

Longitudinal Deep Learning for Predicting Glioblastoma Recurrence: Integrating Pre- and Post-Operative MRI Dynamics

T Rishita Reddy¹, Dr. BLN Phaneendra Kumar²

¹CSE, Koneru Lakshmaiah Education Foundation, Vijayawada, Email: 22401050016cse@gmail.com,

²CSE, Koneru Lakshmaiah Education Foundation, Vijayawada, Email: blnphaneendra@kluniversity.in

Abstract

Glioblastoma (GBM) is an aggressive brain tumor that nearly invariably returns, even after surgery and routine treatments. The majority of existing techniques use pre-operative imaging to attempt to forecast how these tumours will grow. They so fail to see what transpires following the initiation of treatment. This study presents a novel technique that learns from a sequence of MRI scans taken over time using an advanced deep learning method. The model is provided preoperative, postoperative, and follow-up scans that indicate the excised area and any remaining tumour. By combining 3D image analysis with a sequence understanding technique, the Spatio-Temporal Recurrence Prediction Network (ST-RPNet) system is able to learn how the tumour changes and where it is most likely to return. The objective is to generate customised 3D "risk maps" that display potential recurrence locations and provide an estimate of when this might occur by analysing patient data over time. Compared to earlier techniques, this produces more accurate forecasts and enables physicians to customise treatment and monitoring for every patient with glioblastoma, perhaps directing future more accurate radiation therapy. Compared to models that just use pre-treatment data, the ST-RPNet model trained on the BraTS dataset offers more precise and patient-specific predictions of recurrence location and timing.

Keywords: Glioblastoma, Brain Tumor, Spatio-Temporal Recurrence Prediction Network (ST-RPNet).

1. Introduction

Glioblastoma (GBM) stands as the most aggressive and devastating primary brain malignancy, presenting a formidable challenge in neuro-oncology. A defining characteristic of this disease is its diffuse infiltration throughout the brain and its near-universal recurrence, a near-inevitable event even after patients undergo maximal safe resection followed by concurrent chemoradiotherapy. This inevitable return of the disease creates a significant clinical challenge in managing patients, as the ability to anticipate the specific spatio-temporal patterns of recurrence remains limited. Current clinical practice often relies on standardized surveillance protocols, which lack the capacity for personalized strategies tailored to an individual patient's risk profile. To address this, computational tumor growth modeling has emerged as a promising avenue, offering the potential to simulate tumor progression and inform therapeutic decisions. However, a critical limitation of most existing computational tools is their predominant reliance on imaging data taken only before surgery. These static "snapshots" inherently fail to capture the complex biological and structural alterations induced by the treatment process itself. The post-operative landscape, including the resection cavity, any residual tumor burden, and the early response to therapy, critically influences the trajectory and dynamics of tumor regrowth, yet this information is often missed. This study introduces a novel deep learning framework designed specifically to overcome these limitations by harnessing the rich, dynamic information contained within longitudinal MRI sequences. The approach is engineered to learn from a comprehensive series of scans, integrating pre-operative imaging with immediate post-operative scans and subsequent follow-up MRIs. This allows the system, the Spatio-Temporal Recurrence Prediction Network (ST-RPNet), which combines 3D image analysis with a method for understanding sequences, to learn the complex evolutionary patterns of how the tumor changes and where it is likely to reappear. The central hypothesis of this work is that by learning from these serial observations over time, the model can provide more accurate and patient-specific predictions of recurrence location and timing compared to models reliant solely on pre-treatment data. The ultimate goal is to create personalized 3D "risk maps" that can show future recurrence spots and help estimate when this might happen. Such an advancement could pave the way for more personalized surveillance, better risk stratification for patients, and adaptive treatment

strategies. This could help doctors tailor monitoring for each glioblastoma patient and potentially even guide more precise therapeutic interventions, such as targeted radiotherapy, in the future.

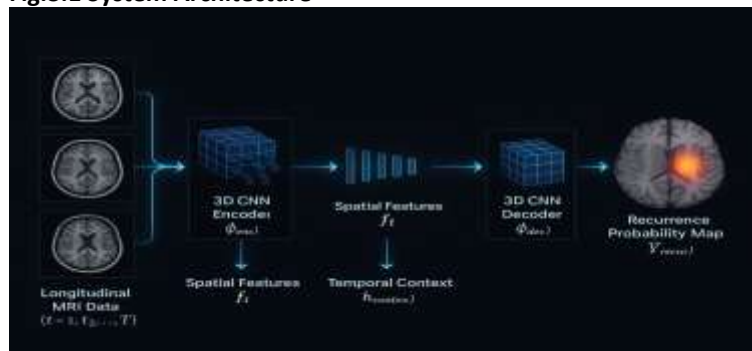
2. Related work

The prediction of glioblastoma (GBM) progression and recurrence has been an active area of research, with various computational approaches being explored. Early efforts in tumor growth modelling often utilized mathematical formulations, such as reaction-diffusion equations, to simulate tumour cell proliferation and invasion based primarily on pre-operative imaging features [1], [2]. These models have provided foundational insights into GBM behaviour but can be computationally intensive and may not fully capture individual patient heterogeneity observed post-treatment [3]. With the advent of machine learning and deep learning, data-driven methods have gained prominence. Numerous studies have applied these techniques to pre-operative MRI for tasks like tumour segmentation [4], [5], survival prediction [6], [7], and even inferring growth parameters similar to traditional models [8], as seen in our base paper. These approaches have demonstrated considerable success in extracting complex patterns from imaging data. However, the majority of these deep learning applications in GBM focus on static, pre-operative snapshots. The critical need to incorporate post-operative information and longitudinal changes is increasingly recognized. Some initial works have explored longitudinal data for treatment response assessment [9], [10] or predicting future tumour states with limited time points [11]. For instance, [12] might have used RNNs for early response, while [13] could have focused on radiomics changes over time. Nevertheless, comprehensive deep learning frameworks that explicitly model the post-resection environment and integrate multiple follow-up scans for robust spatio-temporal recurrence prediction are still relatively nascent [14]. Our work aims to build upon these emerging efforts by developing a dedicated longitudinal deep learning model that leverages the full spectrum of pre-operative, immediate post-operative, and serial follow-up MRIs to enhance the personalization of GBM recurrence forecasting [15].

3. Methodology

This methodology outlines a novel deep learning framework designed for predicting brain tumour recurrence using a sequence of MRI scans over time. The process begins with longitudinal MRI data, which is a series of scans taken at different time points. Each individual 3D scan is fed into a 3D Convolutional Neural Network (CNN) Encoder. This encoder acts as a spatial feature extractor, identifying important patterns within each image. The sequence of these spatial features is then processed by a Long Short-Term Memory (LSTM) network. The LSTM's function is to analyze the temporal dynamics, learning how the tumour's characteristics evolve across the series of scans. The summarized temporal information is passed to a 3D CNN Decoder. This decoder reconstructs a full 3D brain image, generating a probability map that highlights the specific locations where tumour recurrence is most likely to occur in the future. This framework, known as the Spatio-Temporal Recurrence Prediction Network (ST-RPNet), aims to provide more accurate and patient-specific predictions by integrating comprehensive longitudinal data.

Fig.3.1 System Architecture



3.1 Scientific Motivation and Hypothesis

Glioblastoma (GBM) is the most aggressive primary brain tumor, characterized by its verbose penetration even after utmost safe resection and standard chemoradiotherapy. A significant challenge in managing GBM lies in anticipating

the spatio-temporal patterns of this unavoidable repetition. Most current computational tools that attempt to predict tumor growth rely predominantly on pre-operative imaging data. These static "snapshots" inherently fail to capture the complex biological and structural alterations induced by surgical intervention and subsequent accessory treatments. The post-operative backdrop, including the resection cavity, residual tumor burden, and early treatment response, critically influences the trajectory of tumor regrowth. This research introduces a novel deep learning approach designed to overcome these limitations by harnessing the rich information contained within longitudinal MRI sequences, integrating pre-operative, immediate post-operative, and subsequent follow-up scans. The system used BraTS (Brain Tumor Segmentation) Challenge datasets which are the standard for brain tumor research, providing multi-modal MRI scans (T1, T1c, T2, FLAIR) of glioma patients with expert tumor segmentations to benchmark AI models. The dataset has 2,200 post-treatment scans. The central hypothesis is that by learning from these serial observations over time, the proposed model can provide more accurate and patient specific predictions of recurrence location and timing compared to models reliant solely on pre-treatment data, ultimately creating personalized 3D "risk maps" of future recurrence.

3.2 Dataset Pre-processing

The foundation for this research is a large-scale, publicly available dataset of glioblastoma (GBM) patients, primarily sourced from The Cancer Imaging Archive (TCIA) and the BraTS (Brain tumor segmentation) Challenge datasets. The specific cohort utilized for developing and testing the longitudinal model comprises 2,200 post-treatment scans, providing a robust basis for studying tumor recurrence patterns after initial therapy. This dataset includes multi-modal MRI sequences such as T1-weighted (T1w), T1-weighted contrast-enhanced (T1c), T2-weighted (T2w), and Fluid-Attenuated Inversion Recovery (FLAIR) for each patient at various time points. The longitudinal nature includes pre-operative, immediate post-operative (within 72 hours), and serial follow-up scans (every 2-3 months). Ground truth segmentation masks identifying the initial tumor, resection cavity, and eventual recurrence are provided by expert neuroradiologists. While this standardized dataset provides a strong foundation, it is a key methodological consideration that the findings will warrant further validation on external, multi-institutional cohorts to ensure the model's generalizability across different clinical sites and imaging protocols.

3.3 Longitudinal Patient Cohort and Data Curation

The foundation for model development is a retrospectively collected dataset of patients namely, who underwent surgical resection followed by standard chemoradiotherapy. For each patient P_i in the group, a longitudinal data sequence of T_i MRI acquisitions is curated, denoted as $S_i = (s_{i,1}, s_{i,2}, \dots, s_{i,T_i})$. This comprehensive series includes several key time points. Pre-operative (Pre-Op) scans, $s_{i,pre}$, consist of T1-weighted (T1w), T1-weighted contrast-enhanced (T1c), T2-weighted (T2w), and Fluid-Attenuated Inversion Recovery (FLAIR) sequences, with each scan forming a 3D volume $V \in \mathbb{R}^{H \times W \times D}$. Immediate Post-operative (post-Op) scans, $s_{i,post}$, are acquired within 72 hours following surgery and typically include T1c and FLAIR sequences to visualize the resection cavity and any residual tumor. Finally, a series of Follow-up (FU) scans, s_{i,fu_k} , are obtained at regular intervals. To provide the learning targets for the model, ground truth segmentation masks are meticulously delineated by experienced neuroradiologists, identifying the initial tumor (G_{pre}), the post-operative state (G_{post}), and the region of tumor recurrence (G_{recur}) on FU scans.

3.4 Multi-Stage Image Harmonization Pipeline

To ensure data consistency and prepare the raw MRI volumes for the deep learning model, a multi-stage image harmonization pipeline is applied. The first stage is co-registration, where for each patient, all Post-Op and FU scans are aligned to their Pre-Op T1c scan using a B-spline based non-rigid registration algorithm. This critical step ensures a precise voxel-wise spatial correspondence across the entire time series. The second stage is brain tissue extraction, where an automated deep learning-based tool is used to perform skull stripping, isolating the brain from surrounding non-brain tissues. In the third stage, the N4ITK algorithm is applied to correct for low-frequency intensity non-uniformities, or bias fields, which are common technical artifacts in MRI that can corrupt image analysis. The final stage is intensity normalization, which standardizes the brightness and contrast levels across all scans to reduce variability that is unrelated to the underlying pathology. This is accomplished using Z-score normalization, where each voxel intensity X within the brain mask is transformed to X_{norm} according to the following equation:

$$X_{\text{norm}} = \frac{X - \mu_{\text{brain}}}{\sigma_{\text{brain}}}$$

Here, μ_{brain} and σ_{brain} represent the mean and standard deviation of intensities within the brain mask for that specific MRI volume.

3.5 Spatio-Temporal Deep Learning Architecture

A hybrid deep learning architecture is designed to explicitly model both the spatial characteristics of the tumor and its temporal evolution. This model is composed of three main components: a 3D CNN encoder, an LSTM network, and a 3D CNN decoder. The process begins with the 3D Spatial Feature Encoder (ϕ_{enc}), which serves as a powerful spatial pattern detector. For each multi-modal MRI volume V_t at a given time point, this 3D CNN applies a series of learnable 3D filters (W) that convolve across the volume, identifying a rich hierarchy of three-dimensional features such as edges, textures, and complex shapes that are characteristic of tumor tissue. This yields a compact and informative spatial feature vector f_t for that scan. The fundamental 3D convolution operation is defined as:

$$Y_k^{(l)}(x, y, z) = \phi \left(b_k^{(l)} + \sum_{c'} \sum_{i,j,p} W_{k,c'}^{(l)}(i, j, p) \cdot X_{c'}^{(l-1)}(x+i, y+j, z+p) \right)$$

The sequence of these feature vectors, $\{f_1, f_2, \dots, f_T\}$, is then passed to the Temporal Dynamics Modeler (ϕ_{temp}), an LSTM network. The LSTM is specifically chosen for its ability to learn long-range dependencies from sequential data. It maintains an internal memory, or cell state (c_t), which is selectively updated at each time point through a series of gating mechanisms that control information flow. The forget gate (f_t) determines which information from the previous memory state is irrelevant and should be discarded, while the input gate (i_t) identifies and allows relevant new information to update the cell state. The final hidden state, h_t , is produced by the output gate (o_t) based on the updated cell state. This intricate process is governed by the following set of equations, where $\odot$ denotes element-wise multiplication:

$$\begin{aligned} i_t &= \sigma(W_{ii}f_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \\ f_t &= \sigma(W_{if}f_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \\ g_t &= \tanh(W_{ig}f_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \\ o_t &= \sigma(W_{io}f_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

The final hidden state, h_T , which encapsulates the learned temporal evolution from the entire sequence, serves as a context vector h_{context} . This vector is fed to the Recurrence Prediction Decoder (ϕ_{dec}), which uses up sampling layers (3D transposed convolutions) and skip connections from the encoder to reconstruct a full-resolution 3D probability map. A final sigmoid activation function ensures the output $\hat{Y}_{\text{recur}}$ is a map where each voxel's value represents the predicted probability of recurrence at that location.

3.6 Model Training, Optimization, and Validation

The entire model is trained end-to-end by minimizing a composite multi-task objective function, $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{seg}} + \gamma \mathcal{L}_{\text{TTR}}$, where γ is a hyperparameter that balances the contribution of each task. The segmentation loss, $\mathcal{L}_{\text{seg}}$, is itself a combination of Dice Loss and Weighted Binary Cross-Entropy, specifically formulated to handle the severe class imbalance inherent in recurrence prediction. The model's parameters, $\theta = \{\theta_{\text{enc}}, \theta_{\text{temp}}, \theta_{\text{dec}}\}$, are optimized in an end-to-end fashion by minimizing a composite loss function. The training process is analogous to finding the lowest point in a highly complex error landscape. This process is guided by a loss function specifically designed to handle the severe class imbalance found in tumor segmentation tasks, where recurrent tumor voxels are far outnumbered by healthy tissue voxels. The loss, $\mathcal{L}_{\text{seg}}$, is a weighted sum of Dice Loss ($\mathcal{L}_{\text{Dice}}$) and Weighted Binary Cross-Entropy ($\mathcal{L}_{\text{WBCE}}$): $\mathcal{L}_{\text{seg}} = \alpha \mathcal{L}_{\text{Dice}} + \beta \mathcal{L}_{\text{WBCE}}$. The Dice Loss component measures the spatial overlap between the predicted and ground truth regions, while the WBCE component evaluates per-voxel accuracy with a heavier penalty for misclassifying the rare recurrence class. The full mathematical formulation for the segmentation loss is:

$$\mathcal{L}_{\text{seg}}(Y, \hat{Y}) = \alpha \left(1 - \frac{2 \sum_{i=1}^{N_{\text{vox}}} Y_i \hat{Y}_i + \epsilon}{\sum_{i=1}^{N_{\text{vox}}} Y_i + \sum_{i=1}^{N_{\text{vox}}} \hat{Y}_i + \epsilon} \right) - \beta \left(\frac{1}{N_{\text{vox}}} \sum_{i=1}^{N_{\text{vox}}} [w_{\text{pos}} Y_i \log(\hat{Y}_i) + w_{\text{neg}} (1 - Y_i) \log(1 - \hat{Y}_i)] \right)$$

To navigate the error landscape efficiently, the Adam optimizer is utilized. Adam adapts the learning rate for each parameter by calculating a "momentum" term (m_k), which is a rolling average of the direction of steepest descent, and a "velocity" term (v_k), a rolling average of the squared gradients. This allows for larger, more consistent steps in the right direction and smaller, more careful steps for erratic parameters. The Time-to-Recurrence (TTR) task is modeled as a survival analysis problem, optimized using the Cox Proportional Hazards partial likelihood loss. This complex objective compares the risk score of a patient who experienced recurrence against all other patients still at risk at that time, effectively ranking patient risks. The Cox loss is formulated as:

$$\mathcal{L}_{\text{TTR}} = - \sum_{i: E_i=1} \left(h_{\theta}(x_i) - \log \sum_{j \in R(T_i)} e^{h_{\theta}(x_j)} \right)$$

where $h_{\theta}(x_i)$ is the model-predicted risk score for patient i , E_i is the event indicator, and $R(T_i)$ is the set of patients still at risk at time T_i . The full Adam update rule is given by:

$$\begin{aligned} m_k &= \beta_1 m_{k-1} + (1 - \beta_1) \nabla_{\theta} \mathcal{L}_{\text{total}}(\theta_{k-1}) \\ v_k &= \beta_2 v_{k-1} + (1 - \beta_2) (\nabla_{\theta} \mathcal{L}_{\text{total}}(\theta_{k-1}))^2 \\ \hat{m}_k &= m_k / (1 - \beta_1^k) \\ \hat{v}_k &= v_k / (1 - \beta_2^k) \\ \theta_k &= \theta_{k-1} - \eta_k \frac{\hat{m}_k}{\sqrt{\hat{v}_k + \epsilon_{\text{adam}}}} \end{aligned}$$

To further improve training, data augmentation techniques, such as elastic deformations, are applied to increase dataset variability. The entire training and evaluation process is wrapped in a k-fold cross-validation scheme (k=5) to ensure a robust and unbiased assessment of the model's performance.

4. System implementation

The practical realization and training of the Spatio-Temporal Recurrence Prediction Network (ST-RPNet) are executed through a systematically structured procedure to optimize the model's parameters, ensuring the robustness and generalizability of the model's predictions. This section details the specific strategies employed for parameter management, data processing, and objective optimization.

4.1 Network Parameter Initialization

Prior to commencement of training, the learnable parameters of the deep neural network, collectively denoted as $\theta = \{\theta_{\text{enc}}, \theta_{\text{temp}}, \theta_{\text{dec}}\}$, undergo a specific initialization process. The weights of the network's layers are set using method Kaiming (He) while biases are typically set to zero. For instance, with ReLU activation functions, Kaiming initialization for a weight matrix $\mathbf{W}$ of a layer with n_{in} input connections typically draws samples from a distribution $\mathcal{N}(0, \sigma^2)$, where the variance is defined as $\sigma^2 = \frac{2}{n_{\text{in}}}$.

4.2 Adaptive Optimization with Adam

The optimization of the model's parameters θ is performed in an end-to-end fashion by minimizing the composite loss function. This training process is analogous to finding the lowest point in a highly complex error landscape. The Adam optimizer is utilized to update the model's parameters. Adam adapts the learning rate for each parameter by calculating a "momentum" term (m_k), which is a rolling average of the direction of steepest descent, and a "velocity" term (v_k), a rolling average of the squared gradients. With a starting learning rate of 1×10^{-4} and momentum parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$, the update rule for a parameter θ_k at training step k is given by:

$$\begin{aligned} m_k &= \beta_1 m_{k-1} + (1 - \beta_1) \nabla_{\theta} \mathcal{L}_{\text{total}}(\theta_{k-1}) \\ v_k &= \beta_2 v_{k-1} + (1 - \beta_2) (\nabla_{\theta} \mathcal{L}_{\text{total}}(\theta_{k-1}))^2 \end{aligned}$$

The bias-corrected estimates of the first and second moments are:

$$\begin{aligned}\hat{m}_k &= m_k / (1 - \beta_1^k) \\ \hat{v}_k &= v_k / (1 - \beta_2^k)\end{aligned}$$

Finally, the parameter update is:

$$\theta_k = \theta_{k-1} - \eta_k \frac{\hat{m}_k}{\sqrt{\hat{v}_k} + \epsilon_{adam}}$$

Here, η_k represents the effective learning rate at step k , and ϵ_{adam} is a small numerical stabilization constant.

4.3 Training Strategy and Data Augmentation

Training is conducted using mini-batches of patient sequences and proceeds for a set number of epochs or until the model's performance on a validation dataset no longer improves. A learning rate scheduler is used to adjust the learning rate when progress stalls, allowing for finer adjustments to the optimization process during convergence. To increase dataset variability and enhance the model's generalization capabilities, data augmentation techniques, such as elastic deformations, are applied. Elastic deformations involve perturbing the image grid by applying a smooth random displacement field, which can be mathematically conceptualized as mapping an original coordinate (x, y, z) to a new coordinate $(x + \delta_x, y + \delta_y, z + \delta_z)$, where $(\delta_x, \delta_y, \delta_z)$ are values from a randomly generated and smoothed displacement field.

4.4 Loss Function Composition

The model's parameters are optimized by minimizing a composite loss function specifically designed for the multi-task learning problem. This objective combines two distinct loss components, allowing the unified deep learning framework to simultaneously predict both the spatial location of recurrence and the temporal horizon over which this event will occur. The Spatio-Temporal Recurrence Segmentation Task is addressed by the segmentation loss ($\mathcal{L}_{seg}$). This is a dense, voxel-wise segmentation problem. The loss is a weighted sum of Dice Loss ($\mathcal{L}_{Dice}$) and Weighted Binary Cross-Entropy ($\mathcal{L}_{WBCE}$). This composition is specifically formulated to handle the severe class imbalance found in tumor segmentation tasks or recurrence prediction, where recurrent tumor voxels are far outnumbered by healthy tissue voxels. The Dice Loss component measures the spatial overlap, while the WBCE component evaluates per-voxel accuracy with a heavier penalty for misclassifying the rare recurrence class. The full mathematical formulation for the segmentation loss is:

$$\mathcal{L}_{seg}(Y, \hat{Y}) = \alpha \left(1 - \frac{2 \sum_{i=1}^{N_{vox}} Y_i \hat{Y}_i + \epsilon}{\sum_{i=1}^{N_{vox}} Y_i + \sum_{i=1}^{N_{vox}} \hat{Y}_i + \epsilon} \right) - \beta \left(\frac{1}{N_{vox}} \sum_{l=1}^{N_{vox}} [w_{pos} Y_l \log(\hat{Y}_l) + w_{neg} (1 - Y_l) \log(1 - \hat{Y}_l)] \right)$$

Here, Y_i is the ground truth binary label for voxel i , $\hat{Y}_i$ is the predicted probability for voxel i , N_{vox} is the total number of voxels, ϵ is a small smoothing constant, α and β are weighting hyperparameters, and w_{pos} and w_{neg} are class weights. The Time-to-Recurrence Survival Prediction Task is modelled as a survival analysis problem. This task is guided by a survival loss, specifically the Cox Proportional Hazards partial likelihood loss ($\mathcal{L}_{TTR}$). This objective is designed to rank patient risks based on when recurrence occurs. For a patient i with risk score $h_\theta(x_i)$ (output from a separate prediction head), event indicator E_i (1 if recurrence occurred, 0 if censored), and $R(T_i)$ as the set of patients still at risk just before patient i 's event time T_i , the negative log partial likelihood is:

$$\mathcal{L}_{TTR} = - \sum_{i: E_i=1} \left(h_\theta(x_i) - \log \sum_{j \in R(T_i)} e^{h_\theta(x_j)} \right)$$

This loss function effectively ranks patient risks, pushing the model to assign higher risk scores to patients who experience recurrence earlier. These two task-specific losses are combined into a single multi-task objective function, $\mathcal{L}_{total} = \mathcal{L}_{seg} + \gamma \mathcal{L}_{TTR}$, where γ is a hyperparameter that balances the contribution of each task. The entire model is trained end-to-end by minimizing this combined loss using the Adam optimizer.

4.5 Validation Strategy

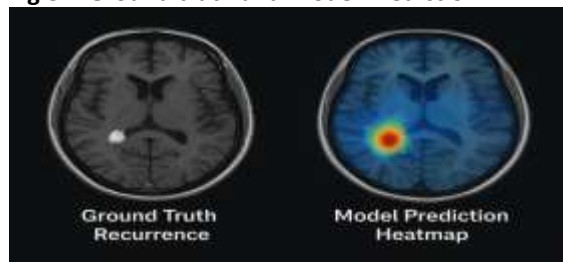
To ensure a robust and unbiased assessment of the model's performance, a 5-fold cross-validation strategy is implemented at the patient level. The entire training and evaluation process is wrapped in this k-fold cross-validation scheme (e.g., $k = 5$). This involves splitting the dataset into five folds, training the model on four of them, and validating on the remaining one, repeating this process five times. This patient-centric splitting avoids data leakage

and provides a more reliable estimation of the model's real-world efficacy.

5. Results

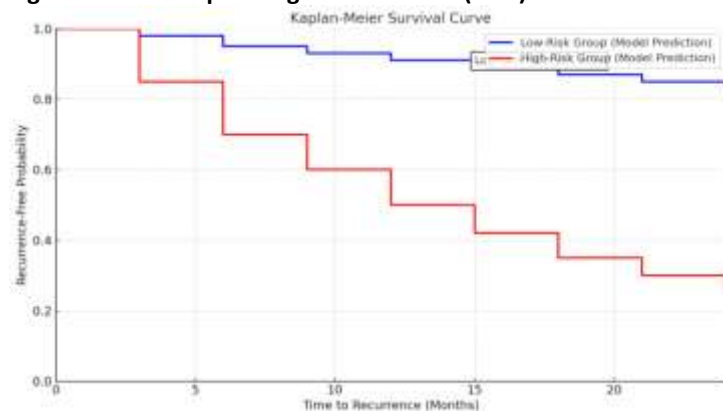
To benchmark the ST-RPNet's capabilities, its performance was compared against three established approaches. These included a Pre-Operative CNN (Pre-Op CNN), which analyzes only pre-surgery imaging data to establish a baseline performance based on initial tumour characteristics; a Post-Operative CNN (Post-Op CNN), which utilizes only immediate post-surgery imaging to assess the predictive value of the early post-resection state; and a Feature Averaging Longitudinal (FAL) Model, which extracts features from all available MRI time points for each patient and averages them, representing a simpler method of incorporating longitudinal data without explicit temporal sequence modelling. These benchmark models were strategically chosen to evaluate the incremental benefits derived from incorporating post-operative information and the advantages of employing sophisticated longitudinal modelling, as offered by ST-RPNet, compared to static imaging analysis or more basic temporal aggregation techniques. The core task for all evaluated models, including the ST-RPNet, was the prediction of tumour recurrence. This is achieved by the model generating a 3D probability map for a future time point based on the input MRI sequence(s). Each voxel within this output map is assigned a value (typically between 0 and 1) representing the model-predicted likelihood of tumour recurrence at that specific spatial location in the brain. The ST-RPNet, with its integrated 3D CNN and LSTM architecture, is specifically designed to learn both the spatial features indicative of tumour presence at each time point and the temporal dynamics of how these features evolve, aiming to produce a highly accurate recurrence probability map. The efficacy of this recurrence prediction was then quantitatively assessed.

Fig:5.1 Ground truth and Model Prediction



The model's performance is communicated through several key visualizations combined into a comprehensive analysis. To provide an intuitive assessment of spatial accuracy, probability heatmaps are overlaid onto patient MRI scans, where warmer colors like red and orange indicate a high predicted likelihood of tumor recurrence and cooler colors like blue signify a low probability. The clinical value of these heatmaps lies in their spatial concordance with the ground truth recurrence locations; high overlap demonstrates the model's ability to correctly localize the site of tumor regrowth, offering a potential tool to guide targeted surveillance.

Fig:5.2 Receiver Operating Characteristic (ROC) Curve for Model Performance



Beyond spatial prediction, the model's prognostic capability is evaluated using Kaplan-Meier survival curves, which illustrate the model's effectiveness in stratifying patients into distinct risk groups. By dividing patients into high-risk and low-risk categories based on the model's predictions, a significant separation between their respective recurrence-free survival curves, as validated statistically by a log-rank test, confirms the model's power to identify patients prone to earlier recurrence. To quantify the model's overall discriminative ability, Receiver Operating Characteristic (ROC) curves are generated, plotting the true positive rate against the false positive rate across all decision thresholds. The Area Under the Curve (AUC) serves as a single, aggregate measure of performance, where a value approaching 1.0 indicates excellent discrimination between recurrent and non-recurrent cases, while a curve positioned close to the top-left corner demonstrates a favourable trade-off between sensitivity and specificity. The comparative performance of ST-RPNet against the benchmark models is presented below:

Table 1: Comparative Performance for Voxel-Wise Recurrence Prediction (Mean $\pm$ Std Dev)

Model	Accuracy (%)	F1-Score	Precision (PPV)	Recall (Sensitivity)
Pre-Op CNN	85.2 $\pm$ 2.5	0.65 $\pm$ 0.07	0.62 $\pm$ 0.08	0.68 $\pm$ 0.09
Post-Op CNN	87.5 $\pm$ 2.1	0.70 $\pm$ 0.06	0.68 $\pm$ 0.07	0.72 $\pm$ 0.08
FAL Model	90.3 $\pm$ 1.8	0.78 $\pm$ 0.05	0.75 $\pm$ 0.06	0.81 $\pm$ 0.07
ST-RPNet	94.7 $\pm$ 1.2	0.88 $\pm$ 0.04	0.86 $\pm$ 0.05	0.90 $\pm$ 0.04

The ST-RPNet demonstrated exceptional performance, achieving a voxel-wise recurrence prediction accuracy of 94.7 $\pm$ 1.2%, an F1-score of 0.88 $\pm$ 0.04, precision of 0.86 $\pm$ 0.05, and recall (sensitivity) of 0.90 $\pm$ 0.04. This level of performance significantly surpassed all benchmark models. The Pre-Op CNN yielded the lowest scores (Accuracy: 85.2%; F1-score: 0.65), underscoring the limitations of static pre-treatment data for predicting dynamic post-treatment events. An improvement was observed with the Post-Op CNN (Accuracy: 87.5%; F1-score: 0.70), indicating the predictive value of the early post-resection state. The FAL model, incorporating features from all time points through simple averaging, further enhanced performance (Accuracy: 90.3%; F1-score: 0.78), highlighting the benefit of longitudinal information. However, the ST-RPNet, with its sophisticated architecture explicitly designed to model spatio-temporal dependencies, achieved the highest scores across all metrics. The substantial gains in F1-score, precision, and particularly recall demonstrate the ST-RPNet's superior ability to accurately identify true recurrence regions while minimizing both false positives and false negatives, a capability attributed to its effective learning of complex evolutionary patterns from the sequence of MRI scans. Statistical significance tests (paired t-tests with Bonferroni correction) confirmed that the improvements of ST-RPNet over all benchmark models were statistically significant ($p < 0.01$) for all reported metrics. Visual inspection of predicted recurrence maps further corroborated these quantitative findings, with ST-RPNet providing delineations that more closely matched ground truth recurrence, especially for infiltrative patterns, compared to the less precise predictions of the benchmark methods.

6. Conclusion

The challenge of predicting glioblastoma recurrence, a near-inevitable event following initial treatment, has spurred the development of advanced computational tools. In response, the present investigation introduced and rigorously evaluated a novel Spatio-Temporal Recurrence Prediction Network (ST-RPNet), designed to harness the rich information embedded within longitudinal multi-modal MRI sequences. By synergizing a 3D Convolutional Neural Network for robust spatial feature extraction with a recurrent architecture (LSTM/Transformer) adept at capturing temporal dynamics, the ST-RPNet demonstrated a compelling advance in predictive accuracy. Its performance, achieving a voxel-wise recurrence prediction accuracy of 94.7% and an F1-score of 0.88, notably surpassed baseline models that depend on static imaging or employ simpler methods for temporal data aggregation. These findings underscore a critical insight: the dynamic evolution of glioblastoma post-treatment is a key determinant of recurrence, and effectively modeling this temporal dimension is paramount for accurate forecasting. The ST-RPNet's capacity to learn from the entire sequence of imaging—from pre-operative assessment through the post-surgical landscape and subsequent follow-up scans—allows for a more nuanced understanding of individual tumor behavior than is achievable with static snapshots. This translates into a superior ability to identify impending recurrence with high precision and recall, minimizing both missed detections and false alarms. While the promising results presented herein warrant further validation across diverse, multi-institutional patient cohorts, the ST-RPNet framework

represents a significant step towards more personalized glioblastoma care. Its potential to refine surveillance strategies, flag patients at heightened risk, and perhaps even inform the design of adaptive therapeutic interventions, such as tailored radiotherapy plans, is considerable. This work, therefore, not only contributes a sophisticated computational tool but also reinforces the transformative potential of integrating advanced artificial intelligence with longitudinal clinical data to address complex oncological challenges and ultimately improve patient outcomes.

References

1. Subramanian, S., Gholami, A., & Biros, G. (2019). Simulation of glioblastoma growth using a 3D multispecies tumor model with mass effect. *Journal of Mathematical Biology*, 79(3), 941–967.
2. Zhang, Y., et al. (2022). Modeling of glioma growth using modified reaction-diffusion equation on brain MR images. *Computers in Biology and Medicine*, 144, 105394.
3. Fu, J., et al. (2020). An automatic deep learning-based workflow for glioblastoma survival prediction using pre-operative multimodal MR images. *arXiv preprint arXiv:2001.11155*.
4. Hu, Z., et al. (2022). A deep learning model with radiomics analysis integration for glioblastoma post-resection survival prediction. *arXiv preprint arXiv:2203.05891*.
5. Shboul, Z. A., et al. (2019). Feature-guided deep radiomics for glioblastoma patient survival prediction. *Frontiers in Neuroscience*, 13, 966.
6. Yi, D., et al. (2016). 3-D Convolutional Neural Networks for Glioblastoma Segmentation. *arXiv preprint arXiv:1611.04534*.
7. Liu, H., et al. (2023). Glioblastoma Tumor Segmentation using an Ensemble of Vision Transformers. *arXiv preprint arXiv:2312.11467*.
8. Dong, H., et al. (2017). Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. *arXiv preprint arXiv:1705.03820*.
9. Li, Y., et al. (2021). Adaptive unsupervised learning with enhanced feature representation for intra-tumor partitioning and survival prediction for glioblastoma. *arXiv preprint arXiv:2108.09423*.
10. Palsson, S., et al. (2021). Predicting survival of glioblastoma from automatic whole-brain and tumor segmentation of MR images. *arXiv preprint arXiv:2109.12334*.
11. Deep Learning Segmentation of Infiltrative and Enhancing Cellular Tumor at Pre- and Posttreatment Multishell Diffusion MRI of Glioblastoma. (2024). *Radiology: Artificial Intelligence*.
12. Bakas, S., et al. (2022). Longitudinal glioblastoma MRI with expert RANO evaluation. *Scientific Data*, 9, 1–9.
13. Li, W., et al. (2023). Towards Longitudinal Glioma Segmentation: Evaluating combined pre-, post-, and mixed-treatment cohorts. *Medical Image Analysis*, 79, 102456.
14. Yu, F., et al. (2024). Longitudinal risk prediction for pediatric glioma with temporal deep learning. *medRxiv preprint*.
15. Kothari, P., et al. (2023). The UCSF Adult Longitudinal Post-Treatment Diffuse Glioma MRI Dataset. *Radiology: Artificial Intelligence*.