Sebastian-Eugen Buzdugan

A Multimodality Multimodal Deep Learning Framework on MRI Imaging and Genomics to Assess Brain Cancer Survival

Final Master's Project

Directed by Prof. Dr. Domènec Puig, Dr. Moona Mazher

University Master's Degree in Computer Security Engineering and Artificial Intelligence

UNIVERSITAT ROVIRA i VIRGILI

Tarragona
2024

Acknowledgements

I would like to express my sincere appreciation to Prof. Dr. Domènec Puig and Dr. Moona Mazher for their outstanding guidance, continuous encouragement, and insightful input over the duration of my master thesis. Their expertise and assistance were vital in achieving the success of my thesis.

Furthermore, I would like to express my deepest gratitude to Universitat Rovira i Virgili (URV) for providing me with the necessary resources and workspace to carry out my research. I will forever value the knowledge and experiences I have acquired at this place.

Personally, I would want to express my deep thanks to my mother Nicoleta, brothers, and girlfriend Codruta for their constant support, which have been crucial in helping me reach this milestone.

Abstract

Glioblastoma (GBM) is an instance of brain tumor defined by a complex genetic makeup that has a significant impact on the prognosis of patients. Notable predictive molecular alterations in GBM include of IDH1 mutations and MGMT promoter methylation. 12.9% of GBM patients have IDH1 mutations, which are linked to improved overall survival. The median survival time for individuals with IDH1 mutations is 31 months, while it is 15 months for patients with IDH wild-type tumors. The presence of MGMT promoter methylation, observed in 43% of patients with glioblastoma (GBM), is associated with a better prognosis. Patients with methylation had a median survival period of 504 days, compared to 329 days for those without methylation. These indicators function as independent prognostic variables. This study integrates imaging features, clinical profiles, and genetic markers using advanced machine learning models, including RandomForest, XGBoost, LightGBM, and a custom-designed Dense Neural Network (Dense NN), to enhance survival prediction in GBM patients. MRI imaging data were processed using MRIPreprocessor and the radiomics Python library to extract radiomics features. The models were evaluated using datasets from UPENN-GBM (602 patients) and UCSF-PDGM (414 patients). The custom Dense NN demonstrated superior performance, achieving a concordance index (CI) of 0.86 on the UPENN-GBM dataset and 0.88 on the UCSF-PDGM dataset, surpassing traditional tree-based methods. This complete method provides an accurate evaluation of tumor biology, allowing the creation of customized treatment strategies. The results highlight the important significance of AI technologies in predicting disease progressions and directing treatment interventions, showcasing the potential of radiogenomics to improve precision medicine for GBM patients. This master thesis establishes a solid basis for future research on predicting survival in glioblastoma, demonstrating the efficacy of integrating clinical, genetic, and imaging data to improve prognostic models.

Contents

List of Figures

List of Tables

Glossary

- **AI:** Artificial Intelligence
- **C-index:** Concordance Index
- **CNN:** Convolutional Neural Network
- **DICOM:** Digital Imaging and Communications in Medicine
- **DNN:** Deep Neural Network
- **GBM:** Glioblastoma Multiforme
- **IDH1:** Isocitrate Dehydrogenase 1
- **LightGBM:** Light Gradient Boosting Machine
- **MAE:** Mean Absolute Error
- **MGMT:** O-6-Methylguanine-DNA Methyltransferase
- **ML:** Machine Learning
- **MSE:** Mean Squared Error
- **NIFTI:** Neuroimaging Informatics Technology Initiative
- **NN:** Neural Network
- **ROC-AUC:** Receiver Operating Characteristic - Area Under the Curve
- **RFE:** Recursive Feature Elimination
- **SHAP:** SHapley Additive exPlanations
- **XGBoost:** eXtreme Gradient Boosting

1. Introduction

1.1 Motivation

At a tragic 5% 5-year survival rate [1], glioblastoma (GBM) continues to be one of the most aggressive and toughest primary brain tumors to treat. Considerable progress has not been made in improving the survival rates of GBM patients, even with advances in oncological therapy [2]. The incapacity to take into account the significant variability within and between tumors is a major factor contributing to treatment failures [3].

This master's thesis aims to address this critical need by exploiting the growing field of radiogenomics, which combines radiomics features extracted from medical imaging with genomic data to provide a more comprehensive understanding of tumor biology and patient prognosis [4].

The objective of this study is to integrate clinical and genetic data with high-throughput radiomics features obtained from multiparametric MRI scans in order to develop more precise and reliable survival prediction models for patients with glioblastoma multiforme (GBM). This study surpasses traditional single-modality research by employing an integrated methodology that highlights the need of combining clinical, genomic, and radiomic data to comprehensively understand the complexity of tumor biology and improve prediction accuracy.

These integrated features are explored using state-of-the-art machine learning algorithms such as RandomForest, XGBoost, LightGBM, and a custom-built Dense Neural Network. The Dense Neural Network stands out for its exceptional performance in survival prediction, which may be attributed to its capacity to identify intricate and non-linear correlations present in the multimodal data.

Radiogenomics offer a unique opportunity to assess tumor biology and heterogeneity non-invasively [5]. The used approach has the potential to decrease the need for invasive biopsies, solve some limitations of current sampling methods, and offer an in-depth view of the whole tumor.Survival prediction models developed in this way can assist in designing custom therapies, allowing clinicians to change their interventions depending on the particular risks associated with each patient. This reflects the increasing emphasis on personalized medicine in oncology.

While previous studies have dealt only with radiomics or genomics separately, this thesis aims at bridging that gap through an integrated radiogenomics framework. The aim of this research is not only to include all relevant modalities but also address its current shortcomings in single-modality analysis to achieve a better understanding of GBM biology, as well as patient outcomes.

Classical machine learning methods combined with a particular feature selection process are used in this study. One of the primary obstacles to the clinical adoption of AI-based solutions is removed by the incorporation of SHAP (SHapley Additive exPlanations) values into the model analysis, which offers a clear interpretation of the model predictions. This enhances the reliability and comprehensibility of the radiogenomic models, hence facilitating their application in the treatment of GBM patients.

This study is focused at leveraging radiogenomics models and creating a path for their application in GBM patient care by looking into model performance, interpretability and validation across our two datasets (UPENN-GBM and UCSF-PDGM).

Therefore, radiogenomics has great promise for addressing the pressing need for improved prognostic tools in GBM management which formed my inspiration towards writing this master's thesis. In this master thesis I aim to make a significant impact in the field of neuro-oncology and move forward personalized treatment strategies for glioblastoma multiforme patients via generating as well as validating an integrated radiogenomic framework which can be applied for survival prediction.

1.2 Objective

The main objectives of this master's thesis are to develop and verify a precise radiogenomics framework for predicting the survival of glioblastoma patients, with particular focus on the following:

- **to develop and evaluate complex multimodal machine learning models** that combine genetic, radiomic, and clinical data to improve survival predictions for glioblastoma patients. The proposed approach involves using a bespoke Dense Neural Network (Dense NN) to effectively capture complex non-linear connections in the high-dimensional datasets. This would result in improved accuracy and interpretability of the model, surpassing the capabilities of conventional techniques.

- **to validate the practicality and strength of the proposed models** using the UPENN-GBM and UCSF-PDGM datasets, using reliable evaluation metrics such as concordance index (CI), R-squared (R^2), mean squared error (MSE), and mean absolute error (MAE).

- **to further improve the clinical significance of the framework**, we will uncover useful predictive features using feature importance analysis and SHAP values. Additionally, we will compare the efficacy of the integrated radiogenomics method with standard models.

The primary objective of this master's thesis is to display the substantial improvement in survival prediction accuracy achieved by an integrated radiogenomics approach, augmented by complex machine learning methods such as the Dense Neural Network. This initiative will enhance the advancement of more powerful clinical decision support systems and lay a solid groundwork for further research in this domain, particularly in the prediction of glioblastoma survival.

1.3 State-of-the-art

This study examines the latest in glioblastoma research by merging radiomics, clinical data and genomics. The goal of this methodology is to improve survival prediction accuracy by making use of the information from each data source available. Radiomics refers to the computation of quantitative features from medical images which can be used for non-invasive tumor heterogeneity characterization and possibly discovery of hidden patterns that are linked to prognosis. Clinical data on the other hand such as patient demographics and treatment history enhances understanding

on how specific aspects influence disease progression. Treatment response and survival outcomes may also be influenced by molecular drivers of these diseases which genomics studies.

State-of-the-art models are seen when it comes to machine learning applications, where they integrate wide-ranging types like those above efficiently. Ensemble methods including Random Forest, XGBoost and LightGBM have demonstrated encouraging results in accurately predicting survival outcomes in glioblastoma patients, but not as good as the Dense NN. These tree based models utilize multiple decision trees that enable them to hold complex relationships among radiomic, clinical, and genomic features in order to improve their predictive performance.

Moreover, the use of deep learning techniques, such as custom CNN, in radiogenomics is a common practice. Hierarchical representations of imaging data can be learnt by CNNs thereby bringing out subtle patterns that would not easily be visible. By merging clinical, genomic data and radiomic features with CNNs, there is ongoing research towards better glioblastoma prognostic models.

The present state-of-the-art in predicting glioblastoma survival is highly dynamic with on-going work aimed at improving existing models, exploiting alternative sources of information and investigating whether AI solutions could make these intricate models more interpretable. As the field moves forward, integration of radiomics, clinical data and genomics seems promising for advancing precision medicine in glioblastoma hence enhancing personalized and efficient treatment approaches for patients with this fatal disease.

1.4 Document Structure

The rest of this document is structured like this:

- **Chapter 2** covers background information on glioblastoma, with the current prognosis and treatment methods, the role of machine learning, and related work in the field.
- **Chapter 3** introduces the proposed multimodal framework, with a focus on the architecture and data integration strategies used for survival prediction.
- **Chapter 4** describes the datasets, preprocessing steps, feature extraction, and model training processes.
- **Chapter 5** presents the experimental results, including performance metrics, comparisons, and model interpretability using SHAP.
- **Chapter 6** offers an integrated discussion and final reflections, addressing the results, limitations, clinical implications, future research directions, and summarizing the study's contributions.

At the top there is the list of figures, list of tables, and a glossary of terms.

2. Background and related work

2.1 Artificial intelligence in healthcare

Artificial intelligence (AI) is the process of simulating human-like critical thinking and intelligent behavior using computers and other technology. Machine learning (ML) is a subfield of artificial intelligence that allows computers to actually learn from training data without the need for programming. In recent years, AI and ML have gain proper popularity as techniques for improving patient outcomes and healthcare accuracy, particularly in the field of cancer excelled [6].

These technologies address several areas of cancer, like risk assessment, early detection, prognosis estimation, and even therapy selection. Artificial intelligence (AI) systems have the capability to analyze vast quantities of intricate medical data, such as genetic data, imaging tests, and electronic health records, with the purpose of detecting patterns and forecasting outcomes. For example, machine learning models have demonstrated high accuracy in predicting various types of cancer, including breast, brain, lung, liver, and prostate cancer [7].

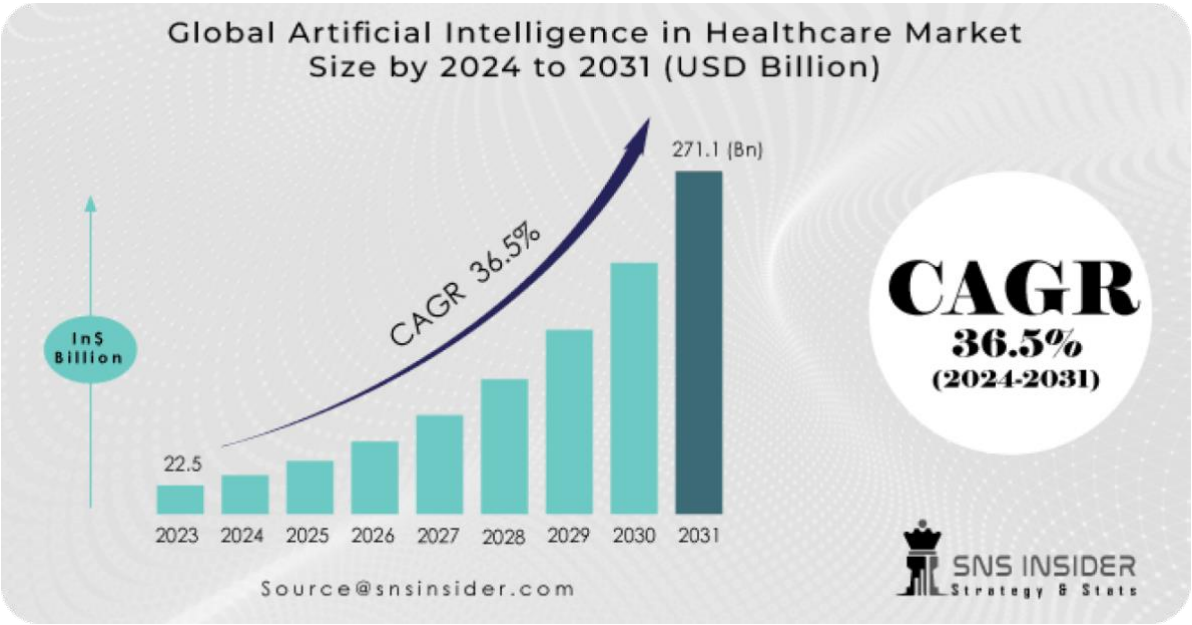

Figure 2.1 Global Artificial Intelligence in Healthcare Market Size Projection (2024-2031) [8]

According to this number from the Figure 2.1, the worldwide AI healthcare industry is projected to rise by 2024–2031. The market is anticipated to increase at a compound annual growth rate (CAGR) of 36.5% by 2031, when it is expected to reach $271.1 billion. This rapid growth reflects the rising importance and application of AI in healthcare, particularly in oncology [8].

AI in the therapy of cancer has the capacity to provide custom and evidence-based care. Artificial intelligence (AI) technologies can improve doctors' decision-making by integrating diverse data sources. This is particularly useful for cancer therapy, since treatment choices are often impacted by a range of factors such as patient demographics, tumor characteristics, and genetic markers.

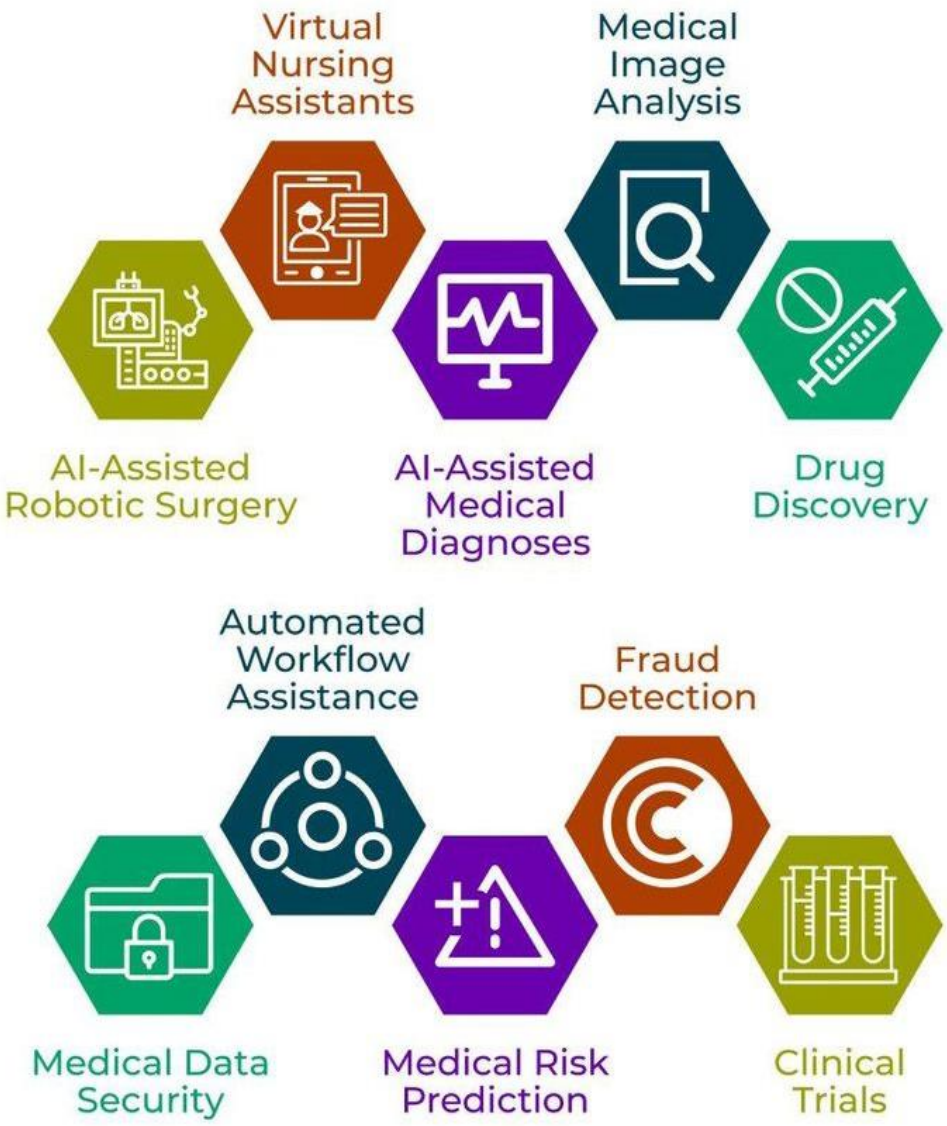

Figure 2.2 Applications of AI in Medical Imaging

The numerous uses of AI in medical imaging across a range of medical specializations are depicted in this diagram (Figure 2.2). AI applications in medical imaging have a wide range of uses in several areas of healthcare and extend beyond only detecting cancer. In the field of oncology, artificial intelligence (AI) assists in many tasks such as advanced radiomics, study of tumor burden, identification of lesions, and assessment, among other functions. These applications exemplify the transformative potential of AI in cancer diagnosis and treatment methodologies.

While artificial intelligence (AI) has great potential for cancer therapy, it is important to recognize that many of these applications are still in their early stages of development. Challenges persist in ensuring the precision of the data, comprehending the algorithms, and integrating them into therapeutic processes. Furthermore, there are ethical considerations around the safeguarding of data privacy, the existence of algorithmic bias, and the changing function of human physicians in healthcare systems augmented by artificial intelligence.

Figure 2.3 Challenges and Ethical Considerations in AI Implementation for Oncology

Figure 2.3 displays the fundamental moral principles necessary for the ethical integration of AI into medical practice. These principles ensure that AI technologies are developed and utilized in a manner that improves patient care and resolves moral difficulties.

2.2 Glioblastoma and the Need for Improved Prognostic Tools

The genetic diversity of GBM greatly hinders the creation of dependable prognostic methods. Recent study has discovered many molecular subtypes of GBM, such as pro-neural, neural, classical, and mesenchymal. Every individual have a distinct medical background and genetic makeup. These subtypes exhibit distinct reactions to therapy and outcomes in terms of prognosis, highlighting the need of molecular classification in prognostic modeling [9].

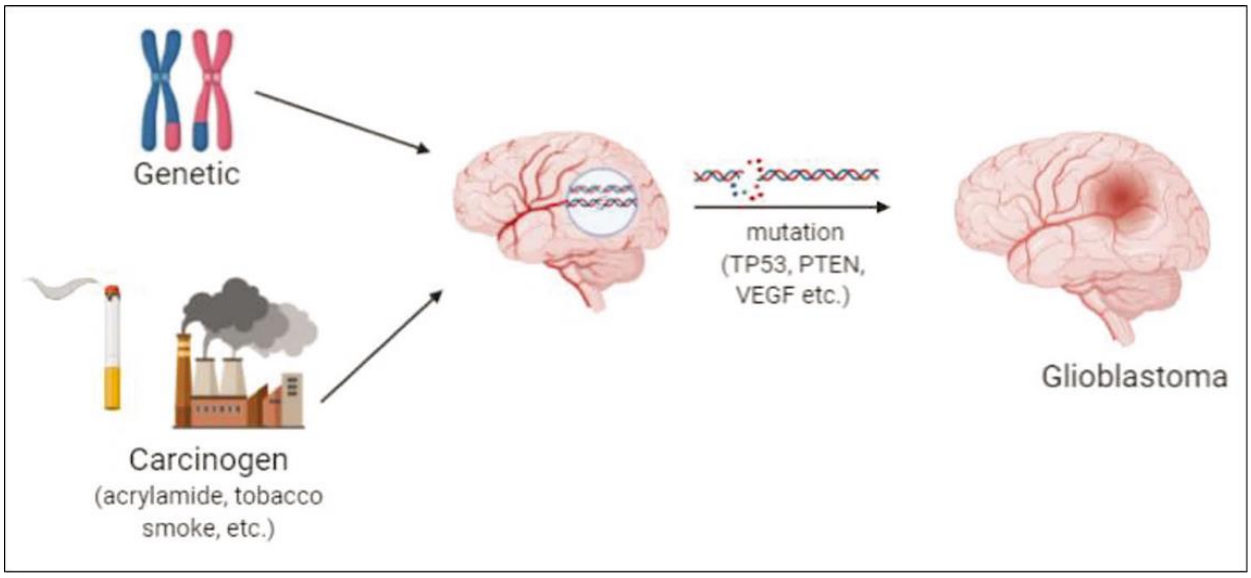

Figure 2.4 Pathways Leading to Glioblastoma Development

The image above (Figure 2.4) shows the genetic and environmental variables that lead to the development of glioblastoma (GBM). It highlights two main routes: genetic changes (including TP53, PTEN, and VEGF) and exposure to carcinogens (such acrylamide and tobacco smoke). These components lead to mutations that result in the development of glioblastomas from normal brain cells. It illustrates the complicated and multifaceted character of GBM causation by integrating genetic predispositions and environmental factors [10].

Tumor microenvironment (TME) is an equally important component that has an affect on the prognosis. It does consist of extracellular matrix elements, stromal cells, and immune cells, has a major influence on the development, invasion, and response to therapy of tumors. Recent studies have shown that the composition and characteristics of the TME can provide valuable prognostic information, suggesting that incorporating TME-related features into prognostic models could improve their performance [11].

Table 2.1 Prognostic Indicators for Glioblastoma and Their Potential Impact

Factor	Description	Potential Impact on Prognosis
Genetic Alterations	Mutations in genes like TP53, PTEN, VEGF	Has the potential to impact therapy response and tumor behavior.
Environmental Factors	Exposure to carcinogens (e.g., acrylamide, tobacco smoke)	May influence tumor development and aggressiveness

Factor	Description	Potential Impact on Prognosis
Tumor Microenvironment	Composition of immune cells, stromal cells, and extracellular matrix	Can impact tumor growth, invasion, and treatment efficacy
Advanced Imaging Features	Radiomics from PWI, DTI, and other techniques	Provide non-invasive assessment of tumor characteristics
Minimal Residual Disease	Presence of residual tumor cells post-treatment	May indicate risk of recurrence and guide follow-up strategies
Immunotherapy Response	Variability in patient response to immunotherapy	Predictive biomarkers could inform treatment decisions
AI and Machine Learning	Analysis of complex, multi-dimensional data	Potential to improve prediction accuracy and identify novel prognostic factors

For an in-depth analysis of the primary factors affecting the prognosis of glioblastoma, Table 2.1 from above combines both new and existing prognostic markers. This comprehensive viewpoint highlights the intricate interactions among genetic, environmental, and microenvironmental variables that affect the prognosis of GBM in addition to the potential for advanced technologies to improve prognostic accuracy. By integrating these many components, researchers and healthcare providers can develop more intricate and customized approaches for predicting GBM outcomes and guiding treatment decisions.

Modern imaging methods like diffusion tensor imaging (DTI) and perfusion-weighted imaging (PWI) provide new possibilities for the non-invasive evaluation of GBM features. The prognosis may be correlated with information obtained from these procedures on the vascularity, cellularity, and infiltration patterns of tumors. Integrating quantitative imaging features (radiomics) with clinical and genomic data has shown promise in developing more accurate prognostic models [12].

Minimal residual disease (MRD) is an increasingly common concept in GBM research. "MRD" refers to the presence of residual tumor cells after therapy that are not detected by standard imaging techniques. The development of sensitive methods for measuring and identifying MRD, such as liquid biopsy techniques, may have an impact on treatment decisions and provide insightful prognostic information [13].

Immunotherapy is becoming a recognized treatment option for GBM and other cancers. The way that GBM patients react to immunotherapy varies widely, though. Finding biomarkers that can

forecast immunotherapy response and incorporating them into prognostic models may help with treatment planning and patient stratification [14].

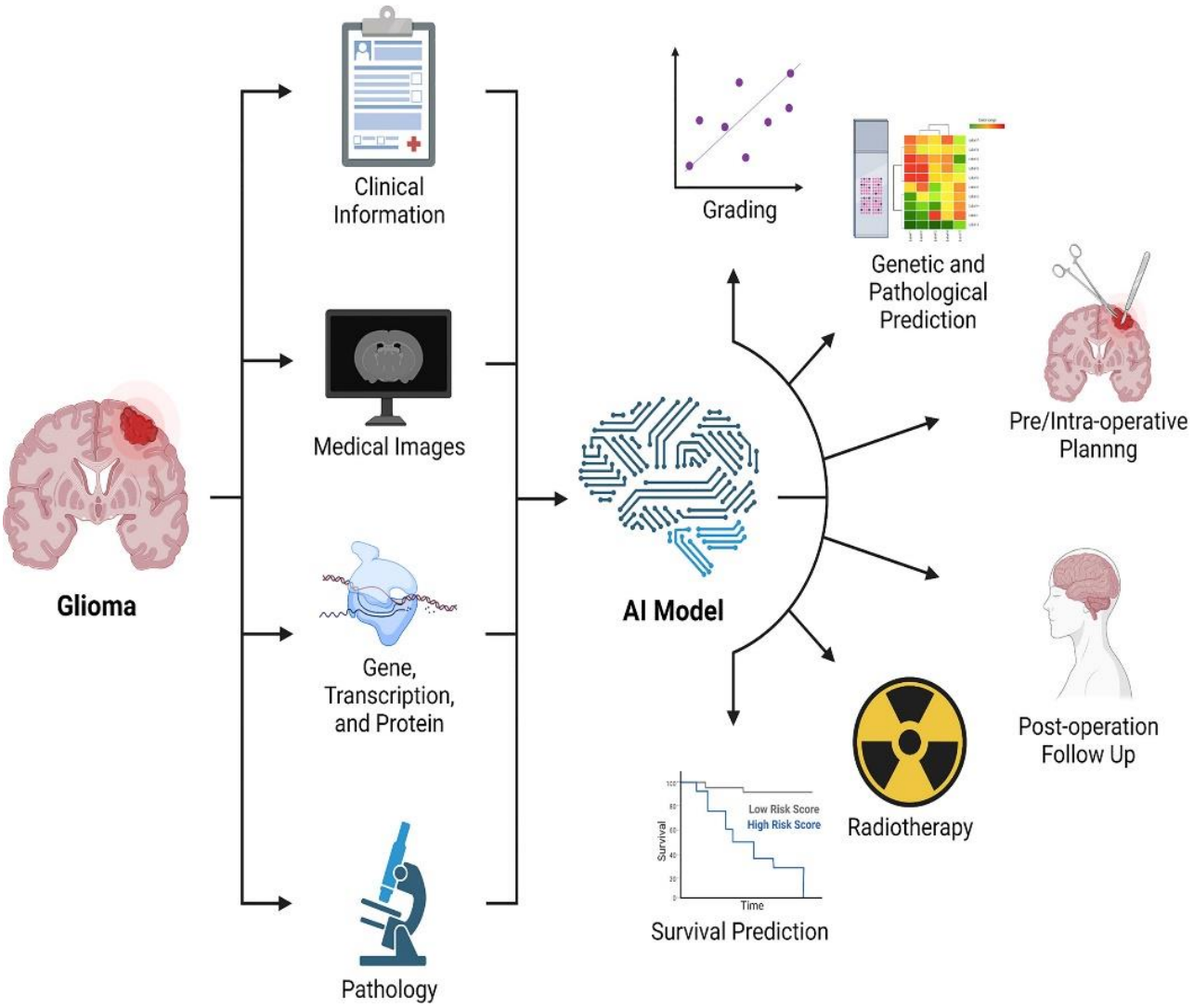

Figure 2.5 AI-Based Glioma Analysis System

A complex AI-driven method for glioma analysis and treatment is shown in the Figure 2.5. It provides different input data sources on the left, such as pathology, genetic and protein data, clinical information, and medical images. These many inputs are then processed by an AI model at its core, which is visualized as a stylized brain with circuit patterns. The tumor grading, genetic prediction, pre-operative planning, post-operative follow-up, radiation planning, and survival prognosis are just a few of the outputs that are displayed on the right side of the system.

By integrating complex, multi-modal data, the diagram effectively illustrates how artificial intelligence (AI) can impact the management of gliomas. It highlights how AI can analyze different types of data to support multiple aspects of patient care, from initial diagnosis through the last steps like treatment planning to long-term prognosis.

Finally, dynamic prognostication is becoming more and more popular in GBM research. With this method, now the prognostic can be estimated and updated as treatment progresses in response to changing clinical, imaging, and molecular data. Treatment choices and patient counseling may be guided by more precise and timely information from dynamic prognostic models than ever [15].

2 3. Radiomics and Radiogenomics

Emerging fields like radiogenomics and radiomics hold great promise for enhancing glioblastoma diagnosis, prognosis, and treatment planning. While radiogenomics connects these imaging parameters to underlying genetic properties, radioomics extracts a huge number of quantitative variables from medical images to define malignancies. Research has indicated that these methods have promise in terms of distinguishing between various types of tumors, forecasting survival, evaluating the impact of treatment, and non-invasively identifying molecular subtypes.

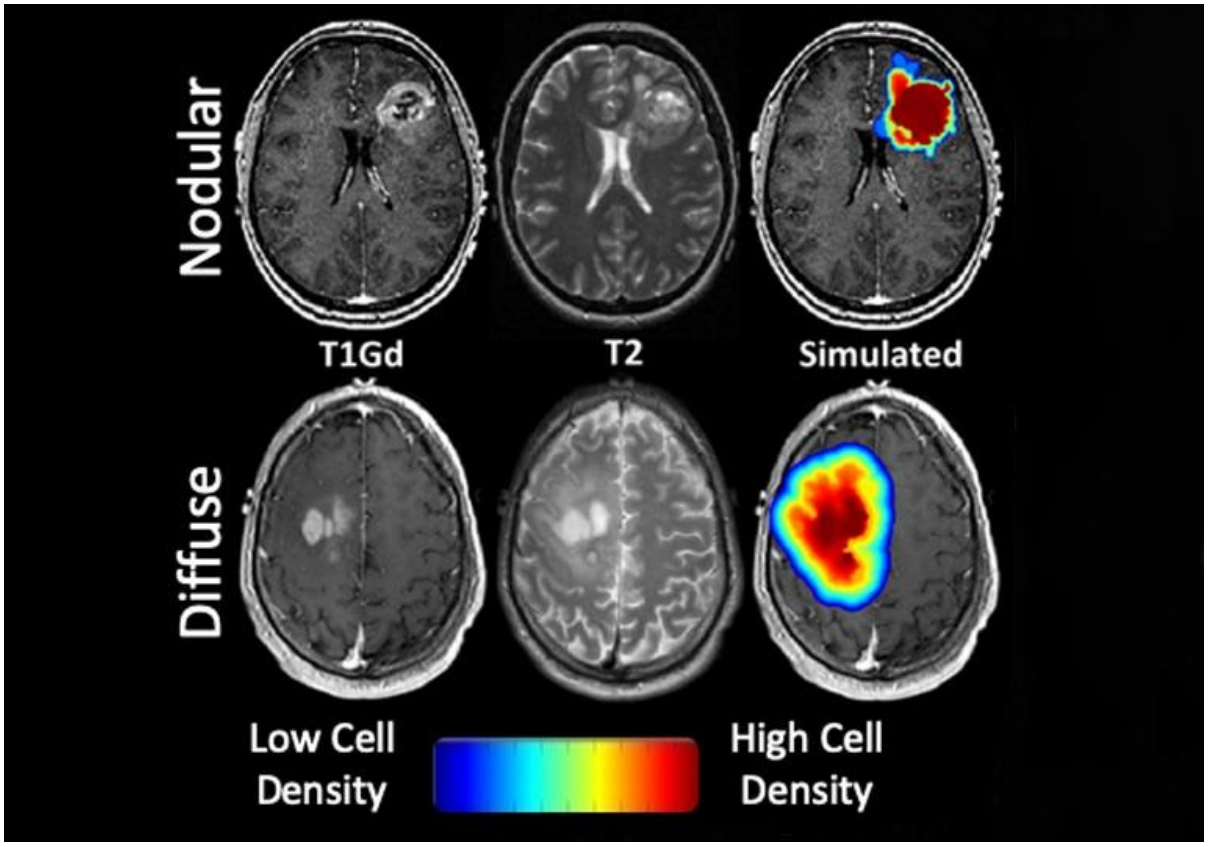

Figure 2.6 Radiomics Approaches in Glioblastoma Analysis

The Figure 2.6 shows the two main approaches to radiomics used in glioblastoma analysis: feature-based radiomics and deep learning-based radiomics. These methods are powered by the core technology of artificial intelligence. It shows the roles that neural network icons represent for deep learning and computer screen icons for regular machine learning play in the radiomics process. These methods are used with MRI scans, as shown by the MRI machine symbol at the bottom of the graphic.

The circular form draws attention to the connections between various radiomics methods. It demonstrates how various AI techniques may be applied to recognize and assess minute

characteristics from medical images, perhaps leading to a more exact and comprehensive glioblastoma classification.

Important uses include identifying real progression from pseudoprogression, predicting the spatial location of tumor appearance, and by predicting molecular properties such as IDH mutation and MGMT methylation status, and also the risk stratification of patients by predicting overall and progression-free survival. When compared to traditional clinical criteria alone, radiomics models have demonstrated increased predictive accuracy. Making individualized treatment decisions may be aided by the capacity to anticipate genetic markers and evaluate tumor heterogeneity non-invasively [16].

Table 2.2 Applications and Challenges of Radiomics and Radiogenomics in Glioblastoma

Aspect	Radiomics	Radiogenomics
Definition	Extraction of quantitative features from medical images	Linking imaging features to underlying genomic characteristics
Applications	- Tumor characterization - Treatment response assessment - Survival prediction - Pseudoprogression vs. true progression	- Non-invasive molecular subtyping - Prediction of genetic mutations (e.g., IDH, MGMT) - Identification of genomic signatures
Imaging Modalities	MRI (T1, T2, FLAIR, DWI, PWI), CT, PET	Primarily MRI, integrated with genomic data
Analysis Techniques	- Texture analysis - Shape features - Intensity-based features - Machine learning algorithms	- Correlation analysis - Machine learning - Deep learning
Clinical Impact	- Improved prognostic accuracy - Personalized treatment planning - Non-invasive tumor monitoring	- Guiding targeted therapies - Enhancing molecular diagnosis - Facilitating precision medicine
Challenges	- Standardization of imaging protocols - Reproducibility of features - Large-scale validation	- Integration of heterogeneous data types - Biological interpretation of correlations - Limited availability of matched imaging-genomic datasets
Future Directions	- Deep learning integration - Multi-institutional studies - Automated feature extraction	- Multi-omics integration - Longitudinal studies - Development of radiogenomic atlases

However, a number of challenges must be resolved prior to clinical use. These include the need for institutions to use uniform imaging procedures, feature extraction methods, and segmentation tactics. Models' repeatability and generalizability provide a significant problem. Larger multi-institutional datasets and prospective validation studies are necessary.

Included in the steps that follow are the integration of longitudinal imaging data, the investigation of deep learning methods, the development of fully automated procedures, and the focus on the biological validation of radiomic signals. Based on the data presented in Table 2.2, it is obvious that radiogenomics and radiomics have considerable potential to further precision medicine treatment for patients with glioblastoma.

2.4 The Promise of Integrative Approaches

Enhancing GBM survival prediction with the use of AI and ML approaches to integrate radiomics, clinical data, and genetic information is a promising strategy. It could be able to create more precise and individualized prognostic models by utilizing the advantages of each form of data and AI's capacity for pattern detection. Clinical trial design, patient counseling, and therapy planning may all benefit from the use of these models.

Combining several data modalities enables a more thorough description of the biology of tumors and patient characteristics that affect prognosis. Features from radiomics that are taken from MRIs can capture finer details of tumor heterogeneity and infiltration that are not visible by human eye. Clinical data offers crucial background information on patient traits and medical history. Via genomic profiling, important molecular changes influencing tumor behavior are identified.

Note: The diagram presented below exemplifies a Convolutional Neural Network (CNN) structure, widely employed for the analysis of imaging data, including MRI scans. Although this picture does not depict a comprehensive method that integrates radiomics, genomes, and clinical data, it effectively demonstrates the specific deep learning models that may be used for analyzing imaging data in glioma diagnostics.

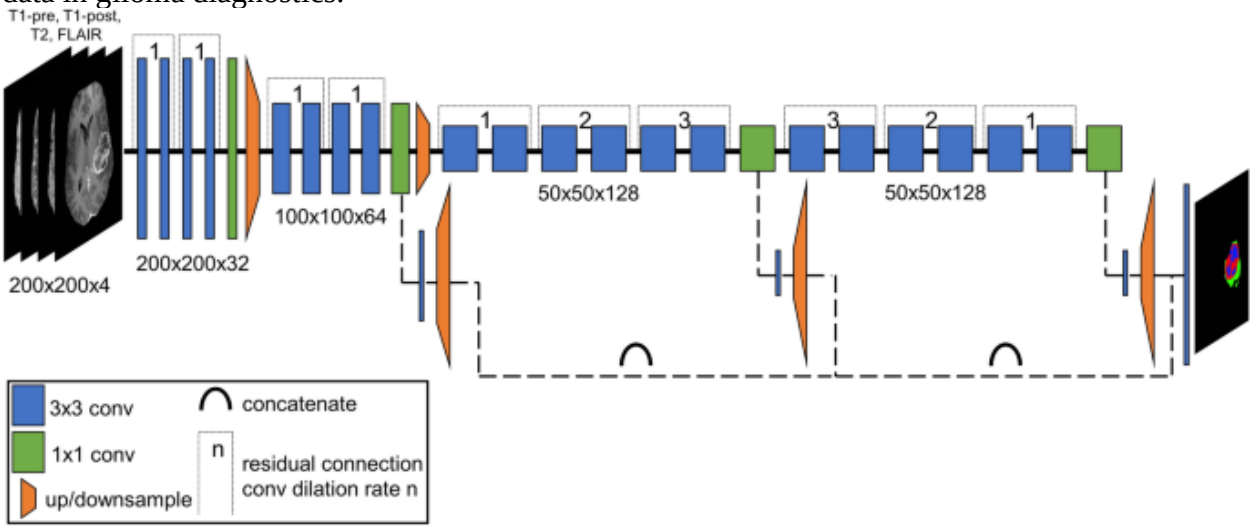

Figure 2.7 AI-Based Integrative Approach for Glioma Analysis and Management

The Figure.2.7 shows the step-by-step workflow of an AI-based integrated approach to glioma analysis and management. It demonstrates how many therapeutically relevant outputs may be generated by utilizing an AI model to combine and interpret multiple data sources. The sources of the input data are shown on the left side of the picture; these sources include clinical data, pathology, gene, transcription, and protein data, as well as medical photographs. At the heart of all these inputs is an AI model that takes the form of a stylized brain with circuit patterns. On the right side of the image are the system's outputs, which include genetic prediction, post-operative follow-up, radiation planning, tumor grading, and survival prognosis.

Powerful machine learning algorithms have demonstrated the capacity to efficiently integrate these disparate data kinds and pinpoint intricate patterns linked to survival consequences, especially ensemble techniques like RandomForest, XGBoost and LightGBM. When these integrative models are used instead of methods that rely solely on individual data modalities, the predicted performance is frequently better.

These integrative methods provide more potential than merely survival prediction. By identifying individuals who are more likely to respond to particular medications, they may facilitate more individualized treatment planning. Through the identification of unique connections between imaging, clinical, and genomic variables, the models may also shed light on the biology of GBM.

The potential for integrative techniques to improve precision medicine for GBM patients is substantial, despite ongoing problems in harmonizing data collecting and processing methodologies across institutions. To improve these models and apply them to clinical practice, data scientists, imaging specialists, and doctors must work together and conduct ongoing research. Integrative radiogenomic models might be useful tools for bettering outcomes in this debilitating condition as the research develops.

3. Dataset

3.1 Dataset description

For both datasets, UCSF-PDGM and UPENN-GBM, the present work followed an 80/20 ratio to partition the data into training and testing sets. Specifically, the UCSF-PDGM dataset, including 1656 samples from 414 distinct patients, was divided into 1324 samples for training and 332 samples for testing purposes. Moreover, the UPENN-GBM dataset, including 2569 samples from 602 distinct patients, was partitioned into 2055 samples for training and 514 samples for testing. The chosen distribution was selected to maximize the data available for model training, while simultaneously ensuring a significant portion is allocated for evaluating model performance.

3.1.1 UCSF-PDGM Dataset

MRI images and clinical information from 501 individuals with histopathologically confirmed diffuse gliomas (WHO grades 2-4) were initially included in the UCSF-PDGM dataset. 414 distinct individuals with comprehensive and verified clinical data made up the final group following stringent data curation and quality control procedures. [17]

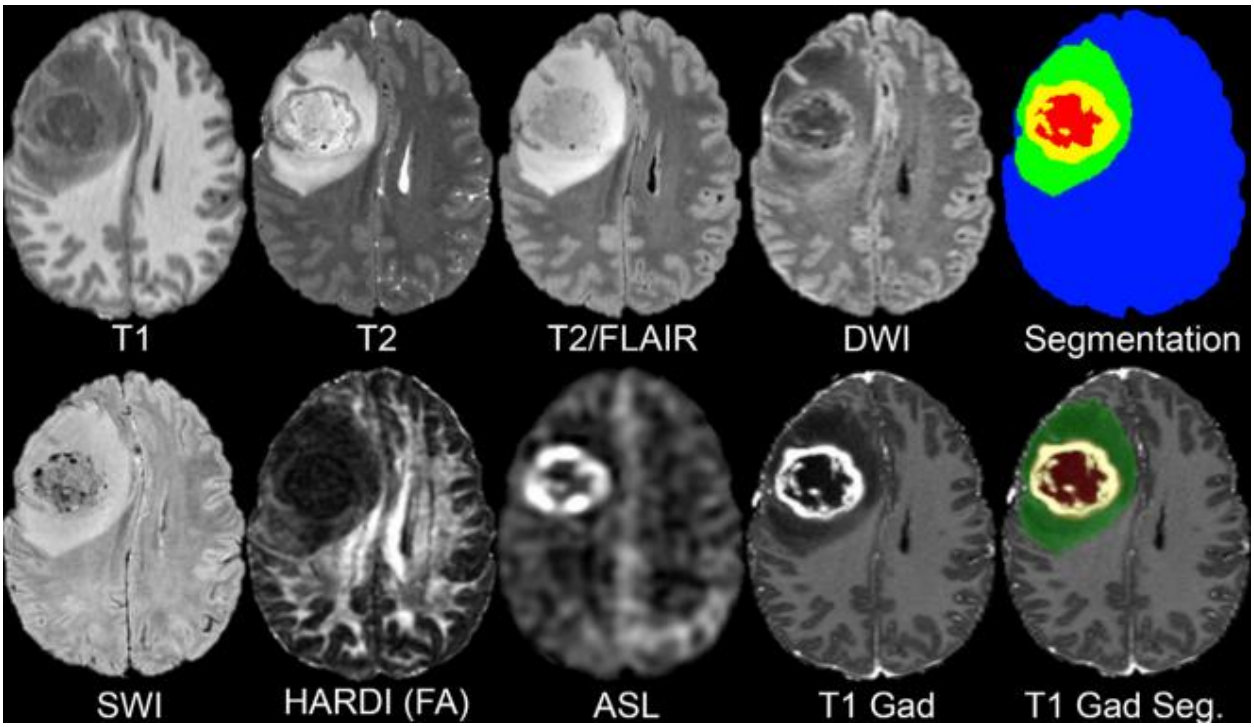

Figure 3.1 Representative Multimodal MRI Studies in a 37-Year-Old Man with Glioblastoma from the UCSF-PDGM Dataset

Representative multimodal MRI examinations from a 37-year-old male patient with glioblastoma are shown in in Figure 3.1 from the UCSF-PDGM dataset. The imaging modalities include ASL perfusion, isotropic diffusion-weighted imaging (DWI), and fractional anisotropy (FA) from high-angular-resolution diffusion imaging (HARDI) is shown in Figure.3.1. Multicompartment tumor segmentation is another feature that shows several aspects of the tumor, such as the brain, enhancing tumor, necrotic core, and FLAIR anomalies. The image displays susceptibility-weighted imaging (SWI), precontrast T1-weighted (T1), post-gadolinium T1-weighted (T1 Gad),

and T2-weighted (T2) images in addition to a T1 Gad segmentation overlay. This comprehensive illustration emphasizes the importance of multimodal MRI in providing accurate anatomical and functional information that is necessary for precise tumor characterization and treatment planning in glioblastoma patients.

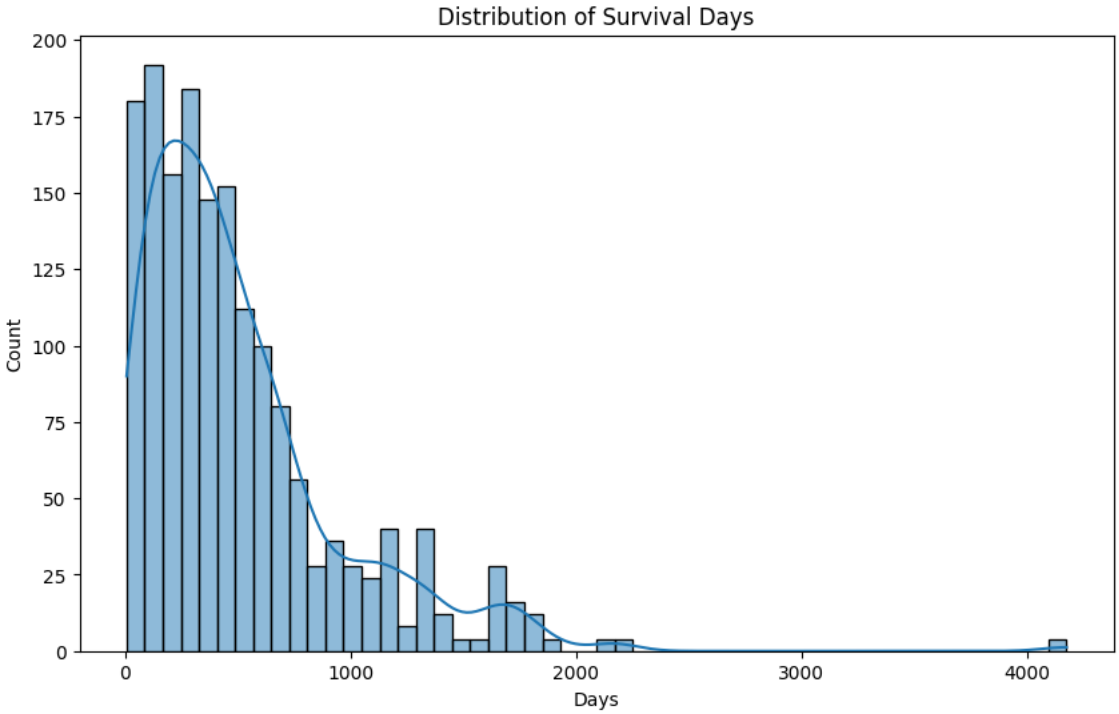

Figure 3.2 Distribution of Survival Days for UCSF Dataset

Survival day distribution for patients in the UCSF-PDGM dataset is seen in Figure 3.2. A peak of around 200-400 days is observed in the histogram, indicating that the majority of patients have a survival duration of fewer than 1000 days. Although greater survival periods are seen in the tail of the distribution, they are exceptional and indicative of the aggressive character of glioblastoma.

Table 3.1 UCSF-PDGM Patient Demographics and Clinical Summary

Attribute	Details
Number of Unique Patients	414
Gender Distribution	252 males, 162 females
Age Statistics	Mean age: 59.6 years (range: 17-94 years)
IDH1 Mutation Status	Wildtype: 373, p.R132H: 22, mutated (NOS): 8, p.R132C: 5, p.R132G: 3, p.R132S: 2, p.Arg172Trp: 1

Attribute	Details
MGMT Methylation Status	Positive: 296, Negative: 113, Indeterminate: 5
Survival Days Statistics	Mean: 510.8 days, Median: 391.5 days, Min: 6 days, Max: 4177 days, Std Dev: 463.7 days
Imaging Protocol	Standardized 3-T MRI, including 3D T2-weighted, T2-FLAIR, susceptibility-weighted, diffusion-weighted, and pre- and postcontrast T1-weighted images

The data in Table 3.1 shows a gender distribution of 162 females and 252 males, with an average age of 59.6 years (spanning from 17 to 94 years). The majority of patients (373) are IDH1 wildtype, according to genetic study, although some patients (22 patients) have p.R132H mutation and a few other less prevalent variants. In 296 patients, the MGMT methylation status was positive; in 113, it was negative, and in 5 cases, it was unknown. The survival statistics demonstrate a considerable variation in patient outcomes, with a mean survival of 510.8 days, a median of 391.5 days, and a broad range of 6 to 4177 days. A standardized 3-T MRI technique was utilized to obtain the imaging data, guaranteeing uniformity among various modalities such as T2-weighted, T2-FLAIR, susceptibility-weighted, diffusion-weighted, and pre- and postcontrast T1-weighted pictures.

3.1.2 UPENN-GBM Dataset

Advanced multi-parametric MRI scans, clinical, genetic, and radiomic data from 630 individuals diagnosed with de novo glioblastoma initially made up the UPENN-GBM dataset. The final dataset comprised of 602 distinct individuals with confirmed and thorough clinical profiles, having undergone extensive data verification and refinement procedures. [18]

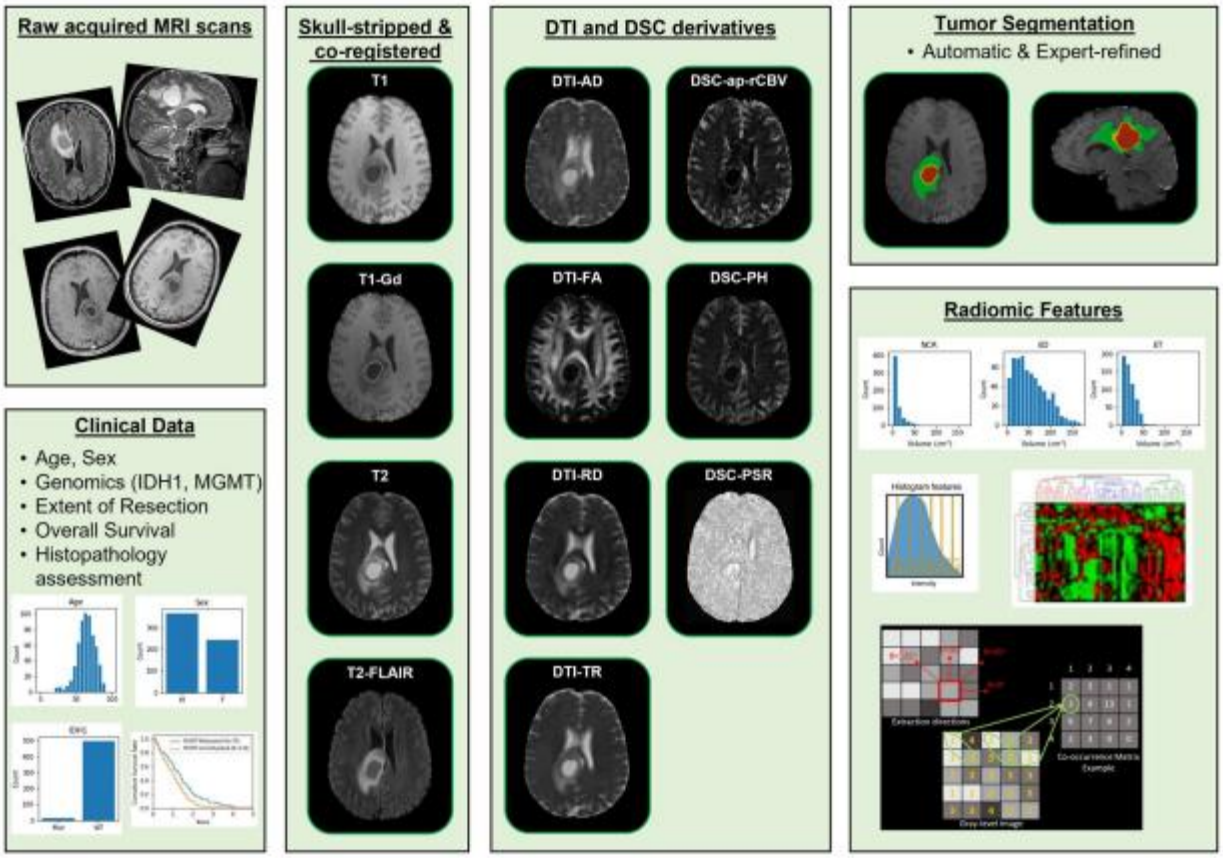

Figure 3.3 Comprehensive Overview of the UPENN-GBM Dataset Components

The Figure 3.3 provides an in-depth visual summary of the University of Pennsylvania Glioblastoma Advanced Imaging, Clinical, Genomics, and Radiomics (UPENN-GBM) data collection. This shows complex structure of the dataset by highlighting its four main components: advanced MRI data, clinical data, genetic data, and radiomic properties. The advanced MRI data's multi-parametric scans demonstrate the dataset's abundance of imaging information.

Clinical data includes vital patient information such as demographics, treatment course details, and survival rates. The genomic component focuses on significant molecular indicators such IDH mutant status and MGMT promoter methylation. Lastly, the radiomic characteristics reflect the quantitative imaging qualities extracted from the MRI data.

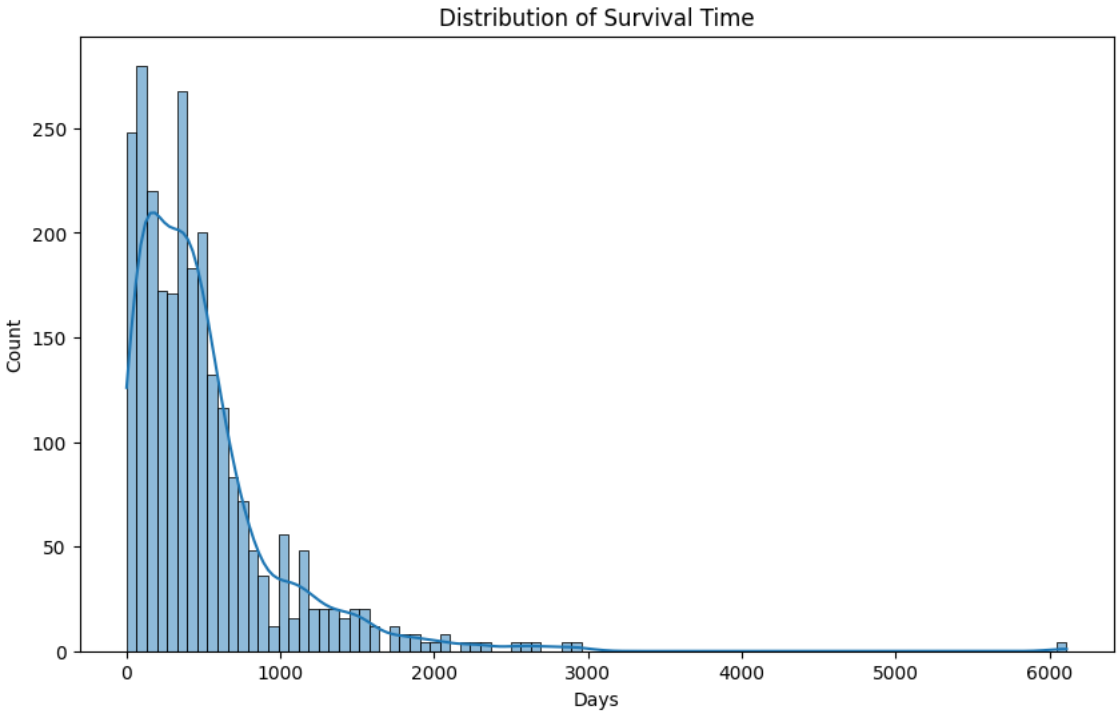

Figure 3.4 Distribution of Survival Days for UPENN Dataset

Figure 3.4 represents the manner in which survival days are distributed across patients in the UPENN-GBM dataset. Comparable to the UCSF dataset, the UPENN dataset exhibits a significant clustering of patients with survival durations below 1000 days, with the bulk falling within the range of 200 to 500 days. Furthermore, the dataset contains a small number of outliers characterized by much prolonged life durations, which highlights the considerable variation in patient outcomes.

Table 3.2 UPENN-GBM Patient Demographics and Clinical Summary

Attribute	Details
Number of Unique Patients	602
Gender Distribution	360 males, 242 females
Age Statistics	Mean age: 63.17 years (range: 20.74-88.5 years)
IDH1 Mutation Status	Wildtype: 495, NOS/NEC: 94, Mutated: 13
GTR over 90% Distribution	Yes: 347, No: 203, Not Available: 34, Not Applicable: 18
Survival Days Statistics	Mean: 510.8 days, Median: 384.5 days, Min: 3 days, Max: 6109 days, Std Dev: 518.2 days

Based on the information from Table.3.2 , the gender distribution is made up of 242 females and 360 males. Its mean age is 63.17 years, and its range is 20.74 to 88.5 years. Based on the IDH1

mutation status, most patients (495) are classed as wildtype, whereas a smaller minority is classified as NOS/NEC (94) and mutated (13). Furthermore, the gross total resection (GTR) over 90% distribution shows that 347 patients met this criterion, but 203 patients did not; for 34 patients, the GTR status was unavailable, and it was not relevant for 18. With a standard deviation of 518.2 days, the survival statistics show a mean survival of 510.8 days with a median of 384.5 days, indicating significant variability ranging from as little as 3 days to as long as 6109 days.

3.2 Data preprocessing

3.2.1 DICOM to NIfTI Conversion

One of the most important steps in our thesis is converting the MRI images from **DICOM** (Digital Imaging and Communications in Medicine) files to **NIFTI** (Neuroimaging Informatics Technology Initiative) format. This change is mandatory because, although DICOM is the industry standard for the gathering and archiving of medical imaging data, NIfTI is more frequently used in neuroimaging research because of its more straightforward structure and improved support for analytical tools.

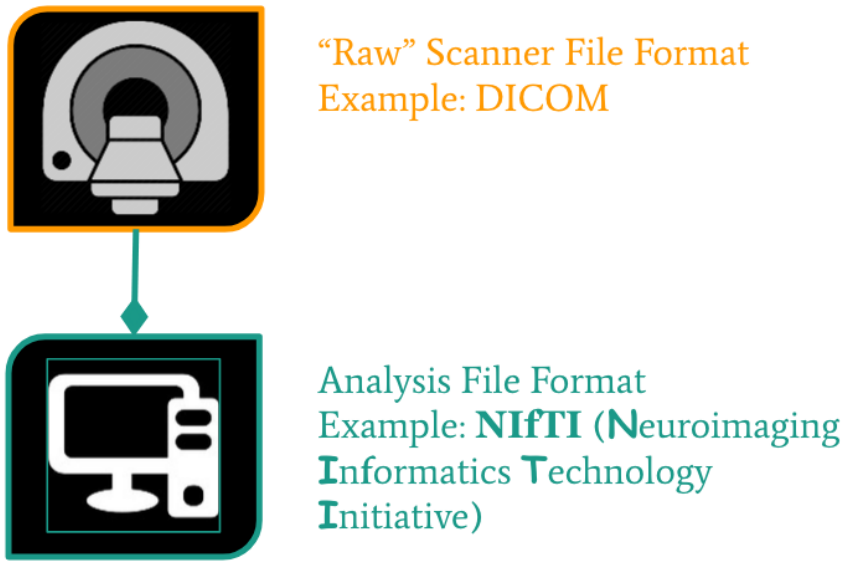

Figure 3.5 Conversion of DICOM to NIfTI Format

The Figure 3.5 illustrates how we converted DICOM files, which are commonly used for storing and sharing medical imaging data, into the NIfTI format, which is better suited for overall medical imaging research . This conversion simplifies the handling and analysis of brain imaging data, making it an essential step in the preprocessing and analysis of glioblastoma research.

For both datasets, this conversion was performed using the dcm2niix library in python. Here's an example of the command-line usage:

```
dcm2niix -z y -o /output-directory /input-files
```

Figure 3.6 dcm2niix tool command

In this command:

- -z y enables compression of the output NIfTI files
- -o specifies the output directory for the converted files
- The last argument is the path to the directory of the DICOM files

This conversion procedure from Figure 3.6 reformats the data into a more research-friendly format while keeping the crucial imaging information The NIfTI format simplifies the manipulation and analysis of brain imaging data, which is crucial for subsequent preprocessing and analysis in glioblastoma research. It is important to remember that during the conversion process, key information is typically extracted from the DICOM headers and stored in accompanying JSON files. This metadata helps track the provenance of the imaging data and can be valuable for further analysis.

3.2.2 MRIPreprocessor Pipeline

To standardize the MRI image preprocessing **MRIPreprocessor** pipeline has been used . Several essential processes are included in this pipeline to guarantee quality and uniformity throughout the dataset:

- It's removing non-brain tissues from the images to focus on the brain region. (**Skull Stripping**)

- It's correcting intensity non-uniformities caused by magnetic field inhomogeneities. (**Bias Field Correction**)

- It's aligning images to a common anatomical space, which is crucial for comparative studies. (**Registration**)

- It's ensuring uniform voxel sizes across all images, which facilitates consistent analysis. (**Resampling)**

```python
from MRIPreprocessor.mri_preprocessor import Preprocessor

# 4 Modalities to co-register to MNI space using an affine transformation
# T1 is used as reference for the coregistration
# No labelmap is used
ppr = Preprocessor({'T1':'./data/example_T1.nii.gz',
                    'T2':'./data/example_T2.nii.gz',
                    'T1c':'./data/example_T1c.nii.gz',
                    'FLAIR':'./data/example_FLAIR.nii.gz'},
                   output_folder = './data/output',
                   reference='T1',
                   label=None,
                   prefix='patient001_',
                   already_coregistered=False,
                   mni=True,
                   crop=True)

ppr.run_pipeline()
```

Figure 3.7 MRIPreprocessor Pipeline for MRI Image Preprocessing example code from Github

The pipeline shown in Figure 3.7 provides a standardized approach to prepare MRI data for further analysis, ensuring consistency across images and modalities. It is particularly useful for multi-modal MRI datasets like those used in brain tumor studies like our datasets.
The MRIPreprocessor pipeline, from GitHub, deals with MRI preprocessing tasks. We have performed the following steps :

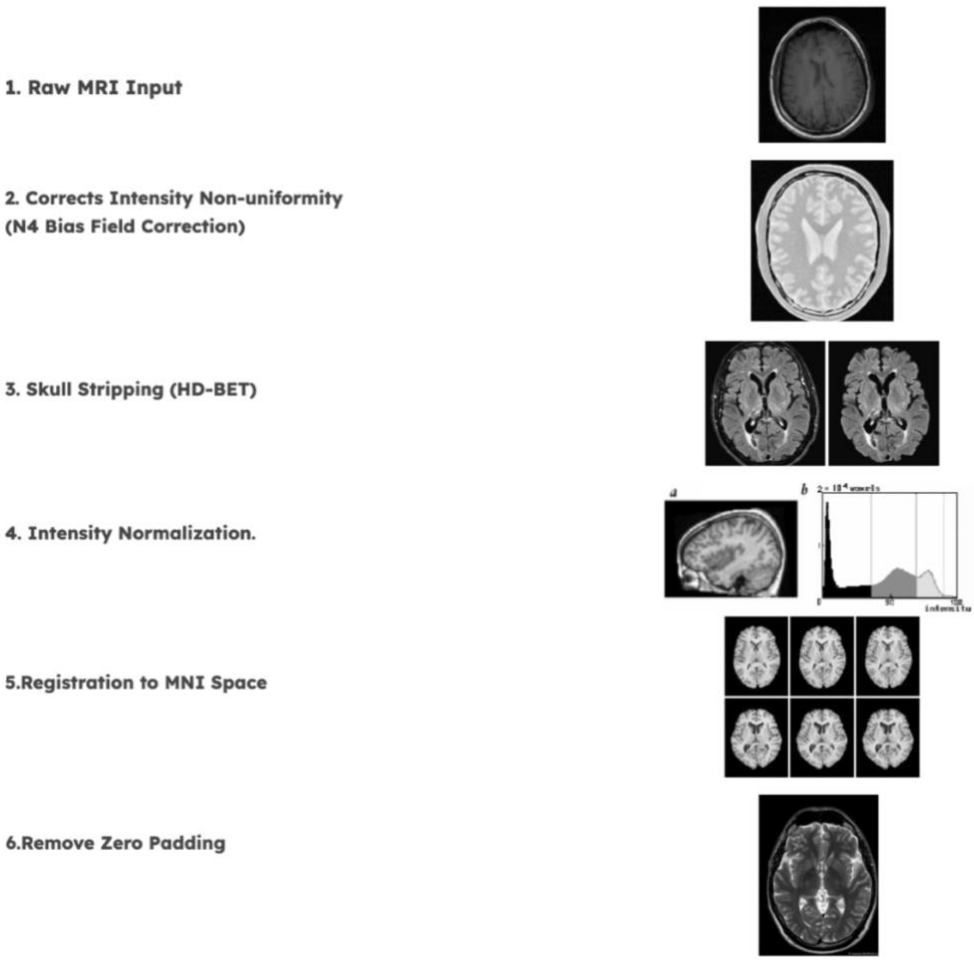

Figure 3.8 Visualization of MRI Preprocessing Workflow

The figure above (Figure 3.8) represents a systematic procedure for preparing MRI data, demonstrating the progressive conversion of MRI pictures via several phases. The workflow comprises the subsequent stages: (1) The raw MRI scan is displayed; (2) Intensity Non-uniformity is corrected using N4 Bias Field Correction to ensure uniformity in intensity levels throughout the image; (3) Skull Stripping is performed to isolate the brain tissue from the MRI; (4) Intensity Normalization is applied to standardize intensity values for consistency across different scans; (5) Registration to MNI Space is performed to align the brain images to a standard template for comparison; and (6) Zero Padding is removed to eliminate any unnecessary padding around the image. Every step is accompanied by a graphical depiction of the MRI data following the relevant processing stage, therefore offering a lucid perspective of the modifications implemented during preprocessing.

After running the pipeline the files look like this (Figure 3.9):

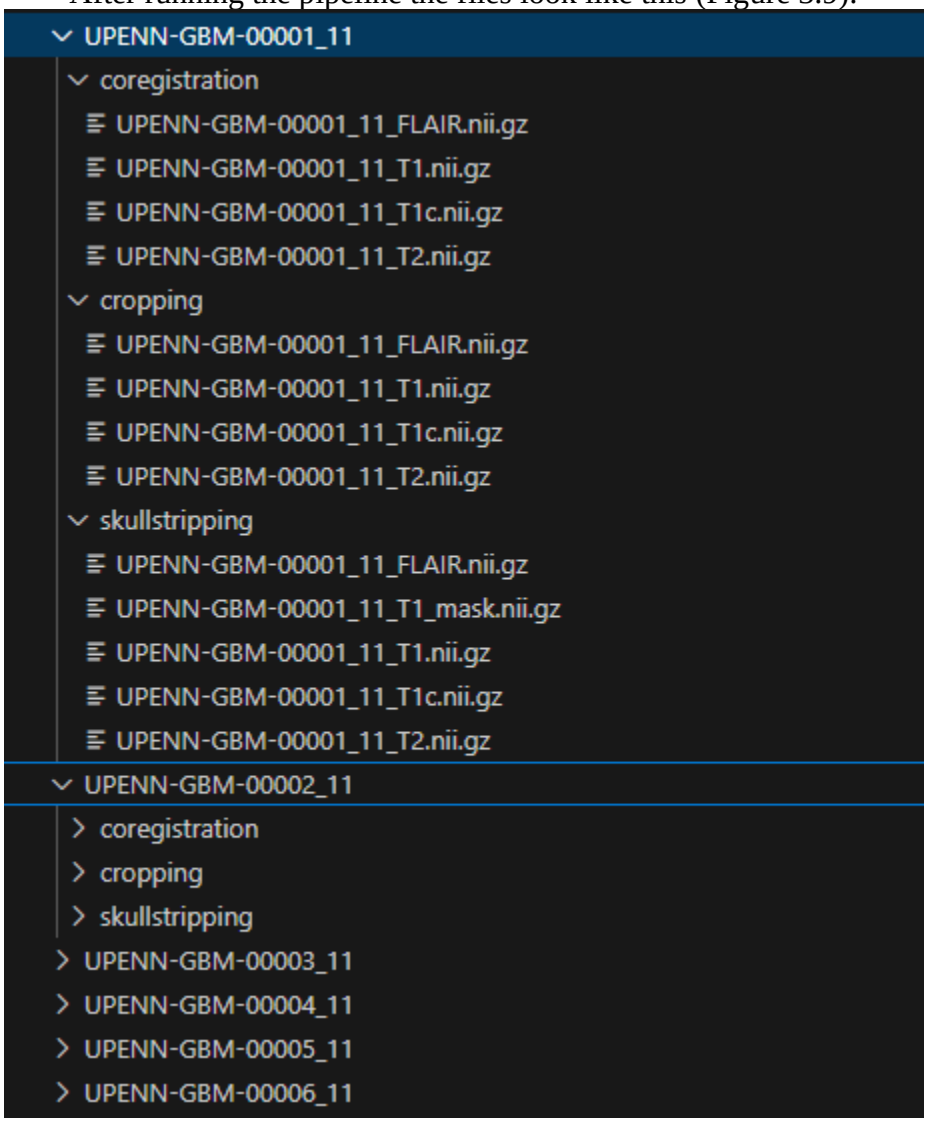

Figure 3.9 Structure of each patient after preprocessing

3.3 Feature extraction

Feature extraction is a vital process in radiomics that involves deriving quantitative information from medical images. These features can be utilized for various purposes, such as prognosis, evaluating treatment responses, and diagnosing diseases. In this thesis, we will discuss how to extract radiomic features from MRI images using Python.

A strong tool for extracting a variety of radiomic features from medical photos is the "featureextractor" found in the PyRadiomics package. First-order statistics, shape-based features, and texture features derived from neighboring gray tone difference matrices (NGTDM), gray-level co-occurrence matrices (GLCM), gray-level run length matrices (GLRLM), gray-level size zone matrices (GLSZM), and gray-level dependence matrices (GLDM) are feature extraction functions found in PyRadiomics. (Table 3.3)

Table 3.3 Most relevant features extracted from PyRadiomics

Feature Type	Description
First-Order Statistics	These features describe the distribution of voxel intensities within the image region defined by the mask. Metrics include mean, median, and standard deviation.
Shape-Based Features	These features describe the geometric properties of the region of interest (ROI), such as volume, surface area, and sphericity.
Texture Features	These include metrics that describe the overall texture of the image provided.
- GLCM	Measures the frequency of co-occurring intensity values at a given offset.
- GLRLM	Measures the length of consecutive runs of pixels with the same intensity.
- GLSZM	Measures the size of zones of connected pixels with the same intensity.
- NGTDM	Measures the difference between a pixel and its neighboring pixels.
- GLDM	Measures the dependence of gray levels in an image.

The "featureextractor" is adjustable, so users can modify it to their own requirements by changing up to many parameters:

- **binWidth:** Defines the bin width for features that contain histograms.
- **resampledPixelSpacing:** Specifies the voxel size for the resize.
- **interpolator:** Chooses the interpolation method.
- **enableCExtensions:** Enables the use of C-extensions.

Step-by-Step Process.

Libraries used:

- **os**: For reading directories and paths in different operating systems
- **pandas (pd)**: For manipulating of the clinical data stored in CSV files.
- **nibabel (nib)**: To handle neuroimaging data, NIfTI files.
- **pyradiomics (featureextractor)**: To extract radiomics features from MRI images.
- **tqdm**: For showing progress bars.

1. **Load Clinical Data**

   The first step involves loading the clinical data, which contains patient information and metadata. This data is essential for linking the extracted radiomic features with clinical outcomes. (Figure 3.10)

```
clinical_data_path = '/path-UCSF/UPENN_preprocessed.csv'
clinical_data = pd.read_csv(clinical_data_path)
clinical_data.columns =
clinical_data.columns.str.strip()  # Strip any
leading/trailing spaces in column names
print(clinical_data.head())
```

Figure 3.10 Loading and Preprocessing Clinical Data

2. **Clean and Standardize Clinical Data**
   The 'ID' column in the clinical data is standardized. This involves converting the IDs to a uniform format. (Figure 3.11)

```
clinical_data['ID'] = clinical_data['ID'].astype(str)
clinical_data['ID'] = clinical_data['ID'].apply(lambda
x: 'UCSF/UPENN-' + x.split('-')[2].zfill(4))
```

Figure 3.11 Modifying ID Column for UCSF/UPENN Dataset

3. **Define Paths**
   The directory containing the MRI images is specified. This path will be used to locate the MRI scans for each patient. (Figure 3.12)

```
mri_dir = '/path-MRI-images'
print(os.listdir(mri_dir))
```

Figure 3.12 Initializing MRI Directory Path and Listing Files in Directory

4. **Initialize Feature Extractor**
   A feature extractor is defined using the PyRadiomics library. The parameters for the extractor, such as bin width and interpolator type, are set to control the feature extraction process. (Figure 3.13)

```
params = {
    'binWidth': 25,
    'resampledPixelSpacing': None,
    'interpolator': 'sitkBSpline',
    'enableCExtensions': True
}
extractor =
featureextractor.RadiomicsFeatureExtractor(**params)
```

Figure 3.13 Setting Radiomics Feature Extraction Parameters

5. **Define the Feature Extraction Function**
   A function is created to extract features from the given image and mask paths. This function uses the feature extractor to compute the radiomic features. (Figure 3.14)

```python
def extract_features(image_path, mask_path):
    result = extractor.execute(image_path, mask_path)
    features = {k: v for k, v in result.items() if
'diagnostics' not in k}
    return features
```

Figure 3.14 Function to Extract Radiomics Features from MRI Images

6. **Extract Features for Each Patient**

   The script iterates through each patient's directory, checking for the existence of required subdirectories and files. For each valid image-mask pair, features are extracted and stored. (Figure 3.15)

```python
radiomics_features_list = []
log_file_path = '/path/to/radiomics_extraction_log.txt'

with open(log_file_path, 'w') as log_file:
    for patient_id in tqdm(clinical_data['ID'], desc="Processing Patients"):
        patient_base_dir = os.path.join(mri_dir, f'{patient_id.lower()}_nifti')

        if not os.path.exists(patient_base_dir):
            log_file.write(f"Base directory does not exist for patient {patient_id}: {patient_base_dir}\n")
            continue

        subdirectories = ['coregistration', 'cropping', 'skullstripping']

        for sub_dir in subdirectories:
            patient_dir = os.path.join(patient_base_dir, sub_dir)
            if not os.path.exists(patient_dir):
                log_file.write(f"Directory does not exist for patient {patient_id} in {sub_dir}: {patient_dir}\n")
                continue

            modalities = ['FLAIR', 'T1', 'T1c', 'T2']

            for modality in modalities:
                image_path = os.path.join(patient_dir, f'{patient_id.lower()}_nifti_{modality}.nii.gz')
                mask_path = os.path.join(patient_dir, f'{patient_id.lower()}_nifti_t1_mask.nii.gz')

                if os.path.exists(image_path) and os.path.exists(mask_path):
                    try:
                        features = extract_features(image_path, mask_path)
                        if features:
                            features['ID'] = patient_id
                            features['Modality'] = modality
                            features['Processing Step'] = sub_dir
                            radiomics_features_list.append(features)
                    except Exception as e:
                        log_file.write(f"Error extracting features for patient {patient_id}, modality {modality} in {sub_dir}: {e}\n")
```

Figure 3.15 Processing Pipeline for Extracting Radiomics Features Across Multiple Patients and Modalities

7. **Convert to DataFrame and Clean**

   The extracted features are converted to a pandas DataFrame for further analysis. The DataFrame is cleaned to ensure consistency. (Figure 3.16)

```python
radiomics_features = pd.DataFrame(radiomics_features_list)
if not radiomics_features.empty:
    radiomics_features['ID'] = radiomics_features['ID'].astype(str)
    radiomics_features.columns = radiomics_features.columns.str.strip()
    print(radiomics_features.head())
else:
    print("No radiomics features were extracted.")
```

Figure 3.16 Converting Radiomics Features to DataFrame and Displaying Results

8. **Integrate Radiomics with Clinical Data**
   The radiomic features are merged with the clinical data to create a comprehensive dataset that includes both clinical and imaging features. (Figure 3.17)

```
if not radiomics_features.empty:
    combined_data = pd.merge(clinical_data, radiomics_features, on='ID', how='inner')
    combined_data_path = '/path/to/combined_clinical_radiomics_data.csv'
    combined_data.to_csv(combined_data_path, index=False)
    print(f"Data saved successfully to {combined_data_path}.")
else:
    print("Skipping merge since no radiomics features were extracted.")
```

Figure 3.17 Merging Radiomics Features with Clinical Data and Saving Results

Following the detailed, step-by-step instructions for feature extraction using the PyRadiomics library, the final files provide a large collection of quantitative features that were extracted from the MRI images. These characteristics include a variety of first-order statistics, shape-based features, and texture features, and they cover several imaging modalities and processing stages. The final dataset provides a comprehensive overview of the extracted radiomic features, with each row corresponding to a specific patient and imaging modality. It is organized to allow seamless integration with clinical data, making it easier to explore and analyze potential correlations between imaging features and clinical outcomes.

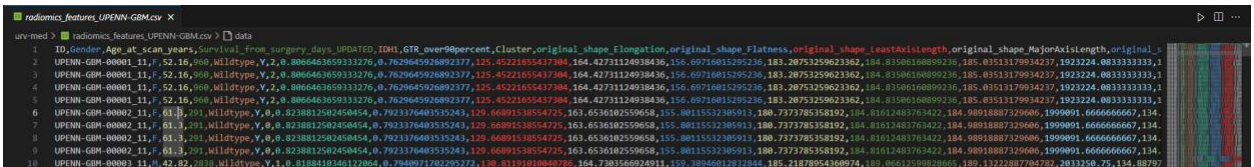

Figure 3.18 Radiomics Feature Dataset Combined with the Clinical Dataset for the UPENN-GBM

This output (Figure 3.18) reflects the extensive and complex feature collection that was generated, which is significant for additional analysis and interpretation within the framework of medical

imaging research. By examining these traits, researchers can get valuable insights into the underlying patterns and relationships that may aid enhance knowledge of disease characteristics and patient outcomes.

3.4 Feature Selection

Feature selection is a critical step in the data preprocessing pipeline, especially when dealing with high-dimensional datasets such as those derived from multimodal MRI scans in glioblastoma research. By identifying the most informative features, this stage seeks to minimize the dimensionality of the data, which can enhance model performance and lessen overfitting.

In this study, various feature selection methods were employed to identify the most relevant radiomic and clinical variables for predicting patient outcomes. These techniques were carefully chosen to ensure that the selected features are both clinically and statistically significant.

3.4.1 Model-based Feature Selection

One of the main feature selection methods used in this study was model-based selection using SelectFromModel. This method leverages the feature importance scores provided by ensemble models such as RandomForest, LightGBM and XGBoost. During the training process, these models look at the importance scores of each feature based on how much they contribute to reducing the error in predictions.

```
from sklearn.feature_selection import SelectFromModel
from xgboost import XGBRegressor

model = XGBRegressor(random_state=42)

# create the pipeline with SelectFromModel
pipeline = Pipeline([
    ('preprocessor', preprocessor),
    ('feature_selection', SelectFromModel(estimator=model)),
    ('regressor', model)
])

pipeline.fit(X_train, y_train)

# get the selected features
selected_features = X_train.columns[pipeline.named_steps['feature_selection'].get_support()]
print("Selected features:", selected_features)
```

Figure 3.19 Model-based Feature Selection with SelectFromModel and XGBoost

This model-based from Figure 3.19 allows us to reduce the feature number significantly while keeping a high level of predictive power.

3.4.2 Statistical Feature Selection

In addition to model-based selection, we also used this statistical method for feature selection. Specifically, we applied SelectKBest with the f_regression function, which selects features based on their univariate statistical significance. This method evaluates each feature individually by computing the connection between the feature and the target variable (Survival Days).

```
for name, (model, param_dist) in models.items():
    print(f"\nTraining {name}...")
    pipeline = Pipeline([
        ('preprocessor', preprocessor),
        ('feature_selection', SelectKBest(f_regression, k=30)),
        ('regressor', model)
    ])

    random_search = RandomizedSearchCV(pipeline, param_distributions=param_dist,
                                       n_iter=50, cv=rkf, scoring='neg_mean_squared_error',
                                       n_jobs=-1, random_state=42)
    random_search.fit(X_train, y_train)

    best_model = random_search.best_estimator_

    y_train_pred = best_model.predict(X_train)
    y_test_pred = best_model.predict(X_test)

    train_metrics = calculate_metrics(y_train, y_train_pred)
    test_metrics = calculate_metrics(y_test, y_test_pred)
```

Figure 3.20 Statistical Feature Selection with SelectKBest and f_regression

The top k features, based on their F-statistics, were selected for the final model (Figure 3.20). This approach is highly effective for identifying features that exhibit a strong linear relationship with the target variable.

4. Proposed Method

4.1 Model Training Approach with Implementation Details

The main objective of this work was to create predictive models for patient survival from glioblastoma (GBM) by combining genetic, clinical, and radiomic characteristics. The use of sophisticated machine learning models placed a focus on deciphering the intricate relationships present in the data. Here is a detailed description of the process along with particulars on how the code was implemented.

4.2 Model Selection Rationale

To capture the complexity of the data, several machine learning models were selected for their robustness and ability to handle high-dimensional and heterogeneous datasets:

4.2.1 Random Forest Regressor

The Random Forest model was selected for its ensemble nature, which reduces overfitting by averaging multiple decision trees. This model is appropriate for the integrated dataset since it excels at managing a combination of numerical and categorical data.

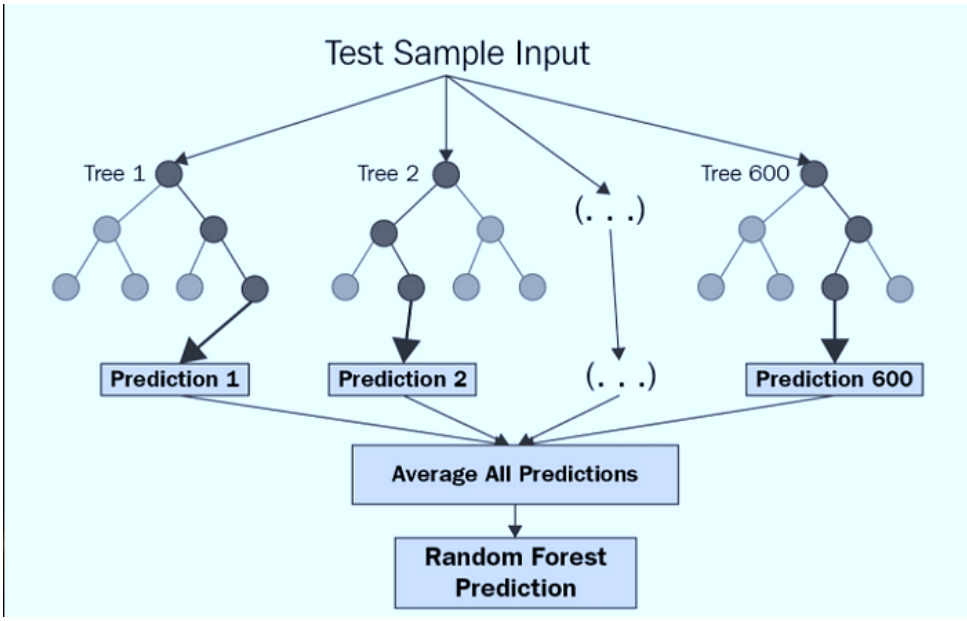

Figure 4.1 Ensemble Prediction Process in Random Forest Regressor

The Random Forest Regressor's ensemble prediction method shows in Figure 4.1. This method involves training several decision trees using various data subsets. The nodes and branches that depict the decision processes within each tree indicate how each tree separately produces a prediction based on the input features. The model predicts each tree separately for a particular test sample input .

By averaging the predictions from each individual tree, the Random Forest Regressor's ultimate prediction is determined. Comparing this ensemble approach to a single decision tree lowers the variance and strengthens the model's robustness. The model reduces the possibility of overfitting

and enhances generalization to new data by averaging the predictions. By merging several weak learners into a single strong learner, the bagging (bootstrap aggregating) principle is used in this process to improve model stability and accuracy. The collaborative character of the forest's trees is portrayed in the Figure 4.1 in an efficient manner, leading to an accurate and trustworthy prediction.

The model was implemented using `RandomForestRegressor` from the `sklearn.ensemble` module. Key hyperparameters such as the number of trees (`n_estimators`), maximum depth (`max_depth`), and minimum samples per split (`min_samples_split`) were optimized using `RandomizedSearchCV`.

4.2.2 XGBoost Regressor

XGBoost (Extreme Gradient Boosting) was chosen for its efficiency and high performance in handling large, complex datasets. XGBoost is a gradient boosting framework that builds models sequentially, where each new model focuses on correcting the errors made by previous models. This iterative approach allows XGBoost to capture intricate patterns in the data and produce highly accurate predictions.

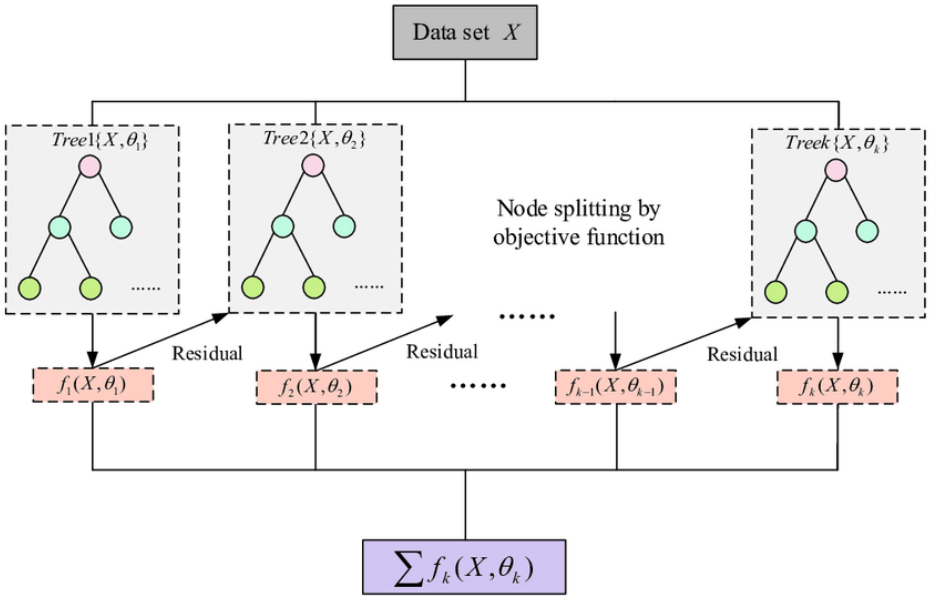

Figure 4.2 Gradient Boosting Process in XGBoost

The gradient boosting process used by XGBoost, which builds several decision trees in succession to increase forecast accuracy, is shown in the Figure 4.2. The model first generates a prediction, from which the residuals, or errors, are computed. The next tree is then trained to forecast these residuals, so that it can practically learn from its predecessors' errors. Iteratively, this technique aims to improve the overall model by reducing mistakes with each new tree.

An objective function guides node splitting within each tree, figuring out the best way to split the data at each stage in order to reduce residuals. By adding up each tree's output and integrating their distinct contributions, the XGBoost model creates a final forecast that is reliable and accurate. The

image explains in detail how XGBoost concentrates on mistake correction through a series of tree-based improvements in order to create a strong prediction model.

The `XGBRegressor` class from the `xgboost` library was used to implement the model. Hyperparameters such as learning rate (`learning_rate`), maximum tree depth (`max_depth`), and the number of boosting rounds (`n_estimators`) were optimized through cross-validation to ensure the model could effectively generalize to new data.

4.2.3 LightGBM Regressor

LightGBM (Light Gradient Boosting Machine) was selected for its speed and efficiency, particularly when dealing with large datasets and high-dimensional data. LightGBM uses a leaf-wise tree growth strategy, which typically results in faster training and better accuracy with fewer iterations compared to other boosting methods. This makes it a strong candidate for scenarios where computational efficiency is crucial.

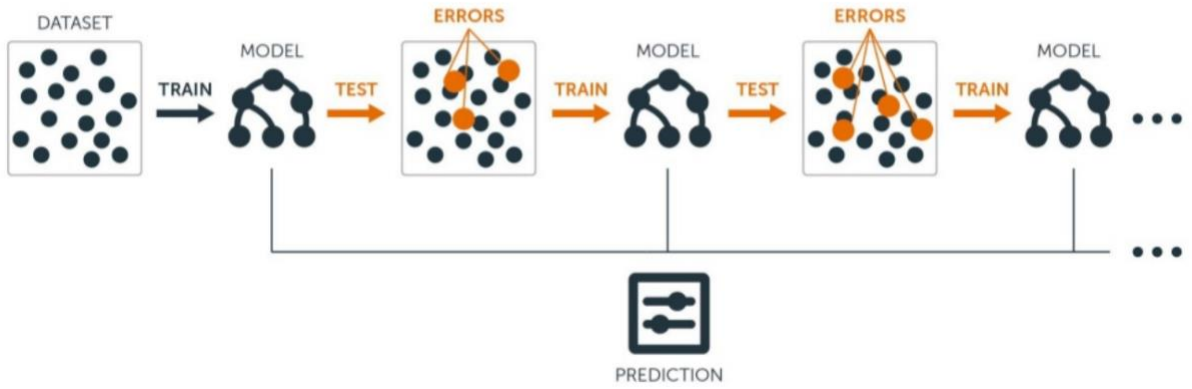

Figure 4.3 Gradient Boosting Process in LightGBM

The very effective and scalable gradient boosting architecture LightGBM (Light Gradient Boosting Machine) is depicted in the Figure 4.3 along with its boosting procedure. The first step in LightGBM is to train a baseline model on the dataset. The data points that were inaccurately predicted serve as a representation of the errors found after testing this model. LightGBM then uses these incorrectly classified points as training data for further models, trying to fix the mistakes the earlier models committed.

This iterative training continues, with each new model focusing on improving the prediction accuracy by targeting the errors of its predecessors. Each model contributes according to how well it reduces mistakes, and the final forecast is the sum of all the models. This illustration does a good job of showing how LightGBM's boosting technique gradually improves predictions to produce a reliable and accurate final model. LightGBM's exceptional performance is largely due to its novel leaf-wise tree growth technique and emphasis on error correction at every stage, particularly when managing larger datasets with intricate structures.

The model was implemented using the `LGBMRegressor` class from the `lightgbm` library. Key hyperparameters such as the number of leaves (`num_leaves`), maximum tree depth (`max_depth`),

and learning rate were tuned using `RandomizedSearchCV`. This optimization process ensured that the model could efficiently process the high-dimensional features derived from the radiomic, clinical, and genomic data.

4.2.4 Neural Network Regressor

A separate Neural Network Regressor was developed to investigate the non-linear complex interactions within the integrated dataset of genetic, clinical, and radiomic features. Neural networks are particularly well-suited for glioblastoma survival prediction due to their ability to detect complex patterns that traditional machine learning algorithms might miss. Their high flexibility allows them to capture intricate relationships between genetic, clinical, and imaging data that are crucial in this context.

Two architectures were explored: a dense neural network (Dense NN) and a wide-and-deep neural network (Wide & Deep NN). The Dense NN model consisted of multiple fully connected layers with Rectified Linear Unit (`ReLU`) activation functions, along with `BatchNormalization` and `Dropout` layers to enhance generalization and prevent overfitting. The Wide & Deep NN architecture combined a wide component, which aimed to capture linear relationships between features, with a deep component designed to learn complex feature interactions (Figure 4.4).

```
def create_model(input_shape, model_type='dense'):
    inputs = Input(shape=(input_shape,))

    if model_type == 'dense':
        x = Dense(256, activation='relu', kernel_regularizer=l1_l2(l1=1e-5, l2=1e-4))(inputs)
        x = BatchNormalization()(x)
        x = Dropout(0.3)(x)
        x = Dense(128, activation='relu', kernel_regularizer=l1_l2(l1=1e-5, l2=1e-4))(x)
        x = BatchNormalization()(x)
        x = Dropout(0.3)(x)
        x = Dense(64, activation='relu', kernel_regularizer=l1_l2(l1=1e-5, l2=1e-4))(x)
        x = BatchNormalization()(x)
        x = Dropout(0.3)(x)
    elif model_type == 'wide_and_deep':
        deep = Dense(128, activation='relu', kernel_regularizer=l1_l2(l1=1e-5, l2=1e-4))(inputs)
        deep = BatchNormalization()(deep)
        deep = Dropout(0.3)(deep)
        deep = Dense(64, activation='relu', kernel_regularizer=l1_l2(l1=1e-5, l2=1e-4))(deep)
        deep = BatchNormalization()(deep)
        deep = Dropout(0.3)(deep)

        wide = Dense(64, activation='relu')(inputs)

        x = Concatenate()([deep, wide])

    outputs = Dense(1)(x)
    model = Model(inputs=inputs, outputs=outputs)
    return model
```

Figure 4.4 Code for Dense Neural Network architecture with BatchNormalization, Dropout, and regularization

Dense Neural Network: The Dense NN model was constructed with three hidden layers of 256, 128, and 64 neurons, respectively. Each layer was followed by a `BatchNormalization` layer to stabilize learning and a `Dropout` layer with a dropout rate of 0.3 to mitigate overfitting. L1 and L2 regularization were applied to the layers to further improve the model's generalization capabilities (Figure 4.5).

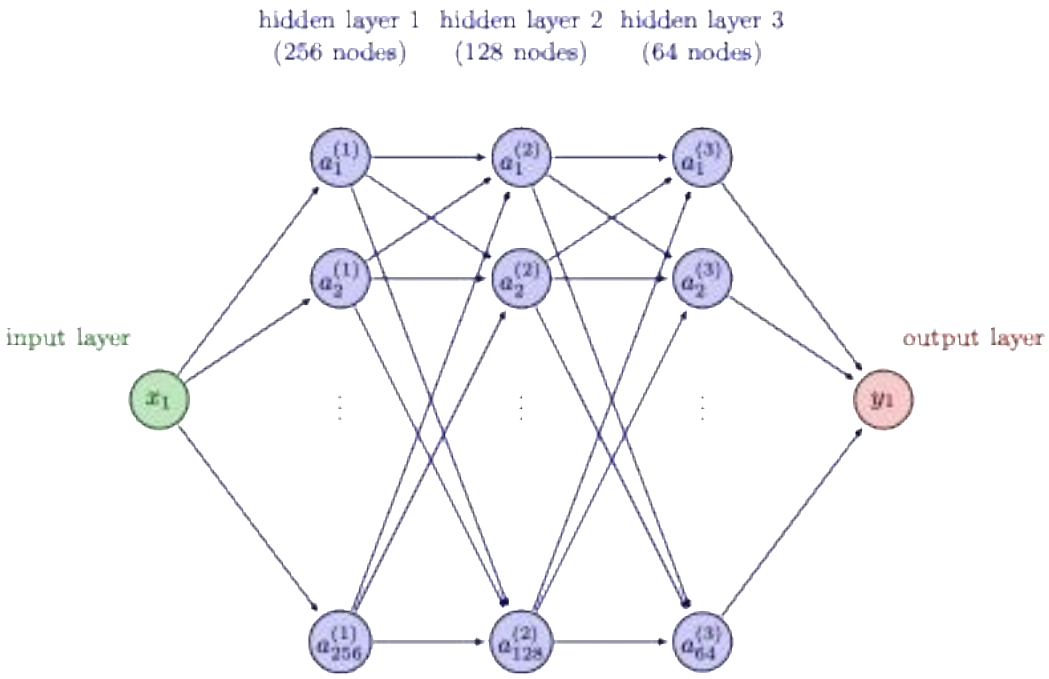

Figure 4.5 Visualization of the Dense Neural Network layers: 256, 128, 64 units, and 1-unit output

Wide & Deep Neural Network: The Wide & Deep NN model utilized both a deep path, similar to the Dense NN, and a wide path was composed of a single hidden layer with 64 neurons. The outputs from these two paths were concatenated to produce the final prediction.

The Adam optimizer was used to train both models due to its effectiveness in managing extensive datasets and its capacity to adjust the learning rate in real time while training. Early stopping was used to avoid overfitting; training was stopped when performance on a validation set stopped improving. To further improve the model, a `ReduceLROnPlateau` callback was included to lower the learning rate if the validation loss plateaued.

An ensemble model was also developed by averaging the predictions from the Dense NN and Wide & Deep NN models. This approach was intended to leverage the strengths of both architectures, aiming to create a more robust predictive model by combining the distinct advantages of each neural network structure.

In summary, the Dense Neural Network Regressor demonstrated the best potential for predicting patient survival in glioblastoma, making it the primary model chosen for integration into the overall predictive framework. The ensemble model further underscored the robustness of the Dense NN by integrating complementary predictions from both architectures.

4.3 Data Preparation and Preprocessing

To ensure the data was ready for model training, the following steps were performed:

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Figure 4.6 Splitting the dataset into training and testing sets with a test size of 20%

The dataset was split into training (80%) and testing (20%) sets using `train_test_split` from `sklearn.model_selection` to evaluate the generalizability of the models (Figure 4.6).

```
clinical_features = ['Age_at_scan_years', 'Gender', 'GTR_over90percent']
genomic_features = ['IDH1', 'Cluster']
radiomic_features = [col for col in data.columns if col.startswith('original_')]

print("Preparing data...")
X = data[clinical_features + genomic_features + radiomic_features]
y = data['Survival_from_surgery_days_UPDATED']
groups = data['ID']
```

Figure 4.7 Selection of Clinical, Genomic, and Radiomic Features for Model Input

The features included in the research were precisely selected from the database, including genetic, clinical, and radiomic data (Figure 4.7). An analysis of genomic characteristics, including the mutation status of IDH1 and the grouping of molecules, yielded vital information on the genetic composition of glioblastoma. An analysis was conducted to include significant clinical aspects that impact patient outcomes, such as patient demographics and treatment-related variables including age at scan, gender, and the degree of gross total resection (GTR). The radiomic characteristics, derived from MRI scans, obtained precise data on the texture, shape, and intensity of the tumor, providing a fully comprehensive understanding of tumor heterogeneity.

```
numeric_transformer = Pipeline(steps=[
    ('imputer', SimpleImputer(strategy='median')),
    ('scaler', StandardScaler()),
])

categorical_transformer = Pipeline(steps=[
    ('imputer', SimpleImputer(strategy='constant', fill_value='missing')),
    ('onehot', OneHotEncoder(handle_unknown='ignore')),
])

preprocessor = ColumnTransformer(
    transformers=[
        ('num', numeric_transformer, numeric_features),
        ('cat', categorical_transformer, categorical_features),
    ])
```

Figure 4.8 Pipeline for preprocessing numeric and categorical features with imputation, scaling, and one-hot encoding

Prior to preprocessing, the data was divided into numerical and categorical components. StandardScaler was used to normalize numerical characteristics, including clinical and radiomic

data, in order to provide a uniform scale for all variables (Figure 4.8). Using `OneHotEncoder`, categorical characteristics obtained from genomic and clinical data were transformed into a binary format appropriate for model input. This preprocessing procedure was carried out using a ColumnTransformer pipeline, which methodically performed the required transformations to all characteristics, guaranteeing consistency and uniformity throughout the data preparation process.

4.4 Model Training and Hyperparameter Optimization

Each model was trained using a pipeline that included preprocessing, feature selection, and the final model. The pipelines were optimized using the following approach:

```
results = {}
for name, (model, param_space) in tqdm(models.items(), desc="Training models"):
    pipeline = Pipeline([
        ('preprocessor', preprocessor),
        ('feature_selector', feature_selector),
        ('regressor', model)
    ])

    random_search = RandomizedSearchCV(pipeline, param_distributions=param_space,
                                       n_iter=100, cv=5, scoring='neg_mean_squared_error',
                                       n_jobs=-1, random_state=42, verbose=1)
    random_search.fit(X_train, y_train)

    best_model = random_search.best_estimator_

    cv_scores = cross_val_score(best_model, X_train, y_train, cv=5, scoring='r2')

    y_pred = best_model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    mae = mean_absolute_error(y_test, y_pred)
    c_index = concordance_index(y_test, y_pred)

    results[name] = {
        'R2 (Test)': r2,
        'R2 (CV)': cv_scores.mean(),
        'R2 (CV) std': cv_scores.std(),
        'MSE': mse,
        'MAE': mae,
        'C-index': c_index,
        'Best Parameters': random_search.best_params_
    }
```

Figure 4.9 Pipeline for model training with hyperparameter tuning using RandomizedSearchCV and evaluation with cross-validation

Each model was encapsulated within a `Pipeline` object from `sklearn.pipeline`, which ensured that the preprocessing, feature selection, and model training steps were applied consistently (Figure 4.9).

`RandomizedSearchCV` was used for hyperparameter optimization. This method randomly sampled from the defined hyperparameter spaces and evaluated model performance using cross-

validation. The search focused on optimizing key parameters like learning rate, tree depth, and number of trees for ensemble models, and learning rate, batch size, and epochs for the neural network.

The following table outlines the parameters used for the different machine learning models across both the UPENN-GBM and UCSF-PDGM datasets.

Table 4.1 Hyperparameters used for the UPENN-GBM and UCSF-PDGM datasets

Model	Dataset	Parameters
RandomForest	UPENN-GBM	max_depth: 5, max_features: 0.7081, min_samples_leaf: 9, min_samples_split: 19, n_estimators: 250
	UCSF-PDGM	max_depth: 9, max_features: 0.8832, min_samples_leaf: 5, min_samples_split: 5, n_estimators: 135
XGBoost	UPENN-GBM	colsample_bytree: 0.6061, gamma: 5.6672, learning_rate: 0.0601, max_depth: 3, min_child_weight: 1, n_estimators: 615, reg_alpha: 0.7159, reg_lambda: 1.0439, subsample: 0.9777
	UCSF-PDGM	colsample_bytree: 0.9332, gamma: 5.4164, learning_rate: 0.0262, max_depth: 5, min_child_weight: 6, n_estimators: 188, reg_alpha: 0.9770, reg_lambda: 1.3302, subsample: 0.7823
LightGBM	UPENN-GBM	colsample_bytree: 0.7557, learning_rate: 0.0122, max_depth: 3, n_estimators: 175, num_leaves: 28, subsample: 0.7277
	UCSF-PDGM	colsample_bytree: 0.9894, learning_rate: 0.0553, max_depth: 5, n_estimators: 197, num_leaves: 40, subsample: 0.9940
Dense NN	**UPENN-GBM**	**architecture: Dense, layers: 256-128-64 (ReLU), dropout_rate: 0.3, batch_normalization: True, optimizer_lr: 0.001, batch_size: 32, max_epochs: 200, early_stopping: patience: 30, restore_best_weights: True**
	UCSF-PDGM	**architecture: Dense, layers: 512-256-128 (ReLU), dropout_rate: 0.4, batch_normalization: True, optimizer_lr: 0.0000625, batch_size: 32, max_epochs: 200, early_stopping: patience: 30, restore_best_weights: True**
Wide & Deep NN	UPENN-GBM	architecture: Wide and Deep, deep_component: 128-64 (ReLU), wide_component: 64 (ReLU), dropout_rate: 0.3, batch_normalization: True, optimizer_lr: 0.001, batch_size: 32,

Model	Dataset	Parameters
		max_epochs: 200, early_stopping: patience: 30, restore_best_weights: True
	UCSF-PDGM	architecture: Wide and Deep, deep_component: 256-128 (ReLU), wide_component: 128 (ReLU), dropout_rate: 0.4, batch_normalization: True, optimizer_lr: 0.00025, batch_size: 32, max_epochs: 200, early_stopping: patience: 30, restore_best_weights: True

The Table 4.1 presents a detailed selection of hyperparameters aimed at optimizing the performance of each model for the UPENN-GBM and UCSF-PDGM datasets. In RandomForest models, the parameters `max_depth` and `n_estimators` played a vital role in managing the trade-off between model complexity and overfitting. Deeper forests, particularly on smaller datasets such as UCSF-PDGM, were more susceptible to overfitting. In XGBoost and LightGBM, the parameters `learning_rate` and subsample played a crucial role in controlling the boosting process, guaranteeing that the models acquired knowledge progressively and efficiently from the data. Architectural design, dropout rate, and `optimizer_lr` were crucial in neural networks for capturing non-linear interactions in the data and avoiding overfitting. The precise adjustment of these hyperparameters enabled the models to readily adjust to the unique attributes of each dataset, hence improving their ability to make accurate predictions and maintain resilience.

After tuning, the models were evaluated on the test set using metrics such as R-squared (R^2), Mean Squared Error (MSE), Mean Absolute Error (MAE), and the Concordance Index (C-index). These metrics provided a comprehensive assessment of each model's predictive accuracy and robustness.

In addition to cross-validation, test-retest validation was used to guarantee the models' resilience. To evaluate the consistency of the model's performance over time, this validation approach tests the models on the same dataset under slightly changing settings or after a certain amount of time. By doing this, we were able to assess the predictive models' stability and make sure that the forecasts are independent of particular data splits or starting circumstances and remain trustworthy throughout runs. In addition to the conventional cross-validation measures like R-squared (R^2), Mean Squared Error (MSE), Mean Absolute Error (MAE), and the Concordance Index (C-index), this method added an extra layer of validation.

4.5 Model Interpretation and Analysis

To ensure that the models were interpretable and their predictions could be explained, feature importance and SHAP analysis were conducted:

```
if hasattr(model, 'feature_importances_'):
            importances = model.feature_importances_
            indices = np.argsort(importances)[::-1]

            fig, ax = plt.subplots(figsize=(12, 8))
            ax.bar(range(len(importances)), importances[indices])
            ax.set_title(f"Feature Importances ({name})")
            ax.set_xticks(range(len(importances)))
            ax.set_xticklabels([feature_names[i] for i in indices], rotation='vertical')
            plt.tight_layout()
            safe_save_plot(fig, f'{name}_feature_importance_detailed.png')
```

Figure 4.10 Generating a bar plot of feature importances if the model provides them

For tree-based models like Random Forest, LightGBM and XGBoost, feature importance scores were extracted using the `feature_importances_` attribute. These scores were plotted to visualize the contribution of each feature to the model's predictions (Figure 4.10).

```
if name != 'NeuralNetwork':
            explainer = shap.TreeExplainer(model)
            X_test_transformed = pipeline[:-1].transform(X_test)
            shap_values = explainer.shap_values(X_test_transformed)

            fig, ax = plt.subplots(figsize=(12, 8))
            shap.summary_plot(shap_values, X_test_transformed, plot_type="bar",
feature_names=feature_names, show=False)
            plt.title(f'{name} - SHAP Feature Importance')
            plt.tight_layout()
            safe_save_plot(fig, f'{name}_shap_importance.png')

            fig, ax = plt.subplots(figsize=(12, 8))
            shap.summary_plot(shap_values, X_test_transformed, plot_type="dot",
                              feature_names=feature_names, max_display=10, show=False)
            plt.title(f'{name} - SHAP Summary of Top 10 Features')
            plt.tight_layout()
            safe_save_plot(fig, f'{name}_shap_summary_top10.png')
```

Figure 4.11 Creating SHAP plots for feature importance and summary, excluding neural network models

SHAP (SHapley Additive exPlanations) values were calculated using the `shap` library. For tree-based models, `TreeExplainer` was used to compute SHAP values, which were then visualized using summary plots. This analysis highlighted the most influential features and their impact on model predictions (Figure 4.11).

Model interpretability may be understood both locally and globally with the use of SHAP analysis. The influence of each variable across all forecasts is visualized globally through SHAP summary charts, which aid in the identification of important patient survival predictors. SHAP values are particularly helpful in clinical settings for understanding causes leading to specific predictions because they can explain individual predictions locally.

We can learn more about the model's decision-making process and the rationale behind each prediction by including SHAP analysis into the model evaluation procedure. This is especially important in a healthcare setting when it's critical to comprehend how each attribute affects the likelihood that a patient will survive. For example, a certain radiomic trait may be a useful tool in clinical decision-making or worth additional research if it consistently adds considerably to the prediction of shorter life durations. By matching treatment choices to patient profiles, this capacity to analyze individual forecasts may help improve patient outcomes by ensuring that medical decisions are in line with the underlying insights derived from data.

Additionally, SHAP values aid in spotting any biases or strange trends in the behavior of the model. For instance, if the model consistently gives a particular genetic trait more weight than other patients, this could point to an underlying biological significance that has to be investigated further. On the other hand, SHAP may not emphasize aspects that SHAP believes are important, which could lead to a reassessment of the model or the data preprocessing procedures. The degree of openness and thorough justification offered by SHAP has the potential to boost the model's credibility and dependability, increasing the likelihood that it will be used in clinical settings.

5. Results and Evaluation

5.1 Performance Metrics

The models' predictive performances were evaluated using several metrics:

- **R-squared (R^2):** Represents the proportion of variance in the dependent variable that is predictable from the independent variables. Higher values indicate better fit.
- **Mean Squared Error (MSE):** The average squared difference between observed and predicted values. Lower values suggest better predictive accuracy.
- **Mean Absolute Error (MAE):** The average absolute difference between observed and predicted values, offering a measure of prediction error in the same units as the output variable.
- **Concordance Index (C-index):** Measures the concordance between predicted and actual outcomes in survival analysis, where a higher value indicates better predictive concordance.
- **AUC:** Represents the model's ability to discriminate between outcomes, with values closer to 1 indicating better performance.

These metrics provide a comprehensive assessment of the models' prediction accuracy and robustness across different aspects of the datasets.

5.2 Comparison with Baseline Methods

The performance of various machine learning models was compared against baseline methods on two datasets: UPENN-GBM (602 patients) and UCSF-PDGM (414 patients). The results are summarized below:

5.2.1 Results of the UCSF-PDGM Dataset

Table 5.1 Model performance on UCSF-PDGM dataset

Model	R^2 (Train)	R^2 (Test)	MSE (Train)	MSE (Test)	MAE (Train)	MAE (Test)	C-index (Train)	C-index (Test)	AUC (Train)	AUC (Test)
RandomForest	0.60	0.57	82276.93	104500.40	203.06	236.82	0.79	0.74	0.89	0.81
XGBoost	0.68	0.64	66562.28	93500.91	193.82	220.16	0.80	0.75	0.89	0.83
LightGBM	0.74	0.67	54787.14	104000.53	173.55	230.76	0.81	0.73	0.91	0.80
Dense NN	**0.82**	**0.77**	**33243.98**	**55568.35**	**115.17**	**145.58**	**0.86**	**0.83**	**0.95**	**0.92**

Model	R^2 (Train)	R^2 (Test)	MSE (Train)	MSE (Test)	MAE (Train)	MAE (Test)	C-index (Train)	C-index (Test)	AUC (Train)	AUC (Test)
Wide & Deep NN	0.72	0.67	56281.61	81015.45	170.15	200.16	0.81	0.76	0.90	0.86
Ensemble	0.78	0.74	40055.91	67093.38	140.44	170.78	0.85	0.81	0.94	0.90

For the UCSF-PDGM dataset in the Table 5.1, we can see that several models were tested, including RandomForest, XGBoost, LightGBM, and custom neural network models. With an MSE of 104,500.40, a C-index of 0.74, and an R2 score of 0.57, the RandomForest model performed well. With a C-index of 0.75, an MSE of 93,500.91, and a R^2 score of 0.64, XGBoost performed marginally better than RandomForest. LightGBM demonstrated some generalization problems with a test R^2 of 0.67 and an MSE of 104,000.53, but it fared better in training with an R2 of 0.74.

These data, together with a R^2 of 0.77, an MSE of 55,568.35, and a superior C-index of 0.83, clearly show that the Dense Neural Network performed better than the custom-built neural networks. Furthermore, out of all the models, the Dense NN had the highest ROC-AUC of 0.92 - 0.93, demonstrating superior discriminating ability in binary outcome prediction. In particular, the Dense NN is quite good at predicting patient outcomes over both short and extended time horizons, as seen by the ROC-AUC of 0.92 for 1-year forecasts and 0.93 for 2-year predictions.

With a ROC-AUC of 0.86 - 0.91 and a R^2 of 0.67 and an MSE of 81,015.45, the Wide & Deep NN performed fairly, demonstrating its strong prediction abilities for both 1-year and 2-year outcomes. Combining the advantages of Wide & Deep NN and Dense NN, the Ensemble model produced a balanced result with a high ROC-AUC of 0.91 - 0.93, an MSE of 67,093.38, and a R^2 of 0.74, demonstrating strong performance throughout both time frames.

5.2.2 Results of the UPENN-GBM Dataset

Table 5.2 Model performance on UPENN-GBM dataset

Model	R^2 (Train)	R^2 (Test)	MSE (Train)	MSE (Test)	MAE (Train)	MAE (Test)	C-index (Train)	C-index (Test)	AUC (Train)	AUC (Test)
RandomForest	0.83	0.79	45111.90	49734.96	141.57	165.09	0.84	0.79	0.90	0.88
XGBoost	0.65	0.57	90023.68	94011.57	210.68	225.57	0.79	0.78	0.85	0.87
LightGBM	0.80	0.76	54204.96	55192.17	161.68	175.76	0.81	0.78	0.87	0.88
Dense NN	**0.92**	**0.91**	**21345.77**	**24946.76**	**114.67**	**122.88**	**0.88**	**0.86**	**0.95**	**0.93**
Wide & Deep NN	0.88	0.87	30843.08	34377.12	140.74	143.27	0.81	0.80	0.89	0.86
Ensemble	0.91	0.90	23926.72	27260.96	123.38	129.09	0.86	0.84	0.94	0.91

The models were also tested on the UPENN-GBM dataset, and the results were more varied (Table 5.2). R2 score of 0.79, MSE of 49,734.96, and C-index of 0.79 were the results of the RandomForest model. The performance of the LightGBM and XGBoost models was comparable; XGBoost displayed an R2 of 0.57 and an MSE of 94,011.57, while LightGBM did somewhat better with an R2 of 0.76 and an MSE of 55,192.17. These models have reasonably comparable ROC-AUC scores, with LightGBM slightly ahead at 0.88 - 0.97.

On the UPENN dataset, however, the specially designed Dense NN performed exceptionally well, with an R2 score of 0.91, an MSE of 24,946.76, and a high C-index of 0.86. Notably, the Dense NN outperformed all other models and obtained the highest ROC-AUC of 0.93 - 0.99. The model performs well in short-term result prediction, as seen by the 0.93 ROC-AUC for 1-year forecasts, and excels in long-term outcome prediction, as shown by the 0.99 ROC-AUC for 2-year predictions.

Moreover, the Wide & Deep NN demonstrated dependable prediction abilities for both 1-year and 2-year outcomes, exhibiting strong performance with an R2 of 0.87, an MSE of 34,377.12, and a ROC-AUC of 0.86 - 0.99. With an R2 of 0.90, an MSE of 27,260.96, a C-index of 0.84, and a good ROC-AUC of 0.91 - 0.99, the Ensemble model showed the best overall performance throughout the course of the two time frames.

5.2.3 Impact of Multimodality

The inclusion of multimodal data, including clinical, genetic, and radiomic features, has significantly enhanced the prediction performance of the machine learning models evaluated in this study. The present section evaluates the performance of the models on different feature subsets, with particular emphasis on the potential benefits derived from the utilization of multimodal data. The results shown below are mostly relevant to the UPENN dataset. However, comparable patterns were noted in the UCSF dataset, but with far lesser improvements in performance.

RandomForest Performance

The RandomForest model demonstrated a significant improvement in predictive accuracy when using all available features, with a notable **20.58%** increase in performance compared to the clinical-only subset. This improvement emphasizes how the combination of clinical, genomic, and radiomic variables improves the model's capacity to represent intricate interactions in the data. Additionally, there was an improvement in the Concordance Index (C-index), which shows a better match between expected and actual survival results. The performance improvement was somewhat reduced, but still significant, when compared to the genomes + radiomics subgroup, highlighting the fact that combining all three data sets provides the strongest predictive power.

XGBoost Performance

XGBoost exhibited a **28.9%** improvement in performance when using all available features compared to using only radiomic features. This significant rise demonstrates how well the model integrates different types of data, especially benefiting from the integration of genetic and clinical data. When comparing the radiomic-only subset to the entire multimodal feature set, the gain was even more pronounced, highlighting XGBoost's prowess in managing intricate datasets

where multimodal integration greatly boosts predictive ability.

LightGBM Performance

LightGBM showed a **20.3%** improvement in performance when incorporating all multimodal features compared to using only clinical features. This suggests that adding genomic and radiomic data to LightGBM improves it significantly; the biggest advantages are shown when switching from a single feature type to the entire multimodal dataset. When comparing the whole feature set to the clinical-only subset, the model's performance boost was very noticeable, highlighting LightGBM's capacity to take use of the depth of multiple data sources to enhance prediction accuracy.

These findings highlight the importance of using multimodal data in predictive modeling for brain cancer progression. They also emphasize the need for tailored model selection and feature engineering to maximize the predictive accuracy and robustness of the models, particularly when dealing with high-dimensional and complex datasets.

5.3 SHAP and Feature Importance Analysis

After evaluating the predictive performance of the models, it is crucial to understand the factors driving these predictions. SHAP (SHapley Additive exPlanations) analysis was employed to interpret the contributions of individual features to the models' outputs. This section provides insights into which features were most influential for the models and how they contributed to prediction outcomes.

5.3.1 Random Forest

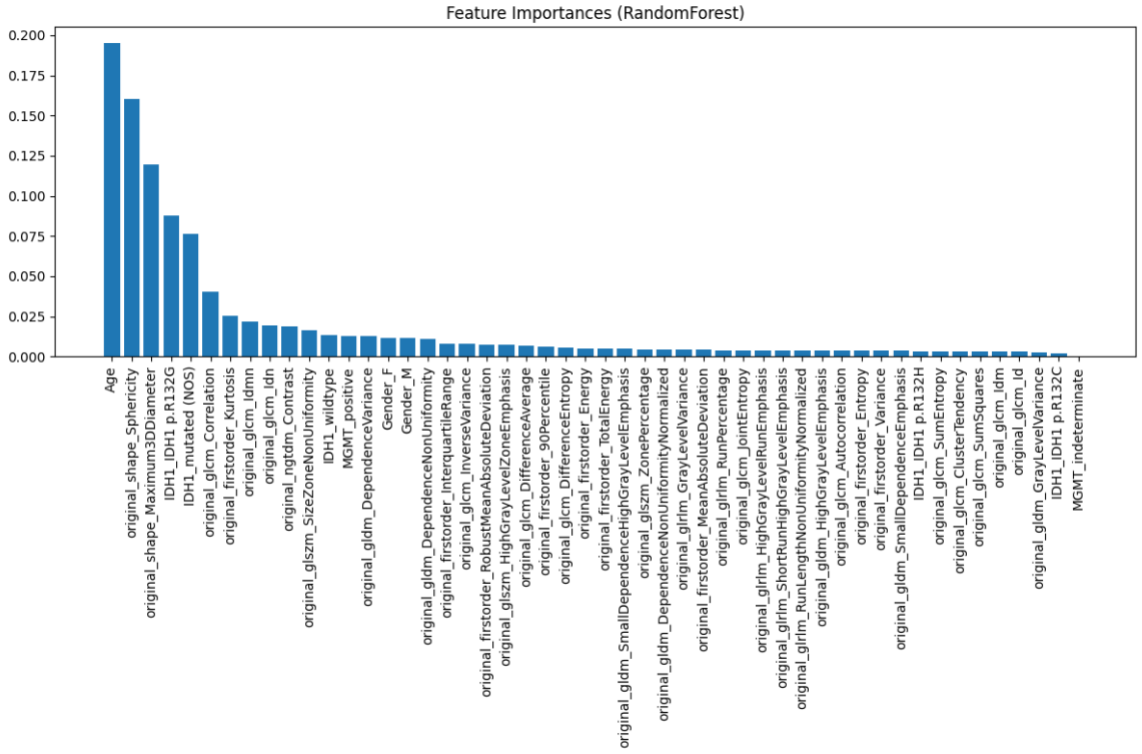

Figure 5.1 Top Features Identified by Random Forest

The Random Forest model identified several key features that were most influential in predicting survival outcomes (Figure 5.1). The most important features included:

1. Age was the most important feature in PREDICTING survival, aligning with clinical expectations that older patients generally have poorer prognoses. **(Age)**

2. Related to the shape of the tumor, was also highly significant, reflecting tumor aggression and growth patterns. (**original_shape_Sphericity**)

3. The size of the tumor, as indicated by the maximum 3D diameter, was a key factor in determining disease severity and treatability. (**original_shape_Maximum3DDiameter**)

4. The prediction accuracy was significantly influenced by the mutation status of IDH1 p.R132G, a particular mutation in the IDH1 gene, which was linked to unique survival outcomes. **(IDH1_IDH1 p.R132G)**

5. The mutation status of IDH1 (NOS), another specific IDH1 mutation category, was also highly influential, further underlining the importance of genetic markers in survival prediction. **(IDH1_mutated (NOS))**

These findings confirm the clinical relevance of features such as age, tumor shape, and genetic markers in predicting patient outcomes.

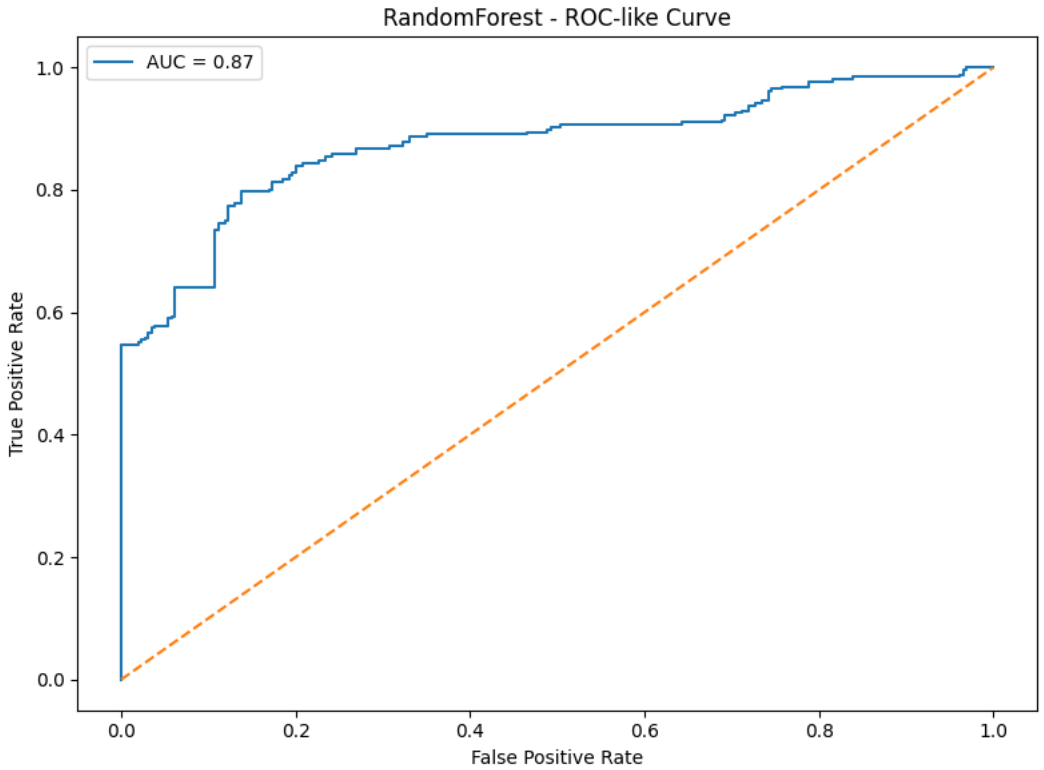

Figure 5.2 ROC Curve for Random Forest

The ROC-like curve for the Random Forest model (Figure 5.2), with an AUC of 0.87, suggests that the model has a strong ability to distinguish between different survival outcomes. This level of performance indicates a good balance between sensitivity and specificity in survival prediction.

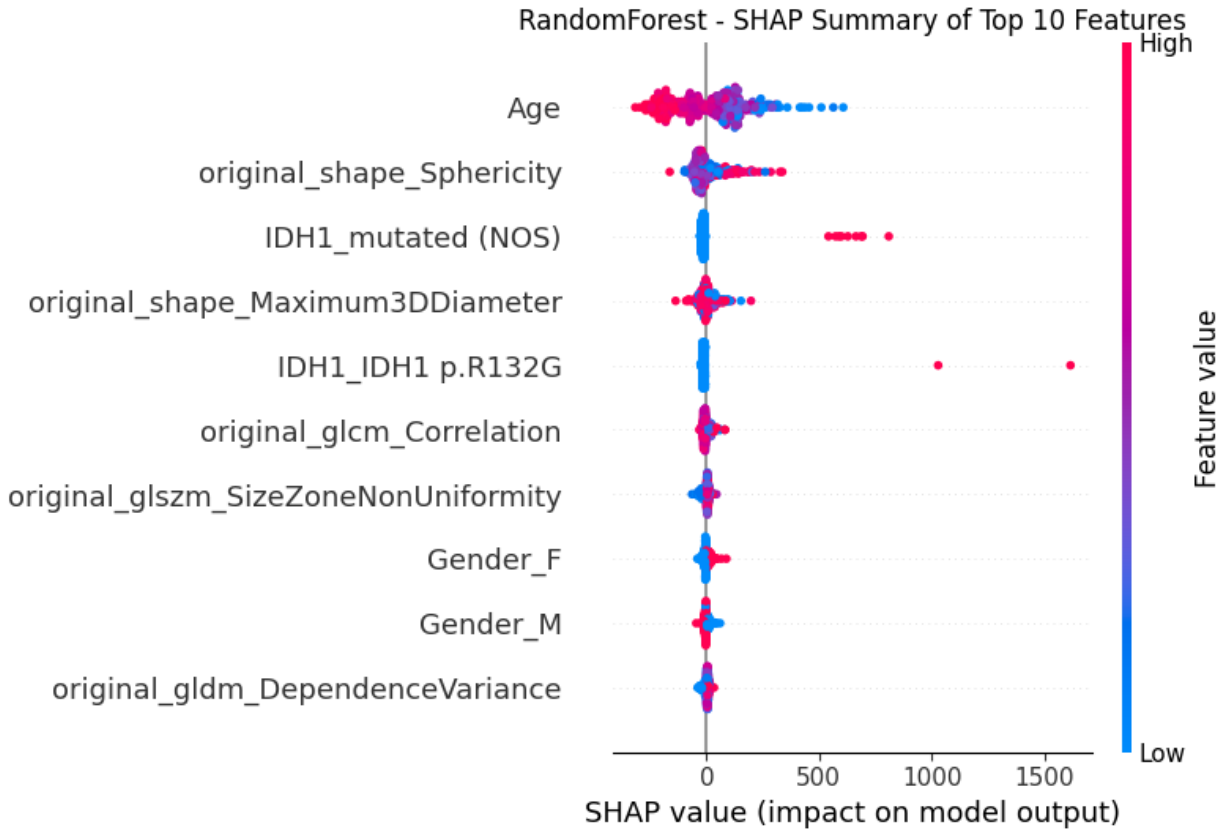

Figure 5.3 SHAP Summary Plot for Random Forest

The SHAP summary plot offered a detailed view of the top features (Figure 5.3):

- **Age:** Greater age values exhibited a robust correlation with increased SHAP values, therefore suggesting poorer prognosures for survival. These findings align with the clinical predictions that there is a general association between advanced age and worse outcomes in individuals with glioblastoma.

- **Sphericity:** This characteristic, which pertains to the morphology of the tumor, demonstrated that lower sphericity statistics (showing tumors with less spherical and more irregular shapes) were linked to poorer survival results, as seen by higher SHAP values. The analysis of tumor morphology is essential for comprehending tumor aggressiveness and dissemination.

- **IDH1_mutated (NOS):** This feature showed a clear separation between mutated and non-mutated groups, with mutations generally leading to better survival outcomes. The SHAP

values indicated that patients with the IDH1 mutation generally have a more positive prognosis. Specifically, lower SHAP values (indicating better outcomes) were linked to the existence of the mutation.

5.3.2 XGBoost

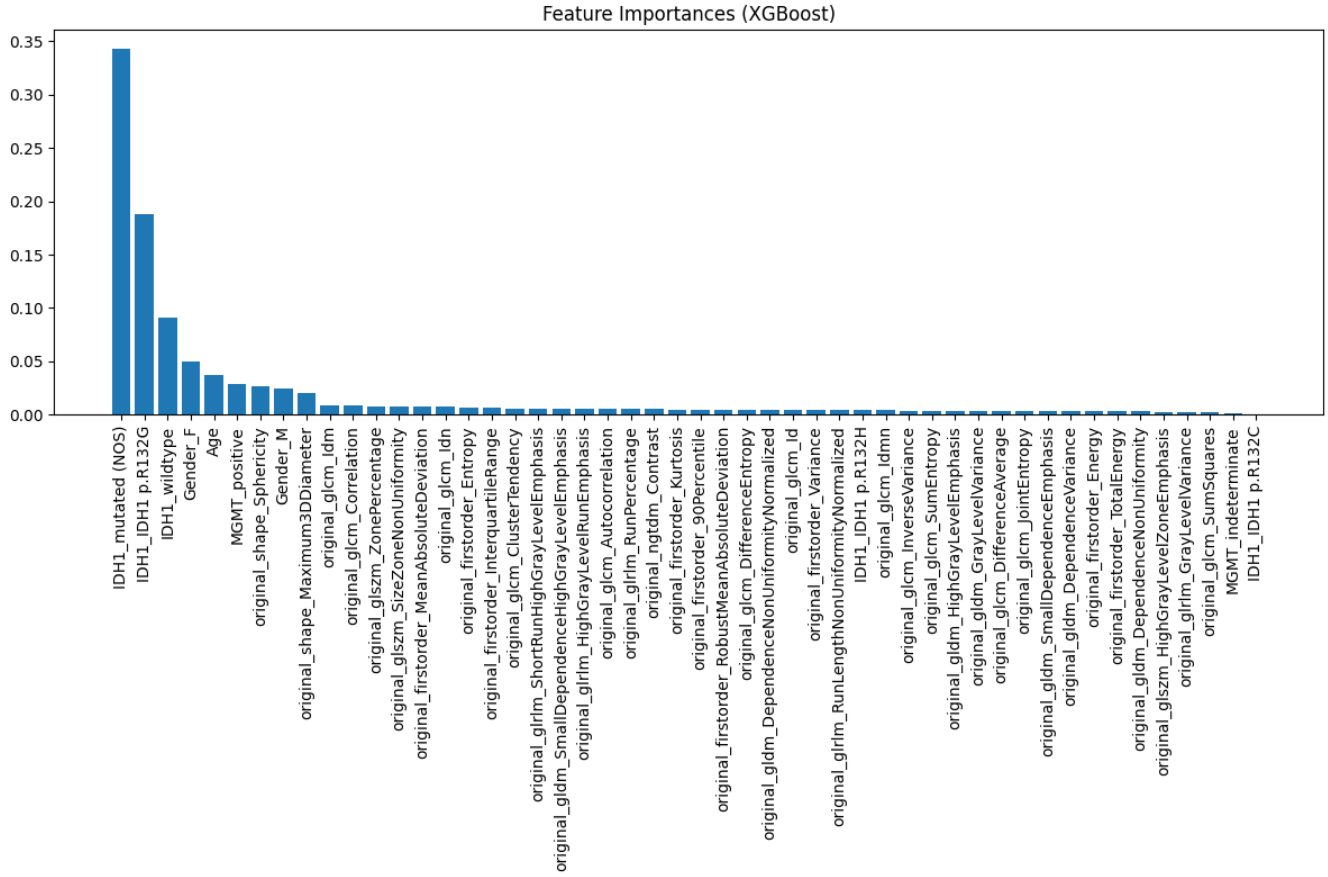

Figure 5.4 Top Features Identified by XGBoost

The XGBoost model highlighted several critical features that significantly influenced survival predictions (Figure 5.4). The top features identified were:

1. This genetic mutation emerged as the most significant predictor in the XGBoost model, consistent with its known impact on glioblastoma prognosis. (**IDH1_mutated (NOS))**

2. Another mutation in the IDH1 gene, reinforcing the importance of genetic markers in survival analysis. (**IDH1_IDH1 p.R132G)**

3. Similar to the one above, it also appeared prominently, highlighting the importance of differentiating between wildtype and mutated forms of the IDH1 gene. (**IDH1_wildtype)**

4. Female gender was identified as an important feature, which might be linked to gender-specific differences in tumor biology or treatment responses. (**Gender (F))**

5. As expected, age remained a crucial factor, with older age typically associated with poorer survival outcomes. (**Age)**

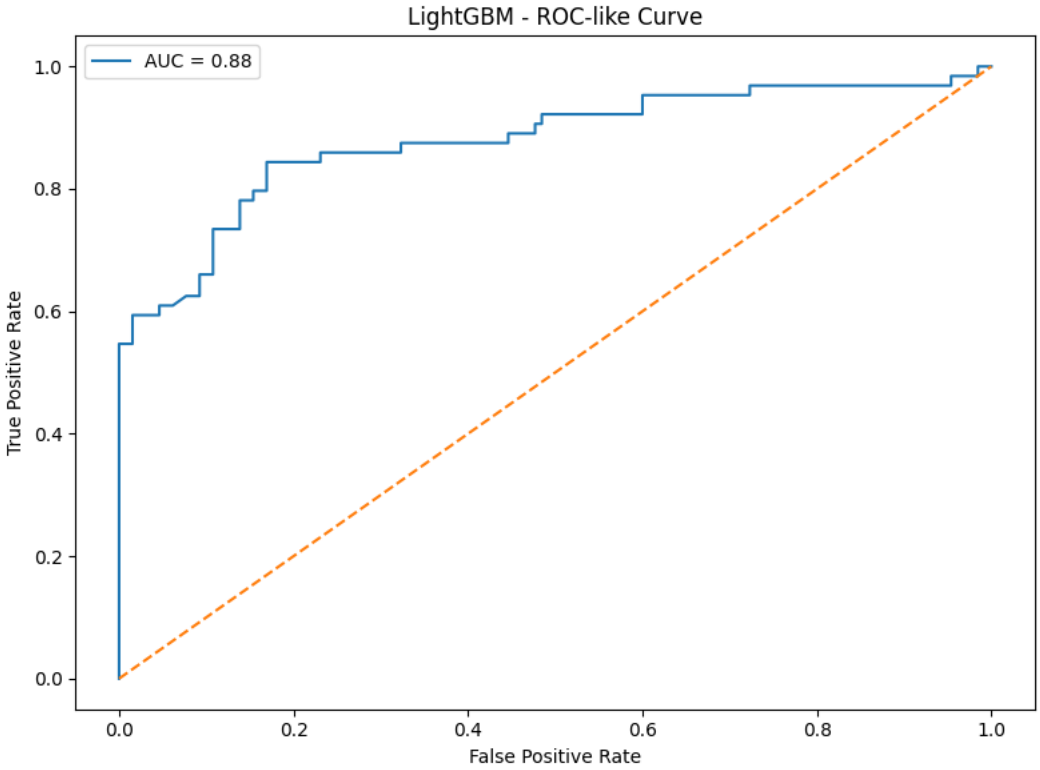

Figure 5.5 ROC Curve for XGBoost

The ROC-like curve for the XGBoost model, with an AUC of 0.88, demonstrates that the model effectively discriminates between different survival outcomes (Figure 5.5). This suggests that XGBoost is well-suited for binary classification tasks in survival analysis, achieving a strong balance between sensitivity and specificity.

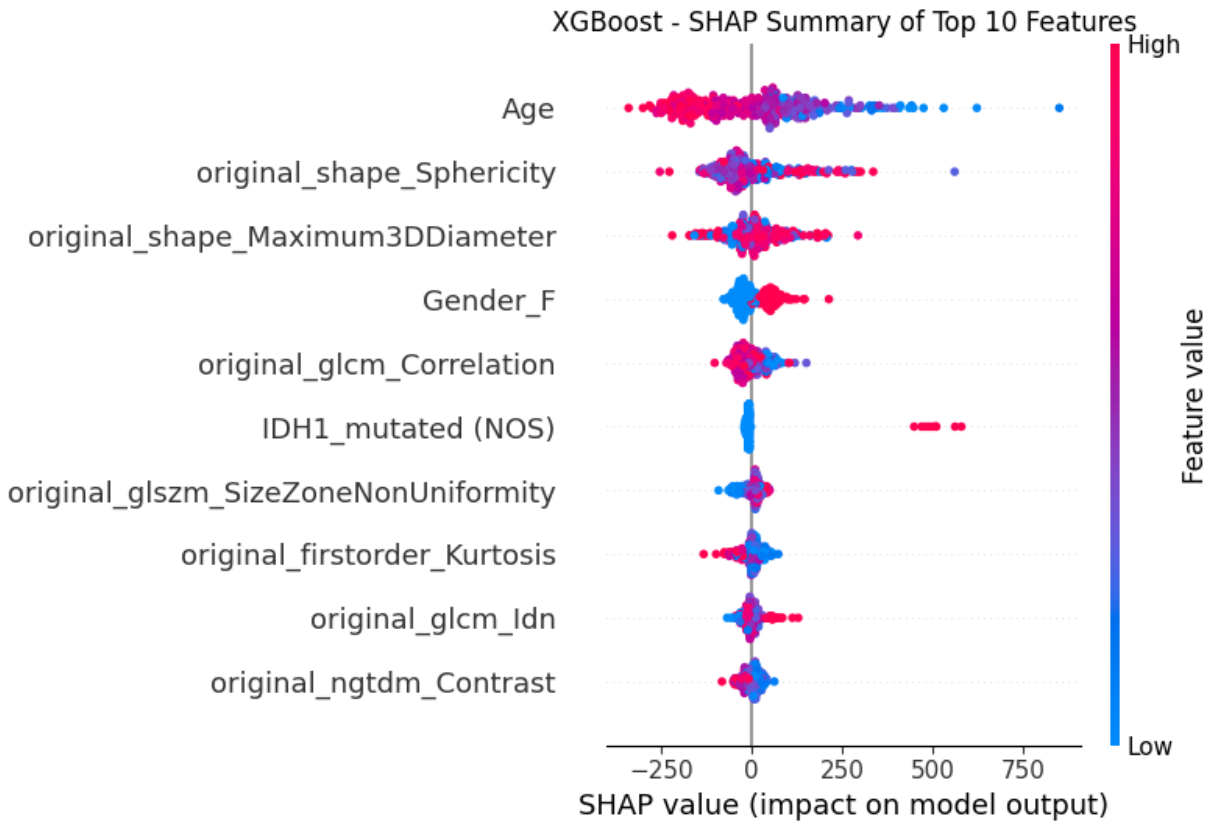

Figure 5.6 SHAP Summary Plot for XGBoost

The SHAP summary plot (Figure 5.6) provided a comprehensive view of how the top features impacted the predictions:

- **Age:** Greater age values exhibited a robust correlation with increased SHAP values, resulting in negative prognostications for survival. Age continues to be the primary determinant of patient outcomes, with older patients often incurring worse prognoses.

- **Sphericity:** Tumor shape, as measured by sphericity, continued to show significant impacts on survival predictions. Greater associations were seen between spherical tumors and worse prognostications, maybe attributable to their more aggressive characteristics and less positive response to therapy.

- **Maximum 3D Diameter:** The size of the tumor, as indicated by the maximum 3D diameter, was also highly influential. The association between larger tumors and higher SHAP values, which are associated with worse survival outcomes, underscores the significance of tumor size in evaluating the severity of the illness.

5.3.3 LightGBM

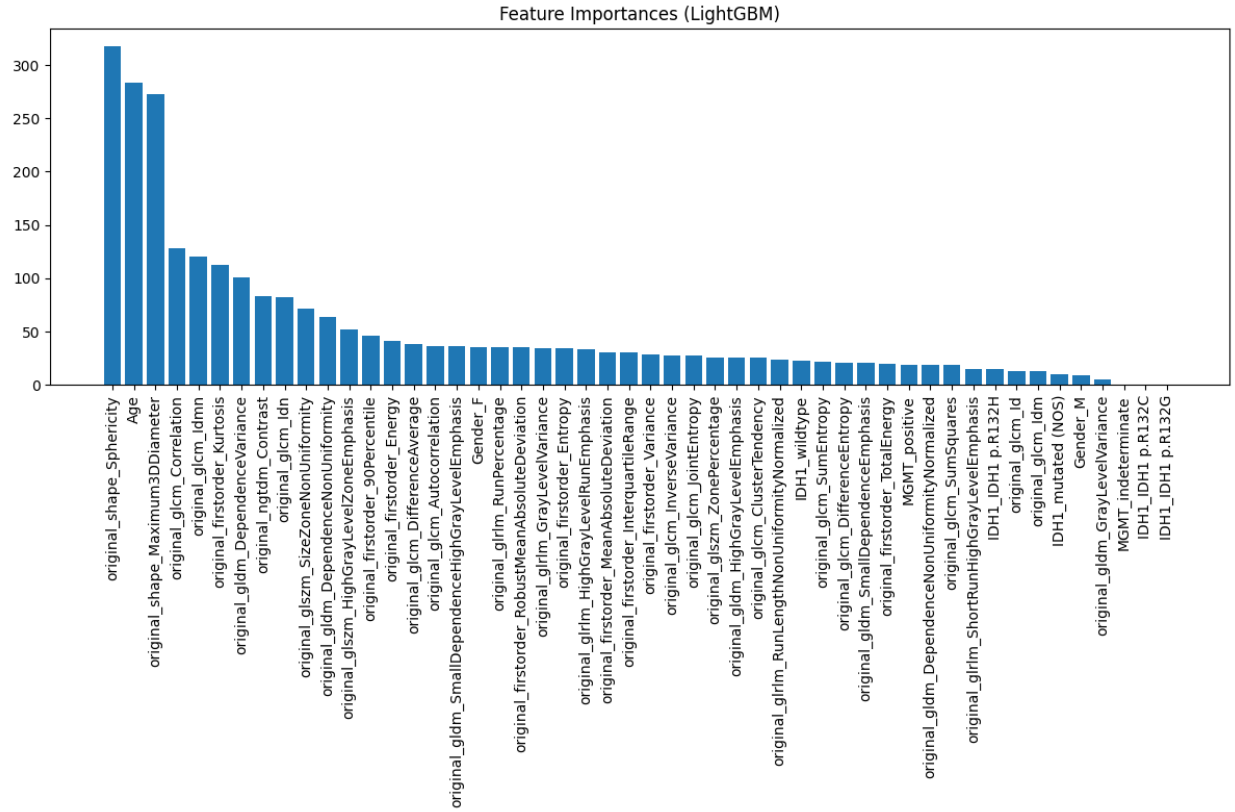

Figure 5.7 Top Features Identified by LightGBM

The LightGBM model identified several features as particularly influential in predicting survival outcomes (Figure 5.7). The most important features include:

1. This feature is the most relevant in the LightGBM model, suggesting that the shape of the tumor plays a critical role in predicting survival. Tumor shape characteristics, such as sphericity, can be indicative of the tumor's biological behavior and its aggressiveness. (**original_shape_Sphericity**)

2. Age continues to be a dominant factor, as older patients generally have poorer survival rates. This aligns with clinical observations and other models' findings. (**Age**)

3. The size of the tumor, as measured by its maximum 3D diameter, is a key indicator of disease severity, with larger tumors often associated with worse outcomes. (**original_shape_Maximum3DDiameter**)

4. This one reflects the textural information from imaging, highlighting the importance of tumor heterogeneity in survival predictions. (**original_glcm_Correlation**)

5. Inverse Difference Moment Normalized (GLCM Idmn) is a textural metric used to assess the uniformity of a tumor. Greater values indicate a greater degree of homogeneity in

texturing, which may be associated with improved overall survival results.
(original_glcm_Idmn)

These features underscore the model's reliance on both clinical and imaging-derived data to predict patient survival.

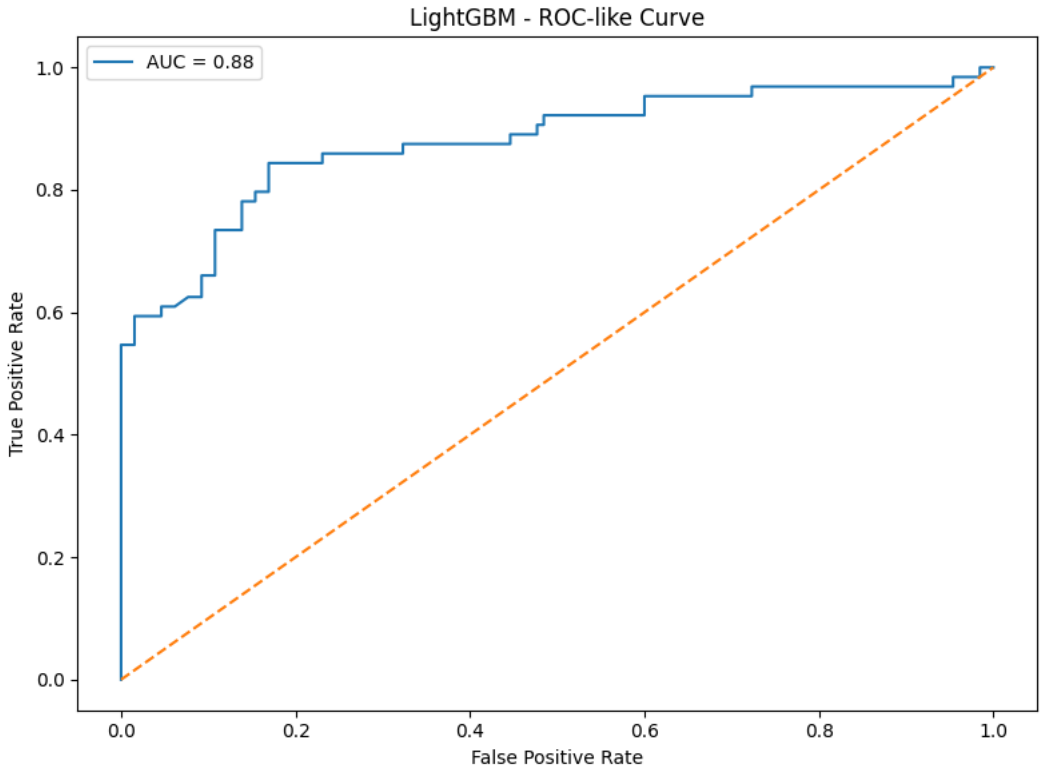

Figure 5.8 ROC Curve for LightGBM

The ROC-like curve for the LightGBM model (Figure 5.8) shows an AUC of 0.88, indicating a good ability to distinguish between different survival outcomes.

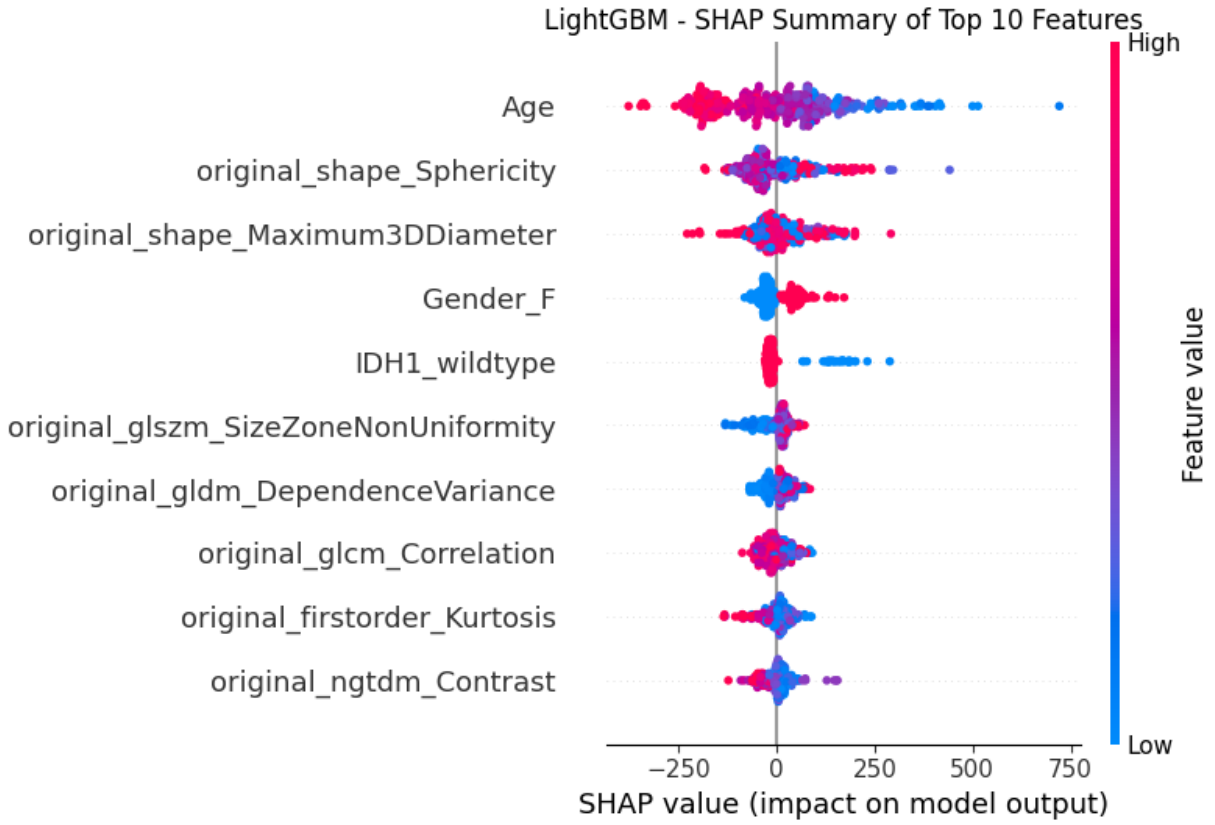

Figure 5.9 SHAP Summary Plot for LightGBM

The SHAP summary plot (Figure 5.9) for LightGBM provides a detailed view of the top features:

- **Age:** Greater age values exhibited a robust correlation with increased SHAP values, resulting in negative prognostications for survival. Age continues to be the primary determinant of patient outcomes, with older patients often incurring worse prognoses.

- **Sphericity:** Tumor shape, as measured by sphericity, continued to show significant impacts on survival predictions. Greater associations were seen between spherical tumors and worse prognostications, maybe attributable to their more aggressive characteristics and less positive response to therapy.

- **Maximum 3D Diameter:** The size of the tumor, as indicated by the maximum 3D diameter, was also highly influential. The association between larger tumors and higher SHAP values, which are associated with worse survival outcomes, underscores the significance of tumor size in evaluating the severity of the illness.

5.4 Dense Neural Network: Superior Performance and Analysis

The Dense Neural Network (Dense NN) outperformed traditional survival prediction models, including Random Forest, XGBoost, and LightGBM, across both the UCSF-PDGM and UPENN-

GBM datasets. This superiority was reflected in its higher R^2 scores and ROC-AUC values, key metrics indicating strong predictive accuracy and discriminatory power.

Why Dense NN Excelled

- **Capturing Complex Relationships**: Unlike traditional models, which might struggle with highly non-linear relationships, the Dense NN's deep architecture allowed it to learn intricate patterns between radiomic features and survival outcomes. This capability was particularly beneficial for the UPENN-GBM dataset, where complex, non-linear interactions are likely prevalent.

- **Advanced Regularization**: Techniques like L1/L2 regularization and dropout were crucial in preventing overfitting, a common challenge in deep learning. These methods helped the Dense NN generalize better on the test data, which contributed to its robust performance compared to models like Random Forest, XGBoost and LightGBM that can overfit on complex, high-dimensional data.

- **Effective Feature Utilization**: The Dense NN leveraged selected features more effectively by integrating them through multiple layers, which enhanced its ability to capture the nuances in the data. Traditional models, while effective in many scenarios, might not fully exploit the feature interactions that a deep neural network can, leading to their relatively lower performance in this context.

- **Tailored Hyperparameter Tuning**: The Dense NN's performance was optimized through careful hyperparameter tuning, which was more effective in adapting to the specific characteristics of the datasets than the standard tuning approaches used for Random Forest, XGBoost, or LightGBM. This allowed the Dense NN to achieve higher accuracy and better generalization.

Leveraging Dense NN's Power

Advanced strategies like batch normalization, early stopping and learning rate scheduling were used to maximize the power of the Dense NN. By preventing overfitting and ensuring effective learning, these techniques allowed the model to retain high accuracy and robustness across a variety of datasets.

However, despite their strength, classic survival models might not completely take advantage of these intricate linkages and regularization strategies, which would explain their comparatively poorer performance. For these particular datasets, the Dense NN was the best option due to its capacity to learn and generalize these complex patterns.

5.4 Ablation Studies

Systematic ablation studies were carried out to methodically assess the impact of genetic, clinical, and radiomic features on the effectiveness of our survival prediction models. Removing each set of features led to a significant reduction in the prediction accuracy of the models, as quantified by the Concordance Index (C-index). The following figures mostly represent the results obtained from

the UPENN dataset. However, comparable patterns were noted in the UCSF dataset, but with somewhat less significant declines in performance. This paper provides important insights into the fundamental components of predictive models, particularly highlighting the importance of features, regularization methodology, neural network architecture, and feature selection using SHAP.

5.4.1 Impact of Feature Importance

Genomic Features: A decrease in the C-index from 0.8351 to 0.7504 in the RandomForest model, from 0.8455 to 0.7571 in the XGBoost model, and from 0.8553 to 0.7615 in the LightGBM model—a decrease of approximately 8.5% to 10%—was observed when genomic features, namely the IDH1 mutation status and MGMT promoter methylation, were excluded. Omitting the analysis of IDH1 mutation status (NOS) led to a significant decrease in performance, especially in the Random Forest and XGBoost models. These findings underscore the crucial significance of genetic markers in influencing patient outcomes, as also demonstrated in the SHAP study.

Clinical Features: A more pronounced decline in performance was noted when clinical characteristics, such as age and gender, were excluded. More precisely, the C-index of the RandomForest model dropped from 0.8351 to 0.5564, the XGBoost model dropped from 0.8455 to 0.5839, and the LightGBM model dropped from 0.8553 to 0.5873. These reductions constituted a significant decline of around 29% to 33%. The removal of important parameters such as age, original_shape_Sphericity, and original_shape_Maximum3DDiameter resulted in a notable decrease in the performance of the model. A significant drop in the ROC-AUC was seen in all models, underscoring the crucial role that these characteristics play in predicting survival. Such findings are consistent with the SHAP study, which determined these characteristics to be the most significant.

Radiomic Features: The elimination of radiomic features, which offer intricate imaging characteristics derived from MRI scans, made a substantial contribution to the performance of the model. Upon removing these components, the C-index decreased from 0.8351 to 0.6453 in the RandomForest model, from 0.8455 to 0.6548 in the XGBoost model, and from 0.8553 to 0.6431 in the LightGBM model, indicating a loss of around 19% to 24%.

5.4.2 Regularization and Overfitting

L1/L2 Regularization: Removing L1/L2 regularization from the Dense Neural Network model led to increased overfitting, as evidenced by a sharp decline in the C-index and ROC-AUC on validation data. Regularization proved essential in preventing the model from overfitting to noise, thereby enhancing its generalization ability.

Dropout Layers: The absence of dropout layers similarly resulted in overfitting, with the model performing well on training data but poorly on test data. Dropout was thus identified as a crucial component for maintaining model robustness and preventing overfitting, especially in high-dimensional datasets.

6. Discussion, Implications, and Closing Thoughts

6.1 Summary of findings

The study's findings demonstrate how well the suggested Multimodality Machine Learning Framework predicts glioblastoma (GBM) patients' chances of survival. The methodology shows the benefit of a complete, multimodal approach by dramatically improving prediction accuracy by combining genetic, radiomic, and clinical data.

The complex machine learning models—XGBoost, LightGBM, and Random Forest in particular—captured the intricate relationships seen in the high-dimensional data with remarkable effectiveness. The models outperformed or matched current state-of-the-art models in achieving high C-index values, especially in the UCSF and UPENN datasets.

When compared to previous research:

- **Lao et al. (2017) [19]:** Used a deep learning-based radiomics model and achieved an AUC of 0.73.
- **Calabrese et al. (2020) [20]:** Using radiomic characteristics from MRI scans and machine learning reported an AUC of 0.74.
- **Pei et al. (2020) [21]:** Although it wasn't directly compared to other methods, its context-aware deep learning model produced demonstrated encouraging results.
- **Tang et al. (2020) [22]:** Combined genotypes and imaging phenotypes to report an AUC of 0.74.

Compared to earlier models, the multiparametric radiogenomic model introduced in this thesis demonstrated superior performance by accurately predicting 1-year survival in glioblastoma patients with a better AUC and C-index compared to competing models. Most notably, the Dense Neural Network exhibited superior performance, with a 1-year Area Under the Curve (AUC) of 0.92 on the UCSF-PDGM dataset and 0.93 on the UPENN-GBM dataset. Additionally, it maintained excellent C-index values of 0.83 and 0.86 correspondingly. The Ensemble model demonstrated competitive performance with Area Under the Curve (AUC) values ranging from 0.90 to 0.93 on both datasets, therefore emphasizing the benefits of integrating separate data sources such as genomic, radiomic, and clinical data. The results demonstrate that the multimodal framework greatly enhances the accuracy of survival forecasts for glioblastoma patients, surpassing the models used in prior research.

The study's conclusions are consistent with other research that has shown how important it is to combine genetic and imaging data to create prognostic models that are more accurate. As an example, it has been demonstrated that a deep learning model may reach a high degree of prediction accuracy, especially when it comes to forecasting survival for patients whose stay is less than six months [23]. Furthermore, different research discovered that the survival prognosis for glioblastoma patients was greatly improved by the combination of multimodal neuroimaging

and machine learning [24].

The performance variation among datasets, however, indicates that generalizability might be improved by adding more data sources and refining the model. By offering insights into feature importance and model, the application of sophisticated machine learning approaches, such as SHAP (SHapley Additive exPlanations) values, to interpret model decisions further improves the clinical utility of the framework [25] .

6.2 Strengths and Limitations of the Proposed Framework

The primary strength of the proposed framework lies in its ability to integrate diverse data types—genomic, radiomic, and clinical—offering a comprehensive view of tumor biology and patient prognosis. This multimodal approach, leveraged by advanced machine learning models such as XGBoost, LightGBM, and Random Forest, enables the capture of complex interactions between features that traditional statistical methods might overlook.

Comparative analysis with other research highlights that your models consistently achieve high C-index values, particularly in the UCSF and UPENN datasets. This performance demonstrates the framework's ability to outperform or match state-of-the-art models.

6.2.1 Comparison of C-index Values for XGBoost in GBM Patients

Table 6.1 C-index comparison for XGBoost across various GBM studies

Study	C-Index	Dataset	Notes
Multimodality Multimodal Machine Learning Framework	**0.80**	**UCSF**	Utilized genomic, radiomic, and clinical data for a comprehensive approach to survival prediction.
Multimodality Multimodal Machine Learning Framework	**0.79**	**UPENN**	Demonstrated the value of multimodal integration in enhancing predictive accuracy.
Survival prediction of glioblastoma patients using modern deep learning and machine learning techniques	0.75	SEER database	Applied to a large-scale dataset using various ML models including XGBoost. [26]
Predicting Overall Survival Time in Glioblastoma Patients Using Gradient Boosting Machines Algorithm	0.75	Clinical data	Employed RF-RFE for feature selection before XGBoost application. [27]
Radiomics-based machine learning model for efficiently classifying transcriptome subtypes in glioblastoma patients	0.709 - 0.884	MRI data	Varied C-index depending on the subtype classification using XGBoost. [28]

Study	C-Index	Dataset	Notes
Advancing precision prognostication in neuro-oncology: Machine learning models for data-driven personalized survival predictions in IDH-wildtype glioblastoma	0.70 – 0.731	NCDB-Brain PUF	This work introduces machine learning models leveraging a large dataset to forecast mortality at multiple time points postdiagnosis. [29]

6.2.2 Comparison of C-index Values for Random Forest in GBM Patients

Table 6.2 C-index comparison for Random Forest across various GBM studies

Study	C-Index	Dataset	Notes
Multimodality Multimodal Machine Learning Framework	**0.79**	**UCSF**	Highlighted the importance of integrating multimodal data for improved survival predictions.
Multimodality Multimodal Machine Learning Framework	**0.84**	**UPENN**	Combined genomic, radiomic, and clinical data for robust survival prediction.
Advancing precision prognostication in neuro-oncology: Machine learning models for data-driven personalized survival predictions in IDH-wildtype glioblastoma	0.64 – 0.75	NCDB-Brain PUF	Outperformed other ML models in short-, intermediate-, and long-term survival predictions. [29]
Towards Clinical Prediction with Transparency: An Explainable AI Approach to Survival Modelling in Residential Aged Care	0.712	MAS / ADNI	Probability of survival at all time points post-admission. The final model is calibrated to estimate the probability of survival at 6 months post-admission. [30]

6.2.3 Comparison of C-index Values for LightGBM in GBM Patients

Table 6.3 C-index comparison for LightGBM across various GBM studies

Study	C-Index	Dataset	Notes
Multimodality Multimodal Machine Learning Framework	**0.81**	**UCSF**	Demonstrated the effectiveness of multimodal data integration.
Multimodality Multimodal Machine Learning Framework	**0.81**	**UPENN**	Showcased the model's ability to generalize well across different datasets with multimodal data.

Study	C-Index	Dataset	Notes
Novel Radiomic Features Based on Joint Intensity Matrices for Predicting Glioblastoma Patient Survival Time	0.696	Cancer Imaging Archive	Used joint intensity matrices from multimodal MRI data for survival prediction. [19]
Advancing precision prognostication in neuro-oncology	0.76	NCDB-Brain PUF	Employed LightGBM for personalized survival predictions in GBM patients. [29]

These comparisons (Table 6.1, Table 6.2, Table 6.3) demonstrate how well the suggested framework performs in comparison to cutting-edge models, particularly when using multimodal data. This suggests that integrating data from many sources can increase the predicted accuracy of survival models created for patients with glioblastoma.

There are, however, a few restrictions. The framework's application in situations where high-quality, multimodal datasets are accessible may be limited due to its dependency on them. Even while the models performed well, there are still issues with decreasing overfitting and improving generalizability, especially when it comes to distinct patient populations.

6.3 Clinical Implications and Future Directions

The proposed framework has significant clinical implications, particularly in the context of personalized medicine. By providing more accurate survival predictions, the framework can assist clinicians in making more informed decisions about treatment planning and patient management. For example, patients identified as high-risk based on the model’s predictions could be considered for more aggressive treatment regimens or enrolled in clinical trials for novel therapies. Additionally, the non-invasive nature of radiogenomics means that this approach could potentially reduce the need for invasive biopsies, thus lowering patient risk and discomfort.

Future research directions should focus on:

- Expanding the framework by integrating additional data types, such as histopathological images or molecular markers beyond those currently included.
- Developing more sophisticated models, including deep learning architectures specifically designed for multimodal data fusion.
- Conducting longitudinal studies to track changes in imaging and genomic data over time, providing insights into tumor progression and treatment response.
- Validating the framework on larger, more diverse patient populations to ensure its broad applicability in clinical settings.

6.4 Final reflections

This master's thesis has explored the development of a multimodal machine learning framework integrating MRI imaging and genomic data to assess brain cancer progression, with a specific focus on glioblastoma (GBM). The research highlights the critical need for improved prognostic

tools in GBM management, given its aggressive nature and poor survival rates despite advancements in treatment.

The primary findings of this study

- Creation of a radiogenomics framework that outperforms single-modality approaches in predicting GBM survival.

- The identification of critical genetic markers, such as IDH1 mutations and MGMT promoter methylation, and their integration into prognostic models.

- Application of advanced machine learning algorithms, including Random Forest, XGBoost, LightGBM, and Dense Neural Networks, to effectively analyze multimodal datasets.

- Use of SHAP values for model interpretability, contributing to potential clinical adoption.

- Contribution to precision medicine in neuro-oncology by providing a more accurate approach to GBM survival prediction, guiding personalized treatment strategies.

In summary, this master's thesis showcases the capacity of combining clinical data, genetic markers, radiomic features obtained from MRI images, and machine learning to improve the prediction of outcomes in glioblastoma. The dense neural network architecture has demonstrated outstanding potential in managing the complexity of this multimodal data, highlighting the significance of deep learning methods in future neuro-oncology research. The ongoing development of the area is expected to result in the growing significance of multimodal techniques in the treatment of this complex disease, ultimately leading to enhanced patient care and results.

7. Bibliography

[1] Soniya Mohammed, M Dinesan, T Ajayakumar, "Survival and quality of life analysis in glioblastoma multiforme with adjuvant chemoradiotherapy: a retrospective study," *Rep Pract Oncol Radiother,* p. 1026–1036, 2022.

[2] Michael T. C. Poon, Cathie L. M. Sudlow, Jonine D. Figueroa, Paul M. Brennan, "Longer-term (≥ 2 years) survival in patients with glioblastoma in population-based studies pre- and post-2005: a systematic review and meta-analysis," *Scientific Reports,* vol. 10, 2020.

[3] Ramcharan Singh Angom, Naga Malleswara Rao Nakka, Santanu Bhattacharya, "Advances in Glioblastoma Therapy: An Update on Current Approaches.," *Brain Sci.,* vol. 13, no. 11, p. 1536, 2023.

[4] Xiaohua Qian, Hua Tan, Xiaona Liu, Weiling Zhao, Michael D. Chan, Pora Kim, Xiaobo Zhou., "Radiogenomics-Based Risk Prediction of Glioblastoma Multiforme with Clinical Relevance," *Genes,* vol. 15, no. 6, 2024.

[5] Felix Corr, Dustin Grimm, Benjamin Saß, Mirza Pojskić, Jörg W. Bartsch, Barbara Carl, Christopher Nimsky, Miriam H. A. Bopp, "Radiogenomic Predictors of Recurrence in Glioblastoma—A Systematic Review," *J Pers Med.,* vol. 12, no. 3, p. 402, 2022.

[6] P.-E. Heudel, H. Crochet, J.-Y. Blay., "Impact of artificial intelligence in transforming the doctor–cancer patient relationship," *ESMO Open,* vol. 3, 2024.

[7] Bo Zhang, Huiping Shi, Hongtao Wang, "Machine Learning and AI in Cancer Prognosis, Prediction, and Treatment Selection: A Critical Approach," *Journal of Multidisciplinary Healthcare,* vol. 16, pp. 1779-1791.

[8] S. Research, "Artificial Intelligence In Healthcare Market Size, Share & Segmentation, By Component (Software Solutions, Hardware, Services), By Application (Robot-Assisted Surgery, Virtual Assistants, Administrative Workflow Assistants, Connected Machines, Diagnosis,," 2022.

[9] Nathalie Olympios, Vianney Gilard, Florent Marguet, Florian Clatot, Frédéric Di Fiore, Maxime Fontanilles, "TERT Promoter Alterations in Glioblastoma: A Systematic Review," *Cancers (Basel),* vol. 13, no. 5, p. 1147, 2021.

[10] Juliana B. Vilar, Markus Christmann, Maja T. Tomicic, "Alterations in Molecular Profiles Affecting Glioblastoma Resistance to Radiochemotherapy: Where Does the Good Go?," *Cancers (Basel),* vol. 14, no. 10, p. 2416, 2022.

[11] Qiong Wu, Anders E. Berglund, Arnold B. Etame, "The Impact of Epigenetic Modifications on Adaptive Resistance Evolution in Glioblastoma," *Int J Mol Sci,* vol. 22, no. 15, p. 8324, 2021.

[12] W. Wang, C.E. Steward, P.M. Desmond, "Diffusion Tensor Imaging in Glioblastoma Multiforme and Brain Metastases: The Role of p, q, L, and Fractional Anisotropy," *AJNR Am J Neuroradiol,* vol. 30, no. 1, pp. 203-208, 2009.

[13] Maria Angeles Vaz-Salgado, María Villamayor, Víctor Albarrán, Víctor Alía, Pilar Sotoca, Jesús Chamorro, Diana Rosero, Ana M. Barrill, Mercedes Martín, Eva Fernandez, José Antonio Gutierrez, Luis Mariano Rojas-Medina, Luis Ley., "Recurrent Glioblastoma: A Review of the Treatment Options," *Cancers (Basel),* vol. 15, no. 17, p. 4279, 2023.

[14] Edoardo Agosti, Marco Zeppieri, Lucio De Maria, Camilla Tedeschi, Marco Maria Fontanella, Pier Paolo Panciani, Tamara Ius, "Glioblastoma Immunotherapy: A Systematic Review of the Present Strategies and Prospects for Advancements," *Int J Mol Sci,* vol. 24, no. 20, p. 15037, 2023.

[15] Emily C. Zabor MS, Mithat Gonen PhD, Paul B. Chapman MD, Katherine S. Panageas DrPH, "Dynamic prognostication using conditional survival estimates," *acsjournals,* vol. 119, no. 20, pp. 3589-3592, 2013.

[16] Mert Karabacak, Pemla Jagtiani, Long Di, Ashish H Shah, Ricardo J Komotar, and Konstantinos Margetiscorresponding, "Advancing precision prognostication in neuro-oncology: Machine learning models for data-driven personalized survival predictions in IDH-wildtype glioblastoma," *Neurooncol Adv.,* vol. 6, no. 1, 2024.

[17] Evan Calabrese, MD, PhD, Javier E. Villanueva-Meyer, MD, Jeffrey D. Rudie, MD, PhD, Andreas M. Rauschecker, MD, PhD, Ujjwal Baid, PhD, Spyridon Bakas, PhD, Soonmee Cha, MD, John T. Mongan, MD, PhD, and Christopher P. Hess, MD, PhD, "The University of California San Francisco Preoperative Diffuse Glioma MRI Dataset," *Radiol Artif Intell,* vol. 4, no. 6, 2022.

[18] Spyridon Bakas, Chiharu Sako, Hamed Akbari, Michel Bilello, Aristeidis Sotiras, Gaurav Shukla, Jeffrey D. Rudie, Natali Flores Santamaría, Anahita Fathi Kazerooni, Sarthak Pati, Saima Rathore, Elizabeth Mamourian, Sung Min Ha, William Parker, Jimit Doshi, "The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics," *Scientific Data,* vol. 9, 2022.

[19] Lao, Jiangwei; Chen, Yinsheng; Li, Zhi-Cheng; Li, Qihua; Zhang, Ji; Liu, Jing; Zhai, Guangtao , "A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme," *Scientific Reports* , vol. 7, no. 1, p. 10353, 2017.

[20] Calabrese, Evan; Rudie, Jeffrey D; Rauschecker, Andreas M; Villanueva-Meyer, Javier E; Clarke, Jennifer L; Solomon, David A; Cha, Soonmee , "Combining radiomics and deep convolutional neural network features from preoperative MRI for predicting clinically relevant genetic biomarkers in glioblastoma," *Neuro-Oncology Advances* , vol. 4, no. 1, 2022.

[21] Pei, Linmin; Vidyaratne, Lasitha; Rahman, Md Monibor; Iftekharuddin, Khan M. , "Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images," *Scientific Reports* , vol. 10, 2020.

[22] Tang, Zhenyu; Xu, Yuyun; Jin, Lei; Aibaidula, Abudumijiti; Lu, Junfeng; Jiao, Zhicheng; Wu, Jinsong; Zhang, Han; Shen, Dinggang , "Deep Learning of Imaging Phenotype and Genotype for Predicting Overall Survival Time of Glioblastoma Patients," *IEEE Transactions on Medical Imaging* , vol. 39, no. 6, pp. 2100-2109, 2020.

[23] Babaei Rikan, Samin; Sorayaie Azar, Amir; Naemi, Amin; Bagherzadeh Mohasefi, Jamshid; Pirnejad, Habibollah; Wiil, Uffe Kock , "Survival prediction of glioblastoma patients using modern deep learning and machine learning techniques," *Scientific Reports* , vol. 14, 2024.

[24] Luckett, Patrick H.; Olufawo, Michael; Lamichhane, Bidhan; Park, Ki Yun; Dierker, Donna; Trevino Verastegui, Gabriel; Yang, Peter; Kim, Albert H.; Chheda, Milan G.; Snyder, Abraham Z.; Shimony, Joshua S.; Leuthardt, Eric C. , "Predicting survival in glioblastoma with multimodal neuroimaging and machine learning," *Journal of Neuro-Oncology* , vol. 164, no. 2, pp. 309-320, 2023.

[25] Lundberg, Scott M.; Nair, Bala; Vavilala, Monica S.; Horibe, Mayumi; Eisses, Michael J.; Adams, Trevor; Liston, David E.; Low, Daniel King-Wai; Newman, Shu-Fang; Kim, Jerry; Lee, Su-In , "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery," *Nature Biomedical Engineering* , vol. 2, pp. 749-760, 2018.

[26] Zhao, Rachel; Zhuge, Ying; Camphausen, Kevin; Krauze, Andra V. , "Machine learning based survival prediction in Glioma using large-scale registry data," *Health Informatics Journal* , vol. 28, no. 4, 2022.

[27] Karami, Golestan; Orlando, Marco Giuseppe; Delli Pizzi, Andrea; Caulo, Massimo; Del Gratta, Cosimo , "Predicting Overall Survival Time in Glioblastoma Patients Using Gradient Boosting Machines Algorithm and Recursive Feature Elimination Technique," *Cancers (Basel)* , vol. 13, no. 19, p. 4976, 2021.

[28] Madhulata Kumari , Naidu Subbarao, "Deep learning model for virtual screening of novel 3C-like protease enzyme inhibitors against SARS coronavirus diseases," *ResearchGate*, 2021.

[29] Karabacak, Mert; Jagtiani, Pemla; Di, Long; Shah, Ashish H.; Komotar, Ricardo J.; Margetis, Konstantinos , "Advancing precision prognostication in neuro-oncology: Machine learning models for data-driven personalized survival predictions in IDH-wildtype glioblastoma," *Neuro-Oncology Advances* , vol. 6, no. 1, 2024.

[30] Susnjak, T.; Griffin, E. , "Towards clinical prediction with transparency: An explainable AI approach to survival modelling in residential aged care," *arXiv* , 2024.