

CHAPTER 2

LITERATURE SURVEY

Hussain (2024) et al. proposed a 4-view multimodal deep learning framework for breast cancer classification. The study used four standard mammogram views, along with structured clinical and radiological data such as age, BI-RADS score, breast density and family history. ResNet50 was used for feature extraction, and a late fusion strategy was used to combine multimodal features for final classification. The results demonstrated improved diagnostic accuracy compared to unimodal systems, with a reported AUC value of 0.965. However, the framework was primarily focused on mammography and did not integrate ultrasound imaging or a screening and diagnosis workflow aligned with real clinical practice [1].

Wu (2020) et al. presented an application of deep neural networks to help radiologists in breast cancer screening using mammogram images. The proposed system was trained on large mammography datasets and displayed improved cancer detection rates compared to standard screening methods. The study highlighted the potential of Artificial Intelligence-assisted diagnosis to reduce workload and improve accuracy. However, the system relied on only mammography images and did not integrate clinical data or explainability to justify predictions [2].

Zhang (2020) et al. pioneered the development of a multimodal breast cancer diagnosis system using deep learning, which combined mammogram images with clinical data. The integration of age and BI-RADS score helped improve accuracy over unimodal systems. The study showcased the importance of clinical data in diagnosis. Despite these improvements, the fusion approach used fixed weights between modalities [3].

Han (2021) et al. focused on the classification of breast lesions using ultrasound images and CNNs. The proposed system could successfully distinguish between benign and malignant lesions with good accuracy. However, the performance of the system was compromised by ultrasound image quality, noise and operator variations, limiting its use in real-world medical practice [4].

Holste (2021) et al. explored multimodal deep learning approaches by combining medical images with clinical features for breast cancer classification. The study showed that multimodal systems improved performance compared to unimodal systems. The authors highlighted the significance of integrating diverse patient information into AI-based diagnostic systems. However, the proposed architecture did not include attention mechanisms to model the contribution of each modality [5].

Song (2021) et al. proposed a multimodal deep learning system using contrast enhanced mammography for breast cancer diagnosis. Multiple imaging views were used to capture lesion features, resulting in improved classification performance. While the approach showed strong results, it relied on specialised imaging equipment, limiting its application in routine and resource-constrained clinical environments [6].

Kyono (2021) et al. presented a Multiview, multitask CNN for mammography breast cancer diagnosis. The model used shared features across different mammogram views to improve generalisation and classification accuracy. The study illustrated the benefit of Multiview learning in breast cancer detection. However, the approach was limited to mammography and did not integrate ultrasound or clinical data [7].

Khater (2023) et al. developed a system for breast cancer classification using Grad-CAM-based visualisation techniques. The study provided visual explanations highlighting regions that influenced model predictions and improved understanding for doctors and radiologists. Although the model improved trust and transparency, it was evaluated on a single modality and did not integrate multimodal data [8].

Wu (2023) et al. proposed a deep learning approach for breast cancer diagnosis using mammography images. The model was based on contrastive learning and improved feature representation by learning similarities across different views, and demonstrated high classification performance. However, the approach required large-scale datasets and did not integrate clinical data [9].

Zhou (2024) et al. pioneered an attention-based multimodal deep learning framework for breast cancer diagnosis by integrating medical images with clinical data. The attention mechanism enabled the model to learn the relative importance of different modalities, resulting in improved performance and interpretability. However, the system lacked a screening and diagnosis pipeline that mimics clinical workflows [10].