ID 182

Advances in Multimodal AI for Breast Cancer Diagnosis: A Comprehensive Review

MTM Dilshan[1#] and GAI Uwanthika[2]

[1,2]Department of Computer Science, Faculty of Computing, General Sir John Kotelawala Defence University

[#]39-bcs-0004@kdu.ac.lk

Abstract

Breast cancer remains a leading cause of mortality among women worldwide. Early and accurate diagnosis is critical to improving survival rates, yet conventional diagnostic techniques, such as mammography, are often limited in integrating diverse clinical data sources. This review explores the transformative potential of multimodal artificial intelligence models, which combine Electronic Health Records (EHRs) and imaging data to enhance diagnostic precision and treatment planning. We analyze advanced architectures, including Convolutional Neural Networks (CNNs), transformers, and fusion layers, evaluating their strengths, limitations, and clinical applicability. Key challenges, such as data heterogeneity, computational demands, and the lack of standardized datasets, are identified and discussed. This review also highlights the gaps in current research, such as inconsistent evaluation criteria and suboptimal fusion techniques, while proposing innovative solutions, including adaptive fusion methods and lightweight architectures, to bridge these gaps. The findings emphasize the need for standardized datasets and efficient multimodal models to foster broader adoption in clinical settings. Future directions underscore the importance of developing scalable and interpretable systems that can integrate seamlessly into oncology workflows, paving the way for improved breast cancer diagnosis and personalized care.

Keywords: *Multimodal AI, Breast cancer diagnosis, Electronic health records (EHRs), Vision-language models, Clinical oncology*