Comparative deep learning CNN architectures for breast cancer detection from thermal imaging

Md. Sumon Hosen, Mustafizur Rahman, Zaid Bin Sajid, Md Naeem Hossan, Apu Biswas, Md. Mijanur Rahman
Department of Computer Science and Engineering, Southeast University, Dhaka, Bangladesh

Article Info

Article history:

Received Dec 9, 2025
Revised Mar 24, 2026
Accepted May 26, 2026

Keywords:

Breast cancer diagnosis
Convolutional neural networks
Deep learning
Medical image analysis
Thermography

ABSTRACT

It has been observed that breast cancer is a severe disease among women globally. Mammography is the most effective screening method for detecting this severe illness. Over the last thirty years, mammography has been widely recognized as a preventive measure against breast cancer. In recent years, convolutional neural networks (CNN) and artificial intelligence (AI) have become more common in digital mammography for automated breast cancer detection. For classifying breast cancer, this study examines the five CNN models: LeNet-5, AlexNet, VGG-16, ResNet-50, and Inception-v3, using the database for mastology research with infrared images (DMR-IR) dataset's thermal image. These models were trained and validated using accuracy, recall, F1-score, specificity, and AUC as evaluation criteria after the dataset was preprocessed using normalization and data augmentation. Among the experimental models, Inception-v3 achieved 99.44% accuracy, outperforming other CNNs by 1–2%, while other models performed with accuracy levels above 97%. These results show the tremendous efficacy of CNN-based deep learning methods for breast thermogram analysis. The research points out thermography as a useful support for traditional imaging and InceptionV3 as a potential option for correctly detecting clinical breast cancer.

Corresponding Author:

Md. Mijanur Rahman
Department of Computer Science and Engineering
Southeast University
Dhaka, Bangladesh
Email: mijanur.rahman@seu.edu.bd

1. INTRODUCTION

Breast cancer continues to be a significant health threat globally because of the uncontrolled cell growth in the mammary tissue that results in a malignant tumor [1]. It is considered as the most common form of cancer world-wide because of having an incidence rate of 12.5% among newly diagnosed cases. In the year of 2020, approximately 2.3 million women suffered from breast cancer, which tragically led to 6,85,000 deaths across the world [2], [3]. Moreover, by the end of 2020, during the last five years, shockingly 7.8 million women were identified as being affected by breast cancer, highlighting the far reaching and long-term effects of the disease [4]. To be more specific, breast cancer continues to be the second most fatal cancer globally among women, stressing the need for its initial detection and timely treatment [5].

Women with identifying early signs of breast cancer have a greater than 93% five-year survival rate, and this high probability of survival indicates that early detection and subsequent diagnosis facilitate positive intervention [6]. Different methods used in breast imaging for detecting early breast cancer consist of ultrasound, mammography, and thermography [7]. Mammography is extensively used, despite issues related to sensitivity and patient discomfort [8]. However, it is associated with other limitations. [9], especially when it comes to dense breast tissues [10] and worries about radiation exposure, particularly in younger women [11]. Also, mammograms may not be able to find lesions that are less than 2 mm, which has led to interest in other ways to screen for them, such as thermography. Thermography provides a radiation-free, non-contact, and non-invasive screening approach for detecting breast cancer [12], [13].

According to the idea that living beings produce infrared emissions, even at temperatures greater than absolute zero, these techniques use thermographic cameras to detect the radiation, and then it is converted into electrical signals, and the thermograms that are produced are examined to identify abnormal temperature patterns. A useful addition to the variety of screening options available, this quick and safe method enables early breast cancer detection without the use of ionizing radiation [7]. Lately, researchers have tried combining thermography and AI with deep learning that can be explained. It seems to make the results both more trustworthy and easier to interpret [14].

In recent years, machine learning (ML) and deep learning (DL) contributed greatly in detecting and assessing breast cancer. convolutional neural networks (CNNs) are among the most widely used techniques in the context of deep learning, and these have sparked considerable interest in the research community because they perform very well on different types of image analysis tasks [15], [16]. Several foundational models, specifically LeNet-5 [17], AlexNet [18], VGG-16 [19], ResNet-50 [20], and Inception-v3 [21], have demonstrated good results on different challenges. They are capable of extracting important features from complex medical images. However, with so many CNN architectures available today [22], it is not straightforward to identify which one is the most suitable for analyzing thermal breast images.

Furthermore, we discuss the field of breast cancer thermal image classification in order to determine which of five unique models, such as LeNet-5, AlexNet, VGG-16, ResNet-50, and Inception-v3, are considered highly advanced CNN models. This study concentrates on employing CNNs for the diagnosis of breast cancer based on thermal breast images collected from the database for mastology research with infrared images (DMR-IR) [23].

This extensive study yields an in-depth comparative analysis of the above-mentioned CNN architectures in the binary classification task of breast cancer diagnosis. The rest of the paper is structured as follows: Details on dataset and CNN model selection, along with the training and evaluation process are covered in Section II. The results are analyzed and given in section III, whereas a deep discussion is given in section IV. We summarize our work’s conclusion in Section V and present our observations and recommendations for future potential research.

Due to CNNs' outstanding performance, researchers have currently become more interested in incorporating them into research on clinical pics [16]. For example, a CNN-primarily based computer-aided diagnostic (CAD) device using thermal photos was proposed in a study with the aid of Zuluaga-Gomez *et al.* [3]. In comparison to earlier diagnostic techniques, their investigation has produced experimental records displaying CNNs' tremendous performance, dependability, and robustness.

Mahoro and Akhloufi [24] examined a novel two-step technique for identifying breast cancer. They used four different models, including DenseNet-201, EfficientNet-B7, VGG-16, and ResNet-50, to identify healthy, unwell, and unknown breast cancer. This study recorded the maximum accuracy 97.26%. The study, however, highlighted that unequal class distribution, a common problem, is a prevalent constraint in medical picture datasets and may have an impact on how the models perform in real-life scenarios. A framework titled "deep multiview breast cancer detection," based on deep transfer learning, is introduced by Tiwari *et al.* [25], which is a fully automated breast cancer diagnosing system. The VGG-16 model was used to arrange thermal images of breast as healthy or unhealthy, and both still and motion thermal image datasets with single-view and multi-view images were used in the study. The results of the comparison showed that VGG-16 was better than other transfer learning models, on dynamic breast thermal images with 99% test accuracy. The study conducted by Ekici and Jawzal [26] focused on the use of thermography, a non-invasive imaging technique, for automated breast cancer detection. They used CNNs optimized using the Bayes algorithm for classification after introducing a novel approach to extract key breast features from thermal pictures, and their method succeeded in finding signs of breast cancer with 98.95% accuracy, which is incredibly high.

Mammoottil *et al.* [27] used the Visual DMR dataset to construct ML models with CNNs, leveraging thermal imaging of the breast to diagnose breast cancer with 93.8% accuracy. Lazo *et al.* [28] employed DL models to identify breast cancer using medical imaging data. VGG-16 and Inception V3 were employed in this study with 947 images, where the highest accuracy was 91% by the VGG-16. Ridha *et al.* [29] suggested using LIME for interpretability, based on an explainable AI-based thermal image analysis framework that combines attention U-Net, K-Means clustering, and EfficientNet-B7. Their system performed effectively and offered clinical data easily accessible, which is a major advancement toward the safe use of AI in healthcare. 96.48% was its training accuracy, and 91.67% was its validation accuracy.

In consideration of the developments in AI for thermography-based breast cancer identification, optimization-based metaheuristics, a field of study that has drawn a lot of public interest, remains an important area for improvement. Moreover, it is now more important than ever to clarify the unique contributions of particular models to forecasts. This study focuses on the responsibility of addressing these crucial elements and aims to provide fresh paths for research by offering several CNN models.

2. METHOD

Figure 1, which depicts the process diagram for classifying breast cancer in thermal images, demonstrates the methodology. It not only uses multiple CNN models but also selects the most reliable model to guarantee the accuracy of the prediction. The process includes gathering datasets, preprocessing them, separating them, extracting features, developing models, training them, and comparing them. In this experiment, 70%, 15%, and 15% of the DMR-IR dataset were used for training, testing, and validation, respectively. A cloud-based platform called Google Colab, which allows GPU acceleration for deep learning applications, was used for all testing. A GPU-capable Python 3 Google Compute Engine backend was used for the runtime configuration. The system's specifications included 112.6 GB of available disk space, 15.0 GB of GPU memory, and 12.7 GB of RAM.

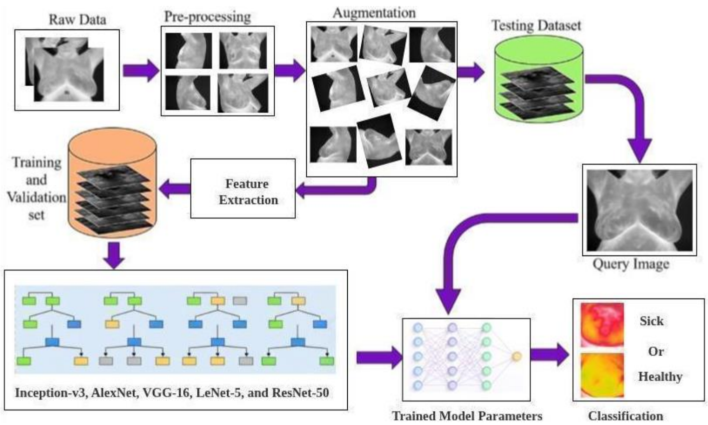

Figure 1. Overview of the methodology for breast cancer detection. This work proposes a technique that uses the Inception-v3, ResNet-50, VGG-16, LeNet-5, and AlexNet to detect diseases of the breast

2.1. Dataset overview

This study's dataset came from the DMR-IR database for mastology research at visual lab, UFF, Niteroi, Brazil [23]. As seen in Figure 2, the FLIR SC-620 thermal camera, which has a 480 × 640 pixel spatial resolution, was used to take these thermal images. The collection includes 3,580 photos from 415 people, 1,944 of whom are healthy and 1,636 of whom are unhealthy. Figure 3 shows the dataset's distribution for training, testing, and validation with 70%, 15%, and 15%.

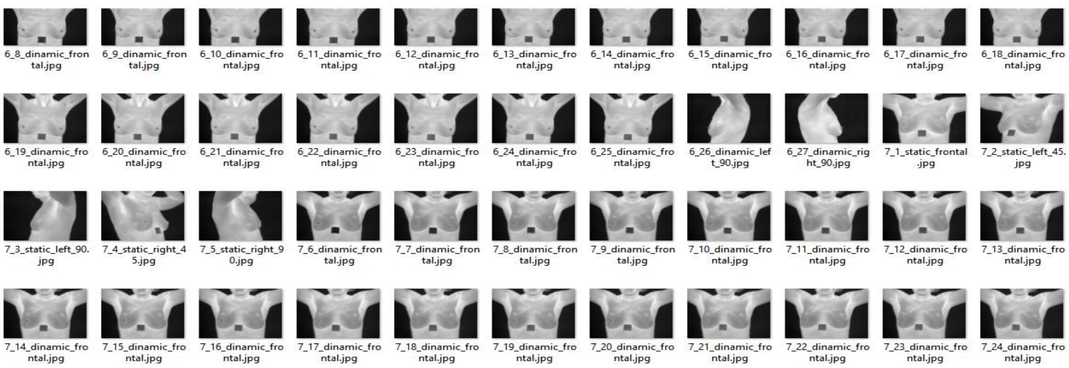

Figure 2. The breast thermal images from DMR-IR database

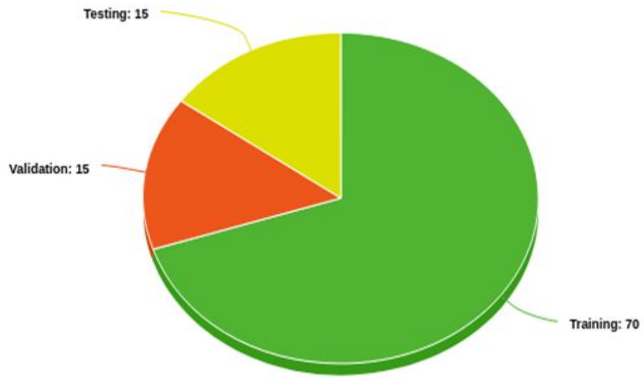

Figure 3. Dataset distribution percentages

2.2. Data preprocessing and augmentation

When applying the breast cancer diagnosis approach, image processing is crucial in the process of getting input images ready for further analysis. To enhance the quality and extract the relevant details, the raw images go through a number of processes. The main data preprocessing procedures and augmentation methods used in this investigation are given in the section that follows, and summary of the training hyperparameters is shown in Table 1.

Data preprocessing: data preprocessing is a crucial stage in order to evaluate breast thermograms and apply machine learning algorithms. Data need to be cleaned, transformed, and organized at this stage in order to increase the precision and efficiency of processes that come after. First, every image in the collection was resized to 299 × 299 pixels. Normalization was then performed to the range [0, 1] by rescaling each image by a factor of 1/255.

Data augmentation: in order to improve robustness and prevent overfitting in the training set, data augmentation was employed in this study. 20-degree random rotations, true horizontal flips, and a 20% width and height shift range were employed. In addition, we used shear transformations with a range of 0.2 and zoom operations of up to 20%.

Table 1. Concise table for training hyperparameters

Hyperparameter	Value
Input image size	299 × 299 × 3
Batch size	32
Epochs	10
Optimizer	Adam
Learning rate	0.001
Loss function	Binary cross-entropy
Activation	Relu, Sigmoid
Weight initialization	ImageNet
Data augmentation	Rotation, flip, shift, zoom, shear
Evaluation metrics	F1-score, specificity, recall, accuracy, precision, AUC

2.3. CNN models architectures

Five reliable CNN models were used in this experiment. Every model has distinct qualities and makes a substantial contribution to the fields of image analysis and deep learning. These models were chosen in order to learn more about how well each one performed on its own and how it handled the complexities of breast cancer detection.

LeNet-5: The LeNet-5 network is an illustration of a convolutional model. There are six convolutional feature maps in the first layer. Convolutional neurons in these maps have receptive fields that are five pixels high and five pixels wide. Due to the 32 × 32 pixel input size and the overlap of adjacent receptive fields. The initial convolutional layer employs six filters, reducing the size to 28x28x6. Subsequently, a subsampling layer further reduces the dimensions to 14x14x6. A second convolutional layer with 16 filters and fewer connections is added in the third layer. Another subsampling layer reduces it to 5x5x16. A fully connected convolutional layer with 120 units makes up the fifth layer, which is succeeded by another fully connected layer with 84 units. The architecture culminates in the final SoftMax output layer, facilitating classification.

AlexNet: With fully connected layers, max pooling, dropout, and 11 × 11 × 5 × 5 × 3 × 3 convolution, AlexNet is a pretrained network. 96 filters are used by the first convolution layer for reducing the image size to 55x55x96 with the rectified linear unit (ReLU) activation functions. The second layer includes local response normalization, and the third layer subsamples to make the size 27x27x96. In order to reach a final size of 13x13x256, the fourth and fifth layers introduce 256 filters with a 5x5 size in addition to subsampling.

VGG-16: In the CNN space, it is renowned for being straightforward and efficient. There are sixteen layers in all in this deep CNN architecture, with three fully linked levels and thirteen convolutional layers. The network maintains a consistent 3x3 filter size throughout the architectural instantiation, which starts with a 224x224 input picture. The number of filters increases gradually during the evolution; there are 64 filters in the first two convolutional layers and 128 filters in the next two. To lower spatial dimensions in a systematic manner, convolutional layers are separated by intervening max-pooling layers. An output layer for image classification, consisting of three fully connected layers, is the result of the architectural trajectory.

ResNet-50: It is part of the ResNet family developed by He et al., which uses a 50-layer deep convolutional neural network design with a focus on residual learning. A 224x224 input image is used to start a five-stage process that includes convolutional layers. Starting with a 7x7 convolutional layer, each stage consists of several convolutional blocks with residual connections, with the number of filters in succeeding blocks progressively rising. Global average pooling is used to reduce spatial dimensions before moving on to a fully linked layer with 1000 units that are specifically designed for classification.

Inception-v3: It works with images that are 299 by 299 pixels in size and is an extension of Google's Inception architecture for image recognition. This architecture consists of a 3x3 convolutional layer at the stem, followed by four Inception modules with various convolutional branches that include 1x1, 3x3, and 5x5 filters. Batch Normalization and ReLU activations are used consistently in these modules. This is followed by global average pooling, fully connected layers, and an output layer for classification.

2.4. Training and validation

15% of the dataset was used for testing, 15% for validation, and 70% for training. The training dataset was also used to train each of the chosen CNN architectures, and standard hyperparameters included using the Adam optimizer, 10 training epochs, a batch size of 32, and a learning rate of 0.001. The CNN architectures were implemented using the pretrained weights, namely ImageNet.

2.5. Performance evaluation

By employing several criteria, the effectiveness of the suggested methods in detecting breast cancer has been examined, which ensures a thorough evaluation. In order to assess the various aspects of model performance, the following metrics were taken into consideration: F1-score, specificity, recall, accuracy, precision, and the area under the receiver operating characteristic curve (AUC). Figure 4 shows that every measure came from the confusion matrix. The matrix has four parts, which are true negatives, false positives, and true positives, false negatives. By examining these indicators, the study demonstrates how alternative choices are made between the models and how various types of prediction errors are balanced.

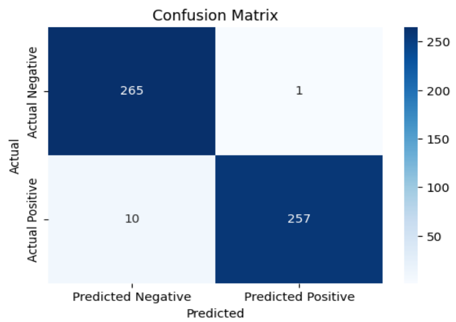

Figure 4. Example performance evaluation matrix

3. RESULT ANALYSIS

Accuracy, recall, F1-score, AUC, precision, specificity, and other important metrics were used in a thorough evaluation of the model's performance. Table 2 lists all of the algorithms' outcomes on the DMR-IR dataset. Figure 5 illustrates the comparative performance of all classifiers across the evaluation metrics.

Table 2. Performance evaluation of the proposed models

Algorithm Names	Accuracy	Precisions	Recall	F1-Score	AUC	Specificity
LeNet-5	98.31%	.99	.97	.98	.99	.99
AlexNet	99.06%	1.00	.95	.98	.98	1.00
VGG-16	98.69%	1.00	.97	.99	1.00	1.00
ResNet-50	97.94%	1	.96	.98	.98	1
Inception V3	99.44%	.99	1.00	.99	1.00	.99

ResNet-50 had the lowest accuracy of 97.94%, while Inception-V3 had the highest accuracy of 99.44%. Both AlexNet and VGG-16 demonstrated the resilience of identifying affirmative situations by achieving a flawless precision score of 1.00. LeNet-5 and Inception-V3 came in second and third, respectively, with precisions of 0.99. The maximum recall rate (100) was achieved by Inception-V3, while LeNet-5 and VGG-16 had somewhat lower recall rates (0.99). AlexNet's rate was the lowest at 0.96, followed by ResNet-50's. ResNet-50 and AlexNet had the lowest levels, at 0.96 and 0.96, respectively, while Inception V3 and VGG-16 received F1 scores of 0.99 and 0.98, respectively. However, Inception V3 also achieved the highest AUC score of 1.00.

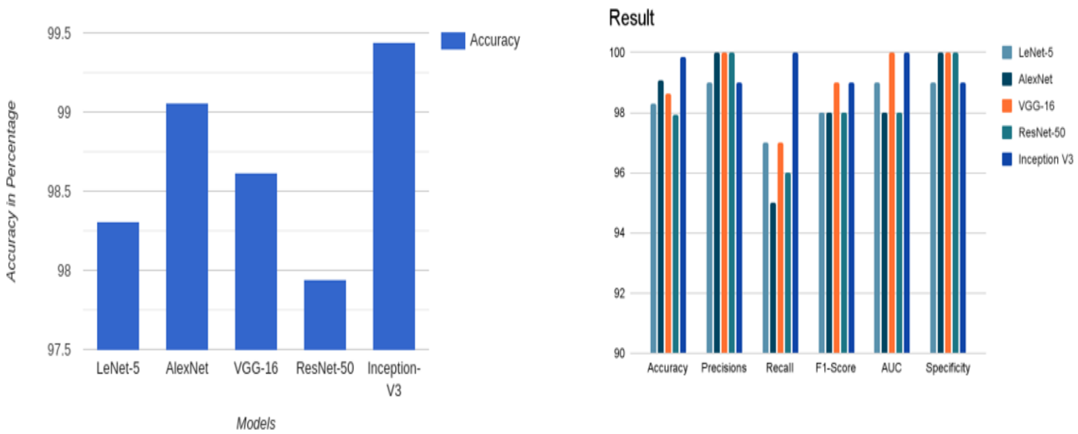

Figure 5. Performance of accuracy, precision, recall, F1-score, AUC, and Specificity using various classifiers

It is possible to evaluate the model's generalization and adaptability over time and obtain useful knowledge about its rate of convergence by comparing the curves of validation and training loss. Examining accuracy trends offers useful data for accessing the model performance by emphasizing its learning curve and predictive capacity which is shown in Figure 6. As shown in Figure 6(a), AlexNet achieved high training accuracy in the early epochs. Figure 6(b) demonstrates that Inception-V3 maintained relatively stable learning behavior. In Figure 6(c), LeNet-5 exhibited a noticeable gap between training and validation performance. Figure 6(d) shows that ResNet-50 achieved smooth convergence with continuously decreasing loss and increasing accuracy, Figure 6(e) illustrates that VGG-16 reached stable accuracy with a gradual reduction in loss.

The accuracy and loss curves show that different models have separate learning behaviors. The situations where validation accuracy is greater than training accuracy are connected to the use of regularization during training and are considered as signs of good generalization. On the other hand, overfitting is present in models that exhibit an increasing difference between training and validation performance. The best learning and successful generalization are represented by stable and convergent curves.

In this study, AlexNet, Inception-V3, VGG-16, ResNet-50, and LeNet-5’s performances were examined using confusion matrices that depict Figure 7. By providing information on the true positive, true negative, false positive, and false negative predictions for every category, the above matrices give a thorough evaluation of a model's classification performance. The matrix for Inception V3 shows its viability for working with image analysis and ability to handle a wide range of feature hierarchies. On the other hand, VGG-16 shows its multi-layer accuracy and its hierarchy for extracting features abilities. The effectiveness of ResNet50, which uses residual learning blocks, in reducing vanishing gradient issues especially in deeper networks, is demonstrated using a confusion matrix. Finally, a confusion matrix is used to examine LeNet 5's straightforward architecture and demonstrate how well it can communicate vital information in a structured way.

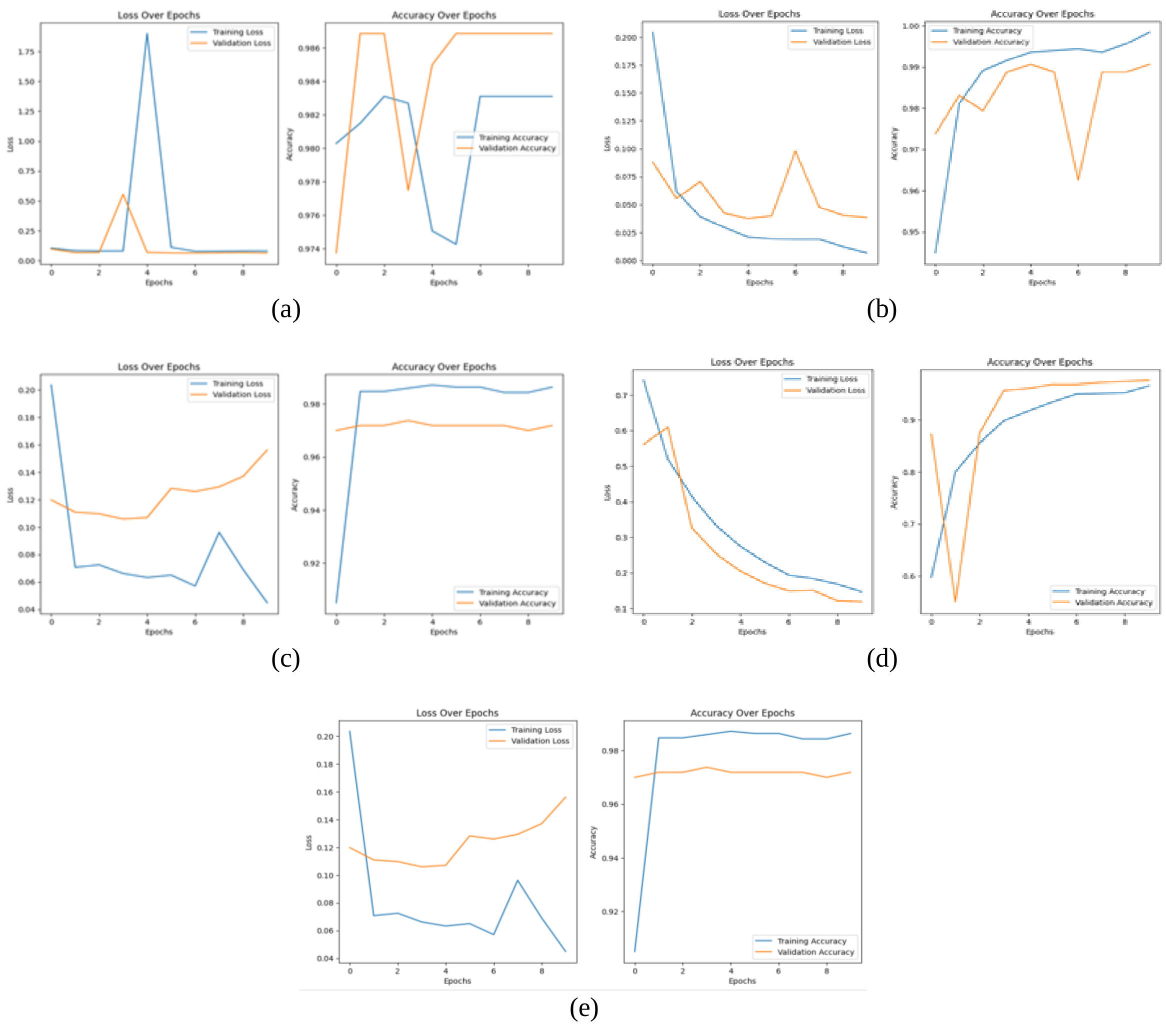

Figure 6. Evolution of loss and accuracy across epochs to (a) alexnet, (b) inception v3, (c) LeNet 5, (d) ResNet 50, and (e) VGG 16

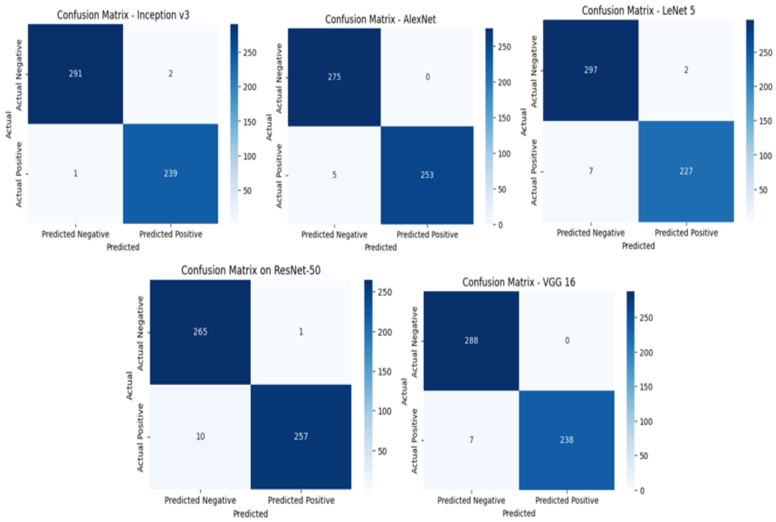

Figure 7. Confusion matrices of the five CNN models

4. DISCUSSION

The results shown in Table 3 demonstrate how different CNN architectures perform for detecting breast cancer based on thermal images with the DMR-IR dataset. In our paper, Inception-V3 achieved the highest accuracy with 99.44%, while AlexNet and LeNet-5 attained accuracy with 99.06% and 98.31%, respectively. With 98.62% and 97.94% accuracy, VGG-16 and ResNet-50 exhibit significant classifying behavior. On the other hand, some earlier studies analyze the same dataset by incorporating machine learning and deep learning models, where they achieve accuracies from 92.28% to 99%. For example, research that is conducted by D. Tiwari *et al.* [25] shows 99% accuracy with VGG-16, followed by 95%, 94%, and 89% with VGG-19, ResNet50V2, and InceptionV3. Nonetheless, our proposed Inception V3 is without a doubt the best performer in this regard, boasting an unmatched accuracy of 99.44%. Capturing features at multiple spatial scales, thermal images typically exhibit low contrast and smooth temperature variations. Due to multi-scale feature extraction ability, which enables parallel convolutional filters of various sizes to process the same input simultaneously, Inception-V3 shows greater accuracy. Additionally, training becomes more reliable as well as less at risk of overfitting when factorized convolution and additional classifiers are used, which improves its performance with new data. These findings highlight that optimal model selection for thermal imaging should be informed by image attributes rather than only by architectural intricacy. A careful look at the data in the table shows a complex view of how the models perform.

Table 3. Comparison analysis with several existing works

SI.NO	Researchers and years	Algorithm names	Dataset	Accuracy
1	Silva *et al.* [30] and 2020	K-Star classifier	DMR-IR dataset	98.57%
2	Raghavan and Sivaselvan [31] and 2024	Xception, InceptionV3, and ResNet101	DMR-IR dataset	95.4 %
3	Demirci *et al.* [32] and 2025	Support Vector Machine (SVM) and Random Forest (RF)	DMR-IR dataset	92.28%
4	D. Tiwari *et al.* [25] and 2021	VGG-16, VGG 19, ResNet50V2, and InceptionV3	DMR-IR dataset	99%, 95%, 94%, and 89%
5	Our Paper	LeNet-5, AlexNet, VGG-16, ResNet-50, and Inception V3	DMR-IR dataset	98.31%, 99.06%, 98.62%, 97.94%, and 99.44%

The multi-metric evaluation strategy is highlighted by the comprehensive discussion table. To the best of our knowledge, this is one of the first comparative studies of multiple CNN models on breast thermography using the DMR-IR dataset, which offers insightful information for thermal-based breast cancer detection. The detailed knowledge gained from accuracy, precision, recall, F1-score, AUC, and specificity all work together to provide a comprehensive view of all CNN architecture's effectiveness in the potential of thermal imaging as a non-invasive, cost effective, and radiation-free screening tool for early breast cancer detection.

5. CONCLUSION

Using thermal imaging, five robust CNN models such as LeNet-5, AlexNet, VGG-16, ResNet-50, and InceptionV3 were investigated for the early detection of breast cancer. Due to the complexity of analyzing medical images, specifically thermal ones, it became critical to thoroughly assess these structures in order to comprehend their computing efficiency and classification accuracy. The InceptionV3 displayed its robustness for image classification tasks by achieving the highest accuracy of 99.44% among all tested models. Because InceptionV3's multi-scale feature extraction effectively captures subtle temperature changes in thermal breast images, it performed extremely well. By incorporating robust CNN models, this research shows the comprehensive accuracy of breast cancer detection with the DMR-IR dataset and gives an insight for selecting CNN architectures for thermal image-based breast cancer detection. Overall, these findings offer valuable information about the behavior of various CNN models in thermal imaging and build foundations for further research into cancer diagnosis using breast thermography.

6. LIMITATIONS AND FUTURE WORK

Despite the outstanding result, there is still a question. This research was conducted on a specific number of thermal imaging dataset with a particular number of patients, which may restrict generalizability. In the future, authors want to integrate explainable AI for further clinical trials and focus on multicenter datasets because in this study they used a single DMR-IR dataset. Authors also want to deploy a user-friendly cross-platform mobile application that will help individuals to diagnose their ailments in low-resource settings.

FUNDING INFORMATION

The authors state that there was no funding for the development of the research.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Md. Sumon Hosen	✓	✓			✓				✓		✓		✓	
Mustafizur Rahman		✓	✓			✓	✓							
Zaid Bin Sajid				✓					✓		✓			
Md Naeem Hossan		✓	✓			✓	✓							
Apu Biswas			✓			✓	✓							
Md. Mijanur Rahman	✓			✓				✓		✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no financial, personal, or professional conflicts of interest that could have influenced the results or interpretations of this research.

DATA AVAILABILITY

The data that support the findings of this study are openly available in data at https://drive.google.com/drive/folders/1F_4QFQcx8-BdgYhPzl43BTAnmGUKhw3p.

REFERENCES

[1] Mayo Clinic, "Breast cancer – symptoms and causes," *Mayo Clinic*, Jun. 2022. Accessed: May. 3, 2026 [Online]. Available: https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470.

[2] S. Lei *et al.*, "Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020," *Cancer Communications*, vol. 41, no. 11, pp. 1183–1194, 2021, doi: 10.1002/cac2.12207.

[3] J. Zuluaga-Gomez, Z. Al Masry, K. Benaggoune, S. Meraghni, and N. Zerhouni, "A CNN-based methodology for breast cancer diagnosis using thermal images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, vol. 9, no. 2, pp. 131–145, 2021, doi: 10.1080/21681163.2020.1824685.

[4] World Health Organization, "Breast cancer," WHO. Accessed: May. 3, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[5] D. Deb, R. Dash, and D. P. Mohapatra, "Customized convolutional neural network with explainable AI for multimodal breast cancer detection," *Sadhana - Academy Proceedings in Engineering Sciences*, vol. 50, no. 3, p. 130, Jun. 2025, doi: 10.1007/s12046-025-02795-x.

[6] J. K. Birnbaum, C. Duggan, B. O. Anderson, and R. Etzioni, "Early detection and treatment strategies for breast cancer in low-income and upper middle-income countries: a modelling study," *The Lancet Global Health*, vol. 6, no. 8, pp. e885–e893, 2018, doi: 10.1016/S2214-109X(18)30257-2.

[7] A. Hossam, H. M. Harb, and H. M. Abd El Kader, "Automatic image segmentation method for breast cancer analysis using thermography," *JES. Journal of Engineering Sciences*, vol. 46, no. 1, pp. 12–32, 2018, doi: 10.21608/jesaun.2017.114377.

[8] R. M. Alzahrani *et al.*, "Early breast cancer detection via infrared thermography using a CNN enhanced with particle swarm optimization," *Scientific Reports*, vol. 15, no. 1, p. 25290, 2025, doi: 10.1038/s41598-025-11218-0.

[9] R. Jalloul, C. H. Krishnappa, V. I. Agughasi, and R. Alkhatib, "Enhancing early breast cancer detection with infrared thermography: a comparative evaluation of deep learning and machine learning models," *Technologies*, vol. 13, no. 1, p. 7, 2025, doi: 10.3390/technologies13010007.

[10] S. T. Kakileti, B. Govindaraju, S. Sampangi, and G. Manjunath, "A density-informed multimodal artificial intelligence framework for improving breast cancer detection across all breast densities," Oct. 2025, [Online]. Available: http://arxiv.org/abs/2510.14340.

[11] K. S. Kacafírková, A. Poll, A. Jacobs, A. Cardone, and J. J. Ventura, "Exploring women's perceptions of traditional mammography and the concept of ai-driven thermography to improve the breast cancer screening journey: mixed methods study," *JMIR Cancer*, vol. 11, p. e64954, 2025, doi: 10.2196/64954.

[12] M. B. Rakhunde, S. Gotarkar, and S. G. Choudhari, "Thermography as a breast cancer screening technique: a review article," *Cureus*. Nov. 08, 2022. doi: 10.7759/cureus.31251.

[13] J. Zuluaga-Gomez, N. Zerhouni, Z. Al Masry, C. Devalland, and C. Varnier, "A survey of breast cancer screening techniques: thermography and electrical impedance tomography," *Journal of Medical Engineering and Technology*, vol. 43, no. 5, pp. 305–322, 2019, doi: 10.1080/03091902.2019.1664672.

[14] H. Dihmani, A. Bousselham, and O. Bouattane, "A new computer-aided diagnosis system for breast cancer detection from thermograms using metaheuristic algorithms and explainable AI," *Algorithms*, vol. 17, no. 10, p. 462, Oct. 2024, doi: 10.3390/a17100462.

[15] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[16] Z. Zhu, S. H. Wang, and Y. D. Zhang, "A survey of convolutional neural network in breast cancer," *CMES - Computer Modeling in Engineering and Sciences*, vol. 136, no. 3, pp. 2127–2172, 2023, doi: 10.32604/cmes.2023.025484.

[17] P. Patro, S. H. Fathima, R. Harikishore, and A. K. Sahu, "Breast cancer image classification by using HCNN and LeNet5," *Discover Sustainability*, vol. 5, no. 1, p. 491, 2024, doi: 10.1007/s43621-024-00725-1.

[18] U. Ergün, T. Çoban, and İ. Kayadibi, "BCECNN: an explainable deep ensemble architecture for accurate diagnosis of breast cancer," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, p. 374, 2025, doi: 10.1186/s12911-025-03186-2.

[19] M. I. Daoud, S. Abdel-Rahman, T. M. Bdair, M. S. Al-Najar, F. H. Al-Hawari, and R. Alazrai, "Breast tumor classification in ultrasound images using combined deep and handcrafted features," *Sensors (Switzerland)*, vol. 20, no. 23, pp. 1–20, Nov. 2020, doi: 10.3390/s20236838.

[20] S. M. Ganie, M. B. Malik, M. Aadil, G. Muteeb, M. Farhan, and M. Aatif, "Explainable AI based hybrid DRM-Net transfer learning model for breast cancer detection and classification using ultrasound images," *Scientific Reports*, vol. 15, no. 1, p. 44170, 2025, doi: 10.1038/s41598-025-27934-6.

[21] P. K. Mondal, M. K. Jahan, and H. Byeon, "Mammogram mastery breast cancer image classification using an ensemble of deep learning with explainable artificial intelligence," *Medicine (United States)*, vol. 104, no. 22, p. e42242, 2025, doi: 10.1097/MD.0000000000042242.

[22] R. Zeng and Y. Zhang, "Comparative analysis on different convolutional neural network (CNN) for classification," in *2022 IEEE 5th International Conference on Information Systems and Computer Aided Education, ICISCAE 2022*, 2022, pp. 598–601. doi: 10.1109/ICISCAE55891.2022.9927565.

[23] L. F. Silva *et al.*, "A new database for breast research with infrared image," *Journal of Medical Imaging and Health Informatics*, vol. 4, no. 1, pp. 92–100, 2014, doi: 10.1166/jmihi.2014.1226.

[24] E. Mahoro and M. A. Akhloufi, "Breast cancer classification on thermograms using deep CNN and transformers," *Quantitative InfraRed Thermography Journal*, vol. 21, no. 1, pp. 30–49, 2024, doi: 10.1080/17686733.2022.2129135.

[25] D. Tiwari, M. Dixit, and K. Gupta, "Deep multi-view breast cancer detection: a multi-view concatenated infrared thermal images based breast cancer detection system using deep transfer learning," *Traitement du Signal*, vol. 38, no. 6, pp. 1699–1711, 2021, doi: 10.18280/ts.380613.

[26] S. Ekici and H. Jawzal, "Breast cancer diagnosis using thermography and convolutional neural networks," *Medical Hypotheses*, vol. 137, p. 109542, 2020, doi: 10.1016/j.mehy.2019.109542.

[27] M. J. Mammoottil, L. J. Kulangara, A. S. Cherian, P. Mohandas, K. Hasikin, and M. Mahmud, "Detection of breast cancer from five-view thermal images using convolutional neural networks," *Journal of Healthcare Engineering*, vol. 2022, no. 4295221, pp. 1–15, Feb. 2022, doi: 10.1155/2022/4295221.

[28] J. F. Lazo, S. Moccia, E. Frontoni, and E. de Momi, "Comparison of different CNNs for breast tumor classification from ultrasound images," *Convegno Nazionale di Bioingegneria*, pp. 560–563, Dec. 2020, [Online]. Available: http://arxiv.org/abs/2012.14517

[29] J. Ridha, K. Saddami, M. Riswan, and R. Roslidar, "An explainable artificial intelligence framework for breast cancer detection," *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, pp. 298–311, Apr. 2025, doi: 10.35882/ijeeemi.v7i2.78.

[30] L. Silva, F. Seixas, C. Fontes, D. Muchaluat-Saade, and A. Conci, "A computational method for breast abnormality detection using thermographs," in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, IEEE, Jul. 2020, pp. 469–474. doi: 10.1109/CBMS49503.2020.00095.

[31] K. Raghavan and B. Sivaselvan, "An explainable ensemble classifier for infrared breast cancer detection," in *2024 3rd International Conference on Artificial Intelligence for Internet of Things, AIIoT 2024*, IEEE, May 2024, pp. 1–6. doi: 10.1109/AIIoT58432.2024.10574722.

[32] B. A. Demirci, M. Engin, and E. Z. Engin, "Comparative analysis of autoencoder architectures for breast cancer detection using dynamic infrared thermography," *Engineering Science and Technology, an International Journal*, vol. 72, p. 102225, Dec. 2025, doi: 10.1016/j.jestch.2025.102225.

BIOGRAPHIES OF AUTHORS

Md. Sumon Hosen completed his Bachelor's degree in Computer Science and Engineering from Southeast University. His interests include Machine Learning, Deep Learning, Data Science, and NLP. He has also contributed to research through his recent publication in the field. He can be contacted at email: csesumon8@gmail.com.

Mustafizur Rahman holds a Bachelor's degree in Computer Science and Engineering from Southeast University. His interests include Data Science, Machine Learning, and Deep Learning. He can be contacted at email: afnanmustafizur@gmail.com.

Zaid Bin Sajid completed his Bachelor's degree in Computer Science and Engineering from Southeast University. His interests include Networking, IoT, Machine Learning, Deep Learning, NLP, and Blockchain. He has published research in the field of Blockchain and is currently working on multiple papers involving Machine Learning and Deep Learning. He can be contacted at email: zaidbinsajid1@gmail.com.

Md Naeem Hossan holds a Bachelor's degree in Computer Science and Engineering from Southeast University. He is interested in Machine Learning and Deep Learning and actively follows the latest advancements in these areas. He can be contacted at email: 2021000000026@seu.edu.bd.

Apu Biswas holds a Bachelor's degree in Computer Science and Engineering from Southeast University. He is interested in Data Science, Machine Learning, and Deep Learning, and stays engaged with emerging developments in these fields. He can be contacted at email: 2022000000088@seu.edu.bd.

Md. Mijanur Rahman is currently an Assistant Professor in the Department of Computer Science and Engineering and Director of the Center for Artificial Intelligence and Robotics at Southeast University. His research interests include deep learning, machine learning, medical image processing, data science, recommendation systems, and blockchain. Rahman has published more than 45 peer-reviewed articles in conferences and journals. He can be contacted at: mijanur.rahman@seu.edu.bd.