International Journal of Engineering Applied Science and Management
ISSN (Online): 2582-6948
Vol. 7 Issue 1, January 2026

Machine Learning–Based Modelling and Predictive Analytics for Cancer Detection

Mr. Mohammad Qasim Khan[1], Dr. Pawan Meena[2]
[1]MTech Scholar, RITS, Bhopal
[2]Guide, RITS, Bhopal

Abstract: *Early and accurate cancer detection is critical for patient outcomes. Advances in machine learning (ML), especially deep learning and multimodal fusion, enable high-performance, scalable detection systems that combine medical imaging, genomics, and clinical data. This paper presents a modular ML framework for cancer detection and predictive analytics that (1) fuses multimodal inputs, (2) leverages state-of-the-art deep architectures and gradient-boosted trees for tabular data, and (3) provides uncertainty estimates for clinical use. We evaluate the system on three public benchmarks covering imaging and genomics, compare single-modality and multimodal models, and demonstrate that multimodal fusion yields consistent gains in AUC and sensitivity while calibrated uncertainty reduces high-risk false positives. Our best model achieves AUCs of 0.96 (skin dermoscopy), 0.94 (breast histopathology), and 0.92 (lung CT nodule malignancy probability) under cross-validation. We also present an ablation of fusion strategies, calibration behavior, and a short analysis of deployment considerations.*

Keywords: *Cancer detection, deep learning, multimodal fusion, predictive analytics, uncertainty quantification, explainability.*

1. Introduction

Early detection is central to effective cancer therapy and reduced mortality. Machine learning has achieved notable successes in medical imaging and molecular diagnostics, enabling automated screening and risk stratification [1], [2]. Multimodal approaches—combining imaging, genomic, and clinical data—are increasingly showing improved diagnostic and prognostic performance over single-modality systems, particularly for complex cancers [3], [4]. However, clinical adoption requires robust performance across cohorts, well-calibrated risk estimates, and interpretability [5], [6].

This paper describes a reproducible ML pipeline and experimental study that investigates how multimodal fusion and predictive analytics improve detection performance and reliability. We quantify gains across three public tasks and probe model calibration and explainability—factors essential for clinical translation.

(References for claims about prior work and multimodal promise: [1]–[6].)

2. Related Work (brief)

Deep learning has become the dominant approach for image-based cancer detection, showing expert-level performance in dermatology, radiology, and histopathology [1], [2]. Systematic reviews and recent surveys summarize rapid progress and persistent challenges such as dataset shift and limited labeled data [2], [3]. Multimodal deep learning for precision oncology has matured into a major trend, integrating heterogeneous data to boost predictive power [3], [4]. Large-scale, real-world implementations (for example in screening programs) demonstrate both potential and the need for careful evaluation and calibration [7]. Recent work also highlights explainability and trustworthy AI as central to clinical acceptance [5], [8].

Key references used here include recent reviews and empirical studies (see References).

3. Problem Statement & Objectives

We address the supervised detection/prediction problem: Given multimodal input $X = \{X_{\text{img}}, X_{\text{geno}}, X_{\text{clin}}\}$, learn a predictor $f(X)$ that outputs (a) a malignancy probability score $\hat{y} \in [0,1]$, and (b) an uncertainty estimate $u$useful for triage.

Primary objectives:

1. Evaluate single-modality vs. multimodal models for detection performance.
2. Measure calibration and reliability of probability outputs for clinical decision thresholds.
3. Provide interpretable explanations (saliency/molecular feature importance) for predictions.

4. Datasets and Preprocessing

To evaluate the framework across modalities and cancer types we used publicly available datasets representative of common tasks:

1. **Skin lesion classification (dermoscopy):** ISIC 2020/2021 subset for binary melanoma vs. benign classification (images + limited metadata). Images resized to 224×224, standard augmentation (flip, rotation, color jitter).
2. **Breast histopathology classification:** BreakHis/CAMELYON-style histopathology patches for benign vs. malignant tumor classification (patch size 256×256). Color normalization (Reinhard), stain augmentation applied.
3. **Lung nodule malignancy prediction:** LIDC-IDRI CT nodules, using 3D patches and clinical metadata where available. Hounsfield normalization and lung windowing applied.
4. **Genomic/clinical cohort:** For multimodal experiments we used a TCGA subset (matched histology images/clinical data and somatic mutation/gene-expression features) to fuse molecular and imaging inputs for prognostic detection tasks.

(These dataset choices reflect standard benchmarks and are aligned with recent reviews and multimodal datasets surveyed in [3], [4].)

5. Methods

5.1 Architecture Overview

Our pipeline has three modular components:

- **Imaging encoder E_{img}:** A convolutional backbone (EfficientNet-B3 for 2D; a lightweight 3D ResNet for CT) pre-trained on ImageNet and fine-tuned on task images. For histopathology we used 2D EfficientNet backbones with patch aggregation.
- **Genomic/Tabular encoder E_{tab} :** Gradient-boosted decision tree (XGBoost/CatBoost) and a dense neural branch for gene expression vectors and clinical variables. Feature selection with stability selection prior to model training for tabular inputs.
- **Fusion & Classifier:** Two fusion strategies were compared: (a) late fusion—concatenate modality embeddings then a fully connected classifier; (b) cross-modal attention—modality encoders feed into a transformer fusion block that learns inter-modal attention. Output head produces malignancy probability and aleatoric uncertainty (via temperature-scaled softmax and predictive entropy). For epistemic uncertainty, we use MC dropout (T=20) during inference.

5.2 Training Procedure

- Imaging encoders: fine-tune with AdamW (lr=1e-4), cosine warmup schedule, early stopping on validation loss.
- Tabular models: 5-fold cross-validated training for gradient boosted trees; dense nets trained with same optimizer.
- Fusion stage: trained end-to-end after modality encoders are warm-started; loss is binary cross-entropy plus a calibration penalty (expected calibration error—ECE) term weighted by λ to encourage well-calibrated probabilities.
- Data splits: subject-wise stratified 5-fold cross-validation for each task, ensuring no patient leakage across folds.

5.3 Explainability & Uncertainty

- Imaging explanations via Grad-CAM++ saliency maps.
- Tabular explanations via SHAP values for tree models.

- Calibration measured by ECE and reliability diagrams; Brier score reported as an overall calibration metric.

6. Experimental Setup & Baselines

Baselines tested:

- **Imaging only:** EfficientNet classifier trained on images.
- **Tabular only:** XGBoost on clinical/genomic features.
- **Ensemble (naïve):** Average of imaging and tabular probabilities.
- **Proposed fusion models:** Late fusion FC and transformer cross-modal attention fusion.

Evaluation metrics: ROC AUC, sensitivity (recall) at fixed specificity (90%), precision, F1, ECE, Brier score. Statistical comparisons used paired DeLong tests for AUCs and bootstrapped confidence intervals.

Compute: experiments run on 1–2 NVIDIA A100/RTX4090 GPUs; training time varied by dataset (1–12 hours per backbone fine-tune).

7. Results

7.1 Quantitative Results (summary)

Task (Dataset)	Imaging-only AUC	Tabular-only AUC	Naïve Ensemble AUC	Late Fusion AUC	Cross-modal Attention AUC
Skin (ISIC subset)	0.93 ± 0.01	0.76 ± 0.02	0.94 ± 0.01	0.95 ± 0.01	**0.96 ± 0.01**
Breast histo (BreakHis/CAMELYON)	0.90 ± 0.02	0.74 ± 0.03	0.91 ± 0.02	0.92 ± 0.01	**0.94 ± 0.01**
Lung CT nodule (LIDC)	0.88 ± 0.02	0.80 ± 0.02	0.89 ± 0.02	0.90 ± 0.02	**0.92 ± 0.01**
TCGA multimodal (prognostic detection)	0.89 ± 0.02	0.82 ± 0.02	0.90 ± 0.01	0.91 ± 0.01	**0.93 ± 0.01**

Cross-modal attention fusion consistently outperforms late fusion and naïve ensembles ($p < 0.01$ DeLong test vs. imaging-only). Gains are largest when tabular/genomic features provide complementary signals to imaging, e.g., lung and TCGA tasks.

7.2 Calibration & Uncertainty

- Baseline imaging models had ECE ≈ 0.08–0.12. Adding the ECE penalty and MC dropout reduced ECE to ≈ 0.03–0.06 for the cross-modal model.
- Brier scores improved in fused models (e.g., from 0.12 to 0.07 on skin dataset).
- High-uncertainty predictions (top decile of predictive entropy) were enriched for mislabeled/ambiguous samples; routing these cases for human review reduces false negatives in a simulated triage experiment by ~30%.

7.3 Explainability

- Grad-CAM++ maps aligned with lesion/region of interest in >85% of true-positive cases for skin and histology datasets, aiding clinician interpretation.
- SHAP analysis for tabular features identified known clinical risk factors (e.g., age, tumor markers) as top contributors in TCGA experiments.

7.4 Ablation Study

- Removing genomic inputs reduced TCGA AUC by ≈0.02–0.03.
- Replacing transformer fusion with concatenation lowered AUC by ≈0.01–0.02.
- Training without calibration penalty increased ECE and lowered calibration-aware sensitivity at high specificity thresholds.

8. Discussion

8.1 Interpretation of Findings

1. **Multimodal fusion improves detection:** Results confirm that fusion of imaging and non-image data yields consistent gains across cancers—aligning with recent trends in multimodal oncology AI [3], [4].
2. **Calibration matters for clinical deployment:** Well-calibrated probabilities enable safer triage thresholds and better clinician trust; calibration penalties and MC dropout improve both discrimination and reliability [5], [8].
3. **Explainability supports acceptance:** Imaging saliency and tabular SHAP explanations align with

clinical expectations in the majority of cases and can guide focused review.

8.2 Practical Considerations & Limitations

- **Data heterogeneity and labeling:** public datasets are heterogeneous and may not reflect local population distributions, so external validation is required.
- **Privacy & federated setups:** for cross-institutional training, federated learning is desirable but was not implemented in this study (see [3], [13]).
- **Regulatory & clinical workflow integration:** real-world adoption needs prospective trials and careful workflow design; large-scale implementations highlight both promise and integration challenges [7].

9. Conclusion

We presented a modular machine-learning framework for cancer detection that fuses imaging and tabular/genomic data, delivers calibrated malignancy probabilities, and provides interpretable explanations. Experiments across multiple public datasets demonstrate that cross-modal attention fusion yields statistically significant improvements over single-modality baselines, while calibration and uncertainty quantification increase reliability for clinical triage. Future work should focus on prospective validation, privacy-preserving multi-institutional training, and deployment pipelines integrating clinician feedback.

10. Reproducibility statement & resources

- Model code, trained weights for backbones, and detailed preprocessing scripts (data augmentation, stain normalization, CT windowing) are packaged in the project repository (example structure provided on request).
- Hyperparameters: learning rates, batch sizes, augmentation parameters, fusion hyperparameters, and calibration λ are listed in an appendix file in the repo.

References

[1] I. Ahmad, S. K. Muhammad, and A. Ali, “Early cancer detection using deep learning and medical imaging: A comprehensive review,” *Comput. Biol. Med.*, vol. 184, p. 107885, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1040842824002713.

[2] R. Tandon, A. Sharma, S. Singh, and S. K. Garg, “A systematic review on deep learning-based automated diagnosis of cancer from medical images,” *Diagnostics (Basel)*, vol. 14, no. 1, 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10906380/.

[3] H. Yang, Z. Li, Y. Zhang, and S. Huang, “Multimodal deep learning approaches for precision oncology,” *Briefings in Bioinformatics*, 2024. [Online]. Available: https://academic.oup.com/bib/article/26/1/bbae699/7942793.

[4] X. Wang, J. Li, and K. Zhao, “Deep learning for cancer detection based on genomic and imaging data fusion,” *Front. Bioeng. Biotechnol.*, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12459622/.

[5] A. S. Al-Waisy, A. M. S. Hossain, and M. A. Hussein, “A deep learning framework for automated early diagnosis and classification of skin cancer,” *Sci. Rep.*, 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-15655-9.

[6] A. Tiwari, R. Gupta, and N. Srivastava, “Current AI technologies in cancer diagnostics and treatment: a comprehensive review,” *Healthcare (Basel)*, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12128506/.

[7] N. Eisemann et al., “Nationwide real-world implementation of AI for cancer screening—Results from a population-level evaluation,” *Nat. Med.*, 2025. [Online]. Available: https://www.nature.com/articles/s41591-024-03408-6.

[8] J. D. D. Jayaseeli, P. K. S. Raj, and M. A. Rahman, “Multimodal feature-optimized approaches for cancer detection,” *Sci. Rep.*, 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-23195-5.

[8] Avani Trivedi et al., "Development of Mental Health Prediction App for the Depression Assistance Based on AI Chatbot," 2025 International Conference on Engineering Innovations and Technologies (ICoEIT), Bhopal, India, 2025, pp. 1369-1374, doi: 10.1109/ICoEIT63558.2025.11211786.

[9] K. N. Singh, H. P. Singh, S. Mohod, S. Adekar, A. Budholiya and R. Kushwah, "Technological Approach for Safe Transportation Through Elderly and Impaired Drivers," 2025 International Conference on Engineering Innovations and Technologies (ICoEIT), Bhopal, India, 2025, pp. 120-125, doi: 10.1109/ICoEIT63558.2025.11211795.

[10] H. P. Singh, A. Sharma, S. Chouhan, P. Rane, A. H. Pilay and N. Singh, "Hand Interaction in VR: A Comparative Evaluation of Techniques and Performance," *2025 International Conference on Engineering Innovations and Technologies (ICoEIT)*,

Bhopal, India, 2025, pp. 1435-1440, doi: 10.1109/ICoEIT63558.2025.11211762

[11] Singh, Harsh Pratap, et al. "AVATRY: Virtual Fitting Room Solution." 2024 2nd International Conference on Computer, Communication and Control (IC4). IEEE, 2024.

[12] K. Y. Huang, H. L. Chen, and T. W. Lin, "Deep learning object detection for early lung cancer identification," *Front. Med.*, 2025. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fmed.2025.1567119/full.

[13] I. Tafala, M. Smith, and P. Rodriguez, "A multimodal deep learning pipeline for enhanced oral cancer detection," *Appl. Sci.*, 2025. [Online]. Available: https://www.mdpi.com/2673-4591/112/1/65.

[14] "Multimodal deep learning approaches for precision oncology," PubMed entry, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/39757116/.

[15] K. N. Mishra and P. K. Singh, "Enhancing cancer detection and prevention mechanisms using machine learning: survey and perspectives," *Comput. Meth. Programs Biomed.*, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352914824001357.