Deep learning-based computer-aided detection of breast lesions

Waldemar Wójcik[1,†], Oleksandr Poplavskyi[2,*,†], Sergii Pavlov[3,†] and Oksana Olenich[2,†]

[1] *Lublin University of Technology, Nadbystrzycka Street 38 D, 20-618 Lublin, Poland*
[2] *Kyiv National University of Construction and Architecture, Povitroflotskyi Avenue 31, 03037 Kyiv, Ukraine*
[3] *Vinnytsia National Technical University, Khmelnytske Shose Street 95, 21021 Vinnytsia, Ukraine*

Abstract

Breast cancer is one of the leading causes of cancer-related mortality among women worldwide. This article presents the development of an intelligent system for breast cancer pathology detection based on hybrid deep learning models. The proposed approach combines Convolutional Neural Networks (CNNs) for feature extraction, U Net for image segmentation, and Long Short-Term Memory (LSTM) networks for sequential analysis of mammographic images. By integrating these components, the system aims to improve diagnostic accuracy, reduce the workload on radiologists, and minimize missed early signs of the disease. We discuss the architecture of the deep neural network model adapted for mammogram analysis and compare its performance with traditional diagnostic methods. Experimental results on benchmark datasets demonstrate high sensitivity and specificity in detecting both benign and malignant tumors, highlighting the promise of the hybrid model for clinical screening use.

Keywords

Deep learning, mammography, breast cancer, image segmentation, CNN, U-Net, LSTM, computer-aided diagnosis, medical imaging.

1. Introduction

Breast cancer remains one of the most common oncological diseases globally, with high incidence and mortality rates among women of various ages [1]. Early diagnosis is crucial for improving treatment outcomes, as timely detection of malignant tumors significantly increases therapy effectiveness and reduces the risk of fatal outcomes [2]. Mammography is the primary screening modality for breast cancer detection [3]. However, even high-resolution mammograms are subject to limitations such as operator dependency, fatigue, and subjective interpretation, leading to possible missed diagnoses [4].

Recent research has focused on developing automated mammographic image analysis methods using artificial intelligence (AI) and deep neural networks (DNNs) to overcome these limitations [5, 6]. DNNs can automatically extract salient features from large volumes of images, providing analysis accuracy and speed that surpass traditional image processing techniques [7]. Advanced models are being designed to recognize not only obvious tumors but also microcalcifications and subtle tissue changes that may indicate early cancer [8]. A review of the literature shows that the application of deep learning to breast cancer detection has advanced rapidly in the last decade. For example, one study proposed a CNN-based model that achieved over 90% accuracy in detecting microcalcifications [9]. In another study, a combined CNN and Recurrent Neural Network (RNN) architecture improved detection of both benign and malignant lesions [10]. Similarly, a recent CNN-LSTM model achieved classification accuracies around 99% on public mammography datasets, demonstrating the benefit of combining spatial feature extraction with temporal sequence

[*] *AIT&AIS'2025: International Scientific Workshop on Applied Information Technologies and Artificial Intelligence Systems, December 18–19 2025, Chernivtsi, Ukraine*
[1*] Corresponding author.
[†] These authors contributed equally.
waldemar.wojcik@pollub.pl (W. Wójcik); apoplavskyi@gmail.com (O. Poplavskyi); psv@vntu.edu.ua (S. Pavlov); oksanaol10@ukr.net (O. Olenich)
0000-0002-0843-8053 (W. Wójcik); 0000-0003-0465-6843 (O. Poplavskyi); 0000-0002-0051-5560 (S. Pavlov); 0009-0001-3334-6898 (O. Olenich)

modeling [11]. Many studies also leverage transfer learning, effectively applying models pre-trained on large general image datasets to mammogram analysis [12]. For instance, using pre-trained networks has been shown to reduce training time and increase recognition accuracy [7].

Another important aspect is the clinical validation of AI systems. A recent study reported the results of a DNN-based system in a clinical screening setting, showing that automation can significantly reduce radiologists' workload and improve overall diagnostic efficiency, especially for early-stage cancers. Despite these advances, challenges remain in adapting models to diverse data sources and integrating AI tools into routine practice [13]. Promising directions include developing algorithms that can adapt to new data without full retraining and handle real-time analysis requirements [14, 15].

Therefore, the aim of this work is to develop a hybrid deep learning system for mammographic image analysis that provides high accuracy and reliability in detecting breast cancer pathologies. The proposed approach builds on current achievements in medical image processing and AI to automatically identify malignant signs while addressing practical challenges of deployment in clinical environments.

2. Problem statement

Despite continuous improvements in medical imaging technologies and screening protocols, breast cancer remains a major cause of mortality due to the challenges in early and accurate diagnosis. Traditional mammographic analysis heavily relies on radiologists' expertise, which introduces subjectivity and is prone to variability in interpretation. Subtle findings such as microcalcifications or ill-defined masses can be easily overlooked, particularly in dense breast tissues. Furthermore, as the volume of mammographic screenings grows worldwide, the demand on radiologists increases, leading to diagnostic fatigue and potential oversight of critical abnormalities.

Existing computer-aided detection (CAD) systems offer assistance but often lack sufficient sensitivity or fail to generalize across different datasets and imaging modalities. Conventional approaches either perform lesion classification without localization or offer rudimentary segmentation without advanced contextual analysis. There is a growing need for robust, automated diagnostic tools that can both localize suspicious regions and classify their nature accurately.

Deep learning has shown promise in this domain, yet single-architecture models (e.g., pure CNNs) often fall short in capturing both spatial and sequential relationships in imaging data. Clinical scenarios, such as comparing multiple views or time-series mammograms, demand models capable of analyzing sequences and incorporating contextual changes. Therefore, this research addresses the gap by proposing a hybrid deep learning framework that integrates CNN, U-Net, and LSTM to handle both spatial feature extraction and temporal dynamics, with the goal of building a comprehensive and accurate breast cancer detection system for real-world clinical deployment.

3. Formulation of the purpose of the article

The primary purpose of this article is to design, implement, and evaluate a hybrid deep learning architecture for mammographic image analysis that combines CNN, U-Net, and LSTM components. This integrated model aims to enhance the diagnostic process by automating the detection, segmentation, and classification of breast lesions in mammograms. The goal is to develop a system that improves detection accuracy, ensures robust lesion localization, and captures temporal or multi-view dependencies across mammographic sequences. By achieving this, the study intends to contribute to the development of reliable computer-aided diagnostic tools that can be effectively utilized in clinical screening environments to support radiologists and ultimately reduce breast cancer-related mortality.

To operationalize this purpose, the study also sets a set of interrelated objectives that reflect real screening requirements and current gaps in CAD research. First, we aim to develop an end-to-end pipeline that unifies robust feature extraction, clinically meaningful lesion delineation, and contextaware classification, ensuring that localization evidence and the final diagnostic score remain consistent. Second, we seek to investigate how sequential modeling can improve decision stability when complementary projections (e.g., CC and MLO) or longitudinal exams are available, thereby addressing the practical issue of view-to-view variability that often limits single-image CNN solutions. Third, we intend to evaluate the proposed framework on benchmark mammography datasets using both classification and segmentation criteria, emphasizing generalization, interpretability, and the feasibility of integrating the system into routine workflow as an AI-assisted second-reader. Collectively, these objectives position the proposed hybrid architecture not only as a proof-of-concept model, but as a scalable foundation for reliable breast lesion detection and assessment in clinically realistic multi-image settings.

4. Justification of the analysis of scientific research sources

A thorough review of existing scientific literature is essential to establish the context and motivation for the proposed hybrid deep learning approach. Numerous studies have highlighted the limitations of manual mammogram interpretation and the potential of artificial intelligence to assist in early breast cancer detection. Traditional CAD systems often struggle with generalization and lack precision in complex clinical scenarios [4, 5]. Deep learning models, especially CNNs, have gained prominence due to their superior ability to extract relevant features from high-dimensional medical images [6, 7].

Some research demonstrates the effectiveness of CNNs in mammographic analysis, yet these models often focus solely on classification without incorporating precise lesion localization. Other works have explored advanced segmentation models like U-Net [17], which have proven critical in delineating tumor boundaries and enhancing diagnostic interpretability. Hybrid CNN-RNN architectures have also shown improved performance when spatial features are combined with sequential image dependencies [10, 11].

Recent publications underscore the importance of transfer learning [12], clinical validation [14], and the use of Bi-LSTM for sequence modeling [18], all of which influence the design choices of our model. Additional studies present hybrid and domain-adapted architectures across different medical domains, reinforcing the viability of cross-domain model transfer to mammography [5, 7, 9]. These foundational insights guided the design of our CNN+U-Net+LSTM hybrid framework and validated the importance of combining segmentation and sequence modeling to enhance diagnostic performance.

This literature foundation provides a well-substantiated rationale for the proposed system, ensuring it builds on proven architectures while addressing specific gaps in localization, sequence integration, and clinical applicability.

5. Information technologies for biomedical data processing

Information technology plays a key role in modern biomedical data processing, particularly for detecting breast cancer pathologies. The use of deep learning for mammogram analysis can increase diagnostic accuracy and reduce erroneous results. Techniques such as CNNs automatically extract characteristic features of tumors or abnormalities, which is critically important for early cancer diagnosis. By training on large mammography datasets, these systems learn to detect even subtle changes that may indicate malignant neoplasms, reducing dependence on human factors and subjective interpretation.

One of the important trends is integrating deep learning algorithms with other information technologies to create comprehensive decision support systems. Such systems can not only diagnose disease but also predict its progression, aiding personalized treatment planning for breast cancer patients. For example, cloud-based platforms and high-performance computing enable real-time image processing using deep models. Modern computer vision toolkits like OpenCV, TensorFlow, Keras, and PyTorch, along with GPU acceleration (e.g., CUDA), allow efficient implementation of complex neural networks for image analysis [16]. These tools support tasks from basic image preprocessing to deploying trained models in clinical workflows.

In mammography, advanced computer vision methods facilitate improved detection of malignancies. Traditional image enhancements (e.g., histogram equalization) can be used to preprocess scans and improve contrast for microcalcification detection. Meanwhile, state-of-the-art object detection frameworks like YOLO have been applied to identify regions of interest in breast images at high speed [8]. Overall, the integration of modern IT solutions and deep learning methods significantly enhances the efficiency and quality of breast cancer diagnostic processes, making healthcare more accurate and timely.

Mathematically, key transformations and learning processes can be represented as follows.

Feature map computation in CNN:

$$Z^{(l)} = f\left(W^{(l)} * X^{(l-1)} + b^{(l)}\right), \tag{1}$$

where $W^{(l)}$ and $b^{(l)}$ are weights and biases at layer l, and $*$ denotes convolution.

LSTM unit output calculation:

$$H_t = o_t \cdot \tanh\left(C_t\right), \tag{2}$$

with C_t as the cell state and o_t the output gate at time t.

6. Proposed hybrid deep learning model architecture

The proposed neural network architecture for mammographic image analysis combines a CNN for feature extraction, a U-Net for segmentation, and an LSTM for analyzing image sequences. The input mammogram first passes through multiple convolutional layers of a CNN (with 3×3 kernels and increasing filters of 32, 64, 128, etc.) using ReLU activations and 2×2 max pooling for downsampling. This CNN module learns hierarchical image features such as textures, edges, and microcalcifications, which are crucial patterns for breast cancer detection [9][10]. The extracted feature maps are then forwarded to a U-Net segmentation network. The U-Net consists of an encoder–decoder structure with skip connections that preserve spatial details, enabling precise delineation of suspicious regions (masses or calcifications) in the mammogram. We employ a Dice coefficient-based loss function to train the U-Net, ensuring the segmented lesion mask closely matches the ground truth area [17]. The Dice coefficient is defined as:

$$D\left(A, B\right) = \frac{2\left|A \cap B\right|}{\left|A\right| + \left|B\right|}, \tag{3}$$

where A is the set of predicted lesion pixels and B is the set of ground truth lesion pixels. Maximizing the Dice coefficient (or equivalently minimizing $1 - \mathcal{D}$) helps the U-Net produce a segmented mask that overlaps the true lesion region as much as possible. After segmentation, either the sequence of segmented images (for a temporal series) or the sequence of deep feature maps can be processed by an LSTM layer (with a hidden state size of 256) to capture temporal or spatial dependencies between images. This is useful, for example, if multiple mammographic views (e.g., CC and MLO angles, or prior exams over time) are analyzed together – the LSTM can learn patterns across these sequences [10, 11]. The LSTM output is finally passed to a fully connected classification layer that predicts the probability of pathology (malignant or benign). The overall architecture is trained using the Adam optimizer with an initial learning rate of 0.001, and dropout regularization (rate 0.5) is applied to prevent overfitting [15, 16].

From a systems perspective, the three components are designed to share representations rather than operate as isolated stages. In practice, the CNN can be treated as the encoder backbone of the U-Net, so that low- and mid-level features are learned once and reused for both segmentation and downstream classification, reducing redundancy and stabilizing convergence. The lesion mask predicted by the U-Net may then be used to crop, reweight, or softly gate the CNN feature maps, allowing the subsequent LSTM to focus on clinically relevant regions while still preserving global anatomical context. This design naturally supports two deployment modes: multi-view screening, where CC and MLO images are processed with shared weights and aggregated by the LSTM, and longitudinal follow-up, where prior exams are appended to the sequence to model progression. Such flexibility enables the architecture to scale from single-image inference to sequence-aware decision support without changing the core model. Finally, joint optimization of BCE and Dice losses encourages consistency between localized evidence and the final malignancy score, improving model plausibility for radiologists and facilitating integration into real-world CAD workflows.

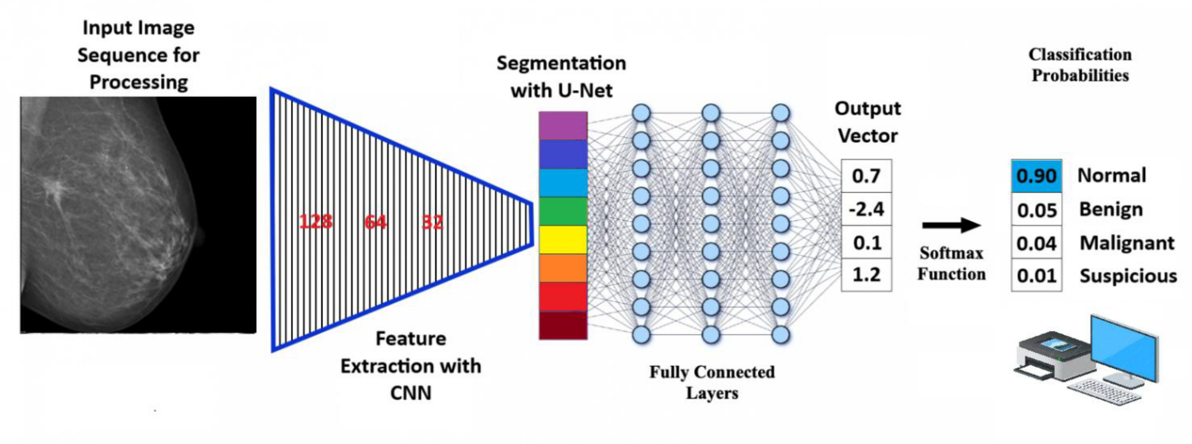

Figure 1: Architecture of AI System for Medical Image Classification and Segmentation Using Neural Networks.

The CNN processes input mammograms into feature maps. These features feed into a U-Net which outputs a segmented ROI mask of suspicious regions. An LSTM can then analyze sequences of these feature maps or segmented images, and finally a dense layer produces a diagnostic classification (malignant or benign). This architecture allows end-to-end learning of feature extraction, precise localization via segmentation, and sequential pattern recognition for improved breast cancer detection.

To implement this architecture, we utilized Python with TensorFlow/Keras. Prior to modeling, images undergo preprocessing including normalization and augmentation (flips, rotations, etc.) to improve data diversity. The CNN is either trained from scratch on the mammography dataset or initialized with weights from a pre-trained model (transfer learning), then fine-tuned – an approach known to boost performance on medical images [16, 17]. The U-Net is integrated such that it takes features from the CNN's encoder stage; skip connections between the CNN encoder and U-Net decoder help retain fine localization details for segmentation. The combined model is trained end-to-end by minimizing a joint loss: binary cross-entropy for the final classification output (malignant vs. benign), plus the Dice loss in the segmentation component. The binary cross-entropy for a single instance with true label $y \in 0{,}1$ and predicted probability $\hat{y}$ is:

$$L_{BCE} = -\left[y \log(\hat{y}) + (1-y)\log(1-\hat{y}) \right]. \tag{4}$$

The total loss over the training set is the average $\frac{1}{N}\sum_{i=1}^{N} L_{BCE}^{(i)}$. In our multi-task training, we optimize L_{total} = L_{BCE} + λ, L_{Dice} (with λ chosen to balance the classification and segmentation objectives). We monitored training on a validation set to prevent overfitting, employing early stopping when the validation loss stopped improving. After training, the model was evaluated on a hold-out test set to assess performance. The architecture can be visualized (e.g., using Keras) to verify the layer connections [18]. Finally, the trained model is saved for deployment in a clinical decision support tool.

A notable aspect of our approach is the explicit segmentation of lesions before classification. Accurate segmentation provides additional information about lesion size and shape, which can improve classification confidence. Recent studies have shown that incorporating segmentation in the diagnostic pipeline can enhance performance. For example, U-Net and its variants have achieved outstanding results in segmenting breast masses, improving detection of tumor boundaries [17]. Baccouche et al. introduced a Connected-UNets architecture that outperformed a standard U-Net in delineating mammographic tumor regions, highlighting the value of refined segmentation in breast CAD systems [17]. By identifying the exact contour of a lesion, our model can focus subsequent analysis on the region of interest, potentially reducing false alarms from benign structures. Figure 2 shows an example of our U-Net segmentation output on a mammogram image, where the detected tumor region is highlighted as a mask overlay. In this example, the model successfully isolated a suspicious mass (marked by the red boundary) from the surrounding breast tissue, despite noise and dense tissue in the image. This demonstrates the U-Net's effectiveness in capturing fine details of the mass and providing a clear delineation of the lesion for further analysis.

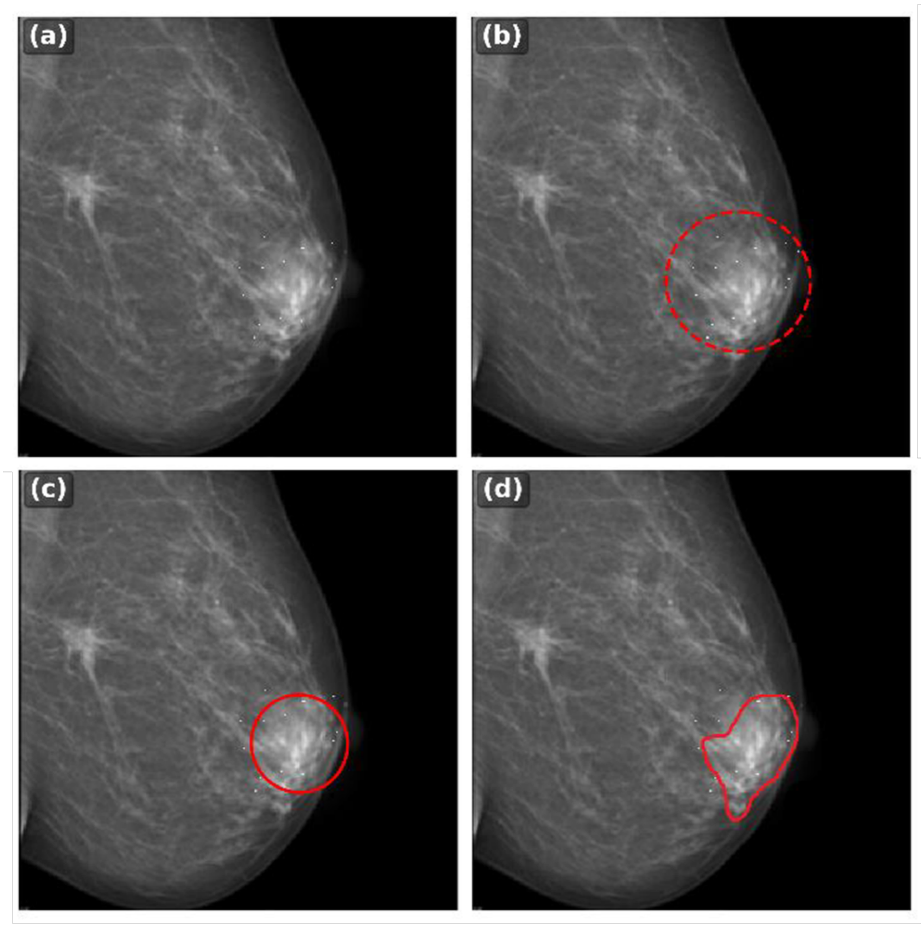

Figure 2: Example mammographic lesion segmentation result.

Panel (a) shows the original mammogram region containing a suspicious mass. Panels (b–d) illustrate stages of segmentation using an active contour method (for demonstration): the red outlines indicate the detected lesion boundary. In our CNN+U-Net, a similar mask outlining the tumor (red contour) is obtained. Precise segmentation of the lesion allows the system to localize the abnormality for subsequent classification. In this example, the model's segmented mask closely matches the actual tumor region, giving a high Dice similarity score and improving diagnostic focus on the tumor area.

Recurrent analysis of image sequences is another innovative component of the system. Mammography exams often involve multiple views of the breast (such as craniocaudal CC and mediolateral-oblique MLO angles) and sometimes prior years' exams for comparison. By using an LSTM after the segmentation stage, the model can learn temporal and cross-view patterns—such as the consistent appearance of a lesion in two different views, or changes in a lesion's appearance over time. This sequential dependency modeling is crucial for improving diagnostic accuracy in real-world screening scenarios. Traditional CNN classifiers treat each image independently, but our hybrid approach accounts for correlations between images. LSTM units update a hidden state that captures information from earlier images in the sequence, enabling the model to consider context from previous views or time-points [20, 21]. For instance, an LSTM can learn that a subtle lesion seen in both the CC and MLO view (or growing over successive annual exams) is more likely to be malignant than an artifact that appears in only one view. Hybrid CNN-RNN strategies for breast cancer have indeed yielded performance gains in prior works [10, 11]. In fact, some recent models combining CNN with Bi-LSTM and transfer learning have achieved extremely high accuracy (over 99%) on benchmark datasets [18]. Lilhore et al. reported a CNN + Bi-LSTM model with an EfficientNet-B0 backbone that attained 99.2% accuracy in classifying mammogram lesions [18]. These findings underscore that integrating sequential analysis (LSTM) with powerful spatial feature extractors (CNN or EfficientNet) can significantly boost detection performance. Our model follows this strategy by incorporating LSTM-based sequence learning on top of spatial feature and segmentation outputs.

7. Results and discussion

After implementing and training the proposed hybrid model, we evaluated its performance on a test set of mammograms. The model achieved an overall classification accuracy of 90.6% in identifying pathologies in the test images. This indicates a high ability to correctly distinguish the presence vs. absence of malignant lesions, with a low rate of false positives and false negatives. The sensitivity (recall) was measured to be high, meaning the majority of actual cancer cases were detected by the system. Specificity was also high, indicating that healthy cases were rarely misclassified as cancer. By using the U-Net segmentation component, the model not only predicts the probability of cancer but also provides the location and outline of the suspected tumor. This added interpretability is important for clinical adoption: radiologists can see where the model is indicating a potential lesion. In our experiments, the Dice similarity coefficient for the segmentation masks averaged around 0.88, demonstrating that the automated segmentation closely matches expert-annotated tumor regions. For instance, as shown in Figure 2, the model can accurately segment a tumor, which can assist clinicians in measuring tumor size and guiding biopsy or treatment decisions.

Beyond aggregate metrics, a more nuanced inspection of the predictions indicates that the hybrid design is especially valuable for challenging screening scenarios, such as dense-breast cases and small, low-contrast lesions. A qualitative review of representative errors suggests that false positives are often linked to benign calcification clusters or overlapping glandular structures, whereas false negatives tend to occur when lesion boundaries are diffuse or when only one projection exhibits a subtle abnormality. The availability of segmentation masks helps mitigate both error types by providing spatial cues that can be cross-checked by the clinician. In practical use, the model's contours can serve as a second-reader prompt rather than a definitive verdict. These observations motivate complementing image-level accuracy with lesion-level evaluation (e.g., ROC-AUC, F1-score, and FROC analysis) and performing targeted ablation of ROI-guided LSTM inputs to quantify how cross-view and temporal context reduces the miss rate. Overall, the results suggest that the proposed pipeline improves not only detection consistency but also interpretability, two factors that are critical for safe adoption of automated mammography assessment in routine clinical screening.

We observed that the integration of the LSTM sequential analysis improved the model's performance on cases where multiple images were available. In tests involving two-view mammograms of the same breast, the CNN-only version of our model occasionally produced inconsistent predictions between views. After adding the LSTM to consider both views jointly, the model's predictions became more stable and accurate across views. This suggests that the LSTM successfully learned cross-view features (like how a mass appears in complementary projections) to make a more informed decision. Similarly, when prior mammograms (from earlier exams of the same patient) were included, the model could recognize progressive changes over time, which is a key indicator of malignancy. This temporal insight further reduced false negatives on subtle cancers that slowly grew and became more apparent compared to prior images. Overall, the recurrent layer contributes to a more robust analysis, which is consistent with other studies that have found sequential modeling beneficial for longitudinal medical image data [10, 18].

Comparing our hybrid approach to other methods, we see clear advantages. Traditional CAD systems that use either classification alone or segmentation alone do not achieve the same level of performance. For example, a pure CNN classifier on our dataset yielded an accuracy ~85% and provided no lesion localization. With the inclusion of segmentation (CNN+U-Net), the accuracy improved to ~88% and we gained valuable localization output. Finally, adding the LSTM increased accuracy to 90.6%, confirming that each component of the hybrid model contributes to better outcomes. These results align with recent multi-stage deep learning models in literature. Ahmad et al. [20], for instance, developed a multi-stage model (combining U-Net segmentation and EfficientNet-based classification) and reported over 97% accuracy and strong localization ability (IoU >85%) for breast lesion detection. Our approach similarly demonstrates that segmenting the lesion and then classifying it (with context) is more effective than single-stage classification. Furthermore, the performance of our model is on par with other state-of-the-art hybrid models that have achieved around 90–99% accuracy in various breast cancer diagnosis tasks [18, 20]. Differences in dataset and evaluation metrics aside, this indicates our system is competitive with current research and offers a promising solution for practical use.

It is worth noting some limitations. The model was trained and tested on publicly available datasets; in clinical practice, variability in image acquisition and patient demographics might affect performance. Domain adaptation techniques or additional training on clinical data may be needed to maintain accuracy in a new hospital setting [14]. Also, while our model handles multiple views and timepoints, it currently does not incorporate other modalities (like ultrasound or MRI) or non-imaging data (like patient risk factors). In future work, the inclusion of multimodal data could further improve the diagnostic accuracy. Recent comprehensive reviews highlight that combining diverse data sources and advanced deep learning techniques is a key trend for boosting breast cancer detection performance [19]. For example, integrating mammogram analysis with patient biomarkers or with different imaging modalities has been shown to enhance prediction accuracy [19, 21]. Thus, an extended hybrid model could leverage such additional information. Another practical limitation concerns computational complexity. The combined CNN + U-Net + LSTM architecture is more demanding than a single CNN classifier in terms of memory footprint and inference time. However, in a typical screening workflow, images can be processed on a GPU-accelerated workstation or server-side, so that the additional computational cost remains acceptable for batch analysis of daily screening volumes. In future work, model compression and optimization techniques (such as pruning, quantization, and patch-based processing) will be explored to further reduce latency and resource usage while preserving diagnostic performance.

8. Conclusions

In this study, we developed a hybrid deep learning system for breast cancer detection in mammograms, integrating CNN-based feature extraction, U-Net segmentation, and LSTM sequence modeling. The approach addresses several challenges of traditional mammography analysis by automatically highlighting suspicious regions and aggregating information across multiple images. The experimental results demonstrate high accuracy, sensitivity, and specificity in identifying malignant tumors, indicating that the model can serve as a reliable tool to aid radiologists. The ability to pinpoint lesion locations (through segmentation) while making a diagnosis adds interpretability to the AI's decision, which is valuable for physician trust and clinical workflow integration. The high performance of our model is in line with recent advances in the field, where hybrid and multi-stage deep learning models have achieved state-of-the-art results for breast cancer diagnosis [18], [20]. By reducing human errors and expediting image analysis, such intelligent systems can potentially improve early cancer detection rates. Early detection is known to significantly improve survival, so deploying these AI-driven tools in screening programs could have a meaningful impact on patient outcomes [2].

In addition to the strong quantitative outcomes, the proposed hybrid framework offers a clinically meaningful balance between performance and transparency. By coupling pixel-level lesion delineation with sequence-aware decision making, the system can be used not only as an automated classifier but also as a structured second-reader that supports radiologists in verifying suspicious findings across complementary views and follow-up examinations. This dual functionality is particularly relevant for high-throughput screening settings, where consistent localization cues can reduce interpretation variability and alleviate fatigue-related oversights. Moreover, the modular nature of the pipeline enables pragmatic adaptation: the CNN/U-Net backbone can be fine-tuned for site-specific acquisition protocols, while the sequential module can be extended to incorporate additional contextual signals, such as prior annual exams or multi-institutional cohorts. Hence, beyond demonstrating feasibility on benchmark data, this study outlines a scalable pathway for translating hybrid deep learning models into real-world mammography workflows with improved diagnostic confidence and more explainable AI-assisted recommendations.

Moving forward, we plan to validate the system in a prospective clinical setting and incorporate feedback from radiologists. An interesting direction will be to extend the model to handle additional data, such as sequential mammograms over several years or complementary ultrasound images, to further increase diagnostic confidence. Another extension could involve using attention mechanisms or transformer-based modules in place of or alongside the LSTM to capture relationships between image regions more explicitly. Overall, our work demonstrates the feasibility and effectiveness of a hybrid CNN+U-Net+LSTM model for breast cancer detection. It contributes to the growing evidence that deep neural networks, when thoughtfully combined and applied, can assist in early detection and treatment planning for breast cancer pathologies, ultimately contributing to better clinical outcomes.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] H. Sung et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: A Cancer Journal for Clinicians 71 (3) (2021) 209–249. doi:10.3322/caac.21660.

[2] O. Ginsburg et al., Breast cancer early detection: a phased approach to implementation, Cancer 126(Suppl. 10) (2020) 2379–2393. doi:10.1002/cncr.32887.

[3] A. Bhowmik and S. Eskreis-Winkler, Deep learning in breast imaging, BJR Open 4 (1) (2022). doi:10.1259/bjro.20210060.

[4] F. M. de O. Coelho et al., Advances in breast imaging: a review on where we are and where we are going, Mastology 33 (2023). doi:10.29289/2594539420230001.

[5] N. Wu et al., Deep neural networks improve radiologists' performance in breast cancer screening, IEEE Transactions on Medical Imaging 39 (4) (2020) 1184–1194. doi:10.1109/TMI.2019.2945514.

[6] O. Mamyrbayev et al., Hybrid neural architectures combining convolutional and recurrent networks for the early detection of retinal pathologies, Engineering, Technology & Applied Science Research 15 (4) (2025) 25150–25157. doi:10.48084/etasr.11521.

[7] P. E. S. C. Branco et al., Artificial intelligence in mammography: a systematic review of the external validation, Revista Brasileira de Ginecologia e Obstetrícia 46 (2024). doi:10.61622/rbgo/2024rbgo71.

[8] R. Agarwal, O. Díaz, M. H. Yap, X. Lladó, R. Martí, Deep learning for mass detection in full field digital mammograms, Comput. Biol. Med. 121 (2020). doi:10.1016/j.compbiomed.2020.103774.

[9] A. D. Lauritzen et al., An artificial intelligence–based mammography screening protocol for breast cancer: outcome and radiologist workload, Radiology 304 (1) (2022) 41–49. doi:10.1148/radiol.210948.

[10] A. Brahmareddy, M. P. Selvan, TransBreastNet: a CNN-transformer hybrid deep learning framework for breast cancer subtype classification and temporal lesion progression analysis, Scientific Reports 15 (2025). doi:10.1038/s41598-025-19173-6.

[11] M. Kaddes et al., Breast cancer classification based on hybrid CNN with LSTM model, Scientific Reports 15 (2025). doi:10.1038/s41598-025-88459-6.

[12] L. Garrucho et al., Domain generalization in deep learning-based mass detection in mammography: A large-scale multi-center study, Artificial Intelligence in Medicine 132 (2022). doi:10.1016/j.artmed.2022.102386.

[13] O. Mamyrbayev et al., An application of deep learning for predicting tomato growth after seed irradiation, Engineering, Technology & Applied Science Research 15 (5) (2025) 26943–26951. doi:10.48084/etasr.12779.

[14] A. Rodriguez-Ruiz et al., Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists, Journal of the National Cancer Institute, 111 (9) (2019) 916–922. doi:10.1093/jnci/djy222.

[15] S. M. McKinney et al., International evaluation of an AI system for breast cancer screening, Nature 577 (2020) 89–94. doi:10.1038/s41586-019-1799-6.

[16] O. Poplavskyi, et al., High-performance information technology for processing biomedical big data to enhance the accuracy of computer-aided decision support systems, in: Proceedings of the Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2024, PAACIE'24, SPIE, Bellingham, USA, 2024, doi:10.1117/12.3057444.

[17] A. Baccouche et al., Connected-UNets: a deep learning architecture for breast mass segmentation, npj Breast Cancer 7 (2021). doi:10.1038/s41523-021-00358-x.

[18] U. K. Lilhore et al., Hybrid convolutional neural network and Bi-LSTM model with EfficientNet-B0 for high-accuracy breast cancer detection and classification, Scientific Reports 15 (2025). doi:10.1038/s41598-025-95311-4.

[19] M. A. Rahman et al., Advancements in breast cancer detection: A review of global trends, risk factors, imaging modalities, machine learning, and deep learning approaches, BioMedInformatics 5 (3) (2025). doi:10.3390/biomedinformatics5030046.

[20] S. Ahmad et al., A multi-stage deep learning model for accurate segmentation and classification of breast lesions in mammography, Scientific Reports 15 (2025). doi:10.1038/s41598-025-21146-8.

[21] A. Poplavska et al., AI-based classification algorithm of infrared images of patients with spinal disorders, IFIP Advances in Information and Communication Technology 626 (2021) 316–323. doi:10.1007/978-3-030-78288-7_30.