A Comprehensive Survey on AI-Driven Mammography-Based Breast Cancer Detection

Pranjali Shinde[1], Swapnil Kamble[2], Srivaramangai R[3]

[1] Department of Information Technology, University of Mumbai, India

Email: pranjalishinde0504@gmail.com

[2] Department of Information Technology, University of Mumbai, India

Email: swapnil.kamble2701[at]outlook.com

[3] Department of Information Technology, University of Mumbai, India

Email: rsrimangai@gmail.com

Abstract: ***Breast cancer remains a major cause of death from tumors in females globally. Detecting it early via mammograms often leads to better results for those affected. Still, challenges like dense tissue, variation between radiologists, too many scans to review, reduce precision in diagnosis. New tools powered by artificial intelligence - particularly systems based on deep learning - are beginning to help address these problems. This examination looks closely at artificial intelligence techniques applied to breast imaging, drawing insights from more than fifty published research papers. Model structures, data characteristics, performance measures, and real-world medical application form the core of the discussion. Convolutional neural networks appear most frequently across the studies reviewed. Boosted by transfer learning, combined approaches, or inputs from multiple image angles, such systems tend to outperform older detection tools in precision, true positive rates, and AUC scores. Yet even with strong test outcomes, movement into daily hospital practice remains limited. This issue stems from multiple weaknesses, including dependence on past or narrow data, insufficient access to broad and varied mammogram collections, weaker results in women with dense breasts, poor model transparency, alongside missing long-term trials across different hospitals. Gathering today's findings reveals key areas left unexplored, highlighting directions essential for creating trustworthy, clear, and medically useful artificial intelligence tools for breast imaging.***

Keywords: Breast Cancer Detection; Mammography; Artificial Intelligence; Deep Learning; Convolutional Neural Networks; Medical Imaging; Computer-Aided Diagnosis

1. Introduction

1.1 Background

Despite progress in medical science, breast cancer still claims many lives among women globally, standing out as a primary source of cancer deaths even with advances in detection tools and treatments [1], [48]. Because catching it sooner often leads to longer survival, less intense therapies, and healthier daily living, structured screening efforts are seen as vital parts of handling the illness well. Detection through mammograms has taken center stage in public health strategies since these scans can reveal tiny tumours that cannot yet be felt by hand [2], [7]. Given their low expense and availability across regions, mammographic screenings remain key components within both local and global preventive healthcare frameworks.

Even though mammograms matter clinically, reading them well is tough. Because breast tissue varies so much, spotting issues gets harder when images show faint signs or cluttered anatomy - especially where cancer blends into healthy areas [3], [4]. One reader might see something another misses, which means some people get called back needlessly while others slip through gaps. Wrong calls like these lead to avoidable procedures, stress, or later detection that changes treatment paths [5], [11]. As more screenings happen globally, radiologists face heavier loads and tiredness builds - raising chances of mistakes over time [8].

Despite aiming to help radiologists detect abnormalities in mammograms, conventional computer-aided diagnosis tools faced challenges due to rigid design choices. Instead of learning patterns, initial versions depended on predefined rules and manually designed traits, making them less effective under varying scan qualities or among different groups of patients [6], [10]. Because outcomes varied and incorrect alarms occurred frequently, trust in these tools remained low. As a result, adoption in everyday medical settings stayed minimal.

Nowadays, artificial intelligence - especially through deep learning - has changed how medical images are analyzed. Starting from basic mammogram data, convolutional neural networks pull out layered features automatically, so specialists do not have to design them by hand; instead, these systems detect subtle signs linked to cancer growth [9], [14]. Work done recently shows machines guided by AI can match or even outperform experienced radiologists when reading mammograms, especially under structured test conditions and evaluation trials [3], [5]. Because of such progress, there is growing confidence that AI might improve precision in diagnosis while aiding real-world medical choices.

Nowadays, artificial intelligence applied to mammograms draws growing attention - not just as an independent tool but as support for doctors reading scans. One goal stands out: cutting down busy workloads while making results more reliable from one expert to another, boosting how smoothly screenings run [7], [21]. Still, differences in data quality, testing methods, and ways researchers confirm findings make it hard to compare what studies claim. Some depend on past cases or selected samples that might not mirror everyday clinic settings. That gap shows why a careful overview matters - to weigh advances so far, spot stubborn hurdles, and point toward where research should go next in using AI to find breast cancer early [10], [22].

1.2 Statement of the Problem

Even though artificial intelligence in mammography performs well, real-world medical use faces persistent hurdles. Differences in data, image handling techniques, testing criteria, and study designs cause wide variation between research findings [10], [15]. Since numerous algorithms rely on limited or narrow data samples, questions emerge about consistency when applied to diverse patient groups or equipment brands [18], [23].

Looking ahead, many AI applications face limited testing across diverse clinical settings [21], [22]. Because deep neural networks often lack transparency, questions emerge - not only around interpretation but also clinician confidence and responsible deployment [35], [46], [53]. What stands out is the importance of systematically analyzing available studies to gauge real advancements, uncover weaknesses, and point toward unmet needs.

1.3 Purpose of the Research

This paper sets out to explore AI-driven techniques applied in mammogram analysis for breast cancer detection. Drawing from scholarly articles published lately, it collects insights into the design, testing, and integration of these systems within clinical imaging processes [2], [7], [10].

This review focuses on how choices in AI architecture relate to types of mammography data and outcomes like accuracy, sensitivity, specificity, and AUC, based on findings from prior work [3], [5], [18]. By comparing studies, it explores influences on results - such as volume of data, clarity of images, tissue density, and methods involving pre-trained models, combined models, or multiple image views - from referenced sources [20], [23], [45].

Looking beyond individual results, this work examines weaknesses often seen across studies - such as skewed datasets, lack of outside testing, poor interpretability, and hurdles in practical use [22], [35], [46]. Instead of just summarizing outcomes, it pulls insights from more than fifty papers to highlight unanswered questions and emerging paths vital for building AI tools in mammography that perform consistently and earn clinical confidence [29], [55].

2. Literature Review

Early detection stands as a key part of breast cancer management, according to existing research. What makes Waks and Winer's work notable is their focus on treatment pathways shaped by timely diagnosis [1]. Screening practices now face changes due to AI, something Houssami et al. examine through practical integration issues and potential benefits [2]. One turning point emerged when Rodríguez-Ruiz and team tested AI independently against many radiologists, revealing how machine results align with expert interpretations in mammograms [3]. Another angle comes from Rodríguez-Ruiz, who examined the role of artificial intelligence in spotting abnormalities on mammograms, revealing its effect on real-world medical judgments - not merely quoting algorithm performance [4]. In much the same way, work by Kim et al. looked at how machine assistance changes both correct diagnoses and unnecessary follow-ups when used across teams of radiologists, grounding lab findings in actual screening results [5]. Beyond individual trials, broader summaries also appear in current papers. For instance, Bharati and team survey techniques using artificial neural networks in breast imaging, drawing lines between older network designs and today's deeper models [6]. What sets Sechopoulos, Teuwen, and Mann apart is their detailed look at mammography alongside digital breast tomosynthesis, spelling out how these imaging forms differ in practice - details that shape both development and testing of artificial intelligence tools [7]. Instead of building models, Raya-Povedano et al. turn attention to how AI might ease radiologist workload, using real-world markers such as reading duration and process flow to measure effect [8]. From another angle, Shah and team map emerging patterns across studies, piecing together where research stands now while pulling forward recurring ideas seen lately in literature [9]. Looking backward with structure, Zebari's group pull together existing computational approaches used in computer-aided detection for mammograms, offering clarity on processing chains, data handling steps, and ways performance gets assessed [10]. Shifting ground entirely, Díaz and others question if AI truly works within population screening efforts, moving past algorithms toward actual program-level results and what stops integration into routine care [11]. Starting with older machine learning approaches, some research still finds value in traditional techniques due to their presence in current mammography workflows involving manual feature design or mixed modeling steps. Instead of focusing solely on modern networks, one team led by Tian evaluates how selecting specific features impacts diagnostic accuracy and consistency across systems, offering insight into foundational model behaviour [12]. Rather than skipping image preparation, Patel along with Hadia shows how refining raw scans before analysis improves signal clarity for artificial neural networks [13]. Shifting perspective, Yoon together with Kim explore deep learning tools through the lens of medical practitioners, aligning algorithmic outputs with real-world interpretation demands and practical constraints in clinics [14]. Moving beyond isolated methods, work by Wang et al. maps out uses of advanced models in breast cancer detection, forming a cohesive backdrop for understanding evolving trends in automated imaging reviews [15]. From 2022 onward, studies show growing attention to clinical application alongside improved research methods. What stands out is how Sechopoulos, Teuwen, and Mann emphasize a modality-sensitive approach, key for coherence and context in review work [16]. In step with broader progress, Karthiga et al. explore artificial intelligence for detecting abnormalities in recent mammogram technologies, fitting into a pattern of refined deep learning systems and stricter validation standards [17]. Instead of treating tumors uniformly, Lee and team investigate how computer-aided detection tools reflect differences tied to cancer type and image traits - revealing that results may shift depending on biological subtype and appearance [18]. Meanwhile, Al-Fahaidy's group proposes a classification framework applying machine learning to digitized scans, building further on basic algorithms and transparent modeling strategies for interpreting mammographic data [19]. A new hybrid deep learning approach emerges through work by Wang et al., aligning with efforts to blend methods for better reliability and precision [20]. Instead of looking back, Chang's team sets up a forward-looking, multi-site framework named AI-STREAM - an unusual move given most past studies analyze old data; such designs help shape how future tests might unfold [21]. Patterns across many trials add depth to how clinicians interpret what artificial intelligence actually delivers in practice. From Yoon and others comes a broad synthesis focused on independent AI use during breast imaging exams, using pooled results to ground claims more firmly than isolated experiments could [22]. In another direction, Badawy's group checks if dense tissue alters how well AI works in mammograms, revealing it may not only influence diagnosis but also expose weaknesses in algorithm behavior and equity issues [23]. A look at CNN performance in mammography comes from Karthik with team, stressing side-by-side testing of network types instead of standalone claims [24]. Work by Dan and others surveys artificial intelligence uses in breast imaging, situating mammography within broader techniques without shifting focus away from its central role [25]. Some studies go beyond mammograms alone, yet remain relevant by revealing how models adapt across breast imaging forms - offering clues about design decisions and data handling that extend further [26]. In one case, Abunasser et al. examine deep learning for image-based tumor sorting, useful when comparing methods even if ultrasound or MRI appear instead of X-ray images [26]. Another path emerges through Trang's group, building a system merging scan data with patient histories - a blend where numbers from charts support visual findings, nudging automated diagnosis closer to actual clinic conditions [27]. One key point reappears often: density affects model results - this remains central in evaluating AI meant for screening tasks, notes Badawy in a repeated entry [28]. Looking across multiple trials, Da Silva along with co-authors examines how artificial intelligence measures up against standard imaging techniques, showing patterns in findings while also pointing out gaps in consistency among published reports [29]. By 2024, questions shift slightly - not just whether these tools function - but more deeply into methods for fair comparison and realistic integration. What stands out in Petchiappan's analysis is the need for structured testing between different models rather than trusting isolated research outcomes [30]. A clearer picture emerges through Khan et al., who gather insights on deep learning applications in breast scans, summarizing common network designs, data sources, and trends in accuracy drawn from current work [31]. Starting with newer forms of imaging, Kinkar et al. examine how artificial intelligence adapts to contrast-enhanced mammography - a shift away from traditional full-field digital methods - raising concerns about consistency across varied image types [32]. Moving into pattern recognition, Ahmad and team outline advances in deep learning applied to identifying and sorting breast cancer cases, highlighting common research strategies despite wide variation in implementation [33]. From an environmental angle, El-Mawla and co-authors propose energy-conscious models for analyzing mammograms, bringing attention to computational cost alongside performance [34]. Ending with practical hurdles, Díaz, Rodríguez-Ruiz, and Sechopoulos explore what stands in the way of real-world integration, touching on evaluation standards and long-term viability of these systems [35]. From 2023 to 2024, research in clinical radiology expands insight into screening practices, moving past basic detection rates. Instead of relying only on theoretical benchmarks, Chen et al. evaluate artificial intelligence through tailored screening protocols, offering clearer alignment with real-world conditions [36]. While routine scans aim to catch tumors early, Nanaa et al. examine cancers that emerge between screenings - a gap revealing where current methods fall short and where AI might prove more reliable [37]. Shifting focus to diagnosis, Krishna and Mahboub explore mammogram interpretation, highlighting adjustments that improve accuracy without overhauling existing workflows [38]. Meanwhile, Zhu and team emphasize preprocessing techniques in imaging, suggesting that foundational steps still shape outcomes, even when advanced algorithms are applied [39]. Looking across recent studies, Nafissi et al. summarize artificial intelligence applications in breast cancer, using reviews to bring together varied outcomes [40]. Moving into imaging, Patra and team examine how AI identifies tumors and evaluates their seriousness, signaling a shift toward more detailed medical analysis rather than just spotting abnormalities [41]. By 2025, newer contributions continue building on earlier knowledge while extending its scope. From another angle, Qureshi and co-authors trace progress from basic image handling through advanced neural networks in mammogram interpretation, offering structure for categorizing research by workflow phases [42]. Elsewhere, Ali and collaborators assess numerous algorithms used in finding and diagnosing lesions, compiling broad insights about existing tools alongside challenges limiting real-world use [43]. Looking at AI in breast imaging, Dave's team presents a narrative overview stressing careful assessment of evidence when evaluating AI-supported mammograms, highlighting how results should be seen using standard measures of diagnostic performance [44]. Instead of single images, Abdikenov's group explores approaches using multiple views, supporting recent observations that such models match real-world workflows more closely while improving reliability [45]. From another angle, Shifa and coworkers investigate tools that make AI decisions clearer in screening, outlining definitions, evaluation methods, and reasons transparency matters for user confidence [46]. Some studies shift focus beyond X-ray based exams to include tissue analysis and sound wave imaging. Still, they contribute useful perspectives by comparing how explanations and testing are handled across diverse AI systems. One example comes from Alom and associates, who built a deep learning model designed to offer insight into its reasoning, bridging pathology slides and ultrasound data. This clarification sheds light on methods for transparency in mammography analysis [47]. Public health priorities around early screening are underlined by the WHO, reinforcing why research often centers on mammography [48]. A proposed method by Tanveer et al. applies machine learning to catch signs earlier. Despite deep learning dominating recent breakthroughs, basic machine learning concepts continue drawing attention [49]. Specialized reviews along with tailored strategies signal movement away from broad frameworks - custom reasoning is becoming more relevant. Instead of just spotting abnormalities, AI now pulls extra insights from mammograms - Hosseinzadeh's team shows how receptors can be profiled directly through imaging [50].

Shifting focus, Pesapane's work highlights customization: screening isn't one-size-fits-all anymore, but adapts per person, guided by AI-driven risk estimates [51]. From another angle, Saeidnia proposes a blueprint where algorithms help shape medical decisions, blending data patterns into real-world judgment [52]. Meanwhile, Ansari explores transparency, building models that reveal their logic when supporting breast cancer diagnoses [51]. What stands out is how often interpretability now appears essential, not just added on by choice [53]. Lately, researchers point toward stronger agreement - consensus matters more than ever before. One contribution comes from SalekShahabi, who pulls together machine learning and deep learning methods in a structured way, building clearer summaries while pushing for head-to-head method evaluations alongside complete documentation [54]. Work led by Añez combines broad literature analysis with extensive testing across datasets. That shift helps close a long-standing mismatch: reviews rarely match real-world performance at scale [55]. A fresh method built on YOLOv8, introduced by Raeisi et al., tweaks network design to improve tumor spotting in mammogram images. While detection methods keep evolving, emphasis now leans into precision when locating abnormalities [56]. Progress in AI for breast imaging moves beyond handcrafted features, embracing deep networks evaluated not just in labs but also in settings resembling real-world screening. Yet alongside gains, hurdles remain visible - uneven data representation and limited population variety trouble reliability [23]; many models lack testing outside original sites [22]. Clinicians question how decisions are made inside black-box systems [46], while fitting tools smoothly into radiology routines proves complex [11]. Because of this mix of promise and friction, comparing different approaches across varied databases becomes essential; so does advancing research that spans multiple hospitals, adapts to tissue density differences, and delivers transparent reasoning matching actual medical practice.

3. Objectives

This paper aims to deliver a clear, data-driven overview of artificial intelligence systems built for detecting breast cancer through mammograms. Drawing from up-to-date scholarly articles, it pulls together findings to highlight patterns in effectiveness, advances in techniques, along with key gaps in current work.

The specific objectives of this study are as follows:

i. To examine AI and deep learning models used in mammography, focusing on convolutional neural network (CNN) structures, ensemble frameworks, and hybrid methods for breast lesion detection and classification.
ii. To compare datasets and performance metrics by looking at public and private mammography datasets, evaluation methods, and reported diagnostic metrics such as accuracy, sensitivity, specificity, and AUC.
iii. To identify limitations and research gaps, including issues related to dataset imbalance, model generalization, and challenges with clinical use.
iv. To understand performance trends and clinical implications by evaluating the factors that affect model reliability, interpretability, and usability in population-based breast cancer screening.
v. To lay a research-informed groundwork for future AI models that focus on explainability, dataset variety, and ethical integration into clinical mammography practices.

4. Methodology / Survey Framework

This work relies on a structured review of existing papers to explore how artificial intelligence supports mammography in identifying and assessing breast cancer. Because the aim involves analyzing and contrasting earlier findings rather than building a fresh forecasting tool, an observational overview strategy guides the process.

4.1 Literature Identification and Selection

Searching major scientific platforms - such as IEEE Xplore, SpringerLink, ScienceDirect, PubMed Central, MDPI, Wiley Online Library, PLOS, and specialty radiology publications - led to identifying pertinent research. Articles selected came after careful examination of peer-reviewed literature, with attention given to systematic reviews, meta-analyses, and applied artificial intelligence investigations centered on mammography for detecting and evaluating breast cancer.

Some of the selected works focused just on how artificial intelligence functions within mammography. Those dealing exclusively with different imaging approaches did not make the cut - unless their methods could clearly benefit breast image evaluation.

4.2 Inclusion and Exclusion Criteria

Studies were included if they:

- Focusing on mammography, researchers applied methods like artificial intelligence. Machine learning approaches were included in some studies. Deep learning models also played a role across various analyses
- Offered a straightforward outline of the model's layout, the data used, or how results were assessed
- Appeared in scholarly journals or established research outlets.

Studies were excluded if they:

- Focused solely on techniques beyond mammography for imaging purposes
- Missing depth in explanation, lacking clear technical insight
- Some lacked peer review, while others omitted original studies or systematic analyses

4.3 Data Extraction

From each selected study, key technical and clinical attributes were extracted to enable structured comparison. These included:

- AI model or algorithm used
- Mammography dataset(s) employed

- Evaluation metrics reported (e.g., accuracy, sensitivity, specificity, AUC)
- Key findings relevant to diagnostic performance
- Reported limitations and clinical constraints

This information formed the basis for comparative evaluation across studies.

4.4 Comparative Analysis Strategy

The extracted information was analysed by grouping studies according to shared methodological and clinical characteristics. Comparisons were conducted across four primary dimensions:

- **AI methodology**, to examine differences between deep learning, ensemble, hybrid, and explainable AI approaches
- **Dataset characteristics**, to assess the influence of dataset size, diversity, and image quality on model performance
- **Evaluation metrics**, to understand how diagnostic performance was measured and reported
- **Reported limitations**, to identify recurring challenges affecting generalisability and clinical adoption

This strategy enabled identification of performance trends, methodological strengths, and research gaps across the literature.

Model / Approach	Dataset Used	Metric Performance	Key Limitation	Ref. No.
Transpara v1.4.0 (Standalone AI)	Large multi-center screening mammography dataset	AUC = 0.840 (95% CI: 0.820–0.860)	Tested only on past data, not real-time screening	[3]
AI-assisted DL CAD system	Screening mammograms (clinical reader study)	AUC improved from 0.87 to 0.89 with AI	Needs radiologist interaction to work well	[4]
Standalone CNN-based AI system	Multi-vendor mammography screening datasets	AUROC = 0.959	Tested on datasets with more cancer cases than normal	[5]
Feature Selection + Classifier model	Mammography images (public dataset)	AUC = 0.867 ± 0.023	Uses limited handcrafted features	[12]
HOFS + ANN CAD framework	Digital mammogram images	Accuracy = 98.97%	Requires manual selection of tumor regions	[13]
CNN-based CAD models (reviewed)	DDSM, CBIS-DDSM, MIAS	Detection accuracy ≈ 85.51%	Mostly tested in laboratory conditions	[14]
Lunit INSIGHT MMG (Commercial AI-CAD)	Screening mammography dataset	Sensitivity = 88.2%	Evaluated on cancer-enriched data	[18]
Hybrid Deep Learning model	Digital mammograms	Accuracy = 97.8%	High computational cost	[20]
Meta-analysis of standalone AI systems	Mammography and DBT studies	Pooled AUC ≈ 0.88–0.90	Results vary across different studies	[22]
AI-aided mammography (density-based study)	Mammograms grouped by breast density	AUC drops from 0.91 to 0.85 in dense breasts	Lower performance in dense breast cases	[23]
CNN architecture comparison study	CBIS-DDSM dataset	Best AUC > 0.90 (ResNet models)	Evaluated on only one dataset	[24]
Bi-xBcNet-96 (Green AI CNN)	Mammography images	Accuracy = 99.12%, Sensitivity = 98.45%	Not tested on data from other hospitals	[34]
Lunit INSIGHT MMG (Screening study)	Large screening mammography dataset	Cancer detection rate increased by ~9%	Retrospective analysis only	[36]
Interval cancer detection AI	Screening mammography dataset	Sensitivity ≈ 84% for interval cancers	Tested on limited population groups	[37]
Multi-view CNN strategy	CC and MLO mammographic views	AUC improved by ~3–5%	Needs more data and processing time	[45]
YOLOv8 with attention modules	Mammography images	Detection accuracy ≈ 94%, mAP > 0.90	Requires many labeled images	[56]

5. Discussion

5.1 Performance Trends Across AI Models

Looking back at recent studies reveals steady gains in how well artificial intelligence works for spotting issues in breast scans. Instead of older tools that relied on preset image traits and basic algorithms, most now use advanced neural networks built for visual data. These newer designs get stronger results because they figure out intricate patterns straight from the images themselves without needing manual input [7], [14].

When examined across broad studies, modern CNN-driven systems show strong results, their AUC scores often ranging from 0.84 to 0.96. Depending on data traits and testing approaches, outcomes shift noticeably [3], [4], [5]. One standout example comes from Kim et al., whose model reached an AUROC of 0.959 by leveraging extensive screening mammography records. In isolated detection scenarios, such accuracy matches - or even exceeds - typical radiologist performance [5]. Under tightly managed conditions, alternative deep learning architectures likewise maintain robust levels of both sensitivity and specificity [18], [22].

Feature extraction happens naturally within CNNs, thanks to their layered design. Because of this, patterns like abnormal tissue texture or unusual growth forms become detectable. Contextual information also gets captured, which helps when judging if a finding is cancerous. Pre-training on vast collections of images gives some models a head start. When fine-tuned for breast imaging, they adapt well, due to prior exposure. Performance gains appear clearly compared to networks built from ground up [24], [34]. Still, even with strong performance in labeling images correctly, models built on convolutional networks react strongly to distortions in pictures, shifts in capture methods, or equipment differences across manufacturers. Their usefulness drops when applied outside controlled settings because of these factors [9], [35].

5.2 Influence of Dataset Size, Diversity and Quality

A key influence on how well AI works in mammography lies in the traits of the data used. Early progress relied heavily on open resources such as DDSM, CBIS-DDSM, and INbreast - these enabled consistent testing across studies. Still, because they cover only a narrow range of imaging setups, include few subjects, and lack variety in patient backgrounds, models built on them often fail to apply widely [10], [18].

On the other hand, research using broad, multisite screenings from actual clinical settings tends to report steadier outcomes. Work by Rodríguez-Ruiz and team, along with Kim's group, indicates systems built on varied hospital data adapt more reliably across different environments - retaining effectiveness even with unfamiliar inputs [3], [5]. What these findings point to is clear: variation in population traits, tissue composition, machines used, and scanning procedures matters just as much as sheer volume of data.

One major hurdle seen in many datasets is uneven class distribution, as abnormal results appear only rarely among routine breast scans. Although techniques such as synthetic data generation, adjusted classification penalties, or repeating minority samples are common, these approaches fail to fully reflect actual population patterns [22]. In broad screening programs, the issue grows more serious - cancer may occur less than once per hundred exams. Such conditions can distort accuracy measurements when models are tested [36].

5.3 Role of Transfer Learning and Data Augmentation

Starting with knowledge gained elsewhere, transfer learning helps adapt powerful image features to mammograms when labeled medical data is scarce. Fine-tuning models initially trained on broad visual tasks improves results in breast imaging. Instead of building networks from scratch, researchers leverage architectures such as ResNet, DenseNet, or EfficientNet - previously exposed to vast image collections. Evidence suggests these approaches reach stable performance more quickly while boosting accuracy across diagnostic benchmarks [24], [34].

Turning images, mirroring them, tweaking brightness, or balancing histograms helps reflect variations in how breasts are positioned during scans. Instead of just relying on real data, some studies explore using GANs to create artificial images when examples are limited or unevenly spread across categories [8], [39]. Such approaches may help models perform better on unseen cases by lowering the risk of memorizing too much from few samples.

However, too much manipulation or clumsy adjustments might blur key details doctors rely on, steering algorithms off track. Relying on models trained outside healthcare could pull results toward unrelated patterns instead. Ahead, work will probably shift toward training systems directly on vast sets of medical scans - sharpening their grasp of breast imaging traits and real-world diagnostics [33], [55].

5.4 Interpretability and Clinical Applicability

Even though these AI tools show strong results, doctors remain cautious because they cannot see how decisions are made. Deep neural networks often work like closed systems - offering answers but not explanations. Without transparency, medical professionals hesitate to rely on them. Regulatory settings demand clarity, making opaque methods difficult to accept [35], [46].

Some tools - such as saliency maps, Grad-CAM, or attention displays - are designed to reveal what parts of an image influence a model's output. Rather than guessing, doctors can see if those regions overlap with known medical signs [46], [53]. When highlighted zones resemble actual lesions marked by experts, trust tends to grow. Evidence shows physicians feel more assured when AI reasoning reflects familiar patterns [7].

Still, being interpretable does not automatically make an AI tool useful in practice. Many top-performing models were assessed after the fact, relying on carefully selected data. Such methods can inflate expectations about how well they work outside controlled settings [22]. Rare are the trials that look ahead, involve multiple readers, or span several sites - approaches widely considered essential for judging real medical impact [21], [37]. Still, getting these tools to work inside hospitals means they must fit with existing imaging storage and reporting systems. Approval from health regulators is also necessary. Who takes legal responsibility if something goes wrong has yet to be decided. Important questions like these have no answers so far [11], [44].

5.5 Limitations and Future Challenges

Even though research advances quickly in AI for mammography, important gaps still exist. Without sizable, uniform, and varied imaging data, creating fair and widely applicable systems becomes difficult. Moreover, limited real-time clinical testing weakens confidence in actual performance and reliability.

Bias shows up when models learn mostly from dominant groups, leaving others behind - this challenges fairness. Populations left out during training might get unreliable results, creating uneven outcomes in practice. When systems lack methods to show their confidence levels, trust becomes harder in medical settings. Uncertainty awareness matters deeply where choices affect health.

Last of all, making models more complex brings issues like higher demands on processing resources, difficulty interpreting outcomes, while also complicating their integration into standard screening workflows. Addressing such hurdles calls for collaboration - between those developing artificial intelligence methods, medical imaging specialists, hospitals, oversight bodies, along with experts in ethics - to support dependable, equitable, clinically useful tools built around automated breast scan analysis.

6. Conclusion and Future Scope

This comprehensive literature survey of recent studies explores how artificial intelligence is being used in mammography to detect and diagnose breast cancer. Attention went toward the structure of models, characteristics of data collections, ways results are measured, besides real-world medical application. A noticeable move has emerged - away from older automated detection tools toward advanced deep learning techniques, particularly those built on convolutional neural networks. Through strategies like reusing pre-trained models, combining multiple models, or analyzing several image angles at once, these systems frequently show stronger outcomes across precision, ability to catch true cases, along with overall predictive strength. Under specific test settings, certain models perform nearly as well as, sometimes even outperforming, typical radiologist readings.

Still, strong results in reports do not guarantee readiness for actual medical use. Some research relies on past data or samples with too many cancer examples, skewing how well systems work when used widely. A shortage of big, consistent, varied collections of mammograms holds back progress toward reliable and equitable tools. Problems like dense breast patterns, uneven case distribution, and variations between scanner brands add layers to the difficulty. These hurdles shape what current methods can truly achieve.

Looking at current evidence, only a handful of multi-site trials have tested AI tools in practice - this leaves major gaps in understanding. Although some algorithms show potential for aiding decisions, actual benefits in diagnosis precision, efficiency gains, or better health results remain unclear. Surprisingly, even advanced deep learning models often lack transparency, while explanations meant to clarify them frequently fail to deliver consistent insights. Trust among doctors and approval by regulators still lag because clarity and dependability fall short.

Looking ahead, work must center on building trustworthy AI tools using broad and varied mammography data. Greater attention is needed in testing these systems within real-world medical settings. Factors like tissue density and equitable performance require careful inclusion during design. Estimating confidence in predictions matters just as much as accuracy itself. Fitting new models into current imaging processes without disruption remains a key hurdle. Progress depends heavily on joint efforts among technologists, clinicians, and ethicists. Only through such cooperation can AI move beyond lab results toward responsible use in detecting breast cancer effectively.

7. References

[1] A. G. Waks and E. P. Winer, “Breast Cancer Treatment: A Review,” JAMA, vol. 321, no. 3, pp. 288–300, Jan. 22, 2019, doi:10.1001/jama.2018.19323.Available: https://doi.org/10.1080/17434440.2019.1610387

[2] N. Houssami, G. Kirkpatrick-Jones, N. Noguchi, and C. I. Lee, “Artificial intelligence (AI) for the early detection of breast cancer: A scoping review to assess AI’s potential in breast screening practice,” Expert Rev. Med. Devices, vol. 16, no. 5, pp. 351–362, 2019, doi: 10.1080/17434440.2019.1610387.

[3] A. Rodríguez-Ruiz et al., “Stand-alone artificial intelligence for breast cancer detection in mammography: Comparison with 101 radiologists,” J. Natl. Cancer Inst., vol. 111, no. 9, pp. 916–922, 2019, doi: 10.1093/jnci/djy222.

[4] A. Rodríguez-Ruiz et al., “Detection of breast cancer with mammography: Effect of an artificial intelligence support system,” Radiology, vol. 290, no. 2, pp. 305–314, 2019, doi: 10.1148/radiol.2018181371.

[5] H.-E. Kim et al., “Changes in cancer detection and false-positive recall in mammography using artificial intelligence: A retrospective, multireader study,” Lancet Digit. Health, vol. 2, no. 3, pp. e138–e148, 2020, doi: 10.1016/S2589-7500(20)30003-0.

[6] S. Bharati, P. Podder, and M. R. H. Mondal, “Artificial neural network based breast cancer screening: A comprehensive review,” arXiv preprint arXiv:2006.01767, 2020.

[7] I. Sechopoulos, J. Teuwen, and R. Mann, “Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: state of the art,” Seminars in Cancer Biology, vol. 72, pp. 214–225, Jul. 2021, doi: 10.1016/j.semcancer.2020.06.002.

[8] J. L. Raya-Povedano et al., “AI-Based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation,” Radiology, vol. 300, no. 1, pp. 57–65, 2021, doi: 10.1148/radiol.2021203555.

[9] S. M. Shah, R. A. Khan, S. Arif, and U. Sajid, “Artificial Intelligence for Breast Cancer Detection: Trends & Directions,” arXiv preprint arXiv:2110.00942, 2021.

[10] D. A. Zebari et al., "Systematic review of computing approaches for breast cancer detection based computer aided diagnosis using mammogram images," Applied Artificial Intelligence, vol. 35, no. 15, pp. 2157–2203, 2021, doi: 10.1080/08839514.2021.2001177.

[11] O. Díaz, A. Rodríguez-Ruiz, A. Gubern-Mérida, R. Martí, and M. Chevalier, "Are artificial intelligence systems useful in breast cancer screening programmes?" Radiología, vol. 63, no. 3, pp. 236–244, 2021, doi: 10.1016/j.rx.2020.11.005.

[12] C.-J. Tian, J. Lv, and X.-F. Xu, "Evaluation of feature selection methods for mammographic breast cancer diagnosis in a unified framework," BioMed Research International, vol. 2021, Art. ID 6079163, 2021, doi: 10.1155/2021/6079163.

[13] J. J. Patel and S. K. Hadia, "An enhancement of mammogram images for breast cancer classification using artificial neural networks," IAES International Journal of Artificial Intelligence, vol. 10, no. 2, pp. 332–345, 2021, doi: 10.11591/ijai.v10.i2.pp332-345.

[14] J. H. Yoon and E.-K. Kim, "Deep learning-based artificial intelligence for mammography," Korean Journal of Radiology, vol. 22, no. 8, pp. 1225–1239, 2021, doi: 10.3348/kjr.2020.1210.

[15] Y. Wang et al., "Application of artificial intelligence based on deep learning in breast cancer screening and imaging diagnosis," Neural Computing and Applications, vol. 33, pp. 9637–9647, 2021, doi: 10.1007/s00521-021-05728-x.

[16] I. Sechopoulos, J. Teuwen, and R. Mann, "Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art," Semin. Cancer Biol., vol. 72, pp. 214–225, 2021, doi: 10.1016/j.semcancer.2020.06.002.

[17] R. Karthiga, K. Narasimhan, and R. Amirtharajan, "Diagnosis of breast cancer for modern mammography using artificial intelligence," Mathematics and Computers in Simulation, vol. 202, pp. 316–330, 2022, doi:10.1016/j.matcom.2022.05.038.

[18] S. E. Lee, K. Han, J. H. Yoon, J. H. Youk, and E.-K. Kim, "Depiction of breast cancers on digital mammograms by artificial intelligence-based computer-assisted diagnosis according to cancer characteristics," European Radiology, vol. 32, no. 11, pp. 7400–7408, 2022, doi:10.1007/s00330-022-08718-2.

[19] F. A. K. Al-Fahaidy, B. A. Al-Fuhaidi, I. Al-Darouby, F. Al-Abady, M. Al-Qadry, and A. Al-Gamal, "A Diagnostic Model of Breast Cancer Based on Digital Mammogram Images Using Machine Learning Techniques," Applied Computational Intelligence and Soft Computing, vol. 2022, 2022, doi:10.1155/2022/3895976.

[20] X. Wang, I. Ahmad, D. Javeed, S. A. Zaidi, F. M. Alotaibi, M. E. Ghoneim, Y. I. Daradkeh, J. Asghar, and E. T. Eldin, "Intelligent Hybrid Deep Learning Model for Breast Cancer Detection," Electronics, vol. 11, no. 17, art. 2767, 2022, doi:10.3390/electronics11172767.

[21] Y.-W. Chang et al., "Artificial Intelligence for Breast Cancer Screening in Mammography (AI-STREAM): A Prospective Multicenter Study Design in Korea Using AI-Based CADe/x," J. Breast Cancer, vol. 25, no. 1, pp. 57–68, 2022, doi: 10.4048/jbc.2022.25.e4.

[22] J. H. Yoon, F. Strand, P. A. T. Baltzer, E. F. C. Johannessen, L. A. Miles, R. M. Nishikawa, et al., "Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: A systematic review and meta-analysis," Radiol. Artif. Intell., vol. 307, 2023, doi:10.1148/radiol.222639.

[23] E. Badawy, R. ElNaggar, S. A. M. Soliman, D. S. Elmesidy, et al., "Performance of AI-aided mammography in breast cancer diagnosis: Does breast density matter?," Egyptian Journal of Radiology and Nuclear Medicine, vol. 54, article 178, 2023, doi: 10.1186/s43055-023-01129-3.

[24] R. S. Karthik, S. Anitha, V. Lakshmi, and P. Ramesh, "Breast Cancer Detection Using Deep Learning on Mammography: A Comprehensive Evaluation of CNN Models," IEEE Xplore, 2023, doi:10.1109/10817580.

[25] Z. Dan, X. He, and J. Jing, "Overview of Artificial Intelligence in Breast Cancer Medical Imaging," J. Clin. Med., vol. 12, no. 2, art. 419, 2023, doi:10.3390/jcm12020419.

[26] B. S. Abunasser, M. R. J. AL-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, "Convolution Neural Network for Breast Cancer Detection and Classification Using Deep Learning," Asian Pac. J. Cancer Prev., vol. 24, no. 2, pp. 531–544, 2023.

[27] N. T. H. Trang, K. Q. Long, P. L. An, and T. N. Dang, "Development of an Artificial Intelligence-Based Breast Cancer Detection Model by Combining Mammograms and Medical Health Records," Diagnostics, vol. 13, no. 3, p. 346, 2023, doi:10.3390/diagnostics13030346.

[28] E. Badawy et al., "Performance of AI-Aided Mammography in Breast Cancer Diagnosis: Does Breast Density Matter?" Egypt. J. Radiol. Nucl. Med., vol. 54, no. 178, 2023, doi: 10.1186/s43055-023-01129-3.

[29] H. E. C. da Silva et al., "The use of artificial intelligence tools in cancer detection compared to the traditional diagnostic imaging methods: An overview of the systematic reviews," PLOS ONE, vol. 18, no. 10, e0292063, 2023, doi: 10.1371/journal.pone.0292063.

[30] Thangaraj Petchiappan, "Evaluation of AI-Driven Models for Breast Cancer Detection: An In-Depth Comparative Study of Algorithms," CMJ Publishers, 2024.

[31] S. A. Khan et al., “Artificial Intelligence–Powered Mammography: Navigating the Landscape of Deep Learning for Breast Cancer Detection,” Cureus, vol. 16, no. 3, 2024.

[32] K. K. Kinkar, B. K. K. Fields, M. W. Yamashita, and B. A. Varghese, “Empowering breast cancer diagnosis and radiology practice: advances in artificial intelligence for contrast-enhanced mammography,” Front. Radiol., vol. 3, Jan. 2024, doi:10.3389/fradi.2023.1326831.

[33] J. Ahmad, S. Akram, A. Jaffar, Z. Ali, S. M. Bhatti, A. Ahmad, and S. U. Rehman, “Deep learning empowered breast cancer diagnosis: Advancements in detection and classification,” PLoS ONE, vol. 19, no. 7, p. e0304757, Jul. 2024, doi:10.1371/journal.pone.0304757.

[34] N. A. El-Mawla, M. A. Berbar, N. A. El-Fishawy, and M. A. El-Rashidy, “A novel deep learning approach (Bi-xBcNet-96) considering green AI to discover breast cancer using mammography images,” Neural Computing and Applications, vol. 36, pp. 12701–12723, 2024, doi:10.1007/s00521-024-09815-7.

[35] O. Díaz, A. Rodríguez-Ruíz, and I. Sechopoulos, “Artificial intelligence for breast cancer detection: technology, challenges, and prospects,” Eur. J. Radiol., vol. 175, p. 111457, 2024, doi:10.1016/j.ejrad.2024.111457.

[36] Y. Chen, A. G. Taib, I. T. Darker, and J. J. James, “Performance of a Breast Cancer Detection AI Algorithm Using the Personal Performance in Mammographic Screening Scheme,” Radiology, vol. 308, no. 3, pp. e223299, 2023, doi: 10.1148/radiol.223299.

[37] M. Nanaa et al., “Accuracy of an Artificial Intelligence System for Interval Breast Cancer Detection at Screening Mammography,” Radiology, vol. 312, no. 2, pp. e232303, 2024, doi: 10.1148/radiol.232303.

[38] T. G. Krishna and M. A. A. Mahboub, “Improving breast cancer diagnosis with AI mammogram image analysis,” Medinformatics, vol. 2, no. 3, pp. 195–204, 2024, doi: 10.47852/bonviewMEDIN42023806.

[39] Z. Zhu, Y. Sun, and B. H. S. Asli, “Early breast cancer detection using artificial intelligence techniques based on advanced image processing tools,” Electronics, vol. 13, no. 17, art. no. 3575, 2024, doi: 10.3390/electronics13173575.

[40] N. Nafissi et al., “The application of artificial intelligence in breast cancer,” Eurasian J. Med. Oncol., vol. 8, no. 3, pp. 235–244, 2024, doi: 10.14744/ejmo.2024.45903.

[41] A. Patra et al., “Transformative insights: Image-based breast cancer detection and severity assessment through advanced AI techniques,” Journal of Intelligent Systems, vol. 33, art. no. 20240172, 2024, doi: 10.1515/jisys-2024-0172.

[42] S. A. Qureshi et al., “Breast cancer detection using mammography: Image processing to deep learning,” IEEE Access, vol. 13, pp. 60776–60801, 2025, doi: 10.1109/ACCESS.2024.3523745.

[43] A. Ali, M. Alghamdi, S. S. Marzuki, T. A. D. A. Tengku Din, M. S. Yamin, M. Alrashidi, I. S. Alkhazi, and N. Ahmed, “Exploring AI Approaches for Breast Cancer Detection and Diagnosis: A Review Article,” Breast Cancer: Targets and Therapy, vol. 17, pp. 927–947, Oct. 2025, doi: 10.2147/BCTT.S550307.

[44] D. Dave, A. Akhunzada, N. Ivković, S. Gyawali, K. Cengiz, A. Ahmed, and A. S. Al-Shamayleh, “Diagnostic test accuracy of AI-assisted mammography for breast imaging: a narrative review,” PeerJ Computer Science, vol. 11, p. e2476, 2025, doi: 10.7717/peerj-cs.2476.

[45] B. Abdikenov, T. Zhaksylyk, A. Imasheva, Y. Orazayev, and T. Karibekov, “Innovative Multi-View Strategies for AI-Assisted Breast Cancer Detection in Mammography,” J. Imaging, vol. 11, no. 8, art. 247, 2025, doi:10.3390/jimaging11080247.

[46] N. Shifa, M. Saleh, Y. Akbari and S. Al Maadeed, “A Review of Explainable AI Techniques and Their Evaluation in Mammography for Breast Cancer Screening,” Clinical Imaging, vol. 110, 2025, doi:10.1016/j.clinimag.2025.110492.

[47] M. R. Alom, F. A. Farid, M. A. Rahaman, A. Rahman, T. Debnath, A. S. M. Miah, and S. Mansor, “An explainable AI-driven deep neural network for accurate breast cancer detection from histopathological and ultrasound images,” Scientific Reports, vol. 15, art. 17531, 2025, doi:10.1038/s41598-025-97718-5.

[48] World Health Organization, “Breast cancer,” WHO Fact Sheets, Aug. 14, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[49] H. Tanveer, M. Faheem, A. H. Khan, and M. A. Adam, “AI-Powered Diagnosis: A Machine Learning Approach to Early Detection of Breast Cancer,” International Journal of Engineering Development and Research, vol. 13, no. 2, pp. 153–166, Apr. 2025.

[50] N. Hosseinzadeh et al., “Non-invasive breast cancer receptor typing from mammograms using artificial intelligence: A systematic review and meta-analysis,” J. Imaging Informatics Med., 2025, doi: 10.1007/s10278-025-01775-1.

[51] F. Pesapane et al., “Artificial intelligence-driven personalization in breast cancer screening: From population models to individualized protocols,” Cancers, vol. 17, no. 17, art. no. 2901, 2025, doi: 10.3390/cancers17172901.

[52] H. R. Saeidnia, C. Sparks, and H. Soleymani Majd, “Conceptual model for breast cancer diagnosis using machine

learning on mammogram data," InfoScience Trends, vol. 2, no. 2, pp. 11–26, 2025, doi: 10.61186/ist.202502.02.02.

[53] Z. A. Ansari, M. M. Tripathi, and R. Ahmed, "The role of explainable AI in enhancing breast cancer diagnosis using machine learning and deep learning models," Discover Artificial Intelligence, vol. 5, art. no. 75, 2025, doi: 10.1007/s44163-025-00307-8.

[54] M. SalekShahabi, "Artificial intelligence for early detection and diagnosis of breast cancer: A systematic review of machine learning and deep learning approaches," Eurasian J. Chem. Med. Pet. Res., vol. 5, no. 1, pp. 64–76, 2026, doi: 10.5281/zenodo.17494535.

[55] D. Añez et al., "Artificial intelligence pipeline for mammography-based breast cancer detection: An integrated systematic review and large-scale experimental validation," Medicina, vol. 61, no. 12, art. no. 2237, 2026, doi: 10.3390/medicina61122237.

[56] Z. Raeisi, A. Roshanzamir, F. A. Lomer, and R. A. Lashaki, "YOLOv8 with innovative dilated residual and attention modules for mammographic tumor detection," Computers and Electrical Engineering, vol. 130, art. no. 110903, 2026, doi: 10.1016/j.compeleceng.2025.110903.