



International Journal of Intelligent Systems Technologies and Applications

ISSN online: 1740-8873 - ISSN print: 1740-8865
<https://www.inderscience.com/ijista>

Brain tumour segmentation for overall survival prediction

Novsheena Rasool, Javaid Iqbal Bhat

DOI: [10.1504/IJISTA.2025.10069003](https://doi.org/10.1504/IJISTA.2025.10069003)

Article History:

Received:	10 August 2024
Last revised:	20 September 2024
Accepted:	25 December 2024
Published online:	30 September 2025

Brain tumour segmentation for overall survival prediction

Novsheena Rasool* and Javaid Iqbal Bhat

Department of Computer Science,
Islamic University of Science and Technology,
Awantipora, Kashmir, 192122, India
Email: novsheena.rasool@iust.ac.in
Email: javaidiqbal@iust.ac.in
Email: Javaid.iqbal@iust.ac.in
*Corresponding author

Abstract: Gliomas present significant challenges due to their heterogeneous and infiltrative nature, making accurate segmentation essential for effective treatment. Manual segmentation methods are highly labour-intensive and often inadequate. This study introduces a novel pipeline for improving glioma management, beginning with advanced MRI preprocessing. We propose two attention-gated UNet architectures, the dual convolution attention gated UNet and the channel attention gated UNet, for precise tumour segmentation. Radiomic features, including the grey-level co-occurrence matrix and grey-level dependence matrix, are extracted to capture detailed tumour characteristics. Clinical data, such as age and resection status, are integrated alongside radiomic features to enhance survival models. A stacking ensemble model, combining a random forest regressor and multilayer perceptron, predicts survival based on integrated data. Validation on the BraTS 2018 dataset shows that dual convolution attention gated UNet excels in both segmentation accuracy and survival prediction, highlighting the potential of these advanced technologies for glioma management.

Keywords: gliomas; segmentation; dual channel attention gated UNet; survival prediction.

Reference to this paper should be made as follows: Rasool, N. and Bhat, J.I. (2025) 'Brain tumour segmentation for overall survival prediction', *Int. J. Intelligent Systems Technologies and Applications*, Vol. 23, No. 3, pp.295–318.

Biographical notes: Novsheena Rasool received her BTech in Computer Science and Engineering from the University of Kashmir in 2016 and MSc degree in Information Technology from the Islamic University of Science and Technology (IUST), Kashmir, Jammu and Kashmir, India in 2018, where she is currently pursuing her PhD at the Department of Computer Science. Her research interests include deep learning, machine learning, and medical image analysis. And she has served as the first author on several papers indexed in Scopus.

Javaid Iqbal Bhat received his PhD in Computer Science from the University of Kashmir, Jammu and Kashmir, India. He is currently an Associate Professor and the Director of IT&SS at the Islamic University of Science and Technology (IUST), Kashmir, Jammu and Kashmir. His research interests include vision computing, social network analysis, medical image analysis and data analytics.

1 Introduction

In the field of medical and biological engineering, finding innovative solutions to pressing health challenges remains crucial (Nitish and Singh, 2020; Wani et al., 2024b). One such challenge of significant concern is the presence of glioma brain tumours, where early detection is vital for effective treatment. Gliomas are characterised by their intricate nature and represent a notable subset within this domain. These tumours are marked by irregular cell proliferation within the brain, posing a substantial threat to neurological function by disrupting surrounding brain tissue. Gliomas are classified by the World Health Organization (WHO) into four grades, ranging from grade I to IV. Grades I and II are referred to as low-grade gliomas (LGG), while grades III and IV are classified as high-grade gliomas (HGG). LGGs comprise a subset of benign brain tumours, whereas HGGs, particularly grade IV tumours commonly known as glioblastomas are highly aggressive (Rasool and Bhat, 2023b). HGGs exhibit distinct regions due to their diverse nature and intricate growth patterns in the brain. Typically, HGGs consist of three primary components, the core tumour (CT), the enhancing tumour (ET) and the oedema (ED). The core or central tumour, represents the innermost and most aggressive part of the glioma, characterised by high cell density and rapid cancer cell proliferation (Osborn et al., 2022; Rasool et al., 2024a). This region often contains necrotic tissue due to limited blood supply, making it the most malignant and challenging part of the tumour. Surrounding the central area is the oedema, marked by swelling of brain tissue caused by the tumours' presence. Accurate segmentation of the oedema area is pivotal, as it facilitates the assessment of the tumour's impact on adjacent healthy brain tissue, thereby informing surgical and therapeutic choices (Santucci et al., 2021). The third component is the enhancing tumour, characterised by the active growth of the tumour and frequently exhibiting heightened blood supply. This region is crucial in surgical planning and treatment decisions, representing the glioma's most actively growing and treatable segment (Rasool and Bhat, 2023a). Together, the core, oedema, and enhancing tumour encompass the entirety of the tumour, also referred to as the whole tumour (WT). Medical imaging techniques, encompassing x-radiations (X-rays), computed tomography (CT) scans, positron emission tomography (PET) scans, and magnetic resonance imaging (MRI), are indispensable for identifying, and delineating gliomas (Yaqub et al., 2023; Rasool and Bhat, 2023c). Among these techniques, MRI is particularly notable for its ability to differentiate various types of brain tissues, offering multiple image contrasts. MRI's superior soft tissue contrast aids in distinguishing between healthy and tumour tissue, streamlining the localisation of tumour boundaries. Furthermore, MRI produces diverse image types, including t1-weighted (T1), t1 contrast-enhanced (T1ce), fluid attenuated inversion recovery (FLAIR), and t2-weighted (T2) sequences (see Table 1), highlighting

specific tissue characteristics such as water content, blood flow, and cellular density, which contribute to precise visualisation (Ronneberger et al., 2015). The MRI-based brain tumour segmentation (BraTS) datasets, specifically BraTS 2018 used in this study, exhibit a significant class imbalance with 98.46% of voxels belongs to background class, 1.02% are oedema, 0.29% are enhancing tumours, and 0.23% are necrotic or non-enhancing tumour regions (Chang et al., 2023; Raza et al., 2023). This imbalance exacerbates the challenges of conventional methods for diagnosing diseases from MRI images, particularly manual segmentation techniques, which remain labour-intensive and time-consuming. These methods rely heavily on the expertise of radiologists or clinicians, who must carefully delineate tumour boundaries on individual image slices, further complicating the process in the presence of imbalanced and subtle tumour regions (Rasool and Bhat, 2024a). This manual approach is not only painstaking but also introduces significant variability among different observers, leading to inconsistencies in the segmentation results (Pathan et al., 2020). Moreover, traditional methods often struggle with accurately segmenting gliomas in multimodal MRI images due to the complex and variable nature of tumour appearances across different imaging modalities. These methods are not well-equipped to handle large-scale imaging data. As the volume of medical imaging data continues to grow, manual segmentation processes become increasingly impractical and time-consuming. Another limitation of traditional approaches is their inability to predict patient survival outcomes effectively, as this task involves analysing intricate patterns and relationships within imaging data and patient features.

Table 1 MRI sequences and mask of one subject from the BraTS dataset

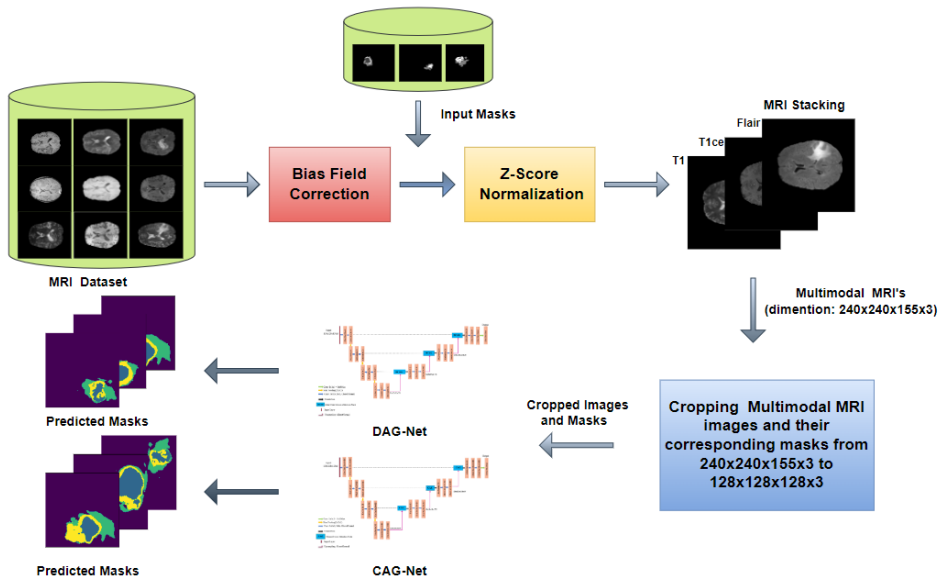
<i>Subject</i>	<i>MRI modalities</i>	<i>Ground truth</i>
1	T1, T2, T1ce, FLAIR, ground truth	0: Background 1: Core tumour 2: Oedema 4: Enhancing tumour

To overcome these limitations, the integration of deep learning (DL) technologies into medical imaging has emerged as a promising solution (Wahlang et al., 2020; Wani et al., 2024c). DL-based automated segmentation algorithms, in particular, offer substantial potential to enhance both the accuracy and efficiency of glioma diagnosis and treatment planning (Rasool et al., 2024b). These technologies significantly reduce the burden on radiologists and clinicians by automating the segmentation process and allowing more precise and quicker diagnoses. Moreover, DL techniques can scale effectively to large datasets, addressing the scalability issues faced by traditional methods. Recent advancements in DL have led to considerable progress in various domains (Sabut et al., 2022; Wani et al., 2024a; Zhou et al., 2022), particularly in glioma segmentation using 3D MRI images, as well as in predicting patient survival. Many researchers have utilised convolutional neural network (CNN)-based architectures to tackle the complex task of multiclass glioma segmentation. Approaches such as encoder-decoder-based 2D U-Net models (Ottom et al., 2022) and 3D U-Net models (Rasool et al., 2024b) have been particularly successful in segmenting different tumour subregions. Other CNN variants, including Edge U-Net models (Allah et al., 2023) and MBANet (Cao et al., 2023) have further improved segmentation performance,

as evidenced by enhanced dice similarity coefficient (DSC) scores, showcasing the adaptability of DL approaches in handling glioma segmentation tasks. Beyond CNN models, other deep learning techniques suggested by (Fang and Wang, 2022; Younis et al., 2022) have shown success in segmenting gliomas by identifying critical regions such as the enhancing core, necrotic core, and oedema. These regions are vital for treatment planning and monitoring disease progression. Parallel to the advancements in segmentation, machine learning (ML) techniques have been extensively explored for predicting patient survival. Predicting survival outcomes involves analysing complex interactions between imaging features and clinical variables. Various ML and statistical models have been applied, including random forest regressor (RFR) (Das et al., 2022), multi-layer perceptrons (MLP) (Mazher et al., 2024), support vector machines (SVM) (Das et al., 2022), AdaBoost (Rikan et al., 2024), and XGBoost (Wani et al., 2024c). These models have demonstrated their ability to handle large feature sets and uncover meaningful relationships between imaging features and survival outcomes, providing valuable prognostic insights.

Our approach goes a step further by prioritising comprehensive preprocessing and innovative model architectures to enhance segmentation accuracy. Specifically, we introduce the dual convolution attention gated UNet (DCAG-Net) and channel attention gated UNet (CAG-Net) models, which play a crucial role in our contribution. Figures 1 and 9 illustrate the key components of our proposed pipeline for precise segmentation and overall survival prediction.

Figure 1 Proposed methodology for brain tumour segmentation (see online version for colours)



The DCAG-Net leverages dual attention mechanisms to capture and integrate multi-scale contextual information, enhancing the precision of tumour boundary delineation by refining features. By selectively attending to relevant features from both encoder and decoder pathways, this network architecture improves segmentation accuracy and

ensures effective utilisation of contextual cues. Meanwhile, the CAG-Net focuses on emphasising the most informative channels in the feature maps, enhancing the model's ability to identify relevant structures within the MRI images. In our methodology, we utilise stacking ensemble model for survival prediction. However, instead of combining the outputs of both segmentation models simultaneously, we adopt a sequential approach. Initially, we feed the segmented outputs from the DCAG-Net into the stacking ensemble model and evaluate the results. Subsequently, we repeat the process with the segmented outputs from the CAG-Net. This sequential integration of segmentation outputs allows us to assess the performance of each model individually within the ensemble framework, providing insights into their respective contributions to survival prediction. Through this approach, we leverage the strengths of each segmentation model while maintaining a clear distinction between their roles within the ensemble.

The paper is structured into four sections, each focusing on different aspects of brain tumour segmentation and its role in overall survival prediction. Section 1 introduces the topic, emphasising the importance of accurate glioma segmentation for improving survival prediction. Section 2 outlines the proposed methodology, detailing preprocessing techniques and introducing two segmentation models, the DCAG-Net and CAG-Net. It also discusses the training process and presents segmentation results. Section 3 focuses on predicting overall survival (OS) based on the segmentation outcomes. Finally, Section 4 summarises key findings and suggests potential avenues for future research in this domain.

1.1 Motivation

Gliomas are the most common primary tumours of the central nervous system and pose significant challenges in neuro-oncology due to their complex characteristics, including heterogeneous subregions, unpredictable locations, and indistinct borders (Rasool and Bhat, 2024b). Manual segmentation is laborious and hinders prompt treatment planning. Automated methods are needed to streamline this process and improve patient outcomes. Deep learning technologies offer promising solutions by enabling precise tumour identification and quantification. By integrating deep learning technologies into gliomas segmentation, we aim to enhance the efficiency of tumour delineation and provide clinicians with valuable insights into patient prognosis, and ultimately support more personalised and effective neuro-oncological treatments.

1.2 Contributions of this article

- 1 We introduce DCAG-Net, a novel dual convolution attention gated UNet, for improved brain tumour segmentation in multimodal MRI images. Our model dynamically weighs feature maps, enhancing information flow across encoder-decoder pathways. Integration of attention mechanisms selectively emphasises relevant features, yielding superior segmentation performance.
- 2 We proposed another CAG-Net, a new channel attention gated UNet to optimise semantic segmentation of brain tumours in multimodal MRI images. These attention gates dynamically prioritise relevant features within the network's channels, refining segmentation accuracy by emphasising crucial information

- 3 We conducted a comprehensive performance comparison between the proposed DCAG-Net and CAG-Net, utilising a range of evaluation metrics. Our analysis reveals that DCAG-Net consistently outperforms CAG-Net across all metrics, demonstrating its superior performance in semantic segmentation of brain tumours in multimodal MRI images.
- 4 We developed a novel stacked ensemble approach, combining predictions from a RFR and a MLP for overall survival prediction of brain tumours. Radiomics features extracted from predicted segmented images obtained from DCAG-Net serve as input in phase 1, while radiomics features from CAG-Net are used in phase 2.
- 5 The proposed stacked ensemble optimises the accuracy of survival predictions in both phases. However, our results indicate that overall survival prediction based on input radiomics features in phase 1 surpasses that in phase 2.

2 Proposed methodology

2.1 Pre-processing

We use the BraTS 2018 dataset for training and validation which has undergone preprocessing by the organisers, including skull-stripping, interpolation to 1 mm³ isotropic resolution, and alignment to a shared anatomical template. Additionally, we applied further preprocessing techniques to enhance image quality. Firstly, we apply the N4 Bias Field Correction (N4ITK) technique, as recommended by Tustison et al. (2010). This technique corrects the bias field in the images, ensuring uniform and accurate intensity levels throughout the dataset. After bias correction, we apply z-score normalisation separately on the T2, T1ce, and FLAIR sequences. The mathematical expression for z-score normalisation is as follows:

$$Z(x, y, z) = \frac{V(x, y, z) - \mu}{\sigma} \quad (1)$$

where $Z(x, y, z)$ denotes the standardised value at a specific voxel in the image, which is the result of the normalisation process. $V(x, y, z)$ represents the original voxel intensity at the corresponding location (x, y, z) in the image, μ is the mean voxel intensity of the entire image, and σ is the standard deviation of voxel intensities in the image.

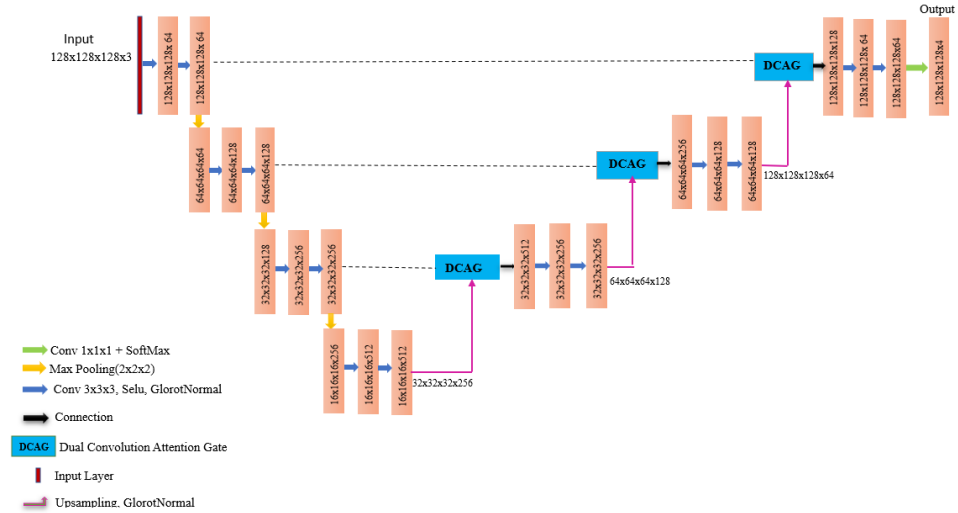
After individually normalising each MRI modality, we combine them into a comprehensive volume, where each voxel contains intensity data from all three modalities. This enhances the data's depth and richness by leveraging the complementary attributes of each modality. We stacked MRI images along the fourth axis and focused on critical areas by cropping to central $128 \times 128 \times 128$ blocks, excluding unnecessary background. This ensured uniformity and highlighted regions of interest. Furthermore, we validated segmentation masks by checking if non-background pixels exceeded a 0.0001 threshold. Although cropping reduced border artefacts, it may risk omitting crucial boundary information and might require adjustments for different datasets or imaging modalities.

2.2 Proposed architectures for segmentation

2.2.1 Dual convolution attention gated UNet

This section introduces a novel dual convolution attention gated UNet (DCAG-Net) architecture for the multiclass segmentation of gliomas, utilising the BraTS 2018 training dataset. The proposed architecture operates on multimodal input MRI volumes with dimensions of $128 \times 128 \times 128 \times 3$. The architecture follows an encoder-decoder pathway paradigm as illustrated in Figure 2.

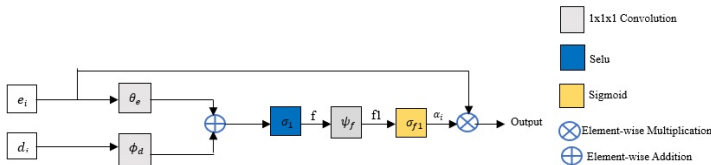
Figure 2 Proposed dual convolution attention gated UNet for multiclass semantic segmentation of gliomas (see online version for colours)



In the encoder segment, the architecture initiates with a primary convolutional layer employing 64 filters, utilising a $3 \times 3 \times 3$ kernel size, the same padding, and the scaled exponential linear units (SELU) activation function. This layer is the initial feature extractor, detecting basic patterns within the input data. Subsequently, the output is passed through a second convolutional layer with identical configurations, enhancing the network's ability to discern intricate features. This iterative process continues with subsequent convolutional layers, each employing SELU activation and same padding to maintain spatial consistency while gradually increasing the number of filters to 64, 128, 256 and finally 512, allowing for the capture of increasingly complex features. Furthermore, after every two convolutional layers within the encoder segment, max-pooling layers are incorporated with a pooling size of $2 \times 2 \times 2$. This design choice facilitates hierarchical feature extraction by downsampling the feature maps, reducing computational complexity, and extracting dominant features from the input data. The network effectively captures local and global features while preserving spatial consistency throughout the encoding process by employing max-pooling layers at regular intervals. In the decoder section of our architecture, we integrate three transpose convolutional layers to facilitate the upsampling of feature maps and reconstruct the segmentation map with finer spatial resolution. The first layer employs 256 filters,

followed by the second layer with 128 filters, and finally, the third layer with 64 filters. These filters determine the depth and complexity of the feature maps generated by each layer, aiding in the reconstruction of detailed spatial information. Furthermore, all transpose convolutional layers utilise a $2 \times 2 \times 2$ kernel size to define the receptive field and extent of upsampling. At the same time, same padding ensures that the output feature maps retain the exact spatial dimensions as the input feature maps. With strides set to $2 \times 2 \times 2$ for all layers, the upsampling factor is effectively doubled along each axis, contributing to the overall spatial resolution enhancement. Additionally, incorporating skip connections from the encoder, concatenated with corresponding feature maps in the decoder, enriches the upsampling process with detailed spatial information, aiding in recovering spatial details lost during encoding. The decoder segment effectively reconstructs high-resolution segmentation maps through these meticulously configured parameters, thereby augmenting the network's segmentation performance and accuracy. Furthermore, our architecture incorporates a dual convolution attention mechanism as a crucial element for dynamically adjusting the contributions of feature maps from both the encoder and decoder pathways. The attention gate comprises three convolutional layers. The initial two layers process input feature maps from the encoder and decoder paths, respectively, utilising $1 \times 1 \times 1$ kernel sizes and SELU activation functions to refine the feature representations. Subsequently, these processed feature maps are merged and forwarded to the third convolutional layer, which is responsible for computing attention scores using a sigmoid activation function. These scores, obtained from the third convolutional layer, indicate the relevance of each feature map to the segmentation task. By employing a sigmoid activation function, these scores are constrained to the range $[0, 1]$, reflecting the relative importance of each feature map. This gate selectively integrate encoder features passed through skip connections with the upsampled features generated by decoder's deconvolution layers. This mechanism assigns weights to features at each spatial location, modulating their contributions. The combined feature maps undergo further convolutional processing to refine the segmentation predictions. Unlike traditional attention mechanisms such as dot-product attention or self-attention, which compute attention scores based on the similarity between different parts of the input data (Younis et al., 2022). The proposed mechanism utilises convolutional layers and element-wise multiplication, offering distinct computational properties tailored specifically for semantic segmentation tasks in MRI images, emphasising the importance of spatial context and feature fusion. The architecture of the dual convolution attention gate unit comprises of several key components to enhance segmentation performance for brain tumour tasks, as shown in Figure 3.

Figure 3 Dual convolution attention gate (see online version for colours)



In this process, an encoder feature map e undergoes a $1 \times 1 \times 1$ convolution to produce a new tensor θ_e , defined as:

$$\theta_e = \text{Conv}1 \times 1 \times 1(e)$$

Similarly, a decoder feature map d undergoes $1 \times 1 \times 1$ convolution to produce a new tensor Φ_d :

$$\Phi_d = \text{Conv}1 \times 1 \times 1(d)$$

These tensors are then added element-wise, and the result is passed through an activation function, typically SELU, producing a tensor σ_1 :

$$\sigma_1 = \text{SELU}(\theta_e + \Phi_d)$$

The SELU activation function is defined as:

$$\text{SELU} = \lambda \begin{cases} x, & \text{if } x < 0 \\ \alpha(\exp(x) - 1), & \text{if } x \geq 0 \end{cases}$$

Here, λ is a scale parameter (typically set to $\lambda \approx 1.0507$), the input to the activation function is represented by x , α is the constant factor, typically set to $\alpha \approx 1.67326$.

Next, another $1 \times 1 \times 1$ convolution is applied to combined feature maps (σ_1), producing a single-channel attention map (Ψ_f):

$$\Psi_f = \text{Conv}1 \times 1 \times 1(f)$$

This attention map is processed using a sigmoid activation function to obtain attention coefficients in the range $[0, 1]$, where values closer to 1 signify increased attention to important regions, and values closer to 0 representing less relevant areas:

$$\text{Attention} = \text{Sigmoid}(\Psi_f).$$

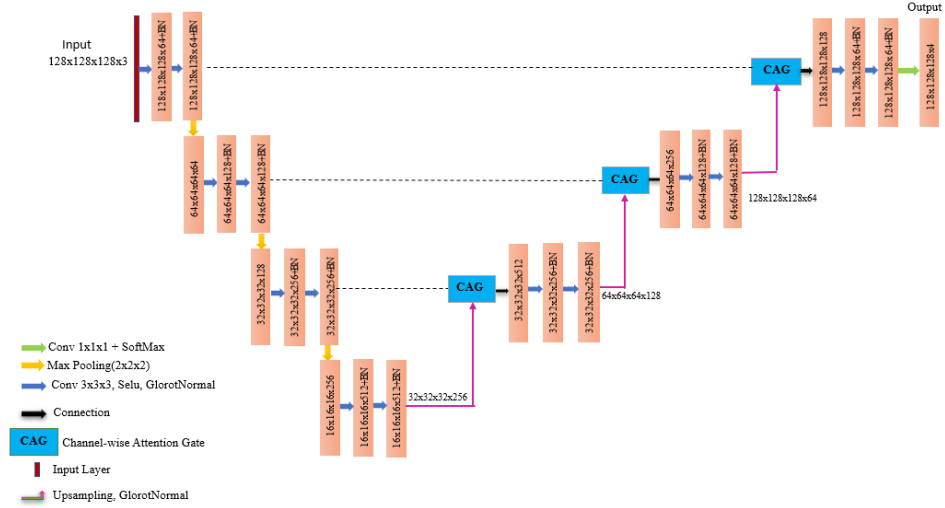
Finally, the original feature map e is multiplied element-wise by the attention coefficients to produce the modulated feature map, which is passed to the next layer of the network:

$$\text{Output} = e \odot \text{Attention}.$$

We apply attention selectively in the decoder to enhance feature refinement during segmentation. This selective attention prioritises relevant encoder features, improving accuracy without extra computational cost. Concentrating attention in the decoder optimises performance and efficiency, making our approach more powerful than conventional methods.

2.2.2 Channel attention gated UNet

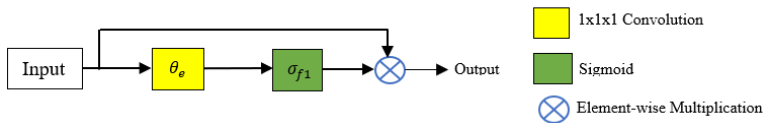
We have designed an additional UNet architecture for the same task, specifically the channel attention gated UNet (CAG-Net). This architecture closely aligns with the DCAG-Net as shown in Figure 4, but introduces crucial modifications, including channel attention gates based on Lee et al. (2023) and batch normalisation (BN) layers.

Figure 4 Proposed channel attention gated unit for multiclass semantic segmentation of gliomas (see online version for colours)

However, in our experimentation, the performance of the UNet with channel attention gates did not surpass that of the UNet with dual convolution attention gates.

We designed channel attention gates to selectively emphasise crucial features. Subsequently, we strategically incorporated these channel attention gates into the decoder. These gates utilise a channel-wise attention mechanism, generating attention maps to guide the model in highlighting specific regions of interest. Additionally, batch normalisation layers were introduced to aid training convergence, facilitating smoother learning as shown in Figure 5. Despite these enhancements, the model did not exhibit improved segmentation accuracy compared to the DCAG-Net. The channel attention gate operates in three key steps. Firstly, a $1 \times 1 \times 1$ convolution is applied to the input e with a weight matrix to produce a new tensor θ_e . A sigmoid activation function is employed on the convolution result, making a sigmoid activation result σ_{f1} . Finally, the original tensor e is element-wise multiplied by the sigmoid activation result, yielding the output of the channel attention gate.

$$\text{Output} = e \odot \sigma_{f1} \quad (2)$$

Figure 5 Channel attention gate (see online version for colours)

The primary objective behind employing two attention models (DCAG-Net and CAG-Net) was to assess the efficacy of distinct attention mechanisms in refining the precision of glioma boundary delineation. This investigation also aligns with a broader

aim of scrutinising and comprehending the diverse roles and effectiveness of attention mechanisms in glioma segmentation tasks.

2.2.2.1 Model training

Our model, implemented using tensorflow (Abadi et al., 2016), utilises MRI data from both LGG and HGG categories within the BraTS 2018 training dataset (Ben Ahmed et al., 2022). The dataset initially comprises 210 HGG and 75 LGG cases, as shown in Table 2. We combine these datasets and split the total into a 70:30 ratio, allocating 70% for training and 30% for validation (see Figures 6 and 7). Each patient's data in this dataset includes multiple imaging modalities such as FLAIR, T1-weighted, T2-weighted, and T1-ce scans, all presented as three-dimensional MRI scans with resolutions of $240 \times 240 \times 155$, formatted in NIfTI format. Manual annotations by organisers accompany each MRI scan, providing original labels for different tumour regions, categorised into four classes: label-0 for background, label-1 for core tumour, label-2 for oedema, and label-4 for enhancing tumour. These labels serve as benchmarks for evaluating and improving tumour delineation and categorisation techniques. Moreover, the dataset encompasses clinical information essential for predicting overall survival, which includes the patient's age, survival duration, and their resection status categorised as gross total resection (GTR), subtotal resection (STR), and not available (NA). At the outset, our dataset comprised 285 (HGG and LGG) images. Before preprocessing, we conducted thorough checks to ensure that each image had its corresponding mask available, and that all mask files contained the required four labels (0, 1, 2, and 4). Upon verification, we identified 27 masks lacking one or more of these labels. Consequently, we removed these incomplete mask files along with their corresponding images from our dataset. As a result of this verification process, our dataset was reduced to 257 images.

To handle the complexity of the data and reduce overfitting risks, we utilise a data generator with a batch size of 2. Our model is optimised using the adaptive moment estimator (Adam), which refines gradient updates by calculating the first and second moments. Through systematic experimentation, we determined that a learning rate of 0.0001 strikes the right balance between precision and convergence speed effectively. We set the epoch count to 250, to ensure sufficient iterations over the dataset, enhancing learning while monitoring performance to avoid overfitting. Dropout regularisation is also applied to further mitigate overfitting and improve model robustness. Recognising the importance of the loss function in deep learning, particularly for brain tumour segmentation, we evaluated several techniques, including focal loss (Ross and Dollár, 2017), cross-entropy loss (Peng and Sun, 2023), and generalised dice loss (Chang et al., 2023). Ultimately, we address class imbalances using two complementary strategies. Firstly, we calculate class weights using the 'balanced' method from scikit-learn, which assigns weights inversely proportional to class frequencies. This approach prioritises learning from minority classes and achieves more balanced segmentation performance across all classes. Specifically, tumour classes with fewer instances receive higher weights, while those with more instances receive lower weights. Secondly, we use the Dice coefficient loss function (Bertels et al., 2019) to improve performance on

imbalanced dataset by measuring spatial overlap between predicted and ground truth masks. Combining these methods ensures more balanced and precise segmentation of tumour classes. The dice coefficient loss function is defined as:

$$\text{Dice coefficient loss} = 1 - \frac{2 \times |X \cap Y|}{|X| + |Y|} \tag{3}$$

Table 2 Training dataset overview and preprocessing information

Original images		After combine	Split ratio		Image size	
HGG	LGG	Total images	Train (70%)	Validation (30%)	Before cropping	After cropping
210	75	285	200	85	240 × 240 × 155	128 × 128 × 128 × 3

Figure 6 Training data distribution (see online version for colours)

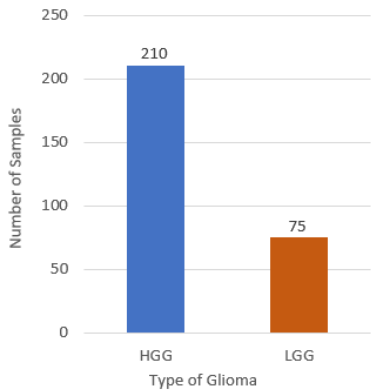
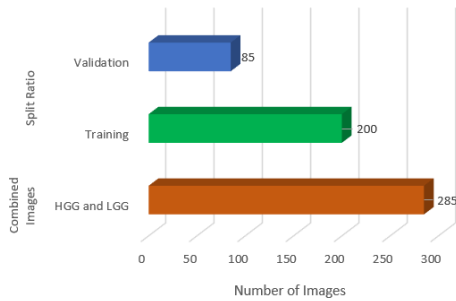


Figure 7 Split ratio of training dataset (see online version for colours)



2.2.2.2 Segmentation results

The proposed DCAG-Net and CAG-Net architectures demonstrate significant improvements in segmenting key regions of brain tumours, such as the whole tumour, core tumour, and enhancing tumour in MRI images. These regions are critical

for accurate diagnosis and treatment, making precise segmentation a fundamental step in improving clinical outcomes. The network hyperparameters of DCAG-Net are detailed in Table 3. Comparative analysis reveals that DCAG-Net outperforms CAG-Net as illustrated in Table 4. To further illustrate the performance of our approach, Figure 8 depicts various screenshots of the ground truth masks alongside their corresponding predicted masks generated by the proposed DCAG-Net architecture, including examples such as BRATS18.TCIA01_147_1, BRATS18.TCIA08_105_1, BRATS18.CBICA.ABW_1, BRATS18_2013_17_1, BRATS18_2013_23_1, BRATS18.TCIA09_141_1, BRATS18_2013_12_1, BRATS18.TCIA08_469_1, and BRATS18.TCIA02_368_1. In these images, different tumour regions are distinctly color-coded: necrotic and non-enhancing areas are represented in blue, oedema regions in yellow, and enhancing tumour regions in green. These visualisations emphasise the model's capability to delineate tumour subregions effectively, ensuring that even complex and subtle distinctions between areas are captured. In addition to the qualitative results, we provide a quantitative comparison of the DCAG-Net with other state-of-the-art methods as shown in Table 5. This comparison shows that our proposed architectures consistently outperform competing methods in segmentation accuracy. The integration of attention mechanisms within DCAG-Net and CAG-Net plays a key role in achieving these results by refining the focus on relevant tumour areas. However, the enhanced performance comes at the cost of increased model complexity. Both DCAG-Net and CAG-Net introduce additional parameters, primarily due to the inclusion of attention gates and convolutional layers. Despite this, the added complexity is effectively managed by NVIDIA Tesla V100 32 GB GPU, allowing the model to maintain computational efficiency without compromising performance. Overall, the increase in parameters is justified by the substantial improvement in segmentation accuracy, ensuring precise delineation of tumour subregions. The results show that the proposed models effectively balance accuracy and computational efficiency.

Table 3 Model hyperparameters and configuration

<i>Hyperparameters</i>	<i>Details</i>
Learning rate	0.0001
Epochs	250
Batch size	2
Activation function	SoftMax
Optimiser	Adam
Stride	$2 \times 2 \times 2$
Filters	64, 128, 256, 512
Kernel size	1, 3
Padding	Same
Kernel initialiser	GlorotNormal()
Pool size	$2 \times 2 \times 2$
Number of classes	4
Input shape	$128 \times 128 \times 128 \times 3$
Output activation	SoftMax
Output shape	$128 \times 128 \times 128 \times 4$

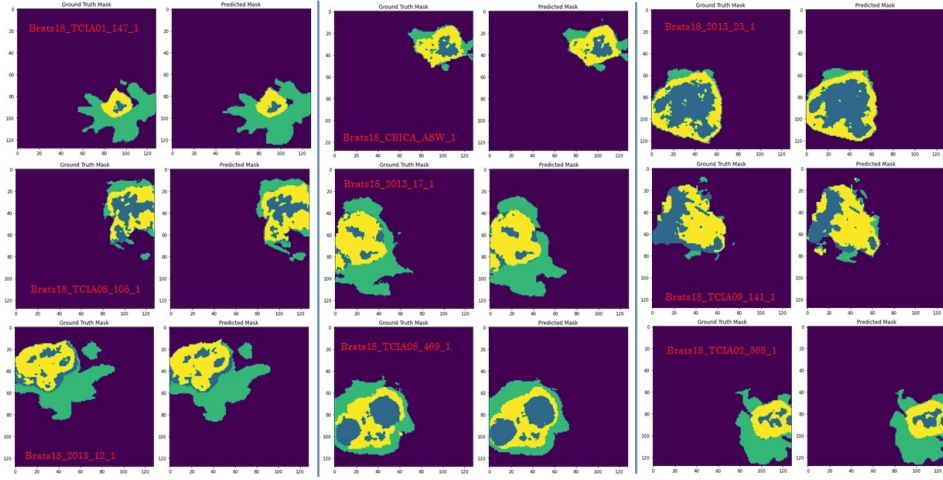
Table 4 Results of the proposed segmentation methods on the training and validation dataset

Methods	Modalities	Dataset	Dice	Accuracy	IOU score	Loss	Label	Precision	Specificity	Sensitivity	DSC
DCAG-Net	T1, T1ce, Flair	Train	0.94	0.99	0.87	0.05	WT	0.99	0.99	0.99	0.92
							CT	0.94	0.99	0.96	0.95
							ET	0.87	0.99	0.93	0.90
	T1, T1ce, Flair	Validation	0.88	0.98	0.76	0.11	WT	0.98	0.99	0.98	0.89
							CT	0.91	0.99	0.90	0.91
							ET	0.76	0.99	0.90	0.82
CAG-Net	T2, T1ce, Flair	Train	0.89	0.99	0.77	0.10	WT	0.99	0.99	0.99	0.86
							CT	0.85	0.99	0.88	0.87
							ET	0.73	0.99	0.86	0.79
	T1, T1ce, Flair	Validation	0.86	0.98	0.73	0.13	WT	0.98	0.99	0.98	0.84
							CT	0.89	0.99	0.86	0.87
							ET	0.75	0.99	0.83	0.80

Table 5 Performance comparison of our proposed method with other approaches

Reference	Architecture	Input image dimension	Dataset	DSC			Sensitivity			Specificity		
				WT	CT	ET	WT	CT	ET	WT	CT	ET
Sun and Wang (2022)	CNN	168×200	BraTS 2018	0.89	0.76	0.77	0.88	0.77	0.79	0.99	0.99	0.99
Ladkat et al. (2022)	DNN	$240 \times 240 \times 155$	BraTS 2019	0.82	0.82	0.72	0.83	0.84	0.82	0.85	0.72	0.93
Wang et al. (2021)	C-dense CNN	$16 \times 160 \times 160$	BraTS 2017	0.89	0.80	0.75	-	-	-	-	-	-
Ali et al. (2024)	3D-UNet	$128 \times 128 \times 128$	Local dataset	0.89	0.83	0.77	0.89	0.82	0.78	0.99	0.99	0.99
Zhou (2024)	Multimodal network	$128 \times 128 \times 128$	BraTS 2010	0.86	0.86	0.77	-	-	-	-	-	-
Ours	DCAG-Net	$12 \times 128 \times 128 \times 3$	BraTS 2018	0.89	0.91	0.82	0.98	0.90	0.90	0.99	0.99	0.99

Figure 8 Segmented tumour regions: whole tumour, core tumour, enhancing tumour (see online version for colours)



Note: colour-coded – core: sky-blue, enhancing: yellow, oedema: green, background: navy-blue.

2.2.2.3 State-of-the-art-techniques

Sun and Wang (2022) introduced a CNN architecture for tumour segmentation, comprising pre-CNN and post-CNN blocks for data reduction and refinement. Experiments utilised the BRATS2018 dataset, yielding mean dice scores of 77.2%, 89.2%, and 76.3% for enhancing tumour, whole tumour, and tumour core, respectively. Ladkat et al. (2022) presented an entirely automated approach for segmenting brain tumours, leveraging a combination of a mathematical model and deep neural networks (DNNs). The proposed method enhances each slice of the 3D image using the mathematical model before passing it through the 3D attention U-Net for tumour segmentation. Wang et al. (2021) introduced a lightweight CNN for brain tumour segmentation from multimodal MRI images. Their approach integrated a novel complementary convolution unit (CCU) to reduce memory usage and employed dense connections to accelerate network training and enable feature reuse. They decomposed the segmentation task into three binary subtasks to mitigate interference among brain tissues, merging segmentation across different axes for enhanced precision. Ali et al. (2024) developed a sophisticated 3D U-Net encoder-decoder system aimed at automatically detecting brain tumours and segmenting their sub-structures. Their proposed model comprises neural units, transition layer blocks, and skip connections. Evaluation of the model's performance utilised a dataset from a local hospital, encompassing 71 subjects. The proposed model achieved Dice scores of 0.891, 0.834, and 0.776 on this private dataset. Zhou (2024) introduced a novel brain tumour segmentation network utilising disentangled representation learning and contrastive learning. Their method involves a feature fusion module to capture multi-modal features, followed by disentangled representation learning to isolate factors related to tumour regions. Contrastive learning aids in extracting tumour region-relevant features, and segmentation decoders generate final results. Experiments on BraTS datasets show

the proposed approach's superiority over existing methods in both quantitative and qualitative evaluations.

3 Overall survival prediction

3.1 Feature extraction and selection

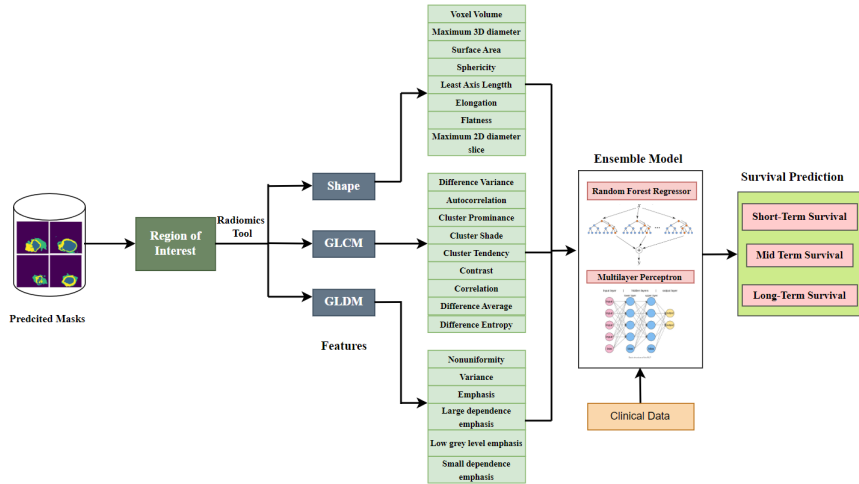
The feature extraction process plays a pivotal role in predicting patient survival. We extract features from predicted masks obtained after segmentation, from the region of interest by employing the radiomics package (van Griethuysen et al., 2017), a potent tool for extracting radiomic features. These features encompass tumours' spatial distribution, heterogeneity, and local image properties from 257 predicted segmented images. The radiomics package proves invaluable in our research, providing essential capabilities for extracting and analysing critical tumour characteristics. After extracting radiomic features from the predicted masks, we eliminate extraneous features from the dataset containing non-essential information such as configurations, python version, and settings. Our primary focus was retaining only the vital features necessary for thorough analysis. The selected features include various shape characteristics, including elongation, flatness, least axis length, major axis length, maximum 2D diameter slice, maximum 3D diameter, sphericity, surface area, and voxel volume. Additionally, the feature set includes properties derived from grey level co-occurrence matrix (GLCM), such as autocorrelation, cluster prominence, cluster shade, cluster tendency, contrast, correlation, difference average, difference entropy, and difference variance. Furthermore, grey level dependence matrix (GLDM) features are integrated, comprising non-uniformity, variance, emphasis, large dependence emphasis, low grey level emphasis, and small dependence emphasis. Our careful feature extraction and selection process refines our dataset to include only pertinent information for subsequent analysis. After feature selection, we combine selected radiomic features with clinical data, using the patient ID as the key. The clinical CSV data file encompasses essential information such as age, survival duration, patient ID, and resection status, providing a comprehensive dataset for survival prediction. In our experimentation, we divided the obtained dataset into 80:20 ratio, with 80% used for training and 20% for testing to predict overall survival. Our primary focus was on evaluating patients classified with gross total resection (GTR) and subtotal resection (STR) statuses. GTR typically denotes complete tumour removal, while STR indicates that a portion of the tumour remains. Although our initial intention was to concentrate solely on GTR status, the limited availability of patient data in this category prompted us to also consider patients with STR status. This decision was made to augment the dataset size and mitigate the potential risk of underfitting.

3.2 Ensemble model

Our model integrates two distinct predictive models, the random forest regressor (RFR) and the multilayer perceptron (MLP), to predict overall survival (OS) time in patients as illustrated in Figure 9. These models are combined through a stacking ensemble approach, which merges the strengths of both models to improve the overall accuracy of predictions. The input data for this ensemble model includes radiomic features extracted

from medical imaging and key clinical data such as patient age, survival duration, and resection status. By incorporating this diverse range of features, the model aims to capture both the imaging-related features and clinical factors that influence patient survival outcomes.

Figure 9 Proposed methodology for overall survival prediction (see online version for colours)



3.2.1 Random forest regressor model

We employed the random forest regressor (RFR) as a foundational model for the regression task, which is particularly well-suited for predicting continuous outcomes like survival time in days. The RFR operates through ensemble learning, constructing a collection of decision trees. Each tree is trained on a unique subset of input features and samples, allowing the model to capture different aspects of the data. This approach enables the RFR to effectively capture complex interactions between radiomic and clinical features and survival duration. In our methodology, we controlled the complexity of individual trees by limiting their depth to a maximum of 5, helping the model focus on significant data patterns while avoiding overfitting to noise or less relevant details. To further enhance robustness, we employed 100 trees (estimators) in the ensemble. A higher number of estimators generally reduces variance in the model's predictions, resulting in more stable and reliable outcomes. We optimised the RFR model using hyperparameter tuning with GridSearchCV, systematically exploring combinations of hyperparameters like tree depth and the number of estimators. This process identified the best configuration for our predictive task. The optimised RFR, as identified by this tuning process, was then applied to both training and testing datasets to predict overall survival times, yielding accurate and robust results across different patient cohorts.

3.2.2 Multilayer perceptron model

The second key component of our ensemble model is the multilayer perceptron (MLP) neural network, which plays a crucial role in capturing complex patterns from both radiomic and clinical data. Our MLP is structured with a deep architecture, comprising

multiple layers designed to transform the input data and uncover intricate relationships among features. It starts with an input layer that processes pre-processed radiomic and clinical data. The first hidden layer contains 256 neurons and uses the rectified linear unit (ReLU) activation function to introduce nonlinearity, allowing the model to learn complex, nonlinear patterns that linear models may miss. To enhance generalisation and prevent overfitting, we applied L2 regularisation to the weights of the first hidden layer, with a penalty term of 0.001. This helps constrain weight magnitudes, reducing the likelihood of overfitting to the training data and improving performance on unseen data. Additionally, a dropout layer follows the first hidden layer, randomly deactivating neurons during training to prevent over-reliance on any single neuron and encouraging the network to learn more robust features. After the first hidden layer, we included two additional hidden layers, with 512 and 256 neurons, respectively. Each of these layers incorporates dropout and L2 regularisation to ensure consistent regularisation across the network and maintain the model's ability to generalise effectively. The final layer of the MLP is the output layer, consisting of three neurons representing the three survival categories: long survivors (survival duration > 540 days), mid survivors (210–540 days), and short survivors (survival < 210 days). The SoftMax activation function transforms the network's output into probabilities, allowing for accurate classification into one of the three categories. To further enhance model performance and avoid overfitting, we implemented early stopping during training. Early stopping monitors the validation loss and halts training when no further improvement is observed, restoring the model to the weights from the epoch with the best validation performance for optimal prediction accuracy.

3.2.3 Stacking classifier (ensemble approach)

After separately training the RFR and the MLP models, we integrate their predictions using a stacking classifier. This ensemble method leverages the strengths of both models by employing a meta-classifier to combine their outputs into a final decision regarding the survival outcome. The stacking classifier operates by using the predictions generated by the RFR and MLP models as input features for the meta-classifier. The RFR is particularly adept at predicting precise survival durations, which are continuous values expressed in days. Its ensemble approach, consisting of multiple decision trees, enables it to capture intricate relationships and patterns in the data that contribute to accurate predictions of survival time. By averaging the predictions of the decision trees, the RFR model provides a robust estimate of survival duration, mitigating potential overfitting and ensuring reliability. On the other hand, the MLP excels at classifying patients into distinct survival categories based on the patterns and correlations it identifies in the data. The MLP's deep neural network architecture allows it to learn complex relationships between the features and survival outcomes, which is especially useful for distinguishing between different survival categories. The stacking classifier effectively combines these complementary strengths. The meta-classifier takes the outputs from both the RFR and MLP models and synthesises them to produce a final prediction. This method enhances the overall prediction accuracy by integrating the RFR's precise duration predictions with the MLP's categorical classifications. As a result, the stacking classifier benefits from the RFR's ability to provide detailed survival duration estimates and the MLP's capacity to classify patients into meaningful survival groups, leading to a more comprehensive and accurate final prediction as shown in Table 6. To

assess the effectiveness of the ensemble model, we use performance metrics including mean squared error (MSE), standard square error (SSE), median squared error (Median SE), and accuracy during both the training and evaluation phases. We compare our survival prediction approach with other methods, as illustrated in Table 7 and visually represented in Figure 10.

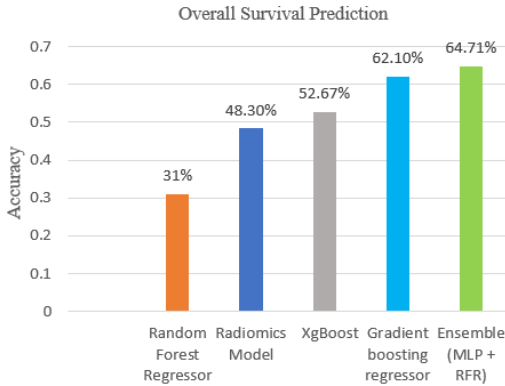
Table 6 Proposed survival prediction methods

Method	Overall survival	Accuracy (%)	MSE	Median SE	Standard SE
DCAG-Net	Ensemble (MLP+RFR)	64.71	170,152	131,068	95,047
CAG-Net	Ensemble (MLP+RFR)	52.94	201,697	159,762	165,293

Table 7 Comparison of proposed DCAG-Net with other survival prediction methods

Reference	Overall survival methods	Accuracy
Das et al. (2022)	XgBoost	52.67%
Rafi et al. (2020)	RFR	31%
Pálsson et al. (2022)	Radiomics model	48.3%
Rajput et al. (2022)	Gradient boosting regressor	62.1%
Proposed	Ensemble (MLP + RFR)	64.71%

Figure 10 Comparative analysis of survival prediction models (see online version for colours)



4 Conclusions and future scope

In our study, we introduced the DCAG-Net and CAG-Net models for glioma segmentation from MRI images, utilising the powerful computational capabilities of the NVIDIA Tesla V100 32 GB GPU. While both models were employed for segmentation, we further explored their utility beyond this task by utilising their segmented outputs separately in stacking ensemble model for survival prediction. The decision to treat the outputs of each model independently in stacking ensemble model was driven by our aim to comprehensively evaluate their individual performance in predicting glioma survival outcomes. Notably, our findings revealed that the DCAG-Net model,

when integrated into stacking ensemble model, consistently yielded superior results compared to the CAG-Net model. This outcome underscores the effectiveness of the DCAG-Net architecture in capturing relevant features crucial for survival prediction in glioma patients. Thus, our approach not only demonstrates methodological rigor but also provides valuable insights into the comparative performance of different segmentation models, ultimately enhancing the reliability and broader applicability of our findings. In the future, our aim is to refine our ensemble technique to enhance predictive performance even further, potentially by integrating the DCAG-Net and CAG-Net architectures. We also plan to delve deeper into methods for enhancing the interpretability of our models, aiming to uncover the imaging features or tumour characteristics driving the superior performance of these architectures. Augmenting our dataset with diverse samples and investigating techniques for improved generalisation will remain a priority, along with conducting prospective clinical validation studies to deploy our models in real-world healthcare settings. Additionally, we aspire to integrate imaging data with multi-omics information to gain a more comprehensive understanding of glioma biology and patient outcomes. Extending our methodology to encompass other types of brain tumours represents another avenue of exploration, with the goal of broadening the impact of our research on tumour imaging and treatment strategies. Through these efforts, we aim to advance glioma care and contribute to improved patient outcomes in clinical practice.

Data availability

The BRATS 2018 dataset is publicly available at <https://www.med.upenn.edu/sbia/brats2018/data.html>.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. et al. (2016) *Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, arXiv preprint arXiv:1603.04467.
- Ali, S., Khurram, R., Rehman, K.U., Yasin, A., Shaukat, Z., Sakhawat, Z. and Mujtaba, G. (2024) 'An improved 3D U-Net-based deep learning system for brain tumor segmentation using multi-modal MRI', *Multimedia Tools and Applications*, Vol. 83, No. 37, pp.1–20.
- Allah, A.M.G., Sarhan, A.M. and Elshennawy, N.M. (2023) 'Edge U-Net: brain tumor segmentation using MRI based on deep U-Net model with boundary information', *Expert Systems with Applications*, Vol. 213, No. 1, p.118833.
- Ben Ahmed, K., Hall, L.O., Goldgof, D.B. and Gatenby, R. (2022) 'Ensembles of convolutional neural networks for survival time estimation of high-grade glioma patients from multimodal MRI', *Diagnostics*, Vol. 12, No. 2, p.345.
- Bertels, J., Eelbode, T., Berman, M., Vandermeulen, D., Maes, F., Bisschops, R. and Blaschko, M.B. (2019) 'Optimizing the dice score and Jaccard index for medical image segmentation: theory and practice', *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019: 22nd International Conference, Proceedings, Part II*, 13–17 October, Springer, Shenzhen, China, pp.92–100.
- Cao, Y., Zhou, W., Zang, M., An, D., Feng, Y. and Yu, B. (2023) 'MBANet: a 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images', *Biomedical Signal Processing and Control*, Vol. 80, No. 1, p.104296.

- Chang, Y., Zheng, Z., Sun, Y., Zhao, M., Lu, Y. and Zhang, Y. (2023) 'DPAFNet: a residual dual-path attention-fusion convolutional neural network for multimodal brain tumor segmentation', *Biomedical Signal Processing and Control*, Vol. 79, No. 3, p.104037.
- Das, S., Bose, S., Nayak, G.K., Satapathy, S.C. and Saxena, S. (2022) 'Brain tumor segmentation and overall survival period prediction in glioblastoma multiforme using radiomic features', *Concurrency and Computation: Practice and Experience*, Vol. 34, No. 20, p.e6501.
- Fang, L. and Wang, X. (2022) 'Brain tumor segmentation based on the dual-path network of multi-modal MRI images', *Pattern Recognition*, Vol. 124, No. 4, p.108434.
- Ladkat, A.S., Bangare, S.L., Jagota, V., Sanober, S., Beram, S.M., Rane, K. and Singh, B.K. (2022) 'Deep neural network-based novel mathematical model for 3D brain tumor segmentation', *Computational Intelligence and Neuroscience*, Vol. 2022, No. 1, p.4271711.
- Lee, J., Wiratama, W., Lee, W., Marzuki, I. and Sim, D. (2023) 'Bilateral attention U-Net with dissimilarity attention gate for change detection on remote sensing imageries', *Applied Sciences*, Vol. 13, No. 4, p.2485.
- Mazher, M., Qayyum, A., Puig, D. and Abdel-Nasser, M. (2024) 'Deep learning-based survival prediction of brain tumor patients using attention-guided 3D convolutional neural network with radiomics approach from multimodality magnetic resonance imaging', *International Journal of Imaging Systems and Technology*, Vol. 34, No. 1, p.e23010.
- Nitish and Singh, A.K. (2020) 'Automatic detection classification and area calculation of brain tumour in MRI using wavelet transform and SVM classifier', *International Journal of Intelligent Systems Technologies and Applications*, Vol. 19, No. 6, pp.526–540.
- Osborn, A., Louis, D., Poussaint, T., Linscott, L. and Salzman, K. (2022) 'The 2021 World Health Organization classification of tumors of the central nervous system: what neuroradiologists need to know', *American Journal of Neuroradiology*, Vol. 43, No. 7, pp.928–937.
- Ottom, M.A., Rahman, H.A. and Dinov, I.D. (2022) 'ZNet: deep learning approach for 2D MRI brain tumor segmentation', *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 10, pp.1–8, Article Sequence Number 1800508.
- Pálsson, S., Cerri, S., Poulsen, H.S., Urup, T., Law, I. and van Leemput, K. (2022) 'Predicting survival of glioblastoma from automatic whole-brain and tumor segmentation of MR images', *Scientific Reports*, Vol. 12, No. 1, p.19744.
- Pathan, M.S., Jianbiao, Z., John, D., Nag, A. and Dev, S. (2020) 'Identifying stroke indicators using rough sets', *IEEE Access*, Vol. 8, pp.210318–210327, Article Sequence Number 9264165.
- Peng, Y. and Sun, J. (2023) 'The multimodal MRI brain tumor segmentation based on AD-Net', *Biomedical Signal Processing and Control*, Vol. 80, p.104336.
- Rafi, A., Ali, J., Akram, T., Fiaz, K., Shahid, A. R., Raza, B. and Mustafa Madni, T. (2020) 'U-Net based glioblastoma segmentation with patient's overall survival prediction', *Intelligent Computing Systems: Third International Symposium, ISICS 2020, Proceedings*, 18–19 March, Springer, Sharjah, United Arab Emirates, pp.22–32.
- Rajput, S., Agravat, R., Roy, M. and Raval, M.S. (2022) 'Glioblastoma multiforme patient survival prediction', *Proceedings of 2021 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2021) Medical Imaging and Computer-Aided Diagnosis*, Springer, pp.47–58.
- Rasool, N. and Bhat, J. (2023a) 'Multimodal brain tumor segmentation using 3D-U-Net', *Indian Journal of Natural Sciences*, Vol. 14, No. 78, pp.57473–57480.
- Rasool, N. and Bhat, J.I. (2023b) 'Glioma brain tumor segmentation using deep learning: a review', *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, IEEE, pp.484–489.
- Rasool, N. and Bhat, J.I. (2023c) 'Unveiling the complexity of medical imaging through deep learning approaches', *Chaos Theory and Applications*, Vol. 5, No. 4, pp.267–280.

- Rasool, N. and Bhat, J.I. (2024a) 'Brain tumour detection using machine and deep learning: a systematic review', *Multimedia Tools and Applications*, Vol. 83, No. 5, pp.1–54.
- Rasool, N. and Bhat, J.I. (2024b) 'A critical review on segmentation of glioma brain tumor and prediction of overall survival', *Archives of Computational Methods in Engineering*, Vol. 31, No. 4, pp.1–45.
- Rasool, N., Bhat, J.I., Aoun, N.B., Alharthi, A., Wani, N.A., Chopra, V. and Anwar, M.S. (2024a) 'ResMHA-Net: enhancing glioma segmentation and survival prediction using a novel deep learning framework', *Computers, Materials & Continua*, Vol. 81, No. 1.
- Rasool, N., Bhat, J.I., Wani, N.A., Ahmad, N. and Alshara, M. (2024b) 'TransResUNet: revolutionizing glioma brain tumor segmentation through transformer-enhanced residual UNet', *IEEE Access*, Vol. 12, No. 99, pp.72105–72116.
- Raza, R., Bajwa, U.I., Mehmood, Y., Anwar, M.W. and Jamal, M.H. (2023) 'dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI', *Biomedical Signal Processing and Control*, Vol. 79, No. 4, p.103861.
- Rikan, S.B., Azar, A.S., Naemi, A., Mohasefi, J.B., Pirnejad, H. and Wiil, U.K. (2024) 'Survival prediction of glioblastoma patients using modern deep learning and machine learning techniques', *Scientific Reports*, Vol. 14, No. 1, p.2371.
- Ronneberger, O., Fischer, P. and Brox, T. (2015) 'U-Net: convolutional networks for biomedical image segmentation', *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th international conference, Proceedings, Part III*, 5–9 October, Springer, Munich, Germany, pp.234–241.
- Ross, T-Y. and Dollár, G. (2017) 'Focal loss for dense object detection', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.2980–2988.
- Sabut, S., Patra, P. and Ray, A. (2022) 'Deep learning approach for classifying ischemic stroke using DWI sequences of brain MRIs', *International Journal of Intelligent Systems Technologies and Applications*, Vol. 20, No. 6, pp.524–535.
- Santucci, D., Faiella, E., Cordelli, E., Calabrese, A., Landi, R., de Felice, C., Zobel, B.B., Grasso, R.F., Iannello, G. and Soda, P. (2021) 'The impact of tumor edema on T2-weighted 3T-MRI invasive breast cancer histological characterization: a pilot radiomics study', *Cancers*, Vol. 13, No. 18, p.4635.
- Sun, Y. and Wang, C. (2022) 'A computation-efficient CNN system for high-quality brain tumor segmentation', *Biomedical Signal Processing and Control*, Vol. 74, p.103475.
- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A. and Gee, J.C. (2010) 'N4ITK: improved N3 bias correction', *IEEE Transactions on Medical Imaging*, Vol. 29, No. 6, pp.1310–1320.
- Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J-C., Pieper, S. and Aerts, H.J. (2017) 'Computational radiomics system to decode the radiographic phenotype', *Cancer Research*, Vol. 77, No. 21, pp.e104–e107.
- Wahlang, I., Sharma, P., Sanyal, S., Saha, G. and Maji, A.K. (2020) 'Deep learning techniques for classification of brain MRI', *International Journal of Intelligent Systems Technologies and Applications*, Vol. 19, No. 6, pp.571–588.
- Wang, Y., Peng, J. and Jia, Z. (2021) 'Brain tumor segmentation via C-dense convolutional neural network', *Progress in Artificial Intelligence*, Vol. 10, No. 2, pp.147–156.
- Wani, N.A., Kumar, R. and Bedi, J. (2024a) 'Deepexplainer: an interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence', *Computer Methods and Programs in Biomedicine*, Vol. 243, No. 1, p.107879.
- Wani, N.A., Kumar, R. and Bedi, J. (2024b) 'Harnessing fusion modeling for enhanced breast cancer classification through interpretable artificial intelligence and in-depth explanations', *Engineering Applications of Artificial Intelligence*, Vol. 136, p.108939.

- Wani, N.A., Kumar, R., Bedi, J., Rida, I. et al. (2024c) 'Explainable AI-driven iomt fusion: unravelling techniques, opportunities, and challenges with explainable AI in healthcare', *Information Fusion*, Vol. 110, No. 4, p.102472.
- Yaqub, M., Jinchao, F., Ahmed, S., Mehmood, A., Chuhan, I.S., Manan, M.A. and Pathan, M.S. (2023) 'DeepLabV3, IBCO-based ALCResNet: a fully automated classification, and grading system for brain tumor', *Alexandria Engineering Journal*, Vol. 76, No. 1, pp.609–627.
- Younis, A., Qiang, L., Nyatega, C.O., Adamu, M.J. and Kawuwa, H.B. (2022) 'Brain tumor analysis using deep learning and VGG-16 ensembling learning approaches', *Applied Sciences*, Vol. 12, No. 14, p.7282.
- Zhou, L., Nandal, A., Ganchev, T. and Dhaka, A. (2022) 'Breast cancer detection by fusion of deep features with CNN extracted features', *International Journal of Intelligent Systems Technologies and Applications*, Vol. 20, No. 6, pp.510–523.
- Zhou, T. (2024) 'Multi-modal brain tumor segmentation via disentangled representation learning and region-aware contrastive learning', *Pattern Recognition*, Vol. 149, No. 6, p.110282.