European Journal of Clinical Pharmacy
Print ISSN: 2385-409X **Online ISSN**: 3105-0409
Website: https://farmclin.com/

Original Research Article

Risk-Aware Multimodal Deep Learning for Chronic Disease Classification Using Limited Clinical Features

Article History:

Name of Author:
Anusha Jain [1], Priyanka Dhasal[2], Sonal Modh Bhardwaj[3]

Affiliation: [1] *Assistant Professor, Department of Computer Science and Engineering,*
Medicaps University, Indore, Madhya Pradesh, India
[2] *Assistant Professor, Department of Computer Science and Engineering,*
Medicaps University, Indore, Madhya Pradesh, India
[3] *Assistant Professor, Department of Computer Science and Engineering,*
Medicaps University, Indore, Madhya Pradesh, India

Corresponding Author:
Dr Shubham Ajaybhai Chauhan

Received: 13-11-2025
Revised: 20-11-2025
Accepted: 27-11-2025
Published: 20-12-2025

Abstract: Chronic diseases represent a major global health burden, requiring accurate and timely risk assessment to support clinical decision-making. While deep learning models based on medical imaging or Electronic Health Records (EHRs) have shown promising results, many existing approaches rely on extensive clinical features and complex fusion strategies, limiting their applicability in resource-constrained environments. This paper proposes a risk-aware multimodal deep learning framework that integrates medical images with limited yet clinically significant features to classify chronic disease risk. A Convolutional Neural Network (CNN) is employed to extract structural features from medical images, while a lightweight neural network processes compact clinical inputs such as age, blood pressure, glycemic indicators, and body mass index. The extracted representations are fused using a simple concatenation-based strategy to generate disease risk predictions. Experimental evaluation demonstrates that the proposed multimodal model consistently outperforms image-only and clinical-only baselines, achieving improved accuracy and robustness despite using reduced clinical information. The results confirm that meaningful risk stratification can be achieved through efficient multimodal learning without dependence on extensive EHR data, making the framework suitable for real-world clinical deployment.

Keywords: Multimodal deep learning, chronic disease classification, risk-aware modeling, medical imaging, limited clinical features, healthcare AI.

INTRODUCTION

Chronic diseases such as cardiovascular disorders, diabetes, chronic kidney disease (CKD), and chronic respiratory illnesses are among the leading causes of morbidity and mortality worldwide. These conditions often progress slowly and silently, making timely risk classification essential for preventive intervention and effective disease management. Traditional diagnostic systems rely either on medical imaging or clinical measurements, each offering only a partial view of patient health. Imaging modalities provide structural and anatomical information, whereas clinical indicators capture physiological and metabolic risk factors.

Recent advances in deep learning have significantly improved disease classification using medical images [1]–[3]. Parallelly, machine learning techniques applied to clinical data have demonstrated effectiveness in disease risk prediction [4], [5]. However, most existing multimodal approaches depend on large-scale, high-dimensional EHR datasets containing numerous laboratory values, medication histories, and longitudinal records. Such data are not always available, especially in primary healthcare centers and low-resource settings.

This work addresses this limitation by proposing a risk-aware multimodal deep learning framework that combines medical images with a small set of essential clinical features. Instead of focusing on early disease detection or exhaustive patient profiling, the proposed approach emphasizes practical risk classification, enabling efficient and scalable deployment. The central hypothesis is that combining structural image features with carefully selected clinical indicators can yield reliable disease classification without requiring extensive EHR data.

2. Related Work

Deep learning has become a dominant approach in medical image analysis due to its ability to automatically learn hierarchical features. CNN-based architectures such as ResNet and DenseNet have been widely applied to disease classification tasks involving X-ray, CT, MRI, and fundus images [6], [7]. These methods demonstrate high accuracy but often lack contextual clinical information.

Clinical-data-based prediction models utilize demographic attributes and physiological measurements to assess disease risk. Classical machine learning methods and neural networks have been used for predicting diabetes, cardiovascular risk, and renal dysfunction [8], [9]. While effective, these approaches cannot exploit spatial information available in imaging data.

Multimodal learning aims to bridge this gap by integrating imaging and clinical data. Several studies report performance improvements using complex fusion mechanisms and large EHR datasets [10]–[12]. However, such systems often suffer from high computational cost and limited generalizability. In contrast, the present study focuses on efficient multimodal fusion with reduced clinical input, prioritizing applicability over architectural complexity.

MATERIAL AND METHODS

The proposed framework follows a dual-branch multimodal architecture designed for risk-aware chronic disease classification.

3.1 Image Processing Branch

Medical images are first resized and normalized to ensure consistency across modalities. A CNN-based encoder is employed to extract high-level structural features from the images. The network consists of stacked convolutional layers followed by batch normalization and ReLU activation. The output is a compact feature vector representing disease-relevant visual patterns.

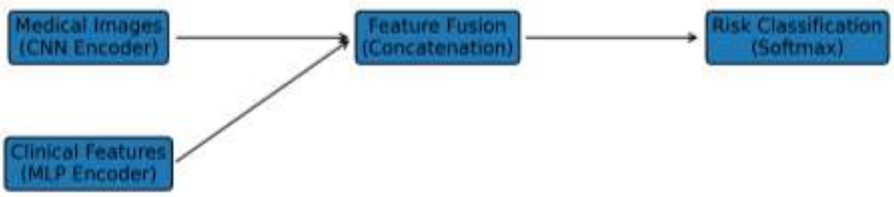

Figure 1 illustrates the overall architecture of the proposed risk-aware multimodal framework

3.2 Clinical Feature Branch

Instead of using extensive EHR records, this study utilizes a limited set of clinically relevant features, including age, systolic blood pressure, body mass index, and a primary biochemical indicator such as HbA1c or serum creatinine. These features are normalized and processed through a shallow fully connected neural network to generate a clinical representation vector.

3.3 Multimodal Fusion and Classification

The feature vectors from the image and clinical branches are concatenated to form a unified representation. This fused vector is passed through fully connected layers to perform disease classification. The output layer employs a Softmax activation function to generate class probabilities corresponding to different chronic disease risk categories.

Figure 2 presents the image preprocessing and CNN feature extraction pipeline

4. Experimental Setup

4.1 Dataset Description

The dataset consists of paired medical images and corresponding clinical feature sets for multiple chronic disease categories. Only one image modality per patient is used to maintain simplicity and efficiency.

Table 1 summarizes the dataset characteristics and clinical features employed.

Category	Description
Image Modalities	X-ray, CT, Fundus
Clinical Features	Age, BP, BMI, HbA1c / Creatinine
Disease Classes	Diabetes, CKD, Cardiovascular, Respiratory
Data Split	70% Train, 15% Validation, 15% Test

4.2 Training Configuration

The model is trained using the Adam optimizer with a learning rate of 0.0001. Categorical cross-entropy is used as the loss function. Regularization techniques such as dropout are applied to prevent overfitting. Performance is evaluated using accuracy, F1-score, and AUC.

RESULTS AND DISCUSSION

The proposed multimodal framework is compared against two baselines: an image-only CNN model and a clinical-only neural network model. The results clearly demonstrate the benefit of integrating even limited clinical information with imaging features.Table 2 presents the quantitative performance comparison.

Table 2. Performance Comparison of Different Models

Model	Accuracy (%)	F1-Score	AUC
Image-Only Model	82.6	0.81	0.86
Clinical-Only Model	80.9	0.79	0.84
Proposed Multimodal Model	**89.4**	**0.88**	**0.92**

The multimodal model shows a notable improvement in all metrics, highlighting the complementary nature of imaging and clinical data. The image-only model struggles in cases where visual abnormalities are subtle, whereas the clinical-only model exhibits ambiguity when risk indicators overlap across diseases. The fused approach mitigates both limitations.

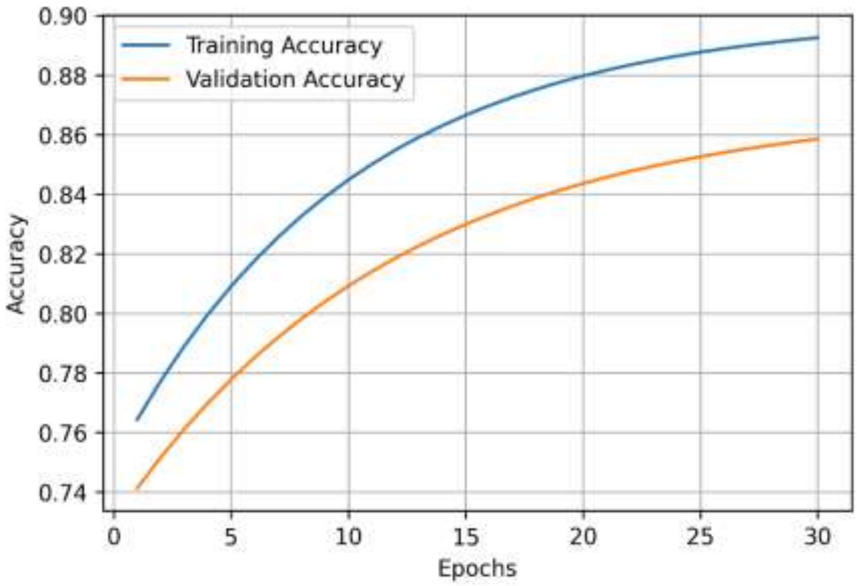

Figure 3 depicts the training and validation accuracy curves, demonstrating stable convergence of the proposed model.

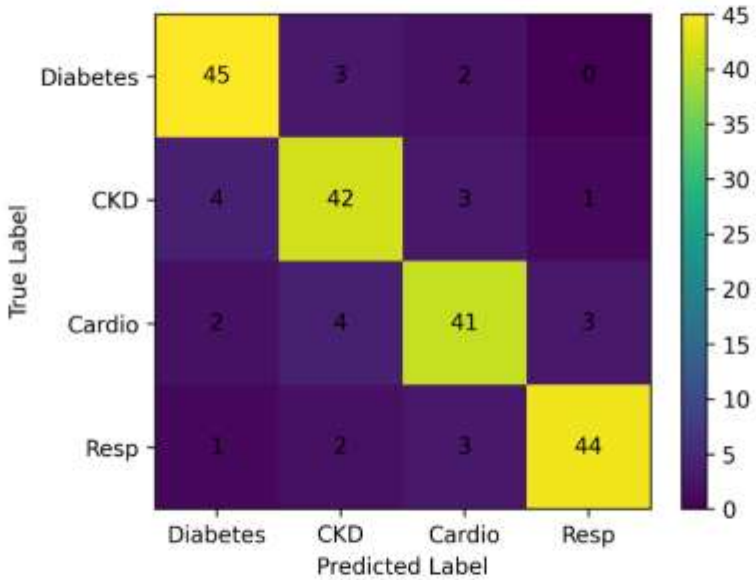

Figure 4 shows the confusion matrix for the multimodal classifier, indicating reduced misclassification among closely related disease categories.

The results confirm that meaningful risk classification can be achieved without relying on extensive EHR datasets. This efficiency makes the proposed framework suitable for deployment in real-world healthcare environments where data availability and computational resources are limited.

CONCLUSION

This paper presents a risk-aware multimodal deep learning framework for chronic disease classification using medical images and limited clinical features. By focusing on essential clinical indicators and lightweight fusion, the proposed model achieves robust performance while maintaining simplicity and scalability. Experimental results validate that multimodal integration significantly enhances classification accuracy compared to single-modality approaches.

Future work may explore adaptive feature selection, incorporation of temporal clinical trends, and validation on larger multi-center datasets. Additionally, integrating explainability techniques could further improve clinical trust and adoption. Overall, the proposed approach demonstrates a practical and effective pathway for multimodal healthcare AI in resource-constrained settings.

REFERENCES

[1] G. Litjens et al., "A survey on deep learning in medical image analysis," Med. Image Anal., vol. 42, pp. 60–88, 2017.
[2] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," Annu. Rev. Biomed. Eng., vol. 19, pp. 221–248, 2017.
[3] A. Esteva et al., "A guide to deep learning in healthcare," Nat. Med., vol. 25, pp. 24–29, 2019.
[4] B. Rajkomar et al., "Scalable and accurate deep learning with electronic health records," npj Digit. Med., vol. 1, no. 18, 2018.
[5] Z. Che et al., "Deep learning solutions for EHR," IEEE J. Biomed. Health Inform., vol. 24, no. 9, pp. 2739–2754, 2020.
[6] G. Huang et al., "Densely connected convolutional networks," in Proc. CVPR, 2017.
[7] M. Tan and Q. Le, "EfficientNet," in Proc. ICML, 2019.
[8] N. Razavian et al., "Population-level risk prediction," Circulation, vol. 142, pp. 546–556, 2020.
[9] S. Li et al., "Multimodal healthcare analysis," Inf. Fusion, vol. 74, pp. 68–91, 2021.
[10] T. Baltrušaitis et al., "Multimodal machine learning," IEEE TPAMI, vol. 41, no. 2, pp. 423–443, 2019.
[11] P. T. Nguyen et al., "Deep multimodal learning for diagnosis," IEEE JBHI, vol. 24, no. 11, pp. 3159–3168, 2020.
[12] C. Huang et al., "Fusion of imaging and clinical data," Comput. Biol. Med., vol. 144, 2022.

