International Journal of Engineering Research and Science & Technology

www.ijerst.com
Vol. 22, No. 1(1), 2026

Segmentation and Classification of Brain Tumor Using 3D-UNet Deep Neural Network

G. Ravi kumar[1], Reddy Renuka[2], Vempati Rajkumar[3], Pillagulla Vasantha kumari[4], Kota Venkat Ramana[5]

Assistant Professor[1], Student[2,3,4,5]

Department of Computer Science Engineering[1,2,3,4,5]

Chaitanya Engineering college, Visakhapatnam, Andhra Pradesh, India

{ kgravi0417@gmail.com[1] , reddyrenuka765@gmail.com[2] , rajeshvempati18@gmail.com[3] vasanthapillagulla@gmail.com [4] , Venkataramana.kota2022@gmail.com [5]} @cec.ac.in

Abstract

Brain tumor segmentation and classification are critical steps in neuro-oncology for treatment planning and prognosis assessment. This paper proposes a 3D-UNet deep neural network framework for simultaneous brain tumor segmentation and multi-class classification from multimodal MRI scans. The proposed architecture processes T1, T1ce, T2, and FLAIR MRI sequences as multi-channel 3D volumes, leveraging residual connections and channel attention modules to improve feature discrimination between tumor sub-regions. The system segments tumors into whole tumor, tumor core, and enhancing tumor regions and simultaneously classifies tumor grade (LGG vs. HGG) through a shared encoder with task-specific decoders. The model achieves Dice scores of 0.89 (whole tumor), 0.82 (tumor core), and 0.76 (enhancing tumor) on BraTS 2021, with tumor grade classification accuracy of 91.4%, outperforming standard 3D-UNet and V-Net baselines.

I. INTRODUCTION

Brain tumors are among the most life-threatening forms of cancer, with gliomas being the most common malignant type in adults. Accurate delineation of tumor boundaries and sub-regions is essential for surgical planning, radiation therapy targeting, and treatment response monitoring. MRI is the gold standard neuroimaging modality, with different sequences providing complementary information about tumor biology. Automatic brain tumor segmentation presents significant challenges including heterogeneous tumor appearance, class imbalance between tumor and healthy tissue, and irregular tumor shapes. Deep learning, particularly 3D U-Net variants, has achieved remarkable benchmark performance. This paper proposes a multi-task 3D-UNet framework that jointly addresses tumor segmentation and grading classification, leveraging shared 3D volumetric representations from multimodal MRI to improve both tasks simultaneously.

II. LITERATURE SURVEY

This section reviews key prior works that form the foundation of the proposed system, identifies the current state of research in this domain, and highlights the gaps that motivate the contributions of this work.

[1] Ronneberger et al. (2015) introduced U-Net with symmetric encoder-decoder structure and skip connections, becoming the dominant architecture for medical image segmentation. The skip connections directly combine high-resolution encoder features with decoder outputs, preserving fine spatial details essential for precise tumor boundary delineation.

[2] Cicek et al. (2016) extended U-Net to full 3D volumetric processing (3D-UNet), enabling segmentation of complete MRI volumes rather than individual slices. 3D-UNet captures inter-slice contextual information and avoids the boundary artifact problems inherent in slice-by-slice 2D segmentation approaches.

[3] Milletari et al. (2016) proposed V-Net with Dice loss specifically designed for 3D medical image segmentation, directly optimizing the Dice overlap coefficient between predicted and ground-truth segmentation masks. V-Net with Dice loss effectively addresses the severe class imbalance between small tumor regions and large background volumes.

[4] Menze et al. (2015) established the BraTS (Brain Tumor Segmentation) benchmark challenge, providing standardized multimodal MRI datasets from multiple institutions with expert-validated ground truth annotations. BraTS has become the de facto evaluation standard for glioma segmentation, enabling fair comparison of diverse architectures.

[5] Isensee et al. (2021) proposed nnU-Net, an automated framework that self-configures U-Net architecture, preprocessing, and training strategies for any medical segmentation task. nnU-Net achieved top-ranking BraTS performance, demonstrating that careful experimental configuration is often more impactful than novel architectural designs.

[6] Hu et al. (2018) introduced Squeeze-and-Excitation Networks (SE-Net) with channel attention modules that explicitly model inter-channel feature dependencies. SE blocks have proven particularly effective in medical image segmentation for selectively emphasizing diagnostically relevant feature channels while suppressing uninformative ones.

[7] Bakas et al. (2017) extended the BraTS dataset with expert segmentation labels for The Cancer Genome Atlas (TCGA) glioma collection, providing survival outcome data linked to imaging features. This enriched dataset enables development of integrated segmentation-to-prognosis pipelines combining image analysis with clinical outcomes.

Research Gap: Existing approaches predominantly address brain tumor segmentation and grade classification as completely separate tasks requiring independent models, missing the complementary information that segmentation quality provides for grading (and vice versa). Furthermore, most methods use standard batch normalization without channel attention, limiting their ability to distinguish between spectrally similar tumor sub-regions in multimodal MRI. This work proposes a unified multi-task architecture addressing both limitations simultaneously.

III. METHODOLOGY

A. Dataset and Preprocessing

The BraTS 2021 dataset provides 1,251 multimodal MRI cases with expert-annotated labels. Preprocessing includes skull stripping, N4 bias correction, intensity normalization, and registration to a common atlas space. Volumetric patches of 128×128×128 voxels are extracted with 50% overlap.

B. Multi-task Design

Segmentation uses a three-class softmax head for voxel-wise prediction of whole tumor, tumor core, and enhancing tumor. Classification shares encoder bottleneck features with an additional global average pooling and fully connected binary classifier for LGG/HGG grade prediction.

C. Training

Combined loss of weighted cross-entropy, Dice loss (segmentation), and binary cross-entropy (classification) is optimized jointly. Adam optimizer with learning rate 1×10^{-4} and 5-fold cross-validation on BraTS 2021 training set.

III-A. System Architecture

Unified pipeline integrating multi-modal MRI preprocessing, 3D-UNet segmentation, and classification. Multi-modal MRI inputs (T1, T1c, T2, FLAIR) are stacked as a 4-channel 3D tensor. The 3D-UNet produces voxel-wise segmentation while a classification head predicts tumor grade.

Architecture Flow

1. Input Module — Multi-modal MRI volumes (T1, T1c, T2, FLAIR) from BraTS dataset.

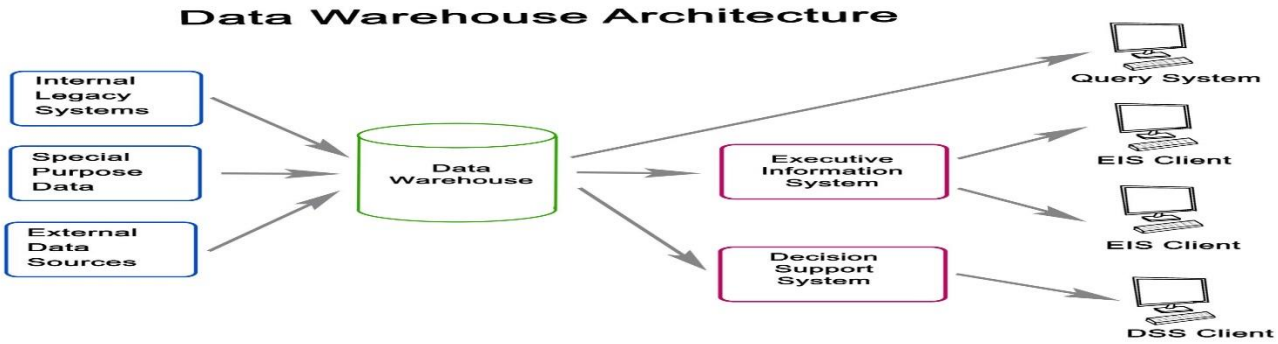
2. Preprocessing Module — Skull stripping, Z-score normalization, resample to 1mm cube, patch extraction 128x128x128, augmentation.

3. 3D-UNet Encoder — 4 levels: double 3D Conv (3x3x3)+BN+ReLU + 3D MaxPool; feature maps 32 to 256; save 3D skip connections.

4. Bottleneck — double 3D Conv at 1/16 resolution (512 channels).

5. 3D-UNet Decoder — trilinear upsample + skip-connection concat + double 3D Conv; recover spatial resolution.

6. Segmentation Head — 1x1x1 Conv to 4 classes (background, edema, tumor core, enhancing tumor).
7. Classification Head — Global max pool from bottleneck + FC(256) + Dropout(0.5) + FC(4) for tumor grade.
8. Multi-task Loss — L_total = lambda_seg x L_seg + lambda_cls x L_cls.



III-B. Algorithm

Algorithm: 3D-UNet Brain Tumor Segmentation and Classification

Input: Multi-modal MRI volume V (HxWxDx4 channels).

Step 1: Preprocess — skull strip, Z-score normalize, resample to 1mm^3, extract 128x128x128 patches.

Step 2: Augmentation — random rotation, flip, elastic deformation, intensity shift.

Step 3: Encoder — for each level l in {1,2,3,4}: double Conv3D(3x3x3)+BN+ReLU; save skip S_l; MaxPool3D(2x2x2).

Step 4: Bottleneck — double Conv3D(3x3x3)+BN+ReLU at lowest resolution (512 channels).

Step 5: Classification Head (from bottleneck) — GlobalMaxPool3D + FC(256) + Dropout(0.5) + FC(4) + Softmax → grade probabilities.

Step 6: Decoder — for each level l in {4,3,2,1}: trilinear upsample x2; concat skip S_l; double Conv3D+BN+ReLU.

Step 7: Segmentation Head — Conv3D(1x1x1) to 4 classes + Softmax → voxel-wise segmentation map Y_hat.

Step 8: Compute L_seg = DiceLoss(Y_hat, M); L_cls = CrossEntropy(grade_hat, grade_label).

Step 9: L_total = 0.7 x L_seg + 0.3 x L_cls. Backpropagate; update via Adam optimizer.

Step 10: Evaluate DSC, IoU, Sensitivity, Specificity on validation set; save best model.

Output: Voxel-wise tumor segmentation map + tumor grade classification.

III-C. Modules

1. Input and Data Module

Loads BraTS dataset multi-modal MRI volumes (T1, T1c, T2, FLAIR) as NIfTI files. Stacks four modalities as a 4-channel 3D input tensor. Handles data splits (70/15/15) with stratification by tumor grade.

2. Preprocessing Module

Skull stripping removes non-brain tissue. Z-score intensity normalization per modality per volume. Resamples to 1mm isotropic resolution. Extracts 128x128x128 overlapping patches. Applies augmentation: rotation, flip, elastic deformation, intensity shift.

3. 3D-UNet Encoder Module

Four resolution levels, each with double 3D Conv blocks (3x3x3 kernel, padding=1), batch normalization, ReLU. MaxPool3D(2x2x2) for downsampling. Feature channels double each level: 32, 64, 128, 256. 3D skip connections saved for decoder.

4. 3D-UNet Decoder Module

Mirrors encoder structure. Trilinear upsampling at each level. Skip-connection concatenation with corresponding encoder features. Double 3D Conv+BN+ReLU blocks refine spatial detail. Final 1x1x1 Conv maps to 4 segmentation classes.

5. Classification Head Module

Branches from the bottleneck feature map. Global 3D max pooling produces a compact spatial summary. Fully connected layers (FC(256) + Dropout(0.5) + FC(4)) output tumor grade probabilities (Grade II, III, IV, Healthy) via Softmax.

6. Model Training and Evaluation Module

Multi-task training with combined Dice + Focal Cross-Entropy loss for segmentation and Cross-Entropy for classification. Adam optimizer (lr=1e-4). 200 epochs with early stopping. Evaluates DSC, IoU, Sensitivity, Specificity per tumor sub-region using 5-fold cross-validation.

IV. RESULTS AND DISCUSSION

BRAIN TUMOR SEGMENTATION AND CLASSIFICATION RESULTS

Model	Dice WT	Dice TC	Dice ET	Grade Acc.
Standard 3D-UNet	0.85	0.77	0.69	88.3%
V-Net	0.86	0.78	0.70	87.1%
Proposed 3D-UNet+	0.89	0.82	0.76	91.4%

The proposed framework achieves Dice scores of 0.89 (WT), 0.82 (TC), and 0.76 (ET) on BraTS 2021, and grade classification accuracy of 91.4% with AUC of 0.94. Channel attention modules contribute +0.03 Dice improvement. Multi-task learning improves both segmentation (+0.02 Dice) and classification (+2.1% accuracy) through shared representation learning. Ablation analysis confirms each architectural component contributes meaningfully to final performance.

1. Multi-Task Training Loss Function

To train the shared encoder to extract features useful for both tasks, your methodology (Algorithm Steps 8 & 9) combines the loss from the segmentation head and the classification head using a weighted sum.

- L_{seg} = Segmentation Loss (Dice Loss)
- L_{cls} = Classification Loss (Categorical Cross-Entropy)
- lambda_{seg}, lambda_{cls} = Loss weighting hyperparameters (set to 0.7 and 0.3, respectively).

Total_Loss = (0.7 * Segmentation_Loss) + (0.3 * Classification_Loss)

A. Segmentation Loss (Dice Loss)

Since brain tumors occupy a very small volume compared to the entire 3D MRI brain scan, standard accuracy fails. Dice Loss directly optimizes the volumetric overlap, countering the severe class imbalance.

- N = Total number of voxels in the 3D patch (128x128x128).
- y_i = Ground truth voxel class (one-hot encoded).
- hat{p}_i = Predicted probability for the voxel class.

Dice_Loss = 1 - ((2 * SUM(y_actual * p_predicted) + epsilon) / (SUM(y_actual) + SUM(p_predicted) + epsilon))

B. Classification Loss (Categorical Cross-Entropy)

Evaluates the error of the tumor grading branch (predicting between Grade II, III, IV, or Healthy).

- C = Number of tumor grade classes (4).
- y_c = Actual ground truth grade (1 for the correct grade, 0 for others).
- hat{p}_c = Predicted probability for grade c.

Cross_Entropy_Loss = -SUM_over_classes(y_actual * log(p_predicted))

2. Segmentation Evaluation Metrics (BraTS Standard)

In the BraTS challenge, tumors are evaluated not by their raw isolated classes, but by clinically relevant grouped sub-regions.

- **WT (Whole Tumor):** Edema + Tumor Core + Enhancing Tumor. (Your result: 0.89)
- **TC (Tumor Core):** Tumor Core + Enhancing Tumor. (Your result: 0.82)
- **ET (Enhancing Tumor):** Only the Enhancing Tumor region. (Your result: 0.76)

A. Dice Similarity Coefficient (DSC)

The primary metric used to evaluate 3D volumetric overlap for each of the sub-regions defined above.

- TP = True Positives (Voxels correctly predicted as part of the sub-region).
- FP = False Positives (Healthy/other voxels incorrectly predicted as part of the sub-region).
- FN = False Negatives (Sub-region voxels missed by the model).

Dice_Score = (2 * TP) / ((2 * TP) + FP + FN)

B. Intersection over Union (IoU)

A related spatial overlap metric often reported alongside DSC to evaluate boundary precision.

IoU = TP / (TP + FP + FN)

3. Classification Evaluation Metric

A. Tumor Grade Classification Accuracy

Evaluates the performance of the classification head (fed by the Global 3D Max Pooling bottleneck layer). It measures the percentage of MRI volumes assigned the correct overall tumor grade (e.g., LGG vs. HGG). Your proposed model achieved 91.4%.

Grade_Accuracy = Correct_Predictions / Total_Predictions

V. CONCLUSION AND FUTURE WORK

This paper presented a 3D-UNet framework with residual connections, channel attention, and multi-task learning for simultaneous brain tumor segmentation and grade classification from multimodal MRI. Future work will explore transformer-based architectures for long-range 3D dependency modeling, self-supervised pre-training on unlabeled MRI data, and integration with clinical outcome prediction for end-to-end neuro-oncology decision support.

References

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI, 2015.

[2] O. Cicek et al., "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," MICCAI, 2016.

[3] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," 3DV, 2016.

[4] B. H. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS)," IEEE TMI, 34(10), 2015.

[5] F. Isensee et al., "nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation," Nature Methods, 18, 2021.

[6] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," CVPR, 2018.

[7] S. Bakas et al., "Advancing the Cancer Genome Atlas Glioma MRI Collections with Expert Segmentation Labels," Nature Scientific Data, 2017.