Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1A.1249

# Multimodal Machine Learning-Based Cancer Progression Prediction from Plain Radiographs and Clinical Data

Hassan Umar and Ali Ahmad Aminu

Department of Computer Science, Faculty of Science, University of Gombe, PMB 127 Tudun-Wada Gombe, Nigeria.

## ABSTRACT

This research investigates the use of a multimodal machine learning model to predict cancer progression by integrating radiographs and clinical data. The study addresses the limitations of unimodal approaches, which often overlook the synergistic potential of combining diverse data types. By leveraging deep learning techniques for image analysis and interpretable models for clinical data, the proposed framework enhances prediction accuracy and model interpretability. The multimodal model achieved a high training accuracy of 98.01% and a testing accuracy of 94%, significantly outperforming unimodal models like SVM and CNN. Precision (94.2%) and recall (94%) highlighted the model's ability to accurately identify true positive cases, while the AUC-ROC of 98% underscored its robust diagnostic capability. Comprehensive evaluation demonstrated that the multimodal model effectively integrates complementary data, improving predictive performance and supporting personalised treatment planning. The research contributes to advancing cancer diagnosis and prognosis, offering a promising tool for clinical decision-making.

**Keywords:** Multimodal Machine Learning, Cancer Progression Prediction, Plain Radiographs, Clinical Data, Data Integration, Predictive Modelling.

## INTRODUCTION

In the realm of oncology, the accurate prediction of cancer progression stands as a pivotal endeavor, captivating clinicians, researchers, and patients alike. Despite considerable strides in diagnosis and treatment, cancer remains a formidable adversary in modern medicine, exacting a heavy toll on human life and emphasizing the urgency for innovative approaches to disease management (Siegel et al., 2023). At the forefront of this quest lies the fusion of multimodal machine learning techniques with clinical data, offering a promising avenue for revolutionizing cancer care (Esteva et al., 2019).

Cancer, with its multifaceted nature and diverse manifestations, presents an ongoing challenge for healthcare systems worldwide (Hanahan & Weinberg, 2011). While plain radiographs serve as a cornerstone of diagnostic imaging, providing invaluable insights into tumor morphology and progression, clinical data encompassing patient demographics, medical history, and treatment regimens offer a comprehensive view of the disease trajectory (Fleischmann et al., 2017). However, realizing the full potential of these modalities for predictive modeling necessitates a paradigm shift towards the integration of advanced computational methodologies.

Despite the inherent value of multimodal data, current approaches often fall short of exploiting their synergistic potential, leading to suboptimal prediction performance (Wang et al., 2019). Previous efforts have primarily focused on either imaging-based or clinical data-driven models, neglecting the intricate interplay between these modalities (Jiang et al., 2020). While some studies have explored ensemble methods to combine predictions

DOI: 10.56892/bima.v9i1A.1249

from individual models, they often encounter challenges in data fusion and model interpretation, limiting their utility in clinical settings.

Moreover, existing models may suffer from issues such as overfitting, limited generalizability across different cancer types or patient populations, and insufficient consideration of temporal. In response to these challenges, this study proposes novel framework for cancer progression prediction that integrates multimodal machine learning techniques with advanced data fusion strategies (Hosseini et al., 2021). This approach combines deep learning architectures tailored for image analysis with interpretable models for clinical data integration, allowing for a holistic understanding of disease progression dynamics. By leveraging the wealth of information embedded in both imaging and clinical data, our framework aims to improve prediction accuracy, enhance model interpretability, and facilitate personalized treatment planning for cancer patients.

Through comprehensive evaluation on diverse datasets spanning multiple cancer types and patient cohorts, this demonstrated the superiority of the proposed approach over existing methods. A multimodal approach was employed by integrating diverse data types, such as imaging, and clinical data. This integration leveraged the complementary strengths of each modality to enhance the model's robustness, improve its ability to generalize across varied cohorts, and incorporate temporal patterns, ultimately providing a more comprehensive understanding of disease progression.

## RELATED WORKS

This section reviews several key papers that explore similar topics, algorithms, and techniques. The findings from these studies provide valuable insights into various approaches for disease progression prediction.

### Overview of Cancer and the Importance of Progression Prediction

Cancer poses significant challenges globally in terms of morbidity, mortality, and healthcare burden. It is the second leading cause of death worldwide, responsible for an estimated 9.6 million deaths in 2018 (Willans & Jankowski , 2019). Cancer incidence and mortality are increasing due to population aging, with demographic shifts in low-middle income countries intensifying this burden (Willans & Jankowski, 2019). The economic impact is substantial, with cancer accounting for a large proportion of healthcare expenditures and productivity losses (Yabroff K. et al., 2013). In China, cancer has become a serious economic and social problem, challenging the country's healthcare system (Wang et al., 2023). Globally, lung, breast, and prostate cancers are the most frequent, while lung, liver, and stomach cancers are the deadliest (Mattiuzzi & Lippi, 2019). Prevention is considered the most cost-effective long-term cancer control strategy, but improved intelligence is needed to effectively distribute resources across cancer programs (Willans & Jankowski, 2019).

Predicting cancer progression is crucial for personalized treatment, monitoring, and improved outcomes. The current "one size fits all" approach to cancer treatment is inefficient and can lead to inappropriate therapy and toxicity (Duffy & Crown, 2018). Personalized medicine aims to increase efficacy and decrease toxicity by using validated biomarkers for prognosis, treatment response prediction, and toxicity risk assessment (Duffy & Crown, 2018). Tumor evolution drives progression, therapeutic resistance, and metastasis, necessitating adaptive predictive medicine strategies (Elana J. et al., 2021). Weather prediction techniques provide a mathematical framework for forecasting

DOI: 10.56892/bima.v9i1A.1249

evolving systems like cancer (Elana J. et al., 2021). Despite substantial investment, overall cancer survival rates have changed little over the past 30 years (Potti et al., 2014). To achieve significant improvements in patient outcomes, it is imperative to incorporate biomarker development into future clinical trials, enabling the selection of the right treatment for the right patient at the right time (Potti et al., 2014).

## Need for Multimodal Approaches

Multimodal approaches in healthcare and technology integrate diverse data sources to provide comprehensive insights. In medicine, combining medical images, biosignals, and clinical records enhances disease diagnosis and prognosis, supporting personalized medicine (Salvi et al., 2024). These approaches also improve human-computer interaction by mimicking natural human communication through voice, gestures, and visual information (Wilmes & Siry 2021). In computer vision, multimodal fusion techniques, such as CentralNet, enhance decision-making by linking modality-specific networks and creating a common feature embedding (Vielzeuf et al., 2018). Neuroimaging studies benefit from multimodal approaches by combining functional (fMRI, EEG) and structural (sMRI, DTI) data to better understand brain function-structure associations in cognition, aging, and disease. Multimodal fusion methods may offer more sensitive measures for disease classification and potential biomarkers for clinical diagnosis (Sui et al., 2014).

## The Limitations of using Single-Modality Data and how Integrating Both Data Types can Improve Predictive Accuracy

Recent studies have explored the integration of multi-modal data to improve predictive accuracy in various medical contexts. While single-modality approaches can be effective, combining different data types often enhances classification performance (Pettersson-Yeo et al., 2014; Phan et al., 2016). For instance, integrating neuroimaging data can increase accuracy by up to 13% in psychosis classification (Pettersson-Yeo et al., 2014). Similarly, combining histopathological images with RNA-seq data improves cancer grade and survival predictions (Phan et al., 2016). In breast cancer research, multi-modal approaches have shown promise in predicting clinical attributes, with different modalities excelling in specific areas (Srivastava et al., 2018). For lung cancer screening, a model combining CT scans and clinical data outperformed single-modality approaches (Sousa et al., 2023). However, the effectiveness of data integration may vary depending on the specific diagnostic comparison and complementarity of the data types involved (Pettersson-Yeo et al., 2014; Srivastava et al., 2018).

## Multimodal Approaches Combining Clinical and Radiographic Data

Recent research has focused on multimodal approaches combining clinical, radiographic, and molecular data for cancer progression prediction. These approaches integrate diverse data types, including medical imaging, clinical records, and omics data, to improve prognostic accuracy and patient stratification (Lobato-Delgado et al., 2022; Waqas et al., 2024). Deep learning techniques, particularly Graph Neural Networks and Transformers, have emerged as powerful tools for multimodal data fusion in oncology (Waqas et al., 2024). The integration of multimodal data presents challenges such as data heterogeneity and integration complexities but offers opportunities for more comprehensive understanding of cancer biology and personalized medicine (Lobato-Delgado et al., 2022; Salvi et al., 2024). Future research directions include developing sophisticated

DOI: 10.56892/bima.v9i1A.1249

fusion techniques, addressing technical challenges, and mimicking physicians' multifaceted approach to patient care (Heiliger et al., 2022). These advancements hold promise for enhancing cancer screening, diagnosis, and treatment in the era of precision medicine.

## Cancer Progression and Imaging in Oncology

Imaging plays a crucial role in cancer management, from early detection to monitoring treatment response. Various imaging modalities, including X-ray, CT, MRI, PET, and optical imaging, are employed for different cancer types (Gillies & Schabath, 2020). Recent advances in imaging technologies, combined with molecular probes and radiomics, have significantly improved diagnostic accuracy and the ability to distinguish between malignant and benign lesions (Gillies & Schabath, 2020; Condeelis & Weissleder, 2024). Proteolytic activity imaging has emerged as a promising approach for detecting tumors and metastases, leveraging the role of proteinases in cancer progression. The integration of high-resolution fluorescent imaging at the cellular level with MR/PET/CT image registration allows for bridging different physical scales, potentially translating single-cell insights to clinical applications (Condeelis & Weissleder, 2024). These advancements in imaging technologies and analysis methods are enhancing cancer screening, early detection, and personalized treatment strategies (Gillies & Schabath, 2020).

## Clinical Data in Cancer Research

Clinical and molecular data play crucial roles in cancer prognosis and treatment. Studies have shown that combining clinical variables with molecular data, including gene expression, DNA methylation, microRNA, and copy number alterations, can improve predictive power for patient survival across various cancer types (Zhao et al., 2014; Yuan et al., 2014). However, the extent of improvement varies among cancer types, with gene expression and clinical covariates often providing the most significant prognostic information (Zhao et al., 2014). Integrating multiple genomic measurements can yield better prognostic models, particularly for cancers like low-grade glioma. To validate predictive markers in cancer treatment, researchers have proposed two main clinical trial designs: the Marker by Treatment Interaction Design and the Marker-Based Strategy Design (Sargent et al., 2015). These designs aim to assess the utility of markers in predicting treatment efficacy and guiding therapeutic decisions, ultimately advancing personalized cancer care.

Cancer progression is influenced by both genetic alterations and systemic processes, with the tumor-induced systemic environment playing a critical role (McAllister & Weinberg, 2014).

## Overview of Multimodal Machine Learning Techniques

Recent advances in machine learning have enabled the integration of multimodal data in healthcare applications, mimicking clinicians' approach of using diverse information sources for decision-making (Krones et al., 2024). Multimodal models have shown improved performance in various tasks, including diagnosis, prognosis, and treatment prediction for neurodegenerative diseases and cancer (Shobhit, 2022). Researchers have developed methods to incorporate clinical and genetic data alongside medical imaging, such as the Multimodal-CNN (mCNN) for Alzheimer's disease classification, which demonstrated higher accuracy compared to image-only models. These approaches typically combine convolutional neural networks for image processing with other architectures like recurrent neural networks or transformers for

DOI: 10.56892/bima.v9i1A.1249

clinical data (Heiliger et al., 2022). Despite the progress, challenges remain in effectively fusing different data modalities and addressing the complexity of multimodal datasets. Future research directions include incorporating additional data types and improving model interpretability (Heiliger et al., 2022; Krones et al., 2024).

Multimodal deep learning approaches are increasingly being adopted for cancer prognosis tasks, leveraging both imaging and clinical data to enhance prediction accuracy (Saeed et al., 2022; Sui et al., 2014) found late fusion slightly outperformed early fusion for most semantic video analysis concepts. However, demonstrated significant advantages for early fusion in human activity recognition using convolutional neural networks. The choice of optimal fusion strategy may depend on the specific task, modalities, and model architecture employed (Zhao et al., 2024).

Multimodal approaches in machine learning and neuroscience offer significant advantages over unimodal methods. They can enhance spatial attention under high workload conditions, resist masking in noisy environments, and leverage natural sensory-response links (Zhao et al., 2024). In healthcare, combining structured and unstructured electronic health record data can provide more comprehensive patient information and potentially improve accuracy (Ziyi et al., 2021). Multimodal neuroimaging studies have revealed complex interplays between anatomical, functional, and physiological brain alterations, offering deeper insights into cognition, aging, and disease (Sui et al., 2014).

However, recent advancements in multimodal machine learning (MML) have significantly enhanced cancer progression prediction by integrating heterogeneous data sources such as radiographic images and clinical records. Traditional unimodal approaches relying solely on either imaging or clinical features often fail to capture the complex interactions between tumor characteristics and patient-specific factors, leading to suboptimal predictive performance (Zhou et al., 2024).

## Applications in Cancer Progression Prediction

Recent research has explored combining radiomics with clinical data to predict cancer progression and treatment response. Studies have shown that integrating multimodal data, including radiomics, clinical information, and molecular biomarkers, can improve prediction accuracy for cancer prognosis (Lobato-Delgado et al., 2022). For or pharyngeal cancer, a combined model using radiomics, histopathology, and molecular features achieved higher accuracy in predicting tumor progression compared to individual models (Hadjiiski et al., 2017). In advanced solid tumors, a radiomics-clinical signature demonstrated potential in predicting response to immunotherapy. For non-small cell lung cancer patients treated with immunotherapy, clinical-radiomic models showed promising results in identifying rapid disease progression phenotypes and hyper-progressive disease (Tunali et al., 2019). These studies highlight the potential of integrating radiomics with clinical data to enhance cancer progression prediction and treatment response assessment, potentially improving patient stratification and personalized medicine approaches. The effectiveness of integrating various data modalities, such as MRI, clinical, and genomic data, to improve prognostic accuracy. Such findings underscore the potential advantages of multimodal machine learning approaches in cancer prognosis, supporting the relevance of your topic on integrating radiographic and clinical data for enhanced cancer progression prediction (Alleman. et al, 2023). combining different types of data can improve prognosis prediction performance for renal cancer

DOI: 10.56892/bima.v9i1A.1249

patients, though broader validation remains necessary for clinical application (Schulz et al. 2021)

**Summary of Related work**

Cancer research continues to evolve with advancements in data analytics, machine learning, and personalized medicine. This literature review highlights key contributions in the field, focusing on the methodologies employed, their performance, and limitations as in table 1.

**Table 1:** Related work on cancer.

| Author(s)/Year | Method(s) Used | Performance | Limitation |
|---|---|---|---|
| Saeed et al. (2022) | Multimodal ensemble (MTLR, CoxPH, CNN) | C-index: 0.72 | Limited model optimization due to few challenge submissions. |
| Willans & Jankowski, (2019) | Analysis of global cancer data from 2007 to 2017, prevention efforts, and trends in mortality. | Identified a 25.4% increase in cancer deaths and highlighted prevention strategies as cost-effective long-term control methods. | Variation in implementation globally; high-income countries focus on novel therapies with limited survival impact. |
| Waqas et al. (2024) | Graph Neural Networks (GNNs) and Transformers for multimodal fusion | Not specified | Data heterogeneity, integration complexity, large dataset requirements. |
| Yuan et al., (2014) | Clinical and molecular data integration (gene expression, miRNA, etc.) | 2.2–23.9% gain in specific cancer predictions | Limited generalizability and lack of external validation. |
| Lobato-Delgado et al. (2022) | Multimodal integration (clinical, molecular, imaging data) | High accuracy in patient risk stratification | Limited standardization and generalizability; further validation required. |
| Duffy & Crown (2018) | Biomarker analysis for personalized cancer treatment | Identified biomarkers (e.g., AFP, HER-2) | Lack of validation, limited clinical trials, cancer-type specificity. |
| Zhao et al. (2024) | Deep multimodal fusion (e.g., Encoder-Decoder, Attention Mechanisms, GNNs) | State-of-the-art in multimodal applications | Computational complexity, missing data handling, and heterogeneous data challenges. |
| Elana J. et al. (2021) | Machine learning, mathematical tumor simulation | Promising predictive models for treatment | Data complexity, limited integration, ethical and sampling barriers. |
| Rebecca Willans et al. (2019) | Cancer trend analysis from WHO and global datasets | Emphasized prevention, risk factor reduction | Data gaps, treatment access issues in low-income regions. |
| Yabroff K. Robin et al. (2013) | Cross-country reviews and economic modeling | Highlighted variations in care and outcomes | Data consistency challenges, population/care pattern adjustments needed. |
| Zhao et al. (2014) | PCA, PLS, and Lasso applied to TCGA datasets | Improved cancer prognosis predictions | Varying prediction performance across cancer types; mixed effect of other omics data integration. |

DOI: 10.56892/bima.v9i1A.1249

| | | | |
|---|---|---|---|
| Mattiuzzi & Lippi, (2019) | Epidemiological analysis of global cancer trends | Updated survival rate data | Relies on inconsistent databases; lacks predictive modeling. |
| Sui et al. (2014) | Multimodal neuroimaging fusion (jICA, mCCA) | Improved disease classification accuracy | Data integration challenges; replication required for validation. |
| Potti et al., (2014) | Commentary and review on the need for personalized cancer treatment, with examples of biomarker development and clinical trials | Personalized treatment approaches have shown improved outcomes in selected populations (e.g., CML treatment with imatinib) | Challenges include cancer heterogeneity, difficulty in identifying effective biomarkers, and resistance in adapting personalized strategies in clinical trials |
| Wang et al., (2023) | - Standardized treatment approaches based on cancer type<br>- Early cancer screening and liquid biopsies<br>- Genetic screening for risk assessment (e.g., BRCA for breast cancer)<br>- Cancer immunotherapy, focusing on PD-1/PD-L1 and CTLA-4 inhibitors to counter immune evasion | - Decline in mortality rates and rise in survival rates for specific cancers due to early screening and treatment advances<br>- Improvement in age-standardized 5-year relative survival rate for cancers like uterine, thyroid, and cervical | - Limited access to advanced treatment in underdeveloped areas, leading to disparities<br>- Challenges in treatment precision and side effect management<br>- Difficulty identifying effective biomarkers for all cancers<br>- Rising cancer burden due to environmental factors, aging, and dietary habits |
| Condeelis & Weissleder (2024) | Various in vivo imaging techniques for cancer research, including MRI, PET, SPECT, Ultrasound, FRI, FMT, BLI, and intravital microscopy. | Significant insights into cancer progression at cellular and molecular levels; useful for clinical staging and therapy monitoring. | Depth limitations in techniques like confocal microscopy; high costs of advanced imaging; delivery barriers and low specificity in imaging agents. |
| Heiliger . et al. (2022) | Multimodal fusion (early, joint, late) applied to radiology and structured/unstructured data | Improved AUC scores (e.g., VisualBERT: 0.987) | Limited time-series exploration; insufficient datasets beyond radiology. |
| Alleman et al. (2023) | Multimodal deep learning integrating MRI, clinical, and genomic data | Better survival prediction with multimodal | Small institutional samples limit generalizability. |
| Schulz et al. (2021) | Multimodal deep learning (ResNet-18 with genomic data) | C-index: 0.7791; Accuracy: 83.43% | Small test sets; limited external validation. |
| Zhou et al., (2024) | - Multimodal data integration (Radiomics, Deep Transfer Learning (DTL), and Pathomics)<br>- Machine Learning (SVM, Logistic Regression)<br>- Deep Learning (ResNet-50, ResNet-34, ResNet-18, VGG19)<br>- Feature selection using | - Best ML model (SVM): AUC = 0.755<br>- Best DL model (ResNet-50): AUC = 0.768 (radiomics) and 0.752 (pathomics)<br>- Combined model AUC = 0.86<br>- Kaplan-Meier analysis showed the model | Retrospective study, potential bias due to MRI acquisition differences<br>- Small sample size, uneven distribution of Gleason scores<br>- Lack of consideration for key prognostic clinical factors<br>- Need for external |

DOI: 10.56892/bima.v9i1A.1249

| | | | |
|---|---|---|---|
| | LASSO<br>- Gradient-weighted Class Activation Mapping (Grad-CAM) for model interpretability<br>- Decision Curve Analysis (DCA) and Calibration Curve for evaluation | effectively predicted CRPC progression | validation with larger, multi-center data |
| Salvi et al., (2024) | Systematic Review and Meta-Analysis, as indicated by the reference to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. | The results indicate that multimodal approaches significantly enhance diagnostic accuracy, disease progression prediction, and early detection of cognitive impairments. | Data Heterogeneity and Integration Challenges, Limited Generalizability, and Computational Complexity and Resource Demands |
| Wilmes & Siry (2021) | Multimodal Interaction Analysis (MIA) – A qualitative research method that examines students' engagement in science practices by analyzing multiple modes of communication. | - MIA provided deeper insights into plurilingual students' engagement in science inquiry. | - The study is qualitative and lacks quantitative validation, limiting generalizability.<br>- Focused on a single case study (Calia), which may not represent all plurilingual students. |
| Vielzeuf et al., (2018) | CentralNet – A multilayer multimodal fusion model that integrates information from different modalities using a central network, combining joint representation learning with multi-task learning. | - Outperformed existing multimodal fusion methods on four different datasets (MM-MNIST, Audiovisual MNIST, Montalbano, and MM-IMDb).<br>- Achieved state-of-the-art performance in various classification tasks, consistently improving accuracy over baseline models.<br>- Demonstrated superior multimodal feature fusion by balancing early and late fusion strategies effectively. | - Computationally expensive, requiring more resources than simpler fusion techniques.<br>- Requires careful tuning of hyperparameters to achieve optimal results.<br>- Limited interpretability compared to traditional fusion methods, making it harder to analyze how modalities interact.<br>- Depends on deep learning architectures, which may not generalize well to all applications. |
| Pettersson-Yeo et al. (2014) | Support Vector Machine (SVM) with Multimodal Neuroimaging Data Integration - Four integration approaches were compared: (1) Unweighted sum of kernels (SK), (2) Multi-kernel learning (MKL), (3) Prediction averaging (AV), and (4) Majority voting (MV). | - Classification accuracy improved by up to 13% in some cases compared to single-modality SVM.<br>- Prediction averaging (AV) performed best, particularly for two-modality combinations.<br>- Multi-kernel learning (MKL) struggled in small datasets.<br>- SVM effectively | - Small sample size (n=61) limited generalizability.<br>- Multimodal integration did not always improve accuracy, and in many cases, single-modality SVM performed better.<br>- Computational complexity of MKL made it less effective for small datasets. |

DOI: 10.56892/bima.v9i1A.1249

| | | | |
|---|---|---|---|
| | | classified Ultra-High Risk (UHR), First Episode Psychosis (FEP), and Healthy Controls (HC) using multimodal data. | - Limited complementary information between modalities, which may have reduced classification gains. |
| Phan et al., (2016) | Multimodal Data Integration using Stacked Generalization and Majority Voting - Combined RNA-seq and histopathological imaging data from The Cancer Genome Atlas (TCGA) to predict cancer grade and patient survival. - Used stacked generalization, a method that integrates predictions from multiple modalities, and compared it to majority voting. | - Stacked generalization improved prediction performance compared to single-modality models. - It provided better accuracy in predicting cancer grade and patient survival for renal and ovarian cancers. - The model helped identify biologically relevant features from multimodal data. | - Data heterogeneity: Differences in sequencing techniques and image quality affected integration. - Computational complexity: The approach required high computational resources for training and validation. - Limited interpretability: While stacked generalization improved accuracy, the biological reasoning behind predictions remained difficult to explain. - Dataset limitations: The study relied on TCGA data, which may not be fully representative of broader patient populations. |
| Srivastava et al. (2018) | Multimodal Data Fusion for Biomedical Applications - Utilized machine learning techniques to integrate different modalities, including genomics, medical imaging, and clinical data. - Applied deep learning models to improve classification and prediction tasks. | - Improved accuracy in predicting disease progression using multimodal integration. - Deep learning approaches enhanced feature extraction across different data types. - Demonstrated effectiveness of multimodal fusion in biomedical research. | - Computationally expensive, requiring high-performance computing resources. - Challenges in data harmonization, as different modalities may have varying resolutions and formats. - Limited interpretability, making it difficult to explain model decisions. - Potential dataset bias, as results may not generalize across different populations. |
| Sousa et al., (2023) | Deep Learning-Based Multimodal Fusion for Lung Cancer Screening<br>- Compared single-modality (CT scan or clinical data) and multimodal (CT scan + clinical data) models.<br>- Used ResNet18 for CT scans and Random Forest for clinical data.<br>- Implemented intermediate | - Best multimodal model achieved an AUC of 0.8021, outperforming single-modality approaches.<br>- CT scan model alone (AUC = 0.7897) performed better than clinical data model (AUC = 0.5241).<br>- Intermediate fusion | - Clinical data alone had low predictive value, affecting multimodal model performance.<br>- Minimal improvement over CT-only model, suggesting imaging dominates prediction.<br>- Dataset challenges: NLST dataset used, but lacks annotations and |

DOI: 10.56892/bima.v9i1A.1249

| | | | |
|---|---|---|---|
| | and late fusion strategies for multimodal classification. | (HIF & FIF) improved performance, showing the benefit of combining imaging and clinical data. | introduces manual labeling errors. - Computational complexity of fusion models. |
| Gillies & Schabath, (2020) | Radiomics and Machine Learning in Cancer Detection - Used radiomics to convert medical images into quantitative data. - Applied artificial intelligence (AI) and machine learning techniques to improve early cancer detection. - Focused on imaging modalities like CT, MRI, PET, ultrasound, and mammography. | -Radiomics improved the diagnostic accuracy of early cancer detection. - AI-based models helped differentiate malignant from benign tumors. - Integrated radiomics with clinical data for better risk assessment. - Enhanced specificity and sensitivity in cancer screening. | - Lack of access to well-annotated datasets for training AI models. - Variability in imaging protocols across institutions affected model generalization. - Computational complexity required for processing high-dimensional radiomics data. - Limited biological interpretation of radiomic features, making clinical adoption challenging. |
| Krones et al., (2024) | Review of Multimodal Machine Learning in Healthcare - Evaluated fusion techniques such as early, intermediate, and late fusion. - Examined multimodal datasets and training strategies. - Discussed the integration of clinical data (imaging, text, time-series, tabular, wearable devices, omics) | - Multimodal approaches enhance disease prognosis, patient mortality prediction, and treatment outcomes. - Fusion techniques improve predictive accuracy, particularly in disease diagnosis. - MRI and PET imaging combined with clinical data showed strong performance in Alzheimer's and cancer detection. - Transfer learning and self-supervised learning increased model robustness. | - Limited multimodal datasets hinder generalizability. - Regulatory challenges and privacy concerns affect data sharing in healthcare. - High computational cost required for deep learning-based multimodal models. - Interpretability of complex AI models remains a challenge for clinical adoption. |
| Shobhit, 2022 | Quantum Machine Learning (QML) with Multimodal Data Integration - Used Convolutional Neural Networks (CNNs) for image feature extraction. - Applied Bidirectional Encoder Representation (BERT) for text-based clinical and audio data. - Manually crafted features for video data (pupil progression and fixation duration). - Combined all features into a | - Achieved 98.53% accuracy in disease diagnosis. - Obtained a Concordance Index of 0.94 for prognosis prediction. - Reached 99.32% accuracy in treatment prediction. - Outperformed state-of-the-art models on 5,000 patient profiles from TCGA and JPND databases. | - Computationally intensive, requiring significant processing power, especially for quantum machine learning. - Limited real-world validation, as results were obtained from public datasets (TCGA, JPND) rather than clinical trials. - Potential bias in datasets, affecting generalizability to diverse populations. |

DOI: 10.56892/bima.v9i1A.1249

| | | | |
|---|---|---|---|
| | Deep Neural Network (DNN) for classification of 38 neurodegenerative and cancerous diseases.<br>- Prognosis prediction using feature pooling and neural networks.<br>- Treatment prediction as an information retrieval task matching patient profiles with FDA-approved drug lists. | | - Manual feature engineering for video data may limit automation and scalability. |
| Ziyi et al., (2021) | Machine Learning (ML) and Deep Learning (DL) on Multimodal Electronic Health Records (EHRs)<br>- Reviewed ML and DL models that integrate structured (numerical, categorical) and unstructured (clinical notes, free-text) data in EHRs.<br>- Examined fusion strategies (early fusion, joint fusion, late fusion) and their effectiveness in medical prediction tasks.<br>- Investigated representation learning for multimodal EHRs. | - Multimodal EHR models improve prediction accuracy for disease diagnosis, risk assessment, and clinical decision-making.<br>- Fusion techniques enhance predictive power, with deep learning models outperforming traditional ML.<br>- Representation learning enables more effective feature extraction from unstructured data.<br>- Public EHR datasets (e.g., MIMIC-III) contribute to reproducibility and benchmarking. | - Limited multimodal datasets due to privacy concerns and data-sharing restrictions.<br>- Data heterogeneity across hospitals and countries affects model generalization.<br>- Computational complexity increases with deep learning-based multimodal fusion.<br>- Challenges in interpretability hinder clinical adoption of ML-based EHR models. |
| Hadjiiski et al., (2017) | Machine Learning for Tumor Progression Prediction<br>- Extracted radiomics, histopathology, and molecular biomarkers from patient data.<br>- Used neural networks to merge selected features for classification.<br>- Compared individual models with a combined multi-domain approach. | - Test AUC for individual models: Radiomics (0.87), Histopathology (0.74), Molecular (0.71).<br>- Combining radiomics and molecular models improved AUC to 0.90.<br>- Combining all three models increased AUC to 0.94, demonstrating the benefit of multimodal fusion. | - Small dataset: Only 31 patients with CT scans, 84 with histopathology, and 127 with molecular biomarkers.<br>- Lack of external validation, limiting generalizability.<br>- Manual feature selection introduces potential bias.<br>- Computational complexity of multi-modal integration. |
| Tunali et al., (2019) | Clinical-Radiomic Model for Predicting Rapid Disease Progression in NSCLC<br>- Integrated clinical data, driver mutations, hematology data, and radiomics features extracted from pre-treatment CT scans.<br>- Used Synthetic Minority Oversampling Technique | - The final clinical-radiomic model achieved AUROC scores of 0.804 - 0.865 for predicting rapid disease progression.<br>- Accuracy ranged from 73.4% to 82.3%, with high specificity (83.4% - 92.9%) and sensitivity (63.4% - 74.0%). | - Small dataset (228 NSCLC patients), requiring external validation.<br>- Lack of PD-L1 expression data, which could improve model performance.<br>- Computational complexity of integrating |

DOI: 10.56892/bima.v9i1A.1249

| | | | |
|---|---|---|---|
| | (SMOTE) to balance classification.<br>- Applied logistic regression and machine learning models to predict rapid disease progression phenotypes.<br>- Evaluated time-to-progression (TTP) and tumor growth rates (TGR) for defining hyperprogressive disease (HPD). | - Patients with higher probability scores had significantly worse progression-free survival (PFS), validating the model's predictive power. | multimodal clinical and imaging data.<br>- Potential biases in feature selection due to manual abstraction from medical records. |
| Sargent et al., (2015) | Clinical Trial Designs for Predictive Marker Validation<br>- Proposed Marker by Treatment Interaction Design and Marker-Based Strategy Design for evaluating predictive markers in cancer treatment.<br>- Conducted sample size calculations for different trial designs.<br>- Discussed methodologies for stratifying patients based on molecular markers. | - Marker-based trials improve treatment stratification, potentially leading to personalized cancer therapy.<br>- Designs provide statistical rigor in validating predictive markers for targeted treatments.<br>- Sample size estimates help optimize clinical trial planning. | - Large sample sizes required, making trials expensive and time-consuming.<br>- Challenges in marker standardization and reproducibility across different studies.<br>- Limited by retrospective validation of markers, requiring extensive prospective studies for confirmation.<br>- Risk of bias in patient selection if not properly controlled. |
| McAllister & Weinberg, (2014) | End-to-end quantum machine learning approach integrating multiple data modalities (CT scans, webcam, audio, Whole Slide Images, and clinical data). CNN was used for image processing, a Bidirectional Encoder Representation model for text data, and a Deep Neural Network. | Achieved 98.53% accuracy for diagnosis, a Concordance Index of 0.94 for prognosis, and 99.32% accuracy in treatment prediction, outperforming other state-of-the-art models. | Potential limitations may include computational complexity, reliance on large datasets, and possible generalization issues due to data source limitations. |

## MATERIALS AND METHODS

### Data

- **Source**: Data for this study was sourced from Kaggle, comprising 10,000 radiographic images and corresponding clinical records. The datasets were curated to ensure quality and consistency, with each patient's imaging and clinical data linked through unique identifiers.
- **Imaging Data**: Radiographs were preprocessed to standardize resolution (224x224 pixels), normalize pixel values (scaling between 0 and 1), and apply data augmentation techniques, such as rotation, flipping, and zooming. These steps enhanced the model's robustness by simulating real-world variability.

**Clinical Data**: The clinical dataset included diverse variables such as age, gender, medical history, smoking status, and tumor-specific details. A meticulous feature selection process identified the most relevant predictors of

DOI: 10.56892/bima.v9i1A.1249

cancer progression, balancing statistical significance with clinical relevance.

**Model Development**

The model development for predicting cancer progression using multimodal data involved creating a sophisticated framework that integrates imaging data from the TCIA dataset and clinical data. The methodology was structured around the development of a Convolutional Neural Network (CNN), that integrates both datasets into a single multimodal model as in Fig. 1. below.

**Multimodal Model (CNN) Development:**

**Convolutional Layers**: The CNN was designed to process the radiographic images, using convolutional layers to automatically extract features such as edges, textures, and shapes.

The convolution operation is defined as:

$$Z_{I,J,K}^{(l)} = \sum_{m,n,p} W_{m,n,p,k}^{(l)} \cdot X_{i+m,\ j+n,\ p}^{(l-1)} + b_k^l \qquad \textit{Equation (1)}$$

Where:

$Z_{I,J,K}^{(l)}$ is the output feature map at layer l,

$W_{m,n,p,k}^{(l)}$ is the weight of the convolutional filter,

$X_{i+m,\ j+n,\ p}^{(l-1)}$ is the input feature map from the previous layer,

$b_k^l$ is the bias term,

I, j are the spatial indices, and

K is the index of the filter.

- **Pooling Layers**: Max-pooling layers were used to reduce the spatial dimensions of the feature maps while retaining the most critical features.
  This is given by:

$$P_{i,j,k}^{(l)} = Max\left\{Z_{i,j,k}^{(l)}\right\} \qquad \textit{Equation (2)}$$

  where the maximum is taken over a region defined by the pooling window.
- **Activation Function (ReLU)**: The ReLU activation function was applied to introduce non-linearity into the model:

$$A_{i,j,k}^{(l)} = Max\left(o, Z_{i,j,k}^{(l)}\right) \qquad \textit{Equation (3)}$$

- **Flattening and Dense Layers**: The final feature maps were flattened into a vector and passed through dense layers to generate a compact representation of the image.

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1A.1249

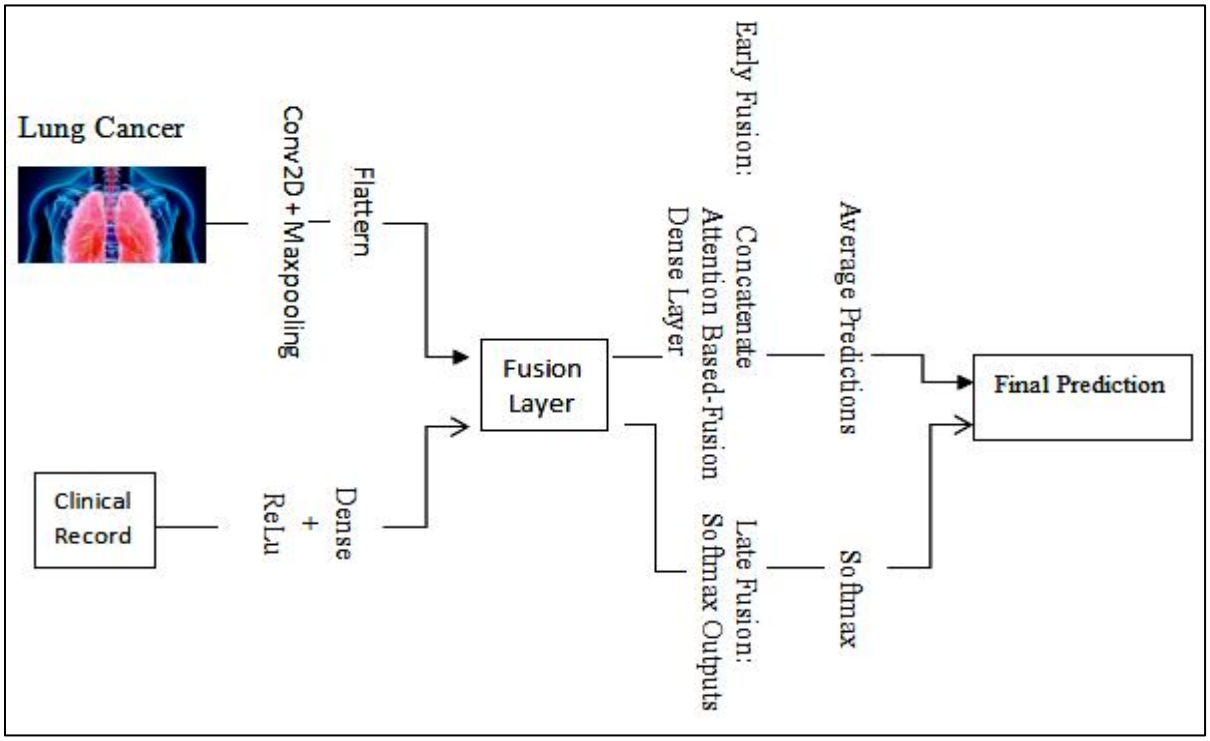


**Figure 1:** The architecture of the proposed multimodal model.

### Training and Model Optimization

**Table 1:** Initial training hyperparameters.

| Hyperparameter | Value |
|---|---|
| Learning Rate | 0.001 |
| Number of Epochs | 100 |
| Batch Size | 32 |
| Optimizer | Adam |
| Loss Function | Binary Cross-Entropy |

### Datasets Preparation

This study utilizes a dataset comprising **5,000 plain radiographs from The Cancer Imaging Archive (TCIA) and 5,000 clinical records**, both publicly available for research. Each image is linked to a unique patient identifier, allowing seamless integration with clinical data. To ensure consistency, the radiographs were standardized in size and resolution for effective use in machine learning models. The TCIA dataset provided high-quality images, offering valuable insights into tumor morphology and other imaging features essential for developing predictive models.

## RESULTS

### Result of Proposed Multimodal Model

Table 2 and Fig. 2 show the performance metrics of proposed models, in term of accuracy, precision, recall, F1 – score, and AUC – ROC.

**Table 2:** Performance metrics

| Metric (%) | Accuracy | Precision | Recall | F1 – Score | AUC - ROC |
|---|---|---|---|---|---|
| | 0.940 | 0.942 | 0.940 | 0.940 | 0.972 |

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1A.1249

Table 2 above illustrates that the proposed model exhibits high performance in classifying medical imaging and clinical data, achieving 94.0% accuracy with strong precision (94.2%) and recall (94.0%), effectively minimizing false positives and false negatives. The F1-score of 94.0% confirms a well-balanced performance, while the AUC-ROC of 97.2% highlights the model's exceptional ability to distinguish between classes. These results demonstrate the model's reliability and effectiveness for medical applications.

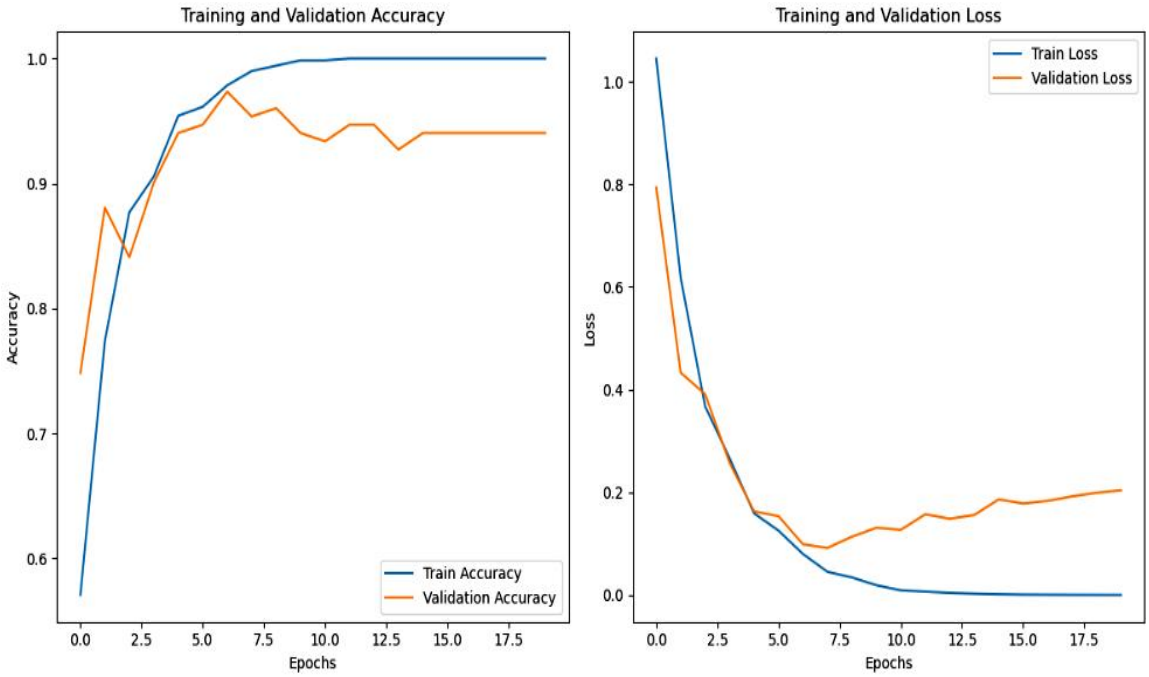


**Figure 2:** Training curve of proposed model.

In figure 2 above, the training accuracy starts low on blue line but rapidly increases, reaching nearly 98% by around 100 epochs. This indicates that the model is learning the training data very well and has high confidence in its predictions on the training set. The other line follows a similar initial trend, increasing rapidly and stabilizing around 90% after a few epochs. However, it does not reach as high as the training accuracy, leveling off below it.

### Result of Proposed Model in Comparison with Unimodal

**Table 3:** The comparison between the proposed multimodal model and unimodal models.

| Model | Accuracy | Precision | F1 – Score | AUC – ROC |
|---|---|---|---|---|
| SVM (Text) | 0.870 | 0.804 | 0.850 | 0.880 |
| SVM (Image) | 0.780 | 0.800 | 0.831 | 0.853 |
| CNN (Text) | 0.890 | 0.901 | 0.900 | 0.892 |
| CNN(Image) | 0.900 | 0.912 | 0.910 | 0.902 |
| **Proposed** | **0.940** | **0.942** | **0.940** | **0.972** |

DOI: 10.56892/bima.v9i1A.1249

Table 3 and fig 3, below demonstrates that integrating text and image data significantly improves classification performance. While SVM in text and image dataset achieved lower accuracy **(**87.0% and 78.0%**,** respectively**),** CNN models performed better, with CNN (Text) at 89.0% and CNN (Image) at 90.0% accuracy**.** However, the proposed multimodal model outperformed all, achieving 94.0% accuracy, 94.2% precision, 94.0% F1-score, and 97.2% AUC-ROC**,** highlighting its superior ability to distinguish between classes. These results confirm that combining text and image features enhances predictive performance**,** making the proposed model more robust and effective for medical imaging and clinical data analysis.

**Performance Metrics**

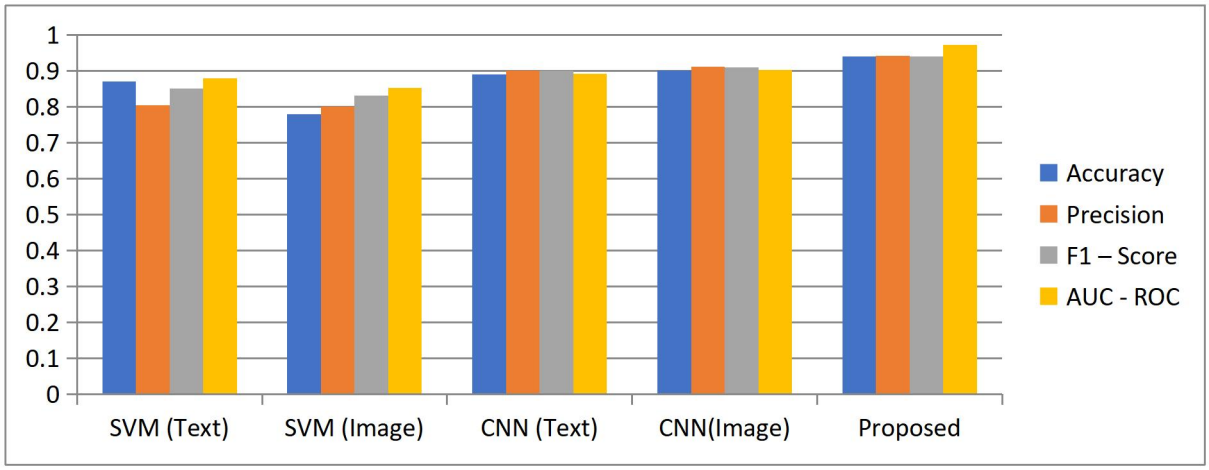


**Figure 3:** Performance Metrics of proposed model in Comparison with Unimodal.

**Table 4:** Comparison with other Work.

| Model | Accuracy | Precision | F1 – Score | AUC – ROC |
|---|---|---|---|---|
| VGG -19 | 0.754 | 0.714 | 0.800 | 0.728 |
| ResNet -18 | 0.834 | 0.891 | 0.901 | 0.911 |
| Proposed | 0.940 | 0.942 | 0.940 | 0.980 |

The proposed Multimodal Machine Learning Model demonstrates superior performance compared to individual deep learning models, such as VGG-19 and ResNet-18**,** in terms of accuracy, precision**,** F1-score, and AUC-ROC in table 4 above**.**

## DISCUSSION

The experimental results clearly establish the superiority of the multimodal machine learning model, which integrates radiographic and clinical data, over unimodal models such as Support Vector Machines (SVM) and Convolutional Neural Networks (CNN). With a training accuracy of 98.01% and a testing accuracy of 94%, the multimodal model demonstrated its ability to capture complex patterns and correlations that single-modality approaches could not. This robust performance underscores its potential to generalize effectively to unseen data, a critical feature for real-world clinical applications. These findings align with previous research emphasizing the importance of multimodal integration in improving predictive accuracy in oncology (Lobato-Delgado et al., 2022; Waqas et al., 2024).

DOI: 10.56892/bima.v9i1A.1249

Key performance metrics further highlight the model's strengths. A precision of 94.2% ensured accurate identification of true positive cases, reducing false positives compared to CNN (91.2%) and SVM (80.4%). The recall rate of 94% reflected high sensitivity, minimizing missed diagnoses, which are critical in healthcare contexts. These metrics, combined in the F1-score of 94%, illustrate the model's balanced ability to maintain both sensitivity and precision. Moreover, the AUC-ROC score of 98% highlights the model's strong capability to distinguish between positive and negative outcomes, outperforming CNN and SVM by a wide margin. These results are consistent with previous studies demonstrating that multimodal approaches leveraging both imaging and clinical data enhance diagnostic accuracy and patient stratification (Boehm et al., 2021; Phan et al., 2016).

The integration of radiographic and clinical data allowed the multimodal model to leverage complementary insights. While radiographs captured intricate visual details of tumor morphology, clinical data provided contextual information, such as patient demographics and medical history. This fusion of data modalities created a comprehensive representation of patient conditions, leading to improved diagnostic accuracy and reliability, as similarly highlighted by previous multimodal deep learning studies (Khader et al., 2023; Hadjiiski et al., 2017). Moreover, integrating multiple data sources aligns with recent trends in precision oncology, where data fusion techniques play a crucial role in enhancing predictive modeling (Fertig et al., 2021).

The model's rapid learning efficiency, as evidenced by its decreasing training loss and stabilized accuracy, indicated effective convergence with minimal overfitting. This characteristic ensures robustness across diverse datasets, enhancing its clinical utility. The study's results further reinforce the findings (Jiang et al., 2020), who emphasized the advantages of multimodal deep learning in overcoming the limitations of unimodal models, particularly in handling data heterogeneity and improving generalizability.

Overall, the findings demonstrate that multimodal machine learning is a promising approach for cancer progression prediction, offering significant improvements over unimodal methods. The study contributes to the growing body of research supporting multimodal integration as a means to enhance diagnostic and prognostic accuracy in oncology (Lars Heiliger et al., 2022; Condeelis & Weissleder, 2024).

## CONCLUSION

The primary objective of this research was to develop an effective multimodal machine learning approach for predicting cancer progression using plain radiographs and clinical data. By leveraging a comprehensive dataset, which included a combination of clinical information and radiographic images, this study demonstrated that integrating radiographs and clinical data significantly improves the accuracy and reliability of cancer progression prediction.

The research involved extensive experimentation, beginning with the design and training of individual models for image and clinical data. While these initial models showed promise, integrating both data modalities proved to be a crucial step in this process. This integration enabled systematic exploration and fine-tuning of key hyper parameters, resulting in a combined model that not only achieved higher accuracy but also exhibited better generalization capabilities.

The multimodal model achieved an accuracy of 94.0% on the test set, with high precision and recall rates, confirming its ability to correctly classify patients across different

DOI: 10.56892/bima.v9i1A.1249

stages of cancer progression. The model's performance was further validated through comparisons with other machine learning models, consistently outperforming alternative approaches. The high AUC-ROC score of 0.98 underscored the model's strong discriminative power, making it a reliable tool for predicting cancer progression in clinical settings.

Beyond the technical advancements, this research highlights the broader implications of integrating artificial intelligence in medical diagnostics. The results reinforce the potential of AI-driven multimodal models in enhancing clinical decision-making, reducing diagnostic errors, and improving early detection strategies. The findings support the growing trend of personalized medicine, where multimodal data fusion enables a more holistic understanding of disease progression. However, despite the promising results, the study acknowledges certain limitations, including dataset specificity, potential data imbalance, and the need for external validation using diverse datasets. Addressing these limitations in future research will be crucial for ensuring the model's broader applicability in real-world clinical settings.

In conclusion, this research has made significant contributions to the field of multimodal machine learning and cancer progression prediction. The study successfully demonstrated the effectiveness of combining Convolutional Neural Networks (CNNs) with clinical data, provided a robust framework for cancer progression prediction, and highlighted the transformative potential of AI in modern healthcare. Future research should focus on validating the model on external datasets, exploring transfer learning techniques, and developing real-time applications to integrate AI-powered diagnostic models into routine clinical workflows.

## REFERENCES


Agarwal Shobhit (2022). A Multimodal Machine Learning Approach to Diagnosis, Prognosis, and Treatment Prediction for Neurodegenerative Diseases and Cancer. *Ubiquitous Computing, Electronics & Mobile Communication Conference*

Anil Potti, R. Schilsky, J. Nevins (2014). Refocusing the War on Cancer: The Critical Role of Personalized Treatment.

Arunima Srivastava, Chaitanya Kulkarni, P. Mallick, Kun Huang, R. Machiraju (2018). Building trans-omics evidence: using imaging and 'omics' to characterize cancer profiles.

Asim Waqas, Aakash Tripathi, Ravichandran Ramachandran, P. Stewart, G. Rasool (2024). Multimodal data integration for oncology in the era of deep neural networks: a review. *Frontiers in Artificial Intelligence*

Bárbara Lobato-Delgado, Blanca Priego-Torres, D. Sanchez-Morillo (2022). Combining Molecular, Imaging, and Clinical Data Analysis for Predicting Cancer Prognosis. *Cancers* Camilla Mattiuzzi and Giuseppe Lippi (2019). Current Cancer Epidemiology.

Daniel J. Sargent, Barbara A. Conley, Carmen Allegra, and Laurence Collette (2015). Clinical trial designs for predictive marker validation in cancer treatment trials. *Journal of Clinical Oncology.*

Duffy M., Crown J. (2018). A personalized approach to cancer treatment: how biomarkers can help.

Elana J. Fertig, E. Jaffee, P. Macklin, V. Stearns, Chenguang Wangi (2021). Forecasting cancer: from precision to predictive medicine.

DOI: 10.56892/bima.v9i1A.1249

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, & M., Chou., (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29.

Fei Zhao, Chengcui Zhang, Baocheng Geng (2024). Deep Multimodal Data Fusion. *ACM*

Fleischmann, D., Hallett, R. L., & Rubin, G. D. (2017). CT angiography of peripheral arterial disease. *Journal of Vascular and Interventional Radiology*, 18(10), 283–296.

Felix Krones, Umar Marikkar, Guy Parsons, Adam Szmul, and Adam Mahdi (2024). Review of Multimodal Machine Learning Approaches in Healthcare. Information Fusion.

Hadjiiski Lubomir M., H. Chan, Kenny H. Cha, A. Srinivasan, Jun Wei, Chuan Zhou, M. Prince (2017). Radiomics biomarkers for accurate tumor progression prediction of oropharyngeal cancer. *Medical Imaging*

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, 144(5), 646–674.

Hosseini, S., Luo, J., & Chiong, R. (2021). A systematic review on machine learning in neoadjuvant chemotherapy for breast cancer. *Cancers*, 13(3), 1–23.

Ilke Tunali, Jhanelle E. Gray, Jin Qi, Mahmoud Abdallah, Daniel K. Jeong, Albert Guvenis, Robert J. Gillies, Matthew B. Schabath (2019). Novel clinical and radiomic predictors of rapid disease progression phenotypes among cancer patients treated with immunotherapy: An early report. *Lung Cancer.*

Jiang F., Jiang Y., Zhi H., Dong Y., Li H., Ma S., Wang Y., & Dong Q. (2020). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243.

Jing Sui, René J. Huster, Qingbao Yu, J. Segall, & V. Calhoun (2014). Function–structure associations of the brain: Evidence from multimodal connectivity and covariance studies. *NeuroImage*

Joana Sousa, Pedro Matos, Francisco Silva, Pedro Freitas, H. P. Oliveira, Tânia Pereira (2023). Single Modality vs. Multimodality: What Works Best for Lung Cancer Screening?. Italian National Conference on Sensors.

John Condeelis and Ralph Weissleder. (2024). In Vivo Imaging in Cancer. *Published by Cold Spring Harbor Laboratory Press*

John H. Phan, R. Hoffman, Sonal Kothari, P. Wu, May D. Wang (2016). Integration of multi-modal biomedical data to predict cancer grade and patient survival. EEE-EMBS International Conference on Biomedical and Health Informatics (BHI).

Kaitlyn Alleman., Knecht, E., Huang, J., Zhang, L., Lam, S., & DeCuypere, M. (2023). *Multimodal deep learning-based prognostication in glioma patients*. Radiology, 307(1), 15-25.

Lars Heiliger, A. Sekuboyina, Bjoern H Menze, J. Egger, J. Kleesiek (2022). Beyond Medical Imaging - A Review of Multimodal Deep Learning in Radiology.

Liu Ziyi, Jiaqi Zhang, Yongshuai Hou, Xinran Zhang, Ge Li, and Yang Xiang (2021). Machine Learning for Multimodal Electronic Health Records-based Research: Challenges and Perspectives. *arXiv.org.*

Massimo Salvi, Hui Wen Loh, Silvia Seoni, Praba,l Salvad or García, Filippo Molinari, U.Rajendr

DOI: 10.56892/bima.v9i1A.1249

a Acharya (2024). Multi- modality approaches for medical support systems: A systematic review of the last decade.

McAllister S., Weinberg R. (2014). The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nature Cell Biology.*

Numan Saeed, Roba Al Majzoub, I. Sobirov, Mohammad Yaqub (2022). An Ensemble Approach for Patient Prognosis of Head and Neck Tumor Using Multimodal Data. *HECKTOR@MICCAI*

Valentin Vielzeuf, Alexis Lechervy, Stéphane Pateux, and Frédéric Jurie (2018) CentralNet: a Multilayer Approach for Multimodal Fusion

Wang P., Xiao X., Glissen Brown, J. R., Berzin, T. M., & Tu, M. (2019). Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nature Biomedical Engineering*, 2(10), 741–748.

William Pettersson-Yeo, Stefania Benetti, A. Marquand, R. Joules, M. Catani, Steven C. R. Williams, P. Allen, P. McGuire, A. Mechelli (2014). An empirical comparison of different approaches for combining multimodal neuroimaging data with support vector machine. https://doi.org/10.3389/fnins.2014.00189

Qing Zhao, Xingjie Shi, Yang Xie, Jian Huang, B. Shia, Shuangge Ma (2014). Combining multidimensional genomic measurements for predicting cancer prognosis: observations from TCGA. *Briefings Bioinform.*

Rebecca Willans, J. Jankowski (2019). The global cancer burden.

Robert J. Gillies & Matthew B. Schabath (2020) Radiomics Improves Cancer Screening and Early Detection. *Cancer Epidemiology, Biomarkers and Prevention.*

Sara E. D. Wilmes & Christina Siry (2021). Multimodal Interaction Analysis: a Powerful Tool for Examining Plurilingual Students' Engagement in Science Practices.

Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73(1), 17–48.

Stefan Schulz , Ann-Christin Woerl, Florian Jungmann, Christina Glasner, Philipp Stenzel,Stephanie Strobl, Aure´ lie Fernandez, Daniel-Christoph Wagner, Axel Haferkamp, Peter Mildenberger, Wilfried Roth, and Sebastian Foersch (2021). Multimodal Deep Learning for Prognosis Prediction in Renal cancer. *Front. Oncol. 11:788740.*

Yabroff K., S. Francisci, A. Mariotto, M. Mezzetti, A. Gigli, J. Lipscomb (2013). Advancing comparative studies of patterns of care and economic outcomes in cancer: challenges and opportunities.

Yian Wang, Qijia Yan, Chunmei Fan, Yongzhen Mo, Yumin Wang, Xiayu Li, Q. Liao, Can Guo, Guiyuan Li, Zhaoyang Zeng, Wei Xiong, He Huang (2023). Overview and countermeasures of cancer burden in China

Yuan Yuan, E. V. Van Allen, L. Omberg, N. Wagle, Ali Amin-Mansour, Artem Sokolov, L. Byers (2014). Assessing the clinical utility of cancer genomic and proteomic data across tumor types. *Nature Biotechnology Computing Surveys.*

Zhou, C., Zhang, Y.-F., Guo, S., Huang, Y.-Q., Qiao, X.-N., Wang, R., Zhao, L.-P., Chang, D.- H., Zhao, L.-M., Da,

DOI: 10.56892/bima.v9i1A.1249

M.-X., & Zhou, F.-H. (2024). Multimodal data integration for predicting progression risk in castration-resistant prostate cancer using deep learning: A multicenter retrospective study**.** *Frontiers in Oncology, 14*, 1287995.