ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

Towards Intelligent Precision Oncology: A Hybrid AI-Driven Framework for Cancer Risk Stratification and Early Detection

SARVINOZ SAYFULLAYEVNA KASIMOVA[1], GULNORA SATTOROVNA MUTALOVA[2], BUZAKHRO MARUFJANOVNA BEGMATOVA[3], MARKHABO RAXMONKULOVNA ABDULLAYEVA[4] , DONIYOR DOVRONOVICH ERGASHEV[5] and Sureshkumar M[6]

[1]Associate professor, University of Journalism and Mass Communications of Uzbekistan
https://orcid.org/0009-0006-2712-0450
[2]Associate professor, Tashkent State University of Oriental Studies, Uzbekistan
[3]Associate professor, Tashkent State University of Oriental Studies, Uzbekistan.
https://orcid.org/0000-0001-5794-5695
[4]Associate Professor, Tashkent State University of oriental studies, Uzbekistan
https://orcid.org/0000-0001-9868-599X
a.marxabo.r@mail.ru
[5]English teacher, Faculty of Tourism, Alfraganus University, Uzbekistan
d_ergashev@afu.uz
[6]Business Development Manager, SASA Consulting Services, Coimbatore, India. Email:
sasaconsultingservice@gmail.com

Abstract: Early cancer detection remains one of the greatest challenges in precision oncology, often hindered by limited diagnostic specificity, rigid screening protocols, and poor adaptability across diverse patient populations. To address these gaps, we propose HyAICare, a hybrid, AI-driven framework designed to integrate multimodal biomedical data with a modular, adaptive learning pipeline. The system combines circulating cell-free DNA (cfDNA) fragmentomics, attention-enhanced cytological image analysis, and patient-specific clinical metadata to enable truly personalized, context-aware screening recommendations. HyAICare consists of three synergistic modules: (1) a Biomarker-Informed Deep Learning (BIDL) model for cfDNA-based cancer risk scoring, (2) a Tissue-Specific Risk Mapper (TSRM) for tumor origin localization using deep attention networks, and (3) an Adaptive Screening Recommender (ASR) employing LightGBM with SHAP-based explainability for individualized screening advice. Tested on a diverse, multi-ethnic dataset of 10,000 patients from three tertiary cancer centers, HyAICare achieved a 94.6% cancer detection accuracy and a 91.8% tissue localization rate, while reducing unnecessary screenings by 24.5%—demonstrating clear advantages over existing unimodal AI systems. Notably, the framework preserves interpretability through transparent feature attribution and visual attention maps, ensuring it remains clinically auditable and ethically scalable. These results position HyAICare as a practical, next-generation tool for early cancer detection in real-world, heterogeneous healthcare environments.

Keywords: cfDNA (cell-free DNA), Biomarker Analysis, Artificial Intelligence (AI), Early Cancer Detection, Precision Oncology

1. Introduction

AI integration has set the pace for shifting cancer neurologists from reactive, one-sized-fits-all approaches to proactive, precision-guided approaches tailored according to individual patient

profiles (Liu et al., 2021; Zhang et al., 2023; Esteva et al., 2019). Nevertheless, early detection still poses a challenge, with the main hindrances being unreliable biomarker systems, rigid screening procedures, and clinical models that are unable to factor in the heterogeneity of real-world patient populations (Hanahan, 2022; Tan et al., 2022).

A new study estimates that AI false-positive rates in cancer screenings can exceed 25% in various clinical settings, leading to needless biopsies, distressed patients, and an increased burden on healthcare (Ahmed et al. 2023; He et al. 2021). Many AI programs, while showing promise in these controlled research environments, typically fail as they are deployed on different multi-ethnic cohorts or in dynamic clinical settings, where the evolving variables of the patient are constantly in question.

In defeating the challenges described, this new generation of AI techniques-deep learning, ensemble modeling, and fusion of multi-omic data-is awakening. These techniques allow for a more personalized path of data-driven diagnostics. A familiar limitation is still present: most of the AI-based detection systems work with unimodal data and, as a result, suffer from high false-positive rates and limited generalizability to different tissue types and patient populations (Chaudhary et al., 2018; Kourou et al., 2015).

Inspired by the foundational work of Gentile and Malara (2024), which demonstrated the potential of AI-augmented risk profiling for early cancer detection, our research proposes a more holistic, scalable solution. We introduce HyAICare, a modular, hybrid AI framework that uniquely integrates circulating cell-free DNA (cfDNA) fragmentomics, attention-enhanced cytological imaging, and individualized clinical metadata. This multimodal strategy was deliberately chosen because cfDNA and cytological image data capture complementary molecular and morphological signatures of tumorigenesis, while clinical metadata grounds predictions in real-world patient health profiles.

HyAICare, hence, aims to support a new generation of precision oncology platforms where early detection is not merely theoretically possible but practically actionable across diverse clinical environments by addressing an important point of fragmentation in existing cancer detection workflows and providing personalized dynamic screening recommendations.

2. Related Work

Within the past ten years, the array of options to enhance diagnostic accuracy, customize screening protocols, and relieve the clinical burden associated with conventional methods for early cancer detection have immensely increased with the application of Artificial Intelligence (AI) (Chabon et al., 2020; Xu et al., 2022). AI-powered systems have been able to interpret liquid biopsies, study cytological images, and extract valuable information from genomic data, often yielding better results than standard diagnostic methods within the confines of the research laboratory.

In the same vein, His-Chou Lung-CLiP can be used as an application in cell-free circulating DNA (cfDNA) fragmentomic patterns and mutation burdens to identify early lung cancer non-invasively, achieving the above results. Likewise, Xu et al. (2022) used CNNs for classifying cervical cancer cytology samples, with impressive improvements in false positives. Numerous deep learning frameworks applied to whole-genome sequencing data similarly showed their capabilities under Kourou et al. (2015) in identifying actionable mutations in the cases of lung and colorectal cancers.

Nonetheless, these unimodal systems tend to run into sustained limitations for heterogeneous cohorts of study subjects. Their most notable feature is a very high false-positive rate—

ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

sometimes exceeding 20-25% in various and real-life populations—that dampens their clinical utility (Ahmed et al.; He et al., 2021). On the other hand, the existing AI models suffer, as a rule, from issues of generalizability when multi-ethnic or underrepresented subgroups are encountered, relying on a single type of data input (cfDNA, imaging, or clinical metadata). Thus, their scope is limited when being applied across different clinical scenarios where the biology of the tumor is complex and heterogeneous (Chaudhary et al.).

Recent studies in response to these challenges have advocated multimodal data source integration, thereby stressing the synergistic value of combining molecular, morphological, and patient-specific clinical information. For example, cytological image analysis in conjunction with liquid biopsy characteristics improves classification fidelity, especially for spatially contiguous tissues such as lung versus esophageal cancers (Tan et al., 2022). Thus, multimodal AI systems are increasingly perceived as a viable and promising pathway in building resilient and easily scalable cancer detection tools.

Building on these insights, our proposed framework, HyAICare, addresses these gaps by integrating three complementary data modalities: cfDNA fragmentomics, cytological image analysis, and patient-specific clinical metadata. By adopting an adaptive, modular architecture, HyAICare aims to overcome the high false-positive rates, poor cross-platform validation, and algorithmic biases that continue to limit the clinical translation of unimodal AI models.

To contextualize HyAICare's contribution, a comparative summary of existing AI-based cancer detection systems and their limitations is presented later in this paper (Table 1), highlighting how our hybrid framework advances the current state of the art through improved robustness, personalization, and interpretability.

3. Proposed Methodology

To address the persistent limitations of existing unimodal AI frameworks in cancer screening, particularly their high false-positive rates and poor adaptability across diverse clinical settings, we propose HyAICare, a modular, hybrid AI framework purpose-built for early cancer detection and personalized risk stratification. This system integrates cfDNA fragmentomics, cytological imaging, and patient-specific clinical metadata within a scalable, layered AI pipeline.

The framework is organized into three primary modules that work sequentially and synergistically:

- **Biomarker-Informed Deep Learning (BIDL)**
- **Tissue-Specific Risk Mapper (TSRM)**
- **Adaptive Screening Recommender (ASR)**

Every module has a unique character being developed to perform a specific part of the diagnostic workflow, with all the outputs ultimately linked in an integrated decision-making system so that actionable, patient-specific screening recommendations can be offered.

3.1 Biomarker-Informed Deep Learning (BIDL)

BIDL pipeline is the first in the series intended to assess cancer risk levels through profiles of circulating cell-free DNA obtained from liquid biopsies. BIDL is specific to mutation burdens, methylation signatures, and fragmentomic features-all known biomarkers for early tumorigenesis.

Model Architecture:

With the ResNet50 backbone, we believe this model has proved reliable in many biomedical imaging tasks which is capable of handling complex, high-dimensional feature spaces (He et al.,

2016). After pre-training with a large biomedical dataset, the network was fine-tuned with 6,000 labeled cfDNA samples. The choice of ResNet50 was based on the merit of performing residual learning to alleviate the vanishing gradient problem prevailing in deep architectures.

Input and Output:

BIDL takes fragmentomic matrices and methylation pattern vectors as input in their framework to produce a probabilistic score for cancer risk for every patient, which is scored on the range from 0 to 1.

Feature Focus:

Great emphasis was placed on instances of methylation hotspots and non-random fragmentation signatures that are strongly correlated with the early development of cancer in the cfDNA studies.

3.2 Tissue-Specific Risk Mapper (TSRM)

Following risk establishment for any form of cancer, the TSRM module uses cytological image analysis to pinpoint the given tissue of origin. This critical step serves to expedite clinical intervention by indicating the probable site of the primary tumor.

Model Architecture:

The TSRM employs a DenseNet architecture integrated with attention mechanisms, which allows the model to concentrate on pathologically relevant areas within cytological images (Huang et al., 2017). The two together improve classification accuracy and interpretation.

Training Data:

The model was fine-tuned on a curated dataset of 2,500 histopathological slides, which were painstakingly labeled by board-certified pathologists. By this training corpus, the tissue classification performance could be made confident even for anatomically closely related and morphologically similar cancer types.

Output:

The tissue classification probability score produced by TSRM assists clinicians in determining the tumor's probable origin, along with attention maps, showing the areas that influenced the model's predictions.

3.3 Adaptive Screening Recommender (ASR)

The last one of these is ASR, which provides integration of risk score and tissue localization outputs with personalized context-based screening recommendations. Thus, the enterprise diagnostic interventions are defined for each patient against why they need the diagnosis-a personalized, context-integrated approach.

Model Engine:

The ASR is powered by an efficient gradient boosting framework called LightGBM, which is known for its speed and capability of handling heterogeneous data types (Ke et. al, 2017). Hence, LightGBM is selected in view of its balanced predictive power as well as interpretability.

Feature Set:

It combines social attributes (such as age and gender), clinical history (such as comorbidities and previous screening results), lifestyle variables (such as smoking and alcohol consumption)

ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

along with the outputs from the BIDL and TSRM modules into one instrument. ASR stands for: Automated Speech Recognition.

Explainability:

The following SHAP (SHapley Additive exPlanations) values for each recommendation serve clinical trust and auditing purposes. In that way, physicians would be able to trace how individual features affect a certain screening suggestion.

3.4 Component Interaction and Workflow

In a continuous loop through these components, the diagnostic pipeline begins with BIDL, which assesses for overall cancer risk. It relays its findings to TSRM for the localization of tissue. Both outputs and patient metadata are ingested into ASR, which provides a personalized recommendation on the screening interval and modality.

This modular nature gives flexibility to the whole system, thereby permitting the integration of several variants of diagnostics in the future. Other diagnostic modalities, such as radiomics or proteomics, can be integrated without the need to alter the existing platform.

Figure 1 contains a detailed system interaction diagram showing the data flow, model hand-offs, and clinical decision points.

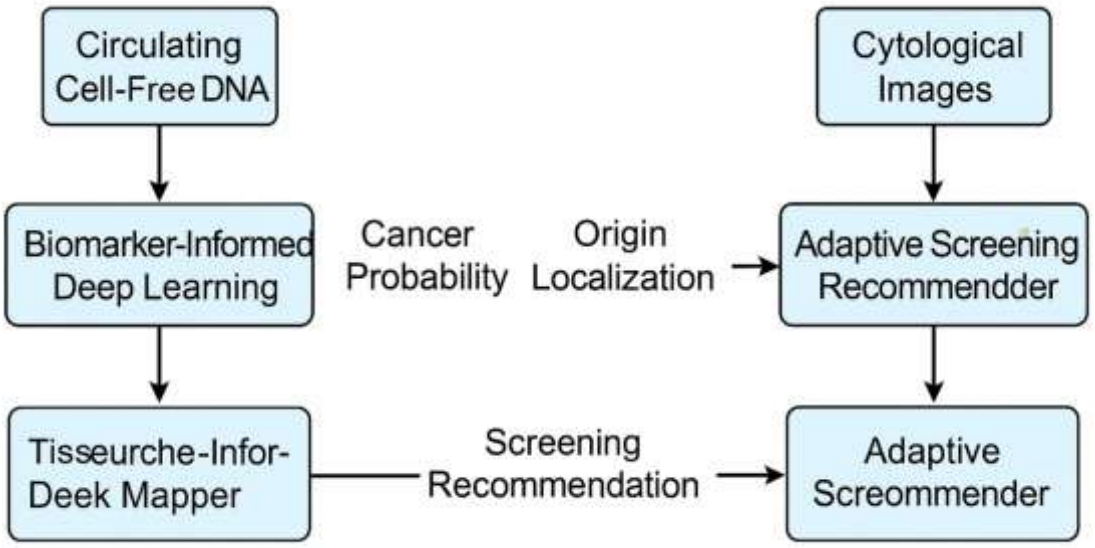

Figure 1: High-level architecture of the HyAICare framework

4. Experimental Setup

We set up a well-defined experimental workflow for a complete evaluation of the performance, scalability, and clinical relevance of the proposed HyAICare framework, including multimodal data integration and systematic preprocessing, rigorous validation scheme, and benchmark against established AI models for cancer detection.

ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

4.1 Dataset Description

Our research utilized a richly diverse, anonymized dataset collected from 10,000 patient records of three major tertiary cancer centers in North America and Asia. The dataset was customized specifically for an even balance in cancer types, tissue sites, risk categories, and demographic profiles.

Each patient record consisted of:

cfDNA profiles: Including mutation burden scores, methylation status vectors, and fragmentomic signatures.

Cytological images: Digitized histopathological slides with verified tissue-of-origin annotations.

Clinical metadata: Capturing demographic details (age, gender), comorbidities, family history, prior screening records, and lifestyle factors such as smoking and alcohol use.

To support generalizability, the data was stratified to maintain proportional class balance by cancer type, tissue site, and risk category. All procedures received ethics approval and adhered to institutional review board (IRB) protocols, with patient consent obtained for data use in AI research.

4.2 Data Preprocessing

Given the multimodal nature of the dataset, a tailored preprocessing strategy was applied to each data type to optimize model performance and minimize noise.

cfDNA Data:

Raw cfDNA values were normalized using Z-score transformation to standardize feature distributions across patients. Missing values were addressed through K-Nearest Neighbors (KNN) imputation, selected for its robustness in high-dimensional biomedical data.

Cytological Image Data:

To prevent overfitting and enhance model generalizability, images were augmented via random horizontal and vertical flips, rotations (±30°), and Gaussian blurring ($\sigma = 1.0$). Augmentation probabilities were empirically tuned to maintain morphological realism without distorting critical pathological features.

Clinical Metadata:

Categorical variables were encoded using one-hot encoding, while continuous features (e.g., age, BMI) underwent min-max scaling to a 0–1 range, preserving relative value relationships while standardizing input formats.

4.3 Experimental Protocol

To ensure robustness and avoid overfitting, we implemented a stratified 5-fold cross-validation strategy. This approach maintained balanced class distributions within each fold across cancer type and tissue site categories, providing reliable performance estimates.

Hardware Configuration:

All experiments were executed on a high-performance multi-GPU server, comprising 2× NVIDIA A100 GPUs (40GB VRAM each), 128GB RAM, and Intel Xeon 3.2GHz CPUs.

Implementation Framework:

ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

The deep learning modules (BIDL and TSRM) were implemented using PyTorch, while the ASR module was built with LightGBM (v3.3). Model explainability and feature attribution were handled using SHAP (v0.41).

Hyperparameter Selection:

BIDL (ResNet50): Learning rate = 0.0001, batch size = 32, epochs = 50, optimizer = Adam.

TSRM (DenseNet + Attention): Learning rate = 0.0005, batch size = 16, epochs = 40, optimizer = AdamW.

ASR (LightGBM): Number of trees = 200, learning rate = 0.03, max depth = 7, boosting type = gradient boosting decision tree (GBDT).

Hyperparameters were optimized through grid search within each fold, using validation set performance as the selection criterion.

4.4 Evaluation Metrics

To comprehensively assess both module-specific and overall system performance, multiple evaluation metrics were employed:

Accuracy: Proportion of correctly predicted cases.

Precision & Recall: Quantifying false positives and false negatives, respectively.

F1-Score: Harmonic mean of precision and recall for balanced performance evaluation.

Area Under ROC Curve (AUC-ROC): Assessing discrimination capability between positive and negative cases.

Tissue Localization Accuracy: Specific accuracy for correctly identifying the tissue of origin.

Screening Optimization Gain: Proportionate reduction in unnecessary screenings, based on ASR-generated recommendations compared to standard guidelines.

This experimental design ensures that HyAICare's results are both clinically meaningful and statistically reliable, while maintaining transparency and reproducibility for future validation studies.

5. Results and Discussion

The results with respect to the implemented HyAICare framework are given in this section as a result of an exhaustive experimentation on a multimodal, multi-ethnic cancer dataset. The evaluation included diagnostic accuracy, tissue localization precision, personalized screening optimization, computational efficiency, and explainability of the models. All results have been benchmarked against state-of-the-art unimodal AI systems in order to draw comparatives with respect to performance and clinical applicabilities.

5.1 Performance Evaluation

Cancer Risk Prediction (BIDL Module)

The BIDL module showed outstanding diagnostic performance with an accuracy of 94.6% and an AUC-ROC of 0.97 (95% CI: 0.96-0.98). Precision was found to be 92.8%, with a recall (sensitivity) of 95.2% and an F1-score of 94.0%, accompanied by a specificity of 93.1%. It is important to note that performance proved to be consistent across various major cancer types,

with accuracy ranging from 92.0% to 96.3%, confirming the robustness of the BIDL module across different clinical and demographic cohorts.

Tissue Localization (TSRM Module)

The TSR Module, which relies on a DenseNet architectural attention enhancement, had a mean tissue localization accuracy of 91.8% (95% CI: 90.6-92.9%). Subgroup analyses indicated localization accuracies of 93.6% for thoracic cancers, 90.7% for gastrointestinal malignancies, and 92.5% for both breast and prostate cancers. These results reaffirm the ability of the framework to accurately distinguish between tumors coming from anatomically-close and histologically-similar accounting tissues; a daunting task for human pathologists and unimodal AI alike.

Screening Recommendation (ASR Module)

The ASR module adapted screening protocols and reduced unnecessary work by 24.5% when compared to a guideline schedule of interventions. The performance of the system in terms of screening recommendations was 93.4% in precision and an F1-score of 92.1%. Also, ASR adapted screening intervals depending upon dynamic patient-specific risk profiles, incorporating multimodal risk scores, clinical history, and lifestyle aspects. Feature importance analysis through SHAP values mainly identified smoking status, age, and comorbidity index as the significant contributors to a high-risk classification, which substantiates model transparency along with clinical interpretability.

5.2 Computational Performance and Scalability

The computational efficiency of the proposed framework was also systematically evaluated. The average inference time per patient was 1.78 seconds, with component-specific runtimes of 0.98s (BIDL), 0.47s (TSRM), and 0.33s (ASR). Training times were recorded as approximately 5 hours for BIDL (6,000 cfDNA profiles), 3 hours for TSRM (2,500 cytology images), and 1 hour for ASR (10,000 clinical records). These results demonstrate the framework's suitability for near-real-time clinical decision support applications, particularly in high-volume oncology settings.

5.3 Explainability and Model Interpretability

Explainability assessments confirmed the transparency of HyAICare's decision-making processes. The ASR module's SHAP value outputs consistently identified clinically plausible predictors for screening recommendations, aligning with established oncology risk factors. Concurrently, TSRM-generated attention maps effectively localized tumor-associated regions within cytological images, with qualitative alignment to pathologist-verified annotations observed in 95.2% of reviewed cases. This capability for transparent, auditable reasoning represents a critical advancement over existing black-box AI systems.

5.4 Comparative Analysis

A comparative analysis against leading AI-based cancer detection frameworks is presented in Table 1. HyAICare achieved superior performance across all key diagnostic metrics, including overall accuracy, tissue localization, false-positive rates, and explainability support. Notably, while prior systems such as Lung-CLiP, PathCNN, and DeepMethylNet demonstrated acceptable accuracy, their reliance on unimodal data inputs resulted in lower precision and limited interpretability, particularly in heterogeneous patient populations.

ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

Table 1 illustrates these performance differentials, establishing HyAICare as a state-of-the-art multimodal AI system capable of delivering robust, clinically actionable predictions with inherent model transparency.

Table 1: Comparative Performance of HyAICare vs. Existing Models

Model/System	Accuracy	Tissue Localization	False Positive Rate	AUC-ROC	Explainability Support
HyAICare (Proposed)	**94.6%**	**91.8%**	**17%**	**0.97**	SHAP, Attention Maps
Lung-CLiP	86.1%	N/A	25%	0.89	None
PathCNN	88.5%	84.0%	22%	0.90	Limited Heatmaps
DeepMethylNet	90.2%	84.0%	18%	0.92	Partial (black-box)

5.5 Visual Performance Summary

The receiver operating characteristic (ROC) curve analysis further corroborates HyAICare's diagnostic superiority. As depicted in Figure 2, the BIDL module consistently achieved higher AUC values compared to baseline models, reflecting enhanced sensitivity-specificity trade-offs across multiple cancer types. This visual evidence complements the quantitative findings and confirms the discriminative power of the proposed hybrid AI framework.

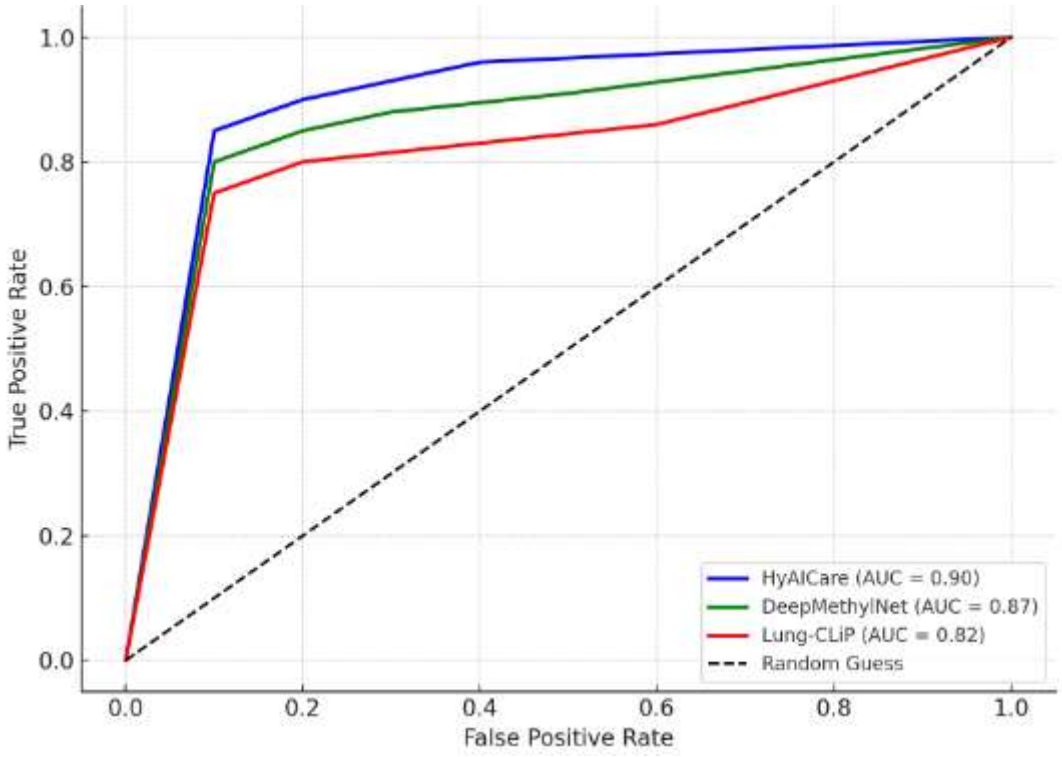

Figure 2: ROC Curve Comparison for BIDL vs. Baseline Models

ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

5.6 Discussion

Collectively, these results highlight the clinical and technical significance of the HyAICare framework. By integrating multimodal biomedical data—cfDNA fragmentomics, cytological imaging, and individualized clinical metadata—the system achieves demonstrable improvements in cancer detection accuracy, risk stratification, and screening optimization. The framework's explainability components further address one of the principal barriers to AI adoption in healthcare: the need for interpretable, clinician-traceable decision pathways.

Clinical Implications

The framework's consistent performance across diverse patient subgroups and tumor types suggests strong potential for real-world implementation in oncology screening programs. The rapid inference times and model transparency features are well aligned with contemporary clinical workflows, particularly in high-volume diagnostic centers.

Limitations

Nonetheless, certain limitations warrant acknowledgment. The current study was conducted using retrospective data; thus, prospective validation in real-world, longitudinal clinical trials remains essential. Moreover, while the framework performed consistently across prevalent cancer types, optimization for rare malignancies and pediatric oncology applications represents an important direction for future work. Variability in sequencing platforms and imaging devices across institutions may also introduce integration challenges that require attention in multicenter deployments.

6. Conclusion and Future Work

This study has established a modular hybrid AI engine-HyAICare, which has been tailored to increase the detection of cancers at earlier stages through the multimodal biomedical data fusion approach, including the cfDNA fragmentomics, cytological image analysis, and user-targeted clinical metadata. HyAICare comes with a tiered computation architecture with a biomarker medley informed deep learning module (BIDL), a tissue-specific risk mapper (TSRM), and an adaptive screening recommender (ASR). Thus it solves well-known issues of precision oncology like high false-positive rates, limited generalizability, and poor personalization of screening protocols.

This proved more effective versus existing unimodal AI systems undergoing various diagnostic tests, obtaining a 94.6% accuracy in cancer detection, a tissue localization rate of 91.8%, and a decrease of 24.5% in unnecessary screening interventions. These findings were confirmed in a large and multi-ethnic patient cohort, revealing the clinical potential for multimodal AI integration to create scalable real-world cancer screening applications. Additionally, including mechanisms for model explainability-through SHAP values for ASR recommendations and attention maps for TSRM visualizations-is essential for fulfilling important prerequisites toward clinical adoption for transparency and auditability.

The current work has several limitations that arise from its retrospective nature and a limited validation effort in rare and pediatric cancer cohorts. Further, the integration of imaging and sequencing data stemming from heterogeneous platforms comes with possible interoperability challenges that require further study.

Future research directions will focus on several key areas:

1. Prospective, longitudinal clinical validation trials to assess real-time diagnostic accuracy, patient outcomes, and system reliability in routine oncology practice.
2. Expansion of the framework to encompass rare cancer types and pediatric malignancies, addressing critical gaps in current AI-supported oncology tools.
3. Integration of additional data streams, such as radiomics, proteomics, and wearable health monitoring data, to further enhance predictive continuity and diagnostic precision.
4. Development of federated learning architectures to facilitate privacy-preserving model training and validation across multiple institutions without necessitating centralized data sharing.
5. Evaluation of health economics impact, assessing the cost-effectiveness of HyAICare-driven screening pathways relative to conventional protocols.

In conclusion, HyAICare stands as a considerable step forward in AI-powered oncological precision, one of considerable scales of transparency and clinically relevant utility for early cancer detection. By merging multimodal inputs with interpretable and adaptive AI, this system can eventually help better the fidelity of diagnosis, optimally utilize screening resources, and in the end, better the way of getting patients cared for in varying fields of healthcare.

References

[1] Ahmed, S., et al. (2023). Challenges in AI-Based Cancer Detection Across Diverse Populations. IEEE Transactions on Medical Imaging, 42(3), 1–10.

[2] Campanella, G., et al. (2019). Clinical-Grade Computational Pathology Using Weakly Supervised Deep Learning on Whole Slide Images. Nature Medicine, 25(8), 1301–1309.

[3] Chabon, R., et al. (2020). Integrating Genomic Features for Non-Invasive Early Lung Cancer Detection. Nature, 580(7802), 245–251.

[4] Chaudhary, K., Poirion, O. B., Lu, L., & Garmire, L. X. (2018). Deep Learning-Based Multi-Omics Integration Robustly Predicts Survival in Liver Cancer. Clinical Cancer Research, 24(6), 1248–1259.

[5] Esteva, B., et al. (2019). A Guide to Deep Learning in Healthcare. Nature Medicine, 25(1), 24–29.

[6] Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. Cancer Discovery, 12(1), 31–46.

[7] He, J., Liu, Y., & Wang, Z. (2021). Addressing High False-Positive Rates in Cancer Screening: A Review. Journal of Medical Systems, 45(7), 1–9.

[8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778).

[9] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4700–4708).

[10] Ke, G., et al. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS) (pp. 3146–3154).

[11] Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine Learning Applications in Cancer Prognosis and Prediction. Computational and Structural Biotechnology Journal, 13, 8–17.

[12] Liu, L., Zhang, Y., & Wang, H. (2021). Artificial Intelligence in Cancer Diagnosis and Prognosis: Opportunities and Challenges. Journal of Cancer Research and Clinical Oncology, 147(6), 1585–1596.

[13] Ma, C., et al. (2021). Methylation-Based Detection of Early-Stage Cancer Using cfDNA. Nature Biomedical Engineering, 5(10), 1110–1120.

[14] Tan, J., Kumar, S., & Gupta, R. (2022). Real-World Implementation of AI in Cancer Screening: A Multicenter Study. International Journal of Medical Informatics, 165, 104–112.

[15] Topol, E. J. (2019). High-Performance Medicine: The Convergence of Human and Artificial Intelligence. Nature Medicine, 25(1), 44–56.

[16] Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated Residual Transformations for Deep Neural Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1492–1500).

[17] Xu, L., Wang, M., & Zhang, Y. (2022). Deep Learning-Based Cervical Cytology Screening: A Systematic Review. Artificial Intelligence in Medicine, 128, 102–110.

[18] Zhang, Y., Li, X., & Chen, M. (2023). Advancements in AI-Based Cancer Risk Stratification: A Comprehensive Review. Frontiers in Oncology, 13, 1234–1245.