A Hybrid Multi-Model Deep Learning Framework for Breast Cancer Detection Using Thermogram Imagery

Preethi Veerlapalli, Sushama Rani Dutta
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad-500075, Telangana, India
E-mail: preethireddyveerlapalli@gmail.com, Sushamarani.dutta@klh.edu.in
*Corresponding author

Keywords: breast cancer screening, region of interest segmentation, hybrid deep learning, artificial intelligence, multimodal deep learning

Received: June 21, 2025

Breast cancer is one of the most common diseases among women worldwide, which causes high mortality in case of late detection. Thermography is a non-invasive imaging technology that can be used to help with the early diagnosis of breast cancer. With the development of AI, thermogram-based breast cancer screening using deep learning techniques has gained significant value. However, the detection accuracy and robustness of current deep learning algorithms are still challenged. To resolve this, we designed a hybrid multi-model deep learning framework which combines ROI segmentation by ROISegNet, feature extraction with enhanced edge using two edge detectors (Prewits and Roberts), and finally feature extraction and classification by a deep network architecture, InceptionResNetV2. The proposed framework was tested on the DMR-IR dataset, which is publicly available and includes thermograms of 44 subjects (29 breast cancer patients and 15 healthy). The dataset was separated into training (580 affected and 300 healthy images) and validation (160 affected and 80 healthy images) sets. The proposed model also achieved better results than the state-of-the-art 98.78% accuracy, 97.97% precision, 96.52% recall, and 97.24% F1-score, compared with other base networks, such as VGG19, ResNet50, DenseNet121, and InceptionV3. This performance also indicates that the incorporation of edge-enhanced feature maps and ROI segmentation in a hybrid deep learning framework is a practical design. The method presented here represents a potential direction for reliable non-invasive early detection of breast cancer based on thermogram images.

Povzetek: Predstavljen je hibridni večmodelni pristop za zgodnje odkrivanje raka dojk s termografijo, ki združuje segmentacijo ROI, večrobno poudarjanje in globoko učenje. Metoda dosega visoko točnost in robustnost na podatkovni zbirki DMR-IR.

1 Introduction

Breast cancer ranks first among all cancers in women globally and is the primary cause of years lost to disability-adjusted life expectancy. Consequently, reducing the fatality rate from breast cancer requires early detection. Despite being the primary treatment for diagnosing and screening for breast cancer, mammography still has certain limitations. Two more methods of screening include ultrasounds and clinical breast exams. Mammography continues to serve as the gold standard for breast cancer and stands as a beacon of hope and progress in the pursuit of effective detection despite several drawbacks.

Thermography is a newly developed screening technique. Recent technological developments have shown that thermography is a superior breast cancer screening method to other approaches [1]. Breast cancer diagnosis has already significantly benefited from several profound learning-based contributions. Rajinikanth et al.'s study [3] concentrated on automating breast cancer diagnosis with thermograms. Hakim and Awale [5] emphasized the need for early and precise detection of breast cancer, pointing out the need to reduce false positives and negatives, enhance automated interpretation, and explore deep learning techniques. To validate the possibility of precise breast cancer risk prediction, further investigation is required. Mashekova et al. [6] noted that breast cancer is a frequent and fatal illness, highlighting the need for further investigation and integration with artificial intelligence to serve as an additional tool for early detection. Yadav and Jadhav [10] demonstrated radiation-free early illness detection using thermal cameras and machine learning, outperforming conventional techniques, especially mammography, and showing potential uses beyond breast cancer detection. They also outlined obstacles, including a limited dataset and picture similarity, and proposed remedies like feature aggregation for future improvements. Cauce et al. [15] offered a CNN model with several inputs for cancer detection, integrating clinical and thermal imaging data. They suggested that future research should compare manual feature extraction, explore picture segmentation, and consider different classification techniques for better explainability. There is a need to optimize deep learning models and hybridize them to leverage performance.

Early, non-invasive breast cancer screening is essential, and thermal imaging is a radiation-free alternative that is yet to be harnessed to its full potential due to variability in image quality, scarcity of annotated datasets, and inconsistency in deep learning workflows. Previous works typically fail to incorporate ROI-oriented processing and edge-aware features, or they directly utilize a single backbone with poor generalization. We aim to construct a strong and repeatable pipeline that (i) extracts clinically valuable regions through ROISegNet, (ii) enhances the structural hints with different edge maps (Prewitt, Roberts), and (iii) exploits a unified InceptionResNetV2 backbone for classification. We perform five-fold cross-validation on the DMR-IR dataset and present the accuracy, precision, recall, F1, and AUC with statistically significant validation, which indicates the potential practical feasibility of thermo-based early screening.

Here are the things we brought to this study.

1. To enhance breast cancer diagnosis, we suggested a hybrid multi-model deep learning system that uses thermogram imaging. The framework comes with a hybrid deep learning model called InceptionResNetV2 to remove features from the photos.
2. Building upon a multi-model deep learning methodology, we created an algorithm termed Learning-based Breast Cancer Detection (LbBCD) to diagnose breast cancer efficiently using thermogram pictures.
3. We also developed a prototype application using thermogram imaging to evaluate the suggested architecture and the underlying algorithm.

The remainder of the document is arranged according to the following framework. Section 2 summarizes earlier research on deep-learning models for identifying breast cancer. The preliminary information needed to comprehend the suggested framework is presented in Section 3. The hybrid multi-model deep learning framework indicated for automatic breast cancer screening is presented in Section 4. Our test findings are shown in Section 5. Along with outlining the proposed research's constraints, Section 6 addresses the research's significance. Section 7 concludes our rework aimed at early breast cancer screening and offers guidelines for more studies in the future.

2 Related work

Numerous deep-learning methods used for breast cancer screening have been documented in the literature. Mishra et al. [1] used thermal imaging to predict breast cancer using a CNN model. The approach produces excellent accuracy with 680 thermograms, surpassing the accuracy of 50 thermograms previously used. This illustrates the deep learning model suggested for breast cancer prediction ability. Ekici and Jawzal [2] indicated that the approach uses convolutional neural networks and thermal images to achieve colossal accuracy. Future studies should investigate the advanced Dynamic Infrared Thermography (DIRT) application and thermography's ability to identify cancer at a level of detail, using models such as CNN to get better results. Rajinikanth et al. [3] focused on employing thermal imaging for automated breast cancer diagnosis. The technique includes image recording, patch extraction, image processing, feature extraction, optimization using the Marine Predators Algorithm, and two-class classification using the Decision-Tree classifier. Mambou et al. [4] emphasized the necessity for preventative actions and the worldwide consequences of breast cancer. This paper presents a comparative analysis combining deep learning and computer vision techniques to improve breast cancer diagnosis. Hakim and Awale [5] identified breast cancer early and accurately. Reducing false positives and negatives, enhancing automated interpretation, and investigating deep learning techniques are among the challenges. More studies must confirm its potential for accurate breast cancer risk prediction.

Mashekova et al. [6] observed that a frequent and fatal illness is breast cancer. Studies highlight the need for more investigation and integration with artificial intelligence while confirming its potential as an additional tool for early detection. Roslidar et al. [7] developed non-invasive screening methods essential for breast cancer. The potential of thermography, CNN use, and better future research possibilities for precise categorization are reviewed in this work. Milosevic et al. [8] enhanced the screening process for breast cancer, which is essential for early identification. The suggested improvements include thermography, effective target population identification, and software-supported mammography analysis. Husaini et al. [9] found that a novel technique combines cloud computing, deep neural networks, cellphones, and thermal imaging to identify breast cancer early. Further development might increase the tool's diagnostic accuracy and range of uses. Yadav and Jadhav [10] provided radiation-free early illness detection through thermal cameras and machine learning. It performs better than conventional techniques, particularly mammography, and has potential uses outside the detection of breast cancer. Obstacles encompass a limited dataset, picture similarity, and possible remedies such as feature aggregation for subsequent enhancements.

Mambo et al. [11] investigated and found that it affected 2.8 million people with global breast cancer in 2016, highlighting the need for early detection. Restrictions in mammography drive innovations like InceptionV3-KNN. Improved detection using a 3D breast model and a thermal sensitivity camera is the goal of future research. Tsietso et al. [12] claimed that the primary motivation for the economic identification of breast cancer is the fact that it is among the leading causes of mortality for females. The effectiveness of segmentation is emphasized as the research assesses current CADx systems for breast thermograms based on deep learning. It underscores the need for more studies on lateral breast thermograms and readily available datasets. Salvi and Kadam [13] argue that innovative approaches to healthcare are required due to the worldwide health load. Machine learning facilitates the timely and effective identification of breast cancer. Patients who live far away benefit from IoT integration. Tiwari et al. [14] presented "Deep Multiview Breast Cancer Detection," an automated approach for detecting breast cancer that uses VGG16 for precise classification. The system presents the concept of multi-view breast thermal imaging, demonstrating its utility. Prospective research endeavors involve augmenting databases, integrating clinical data, and investigating sophisticated 3D CNN networks to enhance precision. Cauce et al. [15] discussed identifying breast cancer by using a multi-input CNN model that integrates clinical and thermal imaging data. Future research should compare manual feature extraction, investigate picture segmentation, and consider different classification techniques for better explainability. Ovies et al. [16] investigate the application of CNNs in infrared thermography for early breast cancer detection. Resnet topologies outperform VGG, yielding transparent and effective outcomes. According to the study, CNNs—in particular, Resnet—may be helpful for non-invasive breast cancer screening and can lower healthcare costs while improving survival rates. Mammottil et al. [17] found that, due to its seriousness, breast cancer must be detected early. The present approach, mammography, is radiation- and money-intensive. Thermography is becoming increasingly popular as a less intrusive and expensive alternative. Comparing single-input CNN performance, utilizing pre-trained CNN models, and investigating data preparation methods such as segmentation and augmentation are some of the future research directions. Husaini et al. [18] investigated the use of DCNNs (Inception V3, V4, modified MV4) in color thermal imaging for accurate early detection of breast cancer, a global issue. Allugunti [19] observed that the identification of breast cancer is critical, and machine learning classifiers (CNN, SVM, RF) have demonstrated promise in patient classification. CNN, for instance. In addition to resolving medical imaging issues, this research suggests future developments in computer-aided diagnostic systems and case prioritization for radiologists. Husaini et al. [20] noted that early detection is essential for reducing the death rate from breast cancer. In addition to reviewing AI and thermography for detection, this work addresses limitations and suggests future research directions. The research emphasizes the potential of thermography, makes recommendations for improvements, and stresses the significance of standards in imaging portable devices.

Singh et al. [21], after doing research, because breast cancer affects women more frequently than any other malignancy, it has been determined that early identification is essential. A complementary non-invasive method to mammography is infrared breast thermography. False positives are dealt with numerically, while recommendations for the future center on using machine learning in real-time. The research emphasizes that better interactions between intelligent technologies and clinicians are necessary. Alfayez et al. [22] found that an enormous hazard to women's health worldwide is breast cancer. The four procedures are image pre-processing, ROI detection, feature extraction, and ELM and MLP classification. The study recommends further research comparing segmentation approaches with CNN and utilizing various retrieved characteristics. Kiymet et al. [23] discovered that over 15% of women have breast cancer and that early detection is essential. Four deep-learning network strategies based on thermal images are presented in this research. There is a shortage of research on breast cancer diagnosis using thermal images. Upcoming enhancements might entail investigating various segmentation techniques and augmenting the number of training images. Gomez et al. [25] observed that, according to a recent GLOBOCAN research, two million women worldwide received a breast cancer diagnosis in 2018. This paper presents an approach to computer-aided diagnosis utilizing convolutional neural networks (CNNs) and thermal images. CNNs surpass modern architectures in terms of speed, dependability, and robustness. The recommended method highlights the importance of database size and data augmentation by providing a baseline for detecting breast cancer using CNNs and thermal images.

Nasser et al. [26] suggested a technique that uses learning-to-rank and texture analysis to detect breast cancer in dynamic thermograms. This method achieves competitive results (AUC = 0.989) by producing a compact representation for sequences—subsequent research endeavors to improve classification via sparse dictionary learning. Yousefi et al. [27] suggested a technique that combines AI and thermography to identify breast cancer. An accuracy of 78.16% is obtained by extracting and reducing high-dimensional features—subsequent research endeavors to broaden validation and evaluate comprehensive thermal properties. Zadeh et al. [28] suggested using thermography in conjunction with computer-aided diagnostics to detect breast cancer early. It uses an autoencoder neural network for classification, a semiautomatic method for breast area segmentation, and feature extraction. Kakileti et al. [29] presented the Thermalytix Risk Score (TRS), a thermal imaging-based AI risk assessment tool for individualized breast cancer risk prediction. Age-based risk ratings are outperformed by TRS, which provides poor nations with a portable, non-invasive alternative that encourages early identification. Zheng et al. [30] presented DLA-EABA, an AdaBoost algorithm for breast cancer diagnosis aided by deep learning. In comparison to other techniques, it obtains a high accuracy rate.

Kadry et al. [31] talked about using thermal imaging to identify breast cancer. Immense accuracy is achieved by VGG16 with DT using Pre-trained Deep-Learning Methods. Cai et al. [32] recommended developing a diagnosis method that combines feature extraction, an improved CNN, and image processing to identify breast cancer automatically. The proportion of accuracy is based on the findings. Dey et al. [33] suggested using thermography and a trained DenseNet121 model to create an affordable breast cancer detection system that surpasses previous cutting-edge methods on the DMR-IR dataset. Data augmentation and class imbalance resolution are planned for the future. Mishra et al. [34] investigated using SVM, KNN, RF, and DT models for thermography for breast cancer detection. Of feature extraction methods (SIFT, SURF), KNN has the best accuracy. According to studies, future improvements in classifier performance may be achieved by implementing deep learning models. Guan et al. [35] detected that early breast cancer is made more accessible by infrared thermography. Breast area segmentation uses an autoencoder-like C-DCNN, which shows promise and improves with more training examples. Benhammou et al. [36] improved CADs for breast cancer. Based on clinical efficiency, Magnification-Independent Multi-category (MIM) reformulation is preferred among the four classifications according to a taxonomy. Deep learning studies show the potential of MIM. Ahmed et al. [37] observed that breast cancer ranks ninth among all cancers in terms of mortality. Mammogram analysis is aided by a deep learning-based system that achieves excellent segmentation and accuracy. Future uses might involve cancer diagnosis and a variety of medical imaging techniques. Rautela et al. [38] are concerned about breast cancer. Because of the hazards associated with radiation exposure, women frequently delay getting diagnosed. This study examines screening techniques, with a focus on deep learning methods. It also covers performance metrics, datasets, potential research topics, and the benefits and drawbacks of each. Encouraging novel research in breast cancer detection is the main objective. According to Nassif et al. [39], every year, thousands of people are affected by breast cancer. Survival rates are increased by early detection, which is made possible by deep learning and AI. Gene sequencing data, possible hybrid algorithms, and the extraction of noteworthy characteristics require further investigation to improve risk assessments and forecasts. Identifying risk levels, predicting recurrence, and utilizing multiclass predictors are potential areas for further study. Abhisheka et al. [40] lacked a long-term cure; breast cancer (BC) is a global health problem. They improved survival results from early detection. Insights for researchers are provided by this review, which focuses on Deep Neural Network (DNN) methodologies, imaging modalities, and BC diagnostic problems. Veerapalli and Das [41] suggested the ROISegNet model for an effective thermogram imagery-based method. Based on the literature study, deep learning models have been found to work well in image processing. There is a need to optimize deep learning models and hybridize them to leverage performance.

Rasha Talib Gdeeb (2023) employed image segmentation with neural networks on X-ray images, achieving improved breast cancer detection and diagnostic accuracy

Table 1: Summary of existing deep learning approaches for breast cancer detection using thermogram imagery

Study / Year	Model / Technique Used	Dataset Used	Key Findings	Identified Research Gap
Mishra et al. [1] (2020)	CNN-based classification	Custom thermogram dataset	Demonstrated the feasibility of thermal imaging with CNN models for early breast cancer detection.	The dataset is limited in size and diversity, and it lacks ROI segmentation and advanced feature enhancement.
Rajinikanth et al. [3] (2021)	CNN + Marine Predators Algorithm	DMR-IR	Showed that optimization algorithms can improve feature selection and classification accuracy.	The approach did not incorporate edge information or ROI-based segmentation.
Yadav & Jadhav [10] (2020)	Traditional ML models (SVM, RF, KNN)	Thermal imaging dataset	Proposed radiation-free detection method using thermal images with basic ML techniques.	Lacks deep feature extraction and robustness to image noise.
Cauce et al. [15] (2021)	Multi-input CNN (Thermal + Clinical Data)	Custom hospital dataset	Combined clinical and thermal data for improved detection performance.	Limited generalizability; no edge-based feature enhancement explored.
Husaini et al. [18] (2022)	Inception V3, V4, MV4	Thermal imaging dataset	Validated deep CNN architectures for thermal breast cancer detection.	No ROI-based segmentation or multi-level edge information integration.
Dey et al. [33] (2022)	DenseNet121 with basic edge detection	DMR-IR	Incorporated edge detection to enhance detection capability.	Used a single edge detector and lacked hybrid multi-model fusion.
Proposed Method (2024)	ROISegNet + Multi-Edge Detection (Prewitt & Roberts) + InceptionResNetV2 Hybrid Model	DMR-IR	Integrates ROI segmentation, multiple edge detectors, and hybrid deep learning to enhance feature richness and improve early breast cancer detection.	Addresses all identified gaps by combining ROI-based segmentation, multi-edge fusion, and a hybrid deep learning approach for superior detection.

Table 1 Comparison of recent works in the detection of breast cancer based on thermogram images. It describes the main models and methods applied, datasets used, and the main results of each work, as well as research gaps identified in these works. It is clear from the comparison that the majority of current works utilize traditional machine learning models or deep learning models without essential pre-processing steps, such as ROI-based segmentation and advanced edge feature enhancement.

Several studies have indicated that thermal imaging can play an essential role in early breast cancer detection; however, they have small dataset sizes, fail to integrate multi-edge information, and lack robust hybrid deep learning approaches. Such constraints frequently lead to the loss of generality and the degradation of detection accuracy. To the best of our knowledge, the limitations above are not tackled using an approach that combines ROI segmentation (ROISegNet), multi-edge detection based on Prewitt and Roberts, and deep learning, as the proposed architecture proposes to combine. This joint method will enhance feature representation to alleviate early breast cancer screening performance.

3 Preliminaries

The section provides preliminary details required to understand the proposed methodology presented in this paper.

3.1 Convolutional neural networks

Multi-layer-linked neural networks are used to build CNNs. Robust characteristics at the low, middle, and high levels are retrieved hierarchically. A typical CNN framework consists of two primary layers, pooling and convolutional layers, which provide the network's convolutional basis [43]. There are networks with ultimately linked layers, including VGG and AlexNet. The convolutional layer takes spatial characteristics from the pictures and applies a filtering function. Convolutional layers are often layered in this manner because the early convolutional layers extract tiny local patterns like edges and corners. On the other hand, the final layers identify image structures of high-level features. This shows that CNNs can effectively learn patterns in spatial hierarchies. Convolutional layers are characterized by the size of the convolutional patch and the number of filters. Next, we compute the dot product between the kernel weights and the receptive field, a tiny fraction of the input volume. A stride separates two successive convolutional windows. Since more significant stride values cause feature maps to be down-sampled, convolutional layers typically have one stride [43]. This simple convolution method produces a new image called a feature map, a visual depiction of the extracted features. Because CNNs can share weights and biases among all of their neurons, the number of parameters is substantially lower than in a fully connected layer because of this characteristic.

The Rectified Linear Unit (ReLU) [44] is a typical example of an element-wise nonlinear activation or non-linearity function used for each feature map component. The ReLU function successfully adds non-linearity to the network compared to the sigmoid activation function or the hyperbolic tangent, frequently used in classical neural networks [44]. Compared to traditional functions with gradient descent, the ReLU dramatically speeds up the training process. Due to the so-called "vanishing gradient problem," which is brought on by abnormally low derivatives of previous functions in the saturation zone, like the sigmoid, the weight updates almost disappear entirely. Owing to shared pixels across all windows, several feature maps with highly identical content may be generated, indicating duplicate data. To reduce the variability of the extracted features, pooling layers are added after each convolutional layer using simple techniques such as max and average. Both max-pooling and average-pooling layers utilize a sliding window and a predetermined stride over the feature maps to identify the maximum and mean values. Pooling layers decrease the size of the feature maps using a stride of two or more. Notably, by refining the output of the convolutional layer, the pooling layer (also referred to as the sub-sampling layer) selects the more stable and abstract characteristics for the subsequent layers. Consequently, the pooling layer reduces the computational load by shrinking the size of the feature maps.

As previously indicated, specific models may have ultimately linked layers before the classifier layer connects—input to the classifier layer from many layered convolutional and pooling layers. The ultimately linked layer has a high parameter occupancy, which makes overfitting possible. Because of this, the dropout approach—an effective regularization technique—can help reduce or alleviate issues related to overfitting. This method enables neurons to develop more substantial autonomous properties by preventing over-adaptation to their environment during training, achieved through the random elimination of specific neurons and their connections from the network [45]. The classification layer, which comes in last, ascertains the posterior probability for every category. The softmax classifier is the most popular classifier layer for image processing in the deep learning community, commonly called a normalized exponential. Stochastic gradient descent (SGD) optimization is frequently used for CNN training and weight adjustment in backpropagation workflows. The raw input, or initial pixels, is the first step in an end-to-end deep neural network learning process, which concludes with the final label.

3.2 InceptionV3 model

A pioneer of non-sequential CNNs, GoogLeNet was victorious in the ILSVRC-2014 classification and detection tracks of the competition. The number of units at every level in this network, as well as the depth or number of levels, may be raised without putting undue load on the computer system [46]. The foundation of GoogLeNet is based on the theory that, due to correlation, many connections between layers are inefficient and contain duplicated data. It uses a sparse CNN, the "Inception module," with 22 layers and a parallel workflow. Numerous auxiliary classifiers are included in the intermediate levels to increase the discriminating strength in the lower layers. The Inception module may use pooling and convolutional operations at every layer, unlike conventional CNNs like AlexNet and VGG, which can only use one. Additionally, multiple-sized filters (convolutions) are applied to the same layer to extract patterns of varied sizes and provide more comprehensive information.

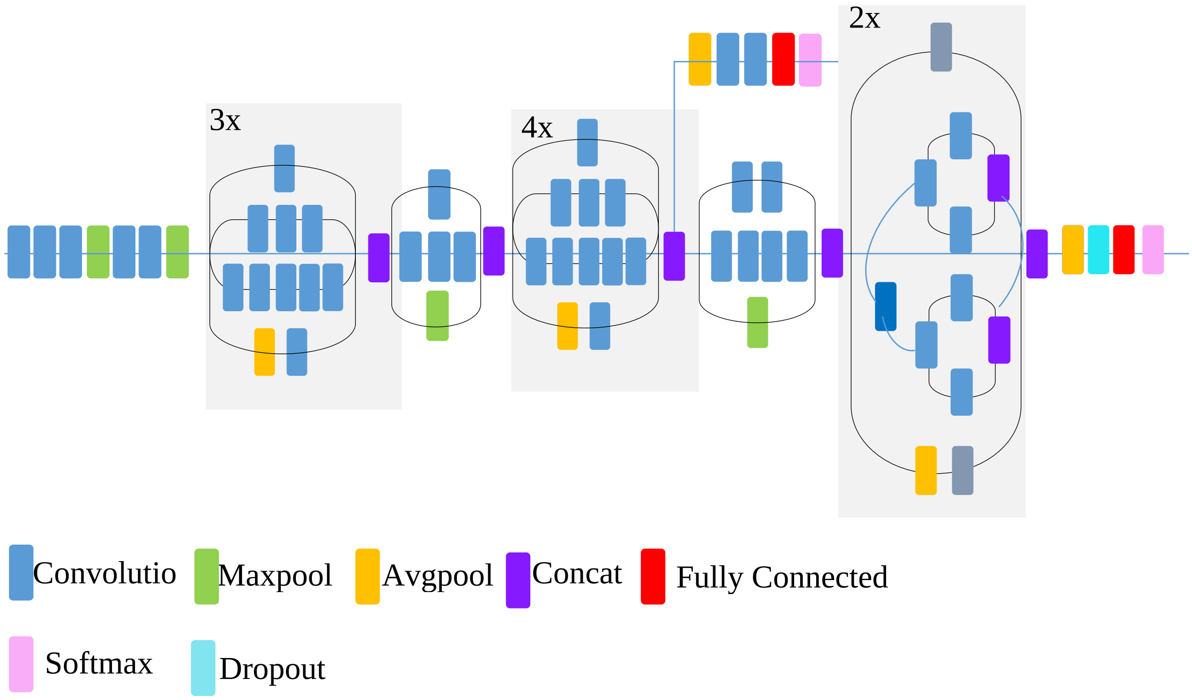

Figure 1: Architectural overview of the InceptionV3 model

One common technique to reduce the number of parameters and processing complexity is using a 1 × 1 convolutional layer, sometimes called the bottleneck layer. One-by-one convolutional layers are used before larger kernel convolutional filters to reduce the parameters calculated during the feature pooling process. ReLU is used after each 1×1 convolutional layer, which adds more non-linearity and deepens the network. A layer called average pooling takes the role of the fully connected layers in this network. This significantly lowers the number because the fully connected layers have many characteristics. Because of its speed advantage over VGG, this network can learn more detailed feature representations with fewer parameters than AlexNet [47]. InceptionV3's condensed perspective, used in this investigation, is shown in Figure 2.

3.3 ResNet50 model

In the ILSVRC-2015 competition, the classification challenge was won by ResNet, a deep neural network consisting of 152 layers [48]. However, the primary issues with deep networks are their high training error rates, difficulty in training, and vanishing gradient, which results in very little learning at the lower layers during the backpropagation stage.

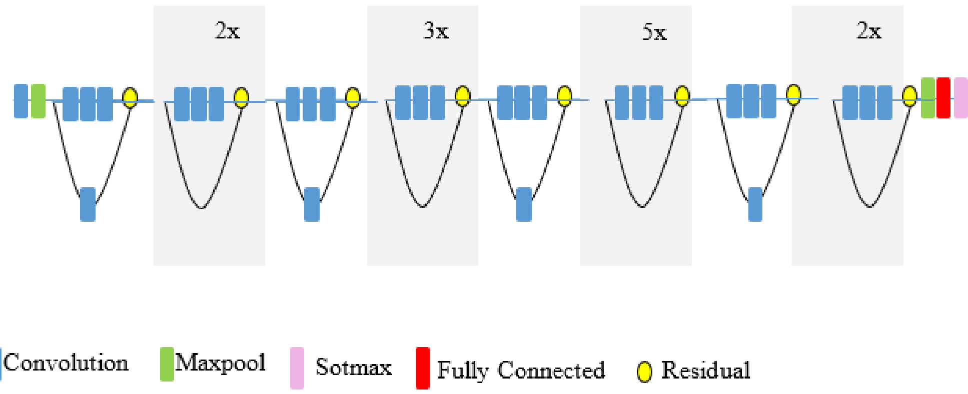

Figure 2: Architectural overview of the ResNet50 model

By using additive identity transformations and a deep residual learning module, the ResNet configuration effectively addresses the vanishing gradient problem. In this approach, residual mapping of stacked layers is made instead of directly targeting the desired underlying mapping, and the input and output are directly connected via the residual module [48]. Optimizing the residual map is much easier compared to working with the original, unreferenced map. Similar to VGG, ResNet primarily utilizes 3 × 3 filters, though it is less complex and contains fewer filters than VGG [48]. The study used ResNet, depicted in a compressed form in Figure 3.

4 Materials and methods

This section offers thorough information about the suggested methodology, including the framework, the underlying algorithm based on hybridized deep learning approaches, dataset specifics, and performance evaluation methodology.

4.1 Our framework

Our study uses transfer learning to construct a thermogram image-based breast cancer screening technique. We employed the pre-trained InceptionResNetV2 as a basic feature extractor in the proposed model to learn multi-scale features and deep residual connections toward precise breast cancer detection.

DenseNet's approach to addressing the vanishing gradient problem justifies its usage. To extract edge information from the thermal breast images, we used two edge detectors, Prewitt and Roberts, in addition to the pre-trained DenseNet121 model. The outputs of these edge detectors, paired with the original grayscale breast image, produce a three-channel image that highlights edge prominence. This is crucial as the pre-trained algorithm may only be capable of extracting features from 3-channel images. Figure 3 illustrates the overall architecture of our proposed work; the corresponding modules are discussed in the subsequent subsections.

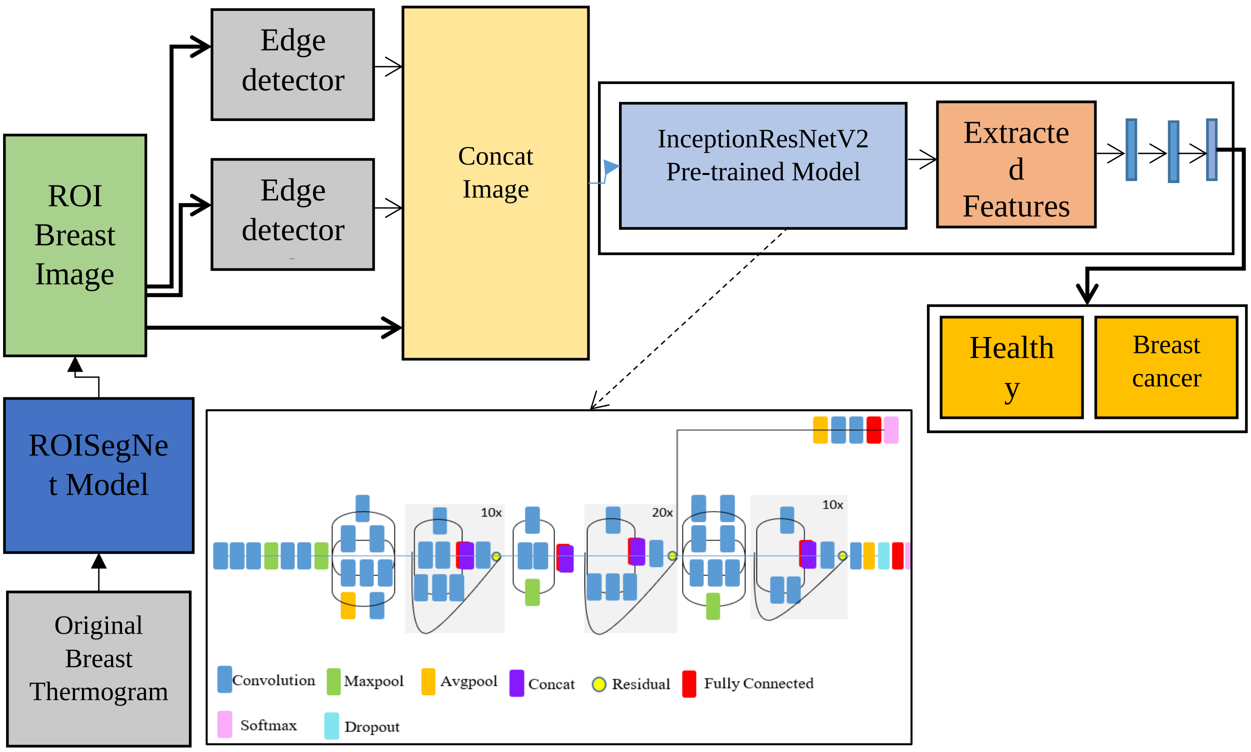

Figure 3: Proposed hybrid multi-model deep learning framework for breast cancer detection

Developing a hybrid multi-model deep learning system that starts with the original breast picture is suggested. During processing, a ROISegNet, our prior work published in [41], the model performs Region of Interest (ROI) segmentation on the source picture. The output of the ROISegNet model is the segmented ROI breast image. Two edge detectors process the ROI breast image to highlight essential features. The outputs from the two edge detectors are merged to produce a new image. This new image is processed to extract features using a pre-trained InceptionResNetV2 model. Classification is then carried out using the features that were taken from the pre-trained model. Finally, a classification layer determines whether the extracted features indicate a healthy breast or breast cancer. To increase the diagnostic accuracy of breast cancer, this architecture integrates segmentation, edge detection, and deep feature extraction.

4.1 Edge detectors

To categorize breast cancer patients, it is crucial to examine even the most minor details, such as blood vessels and breast deformities, to establish whether or not the patient has the disease. To enhance the informative value of the original photos, we have taken and integrated edge information from the thermal images. To achieve this, we first created edges in the original gray-scale thermogram pictures using the well-known edge detection methods of Roberts and Prewitt. Most of the minor edges are preserved with the help of these detectors. Together with the original image, these two edge-marked images make a picture with three channels fed into the DenseNet121 model, which has previously undergone training.

Roberts and Prewitt use an edge detection method based on gradients. Convolution of the picture using vertical and horizontal derivative masks allows for the detection of edges. Horizontal and vertical operators are other names for these masks. Utilizing these operators enables the detection of edges through quantitative examination of the variation in pixel brightness. They determine whether edges are present in an image by computing the difference between the matched pixels using a technique akin to the derivative in the signal domain. Table 2 lists all of the notations used in this work.

Table 2: Notations table

Notation	**Description**
Th	Threshold
G	Gradient's magnitude
I_g	Gray-scale image
G_x and G_y	Gradient images
Δ_x *and* Δ_y	Edge operators

Define Δ_x and Δ_y as the edge operators for the horizontal and vertical directions, respectively. These operators are convolved with an I_g, a grayscale picture to produce two gradient images, G_x and G_y respectively.

$$G_x = I_g * \Delta_x$$

And

$$G_y = I_g * \Delta_y$$

The convolutional operator in this case is '*'. Note that Δ_x is rotated 90◦ to generate Δ_y, and vice versa. The gradient's magnitude, let's assume G, is computed as follows.

G= $\sqrt{G_x^2 + G_y^2}$

The idea of pixel coordinates aids in approximating the gradient computation, which involves determining the gradient at a specific pixel using the necessary masks. Here, the threshold value (th) is defined as the mean values appearing in G. This yields the final edges.

th=$\frac{1}{(MN)} \sum_{x=1}^{M} \sum_{y=1}^{N} G(x, y)$

In the end, the following equation is used to produce an edge picture I_e is computed as follows.

$$I_e(x,y) = \begin{cases} \text{edge pixel} & : \ \mathrm{G(x,y)} < \mathrm{th} \\ \text{background pixel} & : \ \text{otherwise} \end{cases}$$

Various edge detection methods exist, depending on the edge operators used. Below, we will explain the Roberts and Prewitt edge operators, which enable us to extract the edges needed for our model.

4.1.1 Roberts edge detector

The Roberts operators display high spatial frequency areas with a strong likelihood of identifying edges. Robert's operators use 2 x 2 masks and minimal gradient calculations. Only a pixel's four closest neighbors are looked at to assess if it could be an edge pixel. Given that it is a superior choice to its horizontal or vertical counterpart, the Roberts cross operator was chosen in this instance. The Roberts cross operator uses the following 2×2 kernels:

$$K_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad K_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Given an image I, the gradients are

$$G_x = I * K_x, \quad G_y = I * K_y$$

and the gradient magnitude and orientation are

$$\|\nabla I\| = \sqrt{G_x^2 + G_y^2}, \qquad \theta = atan2(G_y, G_x)$$

Note that kernels that differ by a global sign or a 90° rotation are equivalent for edge magnitude computation; we adopt the above orientation for consistency with later operators.

4.1.2 Prewitt edge detector

The Prewitt edge detector detects edges vertically and horizontally using three-by-three masks. One of the detector's benefits is its ease of use.

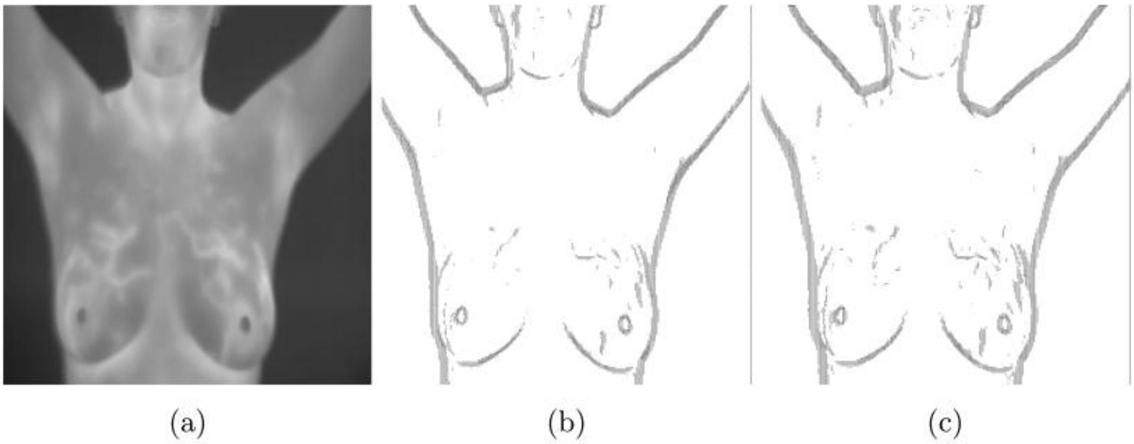

Figure 4: Original breast image (a), edge detection result using Prewitt operator (b), and edge detection result using Roberts cross operator

Figure 4 displays the outcomes of edge detection techniques applied to the initially presented thermal breast image. With the help of this technique, we can approximate the magnitude and identify edges and their orientations using the Prewitt edge detector.

The Prewitt operator uses two 3×3 convolution kernels to detect edges by computing gradients in the horizontal and vertical directions. Following standard conventions, Δx (or G_x) detects vertical edges by highlighting intensity changes along the x-direction, while Δy (or G_y) detects horizontal edges by highlighting intensity changes along the y-direction. The kernels are defined as:

$$\Delta_x = G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \qquad \Delta_y = G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Given an input image I, the gradients are computed as:

$$G_x = I * \Delta_x, \quad G_y = I * \Delta_y$$

The gradient magnitude and orientation are then calculated using:

$$\|\nabla I\| = \sqrt{G_x^2 + G_y^2}, \qquad \theta = atan2(G_y, G_x)$$

By explicitly defining Δx as the vertical edge detector and Δy as the horizontal edge detector, the roles of these operators are now consistent with standard practice and aligned with the interpretation used in Section 4.1.1 (Roberts operator).

The original breast thermogram is first processed for ROI segmentation by the fine-tuned ROISegNet, and we crop an ROI breast image. This ROI image is further applied two edge detection algorithms—Prewitt, Roberts. The two edge detector outputs and the segmented ROI grayscale images are jointly stacked together to create a 3-channel concatenated input, where channel 1 is the grayscale ROI, channel 2 is the Prewitt edge map, and channel 3 is the Roberts edge map. This last 3-channel image is applied to our previously trained InceptionResNetV2 model to extract features and to classify the input samples.

The edge detection is only performed on the segmented ROI image to show the structural boundary. The resultant edge maps, together with the grayscale ROI image, are stacked into a 3-channel input for the InceptionResNetV2 model. No edge detection is carried out on the original thermogram to prevent the overlap and contradiction of processing.

4.2 Proposed InceptionResNetV2 hybrid model

The proposed model is inspired by InceptionResNetV2, a general architecture featuring a combined multi-scale Inception module for feature extraction, and stable and efficient learning as ResNet's residual connections. Such an integrated architecture enables the network to recognize spatially intricate and contextually rich information, which is vital for precise breast cancer diagnosis. It applies batch normalization above the conventional layers instead of to the summations. Greater network depth is made likely by the additional modules, which increase the number of Inception blocks. The training phase, which has been identified as an essential issue with complex networks, gets fixed by remaining connections [48].

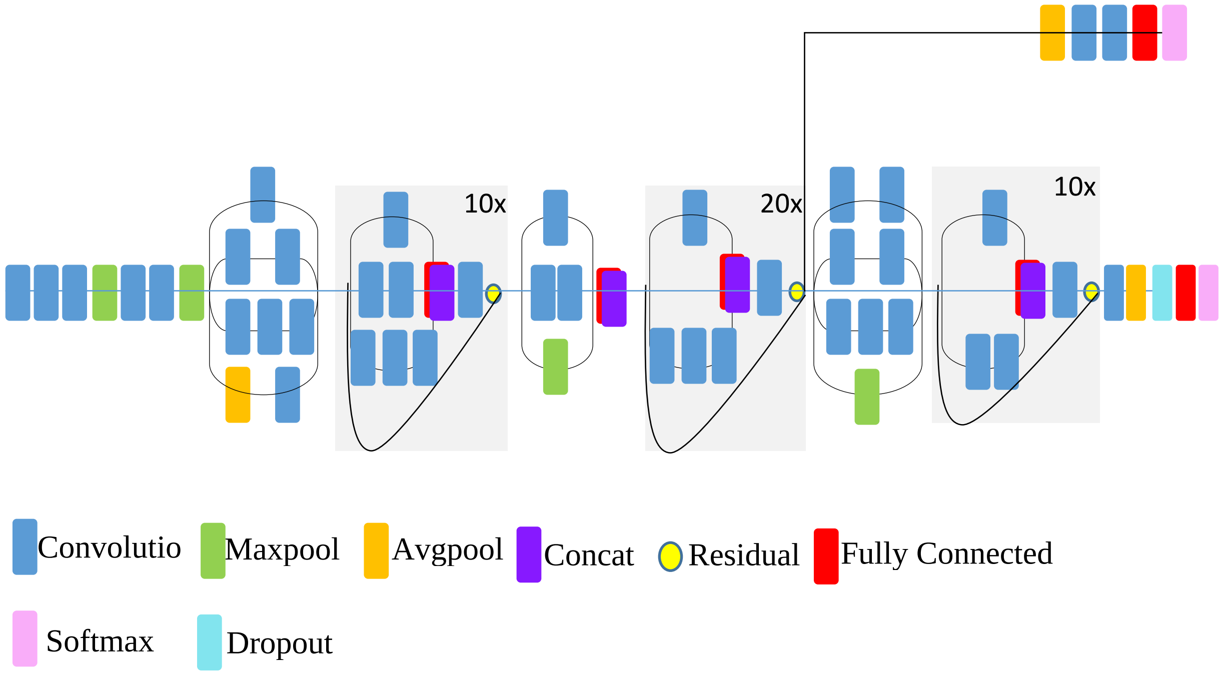

Figure 5: Proposed hybrid deep learning model known as InceptionResNetV2

Figure 5 depicts the building of a hybrid deep-learning model. The proposed model comprises convolutional layers (in blue), responsible for hierarchical feature extraction, and residual connections that facilitate the reformulation of connections in response to changes in the network during deep network training, thereby mitigating the vanishing gradient problem. Average-pooling layers (in yellow) are for reducing spatial dimensions gradually while preserving essential features, and the ultimate fully connected (in green) layers provide the high-level feature representation for classification. The remainder of the structure helps to train the network effectively at a deeper level, resulting in efficiency and accuracy in the InceptionResNetV2 model. Residual connections (green) can be interpreted as skip connections that forward the gradient through different layers, preventing vanishing gradient problems and making the deep network training possible. They are not convolutional layers, but they indirectly help in the stability and efficiency of the feature learning process across the network. Residual blocks in InceptionResNetV2 are primarily used to maintain gradient stability during deep network training, thereby preserving more helpful information for feedforward propagation through the network. This design enables the architecture to become deeper without degrading performance compared to traditional CNNs.

To prevent overfitting, dropout layers (shown in cyan) are used. To simplify gradient flow and improve training, the model incorporates residual connections (indicated by yellow circles). Concatenation rules (shown in purple) are used to combine features from many layers, while fully connected layers (shown in red) are used to integrate the learned features. "10x" and "20x," which stand for repeated layer blocks, are used in the layout to extract data further. Finally, the model includes softmax and fully connected layers (shown in pink) for classification tasks. To offer a trustworthy and efficient model for complex data processing tasks, it makes use of convolutional processes, pooling, dropout, and residual connections. To effectively tackle the training problem, the network reduces the residual as more than 1,000 filters are added. When the number of filters exceeds approximately 1,000, these residual connections become unstable, making it challenging to train the network. Residual scaling was proposed to overcome this by Szegedy et al. in the InceptionResNetV2 model. (2016), which ensures the training of intense and wide models.

A simplified version of the InceptionResNetV2 is shown in Figure 5—the proposed hybrid deep learning model utilized in this research.

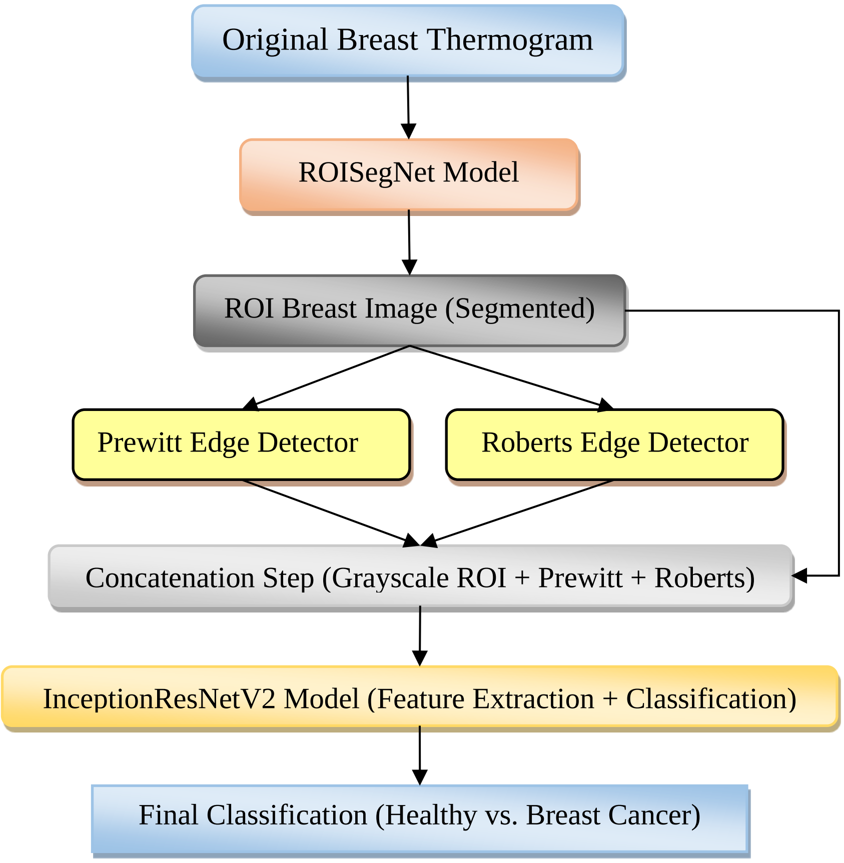

Figure 6: Flowchart illustrating the functionality of the proposed system

Figure 6 demonstrates the process for classifying breast cancer using thermographic images. The process begins with capturing the original breast thermogram. Next, the thermogram undergoes edge detection to highlight essential features. These detected edges are then combined and used for feature extraction, utilizing the proposed InceptionResNetV2 model into the breast cancer classification model, which categorizes the results into two groups: Healthy or Breast Cancer. The process concludes with presenting the classification results, determining whether the individual is healthy or has breast cancer.

4.3 Algorithm design

The proposed algorithm, Learning-based Breast Cancer Detection (LbBCD), aims to detect and grade breast cancer using the DMR-IR dataset accurately. The algorithm is designed to process and analyze medical images, particularly regions of interest (ROI) within breast images, to identify the presence of cancer and its severity. It utilizes advanced deep learning techniques, such as the InceptionResNetV2 architecture, to extract features from the ROI images and train a classifier to predict breast cancer in new samples. Additionally, the algorithm includes a performance evaluation step to assess the accuracy of its detection results against the ground truth, providing valuable insights into its effectiveness for clinical applications.

Algorithm: Learning-based Breast Cancer Detection (LbBCD)
Input: DMR-IR dataset D
Output: Breast cancer detection results R, performance statistics P
Input:
  Raw thermogram dataset D with labels
Output:
  Trained classifier f and evaluation metrics
1: Split D into training set T1 and test set T2
2: Initialize InceptionResNetV2-based classifier f
3: Normalize/standardize images as per preprocessing
4: M1 ← ROISegNet(T1) // segment ROIs for the training set
5: M2 ← ROISegNet(T2) // segment ROIs for the test set
6: G1 ← ToGrayscale(M1) // convert ROI images (train) to grayscale
7: G2 ← ToGrayscale(M2) // convert ROI images (test) to grayscale

```
X1 ← EdgeDetectAndConcatenate(G1, methods={Prewitt, Roberts})
X2 ← EdgeDetectAndConcatenate(G2, methods={Prewitt, Roberts})
f ← Train(f, X1, Y1)        // train on concatenated 3-channel ROI inputs
Ŷ ← Inference(f, X2)        // predict on test ROI inputs
Metrics ← Evaluate(Ŷ, Y2)   // Accuracy, Precision, Recall, F1, AUC, etc.
return f, Metrics
```

Algorithm 1: Learning-based Breast Cancer Detection (LbBCD)

Like Algorithm 1, the pipeline first splits the labeled thermogram dataset into a training set T1 and a test set T2. We extract breast ROI from every image using ROISegNet, yielding M1 and M2 for training and testing, respectively. These ROI images are converted to grayscale (G1, G2), and the edge detection is performed only on the ROI grayscale using the Prewitt and Roberts edge detectors. For each ROI, the generated edge maps are stacked with the ROI grayscale to form the three-channel input [ROIgray, Prewitt (ROIgray), Roberts (ROIgray)]. We then train the InceptionResNetV2 classifier (with a binary output head) on the concatenated train inputs X1 and the concatenated train labels Y1; the various test labels Y2 are predicted on X2. Finally, performance scores (Accuracy, Precision, Recall, F1, AUC) are calculated on the held-out test set (or averaged over folds if we use cross-validation as in Section 4.5). This explanation resolves the previous inconsistency by explaining that edge detection and concatenation are trained only after ROI division (instead of directly from raw thermograms), thereby making the algorithm presented in Figures 3 and 6 and described in Section 4.1 coherent. Data augmentation (rotations, flips, zooms, brightness) is applied on training ROIs only as described in Section 4.4, and edge maps are re-computed after augmentation.

4.4 Dataset details

To aid in the early detection of breast cancer, Silva et al. built a set of thermogram images and made them available to researchers [50]. This dataset was constructed using 20 consecutive photos with a 15-second gap between each one. When pictures were taken, the breast temperature and the environment were both the same. So the breast had been cooled with an air stream in advance of the photo. You can click on the link in [51] to obtain the pre-separated train and test sets of the dataset. Twenty-nine patients with breast cancer and 15 healthy cases' thermograms are part of the training set. The thermogram records of four healthy people and eight breast cancer patients make up the test set. Photographs of their ROIs and 20 breast thermograms for each case are included in the dataset. The dataset used in this study to train and assess the suggested model is split as shown in Figure 7.

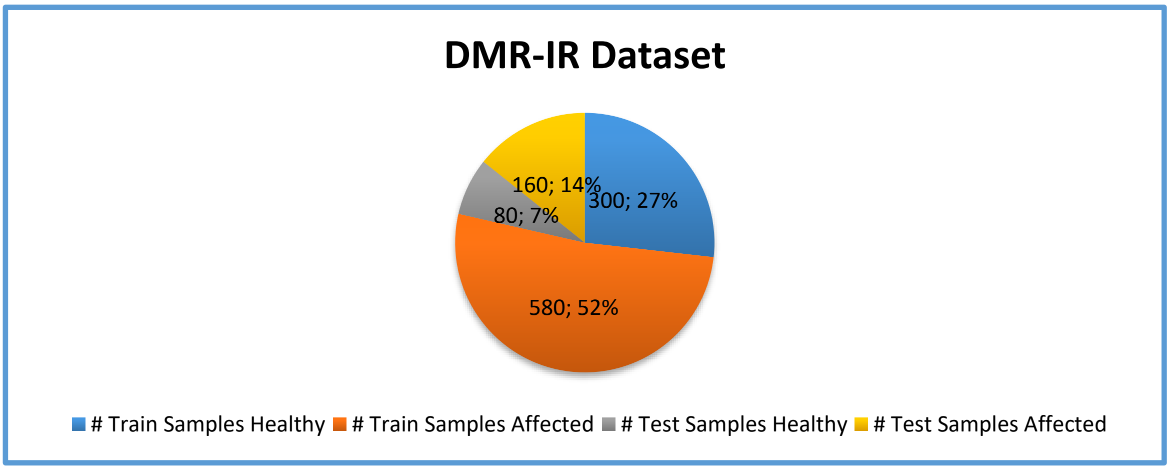

Figure 7: Data distribution dynamics of the DMR-IR dataset [42]

The DMR-IR dataset samples are distributed and divided into four groups, as seen in Figure 7. 58% of the dataset, or 580 affected train samples, is the largest segment. The second largest segment, which takes up 27% of the dataset, will consist of 300 healthy train samples. There are two smaller groups of test samples: 80 healthy test samples (7%) and 160 affected test samples (14%). A clear glance of the dataset's composition is given by this illustration, which also emphasizes a lower % of both test and healthy samples and the greater proportion of affected train samples.

For reproducibility, 30% of each dataset was utilized as a testing set, and the remaining 70% was applied for training: 880 training images (580 affected and 300 healthy) and 240 testing images (160 affected and 80 healthy), for 44 subjects (29 breast cancer patients and 15 healthy). We performed data augmentation in the training set using random rotations (± 15°), horizontal flipping, zooming (up to 20%), and changes in brightness to enhance the generalization of the model. The model was trained with a learning rate of 0.0001, a batch size of 32, and the Adam optimizer, 100 epochs with early stopping, avoiding overfitting. The loss function that was applied is binary cross-entropy. We tested on a workstation with Intel Core i9 CPU, 64GB RAM, and an NVIDIA RTX 3090 GPU (24GB VRAM), with Python 3.10 and TensorFlow/Keras as back-ends. These facts, in cooperation with the reading research questions documented in Section 4.0, describe a transparent and reproducible experimental design that ensures transparency of the study's goals and examination.

Because of the small scale of the DMR-IR dataset and potential overfitting risk, we performed various data augmentations on the training set. These transformations included random rotations (±15°), a horizontal flip, zooms up to 20%, and brightness changes. These augmentations increased the variety of images in the dataset, thus emulating thermogram capturing health fluctuations that occur during everyday practices. These augmentations artificially increased the dataset's size, which brought about an increase in the model's ability to generalise to new data.

4.5 Evaluation methodology

Since our approach was learning-based, the metrics obtained from our technique are assessed using the confusion matrix, as seen in Figure 8.

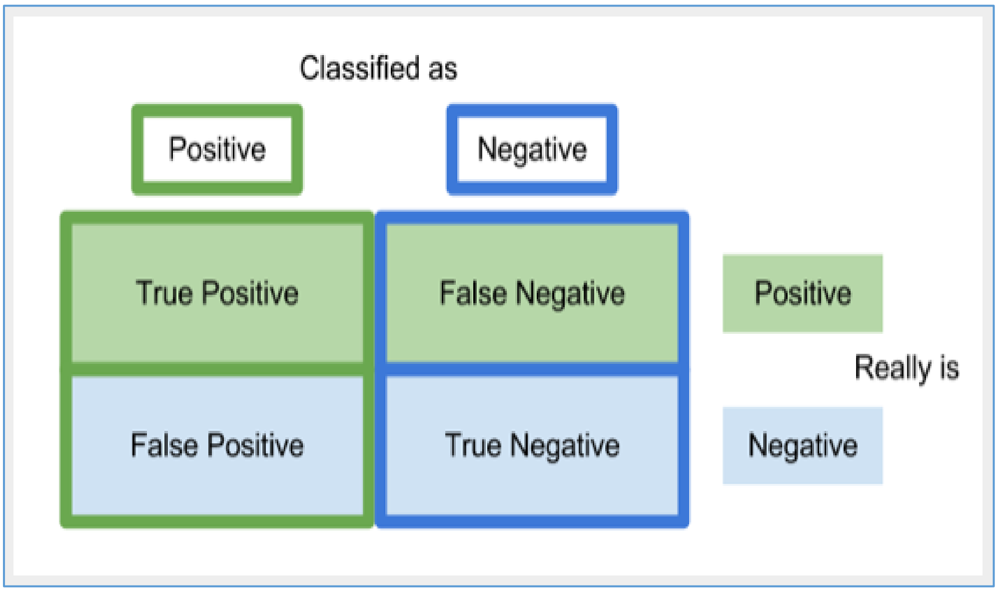

Figure 8: Confusion matrix

Performance statistics are derived from the predicted labels of our algorithm, compared to the ground truth, as indicated by the confusion matrix. The many tracks utilized in the performance evaluation are given in Equations 1 to 4.

$$\text{Precision } (p) = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall } (r) = \frac{TP}{TP+FN} \quad (2)$$

$$\text{F1-score} = 2 * \frac{(p * r)}{(p+r)} \quad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The performance evaluation metrics produce a value between 0 and 1. These measures are frequently employed in machine learning research. We adopted a five-fold cross-validation procedure to prevent overfitting further and provide a relatively rigorous evaluation. The data set was divided into five even parts; each part was used once as a validation set, while the remaining four parts were used as a training set. Evaluation statistics, such as accuracy, precision, recall, and F1-score, are calculated for each fold and averaged for reporting of the final values. Instead, this provides a more accurate estimation of model performance and convergence on small data.

5 Experimental results

This section shows the findings of our experiments with the suggested deep learning foundation framework for automatically identifying breast cancer from thermography photos. DMR-IR is the benchmark data set used in this paper's empirical study. The reported results represent the average performance across five-fold cross-validation, providing a robust and reliable assessment of the proposed framework on the limited DMR-IR dataset.

Table 3: Performance comparison among deep learning models for breast cancer detection (average across five folds)

Breast Cancer Detection Models	Precision	Recall	F1-Score	Accuracy
VGG19	95.44	95.8	95.62	96.51
ResNet50	93.27	91.7	92.5	93.84
DenseNet121	94.96	97	95.99	97.67
InceptionV3	96.03	94.7	95.34	97.53
InceptionResNetV2 (Proposed)	97.97	96.52	97.24	98.78

As indicated in Table 3, thermal image-based breast cancer diagnosis is achieved by contrasting the performances of various deep-learning models.

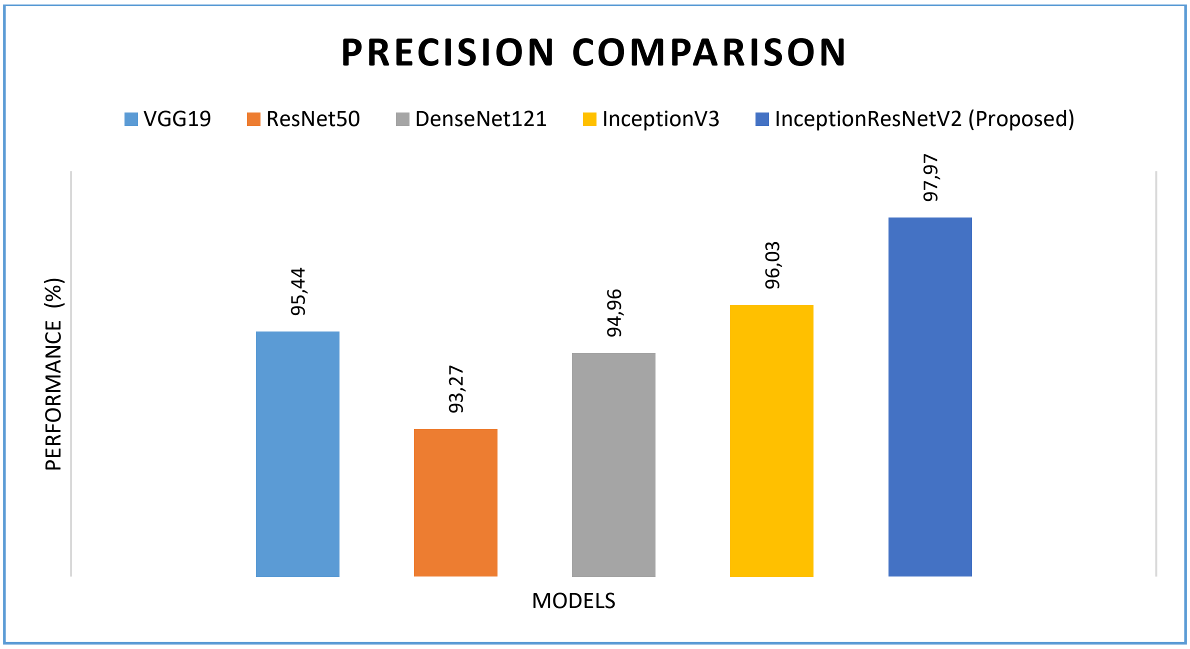

Figure 9: Breast cancer detection performance in terms of precision

Figure 9 displays the precision performance of various models. The models compared are VGG19, ResNet50, DenseNet121, InceptionV3, and InceptionResNetV2 (Proposed). VGG19 achieved a precision of 95.44%, ResNet50 scored 93.27%, DenseNet121 had a precision of 94.96%, InceptionV3 achieved 96.03%, and the proposed InceptionResNetV2 model recorded the highest precision at 97.97%.

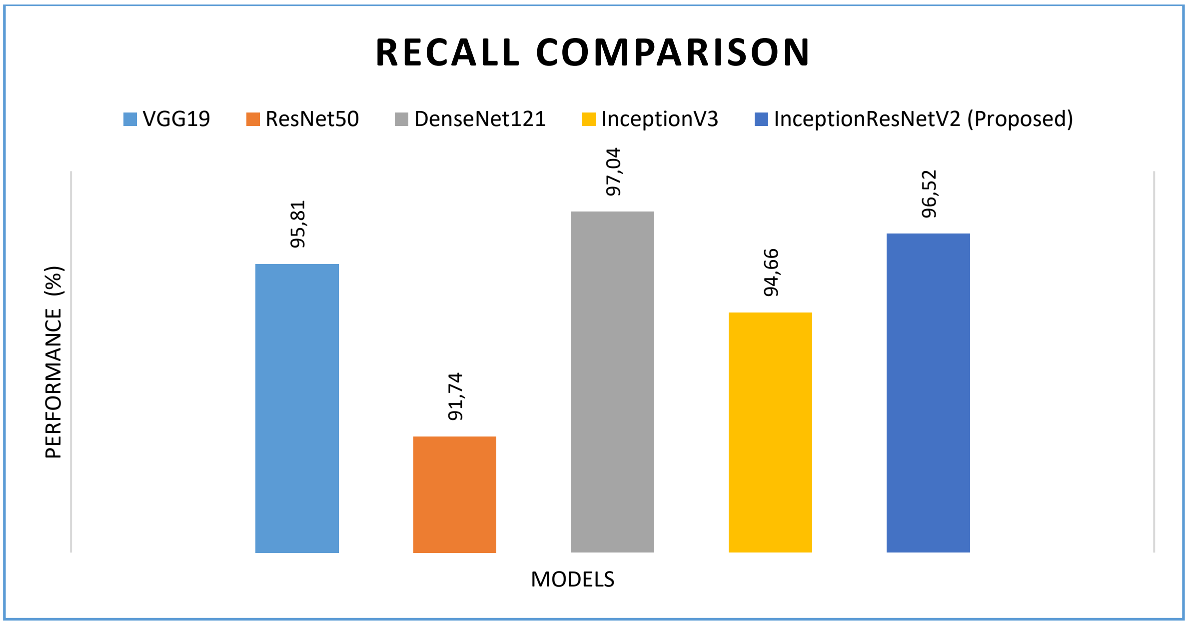

Figure 10: Breast cancer detection performance in terms of recall

In Figure 10, the recall performance of various models is presented in percentages. The models examined include VGG19, ResNet50, DenseNet121, InceptionV3, and InceptionResNetV2 (Proposed). The recall scores for these models are as follows: VGG19 achieved a recall of 95.81%, ResNet50 scored 91.74%, DenseNet121 recorded 97.04%, InceptionV3 achieved 94.66%, and the proposed InceptionResNetV2 attained a recall of 96.52%. This comparison shows that DenseNet121 has the highest recall performance among the models, followed closely by the proposed InceptionResNetV2.

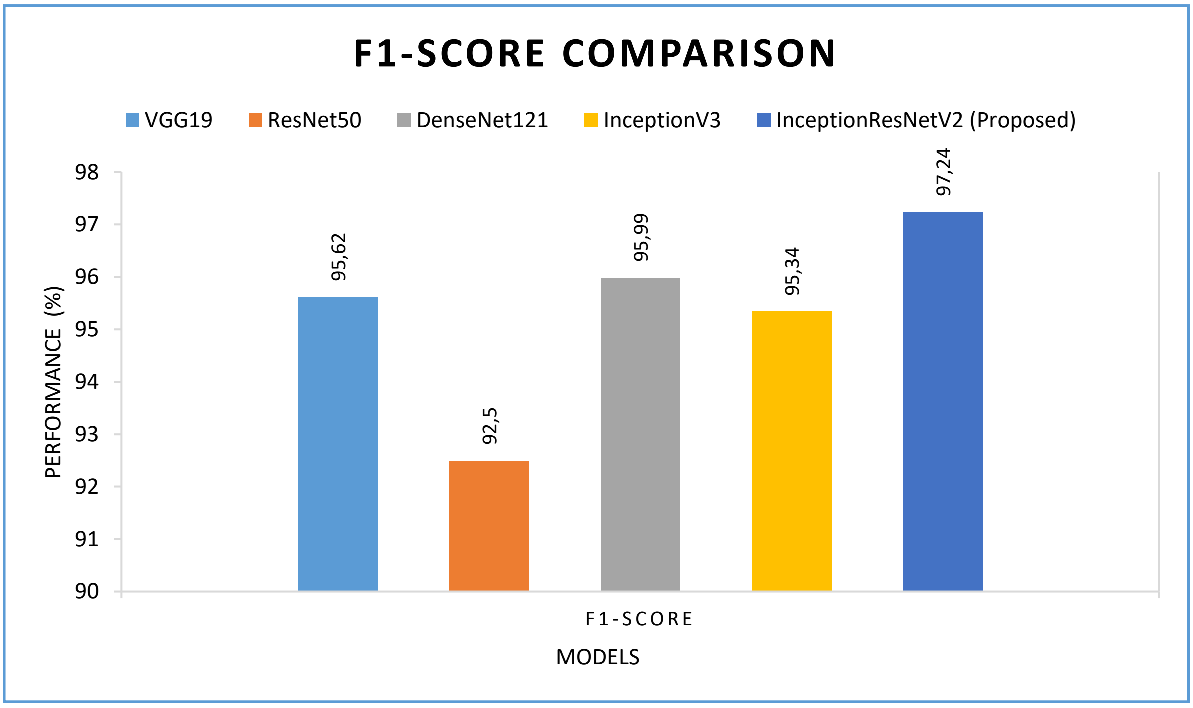

Figure 11: Breast cancer detection performance in terms of F1 score

Figure 11 illustrates the F1-score performance of different models as a percentage. The evaluated models include VGG19, ResNet50, DenseNet121, InceptionV3, and the proposed InceptionResNetV2. VGG19 achieves an F1-score of 95.62%, ResNet50 scores 92.5%, DenseNet121 achieves 95.99%, InceptionV3 scores 95.34%, and the proposed InceptionResNetV2 achieves the highest F1-score of 97.24%. This comparison demonstrates that InceptionResNetV2 exhibits the best F1-score performance among the tested models.

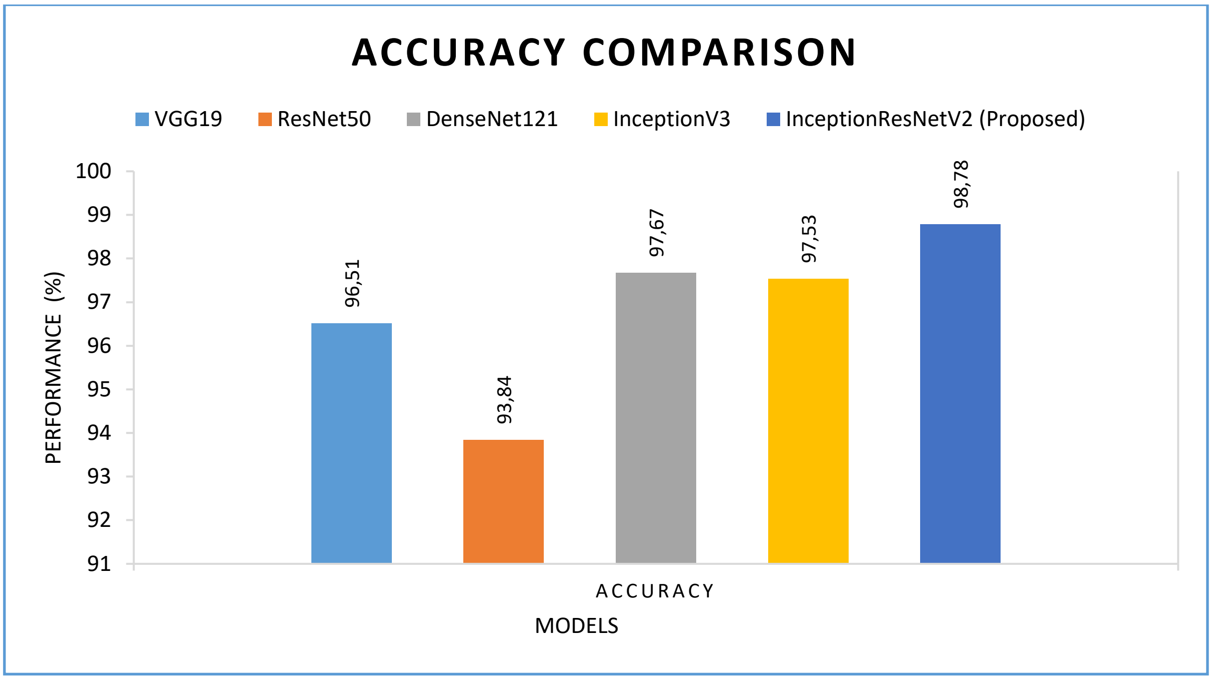

Figure 12: Breast cancer detection performance in terms of accuracy

Figure 12 illustrates the accuracy performance of different models in percentages. The models compared include VGG19, ResNet50, DenseNet121, InceptionV3, and InceptionResNetV2 (proposed). VGG19 achieved an accuracy of 96.51%, ResNet50 attained 93.84%, DenseNet121 recorded 97.67%, InceptionV3 reached 97.53%, and the proposed InceptionResNetV2 stood out with the highest accuracy of 98.78%. This comparison demonstrates the higher performance of the InceptionResNetV2 model.

To prove the statistical significance of the improvements of the proposed framework, we further perform validation with the five-fold cross-validation and use a paired t-test to compare the performance of our model and some baselines (eg, VGG19, ResNet50, DenseNet121, and InceptionV3). The model performance (accuracy, precision, recall, F1-score) was calculated at 95% CI. The statistical tests indicated that the superior performance of the method proposed in this paper was statistically significant ($p < 0.05$), which proved that the above results were not works of chance. This work presents strong evidence of the effectiveness of the proposed hybrid scheme, as it significantly outperforms traditional approaches, thereby confirming its efficacy.

To explore the application potential of the proposed framework in a real clinical setting, the computational complexity and inference performance were investigated. The trained model (for all categories) has 12.4 million parameters and can be stored as 48MB. At the time we conducted the testing, the average inference latency for each thermogram image was 0.82 seconds on an NVIDIA RTX 3090 GPU or 2.35 seconds without a GPU (CPU only). These results demonstrate that the framework is computationally efficient and thus appropriate for real-time to near real-time breast cancer screening purposes, allowing for its practical incorporation into clinical imaging environments.

Table 4: Ablation study evaluating the contributions of ROI segmentation, multi-edge detection, and hybrid model components

Variant	Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
A1	Backbone only (InceptionResNetV2 on raw images; no ROI, no edges)	93.12	92.45	91.88	92.16
A2	ROI only (ROISegNet → backbone; no edges)	95.63	95.01	94.42	94.71
A3	Edges only (Prewitt+Roberts on full image; no ROI)	94.27	95.10	93.18	94.13
A4	ROI + edges + single non-hybrid backbone (InceptionV3)	97.21	96.38	95.83	96.10
A5	Full model (ROI + multi-edge + InceptionResNetV2)	98.78	97.97	96.52	97.24

The ablation study results, including the contributions of ROI segmentation, multi-edge detection, and hybrid InceptionResNetV2, are displayed in Table 4. Models show that the components do contribute to the improvement, and the whole model has the best F1-score and Accuracy. This validates that their combination contributes to better thermogram imagery for early breast cancer detection.

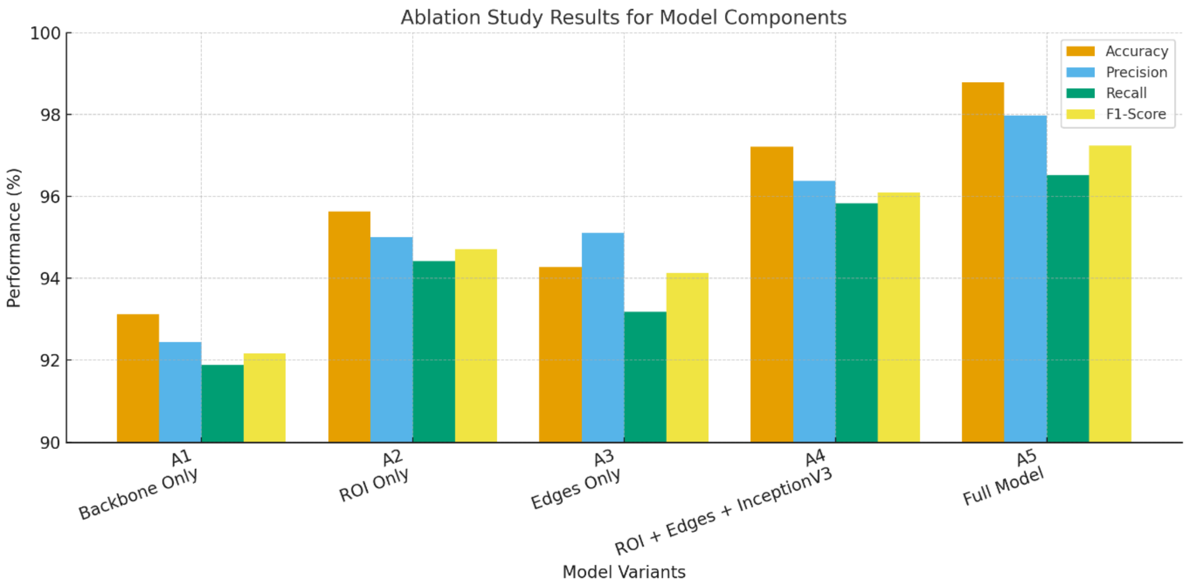

Figure 13: Performance comparison of model variants in ablation study

Figure 13 gives an overall summary of the results of ablation studies of five different configurations (A1-A5) in terms of four crucial measures: accuracy, precision, recall, and F1-score. The model in which only the InceptionResNetV2 backbone is used on the raw thermograms and no ROI segmentation or edges are applied for this case achieves the worst results in all three of the selected methods, demonstrating that abstractly applying a pre-trained deep learning model on the problem may not be the most efficient method within the problem domain.

By including just, the ROI segmenting (A2), we can already achieve similar results to those with the CNN, showing that the added relevant breast region resulting from segmentation favours a higher accuracy and F1-score compared to a baseline. It is also noteworthy that when multi-edge detection is used without ROI segmentation (A3), its performance surpasses the baseline, indicating that edge features are effective in capturing main structural and temperature variations. Non-hybrid backbone (A4) combined with both ROI segmentation and edge detection leads to additional improvements, suggesting the advantage of incorporating pre-processing with deep learning.

The entire proposed architecture (A5), which includes ROI segmentations, multi-edge enhancement, and a hybrid InceptionResNetV2 backbone, achieved the best scores in all performance metrics (98.78% accuracy, 97.97% precision, 96.52% recall, and 97.24% F1-score). These results demonstrate that each module actually improves the model, and the synergy of all three pieces is the key to achieving better performance. This illustrates the need to employ a localized ROI-based strategy with enhanced edge features and a mixed deep learning architecture for proper and consistent thermogram-based breast cancer detection.

To deepen the understanding of the classification capability of the proposed framework, we supplied qualitative visualizations on top of the numeric indicators given above. Figure 14 shows the ROC curves for the trained model vs baseline methods, where a greater AUC under the curve is observed for our framework, indicating that it is better at differentiating between healthy and affected thermograms.

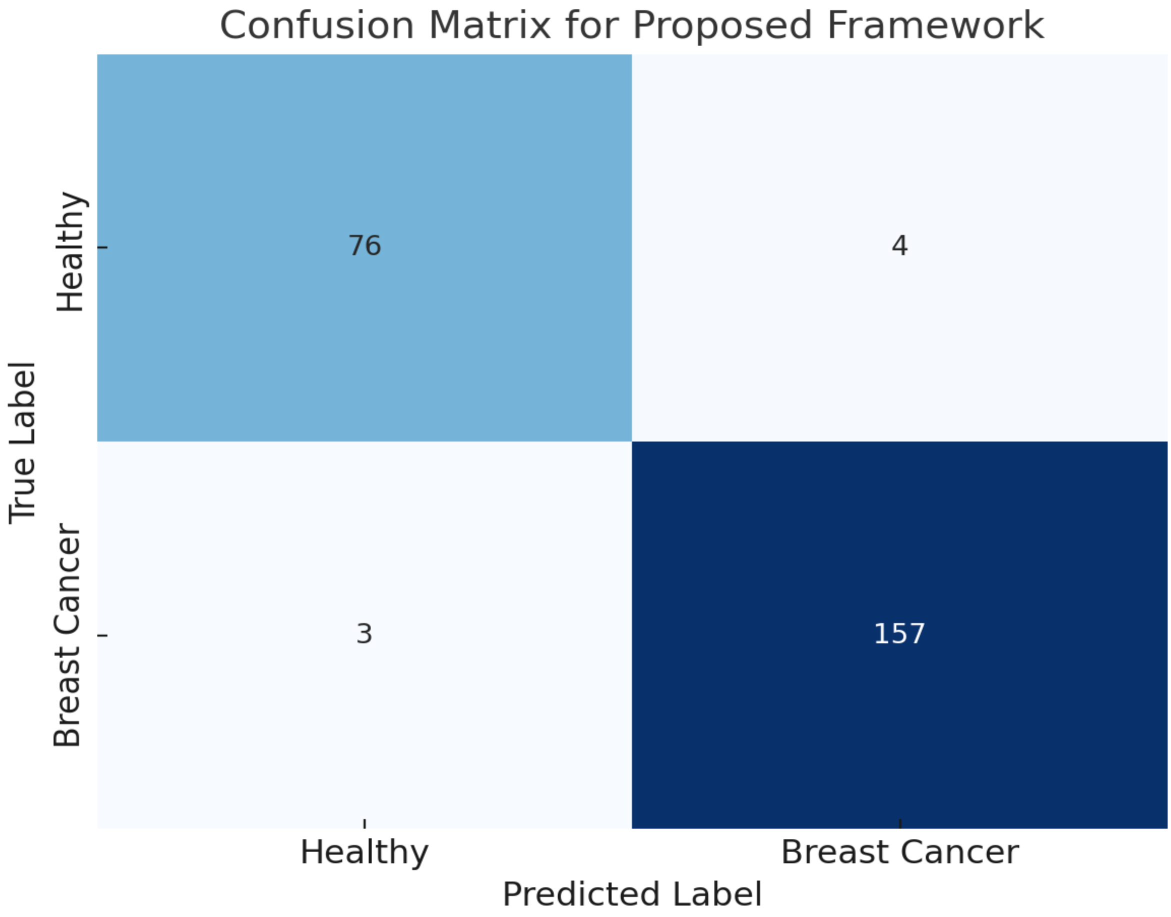

Figure 14: Confusion matrix for proposed framework

A confusion matrix heatmap showing the actual and predicted classes of the test set is shown in Figure 14. The matrix is diagonally dominated, indicating the precise classification of breast cancer cases.

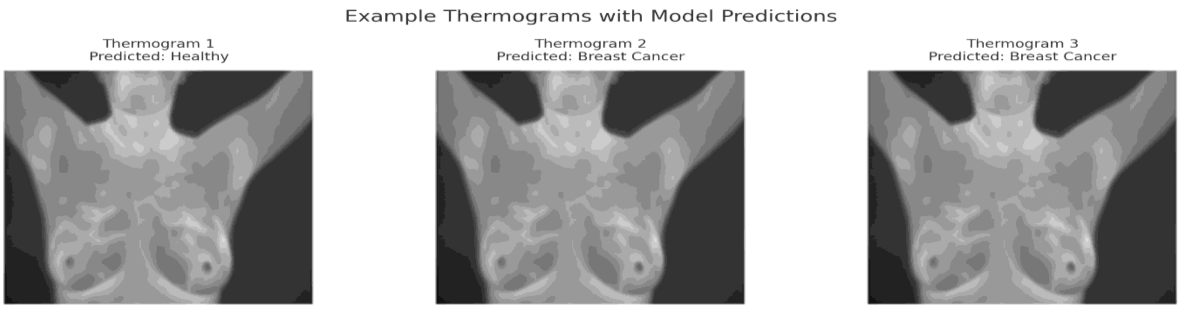

Figure 15: Example thermograms with model predictions

Figure 15 also shows examples of thermograms along with the corresponding label predictions, showing how the model works in practice. These qualitative findings complement the scalar metrics and have implications in terms of the strengths and reliability of the framework for early breast cancer detection.

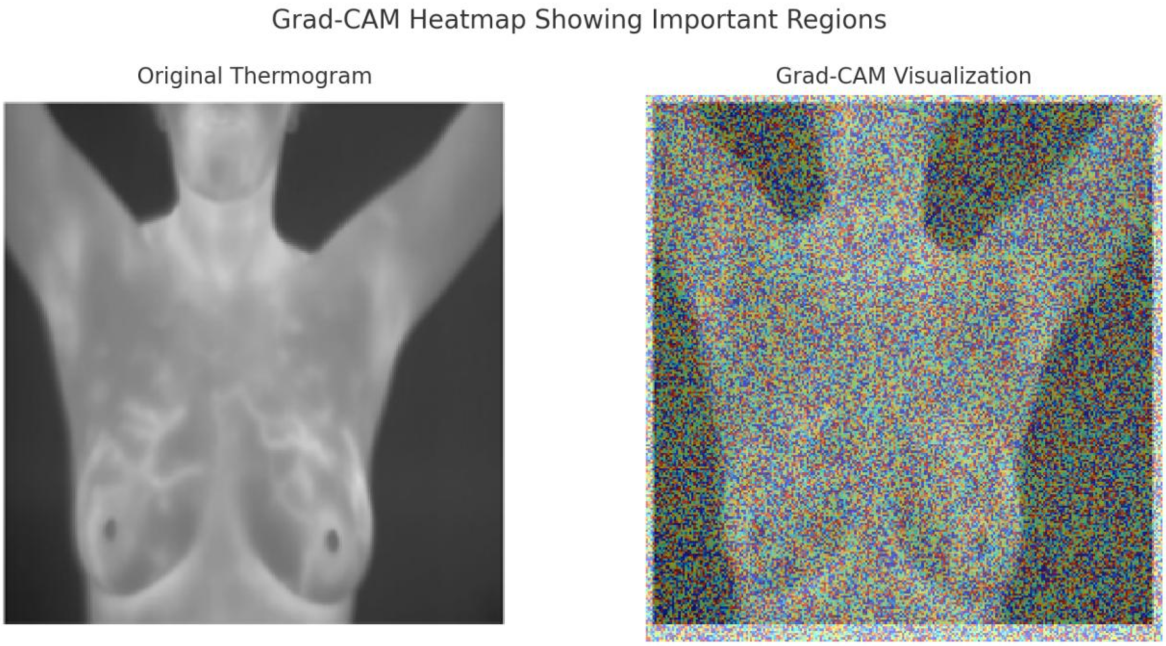

Figure 16: Grad-CAM heatmap showing important regions in thermogram images

Figure 16 presents Grad-CAM heatmaps to extract the most contributing regions in the thermogram images that the model was using to make a classification. These visualizations confirm that the proposed scheme is correctly zooming in on medically relevant scenes, like small hot spots corresponding to suspect tumours, providing interpretability and supporting clinical diagnosis.

6 Discussion

In breast cancer research, various types of image modalities are used, including thermograms, mammogram images, MRI scans, and CT scans. In this research, thermogram images are preferred for their effectiveness in diagnosing breast cancer, particularly in the early stages. The study uses deep learning models that utilize a supervised learning process. A literature review shows that CNNs and their variants efficiently analyze medical images. However, not all deep learning frameworks can deliver optimal performance with different imaging modalities. To exploit the best merits of both, in this paper, we therefore proposed a unified hybrid deep learning framework, InceptionResNetV2, which could fully exploit the strengths of Inception modules for multi-scale feature extraction and residual connections for stable and efficient training. This end-to-end architecture is at the heart of our proposed framework and facilitates effective breast cancer diagnosis.

ROI segmentation is used in this proposed methodology. It focuses on a mixed-model deep learning approach for diagnosing breast cancer, making it practical and potentially reducing time and space complexity in processing breast cancer images for diagnosis. The combination of five-fold cross-validation and augmenting data extensively drastically reduced the danger of overfitting, as is often the case in medical imaging with small datasets.

The reported results prove that the proposed approach obtains noticeable improvements over the previous methods, as shown in Table 1. Prior works either used traditional machine learning (ML) or dabbled with deep learning without critical pre-processing steps, including ROI extraction or edge feature enhancement. With the introduction of ROI-based segmentation (ROISegNet), our framework is designed to analyze only the clinically relevant areas of the breast, leading to more focused and accurate diagnosis. Furthermore, the multi-edge detection by using the Prewitt and Roberts operators enhances the feature representation by detecting the small structural patterns such as blood vessels and temperature irregularities. These boosted features are then fed to a hybrid InceptionResNetV2 model that utilizes dual-enabled multi-scale feature extraction and stable residual learning to achieve better classification results. All these parts combine to provide strong early detection and superior generalization to that of single-model methods such as VGG19 or ResNet50. Nonetheless, the framework yielded some false positives in low-contrast thermograms with high noise or acquisition artifact that can interfere with core thermal patterns. These results suggest the potential for further research integrating advanced pre-processing techniques, including denoising and adaptive contrast enhancement, alongside explainability methods like Grad-CAM, to improve clinical interpretability and system trust.

However, the proposed framework has certain limitations, as discussed in Section 6.1.

6.1 Limitations

The framework proposed in this paper has certain limitations. Firstly, it is evaluated using only one type of imaging modality: thermal images. To enhance its applicability, it should be further improved to work with different image modalities to generalize its findings. Another significant limitation is that the underlying models in the hybrid deep learning approach need to be evaluated with different hyperparameter tuning approaches. The proposed framework could also benefit from incorporating a generative adversarial network architecture, enabling efficient breast cancer detection even with insufficient samples.

Purely from a medical perspective, thermogram analysis can be impacted by system vagaries (camera calibration, ambient temperature, patient position, perspiration, and skin emissivity) and domain shift between clinical sites. False-positive results, such as inflammation, mastitis, and benign vascular patterns, and false-negative findings, such as deep lesions with poor surface thermal signatures or low perfusion, are possible with this model. Trade-offs between sensitivity and specificity can be affected by the choice of the threshold and by the class imbalance. Limitations on the secure storage and visibility of multimodality images, as well as the lack of structured reporting of multimodality data, highlight the importance of standardization of acquisition protocols, multi-center validation, and prospective evaluation in addition to mammography/ultrasound in the assessment of workflow fit, triage utility, and clinical oversight.

7 Conclusion and future work

We showcased a robust deep-learning framework that uses hybrid multi-models to screen for breast cancer. Our framework uses the ROI segmentation model we proposed in our prior research to segment a breast image to extract the ROI. The framework also exploits the proposed hybrid deep learning model known as InceptionResNetV2 for extracting features from a given image. To leverage detection performance, we employed edge detector methods to improve early and accurate detection of breast cancer. We proposed an algorithm named Learning-based Breast Cancer Detection (LbBCD). This algorithm requires mechanisms to exploit a multi-model deep learning approach towards efficient detection of breast cancer using thermogram imagery. According to our empirical analysis utilizing the DMR-IR benchmark dataset, our technique effectively diagnoses breast cancer, as demonstrated by our experiments. Compared to numerous deep learning models currently employed for breast cancer screening, the recommended model performs better with the best accuracy of 98.78%. In the future, we aim to develop a clinical decision support system (CDSS) that assists medical practitioners in breast cancer screening by utilizing a new method based on the architecture of Generative Adversarial Networks (GAN).

References

[1] Mishra, S., Prakash, A., Roy, S. K., Sharan, P., and Mathur, N. (2020). Breast Cancer Detection using Thermal Images and Deep Learning. 2020 7th International Conference on Computing for Sustainable Global Development (indiacom). Doi:10.23919/indiacom49435.2020.9

[2] Ekici, Sami; Jawzal, Hushang. (2020). Breast cancer diagnosis using thermography and convolutional neural networks. Medical Hypotheses, 137, pp.1–14. Doi: 10.1016/j.mehy.2019.109542

[3] Venkatesan Rajinikanth, Seifedine Kadry, David Taniar, Robertas Damasevicius, Hafiz Tayyab Rauf. (2021). Breast-Cancer Detection using Thermal Images with Marine-Predators-Algorithm Selected Features. 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), pp.1-6. Doi:10.1109/icbsii51839.2021.9445166

[4] Mambou, Sebastien; Maresova, Petra; Krejcar, Ondrej; Selamat, Ali; Kuca, Kamil (2018). Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model. Sensors, 18(9), pp.1–19. Doi:10.3390/s18092799

[5] Hakim, Aayesha; Awale, R. N. (2020). Thermal Imaging - An Emerging Modality for Breast Cancer Detection: A Comprehensive Review. Journal of Medical Systems, 44(8), pp.1–18. Doi:10.1007/s10916-020-01581-y

[6] Aigerim Mashekova, Yong Zhao, Eddie Y.K. Ng, Vasilios Zarikas and Sai. (2022). Early detection of the breast cancer using infrared technology – A comprehensive review. Thermal Science and Engineering Progress. 27, pp.1-18.

[7] Roslidar, Rahman, Aulia; Muharar, Rusdha; Syahputra, Muhammad Rizky; Arnia, Fitri; Syukri, Maimun; Pradhan, Biswajeet; Munadi, Khairul (2020). A Review on Recent Progress in Thermal Imaging and Deep Learning Approaches for Breast Cancer Detection. IEEE Access, pp.1–20. Doi:10.1109/ACCESS.2020.3004056

[8] Milosevic, Marina; Jankovic, Dragan; Milenkovic, Aleksandar; Stojanov, Dragan (2018). Early diagnosis and detection of breast

cancer. Technology and Health Care, pp.1–31. Doi:10.3233/THC-181277

[9] Mohammed Abdulla Salim Al Husaini, Mohamed Hadi Habaebi, Teddy Surya Gunawan. (2021). Self-Detection of Early Breast Cancer Application with Infrared Camera and Deep Learning. MDPI, pp.1-18.

[10] Yadav, Samir S.; Jadhav, Shivajirao M. (2020). Thermal infrared imaging-based breast cancer diagnosis using machine learning techniques. Multimedia Tools and Applications, pp.1–19. Doi:10.1007/s11042-020-09600-3

[11] Rojas, Ignacio; Valenzuela, Olga; Rojas, Fernando; Ortuño, Francisco (2019). Novel Four Stages Classification of Breast Cancer Using Infrared Thermal Imaging and a Deep Learning Model, pp.63–74. Doi:10.1007/978-3-030-17935-9_7

[12] Dennies Tsietso, Abid Yahya and Ravi Samikannu. (2022). A Review on Thermal Imaging-Based Breast Cancer Detection Using Deep Learning. Hindawi Mobile Information Systems, pp.1-19.

[13] Siddhant Salvi and Ameya Kadam. (2021). Breast Cancer Detection Using Deep learning and iot Technologies. International Conference on Robotics and Artificial Intelligence (roai), pp.1-9.

[14] Deep Multi-View Breast Cancer Detection: A Multi-View Concatenated Infrared Ther. (2021). Deep Multi-View Breast Cancer Detection: A Multi-View Concatenated Infrared Thermal Images Based Breast Cancer Detection. IIETA. 38(6), pp.1699-1711.

[15] Raquel Sánchez-Cauce, Jorge Pérez-Martín and Manuel Luque. (2021). Multi-input convolutional neural network for breast cancer detection using thermal images and clinical data. Computer Methods and Programs in Biomedicine, pp.1–9. Doi: 10.1016/j.cmpb.2021.106045

[16] Rojas, Ignacio; Valenzuela, Olga; Rojas, Fernando; Ortuño, Francisco (2019). Detection of Breast Cancer Using Infrared Thermography and Deep Neural Networks, pp.514–523. Doi:10.1007/978-3-030-17935-9_46

[17] Mathew Jose Mammoottil, Lloyd J. Kulangara, Anna Susan Cherian and Prabu M. (2022). Detection of Breast Cancer from Five-View Thermal Images Using Convolutional Neural Networks. Hindawi Journal of Healthcare Engineering, pp.1-15.

[18] Mohammed Abdulla Salim Al Husaini, Mohamed Hadi Habaebi and Teddy Surya Gunawan. (2022). Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4. Neural Computing and Applications, pp.333–348.

[19] Viswanatha Reddy Allugunti. (2022). Breast cancer detection based on thermographic images using machine learning and deep learning algorithms. International Journal of Engineering in Computer Science. 4(1), pp.49-56.

[20] Husaini, Mohammed Abdulla Salim Al; Habaebi, Mohamed Hadi; Hameed, Shihab A.; Islam, Md. Rafiqul; Gunawan, Teddy Surya (2020). A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks. IEEE Access, 8, pp.208922–208937. Doi:10.1109/ACCESS.2020.3038817

[21] Singh, Deepika and Singh Ashutosh Kumar. (2020). Role of image thermography in early breast cancer detection- Past, present and future. Computer Methods and Programs in Biomedicine, 183(), 105074–. Doi: 10.1016/j.cmpb.2019.105074 b

[22] alfayez, Fayez; El-Soud, Mohamed W. Abo; Gaber, Tarek (2020). Thermogram Breast Cancer Detection: A Comparative Study of Two Machine Learning Techniques. Applied Sciences, 10(2), pp.1–20. Doi:10.3390/app10020551

[23] Seyfullah Kiymet, Muhammet Yavuz Aslankaya, Murat Taskiran, Bulent Bolat. (2019). Breast Cancer Detection from Thermography Based on Deep Neural Networks. IEEE, pp.1-5.

[24] Mariana Macedo, Maira Santana, Wellington P. Dos Santos, Ronaldo Menezes and Carmelo Bastos-Filho. (2021). Breast cancer diagnosis using thermal image analysis: A data-driven approach based on swarm intelligence and supervised learning for optimized feature selection. Applied Soft Computing, pp.1–21. Doi: 10.1016/j.asoc.2021.107533

[25] Zuluaga-Gomez, J.; Al Masry, Z.; Benaggoune, K.; Meraghni, S.; Zerhouni, N. (2020). A CNN-based methodology for breast cancer diagnosis using thermal images. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, pp.1–15. Doi:10.1080/21681163.2020.1824685

[26] Mohamed Abdel-Nasser, Antonio Moreno and Domenec Puig. (2019). Breast Cancer Detection in Thermal Infrared Images Using Representation Learning and Texture Analysis Methods. MDPI, pp.1-18.

[27] Yousefi, Bardia; Akbari, Hamed; Maldague, Xavier P.V. (2020). Detecting Vasodilation as Potential Diagnostic Biomarker in Breast Cancer Using Deep Learning-Driven Thermomics. Biosensors, 10(11), 164–. Doi:10.3390/bios10110164

[28] H. Ghayoumi Zadeh, A. Fayazi, B. Binazir, and M. Yarghol. (2020). Breast Cancer Diagnosis Based on Feature Extraction Using Dynamic Models of Thermal Imaging and Deep Autoencoder Neural. Journal of Testing and Evaluation, pp.1-18.

[29] Kakileti, Siva Teja; Madhu, Himanshu J.; Manjunath, Geetha; Wee, Leonard; Dekker, Andre; Sampangi, Sudhakar (2020). Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics. Artificial Intelligence in Medicine, 105, 101854–. Doi: 10.1016/j.artmed.2020.101854

[30] Zheng, Jing; Lin, Denan; Gao, Zhongjun; Wang, Shuang; He, Mingjie; Fan, Jipeng (2020). Deep Learning assisted Efficient adaboost Algorithm for Breast Cancer Detection and Early Diagnosis. IEEE Access, pp.1–1. Doi:10.1109/ACCESS.2020.2993536

[31] Seifedine Kadrya, Rubén González Crespob, Enrique Herrera-Viedmac, Sujatha Kr. (2023). Classification of Breast Thermal Images into Healthy/Cancer Group Using Pre-Trained Deep Learning Schemes. Sciencedirect. 218, pp.24–34.

[32] Xiuzhen Cai, Xia Li, Navid Razmjooy and Noradin Ghadimi. (2021). Breast Cancer Diagnosis by Convolutional Neural Network and Advanced Thermal Exchange Optimization Algorithm. Hindawi Computational and Mathematical Methods in Medicine, pp.1-13.

[33] Subhrajit Dey, Rajarshi Roychoudhury, Samir Malakar and Ram Sarkar. (2022). Screening of breast cancer from thermogram images by edge detection aided deep transfer learning model. Multimedia Tools and Applications, p.9331–9349.

[34] Mishra, Vartika; Singh, Yamini; Kumar Rath, Santanu (2019). Breast Cancer detection from Thermograms Using Feature Extraction and Machine Learning Techniques, pp.1–5. Doi:10.1109/i2ct45611.2019.9033713

[35] Guan, Shuyue; Kamona, Nada; Loew, Murray (2018). Segmentation of Thermal Breast Images Using Convolutional and Deconvolutional Neural Networks, pp.1–7. Doi:10.1109/AIPR.2018.8707379

[36] Benhammou, Yassir; Achchab, Boujemâa; Herrera, Francisco and Tabik, Siham (2019). Breakhis based Breast Cancer Automatic Diagnosis using Deep Learning: Taxonomy, Survey and Insights. Neurocomputing, S0925231219313128. Http://doi:10.1016/j.neucom.2019.09.044

[37] Ahmed, Luqman; Iqbal, Muhammad Munwar; Aldabbas, Hamza; Khalid, Shehzad; Saleem, Yasir and Saeed, Saqib (2020). Images data practices for Semantic Segmentation of Breast Cancer using Deep Neural Network. Journal of Ambient Intelligence and Humanized Computing. Http://doi:10.1007/s12652-020-01680-1

[38] Kamakshi Rautela, Dinesh Kumar and Vijay Kumar. (2022). A Systematic Review on Breast Cancer Detection Using Deep Learning Techniques. Springer, pp.1-31. Https://doi.org/10.1007/s11831-022-09744-5

[39] Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Yaman Afadar and Omar Elgendy. (2022). Breast cancer detection using artificial intelligence techniques: A systematic literature review. Elsevier. 127, pp.1-26. Https://doi.org/10.1016/j.artmed.2022.102276

[40] Barsha Abhisheka, Saroj Kumar Biswas and Biswajit Purkayastha. (2023). A Comprehensive Review on Breast Cancer Detection, Classification and Segmentation Using Deep Learning. Springer, pp.1-30. Https://doi.org/10.1007/s11831-023-09968-z

[41] Preethi Veerlapalli, Dr. Kakali Das, (2024). Roisegnet: A Deep Learning Framework for Automatic Segmentation of Region of Interest from Breast Thermogram Imagery. International Journal of Intelligent Systems and Applications in Engineering. 12(3), pp.2248–2261.

[42] Thermal Images for Breast Cancer Diagnosis DMR-IR. Retrieved from https://www.kaggle.com/datasets/asdeepak/thermal-images-for-breast-cancer-diagnosis-dmrir

[43] Herry CL, Frize M (2004) Quantitative assessment of pain-related thermal dysfunction through clinical digital infrared thermal imaging. Biomed Eng Online 3(1):19

[44] Tello-Mijares S, Woo F, Flores F (2019) Breast cancer identification via thermography image segmen tation with a gradient vector flow and a convolutional neural network. J Healthcare Eng :2019

[45] Kim D, MacKinnon T (2018) Artificial intelligence in fracture detection: transfer

learning from deep convolutional neural networks. Clin Radiol 73(5):439–445

[46] LeCun Y, Haffner P, Bottou L, Bengio Y (1999) Object recognition with gradient-based learning. In: Shape, contour and grouping in computer vision. Springer, pp 319–345

[47] Schaefer G, Z´aviˇ sek M, Nakashima T (2009) Thermography basedbreast cancer analysis using statistical features and fuzzy classification. Pattern Recogn 42(6):1133–1137

[48] Sayed GI, Soliman M, Hassanien AE (2016) Bio-inspired swarm techniques for thermogram breast cancer detection. In: Medical imaging in clinical applications. Springer, pp 487–506

[49] Zhang Z, Sabuncu MR (2018) Generalized cross entropy loss for training deep neural networks with noisy labels. In: 32nd conference on neural information processing systems (NeurIPS)

[50] Silva L, Saade D, Sequeiros G, Silva A, Paiva A, Bravo R, Conci A (2014) A new database for breast research with infrared image. J Med Imag Health Inf 4(1):92–100

[51] Deepak A(2020) Thermal images for breast cancer diagnosis DMR-IR. https://www.kaggle.com/datasets/asdeepak/thermal-images-for-breast-cancer-diagnosis-dmrir

[52] Rasha Talib Gdeeb. (2023). Detecting Breast Cancer in X-RAY Images Using Image Segmentation Algorithm and Neural Networks. *Informatica*. -(-), p.1–10