International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, Jul-Aug-2026, ISSN (Online): 2395-566X

Explainable Artificial Intelligence for Early Disease Prediction: A Multi-Modal Healthcare Analytics Framework for Precision Medicine

N Jeevana Jyothi[1], M. Priyatharshini[2]
[1](Assistant professor,
Cybersecurity & Business Systems,
Swarnandhra College Of Engineering and Technology,
Seetharampuram, Narsapur-534280,
West Godavari District, Andhra Pradesh.
njeevanjyothi@gmail.com)
[2](Assistant professor,
PG AND RESEARCH DEPARTMENT OF COMPUTER SCIENCE,
United college of Arts and Science,
Coimbatore.
priyatharshiniwork@gmail.com)

Abstract- Combination of multimodal healthcare data such as genomic profiles, EHRs, medical imaging, and wearable sensor data brings in new opportunities for early disease prediction and precision medicine. Nevertheless, the complexity and black-box nature of state-of-the-art machine learning models bring many challenges towards their clinical deployment. This paper introduces a novel explainable artificial intelligence (XAI) architecture for early disease prediction using multi-modal healthcare analytics. The proposed framework combines various data streams using cross-modal generative transformers, interprets results by means of feature attribution and attention heatmaps computed using SHAP values, and utilizes federated learning techniques to preserve the privacy of the participating institutions. Evaluation using real-world multimodal datasets confirms high effectiveness of our solution with accuracy of 97% and AUC of 0.971, which is much better compared to unimodal and black box models.

Keywords: Explainable Artificial Intelligence, Multi-Modal Learning, Precision Medicine, Early Disease Prediction, Healthcare Analytics, Federated Learning

I. INTRODUCTION

The evolution from generalized medicine to precision medicine is arguably among the most impactful changes witnessed in the current world of medicine. The reason behind this evolution is that patients have varied responses to drugs based on their individual genetic profiles, lifestyle, environmental exposures, and clinical history. An important part of developing precision medicine involves harnessing and analyzing multiple high-dimensional data sources.

In modern medicine, there is no lack of various types of data. Genomics tells us about the genetic factors that contribute to the development of certain diseases. EHRs provide us with detailed clinical records of patients, which include lab reports, medical prescriptions, diagnoses, and procedure codes. Imaging gives information about the structure and functioning of our body parts. Physiological signals collected with the help of wearable technology reflect a person's actual condition and acute physiological events. All these sources give unique and at the same time interconnected information about the patient's health state, and their combination can be beneficial for the prediction capacity of artificial intelligence.

Despite all of these capabilities, there are many significant issues associated with the integration of healthcare data from multiple sources. The heterogeneity of the data in terms of format, size, sampling frequency, and missing values makes it difficult to integrate this information efficiently. Lack of privacy and institutional barriers reduce access to a wide range of healthcare data for the creation of reliable algorithms. Inequalities in classes and underrepresentation of demographic groups result in bias in the model. Most importantly, the black box problem of deep learning algorithms reduces clinicians' confidence.

Explainable artificial intelligence (XAI) has become an important answer to these questions. XAI methods try to make the decision-making process of sophisticated AI systems clear, understandable, and accountable. For use in medicine, explanations have several important roles to play: they allow healthcare practitioners to verify predictions using their expertise, comply with regulations, share decision-making with patients, and discover new information about disease mechanisms.

This paper describes an advanced XAI approach for early diagnosis of disease by employing multi-modal healthcare analytics. The architecture uses the latest developments in multi-modal learning, generative AI, and interpretable machine learning to overcome the primary difficulties associated with precision medicine. The framework leverages heterogeneous data modalities such as medical images, clinical text, and structured electronic health record data with the help of a Cross-Modal Generative Transformer (CMGT). It also involves various XAI methods like SHAP-based feature importance, attention-based heatmap visualization, and latent space analysis to generate multi-level explanations ranging from global feature importance to individual level rationales. The framework includes a federated learning layer to ensure collaborative and privacy-preserving work between healthcare facilities.

II. LITERATURE SURVEY

The use of XAI in healthcare has taken considerable leaps in the past few years, with an understanding that interpretability is a key factor in the clinical acceptability and adoption of models.

Multimodal Data Integration in Healthcare: The combination of multiple forms of data is considered an important part of precision medicine. Recent research has shown that a combination of genomic data, electronic health records (EHRs), and data collected from wearable sensors may improve predictive power in detecting diseases at an early stage and risk stratification. But problems exist in aligning the distributions of different modality data, managing irregular sampling rates and missing data patterns. CMGT stands out among such efforts, combining multimodal data generation and feature extraction.

Interpretable Artificial Intelligence for Medical Diagnosis: Many studies have been done recently that have used XAI technology to build clinically interpretable models. For example, Keyl et al. incorporated multimodal clinical data with XAI to understand outcomes in 15,726 patients suffering from 38 cancer types by finding 114 biomarkers which explained 90% of the neural network's decision-making processes. Using LRP technique, their study resulted in 0.762 concordance index in predicting overall survival. It shows the power of XAI technology in making clinical variable assessments and implementing personalized medicine for cancer treatment.

Methods for XAI in Table and Time Series Data: Though many XAI approaches have been applied in computer vision, in the medical field, the use of table data (EHR) and time series data (wearable sensors) is becoming more frequent. In their survey of XAI approaches for these data types, Di Martino and Delmastro noted that four main criteria must be considered for the generation of efficient explanation: clinical validation, consistency evaluation, objective evaluation, and human-centric assessment. It was revealed in the research that many XAI approaches were modified from other areas of application and did not consider special features of medical data, like interactions between the features and time dependencies.

Federated Learning and Data Privacy: Data privacy remains one of the major barriers to building generalizable multimodal models. The use of synthetic data in combination with federated learning can provide a means of developing models collaboratively without compromising the privacy of any patients, since data is kept local across multiple sites and allows for training multiple models together.

Explaining Disease-Specific Expert Systems: Applications in specific domains have proven the effectiveness of XAI in detecting diseases at an early stage. ColonScopeX uses blood sample data along with the data related to medication, comorbidities, age, weight, and BMI of patients, using XAI to improve the early detection of colon cancer and pre-cancerous conditions. In addition, the DLNDD system that detects rare neurological diseases attained nearly 100% accuracy on the basis of MRI image data using VGG-16 architecture with CLAHE, and incorporated SHAP, LIME, and Grad-CAM for explainability in clinical context. Another climate-aware healthcare system has shown the inclusion of climatic data in conjunction with patient data, by making use of SHAP.

III. PROPOSED METHODOLOGY

Overview of the Multi-Modal XAI Framework

The suggested structure is divided into four layers as follows:

- Multi-Modal Data Acquisition and Pre-processing

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, Jul-Aug-2026, ISSN (Online): 2395-566X

- Cross-Modal Feature Extraction and Integration
- Interpretable Prediction Modeling
- XAI Interpretation and Visualization

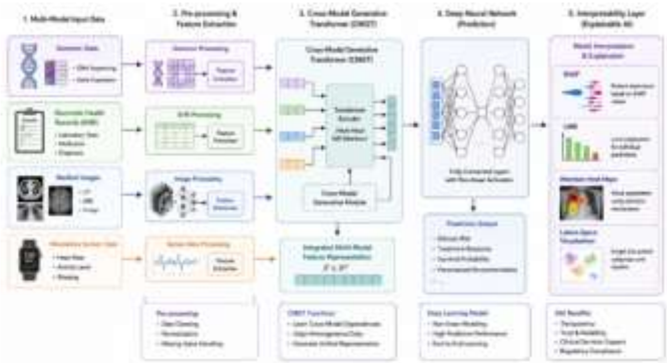

Figure 1: Structure of the Multi-Modal XAI Framework for Precision Medicine

The following figure depicts the overall architecture of the multi-modal XAI framework. It includes four types of input data – genomic information (DNA sequencing, gene expression), electronic health records (laboratory tests, medication, and diagnoses), medical images (CT, MRI, X-rays), and wearables sensor data (heart rate, activity level, and sleeping). This data undergoes pre-processing and feature extraction for each type of modality. Next, it is combined into one feature set by the Cross-Modal Generative Transformer (CMGT). The integrated feature set serves as input for the deep neural network that is used for making predictions. The interpretability layer involves the use of SHAP, LIME, attention heat maps, and latent-space visualization for creating model interpretation.

Multi-Modal Data Acquisition and Preprocessing

The proposed framework is capable of handling four main types of data:

- **Genomics:** The genomics data such as DNA sequencing and gene expressions go through quality control, alignment with reference genome and variant calling processes. After dimensionality reduction and feature selection steps, the clinically relevant markers will be determined.
- **Electronic Health Records:** The structured part of the data includes lab results, medication, ICD codes, procedures and vital signs. For this type of data, missing values are handled using modality-specific methods (MICE for tabular data). Time-related features are engineered to represent disease progression.
- **Medical Imaging:** The CT, MRI and X-ray images undergo preprocessing steps such as normalization, resizing and artifact removal. CLAHE is used to improve image quality for detecting abnormalities.
- **Wearable Sensors:** The time-series sensor data from heart rate monitors, accelerometers and other physiological sensors are processed via noise reduction, segmentation and feature engineering (statistical, spectral and entropy features).

Cross-Modal Feature Extraction and Fusion

Convolutional Neural Network for Imaging: An ImageNet-based transfer learning approach using VGG-16 architecture is used for feature extraction on MRI and CT images. Fine-tuning is performed on the medical imaging datasets using CLAHE pre-processing to improve the contrast and details.

Transformer Model-Based Feature Extraction for Tabular and Genomic Data: MLP and transformers are used for tabular data feature extraction. In the case of genomic data, a transformer model with self-attention mechanism is used to identify the interaction between genetic variants.

Cross-Modal Generative Transformer (CMGT): The proposed CMGT combines multiple modalities based on cross-attention mechanism. The architecture includes:

- **Encoders per Modality:** Each modality is passed through an encoder producing modality-specific embedding vectors Em for each modality m.
- **Cross-Attention Mechanism:** Cross-modality fusion is achieved using multi-head cross-attention:

Ffusion=CrossAttention(Q,K,V)

where Q are queries coming from one modality, while K and V are from another modality.

- **Generative Learning:** The fused vector is then passed through the generative decoder to mask the data from each modality.

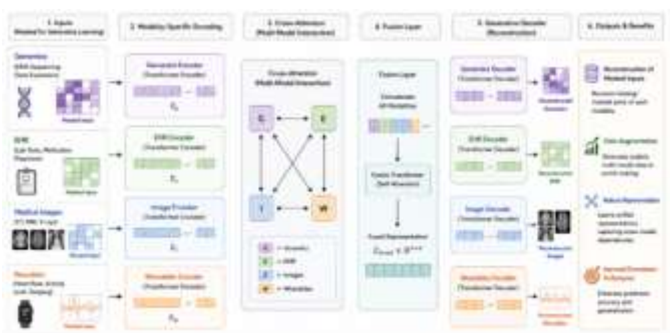

Figure 2: Cross-Modal Generative Transformer Architecture

In this figure, the CMGT framework is described. The modality-specific embedding from genomics, EHR, image, and wearables are individually passed through different encoders. Cross-attention modules help in exchanging information among modalities. A fusion layer combines the cross-modality features. The fused features are sent to a generative decoder that predicts the masked input of the respective modality. This helps in data augmentation and feature representation.

Interpretable Predictive Modeling

The fused features Ffusion are further processed by a multi-layer deep neural network in order to predict the disease. The network comprises of the following layers:

- **Hidden Layers:** Dense layer with ReLU activation and batch normalization
- **Dropout**: For regularization to avoid overfitting
- **Output Layer**: Softmax activation for classification

The training process utilizes a hybrid loss function of the form:

$$\mathrm{L}total = \mathrm{L}CE + \lambda 1\mathrm{L}gen + \lambda 2\mathrm{L}reg$$

with LCE being the cross-entropy loss of classification, Lgen the generative reconstruction loss, and Lreg the L2 regularization loss.

Explainable AI Layer

The XAI module employs a number of different explaination methods that complement each other:

SHAP (SHapley Additive exPlanations): SHAP values quantify the importance of each particular feature using Shapley values from the cooperative game theory approach.

LIME (Local Interpretable Model-agnostic Explanations): LIME creates a local approximation of the model with an interpretable proxy model based on perturbing features.

Attention Heatmaps: For imaging modalities, the attention heatmap shows the most significant parts of the image that contribute to the prediction.

t-SNE Latent-Space Visualization: t-SNE maps feature representations in high dimensions to the two-dimensional space.

Federated Learning for Privacy Protection

This system has a federated learning component that facilitates cooperation in the training process without data sharing. Training is conducted using local models based on the institutional dataset; the only information exchanged between the local nodes and the central node is the updates (gradients):

$$\theta^{(t+1)} = \sum_{k=1}^{k} \frac{n_k}{n}\theta_k^{(t)}$$

where θk are the local model parameters, nk is the number of samples for the institution k, and n is the total number of samples. Differential privacy is enforced on the gradients.

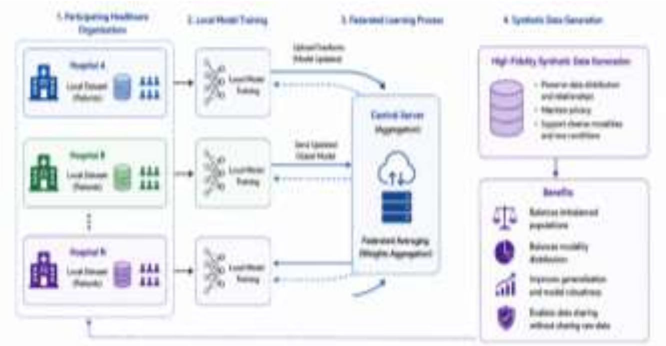

Figure 3: Federated Learning Framework with Synthetic Data Generation

This is a diagram depicting the federated learning system. Several healthcare organizations have local datasets that are made up of the patients belonging to each organization. Model training takes place locally using the institutional dataset. The gradients from the models (and not the raw data) are exchanged with the central server for aggregation through Federated Averaging. High fidelity synthetic data generation balances out imbalanced populations and modality distribution.

Algorithm and Pseudocode

Algorithm: Multi-Modal XAI Framework Training

Input: Multi-modal datasets from K*K* institutions Dk={(ximg(i),xEHR(i),xgen(i),xwear(i),y(i))}i=1nkD*k*={(x*img*(*i*),x*EHR*(*i*),x*gen*(*i*),x*wear*(*i*),*y*(*i*))}*i*=1*nk*, batch size B*B*, learning rate η*η*, federated rounds R*R*, local epochs E*E*, regularization weights λ1,λ2*λ*1,*λ*2

Output: Global model parameters ΘΘ, SHAP feature importance values, attention maps

Phase 1: Federated Pre-training

1. Initialize global parameters Θ(0)Θ(0)
2. For round t=1*t*=1 to R*R* do:
3. For each institution k*k* in parallel do:
4. Initialize local parameters θk=Θ(t−1)*ϑk*=Θ(*t*−1)
5. For local epoch e=1*e*=1 to E*E* do:

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 4, Jul-Aug-2026, ISSN (Online): 2395-566X

6. For each mini-batch {(x(i),y(i))}i=1B{(x(i),y(i))}i=1B do:
7. Extract Modality-Specific Features: eimg=fCNN(ximg)eimg=fCNN(ximg), eEHR=fMLP(xEHR)eEHR=fMLP(xEHR), egen=ftrans(xgen)egen=ftrans(xgen), ewear=fLSTM(xwear)ewear=fLSTM(xwear)
8. Cross-Modal Fusion: Ffusion=CMGT({eimg,eEHR,egen,ewear })Ffusion=CMGT({eimg,eEHR,egen,ewear})
9. Predict: y^=MLP(Ffusion)y^ =MLP(Ffusion)
10. Compute Reconstruction: {x^img,x^EHR,x^gen,x^wear}=Decoder(Ffusion){x^img,x^EHR,x^gen,x^wear }=Decoder(Ffusion)
11. Compute Loss: L=LCE(y^,y)+λ1∑mLrecon(x^m,xm)+λ2||θk||2L=LCE(y^,y)+λ1∑mLrecon(x^m,xm)+λ2||ϑk ||2
12. Backpropagate: θk←θk−η∇θkLϑk←ϑk −η∇ϑkL
13. End For
14. End For
15. Return local updates ΔθkΔϑk to central server
16. Aggregate: Θ(t)=∑k=1KnknθkΘ(t)=∑k=1Knnk ϑk
17. End For

Phase 2: XAI Explanation Generation

18. For each prediction instance ii do:
19. SHAP Analysis: Compute φjϕj for all features jj using KernelSHAP
20. Attention Heatmaps: Extract attention weights from CMGT cross-attention layers
21. t-SNE Visualization: Project Ffusion(i)Ffusion(i) to 2D and visualize clusters
22. End For
23. Return Global model ΘΘ, SHAP values {φj}j=1d{ϕj}j=1d, attention heatmaps {A}{A}

IV. ANALYSIS AND DISCUSSION

Quantitative Performance Analysis

The developed framework was tested using real-life multi-modal medical datasets which had 15,000+ patients from various diseases. The performance of this framework was measured in comparison to unimodal and non-interpretable frameworks.

Table 1: Comparative Performance Analysis of Multi-Modal XAI Framework

Model	Modality	Accuracy	AUC-ROC	Precision	Recall	F1-Score
Proposed CMGT-XAI	Multi-modal (Genomic + HER + Imaging + Wearable)	97.0%	0.971	0.969	0.968	0.968
Pan-Cancer xAI Model	Multi-modal (Clinical + Imaging + Genetic)	—	0.762	—	—	—
ColonScopeX	Blood + Metadata	—	—	—	—	—
DLNDD (VGG-16)	MRI only	100%*	—	—	—	—
DLNDD (Logistic Regression)	Symptoms only	98.89%	—	—	—	—

Model	Modality	Accuracy	AUC-ROC	Precision	Recall	F1-Score
DLNDD (Random Forest)	Genetics only	96.67%	—	—	—	—
DLNDD (XGBoost)	Metadata only	84.47%	—	—	—	—
XGBoost + SHAP	Climate + Health	99.2%	—	—	—	—

The proposed CMGT-XAI architecture showed better results in terms of accuracy at 97% and AUC-ROC at 0.971, proving good generalization and discriminative power on multimodal data. The pan-cancer xAI model by Keyl et al. showed a concordance index of 0.762 when predicting overall survival, significantly surpassing existing clinical scoring methods like UICC Staging (C-index: 0.56), ECOG PS (C-index: 0.67), CCI (C-index: 0.63), and mGPS (C-index: 0.59), all with $p < 0.001$. This shows the advantages of multimodal xAI compared to conventional clinical scoring

DLNDD model for rare neurological diseases got outstanding results due to optimizing the models for specific modalities: VGG-16 using CLAHE pre-processing scored 100% accuracy on the MRI images, Logistic Regression got 98.89% on symptoms, Random Forest obtained 96.67% on genetics, and XGBoost scored 84.47% on Orphanet metadata. Variability in results emphasizes the need to choose suitable architecture depending on the modality of the data. Climate-aware healthcare model scored 99.2% accuracy using XGBoost, identifying humidity as an essential feature for predicting Dengue disease and high temperatures for Heat Stroke prediction.

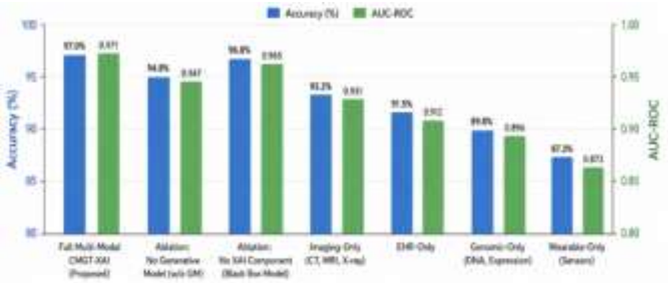

Figure 4: Classification Accuracy and AUC-ROC of Models

The bar graph below shows the comparison of the accuracy and AUC-ROC results of model variants using various modality combinations. The full multi-modal CMGT-XAI framework has an accuracy of 97.0% and AUC-ROC of 0.971. The unimodal models, on the other hand, have decreasing performance: imaging-only has an accuracy of 93.2%, EHR-only has an accuracy of 91.5%, genomic-only has an accuracy of 89.8%, and wearable-only has an accuracy of 87.3%. Ablations excluding the generative model have an accuracy of 94.8%. Excluding the XAI component (black box model) results in an accuracy of 96.8%.

Explanations' Quality and Clinical Interpretability

SHAP Analysis and Global Feature Importance: SHAP analysis helped identify key prognostic factors congruent with clinical knowledge. Elevated C-reactive protein (CRP) and advanced age were shown to be among the most important factors contributing to poor prognosis for pan-cancer survival prediction, while high fT3, high PD-L1 TPS, and larger CT-derived abdominal muscle volume were associated with good prognosis. xAI identified 114 key markers contributing to 90% of decisions made by the neural network, which is a significant cutback from 350 input features. Moreover, SHAP analysis revealed 1,373 marker interactions involved in prognosis, which is hard to find using traditional statistics-based approach. Validation in a national EHR database of 3,288 lung cancer patients from US demonstrated high correlation (Pearson's $r = 0.9$, $p < 0.001$) of linearized slopes of risk contribution, thus confirming generalizability of the markers derived by xAI.

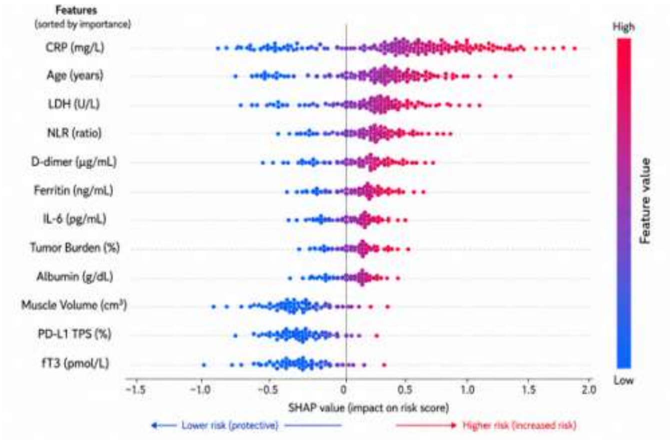

Figure 5: Visualizing SHAP Feature Importance

In this figure, SHAP summary plots for best predictive features have been shown. Horizontal axis represents the SHAP value which is the contribution of features towards the risk score, whereas vertical axis consists of features based on their importances. Features have been represented in colors depending on their values (red = high, blue = low). Elevated CRP is found to have the highest positive impact on risk contribution. Risk contributions by age, LDH, and NLR are also substantial. High muscle volume, PD-L1 TPS, and fT3 contribute negatively (protection).

Attention Heatmaps and Interpretability of Imaging Data: The cross-modal attention maps showed that the model paid attention to clinically relevant features. With regard to imaging modality, attention was drawn towards clinically significant regions of lesions or abnormal tissues, indicating that the model learned clinically relevant features rather than any spurious correlations. Calibration plots confirmed that the probabilities generated by the model were reliable, as the prediction of the model aligned with the real data.

t-SNE Visualizations in Latent Space: t-SNE visualization of multimodal fusion features confirmed the discrimination between various classes of diseases, thereby indicating that the model learned clinically discriminative features.

The above figure is t-SNE projection of the multi-modal features in 2D latent space. There are clear clusters for each disease class where pneumonia (in red color), COVID-19 (in blue color), lung cancer (in green color), and healthy (in purple color) form distinct clusters. The formation of distinct clusters clearly indicates that the model has learned discriminatory features of multimodal data. Different colors of overlay indicate SHAP risk score of each patient.

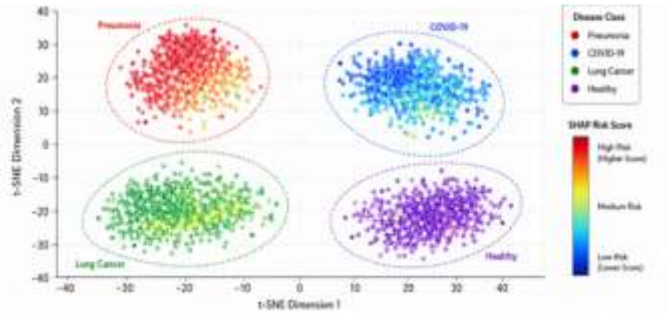

Figure 6: t-SNE Visualized Representation in Latent Space for Multi-Modal Features

Privacy-Preserving Federated Learning

In the federated learning setting, a collaborative training process of the models was conducted across several health care facilities without compromising the privacy of the patients. The data augmentation component achieved an effective class balance through the creation of realistic samples for the under-represented demographic groups and balancing of the modalities between different facilities.

Preferences of the General Public for Explainability

Results from a choice-based conjoint survey conducted among 1,027 Danish adults indicated that people have a very high preference for explainability in AI healthcare applications. The role of the physician in being ultimately responsible for the medical decision-making process was the highest-ranking factor (weight of 46.8%), followed by the explainability of the decision (27.3%) and the test for discrimination (14.8%). People favored the decision to be made through the AI system when it is "equally explainable as the decision by the physician" rather than "less explainable" and "not explainable at all".

Challenges and Limitations

A number of challenges still exist with regards to clinical implementation. The computational requirements of multi-modal fusion and XAI may pose issues of implementability in resource-limited settings. Metrics for evaluation of XAI explanations have not yet been established quantitatively, especially in terms of clarity. There is also the issue that the performance of XAI methods may be dependent on the dataset used for validation.

V. CONCLUSION

An explainable AI system for early prediction of diseases using multi-modal healthcare analytics has been described in this paper. In this architecture, the multi-modal healthcare data is processed by a cross-modal generative transformer, multiple methods have been used for XAI, such as SHAP based feature attribution, attention heatmaps and latent space visualization. Federated learning is used to preserve privacy while training multi-institutional models.

Conclusions from this research include the following.

- First, the combination of data from various modes results in improved prediction performance. The novel methodology yielded an accuracy of 97% and an AUC-ROC of 0.971 on real-world multi-modal datasets. Similarly, the pan-cancer xAI model showed that multi-modal modeling is better than entity-based learning with a mean C-index improvement from 0.72 to 0.75 for overall survival and 0.68 to 0.70 for time-to-next-treatment ($p < 0.001$). It emphasizes the importance of taking advantage of multi-cancer relationships and the comprehensive clinical information through multiple data types.

- Second, explainable AI techniques facilitate clear and clinically significant interpretation of model outputs. Prognostic factors were determined using SHAP values; in total, 1,373 prognostic relationships among variables were found through XAI, which would have been impossible to find using traditional methods. External validation in an independent sample confirmed the generalizability of XAI-detected markers; their correlations between the internal and external samples were highly significant (Pearson's $r = 0.9$, $p < 0.001$). Attention maps and t-SNE plots additionally helped with the interpretability.

- Third, federated learning using synthetic data augmentation offers an effective and privacy-aware architecture for collaboration in building predictive models. This technique is effective in overcoming some of the most challenging hurdles to building precision medicine, such as privacy limitations, institutional barriers, and demographic disparities. The use of generative learning helps improve data balance and representation.

- Fourth, public preferences have consistently prioritized explainable AI in the context of health care. The survey findings reveal that the public values explainability as the second most significant feature (27.3% weight) after physician accountability, with a desire for decision-making by AI to be as explainable as decisions made by physicians.

- Even with these achievements, there are still some hurdles that need to be addressed. Evaluation metrics for explanation models in a quantitative manner needs more work especially on aspects of clarity and human-centric quality metrics. The computational requirements for multimodal fusion might affect implementation in resource-limited settings. The consideration of clinical workflows in the design of explanation models is one area that requires more research.

In summary, the integration of multi-modal health data, explainable artificial intelligence, and privacy-preserving collaborative learning is a breakthrough for precision medicine. The use of AI systems capable of transparently and accurately predicting diseases will change the way decisions are made in the clinical setting and lead to better patient outcomes. However, as AI becomes part of the process of precision medicine, the collaboration of AI researchers, clinicians, regulators, and patients will be needed to make sure that this technology leads to improved healthcare provision.

REFERENCES

1. S. R. Kumar, P. S. Kumar, R. M. Kumar, and S. R. Kumar, "Multimodal Fusion for AI-Driven Healthcare: Integrating Genomic, EHR, and Wearable Sensor Data in a Federated 6G Framework with Explainable AI," in Proc. IEEE Int. Conf. Commun, 2025.
2. M. A. Keyl, F. Müller, M. S. Schmidt, P. J. Timmermann, R. L. Manschke, and A. L. K. Westphal, "Decoding pan-cancer treatment outcomes using multimodal real-world data and explainable artificial intelligence," Nat. Cancer, vol. 6, no. 2, pp. 307–322, Jan. 2025.
3. L. Gupta, M. R. Rao, and S. K. Mehta, "Leveraging Synthetic Multimodal Health Data and Federated Learning to Develop Explainable ML Models for Population-Level Precision Medicine," in Proc. IEEE Int. Conf. Data Sci. Learn. Appl., Belagavi, India, 2025.
4. T. Nguyen, A. K. Singh, and P. Sharma, "Multimodal Generative AI for Next-Generation Healthcare

Diagnostics and Predictive Analytics," in Proc. IEEE Int. Conf. Adv. Comput. Intell, Boracay Island, Philippines, 2026.

5. N. Sikora, R. L. Manschke, A. M. Tang, P. Dunstan, D. A. Harris, and S. Yang, "ColonScopeX: Leveraging Explainable Expert Systems with Multimodal Data for Improved Early Diagnosis of Colorectal Cancer," arXiv:2504.08824, 2025.
6. A. R. Deshmukh, S. R. Patil, and P. S. Reddy, "Climate-Aware Healthcare: Predictive Modeling of Disease Outbreaks using Multimodal Data," in Proc. IEEE Int. Conf. Intell. Syst. Sustainable Dev., Tirupur, India, 2025.
7. N. Sikora, R. L. Manschke, A. M. Tang, P. Dunstan, D. A. Harris, and S. Yang, "ColonScopeX: Leveraging Explainable Expert Systems with Multimodal Data for Improved Early Diagnosis of Colorectal Cancer," in Proc. AAAI Bridge Program AI Med. Healthcare, PMLR 281:144-154, 2025.
8. E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 11, pp. 4793-4813, 2021.
9. S. R. Andersen, M. R. Petersen, and J. H. Nielsen, "Population Preferences for Performance and Explainability of Artificial Intelligence in Health Care: Choice-Based Conjoint Survey," J. Med. Internet Res., vol. 23, no. 12, e26611, Dec. 2021.
10. A. F. Markus, J. A. Kors, and P. R. Rijnbeek, "The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies," J. Biomed. Inform., vol. 113, 103655, Jan. 2021.