



OPEN ACCESS

EDITED BY

Sokratis Makrogiannis,
Delaware State University, United States

REVIEWED BY

Domenico Pomarico,
University of Bari Aldo Moro, Italy
Amrutanshu Panigrahi,
Siksha O Anusandhan University, India
Meng Wang,
Liaoning Technical University, China

*CORRESPONDENCE

Nazmul Ahasan Maruf
✉ nmaruf0001@stu.kau.edu.sa

RECEIVED 09 June 2024

ACCEPTED 10 March 2025

PUBLISHED 01 April 2025

CITATION

Maruf NA and Basuhail A (2025) Breast cancer diagnosis using radiomics-guided DL/ML model-systematic review and meta-analysis. *Front. Comput. Sci.* 7:1446270. doi: 10.3389/fcomp.2025.1446270

COPYRIGHT

© 2025 Maruf and Basuhail. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Breast cancer diagnosis using radiomics-guided DL/ML model-systematic review and meta-analysis

Nazmul Ahasan Maruf* and Abdullah Basuhail

Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

Cancer is one of the leading causes of death on a global scale, whereas breast cancer is the type of cancer that affects the most women. Early detection and accurate staging are essential for effective cancer treatment and improved patient outcomes. Recent developments in medical imaging and artificial intelligence (AI) have created new opportunities for breast cancer detection and staging. Medical image analysis techniques, including radiomics, machine learning and deep learning, have shown promise for breast cancer detection and stage estimation. The goal of the systematic review and meta-analysis is to evaluate and examine the state-of-the-art implications of radiomics-guided deep learning (DL) approaches for breast cancer early detection utilizing different medical image modalities. The selection criteria were established on the basis of the PRISMA statement. Our research employs a PICO structure and text mining technique (Topic Modeling) using Latent Dirichlet allocation (LDA) approach. The primary objective of the search was to conduct a thorough evaluation of the literature related to radiomics analysis and breast cancer in the fields of medical informatics, computer vision, and cancer research. Subsequently, the investigation concentrated on the fields of medical science, artificial intelligence, and computer science. The inquiry encompassed the years 2021 to 2024. The QUADAS-2 instrument is employed to evaluate the articles to ensure their quality and eligibility. Feature extraction methods that employ radiomics and deep learning are extracted from each study. The sensitivity value was pooled and transformed using a random-effects model to estimate the performance of DL techniques in the classification of breast cancer. The systematic review comprised 40 studies, while the meta-analysis consisted of 23 studies. The research studies employed a variety of image modalities, radiomics, and deep learning models to diagnose breast cancers. Ultrasound and DCI-MRI are the most frequently employed image modalities. The pyradiomics python package is employed to extract the radiomic features, and CNNs, ResNet, and DenseNet models are employed to extract the deep features. The LASSO (13) and *T*-test (9) statistical models are the most commonly used for feature selection. The most widely used deep learning models for breast cancer classification are ResNet and VGG. This systematic review and meta-analysis examined the feasibility of employing radiomics-guided deep learning/machine learning models for identifying breast cancer. The studies yielded positive results, as specific models demonstrated remarkable precision in distinguishing between malignant and benign breast tumors. However, there is a wide variety of variations in the designs of studies, the architectures of models, and the methodologies used for validation. Further research is required to verify the results of this study and to

investigate the potential of deep learning models guided by radiomics in the early detection of breast cancer.

KEYWORDS

breast cancer, medical image processing, radiomics, deep learning, AI, systematic review and meta-analysis

1 Introduction

Cancer is one of the primary causes of mortality worldwide, whereas breast cancer is the type of cancer that affects the greatest number of women worldwide (Sung et al., 2021). In 2020, approximately 2.3 million new cases of breast cancer were diagnosed, making it the most prevalent cancer among women globally (Lei et al., 2021). Although the death rates have decreased due to innovations in breast cancer early detection and diagnosis, it remains the second-leading cause of death for women (Arnold et al., 2022). However, effective cancer treatment and enhanced patient outcomes are contingent upon early detection and accurate staging. Staging is a comprehensive procedure that involves assessing the extent of cancer progression, including whether it has progressed from the breast to other regions of the body. Accurate staging could be useful in choosing the most effective treatment and determining the patient's prognosis.

In the past decade, there has been a growing topic of discussion on the significance of medical imaging techniques in breast cancer staging (Balkenende et al., 2022). Mammography is one of the most widely used image modalities for detecting breast cancer, and recently, with enhanced imaging approaches, it's become a potential tool for detection and diagnosis (Tsarouchi et al., 2023). A follow-up examination is crucial for detecting abnormalities identified during mammography and assessing the degree of dense breast tissue (Gatta et al., 2023). Mammography may be supplemented with other imaging modalities like computed tomography (CT), ultrasound, magnetic resonance imaging (MRI), and positron emission tomography (PET), as mammography has some limitations (Ha et al., 2023). Dynamic MRI (DCE-MRI) is a process of contrast enhancement based on the usage of images, wherein the time-dependent transformation in contrast enhancement is observed precisely (Ingrisch and Sourbron, 2013). This technique can help identify areas of the body where blood vessels associated with cancer are growing and determine the size and depth of the tumor. MRI also has one main negative aspect concerning unsatisfactory specificity that might lead to probable false positives, in addition to raising the overall cost compared to mammography and ultrasound (Zhang et al., 2023). Besides, it takes a long evaluation period (Wang L. C. et al., 2023). Elastography with ultrasound is a technique that is used to determine the body tissues' stiffness (Ditunno et al., 2023), which might be the symbol of malignant growth. Image processing is an essential part of the ultrasound screening technique used to examine and expose soft tissues. It enables the detection and description of breast abnormalities (Cè et al., 2023). Studies indicate that ultrasound

imaging techniques are perfectly safe for their frequent use and have the unique property of being free from radiation (Abhisheka et al., 2023).

Recent advances in medical imaging and artificial intelligence (AI) have created new opportunities for the detection and staging of breast cancer. Techniques of medical image analysis, such as radiomics and machine learning, have demonstrated promise for breast cancer detection and stage estimation (Wang Q. et al., 2023). Radiomics as a field of study is about obtaining quantitative features from medical images, which help to visualize and identify the patterns and biomarkers related to cancer growth (Rizzo et al., 2018). These attributes process information about the shape, texture, intensity levels, and spatial relations between the data in the images (Gupta et al., 2024). Radiomics analysis embraces most of the medical imaging modalities like computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, positron emission tomography (PET), and others (Peng et al., 2023). Those modalities should be an integral part of any analysis that seeks to establish the impact radiomics has on imaging for diagnosis, prognosis, treatment planning, or disease monitoring, specifically in cancer.

Deep learning (Guo et al., 2016; Dhar et al., 2023; Li X. et al., 2023), a subfield of machine learning, has shown enhanced performance in medical image analysis. In the research of Yuan et al. (2023), they emphasize that the CNN (convolutional neural network) can be trained to accurately recognize patterns and markers associated with breast cancer using medical images. Based on radiomics characteristics, machine learning and deep learning algorithms can be trained to develop predictive models for breast cancer detection and staging. Such models have demonstrated high accuracy rates, outperforming conventional diagnostic techniques. Early detection and accurate staging of breast cancer are crucial for selecting the optimal treatment and enhancing patient outcomes. Through using deep learning methods for medical image analysis, we can make our approach more accurate and go into the higher stages too, which in turn can lead to improved treatment outcomes and survival rates.

Despite the fact that a substantial amount of work has been done in this area in recent years, there is still potential for improvements, as the accuracy of radiomics-guided breast cancer detection and stage prediction can significantly impact the diagnosis of this disease. Hence, to lead our research on radiomics-guided breast cancer detection, we are now assessing the current state-of-the-art literature through a systematic review and meta-analysis. Our systematic review focused on the following questions:

- What are the most prominent deep learning and machine learning models that have been developed in state-of-the-art literature, and how are those models analyzed with performance metrics?
- What are the widely used feature extraction and selection processes in breast cancer detection-based medical image processing?
- How is radiomics-guided feature extraction made in cutting-edge research?

The forthcoming sections of the paper will be presented in the following arrangement: Section 2 describes the methodology of a systematic review taken. In the result Section 3, the findings of the study are provided. Section 4 discusses the findings. Lastly, Section 5 concludes the inferences from the analysis.

2 Methodology

Systematic reviews and meta-analyses serve as an essential conduit in breast cancer research for appraising the efficacy of radiomics analysis using deep learning techniques. This systematic approach carefully weighs and combines ongoing research to get an overall verdict on the effectiveness of deep learning radiomics in improving the early detection, diagnostic, and treatment functions of breast cancer.

2.1 Search strategy

In order to conduct a comprehensive and systematic search, we devised a search strategy to locate pertinent material. The search technique was customized for five databases: Scopus, Web of Science, Science direct, IEEE and Google scholar. The search terms used were “Radiomics” AND “Radiomics Analysis” AND “Deep learning” AND “Breast Cancer.” The search included all data from the database’s establishment until 2021, specifically focusing on journal articles, conferences and review articles that were published only in English.

2.2 Eligibility criteria

To ensure the eligibility of studies for our present study, we established precise criteria for inclusion and exclusion. Additionally, we created a PICO structure specific to our research, which comprises the following components: P, Breast Cancer Patient; I, Deep Learning and Image Processing Approaches; C, Radiomics Analysis; Radiomics Guided; O, Image Classification, Segmentation, Prediction, Detection, and Medical Image Analysis. In order to minimize publication bias, we also implement the text mining technique (topic modeling) for knowledge discovery using the latent Dirichlet allocation (LDA) approach.

2.3 Inclusion criteria

In our systematic review and meta-analysis, we established inclusion criteria based on many factors. All the research conducted was centered on the English language. The studies specifically targeted breast cancer, and the reported findings were based on the analysis of deep learning and radiomics features in breast cancer images.

2.4 Exclusion criteria

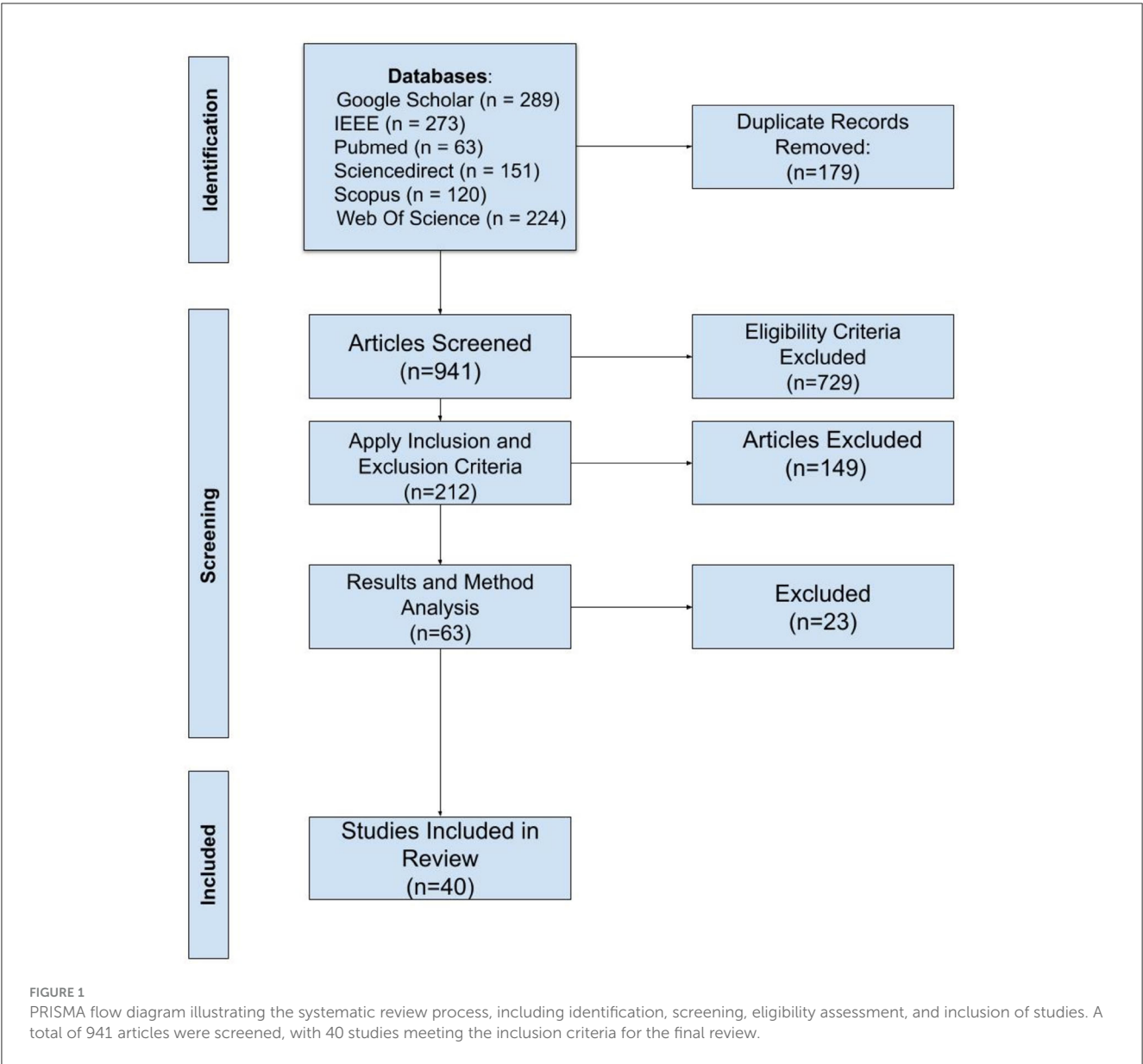
We have excluded from our systematic review and meta-analysis any studies that meet particular criteria. These requirements include research that does not analyze the results of deep learning approaches, case report studies, book chapters, conference abstracts, comments, letters to the editor, or studies published in languages other than English. In addition, research focusing on topics other than breast cancer is also eliminated.

2.5 Selection criteria

The selection criteria were derived from the PRISMA Statement (Moher et al., 2009). The search primarily focused on a comprehensive assessment of the literature pertaining to Radiomics Analysis and Breast cancer within the domains of cancer research, medical informatics, and computer science. The inquiry thereafter focused on the subject areas of computer science, Artificial Intelligence, and medical science. The search included the time period from 2021 to 2024. Articles published before to 2021 were omitted from the search results. The search mostly focused on certain keywords such as Deep learning, Radiomics, Radiomics Analysis, Breast cancer, and image processing. The selection method started by eliminating any redundant items from the study selection process. The studies underwent a two-step procedure, which included screening the titles and abstracts, followed by an in-depth review of the entire texts. Only pertinent papers that satisfied the predetermined criteria for inclusion were selected for comprehensive evaluation after conducting a thorough examination of the title and abstract. Figure 1 presents the PRISMA flow chart, which visually represents the process of selecting studies for the research.

2.6 Quality assessment

A quality appraisal tool called QUADAS-2 was used to assess the quality of studies on diagnostic tests (Whiting, 2011). This tool contains four questions, which are categorized into four different domains. The categories include patient selection, index tests, time and flow, and reference standards. Here are the questions: (1). Could the selection of patients have introduced bias? (2). Was the threshold value pre-specified, if threshold value used? (3). Is the reference standard likely to correctly classify the target condition? (4). Were all patients included in the analysis?



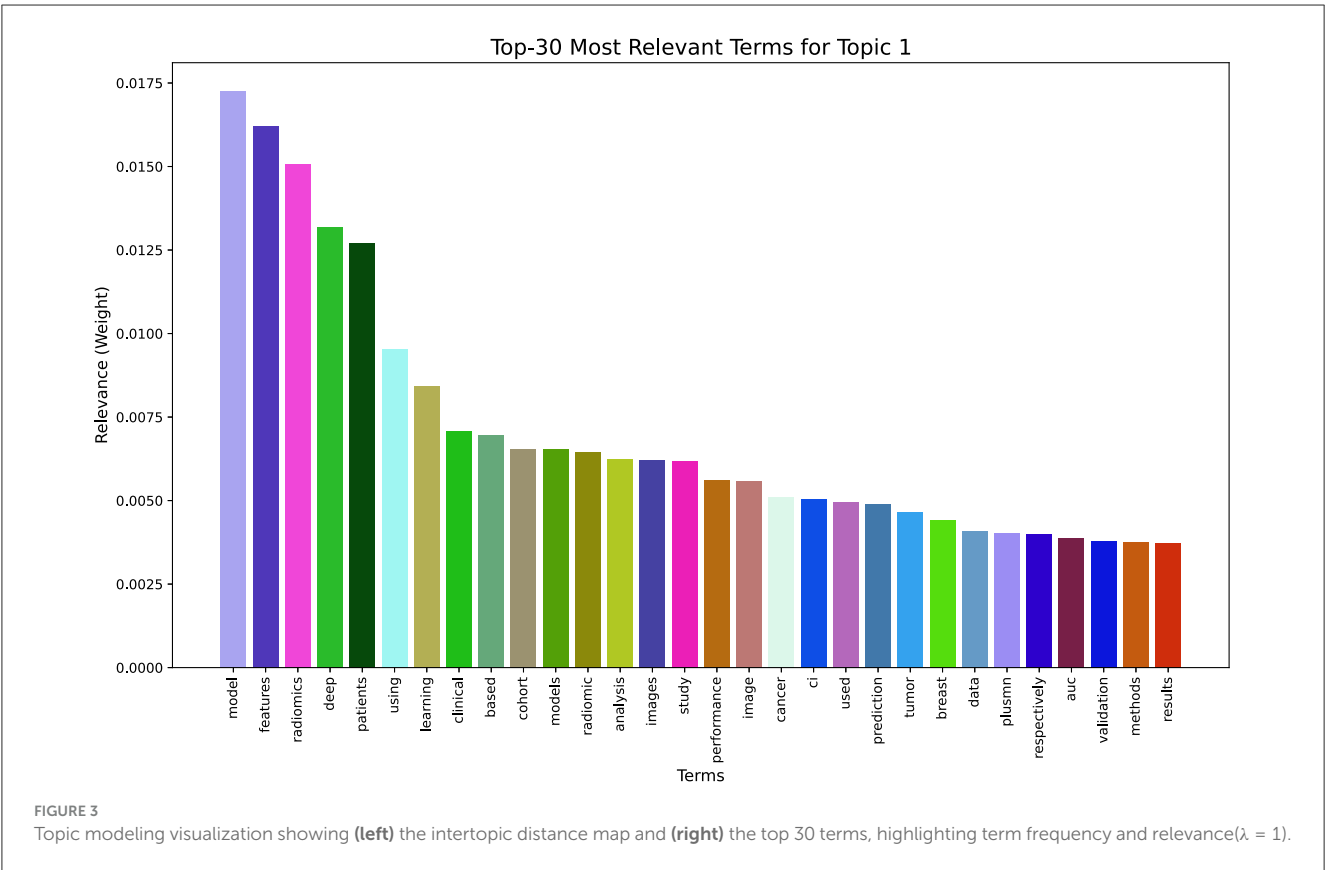
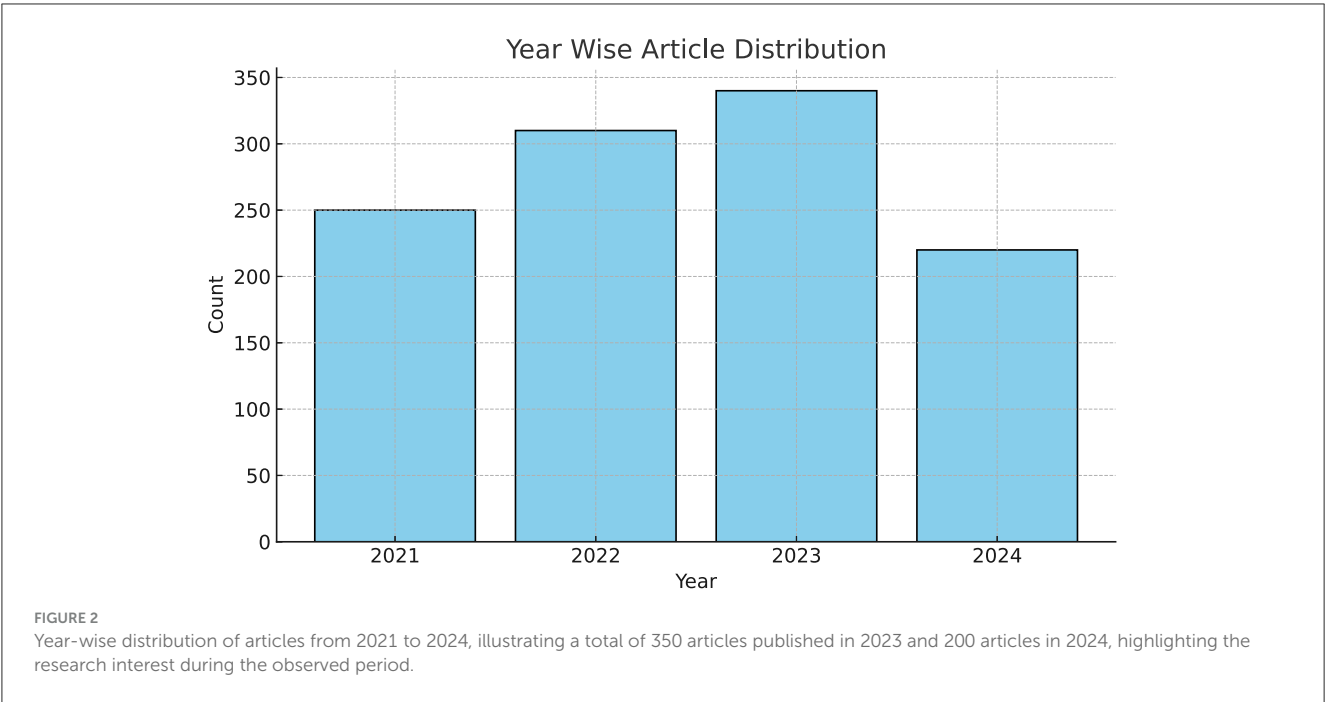
2.7 Meta-analysis

We employed a correlation coefficient-specific random effect model for correlation data with the application of sensitivity and sample size analysis. The web application [Jamovi (Version 2.5), 2024; Hornik, 2012; Viechtbauer, 2010] was used to effect and conduct the meta-analysis of all the data.

3 Result

The initially conducted electronic systematic search yielded 1120 studies from databases such as PubMed, Google Scholar, IEEE, Science Direct, Web of Science, and Scopus. Following the elimination of the duplicate, we discovered 941 articles. Figure 2 depicts the year-wise distribution of articles. The x-axis

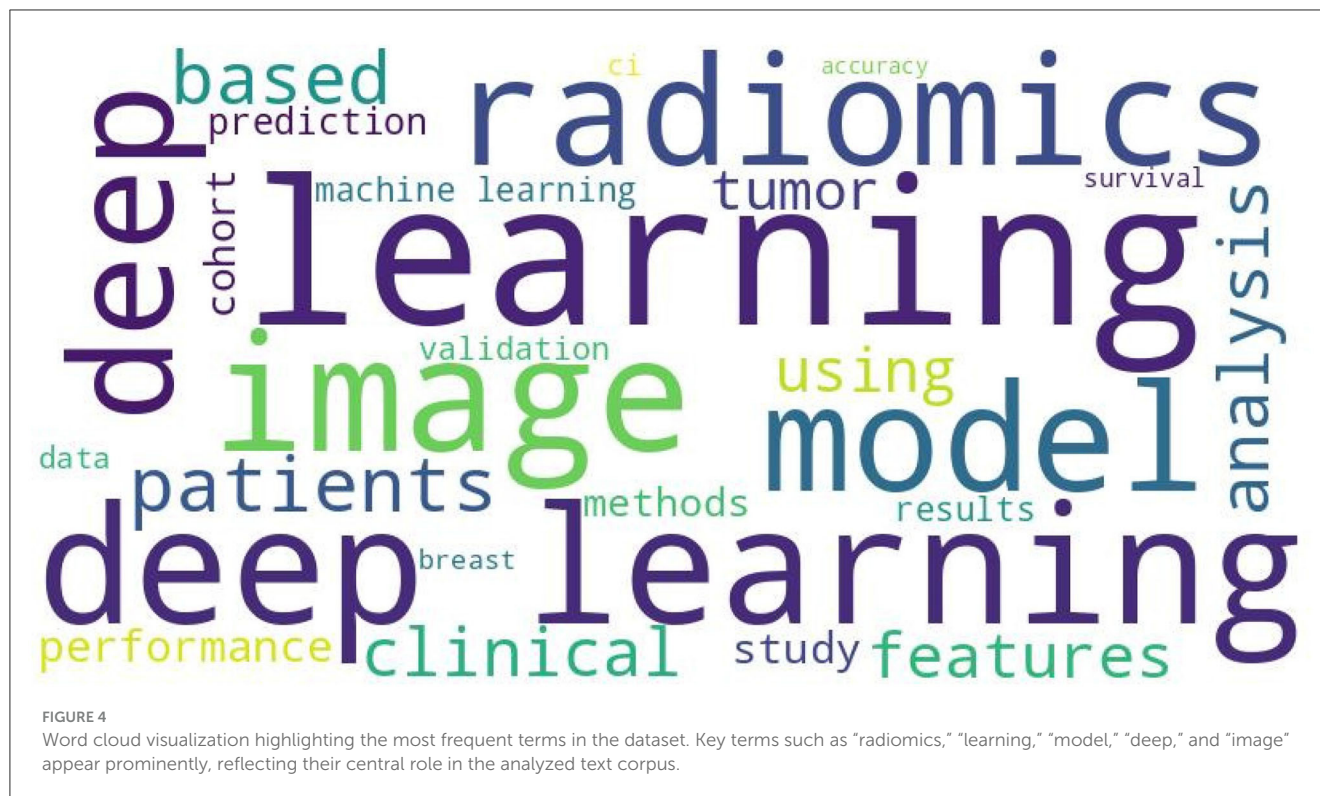
represents the time from 2021 to 2024, with each year being shown. The vertical axis with blue bars represents the number of papers published each year. There has been a consistent growth in the quantity of published papers over the last four years. Subsequently, we use inclusion and exclusion criteria to assess all titles, abstracts, and keywords using our unique keywords. As a result, we identified 63 papers for additional investigation. We use a custom Python script to implement screening criteria and conduct the actions. Out of the total number of articles screened, 23 were removed because they did not have sufficient research relevance to our research and did not provide clear information on the sensitivity, specificity, accuracy, precision, and AUC of outcomes. Ultimately, 40 studies satisfied all the requirements and were subsequently added for additional examination. Figure 1 depicts the PRISMA flowchart used for the selection procedure.



3.1 Knowledge discover from literature abstract

Topic modeling is a text mining technique that has demonstrated its effectiveness as a method for doing systematic

literature reviews (Asmussen and Muller, 2019). The most common topic modeling technique for knowledge discovery, Latent Dirichlet Allocation (LDA), finds meaningful topics in multiple literature by calculating the probability of words from each topic. LDA reveals latent topics for papers by extracting



a set of words with high probabilities (Jelodar et al., 2019). In our systematic review and meta-analysis, we extract the abstracts from selected 40 literature and identify the latent knowledge discoveries from each paper by analyzing the frequency of words and their probabilities using LDA. The results from Figure 3 shows that certain words like radiomics, image, breast, deep, learning appear more frequently, and have higher probabilities. In our topic modeling we followed a specific methodology. After extracting the abstracts we first preprocess the text by removing stop words and punctuation after that we convert the abstract to lower case. We create a dictionary and document-term matrix to identify the most frequent words and topics that appear in the abstracts. Then, we use LDA for topic modeling. The Figure 4 shows most frequent words used in abstract. This word cloud supports the relevance of our paper’s focus on deep learning, machine learning, and radiomics analysis by highlighting these topics as prominent themes in recent research. The prominence of terms like “deep learning,” “machine learning,” and “radiomics” reflects a significant trend in applying advanced learning techniques to medical imaging, which aligns closely with our study’s approach. Additionally, terms such as “patients,” “clinical,” and “cohort” underscore a focus on real-world clinical applications and patient outcomes, validating our research’s emphasis on improving healthcare through radiomics-based predictive modeling. The presence of words like “model,” “validation,” “performance,” and “prediction” shows that evaluating model accuracy is crucial in this field, which supports the significance of our work in assessing and validating predictive models. Furthermore, the word “image” emphasizes the role of medical imaging, demonstrating the importance of our findings in enhancing diagnostic capabilities through radiomics analysis. Overall, the word cloud illustrates that our study not only

aligns with high-interest areas but also contributes meaningfully to the ongoing discourse in medical machine learning, reinforcing our research’s novelty and practical impact.

3.2 Study classification

A collection of 40 papers, with publication dates ranging from 2021 to 2024, was used in this study. Overall, 12,685 patients were in the study, which trained, tested, and validated deep learning models for processing outcomes. In this review, the participants were between 48 and 70. Figure 5a shows the distribution of patients among different publications. The articles that were selected included an average of 507 patients. The imaging modalities highlighted in the literature focus on DCI-MRI, a commonly utilized technique. The Figure 5b showcases the imaging techniques frequently used in papers where ultrasound ranked as the second most popular modality. Twelve studies utilized DCI-MRI image modalities. Out of these, five studies utilize various techniques to forecast breast cancer. Various methods involve the prediction of cancer states, classification of sentinel lymph nodes (SLN) and metastasis (SLNM), estimation of HER2 expression in breast cancer, and forecasting preoperative axillary lymph node (ALN) status. Other imaging techniques are commonly employed to predict the condition of axillary lymph nodes (ALN). In addition, researchers use these techniques to predict the probability of achieving a pathological complete response (pCR) to neoadjuvant chemotherapy (NAC), evaluate lymphovascular invasion (LVI) in patients, and make predictions based on the proliferation marker Ki-67. Figure 5c in this study illustrates the frequency distribution of cancer biomarker values and treatment

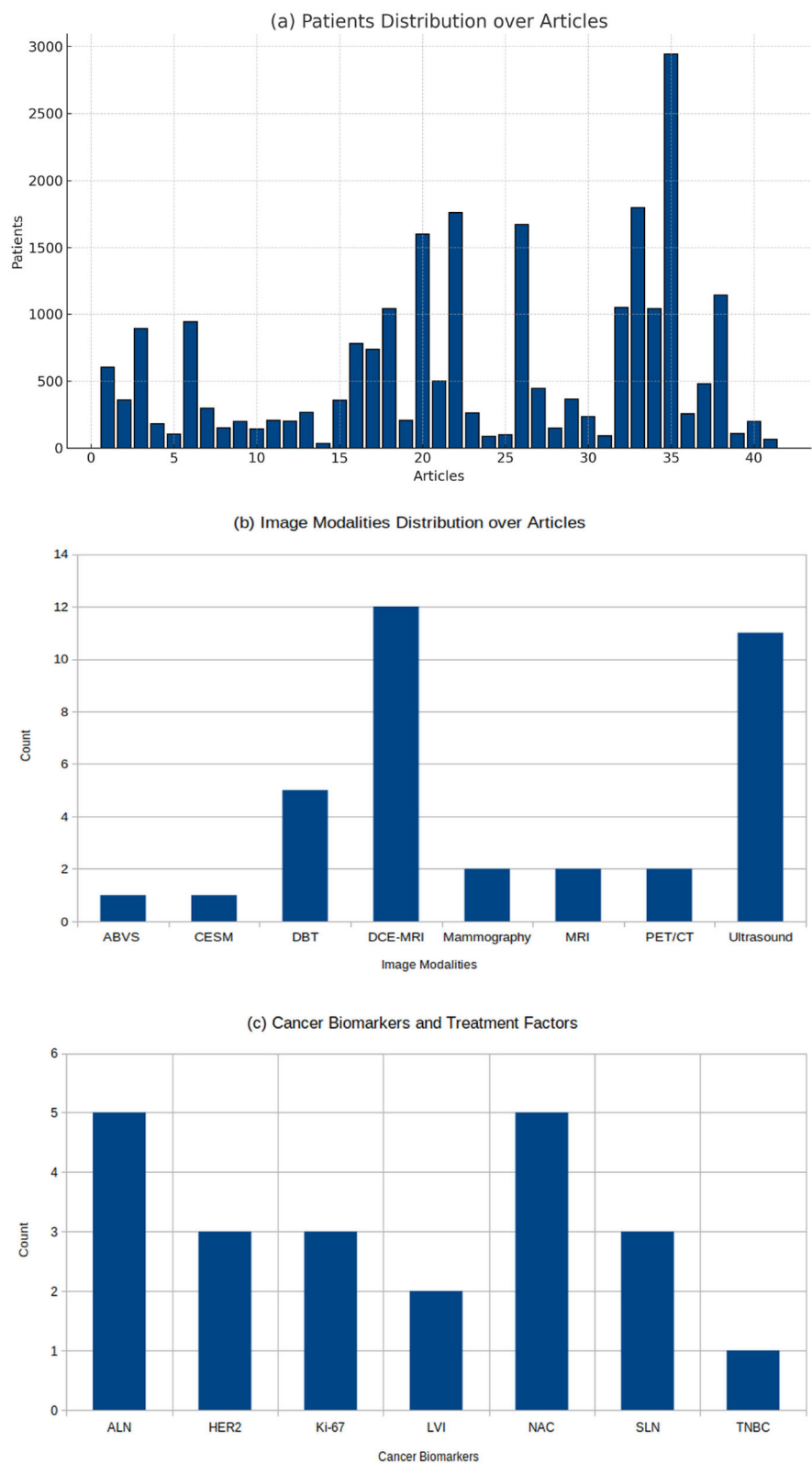


FIGURE 5 Illustrating key distributions across analyzed articles: **(a)** Patient distribution over 40 articles, with the maximum patient count reaching approximately 3000 in individual studies. **(b)** Image modality distribution, where DCE-MRI and Ultrasound are the most frequently utilized techniques. **(c)** Distribution of cancer biomarkers and treatment factors, highlighting NAC and ALN as the most reported, followed by HER2 and SLN, reflecting the emphasis on these factors in the studies.

factors in the selected literature. The 13 reviewed studies centered their analyses on ALN, SLN, and NAC biomarkers. Nine of the articles used specific cancer biomarkers, including HER2, Ki-67, triple-negative breast cancer (TNBC), and LVI, employing diverse methodologies. Table 1 explains the processing method for particular biomarkers and the process for choosing radiomics features using feature selection. In radiomics analysis, several easily detectable changes are qualitative characteristics extracted from CT, MRI, or PET scans. Different data types are sent through various stages of analysis to seek patterns and correlation (Kumar et al., 2012). Likewise, such patterns provide information that is vital for the diagnosis and prediction of prognosis. They act closely with the clinical outcomes, aligning with them (Traverso et al., 2018). The major part of the chosen articles employs the usual features of radiomics, which are first-order Features, Shape features (3D and 2D), Gray Level Co-occurrence Matrix too (GLCM), Gray Level Run Length Matrix (GLRLM), and Gray Level Size Zone Matrix (GLSZM), Neighboring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM). After reviewing 40 studies (Table 1), DCI-MRI and Ultrasound emerged as the most widely used imaging modalities for breast cancer detection. Key cancer biomarkers included ALN, NAC, and SLN, central to various diagnostic and prognostic analyses. Advanced radiomics features were frequently selected to enhance predictive accuracy, demonstrating the essential role of these tools in personalized cancer assessment.

3.3 Feature extraction and selection model

The Table 2 presented all extraction-feature and selection methods used in the selected studies. This knowledge will be a facilitator for determining the ways in which case data analysis can be conducted and the results can be deduced. This table provides an in-depth review through which we can learn trend details of feature selection methods and also get a grasp of the significance of feature features in different tasks. The existing radiomics method is PyRadiomics, a highly reliable approach used for feature extraction with the CNNs, ResNet, and DenseNet models following. There were twenty four studies that used pyradiomics in order to extract radiomics features for different imaging modalities. Furthermore, out of the five papers that utilized deep learning techniques, three of them used CNN, ResNet, or VGG-16 to extract their features. The number of features extracted varies significantly across studies, from as few as 25 to over 11,000, highlighting differences in feature granularity and study focus. The table indicates that many studies in radiomics and breast cancer research utilize privately collected datasets from different hospitals, reflecting a common approach where institutions collect and analyze their imaging data. This reliance on private datasets allows for tailored data that fits specific study objectives. It also introduces variability, making it challenging to replicate results or compare findings across studies due to differences in data acquisition protocols, equipment, and patient demographics. On the other hand, the Duke-Breast-Cancer-MRI (Saha et al., 2018) dataset is one of the most frequently used publicly available datasets in this domain. The technique of feature selection generally involves the use

of an array of methods, with machine learning approach like LASSO being the most popular option that is frequently used. Statistical methods like the ANOVA, t-test, the Spearman rank correlation coefficient, and correlation analysis are the methods often selected for feature selection. Other advanced selection models include neural networks, ensemble methods like Random Forest and LightGBM, as well as ranking algorithms such as Mutual Information, Gain Ratio, and Information Gain. These techniques help manage complex data relationships effectively. This was observed as 15 studies applied the LASSO, 5 studies adopted the ANOVA, and other studies used the Spearman rank correlation coefficient, correlation analysis, T-test, PCA, U-Test. This overview highlights the wide array of tools and methods utilized in radiomics research, reflecting a trend toward integrating statistical rigor with machine learning capabilities for effective feature selection.

3.4 Characteristics of DL/ML models

The Table 3 presents a comprehensive overview of the performance of various deep learning (DL) and machine learning (ML) models utilized in studies, highlighting their sensitivity, specificity, accuracy, and area under the curve (AUC) metrics. The performance metrics presented here provide a rounded evaluation of each model's effectiveness. Sensitivity, or recall, reflects a model's ability to correctly identify positive cases, making it essential in scenarios where missing a positive result is costly, such as medical diagnoses. Specificity measures how accurately a model identifies negative cases. Reducing false positives and avoiding misclassifying negative cases as positive is crucial. Conversely, accuracy assesses the model's overall efficacy by determining the correct positive and negative prediction ratio to the total predictions made. Lastly, the Area Under the Curve (AUC) provides a holistic measure of the model's discriminatory power across various threshold settings, with higher values indicating improved performance in differentiating between positive and negative outcomes. These metrics comprehensively understand each model's performance, highlighting their strengths and limitations in various contexts. The Table 3 indicates that ResNet (10 of the studies) is a preferred model in breast cancer analysis, with numerous studies employing its different variations due to its effectiveness in extracting detailed imaging features. For instance, Beuque et al. (2023) use ResNet101 along with Mask R-CNN, achieving a strong balance of high sensitivity and AUC, highlighting ResNet's robustness in capturing intricate details in imaging data. The study employs a data split rather than cross-validation, dividing the dataset into training, test, and external validation sets. Specifically, the study allocated 850 patients to the training set, 212 patients to the test set, and 279 patients to an external validation set from a separate institution. This setup does not involve cross-validation folds; instead, using an independent external dataset is an additional test for the model's generalizability. The external validation provides a strong benchmark for testing the model on unseen data from a different population. The wide adoption of ResNet's variations, including ResNet18, ResNet34, ResNet50, and ResNet101, across studies reassures researchers about its adaptability to specific data sizes and computational resources. It also found that authors leverage

TABLE 1 Summary of imaging modalities, cancer biomarkers, and radiomics feature extraction techniques reported across reviewed studies, highlighting the diversity in methodologies and patient cohorts.

References	Year	Image modality	Patients	Cancer bio markers	Radiomics features
Yang et al., 2023	2023	DCE-MRI	206	Lymphovascular Invasion(LVI)	7 Radiomics features
Abbasian Ardakani et al., 2023	2023	Ultrasound	200	None	3 Radiomics features
Wang C. et al., 2023	2023	Ultrasound	359	Sentinel Lymph Node(SLN)	DL-Grayscale DL-CDFI DL-elastography
Yu et al., 2023	2023	Ultrasound	603	Neo Adjuvant Chemotherapy (NAC)	Deep learning radiomics
Wei et al., 2023	2023	Ultrasound	892	Axillary Lymph Node(ALN)	5 radiomics features
Zheng et al., 2023	2023	PET/CT	105	Neo adjuvant Chemotherapy (NAC)	7 Radiomics features
Gao et al., 2023	2022	DCE-MRI	941	Axillary lymph Node(ALN)	7 Radiomics features
Xu et al., 2022	2022	Ultrasound	144	Human epidermal Growth factor receptor 2 (HER2)	None
Peng et al., 2022	2022	DCE-MRI	356	Neo adjuvant Chemotherapy (NAC)	7 Radiomics features
Vigil et al., 2022	2022	Ultrasound	780	None	7 Radiomics features
Rashid et al., 2022	2022	MRI	35	None	7 Radiomics features
Chen et al., 2022	2022	Digital breast Tomosynthesis (DBT)	298	None	7 Radiomics features
Jiang et al., 2022	2022	Digital breast Tomosynthesis (DBT)	266	K-67	7 Radiomics features
Wang et al., 2022	2022	DCE-MRI	151	Axillary Lymph Node(ALN)	7 Radiomics features
Cattell et al., 2022	2022	DCE-MRI	198	Sentinel Lymph Node(SLN)	5 Radiomics features
Wu et al., 2022	2022	Contrast-enhanced Cpectral Mammography(CESM)	182	Sentinel lymph Node(SLN)	7 Radiomics features
Bong et al., 2023	2023	Ultrasound	1,024	None	None
Sharmin et al., 2023	2023	Histology images	None	None	Deep feature extraction
Beuque et al., 2023	2023	Mammography	850	None	Handcrafted radiomics
Liu et al., 2024	2024	MRI	1,760	None	7 Radiomics features
Wang J. et al., 2023	2023	Ultrasound	263	Ki-67	7 Radiomics features
Ferre et al., 2023	2023	Ultrasound	88	HER2	7 Radiomics features
Haraguchi et al., 2023	2023	Histology images	100	ALN	7 Radiomics features
Jailin et al., 2023	2023	Mammography	1673	None	None
Quan et al., 2023	2023	Ultrasound	445	None	7 Radiomics features
Murtas et al., 2023	2023	DBT	150	None	7 Radiomics features
Nicosia et al., 2023	2023	Ultrasound	365	None	7 Radiomics features
Chen W. et al., 2023	2023	PET/CT	236	None	Deep radiomics
Sun et al., 2023	2023	DCE-MRI	91	TNBC	7 Radiomics features
Wei Li et al., 2023	2023	DCE-MRI	1,048	NAC	7 Radiomics features
Xiang et al., 2023	2023	Histology images	8,797	None	None
Bangalore et al., 2024	2024	Histology images	None	None	Deep feature extraction
Romeo et al., 2023	2023	DCE-MRI	248	HER-2	7 Radiomics features
Li Y. et al., 2023	2023	DCE-MRI	95	None	7 Radiomics features
Caballo et al., 2023	2023	DCE-MRI	922	NAC	7 Radiomics features
Chen Y. et al., 2023	2023	DCE-MRI	479	ALN	Deep feature extraction
Oladimeji et al., 2024	2024	DBT	1,140	None	7 Radiomics features
Gamal et al., 2024	2024	DCE-MRI	109	NAC	7 Radiomics features
Jiang et al., 2024	2024	DCE-MRI	198	LVI	7 Radiomics features
Del Corso et al., 2024	2024	ABVS and DBT	66	None	7 Radiomics features

TABLE 2 Overview of feature extraction methods, extracted feature counts, and feature selection techniques used across reviewed studies, emphasizing the diverse approaches to data processing and analysis.

References	Feature extraction	Extracted features count	Feature selection
Yang et al., 2023	PyRadiomics	2,478	T-tests, Correlation analysis, LASSO
Abbasian Ardakani et al., 2023	PyRadiomics CNN	27 features for each Image	Kolmogorov-Smirnov test, <i>t</i> -test ROC analysis and curve
Wang C. et al., 2023	Annotation Tool		ResNet-18
Yu et al., 2023	Annotation Tool		DCNN
Wei et al., 2023	PyRadiomics	1,403	Spearman's rank correlation coefficient, LASSO
Zheng et al., 2023	PyRadiomics ResNet101	7,740	Boruta-shap Method, Univariate analysis Multivariate Cox analysis
Gao et al., 2023	PyRadiomics	1,409	ANOVA, LASSO
Xu et al., 2022	DenseNets		
Peng et al., 2022	PyRadiomics	851	LASSO
Vigil et al., 2022	Radiomics Deep Radiomics	354	Deep Neural Network Spectral embedding method Laplacian Eigenmaps
Rashid et al., 2022	PyRadiomics	75	
Chen et al., 2022	PyRadiomics	107 (Saha et al., 2018)	Multiple SVM classifier
Jiang et al., 2022	Radiomics Deep Radiomics	2,479	Intraclass correlation coefficient (ICC) Mann-Whitney U-test LASSO
Wang et al., 2022	PyRadiomics VGG-16	120	T-test LASSO
Cattell et al., 2022	PyRadiomics VGG-16	1,105	z-score Normalization Spearman, U-test Principal Component Analysis(PCA), LASSO
Wu et al., 2022	PyRadiomics ResNet-18	3,738	Correlation Analysis LASSO ANOVA
Bong et al., 2023	ResNet101	Deep features	Clinical factors and R-CNN
Sharmin et al., 2023	ResNet50	Deep Features (Janowczyk and Madabhushi, 2016)	Random forest and LightGBM
Beuque et al., 2023	Handcrafted radiomics ResNet101		
Liu et al., 2024	PyRadiomics	11,342	LASSO ANOVA
Wang J. et al., 2023	PyRadiomics	849	LASSO
Ferre et al., 2023	PyRadiomics		
Haraguchi et al., 2023	PyRadiomics		LASSO
Jailin et al., 2023	YOLO		
Quan et al., 2023	PyRadiomics	91	
Murtas et al., 2023	LIFEx	58	K-best, ANOVA, Sequential Embedded method, Random Forest
Nicosia et al., 2023	LIFEx		LASSO
Chen W. et al., 2023		Deep Features	Attention-based aggregate CNN
Sun et al., 2023	PyRadiomics	2,832	ANOVA Recursive feature elimination
Wei Li et al., 2023	PyRadiomics		
Xiang et al., 2023		Deep Features	DenseNet-121
Bangalore et al., 2024	Pruned CNN	550 (Janowczyk and Madabhushi, 2016 ; Balkenende et al., 2022)	Information Gain (IG) Supervised Relative Reduct (SRR) Gain-Ratio (GR), OneR mRMR Chi-Squared (CS)

(Continued)

TABLE 2 (Continued)

References	Feature extraction	Extracted features count	Feature selection
Romeo et al., 2023	PyRadiomics	1,288 (Saha et al., 2018)	ICC Analysis
Li Y. et al., 2023	PyRadiomics	1,130	LASSO
Caballo et al., 2023	PyRadiomics	348 (Saha et al., 2018)	U-test
Chen Y. et al., 2023	DenseNet121	1,000 (Saha et al., 2018)	mRMR LASSO SMOT
Oladimeji et al., 2024	PyRadiomic	857 (Buda et al., 2020)	Mutual information U Test
Gamal et al., 2024			Genetic algorithm (GA) Standardization
Jiang et al., 2024	PyRadiomics	6,195	mRMR LASSO
Del Corso et al., 2024	PyRadiomics	25	PCA Adaptive feature selection

two deep learning models like (Yu et al., 2023), leveraging VGG16 and ResNet50 models where authors use a split-sample validation approach rather than cross-validation. Specifically, author divides the data from three medical centers, using patients from two centers as a training cohort (420 patients) and patients from the third center as an external validation cohort (183 patients), attaining robust accuracy metrics highlighting deep learning capacity for breast cancer classification. Other famous models, such as Inception V3, DenseNet, and SqueezeNet, are frequently used by researchers. In contrast, traditional ML models are often used as benchmarks or in cases of limited data availability; for instance, Liu et al. (2024) utilize Support Vector Machine (SVM) and Gaussian Process models where authors employ a split-sample approach across five distinct cohorts to evaluate the model's robustness. They divide the data into training (775 patients) and validation (518 patients) cohorts for model development. For further testing, they use three independent testing cohorts: an internal retrospective cohort (167 patients), an internal prospective cohort (188 patients), and an external retrospective cohort (112 patients), reporting moderate sensitivity and specificity values. Ensemble methods, particularly XGBoost, are frequently implemented, with Quan et al. (2023) demonstrating XGBoost's impressive high specificity and accuracy. The authors apply a 4:1 split ratio, dividing patients into a training set (357 patients) and an independent test set (88 patients), highlighting the advantages of ensemble learning in achieving better generalizability. Studies using hybrid approaches, such as Nicosia et al. (2023), incorporate attention-based CNNs where authors apply 10-fold cross-validation to select radiomics features using the LASSO (Least Absolute Shrinkage and Selection Operator) logistic regression model. Additionally, they split the data into a training set (70%, with 255 lesions) and a test set (30%, with 110 lesions) to further validate model accuracy and robustness, achieving high AUC values. Moreover, studies such as those by Ferre et al. (2023) and Gamal et al. (2024) show the continued relevance of logistic regression and random forests, particularly when paired with feature selection techniques, to achieve competitive performance. In the study by Rashid et al.

(2022), CNN-SVM achieved the highest AUC of 0.974. Notably, this model also achieves higher accuracy (98.83%) compared to other models. Yang et al. (2023) achieved the highest sensitivity of 0.889 using 3DResNet. This model also demonstrates a remarkable level of lower specificity, with a value of 0.692. The study conducted by Wu et al. (2022) attained the nearest sensitivity of 0.88 utilizing a radiomics model. Accuracy is crucial in cases where detecting all positive instances, such as cancerous tumors, is vital. In a study by Rashid et al. (2022) and Bangalore et al. (2024), the CNN-SVM model and EfficientNet-Transformer models respectively achieved the highest accuracy of 0.9883 and 0.9884. Specificity is vital in situations where there is a chance that a negative prediction is truly negative. Cattell et al. (2022) demonstrated that VGG-16 achieved the highest level of specificity, with a score of 0.87. However, the AUC value for this study is 0.83, which is notably higher than other models.

3.5 Literature methodology analysis

The primary objective of the study will be the detailed analysis of the methods sections of 16 selected papers. We evaluated and studied different research designs for our examine the methodology used in recent literature. This approach will ensure the exploration of each methodology, developing a more detailed understanding of the research process. Choosing an appropriate image dataset is usually the first step in research methodology where researcher typically use one specific type of image modality. After doing several image processing steps the features are extracted using radiomics models or deep learning models. Then, the next stage will be selecting the best features using several appropriate statistical models. Finally, different DL/ML models are proposed or evaluated, either individually or in combination. The authors of those studies (Yang et al., 2023; Wu et al., 2022; Abbasian Ardakani et al., 2023; Gao et al., 2023) implemented new deep learning or radiomics models in their research. Yang

TABLE 3 Performance analysis of DL/ML models across studies, including metrics such as sensitivity, specificity, accuracy, and area under the curve (AUC), highlighting variations in model effectiveness and application.

References	ML/DL	Sensitivity	Specificity	Accuracy	Area under curve (AUC)
Yang et al., 2023	3DResNet	0.889	0.692	0.743	0.857
Abbasian Ardakani et al., 2023	SqueezeNet, InceptionV3 VGG16, VGG19				0.861, 0.843 0.847, 0.867
Wang C. et al., 2023	ResNet18	DL-grayscale - 0.82, DL-CDFI - 0.77 DL-elastography - 0.88	DL-grayscale - 0.84, DL-CDFI - 0.71 DL-elastography - 0.82	DL-grayscale - 0.83, DL-CDFI - 0.76 DL-elastography - 0.86	DL-grayscale - 0.85, DL-CDFI - 0.76 DL-elastography - 0.87
Yu et al., 2023	ResNet34, ResNet50 VGG16, DenseNet121	ResNet34 - 0.788, ResNet50 - 0.773 VGG16 - 0.606, DenseNet121 - 0.697	ResNet34 - 0.684, ResNet50 - 0.855 VGG16 - 0.786, DenseNet121 - 0.684	ResNet34 - 0.721, ResNet50 - 0.825 VGG16 - 0.716, DenseNet121 - 0.689	ResNet34 - 0.785, ResNet50 - 0.879 VGG16 - 0.707, DenseNet121 - 0.724
Wei et al., 2023	ResNet50, ResNet101 Inception_v3, VGG19	ResNet50 - 0.71, ResNet101 - 0.57 Inception_v3 - 0.71, VGG19 - 0.71	ResNet50 - 0.67, ResNet101 - 0.80 Inception_v3 - 0.74, VGG19 - 0.49	ResNet50 - 0.72, ResNet101 - 0.76 Inception_v3 - 0.72, VGG19 - 0.83	ResNet50 - 0.69, ResNet101 - 0.69 Inception_v3 - 0.71, VGG19 - 0.58
Zheng et al., 2023	ResNet101				0.8595
Gao et al., 2023	ResNet	0.792	0.853	0.828	0.852
Xu et al., 2022	DenseNets	0.727	0.84	0.805	0.84
Peng et al., 2022	ResNext50	0.566	0.556	0.558	0.554
Vigil et al., 2022	Unet			0.853	
Rashid et al., 2022	Inception_V3, CNN-SVM			Inception_V3 - 0.9196, CNN-SVM - 0.9883	Inception_V3 - 0.932, CNN-SVM - 0.974
Chen et al., 2022	ResNet50	0.68	0.78	0.74	0.82
Jiang et al., 2022	ResNet50	0.864	0.500	0.500	0.714
Wang et al., 2022	VGG-16			0.78	0.87
Cattell et al., 2022	VGG-16	0.58	0.87	0.77	0.83
Wu et al., 2022	Radiomics Model	0.88	0.68	0.75	0.85
Bong et al., 2023	Mask R-CNN ResNet101	0.78	0.85		
Sharmin et al., 2023	ResNet50V2			0.95	
Beuque et al., 2023	Mask R-CNN ResNet101	0.90	0.86		0.95
Liu et al., 2024	Gaussian Process Support vector machine(SVM)			0.84	
Wang J. et al., 2023	Logistic regression Random forests Support vector machine Extreme gradient boosting (XGBoost)				0.88
Ferre et al., 2023	Logistic regression k-nearest neighbor Näve Bayes	0.818	0.742		0.824
Haraguchi et al., 2023	Support vector machine				0.765
Jailin et al., 2023	YOLO				0.964
Quan et al., 2023	Logistic regression Random forests Support vector machine Extreme gradient boosting (XGBoost)	0.991	0.996		
Murtas et al., 2023	Random forest			0.72	

(Continued)

TABLE 3 (Continued)

References	ML/DL	Sensitivity	Specificity	Accuracy	Area under curve (AUC)
Nicosia et al., 2023	Multivariable logistic				0.914
Chen W. et al., 2023	Attention-based aggregate CNN	0.809	0.85	0.83	0.886
Sun et al., 2023	Support vector machine			0.82	0.85
Wei Li et al., 2023	Logistic regression (LR) Linear discriminant analysis (LDA) Support vector machine (SVM)			0.843	0.893
Xiang et al., 2023	Attention-based CNN	0.864	0.866	0.865	0.95
Bangalore et al., 2024	Triplet attention-based Efficient network (TAENet)	0.9718	0.9812	0.9884	
Romeo et al., 2023	Logistic regression	0.80	0.43	0.63	0.66
Li Y. et al., 2023	Deep learning radiomic (DLR) Logistic regression	0.875	0.90	0.893	0.914
Caballo et al., 2023	Logistic regression model with leave-one-out cross-validation (LOOCV)				0.707
Chen Y. et al., 2023	DenseNet121	0.65	0.80	0.75	0.80
Oladimeji et al., 2024	Random forest Extreme gradient boosting (XGBoost) Logistic regression Decision Tree			0.92	
Gamal et al., 2024	LightGBM Extreme gradient boosting (XGBoost)	0.88	0.98	0.88	
Jiang et al., 2024	Logistic regression (LR) Support vector machine (SVM) Classification and Regression Trees (CART) K-Nearest Neighbors (KNN) Gradient Boosting Machine (GBM)	0.90	0.72	0.87	
Del Corso et al., 2024	Logistic regression (LR) Support vector machine (SVM) Logistic regression	0.842	0.807	0.807	0.921

[et al. \(2023\)](#) introducing three innovative deep-learning models that leverage the power of radiomics: DL-gray scale, DL-CDFI, and DL-elastography. [Wu et al. \(2022\)](#) developed two radiomics nomograms capable of reliably showing the presence or absence of NSLN metastasis and the extent of axillary tumoral burden. A new filter based on deep learning and adaptive residual learning has been proposed by [Abbasian Ardakani et al. \(2023\)](#). [Gao et al. \(2023\)](#) explore the capabilities of their attention-based DL model in distinguishing ALN metastasis in breast cancer using dynamic contrast-enhanced MRI (DCE-MRI) before surgery. The remaining authors of the papers utilize various cancer bio-markers to make predictions about breast cancer. Authors employ feature

extraction and selection model that is based on radiomics and deep learning. The authors also conducted various statistical analyses to select features. To develop their DL/ML model, authors created separate models for radiomics and deep learning. Furthermore, several of them utilized radiomics and deep learning features to train DL/ML models. In the end, the authors provided an affirmation of accuracy by performing both internal and external validation techniques. Research from [Jiang et al. \(2022\)](#) exploited radiomics and deep learning features from different tumor areas. These features were then sequentially selected by LASSO regression and finally produced radiomics signatures. AUROC, accuracy, sensitivity, and specificity were calculated to estimate

the radiomics signatures. After doing extensive research on 40 articles we developed a widely used methodology which is shown in Figure 6.

3.6 Risk of bias

The quality assessment of 16 publications was assessed using QUADAS-2. We prioritized studies that provided high-quality and comprehensive data on patient selection and target classification for breast cancer to ensure the robustness and applicability of our findings. The selected studies represented the most relevant and rigorous evidence available for quality assessment. For the risk of bias, there is a higher number of unclear answers in the reference standard selection due to the specific modality used. The flow and timing of high-risk responses may suggest different reference standards being used and ambiguous intervals with index tests. There are some concerns regarding the applicability of the reference standard due to the use of a particular modality, which has resulted in a higher number of unclear answers. The increased risk of bias in index tests primarily stems from the design of the validation process. There is some ambiguity when it comes to patient selection due to the lack of clear inclusion and exclusion criteria. Figure 7 illustrates the process of quality assessment, utilizing the robvis tool (McGuinness and Higgins, 2021).

3.7 Correlation coefficient

Correlation coefficients are thus pivotal in the distinction between correlation types, revealing whether two variables show synchronized movements, diverge in opposite directions, or are completely unrelated. By adopting the random effect model, we endeavor to find out the values of the correlation coefficient (n , r) between the sample size and sensitivity in our meta-analysis study.

This meta-analysis investigates the correlation coefficients across 23 studies using a random-effects model. Fisher's r -to- z transformed correlation coefficient is used as the primary outcome measure, allowing for a standardized effect size across studies with varying sample sizes. The model uses the restricted maximum likelihood (REML) estimator for study-level variability. The Table 4 and Figure 8 describe our meta analysis. The random-effects model estimated an average effect size of 1.20 with a standard error of 0.0869. A Z-test yielded a value of 13.8, with a p -value of <0.001 , indicating statistical significance. The 95% confidence interval (CI) for the effect size ranges from 1.033 to 1.373, suggesting a robust, statistically significant association. Tau τ , representing the standard deviation of true effect sizes, is estimated at 0.412, and Tau² τ^2 , the variance of effect sizes, is 0.1694 (with a standard error of 0.0524). I^2 is very high at 98.63%, indicating substantial heterogeneity across studies. This indicates that almost all variability in observed effect sizes is due to differences between studies rather than random chance. The Q-test for heterogeneity also supports this high level of heterogeneity, with a Q value of 1,574.190 ($df = 22$) and a p -value < 0.001 . These metrics imply that the included studies are highly heterogeneous, and the overall effect size should be interpreted with caution, as it may not represent a single underlying population

effect. The funnel plot shows asymmetry, suggesting the potential for publication bias or small-study effects. Studies appear unevenly distributed around the mean effect size, especially toward the left side, which might indicate an underrepresentation of studies with small or null effects. The forest plot displays individual effect sizes for each study along with their 95% confidence intervals. Effect sizes range considerably across studies, from approximately 0.5 to 2.7, reinforcing the high heterogeneity ($I^2 = 98.63\%$). The pooled effect size obtained from the random-effects model is 1.20, with a 95% confidence interval of 1.03 to 1.37, displayed at the bottom of the forest plot. Log-likelihood, Deviance, AIC, BIC, and AICc values are reported for both the maximum-likelihood and restricted maximum-likelihood models. The REML model has a slightly better fit, with lower values across these criteria (AIC = 27.743, BIC = 29.925, AICc = 28.374), suggesting it is more appropriate for this data due to better handling of the high heterogeneity. The meta-analysis reveals a significant average effect size, suggesting a meaningful correlation across studies. However, the high heterogeneity, as evidenced by I^2 and the Q-test, indicates that effect sizes vary widely among the studies, possibly due to differences in study populations, methodologies, or contextual factors. The presence of funnel plot asymmetry further suggests the possibility of publication bias, which should be considered when interpreting these findings.

4 Discussion

In the discussion section of this systematic review and meta-analysis, we explore the detailed insights derived from the comprehensive body of evidence examined. By carefully analyzing the data, we aim to present an in-depth understanding of the research field and learn the research patterns.

After surveying the studies, DL/ML technology shows remarkable accuracy in forecasting results. Through an in-depth and methodical review, we have assessed the efficacy of existing techniques, identified promising areas for future research, and acquired valuable insights into their accuracy. The radiomics-guided DL/ML model has shown promising potential for improving accuracy. In some instances, the radiomics model alone has shown the ability to achieve the highest accuracy in identifying breast cancer. In our research, 12,685 patients were enrolled, and the authors used several strategies to identify these individuals. Since our review question primarily pertains to image processing methods using radiomics-guided deep learning models, we will only concentrate on image processing using radiomics-guided deep learning models. Ultrasound is a widely used tool for imaging cancer biomarkers. Sentinel lymph nodes (SLN) and axillary lymph nodes (ALN) are mainly used as biomarkers. Most of the articles we picked used seven radiomics attributes for radiomics analysis. The radiomics characteristics are extracted using the pyradiomics python package. Deep learning models are famous for extracting deep learning information from images. Following the extraction process, writers often use LASSO statistical analysis for feature selection. In addition, the researchers used several statistical analyses, such as U-test, T-test, and correlation coefficient, in their study. The authors used many statistical approaches

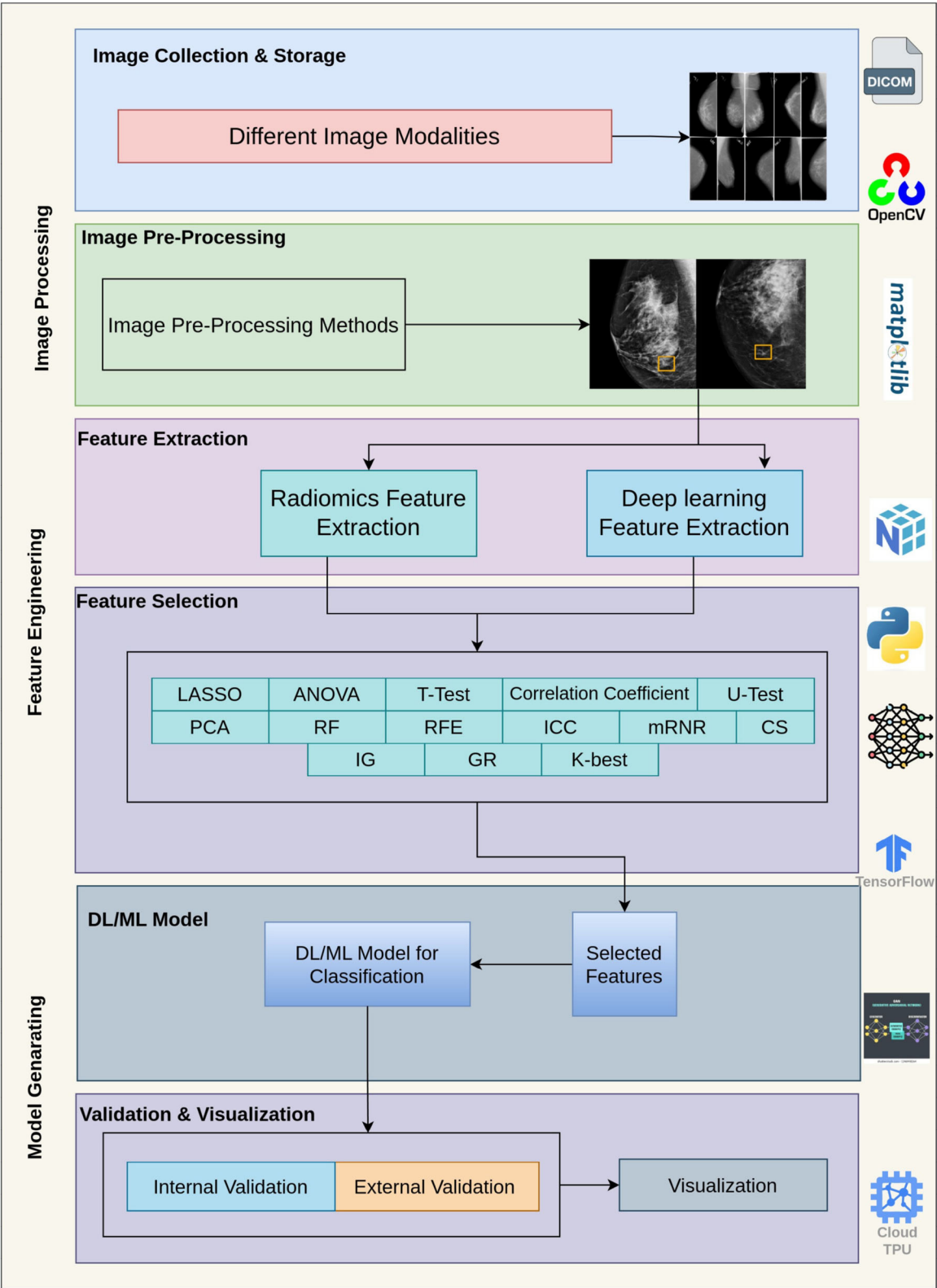


FIGURE 6
Workflow for medical image analysis: image collection, pre-processing, feature extraction, feature selection, model generation, and validation.

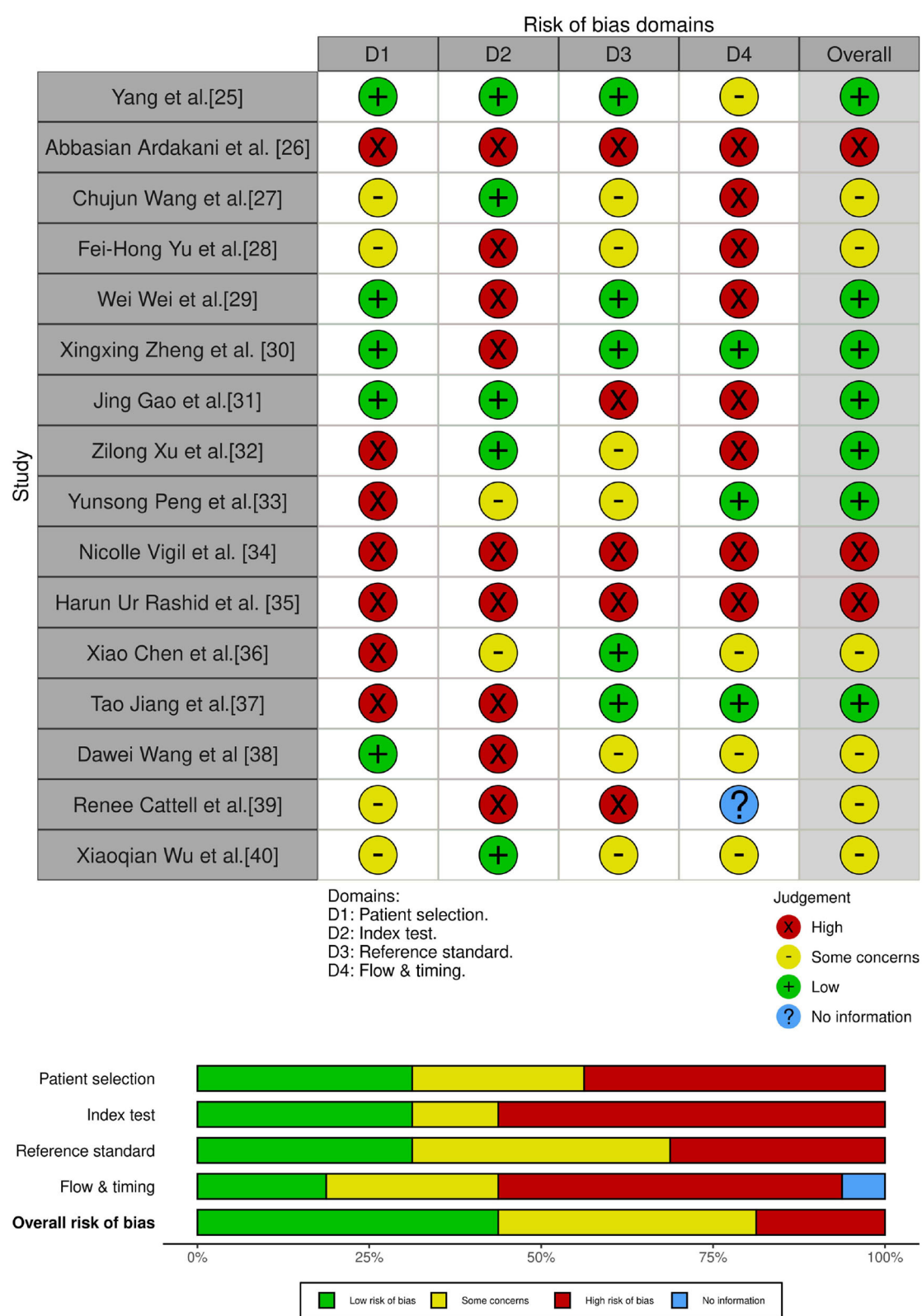


FIGURE 7
Quality assessment according To QUADAS-2.

collectively for their investigations. Our research revealed that authors independently use statistical approaches for selection in their radiomics and deep learning models. When generating DL/ML models, authors often use various variations of ResNet, including ResNet 18, ResNet 34, ResNet 50, and ResNet 101. Even though the CNN-based model achieved the highest AUC of

TABLE 4 Random-effects model summary (R = 2.5, K = 23) with heterogeneity statistics and model fit metrics for ML and REML methods.

	Estimate	se	Z	p	CI lower bound	CI upper bound
Random-effects model (k = 23)						
Intercept	1.14	0.0869	13.8	0.001	1.033	1.373
Tau	τ^2	I^2	H^2	IR^2	df	Q
Heterogeneity Statistics						
0.412	0.1694 (SE = 0.00524)	98.63%	72.766	-	22.000	1,574.190
	Log-likelihood	Deviance	AIC	BIC	AICc	
Model Fit Statistics and information Criteria						
Maximum-Likelihood		-11.915	113.448	27.831	30.102	28.431
Restricted Maximum-Likelihood		-11.871	23.743	27.743	29.925	28.374

97.4%, with the same model, we achieved an impressive accuracy of 98.83%.

A previous systematic review conducted by Taddese and Tilahun (2024) consisted of 48 studies, while their meta-analysis comprised 24. The studies utilized various images and models to diagnose various gynecological cancers. The authors emphasized that DL algorithms demonstrated higher sensitivity but lower specificity than machine learning (ML) methods.

This systematic review thoroughly examined the topic using a well-defined methodology and strict inclusion criteria. We also customized the quality assessment tools to suit the included studies. We explore image feature extraction and selection methods, utilize radiomics and deep learning models, and compare their performance. It is worth noting that previous studies and current guidelines have highlighted the importance of internal validation. This involves training and validating models using the same dataset, often through techniques like cross-validation. However, it is essential to be cautious with the results obtained through internal validation, as they tend to overestimate accuracy and may not be easily generalized due to overfitting. Therefore, only studies that utilized external validation of test sets were considered during the initial phase of literature identification. Thus, our research offers valuable insights into the use of DL for diagnosing breast cancer. However, this systematic review did not include publications in languages other than English, which could introduce bias in the selection process. Furthermore, the lack of sufficient data hindered the calculation of comprehensive diagnostic measures.

We recommend conducting more precise research on feature extraction and selection based on radiomics and deep learning models. After extracting the features, we should find the most valuable features by conducting statistical analysis using both categories of features combined to obtain more accurate prospective studies. We recommend using externally validated data to conduct a more thorough assessment of the DL/ML model for both ruptured and unruptured aneurysms. To ensure that the results of this promising technology can be replicated and applied to a broader range of cases, we suggest developing standardized research guidelines for further investigations.

This systematic review investigated the potential of utilizing radiomics-guided deep learning/machine learning models to identify breast cancer. The studies provided encouraging findings,

as specific models showed impressive accuracy in distinguishing between malignant and benign breast tumors. Nevertheless, there is a wide range of variations in the designs of studies, architectures of models, and techniques used for validation. In literature, a variety of imaging methods are employed. Upon evaluation, we found that the most frequent imaging modality is ultrasound imaging. In addition, DBT, CT/PET, and DCI-MRI are utilized for early identification of breast cancer. Extracting features mainly combines radiomics and deep learning feature extraction methods. The Pyradiomics Python package extracts Radiomics characteristics, and the most used deep learning model for extracting medical image features is ResNet. In this review, we observed that while there is a considerable use of other statistical models such as T-test, ANOVA, and correlation analysis, researchers mainly utilize LASSO for feature selection. The most popular deep-learning models for classifying breast cancer are ResNet and VGG. Additional research is necessary to establish uniform techniques, enhance applicability, and explore the practical implications of these models. Future research in radiomics-guided deep learning (DL) and machine learning (ML) for breast cancer detection should prioritize several key areas to build upon the promising yet varied findings highlighted in this review. First, standardized model development, validation, and evaluation guidelines are crucial. The wide variability in model architectures, feature extraction techniques, and validation methods across studies has led to inconsistent performance metrics, challenging the generalizability of findings. Establishing a common framework will allow researchers to compare results more effectively and ensure that the models developed are robust, reproducible, and clinically applicable. Second, the future of our research should be built on a foundation of external validation through multi-center prospective trials. Many current studies rely on internal validation, which may introduce overfitting and overestimate model accuracy. It's urgent that we evaluate model performance in real-world clinical environments by conducting trials across diverse populations and imaging settings. This step is crucial in increasing confidence in the models' diagnostic accuracy and generalizability, and it's a key part of our journey toward more reliable and effective breast cancer detection. Additionally, the field would benefit from further investigation into optimal feature selection techniques that combine radiomics and DL features. Current methods, such

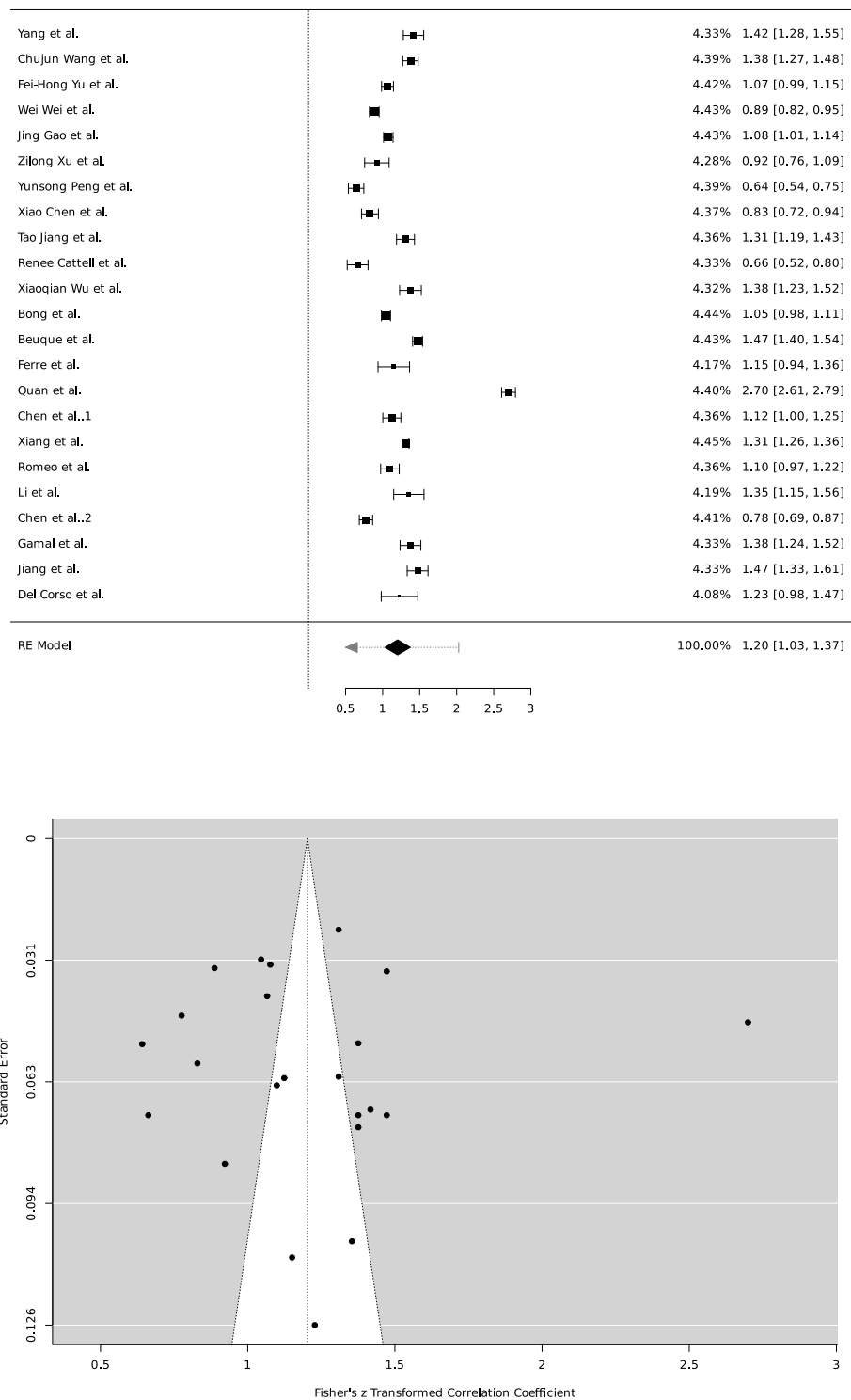


FIGURE 8 Forest plot and funnel plot presenting the meta-analysis results. The forest plot illustrates individual study effect sizes and 95% confidence intervals, with a pooled effect size of 1.20 (95% CI: 1.03, 1.37) obtained using a random-effects model. The funnel plot indicates an asymmetric distribution, suggesting potential publication bias or small-study effects.

as LASSO, U-test, and *T*-test, show promise, but additional methods that integrate both categories of features could enhance the predictive power of DL/ML models. Exploring new feature selection algorithms or hybrid approaches could yield insights into the most predictive attributes for distinguishing between benign and malignant tumors. Lastly, our research should explore the practical application of these advanced imaging techniques with existing diagnostic workflows. We should study the potential

roles of DL/ML models not as standalone diagnostic methods, but as complementary tools. This approach could facilitate their practical application in early breast cancer detection, monitoring, and treatment planning, making them an integral part of our clinical workflows.

5 Conclusion

This systematic review investigated the potential of utilizing radiomics-guided deep learning/machine learning models to identify breast cancer. The studies provided encouraging findings, as certain models showed impressive accuracy in distinguishing between malignant and benign breast tumor. Nevertheless, there is a wide range of variations in the designs of studies, architectures of models, and techniques used for validation. In literature, a variety of imaging methods are employed. Upon evaluation, we found that the most frequent imaging modality is ultrasound imaging. In addition, DBT, CT/PET, and DCI-MRI are utilized for early identification of breast cancer. Extracting features are mostly done with the combination of radiomics and deep learning feature extraction methods. The Pyradiomics Python package is used to extract Radiomics characteristics, and the most used deep learning model for extracting medical image features is ResNet. In this review, we observed that while there is a considerable use of other statistical models such as *T*-test, ANOVA, and correlation analysis, researchers mostly utilize LASSO for feature selection. The most popular deep learning models for classifying breast cancer are ResNet and VGG. The review identifies notable challenges, such as variability in model architectures, feature selection techniques, and validation approaches across studies, which have led to inconsistencies in model performance and generalizability. Although widely used, internal validation methods are limited in assessing accurate diagnostic accuracy due to the risk of overfitting. The recommendation is to prioritize external validation through multi-center, prospective trials, enabling more accurate and generalizable assessments and supporting broader clinical applicability. Furthermore, combining radiomics and DL/ML features through optimized feature selection techniques, such as LASSO, and exploring hybrid approaches could enhance model precision. Standardizing model development, validation, and evaluation protocols is essential to improve the comparability and reliability of findings across studies. Looking forward, radiomics-guided DL/ML models show great promise as complementary diagnostic tools rather than standalone methods, potentially enhancing early breast cancer detection, monitoring, and treatment planning. However, the findings underscore the need for standardized guidelines, external validation, and more rigorous prospective studies to realize the full potential of these models

in clinical settings. This review provides a foundation for future research, which should continue to refine and integrate DL/ML methodologies into diagnostic pathways, ultimately improving patient outcomes in breast cancer care.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

NM: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. AB: Supervision, Funding acquisition, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The Deanship of Scientific Research (DSR) at King Abdu-laziz University in Jeddah, Saudi Arabia, provided funding for this project under grant number (GPIP: 1821-611-2024). The authors therefore acknowledge with thanks DSR for technical and financial support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abbasian Ardakani, A., Mohammadi, A., Vogl, T. J., Kuzan, T. Y., and Acharya, U. R. (2023). AdaRes: a deep learning-based model for ultrasound image denoising: results of image quality metrics, radiomics, artificial intelligence, and clinical studies. *J. Clin. Ultrasound* 52, 131–143. doi: 10.1002/jcu.23607
- Abhisheka, B., Biswas, S. K., Purkayastha, B., Das, D., and Escargueil, A. (2023). Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. *Multim. Tools Applic.* 83, 43035–43070. doi: 10.1007/s11042-023-17326-1

- Arnold, M., Morgan, E., Rumgay, H., Mafrá, A., Singh, D., Laversanne, M., et al. (2022). Current and future burden of breast cancer: global statistics for 2020 and 2040. *Breast* 66, 15–23. doi: 10.1016/j.breast.2022.08.010
- Asmussen, C. B., and Möller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *J. Big Data* 6:93. doi: 10.1186/s40537-019-0255-7
- Balkenende, L., Teuwen, J., and Mann, R. M. (2022). Application of deep learning in breast cancer imaging. *Semin. Nucl. Med.* 52, 584–596. doi: 10.1053/j.semnucmed.2022.02.003
- Bangalore, M., Hiremath, B. S., and Polnaya, A. M. (2024). Classification of breast cancer using ensemble filter feature selection with triplet attention based efficient net classifier. *Int. Arab J. Inf. Technol* 21, 17–31. doi: 10.34028/iajit/21/1/2
- Beuque, M. P., Lobbes, M. B., van Wijk, Y., Widaatalla, Y., Primakov, S., Majer, M., et al. (2023). Combining deep learning and handcrafted radiomics for classification of suspicious lesions on contrast-enhanced mammograms. *Radiology* 307:221843. doi: 10.1148/radiol.221843
- Bong, J. H., Kim, T. H., and Jeong, S. (2023). Deep learning model for the diagnosis of breast cancers smaller than 1 cm with ultrasonography: integration of ultrasonography and clinical factors. *Quant. Imaging Med. Surg.* 13, 2486–2495. doi: 10.21037/qims-22-880
- Buda, M., Saha, A., Walsh, R., Ghate, S., Li, N., Swiecicki, A., et al. (2020). *Breast cancer screening digital breast tomosynthesis (bcs-dbt)*. Data obtained from Duke Health System.
- Caballo, M., Sanderink, W. B., Han, L., Gao, Y., Athanasiou, A., and Mann, R. M. (2023). Four-dimensional machine learning radiomics for the pretreatment assessment of breast cancer pathologic complete response to neoadjuvant chemotherapy in dynamic contrast-enhanced MRI. *J. Magn. Reson. Imag.* 57, 97–110. doi: 10.1002/jmri.28273
- Cattell, R., Ying, J., Lei, L., Ding, J., Chen, S., Serrano Sosa, M., et al. (2022). Preoperative prediction of lymph node metastasis using deep learning-based features. *Visual Comput. Ind. Biomed. Art* 5:8. doi: 10.1186/s42492-022-00104-5
- Cè, M., D'Amico, N. C., Danesini, G. M., Foschini, C., Oliva, G., Martinenghi, C., et al. (2023). Ultrasound elastography: basic principles and examples of clinical applications with artificial intelligence—a review. *BioMedInformatics* 3, 17–43. doi: 10.3390/biomedinformatics3010002
- Chen, W., Liu, F., Wang, R., Qi, M., Zhang, J., Liu, X., et al. (2023). End-to-end deep learning radiomics: development and validation of a novel attention-based aggregate convolutional neural network to distinguish breast diffuse large B-cell lymphoma from breast invasive ductal carcinoma. *Quant. Imaging Med. Surg.* 13, 6598–6614. doi: 10.21037/qims-22-1333
- Chen, X., Zhang, Y., Zhou, J., Wang, X., Liu, X., Nie, K., et al. (2022). Diagnosis of architectural distortion on digital breast tomosynthesis using radiomics and deep learning. *Front. Oncol.* 12:991892. doi: 10.3389/fonc.2022.991892
- Chen, Y., Wang, L., Dong, X., Luo, R., Ge, Y., Liu, H., et al. (2023). Deep learning radiomics of preoperative breast MRI for prediction of axillary lymph node metastasis in breast cancer. *J. Digit. Imaging* 36, 1323–1331. doi: 10.1007/s10278-023-00818-9
- Del Corso, G., Germanese, D., Caudai, C., Anastasi, G., Belli, P., Formica, A., et al. (2024). Adaptive machine learning approach for importance evaluation of multimodal breast cancer radiomic features. *J. Imag. Inform. Med.* 37, 1642–1651. doi: 10.1007/s10278-024-01064-3
- Dhar, T., Dey, N., Borra, S., and Sherratt, R. S. (2023). Challenges of deep learning in medical image analysis improving explainability and trust. *IEEE Trans. Technol. Soc.* 4, 68–75. doi: 10.1109/TTTS.2023.3234203
- Dittonno, F., Franco, A., Manfredi, C., Vecchia, A., Valerio, M., Bukavina, L., et al. (2023). Novel non-MRI imaging techniques for primary diagnosis of prostate cancer: micro-ultrasound, contrast-enhanced ultrasound, elastography, multiparametric ultrasound, and PSMA PET/CT. *Prostate Cancer Prostatic Dis.* 27, 29–36. doi: 10.1038/s41391-023-00708-9
- Ferre, R., Elst, J., Senthilnathan, S., Lagree, A., Tabbarah, S., Lu, F. I., et al. (2023). Machine learning analysis of breast ultrasound to classify triple negative and HER2+ breast cancer subtypes. *Breast Dis.* 42, 59–66. doi: 10.3233/BD-220018
- Gamal, A., Sharafeldien, A., Alnaghy, E., Alghandour, R., Saleh Alghamdi, N., Ali, K. M., et al. (2024). A novel machine learning approach for predicting neoadjuvant chemotherapy response in breast cancer: integration of multimodal radiomics with clinical and molecular subtype markers. *IEEE Access* 12, 104983–105003. doi: 10.1109/ACCESS.2024.3432459
- Gao, J., Zhong, X., Li, W., Li, Q., Shao, H., Wang, Z., et al. (2023). Attention based deep learning for the preoperative differentiation of axillary lymph node metastasis in breast cancer on DCE-MRI. *J. Magn. Reson. Imag.* 57, 1842–1853. doi: 10.1002/jmri.28464
- Gatta, G., Somma, F., Sardù, C., De Chiara, M., Massafra, R., Fanizzi, A., et al. (2023). Automated 3D ultrasound as an adjunct to screening mammography programs in dense breast: literature review and meta-analysis. *J. Pers. Med.* 13:1683. doi: 10.3390/jpm13121683
- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., and Lew, M. S. (2016). Deep learning for visual understanding: a review. *Neurocomputing* 187, 27–48. doi: 10.1016/j.neucom.2015.09.116
- Gupta, H., Singh, H., and Kumar, A. (2024). Texture and radiomics inspired data-driven cancerous lung nodules severity classification. *Biomed. Signal Process. Control* 88:105543. doi: 10.1016/j.bspc.2023.105543
- Ha, S. M., Yi, A., Yim, D., Jang, M.-J., Kwon, B. R., Shin, S. U., et al. (2023). Digital breast tomosynthesis plus ultrasound versus digital mammography plus ultrasound for screening breast cancer in women with dense breasts. *Korean journal of radiology* 24, 274–283. doi: 10.3348/kjr.2022.0649
- Haraguchi, T., Kobayashi, Y., Hirahara, D., Kobayashi, T., Takaya, E., Nagai, M. T., et al. (2023). Radiomics model of diffusion-weighted whole-body imaging with background signal suppression (DWIBS) for predicting axillary lymph node status in breast cancer. *J. Xray. Sci. Technol.* 31, 627–640. doi: 10.3233/XST-230009
- Hornik, K. (2012). *The Comprehensive R Archive Network*. John Wiley & Sons, Ltd, 394–398. doi: 10.1002/wics.1212
- Ingrisch, M., and Sourbron, S. (2013). Tracer-kinetic modeling of dynamic contrast-enhanced MRI and CT: a primer. *J. Pharmacokinet. Pharmacodyn.* 40, 281–300. doi: 10.1007/s10928-013-9315-3
- Jailin, C., Mohamed, S., Iordache, R., Milioni De Carvalho, P., Ahmed, S. Y., Abdel Sattar, E. A., et al. (2023). AI-based cancer detection model for contrast-enhanced mammography. *Bioengineering* 10:974. doi: 10.3390/bioengineering10080974
- Jamovi (Version 2.5). (2024). *The jamovi project*. Available online at: <https://www.jamovi.org>
- Janowczyk, A., and Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J. Pathol. Inform.* 7. doi: 10.4103/2153-3539.186902
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., et al. (2019). Latent dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimed. Tools Appl.* 78, 15169–15211. doi: 10.1007/s11042-018-6894-4
- Jiang, T., Jiang, W., Chang, S., Wang, H., Niu, S., Yue, Z., et al. (2022). Intratumoral analysis of digital breast tomosynthesis for predicting the Ki67 level in breast cancer: a multi-center radiomics study. *Med. Phys.* 49, 219–230. doi: 10.1002/mp.15392
- Jiang, Y., Zeng, Y., Zuo, Z., Yang, X., Liu, H., Zhou, Y., et al. (2024). Leveraging multimodal MRI-based radiomics analysis with diverse machine learning models to evaluate lymphovascular invasion in clinically node-negative breast cancer. *Heliyon* 10:e23916. doi: 10.1016/j.heliyon.2023.e23916
- Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B., et al. (2012). Radiomics: the process and the challenges. *Magn. Reson. Imag.* 30, 1234–1248. doi: 10.1016/j.mri.2012.06.010
- Lei, S., Zheng, R., Zhang, S., Wang, S., Chen, R., Sun, K., et al. (2021). Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun.* 41, 1183–1194. doi: 10.1002/cac2.12207
- Li, X., Li, M., Yan, P., Li, G., Jiang, Y., Luo, H., et al. (2023). Deep learning attention mechanism in medical image analysis: basics and beyonds. *Int. J. Netw. Dyn. Intell.* 2, 93–116. doi: 10.53941/ijndi0201006
- Li, Y., Fan, Y., Xu, D., Li, Y., Zhong, Z., Pan, H., et al. (2023). Deep learning radiomic analysis of DCE-MRI combined with clinical characteristics predicts pathological complete response to neoadjuvant chemotherapy in breast cancer. *Front. Oncol.* 12:1041142. doi: 10.3389/fonc.2022.1041142
- Liu, Y., Jia, X., Zhao, J., Peng, Y., Yao, X., Hu, X., et al. (2024). A machine learning-based unenhanced radiomics approach to distinguishing between benign and malignant breast lesions using T2-weighted and diffusion-weighted MRI. *J. Magn. Reson. Imag.* 60, 600–612. doi: 10.1002/jmri.29111
- McGuinness, L. A., and Higgins, J. P. T. (2021). Risk-of-bias visualization (robvis): An R package and shiny web app for visualizing risk-of-bias assessments. *Res. Synth. Methods* 12, 55–61. doi: 10.1002/jrsm.1411
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 339:b2535. doi: 10.1136/bmj.b2535
- Murtas, F., Landoni, V., Ordóñez, P., Greco, L., Ferranti, F. R., Russo, A., et al. (2023). Clinical-radiomic models based on digital breast tomosynthesis images: a preliminary investigation of a predictive tool for cancer diagnosis. *Front. Oncol.* 13:1152158. doi: 10.3389/fonc.2023.1152158
- Nicosia, L., Pesapane, F., Bozzini, A. C., Latronico, A., Rotili, A., Ferrari, F., et al. (2023). Prediction of the malignancy of a breast lesion detected on breast ultrasound: radiomics applied to clinical practice. *Cancers* 15:964. doi: 10.3390/cancers15030964
- Oladimeji, O. O., Ayaz, H., McLoughlin, I., and Unnikrishnan, S. (2024). Mutual information-based radiomic feature selection with SHAP explainability for breast cancer diagnosis. *Results Eng.* 24:103071. doi: 10.1016/j.rineng.2024.103071
- Peng, J., Wang, W., Song, Q., Hou, J., Jin, H., Qin, X., et al. (2023). 18f-fdg-pet radiomics based on white matter predicts the progression of mild cognitive impairment to Alzheimer disease: a machine learning study. *Acad. Radiol.* 30, 1874–1884. doi: 10.1016/j.acra.2022.12.033

- Peng, Y., Cheng, Z., Gong, C., Zheng, C., Zhang, X., Wu, Z., et al. (2022). Pretreatment DCE-MRI-based deep learning outperforms radiomics analysis in predicting pathologic complete response to neoadjuvant chemotherapy in breast cancer. *Front. Oncol.* 12:846775. doi: 10.3389/fonc.2022.846775
- Quan, M. Y., Huang, Y. X., Wang, C. Y., Zhang, Q., Chang, C., and Zhou, S. C. (2023). Deep learning radiomics model based on breast ultrasound video to predict HER2 expression status. *Front. Endocrinol.* 14:1144812. doi: 10.3389/fendo.2023.1144812
- Rashid, H. U., Ibricki, T., Payda, S., Binokay, F., and Çevik, U. (2022). Analysis of breast cancer classification robustness with radiomics feature extraction and deep learning techniques. *Expert Syst.* 39, 1–12. doi: 10.1111/exsy.13018
- Rizzo, S., Botta, F., Raimondi, S., Origgi, D., Fanciullo, C., Morganti, A. G., et al. (2018). Radiomics: the facts and the challenges of image analysis. *Eur. Radiol. Exper.* 2:36. doi: 10.1186/s41747-018-0068-z
- Romeo, V., Cuocolo, R., Sanduzzi, L., Carpentiero, V., Caruso, M., Lama, B., et al. (2023). MRI radiomics and machine learning for the prediction of oncotype dx recurrence score in invasive breast cancer. *Cancers* 15:1840. doi: 10.3390/cancers15061840
- Saha, A., Harowicz, M. R., Grimm, L. J., Kim, C. E., Ghate, S. V., Walsh, R., et al. (2018). A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br. J. Cancer* 119, 508. doi: 10.1038/S41416-018-0185-8
- Sharmin, S., Ahammad, T., Talukder, M. A., and Ghose, P. (2023). A hybrid dependable deep feature extraction and ensemble-based machine learning approach for breast cancer detection. *IEEE Access* 11, 87694–87708. doi: 10.1109/ACCESS.2023.3304628
- Sun, K., Zhu, H., Chai, W., and Yan, F. (2023). Multimodality MRI radiomics analysis of TP53 mutations in triple negative breast cancer. *Front. Oncol.* 13:1153261. doi: 10.3389/fonc.2023.1153261
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249. doi: 10.3322/caac.21660
- Taddese, A. A., and Tilahun, B. C. (2024). Deep-learning models for image-based gynecological cancer diagnosis : a systematic review and meta- analysis. *Front Oncol.* 13:1216326. doi: 10.3389/fonc.2023.1216326
- Traverso, A., Wee, L., Dekker, A., and Gillies, R. (2018). Repeatability and reproducibility of radiomic features: a systematic review. *Int. J. Radiat. Oncol.* 102, 1143–1158. doi: 10.1016/j.ijrobp.2018.05.053
- Tsarouchi, M. I., Hoxhaj, A., and Mann, R. M. (2023). New approaches and recommendations for risk-adapted breast cancer screening. *J. Magn. Reson. Imag.* 28, 1–24. doi: 10.1002/jmri.28731
- Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *J. Stat. Softw.* 36, 1–48. doi: 10.18637/jss.v036.i03
- Vigil, N., Barry, M., Amini, A., Akhloufi, M., Maldague, X. P. V., Ma, L., et al. (2022). Dual-intended deep learning model for breast cancer diagnosis in ultrasound imaging. *Cancers* 14:2663. doi: 10.3390/cancers14112663
- Wang, C., Zhao, Y., Wan, M., Huang, L., Liao, L., Guo, L., et al. (2023). Prediction of sentinel lymph node metastasis in breast cancer by using deep learning radiomics based on ultrasound images. *Medicine* 102:E35868. doi: 10.1097/MD.00000000000035868
- Wang, D., Hu, Y., Zhan, C., Zhang, Q., Wu, Y., and Ai, T. (2022). A nomogram based on radiomics signature and deep-learning signature for preoperative prediction of axillary lymph node metastasis in breast cancer. *Front. Oncol.* 12:940655. doi: 10.3389/fonc.2022.940655
- Wang, J., Gao, W., Lu, M., Yao, X., and Yang, D. (2023). Development of an interpretable machine learning model for Ki-67 prediction in breast cancer using intratumoral and peritumoral ultrasound radiomics features. *Front. Oncol.* 13:1290313. doi: 10.3389/fonc.2023.1290313
- Wang, L. C., Rao, S., Schacht, D., and Bhole, S. (2023). Reducing false negatives in biopsy of suspicious MRI findings. *J. Breast Imag.* 5, 597–610. doi: 10.1093/jbi/wbad024
- Wang, Q., Xu, J., Wang, A., Chen, Y., Wang, T., Chen, D., et al. (2023). Systematic review of machine learning-based radiomics approach for predicting microsatellite instability status in colorectal cancer. *Radiol. Med.* 128, 136–148. doi: 10.1007/s11547-023-01593-x
- Wei Li, H. Huang, Y., Zhu, T., Wu, Z., Lin, Y., Ye, G., and Wang, K. (2023). Abstract P6-01-19: multiparametric MRI-based Longitudinal-radiomics analysis for early prediction of treatment response of breast cancers to neoadjuvant chemotherapy: a multicenter study. *Cancer Res.* 83, P6-01-19. doi: 10.1158/1538-7445.SABCS22-P6-01-19
- Wei, W., Ma, Q., Feng, H., Wei, T., Jiang, F., Fan, L., et al. (2023). Deep learning radiomics for prediction of axillary lymph node metastasis in patients with clinical stage T1–2 breast cancer. *Quant. Imaging Med. Surg.* 13, 4995–5011. doi: 10.21037/qims-22-1257
- Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., et al. (2011). Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Int. Med.* 155, 529–536. doi: 10.7326/0003-4819-155-8-201110180-00009
- Wu, X., Guo, Y., Sa, Y., Song, Y., Li, X., Lv, Y., et al. (2022). Contrast-enhanced spectral mammography-based prediction of non-sentinel lymph node metastasis and axillary tumor burden in patients with breast cancer. *Front. Oncol.* 12:823897. doi: 10.3389/fonc.2022.823897
- Xiang, H., Wang, X., Xu, M., Zhang, Y., Zeng, S., Li, C., et al. (2023). Deep learning-assisted diagnosis of breast lesions on US images: a multivendor, multicenter study. *Radiology* 5:e220185. doi: 10.1148/ryai.220185
- Xu, Z., Yang, Q., Li, M., Gu, J., Du, C., Chen, Y., et al. (2022). Predicting HER2 status in breast cancer on ultrasound images using deep learning method. *Front. Oncol.* 12:829041. doi: 10.3389/fonc.2022.829041
- Yang, X., Fan, X., Lin, S., Zhou, Y., Liu, H., Wang, X., et al. (2023). Assessment of lymphovascular invasion in breast cancer using a combined MRI morphological features, radiomics, and deep learning approach based on dynamic contrast-enhanced MRI. *J. Magn. Reson. Imaging* 102:e35868. doi: 10.1002/jmri.29060
- Yu, F.-H., Miao, S.-M., Li, C.-Y., Hang, J., Deng, J., Ye, X.-H., et al. (2023). Pretreatment ultrasound-based deep learning radiomics model for the early prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. *Eur. Radiol.* 33, 5634–5644. doi: 10.1007/s00330-023-09555-7
- Yuan, F., Zhang, Z., and Fang, Z. (2023). An effective CNN and Transformer complementary network for medical image segmentation. *Pattern Recognit.* 136:109228. doi: 10.1016/j.patcog.2022.109228
- Zhang, M., Mesurolle, B., Theriault, M., Meterissian, S., and Morris, E. A. (2023). Imaging of breast cancer beyond the basics. *Curr. Probl. Cancer* 47:100967. doi: 10.1016/j.currprobcancer.2023.100967
- Zheng, X., Huang, Y., Lin, Y., Zhu, T., Zou, J., Wang, S., et al. (2023). 18F-FDG PET/CT-based deep learning radiomics predicts 5-years disease-free survival after failure to achieve pathologic complete response to neoadjuvant chemotherapy in breast cancer. *EJNMMI Res.* 13:105. doi: 10.1186/s13550-023-01053-7