learning approach for brain tumor classification using MRI images. *Biomedical Signal Processing and Control*, 101, p.107199.
 29. Memari, M., Shekaramiz, M., Masoum, M.A. and Seibi, A.C., 2024. Data fusion and ensemble learning for advanced anomaly detection using multi-spectral RGB and thermal imaging of small wind turbine blades. *Energies*, 17(3), p.673.
 30. Katzman, J.L., Shaham, U., Cloninger, A., Bates, J., Jiang, T. and Kluger, Y., 2018. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology*, 18(1), p.24.
 31. Kumar, V.V.N.S., Reddy, G.H. and GiriPrasad, M.N., 2023. A novel glaucoma detection model using Unet++-based segmentation and ResNet with GRU-based optimized deep learning. *Biomedical Signal Processing and Control*, 86, p.105069.
 32. Van Zyl, C., Ye, X. and Naidoo, R., 2024. Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP. *Applied Energy*, 353, p.122079.