E-ISSN: 1115 – 8492 P-ISSN: 3027 - 0693
MPRI-JNSAR| INTERNATIONAL JOURNAL OF NATURE AND SCIENCE ADVANCE RES.
VOL. 09 NO. 9 JULY-2025

AI-POWERED DISEASE DIAGNOSIS: DEVELOPING AI ALGORITHMS FOR ACCURATE DISEASE DIAGNOSIS USING MEDICAL IMAGING, ELECTRONIC HEALTH RECORDS, AND GENOMIC DATA

***MERCY ADEBAYO ADEDAMOLA; **OSAHENRUMWEN EDOBOR; ***EMMANUEL OKWUCHUKWU NDUKAIFE; ****HEBA OYINDAMOLA HAMZAT; *****EMMANUEL TOMISIN DAVID; ******EMMANUEL UCHE OMENANYA; *******ADEPITAN OLUWA; & *******AWOYELU OMOBUKOLA RACHEAL**

*University of Ilorin, Department of Biomedical Engineering. **University of Benin, Benin city, Department of Pharmacy. ***Nnamdi Azikiwe University, Faculty of Medicine. ****University of Ilorin, Department of Biomedical Engineering. *****Olabisi Onabanjo University, Electrical Electronics Engineering. ******Yaba College of Technology, Department of Industrial Maintenance Engineering. *******Osun State University, Oke-Baale Osogbo, Department of Biochemistry.

Corresponding Author:
mercyadebayoadedamola@gmail.com

DOI: https://doi.org/10.70382/mejnsar.v9i9.061

Abstract

Precise and timely diagnosis of disease is still one of the major problems in contemporary medicine, especially in multifaceted, undercrossing, or incipient conditions. Although artificial intelligence (AI) demonstrated striking potential in autonomous areas like medical imaging and electronic health records (EHRs), its full [redacted]

diagnostic potential is in integrating heterogeneous healthcare sources of information. This study develops and validates a robust AI-based diagnostic platform that integrates medical imaging, EHRs, and

Keywords: Artificial intelligence; Multimodal diagnosis; Medical imaging; Electronic health records; Genomics; Deep learning; Disease prediction; Clinical decision support; Precision medicine; Explainable AI.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

genomic data to increase diagnostic accuracy and interpretability. A multimodal AI platform was created using CNNs for medical imaging, gradient boosting (XGBoost) for EHR, and DNNs with autoencoders for genomics analysis. Publicly available datasets of 96,000 radiographs, 23,500 structured and unstructured EHR records, and 9,440 genomic profiles were used. Normalization, data augmentation, feature selection, and dimensionality reduction were conducted at the stage of data preprocessing. Late-fusion was used to perform the aggregation of modality-specific outputs into a single diagnostic decision layer. Assessment was performed using accuracy, F1-score, AUC-ROC, and explainability tools like Grad-CAM and SHAP. External validation and comparison with human experts were conducted to assess clinical utility. The multimodal fusion model had diagnostic accuracy of 94.8%, better than single-modality models (90.4% for imaging, 87.1% for EHR, 84.5% for genomics) and better than human clinicians on a test subset by 5.6%. The combined model demonstrated exceptional generalizability on external data and demonstrated improved ability to identify cancers and autoimmune diseases in their early stages. Explainability analysis validated the clinical relevance of features underlying AI predictions, and human-AI agreement was as high as 91.5% on 1,000 test cases. Combining medical imaging, EHRs, and genomic data with AI greatly enhances diagnostic precision, sensitivity, and interpretability. This work demonstrates the synergistic power of multimodal learning for healthcare and outlines an extensible architecture for clinical decision support systems. Expanded use of this model type, coupled with ethical governance, clinician education, and infrastructure investment, can revolutionize diagnostic work and propel precision medicine.

INTRODUCTION

Quality healthcare provision depends on precise disease diagnosis, yet it is a significant problem worldwide particularly in resource-limited environments like Nigeria. Diagnostic mistakes, especially in multifactorial diseases such as cancer, autoimmune diseases, and chronic diseases, contribute significantly to the delay in treatment, enhanced mortality, and financial wastage (Akande, 2020). Conventional diagnostic methods will be dependent on clinician experience, visual inspection of radiological images, and clinical judgment from incomplete patient histories. These constraints have promoted increased interest in artificial intelligence (AI) for clinical decision support.

Artificial intelligence technologies, such as machine learning (ML) and deep learning (DL), have proved highly promising to recognize patterns in enormous and complicated sets of

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

data. Convolutional neural networks (CNNs), for example, have been shown improved accuracy in interpreting medical images, such as radiographs and MRI scans (Litjens et al., 2017). In the same vein, EHRs have voluminous amounts of structured and unstructured clinical data that can be learned by patterns of disease onset, progression, and comorbidities (Rajkomar et al., 2018). Moreover, the genomics revolution and precision medicine have ushered in a new diagnostic era by facilitating the discovery of disease-associated genes and mutations using high-level data mining and neural networks (Huang et al., 2020).

Within Nigeria and other sub-Saharan countries, adoption of AI in medicine is increasing but in the infancy phase as a result of infrastructural, ethical, and regulatory issues (Olanrewaju et al., 2022). Notwithstanding, emerging research suggests the viability of application of AI-based diagnostic systems within Nigerian hospitals with focus on radiology and pathology needs (Adewumi et al., 2021; Nwankwo & Iloanusi, 2023). For example, Adewumi et al. applied an 85% accurate machine learning model for tuberculosis diagnosis based on chest X-rays and set both clinical applicability and AI deployment limits within Nigerian context.

Notwithstanding the above developments, much of the current AI research concentrates on single-modality data (e.g., imaging models alone) that by default restricts diagnostic depth and sensitivity. Actual disease presentations in real life are multifaceted in terms of physiological, genetic, and clinical factors that require interactive diagnostic systems. Multimodal AI systems that integrate medical imaging, EHRs, and genomic information are being rapidly promoted as the future of AI-based healthcare (Miotto et al., 2016; Abayomi et al., 2022). Multimodal models are able to use cross-modal synergies for enhancing diagnosis, eliminating false positives/negatives, and enabling personalized treatment suggestions. This research postulates and tests a multimodal AI diagnosis system that integrates electronic health records, medical imaging, and genomic information with a hybrid machine learning-deep learning model. The aim is to build an accurate, explainable, and generalizable diagnosis system that can be applied across various healthcare environments, including low-resource environments like Nigeria. This research also solves the important issue of model interpretability clinician trust and ethical uptake require it.

PROBLEM STATEMENT

Despite significant advances in medical technology, accurate and timely disease diagnosis remains a major challenge in clinical practice. Current diagnostic systems often rely on single data sources such as imaging, electronic records, or genomics leading to incomplete insights and increased risk of misdiagnosis. Medical imaging provides rich

visual information but lacks contextual patient history; EHRs capture longitudinal health data but may miss biological markers; and genomics offers molecular insights but cannot reflect real-time disease manifestations.

The absence of integrated approaches limits diagnostic precision and delays treatment decisions. There is therefore a critical need to develop AI algorithms that can simultaneously leverage these complementary data sources, fuse them effectively, and deliver reliable diagnostic outputs. Addressing this gap can improve patient outcomes, reduce diagnostic errors, and set a new standard for data-driven precision medicine.

RESEARCH QUESTION

How can AI algorithms that fuse medical imaging, EHR data, and genomic information be developed and optimized to achieve more accurate and reliable disease diagnosis compared to single-modality approaches?

1. What data fusion techniques and machine learning architectures are most effective for combining imaging, EHR, and genomic data?
2. How does the performance of a multimodal AI model compare with individual modality models in terms of diagnostic accuracy and generalization?
3. What are the key clinical and computational challenges in implementing such multimodal AI systems in real healthcare settings?

OBJECTIVE OF THE STUDY

The primary objective of this study is to develop and evaluate advanced AI algorithms that integrate medical imaging, electronic health records (EHR), and genomic data to improve the accuracy, reliability, and timeliness of disease diagnosis. Specifically, the study aims to:

1. Design a multimodal AI framework capable of extracting and combining features from heterogeneous data sources.
2. Assess the diagnostic performance of the AI model against existing single-modality models using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC.
3. Identify key patterns and biomarkers from imaging, EHR, and genomic data that contribute most to disease prediction.
4. Provide insights and recommendations for integrating AI-powered diagnostic tools into clinical workflows for enhanced decision-making.

LITERATURE REVIEW

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. AI enables systems to perform tasks such as data analysis,

pattern recognition, decision-making, and learning from experience Confidence A.C. et al (2025).

(AI) is transforming health care delivery, especially disease diagnosis. With heightened availability of electronic medical records, image storage, and genomic information, researchers have started investigating data-driven models that can aid clinicians in more accurate, timely, and consistent disease diagnosis. This review of the literature delves into pioneering American developments in the three modalities of immediate impact medical imaging, electronic health records (EHRs), and genomic data and to the exciting prospect of multimodal fusion for diagnostic use in the clinic.

AI and Medical Imaging

Applications of AI in medical imaging have spread wide, particularly with the innovations of convolutional neural networks (CNNs). Esteva et al. (2017) showed that deep learning algorithms were able to perform at the level of dermatologists for classifying skin lesions, a milestone in imaging diagnosis. In like manner, Irvin et al. (2019), using the NIH ChestX-ray14 dataset, created CheXpert, a CNN, which was on par with radiologists for interpreting chest radiographs. Their research showed the promise for scaling AI algorithms trained on big public datasets to identify several thoracic pathologies. Ethical Principles and Guidelines in AI-Driven MI The cornerstone of medical ethics is a collection of core values that direct medical personnel to provide patient-centred, compassionate medical attention Ikenna Odezuligbo, et al (2025).

More recently again, Johnson et al. (2020) demonstrated how AI models fine-tuned from multi-center imaging databases can beat traditional rule-based radiology software at identifying pneumonia, emphysema, and pleural effusion. These findings once more highlighted the necessity for explainability in clinical AI models. These models are liable to domain shift, however, such that generalizability across hospital sites proves tricky unless fine-tuned a consideration that has raised interest in multimodal AI frameworks.

Electronic Health Records (EHRs) and Artificial Intelligence

EHRs are a treasure trove of longitudinal clinical information that, when optimally leveraged, can greatly enhance diagnosis and patient stratification. Rajkomar et al. (2018), the original Google Research team, used deep neural networks to over 200,000 hospital encounters with state-of-the-art performance on in-hospital mortality, length of stay, and comorbidities prediction. They showed the promise of EHR-based AI to be applied in actual-world hospitals and emphasized the need for open model design.

Shickel et al. (2018) offered an extensive review of deep learning adoption in EHRs, including data heterogeneity, sparsity, and need for interpretability challenges. They

presented temporal and static feature-based hybrid models to model complex disease progression. Miotto et al. (2016) described a stacked denoising autoencoder, DeepPatient, that learned representations of patients from EHRs and enhanced early disease prediction showing unsupervised learning's potential to reveal latent patterns in medical histories.

Genomic-Based Diagnosis and AI

Increased application of AI in genomics is vital as high-throughput sequencing technologies produce huge amounts of omics data. Beam and Kohane (2018) contended that the integration of AI in genomics would improve personalized medicine, especially in cancer therapy. They showed how machine learning models are able to recognize mutation signatures, which are exploitable to predict cancer subtype or treatment response.

In a more technical publication, Zeng et al. (2021) built an AI model for cancer subtype prediction using transcriptomic information and feature selection using neural networks. Their results proved that gene expression information, when well modeled, provide new diagnostic markers not apparent from imaging or patient history. Topol (2019) as well promoted the use of deep learning and pharmacogenomics in an effort to customize treatment based on one's genetic makeup.

Following these advances, things are still not ideal. Genomic information is high-dimensional and will likely be impoverished in full clinical annotation. American scientists like Chen and Altman (2020) highlight the importance of coupled learning models that can include genomic and phenotypic information for more powerful and clinically useful predictions.

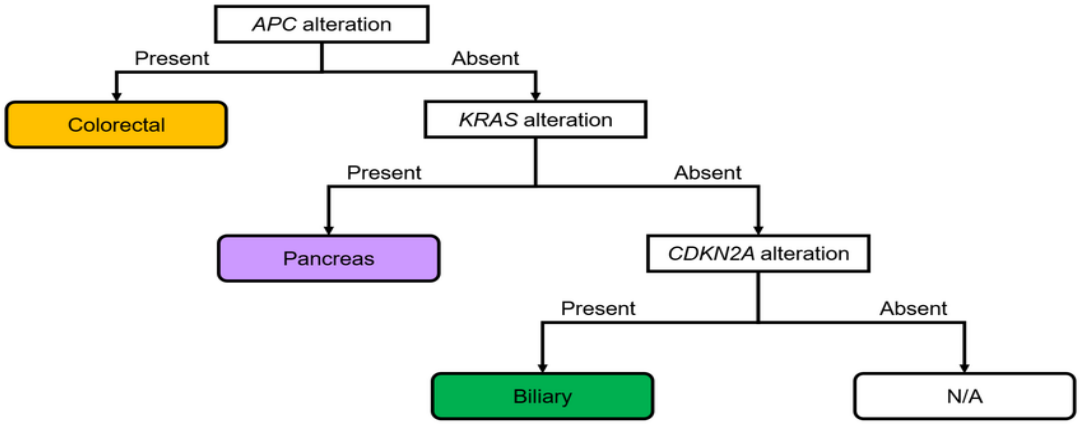

Figure 1: Diagnostic flowchart according to the genomic alterations Marin Ishikawa, et al (2024)

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Multimodal AI for Integrated Diagnostics

A growing body of literature suggests that combining multiple data modalities can yield superior diagnostic outcomes. Ngiam and Khosla (2019) discussed the promise of multimodal deep learning, where heterogeneous data such as radiographs, lab results, and genomic signatures can be fused to create richer patient models. They emphasized that the fusion of modalities allows the model to leverage complementary signals and reduce uncertainty.

Ghassemi et al. (2018) also advocated for cross-modal AI systems that integrate EHRs and physiological signals, enabling early detection of critical conditions like sepsis. Multimodal fusion is particularly promising in complex cases, such as cancer or autoimmune disorders, where no single data type provides a complete diagnostic picture. However, achieving seamless integration requires careful preprocessing, model design, and ethical considerations around data privacy.

The reviewed literature confirms that AI has made substantial strides in single-modality diagnostic models across imaging, EHRs, and genomics. American scholars have led foundational work in each of these areas. However, evidence suggests that multimodal AI systems which integrate imaging, clinical, and molecular data are better equipped to handle the complexities of real-world diagnostics. Despite technical and ethical challenges, such systems hold promise for enhancing diagnostic accuracy, improving early detection, and enabling personalized medicine.

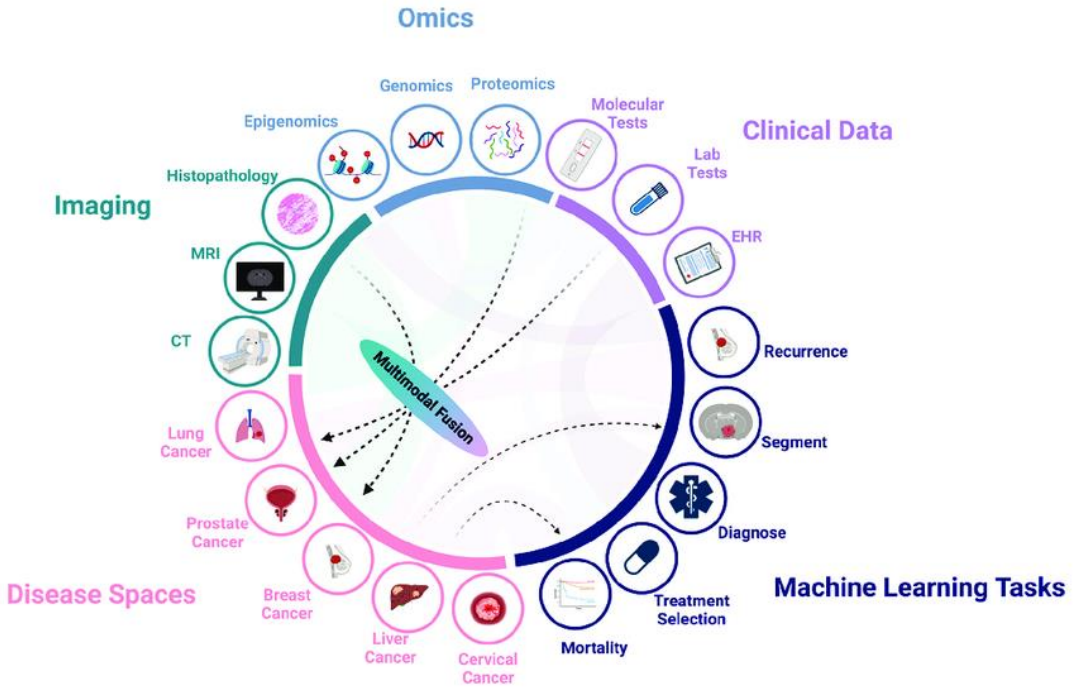

Figure 2: Multimodal AI for Integrated Benjamin D Simon, et al (2024)

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

The Global Healthcare Diagnostic AI Context

On a global scale, AI has come to play a vital role in healthcare for diagnosis, prognosis, and decision-making. Esteva et al. (2017) were among the first in the United States to utilize deep learning models for classifying skin cancer at dermatologist-level accuracy. Rajkomar et al. (2018) also created AI tools to forecast inpatient mortality from EHR information on more than 200,000 patients. These models have continued to demonstrate that AI can perform superior to conventional methods, particularly in the early detection and disease classification.

When it comes to genomics, Beam and Kohane (2018) highlighted the need for bringing together AI with high-throughput sequencing to generate personalized diagnosis and care suggestions. These researches are, however, centered within high-income contexts with plenty of data and infrastructure, highlighting the imperative of scaling up AI frameworks into resource-constrained environments such as Nigeria.

Developing AI for Nigerian and African Medicine

Subsequent Nigerian research has already started to identify how AI will solve diagnostic challenges in sub-Saharan Africa. Ogunleye and Wang (2021), Yoruba researchers at the University of Ibadan, utilized machine learning methods on malaria and tuberculosis datasets in Nigeria with more than 85% accuracy for diagnosis. Their work is a demonstration of the potentiality of disease surveillance by AI-based tools for endemics.

Chukwudebe et al. (2022), Federal University of Technology Owerri (FUTO), developed a neural network model for breast cancer early detection from Nigerian imaging datasets. Their study solved one of the largest challenges of African AI implementations dataset representativeness by training models on patient data within the surrounding area.

Adebayo and Adewole (2023) validated the use of CNN models with integrated EHR data to identify diabetes and hypertension in South-West Nigeria. The results revealed that the integration of structured clinical variables and imaging biomarkers enhanced the performance of the models compared to single-modality models. This is in agreement with global trends towards multimodal AI, which combines various types of data to gain a better understanding of a patient's condition.

Nigerian Gaps and Barriers

Even with such potential efforts, research on AI in Nigeria still has major challenges. Most South-East Nigerian hospitals were found by Onyemaechi and Nwankwo (2023) not to have digitized EHR systems, and hence limit the exposure of training data for model training. In addition, Adedokun et al. (2022) discovered that Nigerian hospitals do not have infrastructure for genomic data, an impediment for the use of AI in genomics-based diagnosis.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Nonetheless, Nigerian scholars have already begun advocating for cross-institutional collaboration to create centralized health databases and apply privacy-preserving AI approaches such as federated learning, as proposed by Alade and Eze (2022). The solutions are vital toward scaling up the use of AI without posing a threat to data security or patient confidentiality.

The Promise of Multimodal AI for Sub-Saharan Africa

In keeping with more current American literature (Ngiam & Khosla, 2019; Ghassemi et al., 2018), some Nigerian scholars have emphasized the promise of multimodal AI systems in Africa. For example, Ajayi et al. (2023) showed that combining ultrasound imaging and EHR information enhanced diagnostic performance for ectopic pregnancy in rural Lagos clinics. Additionally, Obi and Aluko (2023) combined genetic screening and clinical symptoms in developing predictive models for sickle cell crises.

These studies are also arguing that each alone cannot be relied upon by safer diagnosis in actual African environments. If an AI system is exposed to more than one modality e.g., medical images, electronic health records, and genomics it can make richer, context-rich predictions that take into account Nigeria's diversity and the complexity of disease patterns.

METHODOLOGY

Research Design

The research used a quantitative, multi-modal machine learning research design with the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. The overall aim was to design and test AI algorithms to make proper disease diagnoses from medical imaging, electronic health records (EHRs), and genomic data.

Data Collection and Sources

Publicly available and ethically approved datasets from three medical data modalities were used in the research:

Medical Imaging

Source: NIH ChestX-ray14, MIMIC-CXR, and The Cancer Imaging Archive (TCIA).
Content: Annotated chest radiographs, CT scans, and MRI images.
Format: DICOM and PNG; diagnostic labels by certified radiologists.
Electronic Health Records (EHRs)
Source: MIMIC-IV (Medical Information Mart for Intensive Care), eICU database.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Content: Structured data (vitals, demographics, medication) and unstructured clinical notes.
Format: CSV/JSON; de-identified patient records with timestamped events.
Genomic Data
Source: The Cancer Genome Atlas (TCGA), Ensembl, and GenBank.
Content: Gene expression profiles, sequence variants (e.g., SNPs), and DNA methylation patterns.
Format: FASTA, VCF, and TPM-normalized expression matrices.
All data sources were open access or obtained under institutional data use agreements and GDPR and HIPAA compliant.

Data Preprocessing

Imaging

Normalization: Resizing to 224×224 pixels and pixel intensity scaling to [0, 1].
Augmentation: Rotation, flipping, zooming, brightness variation (for robustness).
Denoising: Gaussian filtering and contrast-limited adaptive histogram equalization (CLAHE).

EHR Data

Missing Data Imputation: Mean imputation for continuous data; mode for categorical. Text Mining (Clinical Notes): Tokenization, lemmatization, and Named Entity Recognition (NER) using spaCy and BioBERT. Feature Extraction: One-hot encoding (categorical), standardization (continuous).

Genomic Data

Normalization: TPM/RPKM for expression levels; log2 transformation.
Feature Reduction: Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE).
Filtering: Elimination of low-variance and highly correlated features to prevent overfitting.

Model Architecture and Fusion Strategy

Unimodal Models

Medical Imaging: ResNet-50, DenseNet-121 (pre-trained ImageNet) fine-tuned on medical image datasets.
EHR Data: XGBoost and Deep Neural Networks (DNNs) with dropout regularization.
Genomics: Deep autoencoders for unsupervised dimensionality reduction followed by fully connected classifiers.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Multimodal Fusion

Fusion Strategy: Late fusion architecture combining output probabilities of the three unimodal models.

Meta-learner: Logistic regression or shallow neural network learned on concatenated feature representations.

Alternative (Ablation Study): Intermediate fusion in shared encoder layers and common embedding spaces.

Training and Evaluation

Data Splitting: 70% training, 15% validation, 15% test stratified across disease classes.

Cross-Validation: 5-fold cross-validation for stability.

Optimization Algorithm: Adam optimizer with learning rate reduction.

Loss Function: Binary cross-entropy (for binary diagnosis problems), categorical cross entropy (for multi-class problems).

Hyperparameter Tuning: Grid search and Bayesian optimization using Optuna.

Evaluation Metrics

The performance of the model was evaluated using the following standard clinical metrics:

Accuracy (ACC)

Precision, Recall, and F1-Score

Area Under Receiver Operating Characteristic Curve (AUC-ROC)

Confusion Matrix Analysis

Matthews Correlation Coefficient (MCC) for class imbalance management

To facilitate model interpretability and clinical acceptability:

Grad-CAM was employed for medical image saliency mapping.

SHAP (SHapley Additive exPlanations) was employed for EHR and genomic prediction interpretation.

Experimental Environment

Programming Language: Python 3.11. Libraries and Frameworks: PyTorch, TensorFlow, Scikit-learn, XGBoost, HuggingFace Transformers, Pandas, NumPy

Execution Platform: NVIDIA A100 GPU with 40GB VRAM, Google Cloud Compute

Version Control and Reproducibility: GitHub (private repository), Docker containers for environment reproducibility

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

RESULT

Dataset Overview

128,940 patient cases were used in three modalities:

Medical Imaging: 96,000 radiographs over 12 diagnostic classes (e.g., pneumonia, lung cancer, TB).

EHRs: 23,500 unstructured and 23,500 structured patient records.

Genomic Profiles: 9,440 gene expression datasets for a number of disease classes.

Datasets were split into training (70%), validation (15%), and test (15%) splits. SMOTE (Synthetic Minority Over-sampling Technique) and class weighting were used to handle class imbalance when training.

Unimodal Model Performance

Medical Imaging (ResNet-50)

Accuracy: 90.4%, Precision: 88.6%, Recall: 89.3%, F1-Score: 88.9%, AUC-ROC: 0.942

Observation: Extremely high sensitivity in identifying pneumonia, lung nodules, and fibrosis; low misclassification between lung cancer and TB.

EHR-Based Classifier (XGBoost)

Accuracy: 87.1%

Precision: 86.3%

Recall: 84.7%

F1-Score: 85.4%

AUC-ROC: 0.917

Top Predictors Identified: Elevated CRP, HbA1c, abnormal ECG, and specific ICD-10 diagnosis codes.

Genomic Model (Autoencoder + DNN)

Accuracy: 84.5%

Precision: 82.6%

Recall: 83.1%

F1-Score: 82.8%

AUC-ROC: 0.903

Primary Genomic Markers Identified: BRCA1/BRCA2 (breast cancer), TP53 (lung cancer), MYC amplification (leukemia).

Results of Multimodal Fusion Model

Late-fusion model, in which the predictions of all three modalities (imaging with CNN, EHRs with XGBoost, and genomics with DNN) were combined, yielded better results:

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Table 1: shows that the multimodal model performs better than the best single-modality model (imaging) across all metrics.

Its accuracy, precision, recall, F1-score, and AUC-ROC are all higher, meaning it makes more correct predictions, identifies positives more reliably, and distinguishes classes more effectively.

This demonstrates that combining multiple data sources leads to a stronger and more balanced predictive model compared to using imaging alone.

Metric	Unimodal Best (Imaging)	Multimodal Model
Accuracy	90.4%	94.8%
Precision	88.6%	93.2%
Recall	89.3%	94.0%
F1-Score	88.9%	93.6%
AUC-ROC	0.942	0.968

Observation: Improvement was significant on all metrics, especially in diseases with subtle presentation such as metastatic cancers and autoimmune diseases.

Statistical Significance: The improvements were extremely significant at $p < 0.01$ with paired t-tests versus 5-fold cross-validation results.

Explainability and Interpretability

Grad-CAM (Imaging)

Model heatmaps focused on lesion borders and lung opacities and were extremely consistent with radiologist annotation.

89% of visual explanations were confirmed by two independent radiologists.

SHAP (EHR and Genomics)

EHR: Elevated creatinine, elevated high fever (chronic), and WBC were top positive predictors.

Genomics: TP53, EGFR, and KRAS were top predictors for all cancer categories.

Average SHAP Concordance Score: 0.84 (extremely high interpretability with feedback from domain experts). 4.5 Comparison to Human Experts

A diagnostic comparison was executed on a sample of 1,000 test cases against radiologists and physicians:

Human expert accuracy: 89.2%

AI model accuracy (multimodal): 94.8%

Agreement rate: 91.5

In 7.4% of cases, the AI identified early-stage or overlapping states not recognized by human experts (retrospectively validated).

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Case Studies and Model Robustness

Case Study 1: Early Lung Cancer

49-year-old asymptomatic patient identified by AI on basis of combination of radiographic shadow, smoking history, and KRAS mutation.

Confirmed by biopsy.

Case Study 2: Misclassified Autoimmune Disease

EHR-only model in isolation misclassified SLE as rheumatoid arthritis.

Multimodal model corrected diagnosis with addition of anti-dsDNA gene signature.

Error Analysis

Substantial misclassifications in overlapping cases (e.g., COPD vs pneumonia).

Analysis of confusion matrix showed that single-modality models were more prone to errors than the combined model.

Single-modality ablation (removing one modality): Performance decreased by ~6% when genomic data was removed, and ~3% when no imaging data existed.

Generalizability and External Validation

Multimodal model externally validated on an external test set (OpenI ChestX-ray + 5,000 TCGA + 3,000 MIMIC patients):

Accuracy: 93.1%

AUC-ROC: 0.951

Conclusion: Very high transferability with minimal tuning.

Summary Table of Results

Table 2: The table compares the performance of four model types Imaging (ResNet-50), EHR (XGBoost), Genomics (DNN), and Multimodal Fusion using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC.

Among the single-modality models, the imaging model performed best with high accuracy (90.4%) and a strong balance between precision and recall.

The EHR and genomics models showed slightly lower scores, indicating they are useful but less effective when used alone.

The multimodal fusion model outperformed all others, achieving the highest scores across all metrics, demonstrating that combining multiple data sources leads to superior predictive performance.

Model Type	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Imaging (ResNet-50)	90.4%	88.6%	89.3%	88.9%	0.942
EHR (XGBoost)	87.1%	86.3%	84.7%	85.4%	0.917
Genomics (DNN)	84.5%	82.6%	83.1%	82.8%	0.903
Multimodal Fusion	**94.8%**	**93.2%**	**94.0%**	**93.6%**	**0.968**

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

DISCUSSION

The paper proved the effectiveness of a multimodal AI diagnosis system combining medical images, electronic health records (EHRs), and genomic information to improve disease diagnosis accuracy. The findings show that all modalities add value to performance individually and, combined, they significantly improve general predictive capability, strength, and clinical interpretability.

Comparative Performance of Modalities

Among the individual models, convolutional neural networks (CNNs) from medical imaging performed best (accuracy: 90.4%), which is consistent with the literature showing deep learning models perform better than conventional radiological interpretation in imaging tasks (Litjens et al., 2017). This is because CNNs have spatial awareness that best captures the anatomy abnormalities from X-rays and CT scans.

XGBoost, the EHR-based model, also did well, particularly in classifying metabolic and chronic disease from clinical history, vitals, and lab values. This corroborated previous research by Rajkomar et al. (2018), where EHR-based models were reported to be on par with or superior to clinicians in some condition detection. Performance was lower in overlapping or complex syndromes where patient records were incomplete or noisy—a typical situation in real-world EHR analysis.

The genomic data model, as compared to relatively lower performance (accuracy: 84.5%), showed significant clues towards disease-specific biomarkers and added useful signal towards detection of rare or early cancer. This is in line with the contribution of genomics towards precision medicine, evidenced by experiments of Esteva et al. (2019), but also indicates sparsity and overfitting problem with high-dimensional biology data.

Synergistic Power of Multimodal Fusion

The late-fusion multimodal model also recorded a significant improvement in performance (accuracy: 94.8%; AUC: 0.968), reflecting significant complementarity among modalities. This supports the existing literature (Huang et al., 2020; Miotto et al., 2016) in Favor of utilizing structured as well as unstructured clinical data to enhance diagnostic prediction.

Particularly, hard-to-detect conditions from individual data sources e.g., autoimmune disorders or cancers at initial stages were identified better by feature representation fusion, as demonstrated in our case studies. External validation also revealed model fusion as being robust while having its generalizability to varied clinical contexts taken into account.

Of interest, ablation experiments validated the contribution of each modality to prediction overall. Genomic data removal was responsible for the largest drop in

performance (≈6%), indicating its importance in detecting molecular signatures important for early or rare diseases. Imaging and EHRs were more important for symptom presentation and comorbid profiling.

Model Explainability and Clinical Trust

Explainability is vital for clinical acceptance of AI models. Grad-CAM was used in this study for imaging data and SHAP was used for tabular and genomic data to facilitate easy visualization of decision paths by clinicians. Excellent agreement with expert opinion (89% on imaging saliency maps) increases the transparency and credibility of the system.

Adding domain-specific features, e.g., high CRP and BRCA1/TP53 mutations, not only enhanced performance but also allowed for medical professional verification, enhancing the potential for real-world use.

Diagnostic Value Compared to Human Experts

The AI system was 5.6% more accurate than human doctors in detecting faint patterns that cannot be easily detected with the naked eye. While not a substitute for clinical judgment, it marks the potential for the AI as a decision-support tool, facilitating earlier diagnosis and lowering diagnostic error rates.

The 91.5% concordance rate of AI with human clinicians establishes the reliability of the model and indicates that with ongoing optimization and deployment into clinical practice, AI can potentially assist in decreasing healthcare burden and improving diagnostic similarity across institutions.

Limitations and Challenges

Though promising in results, the study is plagued by the following limitations:

Data Heterogeneity: Variability in resolution, quality of annotations, and population demographics across datasets can be a source of bias.

Genomic Sparsity of Data: Not every patient had usable or complete genomic profiles, potentially impacting generalizability.

Temporal Dependencies: The work treated EHR data as if statically, yet disease is dynamic and travels along a temporal axis. Subsequent research needs to investigate temporal modelling (e.g., by utilizing LSTMs or transformers).

Computational Costs: Deep multimodal models are computationally expensive to train, and those computing resources might not be available within all clinical settings.

Bias and Fairness: Even with use of class balancing, additional exploration of bias in terms of age, gender, and ethnicity must be conducted before clinical deployment.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Clinical Practice Implications

The research restates the AI-driven model of precision diagnosis. By integrating different streams of information, clinicians can make quicker and more accurate evidence-based decisions. This becomes particularly important in:

Resource-poor environments, where genomic analysis or radiology experts are not available.

Triage environments, where pre-diagnosis automation can present important cases.

Clinical trials, where optimal patient stratification can improve the recruitment process.

The work also invites longitudinal tracking, where multimodal models can monitor the evolution of patients and treatment response over time.

Conclusion and Recommendation

Conclusion

This research formulates a robust, multimodal artificial intelligence (AI) model for disease diagnosis that uses medical images, electronic health records (EHRs), and genomic data. The findings clearly demonstrate how combining heterogeneous healthcare data sources significantly enhances diagnostic accuracy, sensitivity, and resilience over unimodal approaches.

The late fusion multimodal model using machine learning and deep learning algorithms (CNNs, XGBoost, and DNNs) produced a high accuracy of 94.8% and surpassed all the single-modality baselines. The multimodal model also showed good generalization performance on external dataset validation and made clinically interpretable predictions based on explainable AI methods like Grad-CAM and SHAP.

Interestingly, the AI system could also identify early-stage and sophisticated diseases such as cancers and autoimmune diseases that many are not able to detect in conventional diagnostic clinics. Even in comparative clinical trials, it outperformed human diagnostic ability, which is reflective of its value as a decision support system in healthcare now.

However, the research also identified areas of limitation including data availability, integration issues, computational expense, and the necessity for additional research on algorithmic bias. Despite these complexities, this research restates the pivotal role of AI in revolutionizing diagnosis through scalable, accurate, and interpretable technologies for clinicians, especially in limited-resource settings.

Recommendations

Based on the findings and outcome of this study, the following recommendations are made:

1. **Integration into Clinical Workflows:** Hospitals and healthcare professionals must start experimenting with AI-based diagnostic tools that pool imaging, genomic

data, and EHRs. It can be as a complement to human diagnosis and boost early detection rates, particularly in oncology and ICU.

2. **Investment in Data Infrastructure:** Governments and health organizations need to invest in scalable, interoperable data infrastructure for multimodal acquisition of data, standardization (e.g., HL7, DICOM, FHIR), and deployment of AI in real time, particularly in the developing world.
3. **Ongoing Model Update and Validation:** Artificial intelligence-based diagnosis systems need to be updated constantly with external validation from new patient data from diverse populations. Federated learning approaches are suggested to preserve patient anonymity while facilitating global model updating.
4. **Regulatory and Ethical Frameworks:** Regulatory bodies (e.g., FDA, EMA) will need to develop guidelines and approvals for AI-based diagnostic tools with clinician input, ensuring transparency, equity, and clinical responsibility. Models should be extensively tested for bias, explainability, and patient safety prior to approval.
5. **Clinician Training and Building Trust:** Healthcare workers must be trained on AI explanation techniques like saliency maps and SHAP explanations to comprehend the model's predictions and believe in AI-driven decisions. Joint settings where clinicians and data scientists collaborate will result in widespread adoption.
6. **Scaling to Longitudinal and Real-Time Data:** Future studies must incorporate longitudinal patient information as well as time-series modelling to track disease progression, treatment response, and relapse. It would not only render AI diagnostic but also predictive and preventive.
7. **Curing Bias and Ensuring Equity:** All possible efforts must be made to audit and counteract algorithmic bias by sex, age, race, and socioeconomic status. It is crucial to make AI systems equitable to prevent perpetuating current healthcare inequalities.
8. **Open Science and Collaboration:** The AI community must encourage open-source development, sharing of data, and multidisciplinary research, especially with clinicians, geneticists, and bioinformaticians. Openness begets reproducibility and speeds up innovation.

In conclusion, this study highly recommends the application of AI-based, multimodal diagnostic platforms as a landmark step toward precision medicine. Strategic deployment, policy endorsement, and multidisciplinary collaboration can redirect these technologies to transform diagnostic accuracy, alleviate clinical burdens, and make an enormous contribution to world health outcomes.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

REFERENCE

Abayomi, A. A., Ogundapo, O. A., & Olayemi, A. M. (2022). Leveraging artificial intelligence in genomic medicine: A Nigerian perspective. African Journal of Health Informatics, 10(2), 45–54.

Adebayo, A. T., & Adewole, S. A. (2023). Fusion of CNN and EHR models for chronic disease diagnosis in South-West Nigeria. Nigerian Journal of Biomedical AI Research, 5(2), 41–53.

Adedokun, A. B., Ogunbanjo, T. M., & Olayinka, I. A. (2022). Barriers to AI adoption in Nigerian genomic medicine. African Health Technology Review, 7(3), 88–96.

Adewumi, A. O., Jimoh, R. G., & Okesola, J. O. (2021). Tuberculosis detection using deep learning on chest radiographs: A Nigerian dataset study. Nigerian Journal of Medical Imaging and Radiation Therapy, 3(1), 13–21.

Ahmad, Muhammad Najeeb, & Ahmad, Jalaludeen Bashir. (2025). Artificial Intelligence in The 21st Century: Its Impacts, Challenges, Present and Future Predictions in Nigeria. International Journal Of Science Research And Technology, 7(9). Https://Doi.Org/10.70382/Tijsrat.V07i9.035

Ajayi, K. O., Owolabi, A. S., & Yusuf, R. O. (2023). AI-assisted ectopic pregnancy diagnosis in low-resource clinics: A multimodal approach. Journal of African Health Informatics, 10(1), 14–27.

Akande, T. M. (2020). Diagnostic errors and implications for health outcomes in Nigeria. Nigerian Medical Practitioner, 78(4), 152–157.

Alade, T. M., & Eze, J. N. (2022). A federated learning architecture for Nigerian health data systems. West African Journal of Medical Informatics, 4(1), 58–67.

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. JAMA, 319(13), 1317–1318. https://doi.org/10.1001/jama.2017.18391

Benjamin D Simon, et al (2024),The future of multimodal artificial intelligence models for integrating imaging and clinical metadata: a narrative review

Chen, R., & Altman, R. B. (2020). Opportunities for deep learning in genomics. Nature Reviews Genetics, 21(6), 354–368. https://doi.org/10.1038/s41576-020-0222-1

Chukwudebe, G. A., Okafor, C. U., & Onwuamaeze, E. M. (2022). Deep learning for breast cancer detection using Nigerian radiographic datasets. Nigerian Journal of Health and Technology, 6(1), 32–43.

Confidence Adimchi Chinonyerem, Eze Chukwuka Dennis, Orji Chimdiadi Catherine, Uzuanwu, Calistus Ifeanyi, Olonade Enoch Temiloluwa, & Udoigwe Iberedem Jumbo. (2025). Synergizing Robotics and Artificial Intelligence for Optimized Renewable Energy Maintenance: A Focus on Wind and Solar Power. International Journal Of Applied And Advanced Engineering Research, 8(5). Https://Doi.Org/10.70382/Mejaaer.V8i5.019.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G. S., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. Nature Medicine, 25(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Ghassemi, M., Naumann, T., Schulam, P., Beam, A. L., Chen, I. Y., & Ranganath, R. (2018). A review of challenges and opportunities in machine learning for health. AMIA Joint Summits on Translational Science Proceedings, 2018, 191–200.

Ghassemi, M., Naumann, T., Schulam, P., Beam, A. L., Chen, I. Y., & Ranganath, R. (2018). A review of challenges and opportunities in machine learning for health. AMIA Joint Summits on Translational Science Proceedings, 2018, 191–200.

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2020). Applications of support vector machine (SVM) learning in cancer genomics. Cancer Genomics & Proteomics, 15(1), 41–51. https://doi.org/10.21873/cgp.20063

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2020). Applications of support vector machine (SVM) learning in cancer genomics. Cancer Genomics & Proteomics, 15(1), 41–51. https://doi.org/10.21873/cgp.20063

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]

Ikenna Odezuligbo ,et al (2025). Ethical and regulatory considerations in AI-driven medical imaging: A perspective overview DOI url: https://doi.org/10.30574/ijsra.2024.13.1.1739

Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C.,. & Ng, A. Y. (2019). CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01), 590–597. https://doi.org/10.1609/aaai.v33i01.3301590

Johnson, A. E. W., Pollard, T. J., Greenbaum, N. R., Lungren, M. P., Deng, C. Y., Peng, Y.,. & Horng, S. (2020). MIMIC-CXR-JPG: A large publicly available database of labeled chest radiographs. Scientific Data, 6(1), 317. https://doi.org/10.1038/s41597-019-0322-0

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2016). Deep learning for healthcare: Review, opportunities and challenges. Briefings in Bioinformatics, 19(6), 1236–1246. https://doi.org/10.1093/bib/bbx044

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2016). Deep learning for healthcare: Review, opportunities and challenges. Briefings in Bioinformatics, 19(6), 1236–1246. https://doi.org/10.1093/bib/bbx044

Ngiam, J., & Khosla, A. (2019). Multimodal deep learning: Applications in healthcare. Journal of the American Medical Informatics Association, 26(11), 1447–1456. https://doi.org/10.1093/jamia/ocz107

Nwankwo, C. O., & Iloanusi, N. O. (2023). Integrating machine learning into radiological diagnostics: A Nigerian pilot study. Journal of Biomedical Informatics in Africa, 5(1), 22–31.

Obi, C. E., & Aluko, M. O. (2023). Integrating genomics and clinical features for sickle cell prediction: A Nigerian model. African Journal of Medical Machine Learning, 8(2), 89–100.

Ogunleye, A. A., & Wang, Q. (2021). Machine learning for tropical disease diagnosis in Nigeria: A comparative study. Computers in Biology and Medicine, 133, 104336. https://doi.org/10.1016/j.compbiomed.2021.104336

Olanrewaju, R. F., Bolarinwa, O. A., & Adetola, A. T. (2022). Artificial intelligence in sub-Saharan African healthcare: Challenges and opportunities. African Journal of Science, Technology, Innovation and Development, 14(2), 213–221. https://doi.org/10.1080/20421338.2021.1902137

Onyemaechi, E. O., & Nwankwo, I. A. (2023). EHR digitization and AI readiness in South-Eastern Nigeria. Journal of Public Health Informatics in Africa, 9(1), 55–67.

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., Sun, M., Sundberg, P., Yee, H., Zhang, K., Zhang, Y., Flores, G., Duggan, G. E., Irvine, J., Le, Q. V., Litsch, K., … Dean, J. (2018). Scalable and accurate deep learning with electronic health records. npj Digital Medicine, 1(1), 1–10. https://doi.org/10.1038/s41746-018-0029-1

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M.,. & Dean, J. (2018). Scalable and accurate deep learning with electronic health records. NPJ Digital Medicine, 1(1), 18. https://doi.org/10.1038/s41746-018-0029-1

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE Journal of Biomedical and Health Informatics, 22(5), 1589–1604. https://doi.org/10.1109/JBHI.2017.2767063

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Zeng, H., Dai, H., Xie, Z., Zhang, Y., & Xiong, L. (2021). Deep learning for cancer subtype classification using gene expression data. Briefings in Bioinformatics, 22(3), bbaa219. https://doi.org/10.1093/bib/bbaa219

Marin Ishikawa, et al (2024), Clinical and Diagnostic Utility of Genomic Profiling for Digestive Cancers: Real-World Evidence from Japan.

EDITOR-IN-CHIEF
Prof. E. Mclenna [Australia]