



OPEN

Integrative hybrid deep learning for enhanced breast cancer diagnosis: leveraging the Wisconsin Breast Cancer Database and the CBIS-DDSM dataset

Patnala S. R. Chandra Murty¹, Chinta Anuradha², P. Appala Naidu³, Deenababu Mandru⁴, Maram Ashok⁵, Athiraja Atheeswaran⁷, Nagalingam Rajeswaran⁸ & V. Saravanan⁶✉

The objective of this investigation was to improve the diagnosis of breast cancer by combining two significant datasets: the Wisconsin Breast Cancer Database and the CBIS Curated Breast Imaging Subset (CBIS-DDSM). The Wisconsin Breast Cancer Database provides a detailed examination of the characteristics of cell nuclei, including radius, texture, and concavity, for 569 patients, of which 212 had malignant tumors. In addition, the CBIS-DDSM dataset—a revised variant of the Digital Database for Screening Mammography (DDSM)—offers a standardized collection of 2,620 scanned film mammography studies, including cases that are normal, benign, or malignant and that include verified pathology data. To identify complex patterns and trait diagnoses of breast cancer, this investigation used a hybrid deep learning methodology that combines Convolutional Neural Networks (CNNs) with the stochastic gradients method. The Wisconsin Breast Cancer Database is used for CNN training, while the CBIS-DDSM dataset is used for fine-tuning to maximize adaptability across a variety of mammography investigations. Data integration, feature extraction, model development, and thorough performance evaluation are the main objectives. The diagnostic effectiveness of the algorithm was evaluated by the area under the Receiver Operating Characteristic Curve (AUC-ROC), sensitivity, specificity, and accuracy. The generalizability of the model will be validated by independent validation on additional datasets. This research provides an accurate, comprehensible, and therapeutically applicable breast cancer detection method that will advance the field. These predicted results might greatly increase early diagnosis, which could promote improvements in breast cancer research and eventually lead to improved patient outcomes.

Keywords Wisconsin Breast Cancer Database, CBIS-DDSM, CNN, AUC-ROC

Breast cancer continues to be among the most prevalent and severe types of cancer that affect women globally. Patient survival and effective treatment depend on early discovery¹. Conventional diagnostic techniques, such as physical examinations and mammography, might produce false negatives or positives because of their limited accuracy². With the development of artificial intelligence and machine learning, there is an enormous opportunity to improve the precision and effectiveness of breast cancer diagnosis. To improve breast cancer detection, this investigation presents a hybrid deep learning strategy that utilizes the Wisconsin Breast Cancer Database and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM). Together with innovative computational techniques,

¹Department of CSE, Malla Reddy Engineering College (Autonomous), Maisammaguda, Secunderabad 500100, Telangana, India. ²Department of CSE, Velagapudi Ramakrishna Siddhartha Engineering College (Deemed to be University), Kanuru, Vijayawada 520007, Andhra Pradesh, India. ³Department of CSE, Raghu Engineering College (Autonomous), Visakhapatnam 531162, Andhra Pradesh, India. ⁴Department of IT, Malla Reddy Engineering College (Autonomous), Maisammaguda, Secunderabad 500100, Telangana, India. ⁵Department of CSE, Malla Reddy College of Engineering, Maisammaguda, Secunderabad 500100, Telangana, India. ⁶Department of Computer Science, Dambi Dollo University, Dambi Dollo, Ethiopia. ⁷Department of CSE (AIML), Malla Reddy College of Engineering, Secunderabad, India. ⁸Department of IQAC, IMS Unison University, Dehradun, India. ✉email: saravanan@dadu.edu.et

the combination of these two datasets, each with distinct and complimentary qualities, represents a significant advancement in the struggle against this prevalent disease.

Background

Breast cancer continues to be a major worldwide health problem; thus, improved outcomes for patients will require ongoing advances in detection techniques. Innovative techniques have been created by combining computational methods and modern imaging technology³. One such promising strategy is the integration of the Wisconsin Breast Cancer Database with the Curated Breast Imaging Subset of DDSM (CBIS-DDSM). The radius, smoothness, and concavity are just a few of the 30 characteristics of the innovative Wisconsin Breast Cancer Database that offer a comprehensive understanding of cellular properties. Among the 569 patients in this dataset, 212 had significant differences because they had malignant tumors⁴. All-encompassing examinations of breast cancer are based on these characteristics, which are important indicators of cellular abnormalities and anomalies. Data augmentation techniques have been widely used to improve the performance of deep learning models in medical image analysis tasks. However, traditional data augmentation methods often to maintain bounding box and segmentation precision.

In addition, the CBIS-DDSM dataset is a modern, standardized development of the Digital Database for Screening Mammography (DDSM). The DDSM, which starts with 2620 scanned film mammography examinations, provides a wide range of cases, from benign to malignant, along with carefully checked pathological data for each case⁵. This repository is updated and refined by the CBIS-DDSM dataset, which serves as a standard for modern breast imaging research. This research used a state-of-the-art hybrid deep learning technique to obtain accurate detection of breast cancer. By combining the advantages of both approaches, Convolutional Neural Networks (CNNs) and the stochastic gradients method enable the model to identify subtle characteristics and complex patterns that may be symptoms of breast cancer. The Wisconsin Breast Cancer Database's rich attribute information is an ideal starting point for the CNN, which may then be further enhanced using the CBIS-DDSM dataset to guarantee flexibility across a range of mammography investigations. Using a hybrid deep learning method and combining these datasets, this investigation intends to overcome the challenges associated with diagnosing breast cancer. The expected results show promise for a more precise, understandable, and practically applicable diagnostic method, which will support further efforts to enhance early detection techniques and improve the prognosis of breast cancer patients.

Previous work

Large-scale imaging datasets and innovative computational techniques have been widely used in previous research on breast cancer diagnosis. Prominent research efforts have established a foundation for the use of deep learning methods, namely, CNNs, to improve the precision and effectiveness of identifying breast cancer⁶. The Wisconsin Breast Cancer Database has been widely used by researchers to investigate the complex relationships among cellular characteristics and breast cancer pathology. The development of machine learning models to distinguish between benign and malignant tumors has been made possible by the 30 properties of the dataset, which include radius, texture, and concavity⁷. Previous research has highlighted the importance of these traits in identifying subtle patterns suggestive of breast cancer, indicating the possibility of improved diagnostic precision.

As an essential tool for breast imaging research, the Digital Database for Screening Mammography (DDSM) offers a wide range of 2,620 scanned film mammography investigations. A significant amount of work has been done previously using DDSM to create models that can identify situations that are benign, malignant, or normal⁸. The significance of authentic pathology data accompanying every instance in guaranteeing the dependability of the training and assessment procedure has been emphasized by these investigations. Recent years have made tremendous progress in the detection of breast cancer, with researchers increasingly utilizing sophisticated computer algorithms and extensive imaging datasets to improve the precision and efficacy of these methods. An innovative technique for identifying more accurate breast cancer diagnoses is the combination of the Wisconsin Breast Cancer Database and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) with a hybrid deep learning algorithm⁹. The main challenge is to develop multimodality process that are low accuracy, and efficient within a clinically acceptable period. The source and target images, the majority of medical image processing with low clarity.

Previous investigations have focused on the Wisconsin Breast Cancer Database, which is well known for its comprehensive features related to cell nuclei. Researchers have investigated the discriminative ability of various characteristics, such as the radius, roughness, and concavity. Notable works include (reference), where machine learning algorithms used this dataset to show encouraging results in differentiating between benign and malignant cancers¹⁰. By providing a wide range of 2,620 scanned film mammogram investigations, DDSM has been a mainstay in breast imaging research. To create models for classifying instances as normal, benign, or malignant, early research (reference) used the DDSM^{11–13}. To guarantee the validity of these models and provide a standard for future studies, confirmed pathology data are essential. The Created Breast Imaging Subset of DDSM (CBIS-DDSM) is a development in the field of imaging. Upgrades and standardization of DDSM have produced a more recent dataset for use in breast imaging research. The work of (Maqsood, S, et al. 2022)¹⁴ demonstrated the importance of this dataset for enhancing model generalizability and guaranteeing relevance to contemporary diagnostic difficulties^{15–19}.

An innovative approach is introduced by integrating a hybrid deep learning algorithm that combines the stochastic gradients method with CNNs. Previous investigations^{20–29} have indicated that hybrid models are effective at extracting complex patterns from various datasets, exhibiting flexibility in various imaging properties and enhanced diagnostic capabilities. Although several datasets and methods have demonstrated potential, there is still a significant lack of comprehensive techniques that include several datasets. Managing bias and overfitting concerns as well as the requirement for interpretability in deep learning models are challenges. The

complementary qualities of the datasets and algorithms could be used to produce more reliable and broadly applicable breast cancer diagnosis systems.

The literature review summarizes the important contributions made by earlier studies in the use of specific datasets and sophisticated techniques for the detection of breast cancer. On that basis, this work used a hybrid deep learning algorithm to integrate the Wisconsin Breast Cancer Database and CBIS-DDSM datasets to solve existing issues and further the continuous development of accurate and efficient breast cancer diagnostic systems.

Significance of the study

Improving diagnostic accuracy: By combining two significant datasets, the Wisconsin Breast Cancer Database and CBIS-DDSM, this work significantly improves the area of breast cancer detection. Together with in-depth mammography data, a thorough examination of 30 cell nuclei properties provides a wealth of data for creating a highly reliable diagnostic model. To improve the health of patients, a hybrid deep learning system that combines a CNN and stochastic gradients is being integrated with the objective of increasing diagnostic accuracy³⁰. The requirement for updated and standardized datasets has been addressed by the integration of the CBIS-DDSM with the Wisconsin Breast Cancer Database. Incorporating a range of examples, such as benign, malignant, and normal cases, provides a thorough portrayal of breast pathology. This approach enhances the dependability of the created diagnostic system while allowing extensive comparisons with other investigations and datasets. The CNN and stochastic gradient hybrid deep learning approach is sufficiently adaptable to accommodate a range of imaging properties observed in both datasets. This flexibility is essential to the model's capacity to recognize inadequate developments across many imaging modalities that are suggestive of breast cancer. Utilizing the advantages of both methods, this research aspires to develop a reliable and efficient diagnostic instrument for use in a variety of mammography investigations.

Ultimately, this investigation aspires to directly influence patient outcomes and therapeutic practices. Early detection can be greatly aided by an accurate and reliable breast cancer testing system, which enables early intervention and treatment. Increased diagnostic precision is especially important for differentiating between benign and malignant cases, minimizing unnecessary procedures and treatments, and eventually increasing the standard of patient care. This investigation advances the field of breast cancer research by establishing an integrated methodology that integrates conventional mammography investigations with intricate cellular characteristics. These results might further the continuing attempts to gain a deeper understanding of medical conditions by offering new insights into the connections between certain traits and breast cancer pathology. Benchmarking and comparative analysis are made possible by the hybrid deep learning method and dataset integration. To demonstrate the superiority of the created model in terms of accuracy and efficiency, the results were compared to those of current diagnostic techniques. Future research can benefit from this comparison study, which can also help choose the best diagnostic instruments for use in clinical settings. In conclusion, the relevance of this approach stems from its potential to aid in the detection of breast cancer. The objective of this project is to contribute to the creation of a reliable and accurate diagnostic system that may have a favourable influence on patient care and outcomes by combining extensive datasets, utilizing cutting-edge deep learning algorithms, and emphasizing clinical relevance. The research gap highlights challenges including prolonged calculation times due to extensive parameter tuning, late convergence, and high complexity. Additionally, the approach has limited scalability with high-dimensional data, as it heavily depends on advanced parameter tuning and numerical optimization.

The article makes the following key contributions to the field of breast cancer diagnosis:

- Developed a novel hybrid deep learning model that combines CNN with stochastic gradient methods, tailored specifically for breast cancer detection.
- Successfully integrated two diverse and significant datasets—the Wisconsin Breast Cancer Database and the CBIS-DDSM Dataset—to enhance the robustness and accuracy of the diagnostic model.
- Conducted an extensive performance analysis using metrics such as accuracy, sensitivity, specificity, and AUC-ROC, demonstrating the model's superiority over existing methodologies.
- Implemented strategies to improve the model's generalizability across different datasets and provided insights into its decision-making process, enhancing its clinical applicability.
- Addressed ethical dimensions related to patient data usage, including privacy and consent, and discussed the model's potential impact and integration within clinical practice.

The article is organized into following key sections: Sect. 2 discusses on the literature review, offering a detailed examination of recent advancements in deep learning techniques and their application in medical imaging. Section 3 then elaborates on the hybrid deep learning model, detailing the integration of the Wisconsin Breast Cancer Database and the CBIS-DDSM Dataset, along with the specific techniques used for data processing, model training, and performance evaluation. Next, Sect. 4 presents the methodology and framework used in the study. Section 5 focuses on the findings of the study, including a thorough performance analysis and comparison with existing methods, alongside a discussion of the model's generalizability, interpretability, and clinical relevance in provided in Sect. 6. Finally, the article concludes in Sect. 7, summarizing the contributions of the study and proposing areas for further research and improvement.

Description of the hybrid deep learning algorithm

In this section, we provide a detailed description of the hybrid deep learning algorithm employed in this study to enhance breast cancer diagnosis. The section begins by outlining the core components of the CNN architecture, which forms the foundation of the model. Then the integration of the SGD method is elaborated, explaining how it is utilized to optimize the model's parameters effectively.

Description of the convolutional neural network algorithm

CNNs are a class of deep learning models specifically designed for processing structured grid data, such as images, as shown in Fig. 1³¹. They have proven highly effective in tasks such as image classification, object detection, and segmentation. The following is a description of the key components of CNNs and the corresponding mathematical equations:

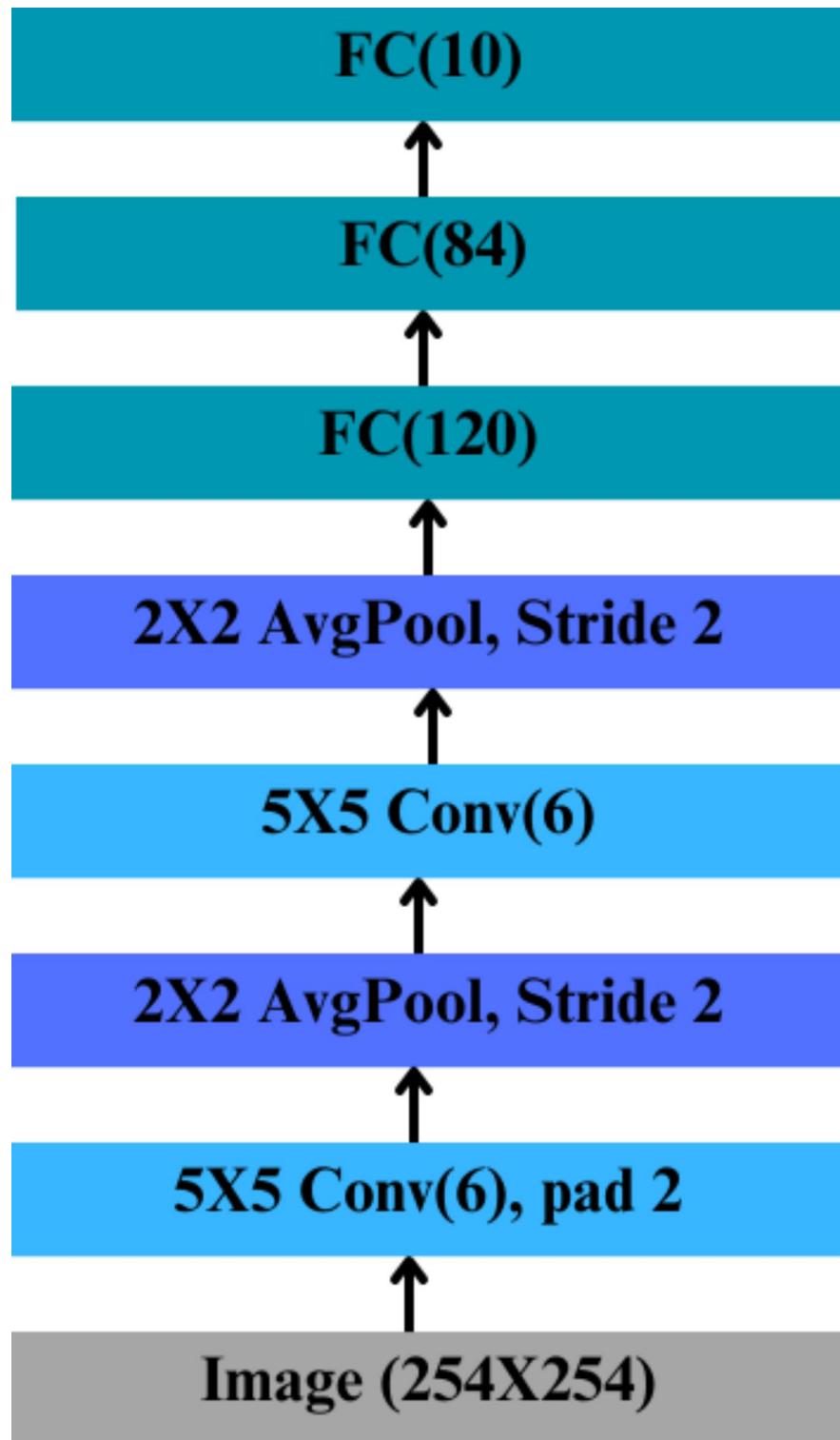


Fig. 1. Convolutional neural network model.

Convolutional layer

The core operation in CNNs is convolution. This operation involves sliding a small window called a kernel or filter over the input data, performing element-wise multiplication, and summing the results to produce a feature map. The convolution operation is expressed as follows:

$$C(i, j) = (Kv * Iv)(i, j) \quad (1)$$

$$C(i, j) = \sum_m \sum_n Iv(i+m, j+n) \cdot Kv(i+m, j+n) \quad (2)$$

where: $C(i, j)$ is the value at position (i, j) in the feature map. $Iv(i+m, j+n)$ is the input value at position $(i+m, j+n)$. $Kv(m, n)$ is the kernel value at position (m, n) .

Activation function

After convolution, an element wise activation function is applied to introduce nonlinearity to the model. Common activation functions include the Rectified Linear Unit (ReLU). The ReLU activation function is defined as follows:

$$f(x) = \max(0, x) \quad (3)$$

where, x is the input value.

Pooling layer

Pooling layers are used to down sample the spatial dimensions of the feature maps, reducing computation and controlling overfitting and applies a value per pixel called intensity or brightness. Max pooling is a common pooling technique defined as follows:

$$Pv(i, j) = \max(Iv(2i, 2j), Iv(2i+1, 2j), Iv(2i+2, 2j), Iv(2i+3, 2j)) \quad (4)$$

where $Pv(i, j)$ is the value at position (i, j) in the pooled feature map.

Fully connected layer

The fully connected layer connects every neuron in one layer to every neuron in the next layer, typically followed by an activation function. The output of a fully connected layer is computed as follows:

$$Op = f(Wm \cdot X + Bi) \quad (5)$$

where, Op is the output. X is the input vector. Wm is the weight matrix. Bi is the bias vector. f is the activation function.

Softmax layer

In classification tasks, a softmax layer is often used to produce probability distributions over multiple classes. The softmax function is defined as follows:

$$Pv(yi) = \frac{e^{yi}}{\sum_j e^{yj}} \quad (6)$$

where $P(yi)$ is the probability of class i and yi is the input to the softmax layer for class i .

These mathematical representations provide a foundation for understanding the operations performed by CNNs, allowing them to effectively learn hierarchical features from input data, particularly in image-based tasks.

Description of CNNs: VGG network

The VGG (Visual Geometry Group) network is a widely recognized architecture for image classification. VGG is known for its simplicity and uniformity, with most layers consisting of 3×3 convolutional filters followed by max pooling layers, as shown in Fig. 2³². Here, is an overview of the VGG network architecture and mathematical equations for its key components:

Convolutional block

The VGG network consists of multiple convolutional blocks, each comprising multiple convolutional layers with small 3×3 filters. The convolutional block can be represented as follows:

$$ConvBlock(X) = Conv(X) \rightarrow Conv(X) \rightarrow Conv(X) \rightarrow MaxPool(X) \quad (7)$$

where X is the input feature map. $Conv$ represents a 3×3 convolutional layer. $MaxPool$ represents a 2×2 max pooling layer.

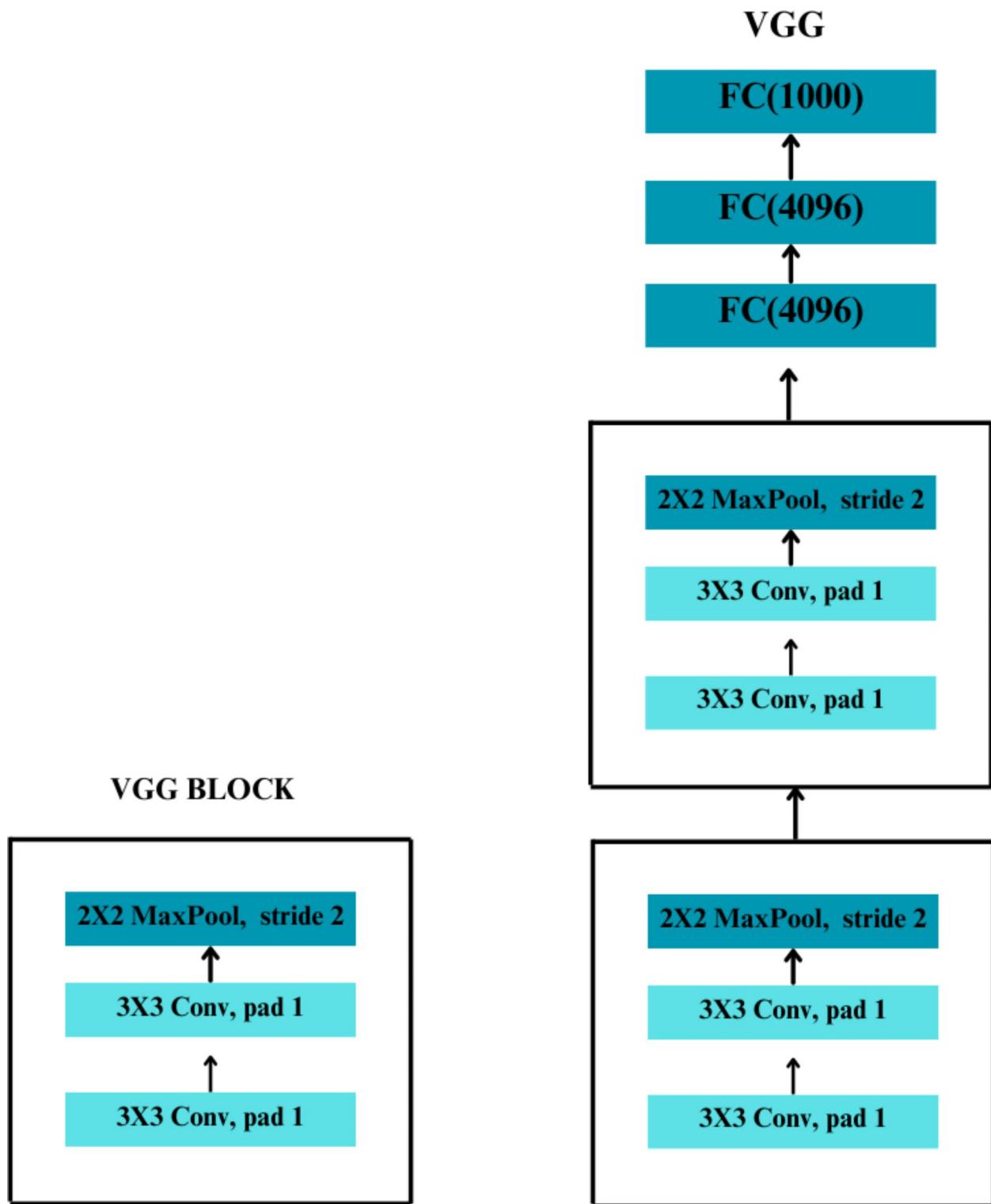


Fig. 2. CNN-VGG architecture.

Fully connected block

Towards the end of the network, fully connected layers are employed for final classification. The fully connected block is expressed as follows:

$$FCBlock(X) = FC(X) \rightarrow FC(X) \rightarrow Softmax(X) \quad (8)$$

where FCFC denotes a fully connected layer. SoftmaxSoftmax is the softmax activation function for classification.

Overall VGG architecture

The complete VGG architecture is a stack of convolutional blocks followed by fully connected blocks. For example, the VGG16 architecture can be represented as follows:

$$VGG(X) = ConvBlock(X) \rightarrow ConvBlock(X) \rightarrow ConvBlock(X) \rightarrow ConvBlock(X) \rightarrow FCBlock(X) \quad (9)$$

Convolutional layer

The convolutional layer can be represented mathematically as follows:

$$Y_{i,j,k} = \sigma \left(\sum_{l=1}^c W_{i,j,l,k} \cdot X_{i,j,l} + B_k \right) \quad (10)$$

where $Y_{i,j,k}$ is the output at position (i, j) in the k -th feature map. $X_{i,j,l}$ is the input at position (i, j) in the l -th input feature map. $W_{i,j,l,k}$ is the weight for the convolutional kernel at position (i, j) for input channel l and output channel k . B_k is the bias term for the k -th feature map. σ is the activation function, commonly referred to as ReLU.

Max-pooling layer

The max pooling layer is represented as follows:

$$Y_{i,j,k} = \max_{(a,b) \in R_{i,j}} X_{a,b,k} \quad (11)$$

where $R_{i,j}$ is the receptive field at position (i, j) in the input feature map.

Fully connected layer

The fully connected layer is expressed as follows:

$$Y_k = \sigma \left(\sum_{l=1}^N W_{k,l} \cdot X_l + B_k \right) \quad (12)$$

where Y_k is the output of the k -th neuron in the fully connected layer. X_l is the input from the l -th neuron in the preceding layer. $W_{k,l}$ is the weight connecting the l -th neuron to the k -th neuron. B_k is the bias term for the k -th neuron. σ is the activation function, often ReLU.

These equations provide a mathematical representation of the key components of the VGG network, offering insights into the operations performed by this architecture during image classification tasks.

Description of CNNs: GoogleNet (inception v1)

GoogleNet, also known as Inception v1, is a deep convolutional neural network developed by Google. It is renowned for introducing the concept of “inception modules,” which utilize multiple convolutional filters of different sizes within the same layer³³. This design allows the network to capture features at various scales efficiently. Here, is an overview of the GoogleNet architecture and mathematical equations for its key components:

Inception module

The inception module is the cornerstone of GoogLeNet and incorporates multiple convolutional filters of different sizes in parallel, as shown in Fig. 3. The mathematical representation of the Inception module can be expressed as:

$$Inception(X) = Concat(Conv1X1(X), Conv3X3(X), Conv5X5(X), MaxPool3X3(X)) \quad (13)$$

where X is the input feature map. $Conv1 \times 1$, $Conv3 \times 3$, and $Conv5 \times 5$ represent convolutional layers with filter sizes of 1×1 , 3×3 , and 5×5 , respectively. $MaxPool3 \times 3$ denotes a max pooling layer with a filter size of 3×3 . $Concat$ signifies the concatenation of the outputs from all these operations.

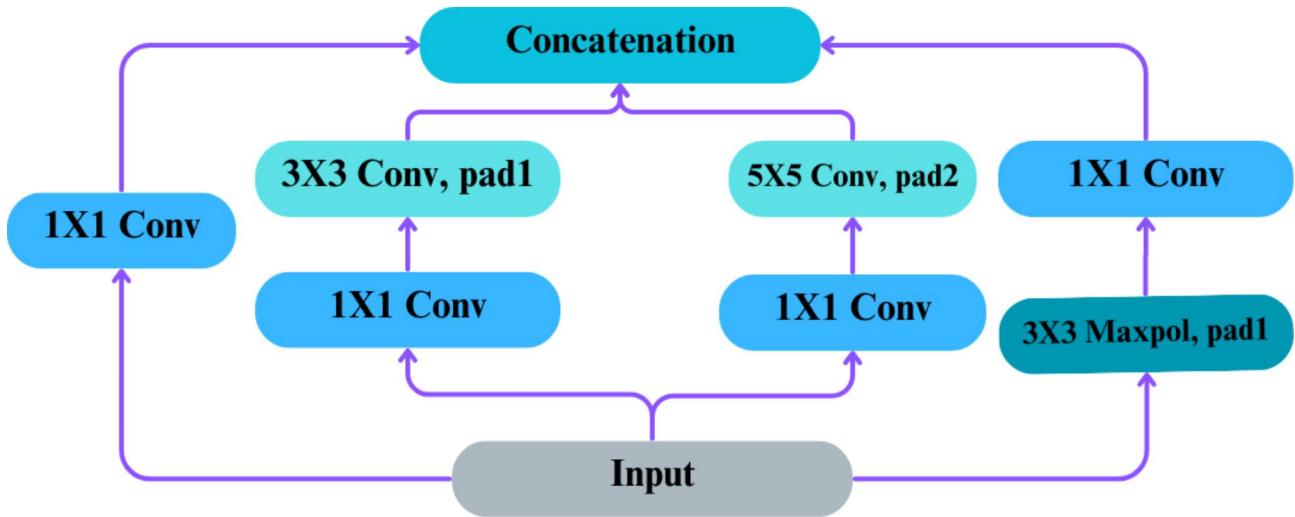
GoogLeNet architecture

The complete GoogLeNet architecture is composed of multiple inception modules stacked together interspersed with other layers, such as max pooling and fully connected layers. A simplified representation of the architecture can be expressed as in Fig. 4:

$$GoogleNet(X) = Conv7X7(X) \rightarrow MaxPool3X3(X) \rightarrow Inception1(X) \rightarrow Inception2(X) \rightarrow \dots \rightarrow FC(X) \quad (14)$$

Convolutional layer

The convolutional layer is represented mathematically as follows:

**Fig. 3.** Inception module.

$$Y_{i,j,m} = \sigma \left(\sum_{l=1}^c W_{i,j,l,m} \cdot X_{i,j,m,l} + B_m \right) \quad (15)$$

where $Y_{i,j,m}$ is the output at position (i, j) in the k -th feature map. $X_{i,j,m,l}$ is the input at position (i, j) in the l -th input feature map. $W_{i,j,l,m}$ is the weight for the convolutional kernel at position (i, j) for input channel l and output channel m . B_m is the bias term for the m -th feature map. σ is the activation function, commonly referred to as ReLU.

Max-pooling layer

The max pooling layer is represented as follows:

$$Y_{i,j,m} = \max_{(a,b) \in R_{i,j}} X_{a,b,m} \quad (16)$$

where $R_{i,j}$ is the receptive field at position (i, j) in the input feature map.

Fully connected layer

The fully connected layer is expressed as follows:

$$Y_m = \sigma \left(\sum_{l=1}^N W_{m,l} \cdot X_l + B_m \right) \quad (17)$$

where Y_m is the output of the m -th neuron in the fully connected layer. X_l is the input from the l -th neuron in the preceding layer. $W_{m,l}$ is the weight connecting the l -th neuron to the m -th neuron. B_m is the bias term for the m -th neuron. σ is the activation function, often ReLU.

These equations provide a foundation for understanding the operations performed by GoogleNet, particularly the Inception modules, enabling effective feature extraction at multiple scales within the network.

Description of CNNs: densely connected Network (DenseNet)

DenseNet is a convolutional neural network architecture that introduces dense connections between layers. In DenseNet, each layer receives input not only from the preceding layer but also from all preceding layers. This dense connectivity is designed to address issues such as vanishing gradients, promote feature reuse, and enhance model compactness³⁴. Here, is an overview of the DenseNet architecture and mathematical equations for its key components:

Dense block

The DenseNet architecture is built around dense blocks, each consisting of multiple densely connected layers, as shown in Fig. 5. Mathematically, the output of a dense block can be expressed as:

$$X_{dense} = \text{Concat}(X_0, H_1, H_2, \dots, H_n) \quad (18)$$

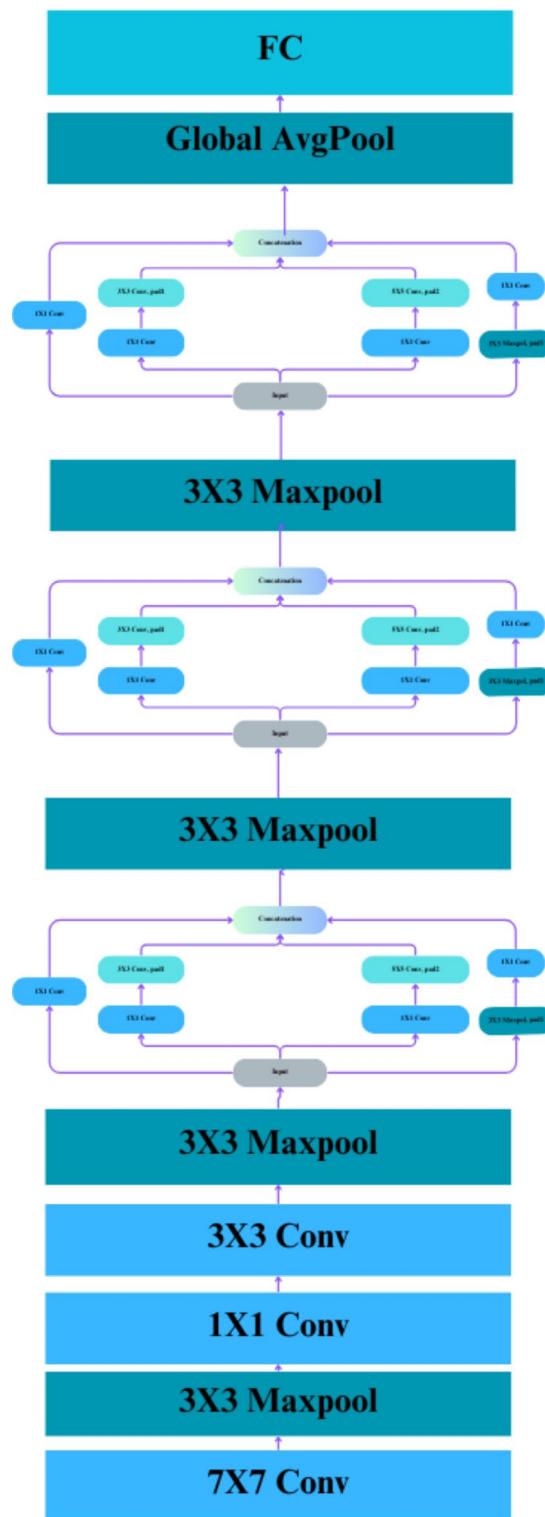


Fig. 4. CNN-GoogLeNet Architecture.

where X_0 is the input to the dense block. H_i is the output of the i -th layer in the dense block. Concat denotes the concatenation operation.

Overall DenseNet architecture

The complete DenseNet architecture is a sequence of dense blocks separated by transition layers, as shown in Fig. 6. A simplified representation of the architecture can be expressed as:

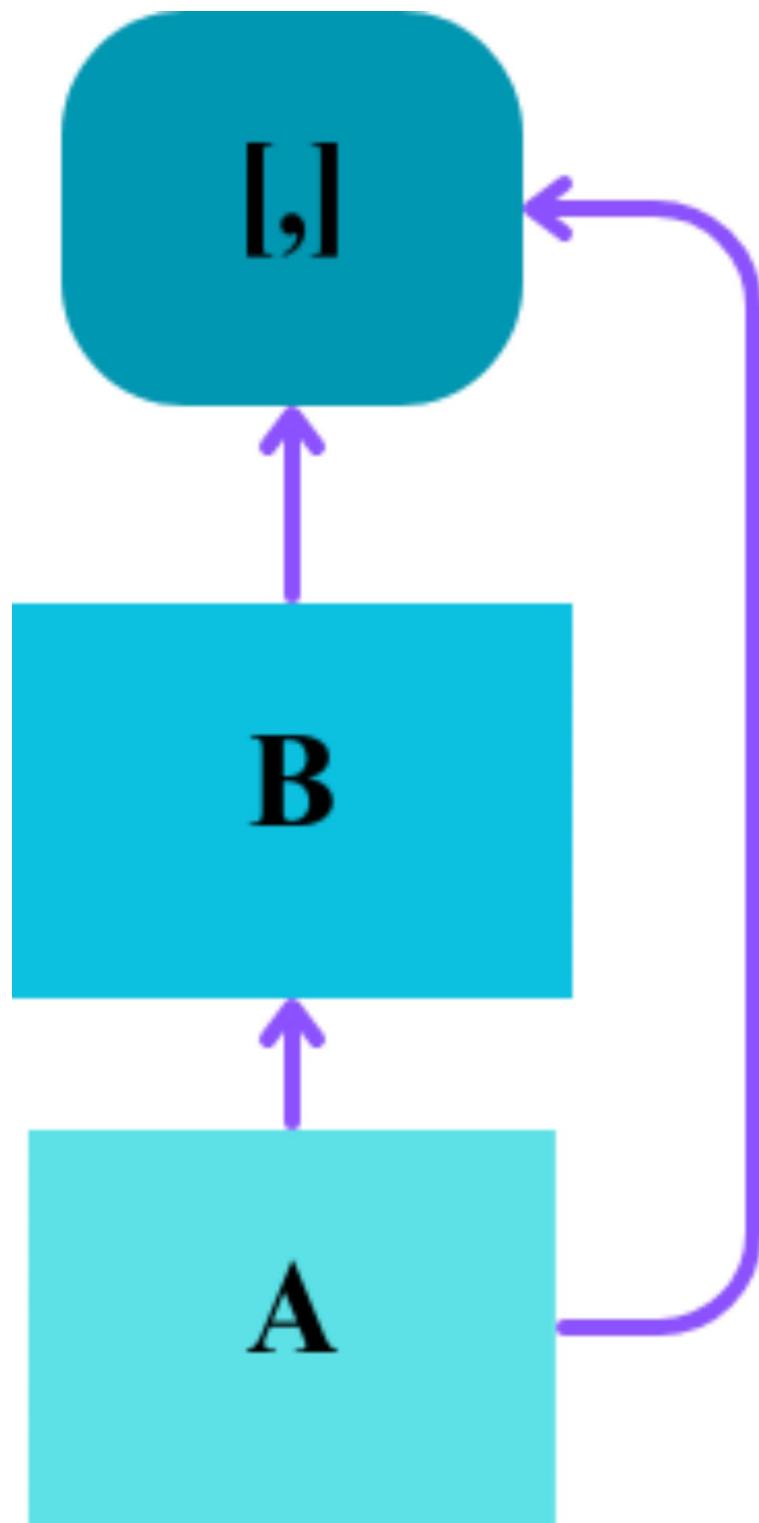


Fig. 5. DenseNet architecture.

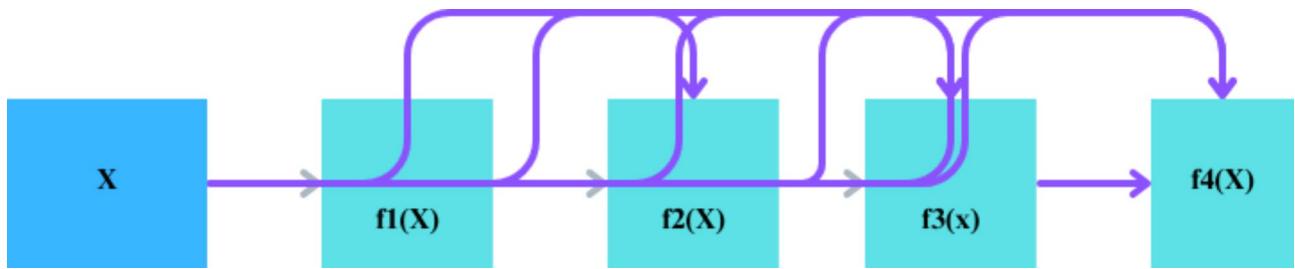


Fig. 6. CNN-DenseNet architecture.

$$\begin{aligned} DenseNet(X) = & \text{Conc7x7}(X) \rightarrow \text{DenseBlock1}(X) \rightarrow \text{Transition1}(X) \rightarrow \text{DenseBlock2}(X) \\ & \rightarrow \text{Transition2}(X) \rightarrow \dots \rightarrow \text{GlobalAvgPool}(X) \rightarrow \text{Softmax}(X) \end{aligned} \quad (19)$$

Transition layer

Transition layers are used to control the growth of feature maps between dense blocks. Mathematically, the output of a transition layer can be expressed as follows:

$$X_{\text{Transition}} = \text{BatchNorm}(\text{ReLU}(\text{Conv}1 \times 1(X_{\text{dense}}))) \quad (20)$$

where $\text{Conv}1 \times 1$ represents a 1×1 convolutional layer. BatchNorm is a batch normalization method. ReLU is the rectified linear unit activation function.

Convolutional layer

The convolutional layer is represented mathematically as follows:

$$Y_{i,j,z} = \sigma \left(\sum_{l=1}^c W_{i,j,l,z} \cdot X_{i,j,l} + B_z \right) \quad (21)$$

where $Y_{i,j,z}$ is the output at position (i, j) in the z -th feature map. $X_{i,j,l}$ is the input at position (i, j) in the l -th input feature map. $W_{i,j,l,z}$ is the weight for the convolutional kernel at position (i, j) for input channel l and output channel z . B_z is the bias term for the z -th feature map. σ is the activation function, commonly referred to as ReLU .

Global average pooling layer

The GAP layer computes the average value of each feature map across its spatial dimensions. Mathematically, the global average pooling is expressed as:

$$Y_k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{i,j,k} \quad (22)$$

where H and W are the height and width of the feature map, respectively. Y_k is the output of the k -th neuron in the global average pooling layer.

These equations provide a foundation for understanding the operations performed by DenseNet, particularly the dense connections that facilitate feature reuse and contribute to the network's parameter efficiency.

Stochastic gradient descent (SGD)

The term "Stochastic Gradient Descent" (SGD) is a general optimization algorithm widely used in training CNNs and other machine learning models³⁵. SGD optimizes the model parameters by updating them iteratively based on the gradients of the loss function with respect to those parameters. A description of the SGD method for CNNs and the mathematical equations involved are provided below:

Objective function:

In CNNs, as in many machine learning models, the objective is to minimize a loss function $J(\theta)$, where θ represents the model parameters.

Update rule:

The update rule for the parameters θ during each iteration of SGD is given by:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla J(\theta_t) \quad (23)$$

where η is the learning rate, a hyperparameter controlling the size of the step. $\nabla J(\theta_t)$ is the gradient of the loss function with respect to the parameters at iteration t .

Batch Stochastic Gradient Descent:

In practice, SGD is often applied to minibatches of data rather than to the entire dataset. In this case, the update rule becomes:

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{1}{|B|} \sum_{i \in B} \nabla J(\theta_t) \quad (24)$$

where $|B|$ is the batch size. $\nabla J_i(\theta_t)$ is the gradient of the loss with respect to the parameters computed on the i -th mini-batch.

Mathematical Equations:**Objective Function:**

The objective function is typically defined as the average loss over the training dataset:

$$J(\theta) = \frac{1}{N} \sum_{i=1}^N L(f(x_i; \theta), y_i) \quad (25)$$

where N is the number of training samples. L is the loss function. $f(x_i; \theta)$ is the predicted output of the CNN for input x_i with parameter θ . y_i is the true label of the i -th sample.

Gradient calculation.

The gradient of the loss with respect to the parameters is computed using the chain rule:

$$\nabla J(\theta) = \frac{1}{N} \sum_{i=1}^N \nabla L(f(x_i; \theta), y_i) \nabla f(x_i; \theta) \quad (26)$$

This involves computing the gradient of the loss function and the gradient of the model's output with respect to the parameters.

Update rule:

The parameters are updated using Eqs. (22) and (23), which are the calculated gradients. These equations form the basis of the SGD optimization algorithm for training CNNs³⁶. The learning rate (η) is a critical hyperparameter that requires tuning for effective training.

Clinical data sources

This section provides an overview of the clinical data sources utilized in the study, specifically focusing on the Wisconsin Breast Cancer Database and the CBIS-DDSM Dataset. It details the characteristics, relevance, and integration of these datasets in developing and validating the hybrid deep learning model for breast cancer diagnosis.

Wisconsin breast cancer database

A significant and widely used dataset in the field of breast cancer research is the Wisconsin Breast Cancer Database. The source material is obtained from breast mass Fine Needle Aspirates (FNAs), which offer important insights into cellular properties. The prominent features of this dataset include the following:

Attributes of Cell Nuclei: The collection provides comprehensive information on 30 features of cell nuclei, describing them in excellent detail. These characteristics provide a thorough understanding of cellular morphology by encompassing a variety of features, such as diameter, texture, and concaveness, as shown in Table 1.

Instances and class distribution: The dataset, which includes 569 cases in total, includes a wide range of breast cancer cases. A total of 212 of these patients were categorized as malignant tumors, suggesting that diagnostic models can be used to accurately distinguish between benign and malignant diseases, as shown in Table 2.

| id | Radius_mean | Texture_mean | Concavity_mean | Radius_se | Texture_se | Concavity_se | Radius_worst | Texture_worst | Concavity_worst |
|------------|--------------------|---------------------|-----------------------|------------------|-------------------|---------------------|---------------------|----------------------|------------------------|
| 842,302 | 17.99 | 10.38 | 0.3001 | 1.095 | 0.9053 | 0.05373 | 25.38 | 17.33 | 0.7119 |
| 842,517 | 20.57 | 17.77 | 0.0869 | 0.5435 | 0.7339 | 0.0186 | 24.99 | 23.41 | 0.2416 |
| 84,300,903 | 19.69 | 21.25 | 0.1974 | 0.7456 | 0.7869 | 0.03832 | 23.57 | 25.53 | 0.4504 |
| 84,348,301 | 11.42 | 20.38 | 0.2414 | 0.4956 | 1.156 | 0.05661 | 14.91 | 26.5 | 0.6869 |
| 84,358,402 | 20.29 | 14.34 | 0.198 | 0.7572 | 0.7813 | 0.05688 | 22.54 | 16.67 | 0.4 |
| 843,786 | 12.45 | 15.7 | 0.1578 | 0.3345 | 0.8902 | 0.03672 | 15.47 | 23.75 | 0.5355 |
| 844,359 | 18.25 | 19.98 | 0.1127 | 0.4467 | 0.7732 | 0.02254 | 22.88 | 27.66 | 0.3784 |
| 84,458,202 | 13.71 | 20.83 | 0.0936 | 0.5835 | 1.377 | 0.02488 | 17.06 | 28.14 | 0.2678 |
| 844,981 | 13 | 21.82 | 0.1859 | 0.3063 | 1.002 | 0.03553 | 15.49 | 30.73 | 0.539 |
| 84,501,001 | 12.46 | 24.04 | 0.2273 | 0.2976 | 1.599 | 0.07743 | 15.09 | 40.68 | 1.105 |

Table 1. Attributes of the cell nuclei.

| id | Diagnosis | id | Diagnosis | id | Diagnosis | id | Diagnosis | id | Diagnosis | id | Diagnosis |
|-------------|------------------|------------|------------------|-----------|------------------|------------|------------------|------------|------------------|------------|------------------|
| 842,302 | M | 845,636 | M | 8,510,824 | B | 85,382,601 | M | 856,106 | M | 857,438 | M |
| 842,517 | M | 84,610,002 | M | 8,511,133 | M | 854,002 | M | 85,638,502 | M | 85,759,902 | B |
| 84,300,903 | M | 846,226 | M | 851,509 | M | 854,039 | M | 857,010 | M | 857,637 | M |
| 84,348,301 | M | 846,381 | M | 852,552 | M | 854,253 | M | 857,13,702 | B | 857,793 | M |
| 84,358,402 | M | 84,667,401 | M | 852,631 | M | 854,268 | M | 857,15 | M | 857,810 | B |
| 843,786 | M | 84,799,002 | M | 852,763 | M | 854,941 | B | 857,155 | B | 858,477 | B |
| 844,359 | M | 848,406 | M | 852,781 | M | 855,133 | M | 857,156 | B | 858,970 | B |
| 844,458,202 | M | 84,862,001 | M | 852,973 | M | 855,138 | M | 857,343 | B | 858,981 | B |
| 844,981 | M | 849,014 | M | 853,201 | M | 855,167 | M | 857,373 | B | 858,986 | M |
| 84,501,001 | M | 8,510,426 | B | 853,401 | M | 855,563 | M | 857,374 | B | 859,196 | B |

Table 2. Instances and class distribution.

| id | Perimeter_mean | Area_mean | Smoothness_mean | Perimeter_se | Area_se | Smoothness_se | Perimeter_worst | Area_worst | Smoothness_worst |
|------------|-----------------------|------------------|------------------------|---------------------|----------------|----------------------|------------------------|-------------------|-------------------------|
| 842,302 | 122.8 | 1001 | 0.1184 | 8.589 | 153.4 | 0.006399 | 184.6 | 2019 | 0.1622 |
| 842,517 | 132.9 | 1326 | 0.08474 | 3.398 | 74.08 | 0.005225 | 158.8 | 1956 | 0.1238 |
| 84,300,903 | 130 | 1203 | 0.1096 | 4.585 | 94.03 | 0.00615 | 152.5 | 1709 | 0.1444 |
| 84,348,301 | 77.58 | 386.1 | 0.1425 | 3.445 | 27.23 | 0.00911 | 98.87 | 567.7 | 0.2098 |
| 84,358,402 | 135.1 | 1297 | 0.1003 | 5.438 | 94.44 | 0.01149 | 152.2 | 1575 | 0.1374 |
| 843,786 | 82.57 | 477.1 | 0.1278 | 2.217 | 27.19 | 0.00751 | 103.4 | 741.6 | 0.1791 |
| 844,359 | 119.6 | 1040 | 0.09463 | 3.18 | 53.91 | 0.004314 | 153.2 | 1606 | 0.1442 |
| 84,458,202 | 90.2 | 577.9 | 0.1189 | 3.856 | 50.96 | 0.008805 | 110.6 | 897 | 0.1654 |
| 844,981 | 87.5 | 519.8 | 0.1273 | 2.406 | 24.32 | 0.005731 | 106.2 | 739.3 | 0.1703 |
| 84,501,001 | 83.97 | 475.9 | 0.1186 | 2.039 | 23.94 | 0.007149 | 97.65 | 711.4 | 0.1853 |

Table 3. Ground truth pathology information.

| id | Compactness_mean | Concave points_mean | Symmetry_mean | Fractal_dimension_mean | Compactness_se | Concave points_se | Symmetry_se | Fractal_dimension_se | Compactness_worst | Concave points_worst | Symmetry_worst |
|------------|-------------------------|----------------------------|----------------------|-------------------------------|-----------------------|--------------------------|--------------------|-----------------------------|--------------------------|-----------------------------|-----------------------|
| 842,302 | 0.2776 | 0.1471 | 0.2419 | 0.07871 | 0.04904 | 0.01587 | 0.03003 | 0.006193 | 0.6656 | 0.2654 | 0.4601 |
| 842,517 | 0.0786 | 0.07017 | 0.1812 | 0.05667 | 0.01308 | 0.0134 | 0.01389 | 0.003532 | 0.1866 | 0.186 | 0.275 |
| 84,300,903 | 0.1599 | 0.1279 | 0.2069 | 0.05999 | 0.04006 | 0.02058 | 0.0225 | 0.004571 | 0.4245 | 0.243 | 0.3613 |
| 84,348,301 | 0.2839 | 0.1052 | 0.2597 | 0.09744 | 0.07458 | 0.01867 | 0.05963 | 0.009208 | 0.8663 | 0.2575 | 0.6638 |
| 84,358,402 | 0.1328 | 0.1043 | 0.1809 | 0.05883 | 0.02461 | 0.01885 | 0.01756 | 0.005115 | 0.205 | 0.1625 | 0.2364 |
| 843,786 | 0.17 | 0.08089 | 0.2087 | 0.07613 | 0.03345 | 0.01137 | 0.02165 | 0.005082 | 0.5249 | 0.1741 | 0.3985 |
| 844,359 | 0.109 | 0.074 | 0.1794 | 0.05742 | 0.01382 | 0.01039 | 0.01369 | 0.002179 | 0.2576 | 0.1932 | 0.3063 |
| 84,458,202 | 0.1645 | 0.05985 | 0.2196 | 0.07451 | 0.03029 | 0.01448 | 0.01486 | 0.005412 | 0.3682 | 0.1556 | 0.3196 |
| 844,981 | 0.1932 | 0.09353 | 0.235 | 0.07389 | 0.03502 | 0.01226 | 0.02143 | 0.003749 | 0.5401 | 0.206 | 0.4378 |
| 84,501,001 | 0.2396 | 0.08543 | 0.203 | 0.08243 | 0.07217 | 0.01432 | 0.01789 | 0.01008 | 1.058 | 0.221 | 0.4366 |

Table 4. Fine needle aspirate details.

Ground truth pathology information: Ground truth labels are reliable for investigators and clinicians since the pathology information has been validated for every patient in the Wisconsin Breast Cancer Database. The training, verifying, and testing machine learning models guarantee the dataset's dependability, as shown in Table 3.

Fine needle aspiration (FNA): Given that FNA is a frequently used diagnostic technique in actual clinical settings, the use of FNA samples provides an extra degree of clinical relevance. Therefore, the dataset is especially useful for comprehending and identifying breast cancer based on cellular features, as shown in Table 4.

CBIS-DDSM dataset description

This dataset is the original dataset (163GB) that has been converted to jpeg format. The resolution is the same as that of the original dataset.

No. of courses: 6775.

No. of rows: 6775.

Number of participants: 1,566 (note).

No of images: 10,239.

Image size (GB): 6(.jpg).

The imaging data for this set is structured such that each participant has multiple patient identities. For example, pat_id 00038 contains 10 distinct patient IDs (e.g. Calc-Test_P_00038_LEFT_CC, Calc-Test_P_00038_RIGHT_CC_1) that provide information about the persons within the ID, and according to the actual DICOM metadata, only 156, 6 IDs appear. Participants in the group³⁸.

Curated breast imaging subset of DDSM (CBIS-DDSM)

An improved and standardized version of the Digital Database for Screening Mammography (DDSM) is the CBIS-DDSM dataset, as shown in Fig. 7. It offers improvements over its predecessors and an array of mammography investigations. The important characteristics of the CBIS-DDSM dataset include the following:

The CBIS-DDSM is a standardized and updated version of the DDSM that takes into account improvements in breast imaging techniques and technology, as shown in Table 5. The dataset was chosen carefully to guarantee coherence and pertinence within the framework of modern mammography research. A total of 2,620 scanned film mammography investigations encompassing a wide range of breast diseases, including benign, malignant, and normal instances, were included in the initial DDSM. Because of its variety, the CBIS-DDSM is an invaluable resource for developing and accessing diagnostic models for breast cancer. The CBIS-DDSM preserves pathology data for every mammography investigation, much like does the Wisconsin Breast Cancer Database. By doing so, the dataset's clinical validity is preserved, facilitating the creation of deep learning models that correspond with actual diagnostic instances.

The CBIS-DDSM provides an accurate representation of breast diseases found in clinical practice by including lesions that are benign, malignant, or normal. This diversity is important for training models that can reliably diagnose various breast diseases in Table 5. The therapeutic relevance of breast cancer research has been greatly

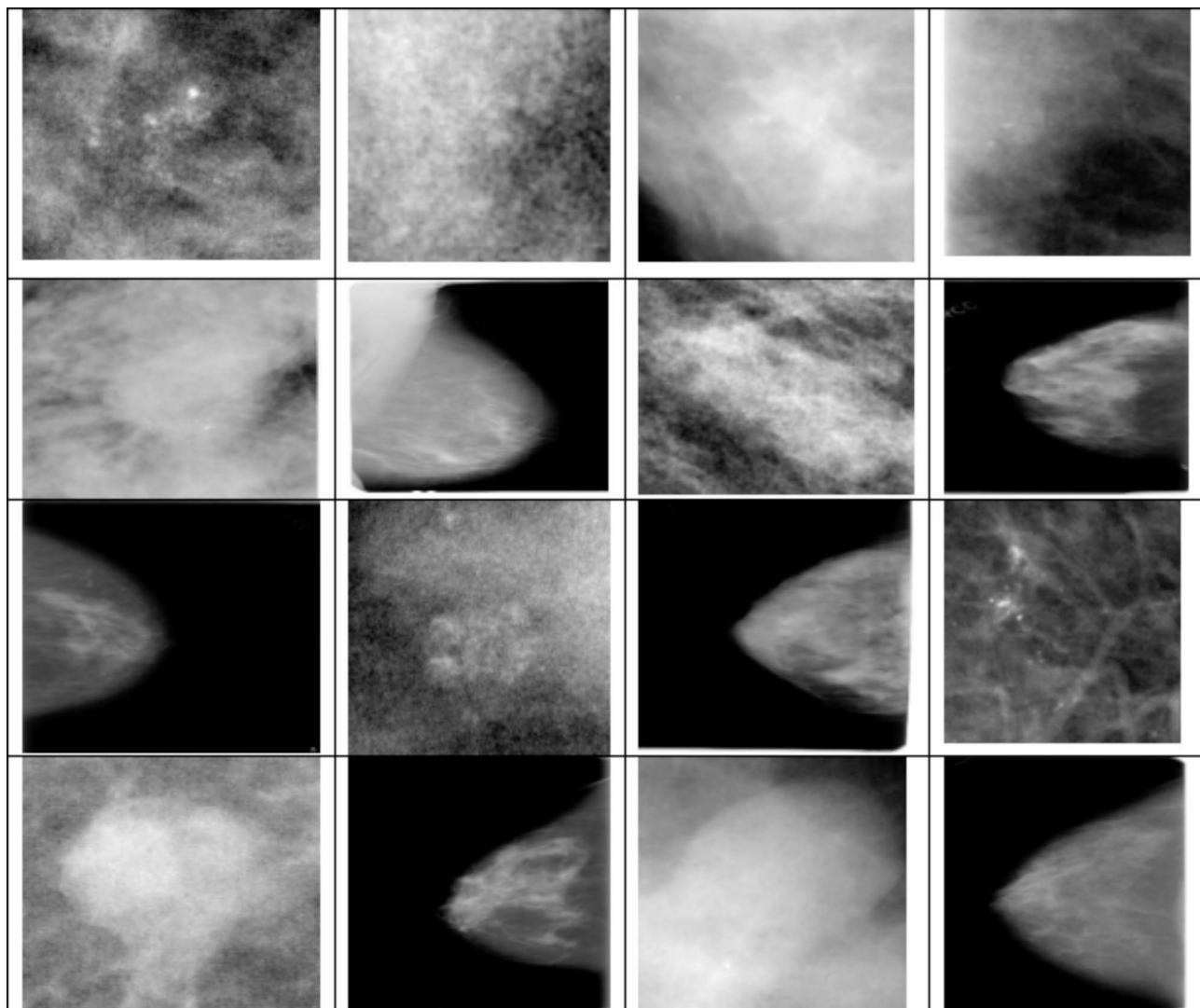


Fig. 7. Digital database for screening mammography.

| Patient_id | Breast density | Left or right breast | Image view | Abnormality id | Aabnormality type | Calc type | Calc distribution | Assessment | Pathology | Subtlety |
|------------|----------------|----------------------|------------|----------------|-------------------|----------------------|-------------------|------------|-------------------------|----------|
| P_00038 | 2 | Left | CC | 1 | Calcification | PUNCTATE-PLEOMORPHIC | CLUSTERED | 4 | BENIGN | 2 |
| P_00038 | 2 | Left | MLO | 1 | Calcification | PUNCTATE-PLEOMORPHIC | CLUSTERED | 4 | BENIGN | 2 |
| P_00038 | 2 | Right | CC | 1 | Calcification | VASCULAR | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 5 |
| P_00038 | 2 | Right | CC | 2 | Calcification | VASCULAR | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 5 |
| P_00038 | 2 | Right | MLO | 1 | Calcification | VASCULAR | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 5 |
| P_00038 | 2 | Right | MLO | 2 | Calcification | VASCULAR | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 5 |
| P_00041 | 1 | Left | CC | 2 | Calcification | LUCENT_CENTER | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 5 |
| P_00041 | 1 | Left | MLO | 2 | Calcification | LUCENT_CENTER | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 5 |
| P_00077 | 2 | Left | CC | 1 | Calcification | ROUND_AND_REGULAR | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 3 |
| P_00077 | 2 | Left | MLO | 1 | Calcification | ROUND_AND_REGULAR | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 3 |
| P_00077 | 2 | Right | CC | 1 | Calcification | PUNCTATE | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 3 |
| P_00077 | 2 | Right | CC | 2 | Calcification | EGGSHELL | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 4 |
| P_00077 | 2 | Right | MLO | 1 | Calcification | EGGSHELL | N/A | 2 | BENIGN_WITHOUT_CALLBACK | 4 |

Table 5. CBIS-DDSM resource <https://www.kaggle.com/datasets/awsaf49/cbis-ddsm-breast-cancer-image-dataset>.

enhanced by the contributions of both the CBIS-DDSM and the Wisconsin Breast Cancer Database databases. The former incorporates cellular features, while the latter includes a variety of mammography investigations to guarantee that machine learning models trained on these datasets can accurately capture the nuances of diagnosing breast cancer in practical settings. The clinical utility of these datasets is further enhanced by the availability of ground truth pathology information, which makes them invaluable tools for improving breast cancer research and diagnosis.

Methodology

The objective of the proposed method is to integrate the CBIS-DDSM and Wisconsin Breast Cancer Database datasets to create an innovative breast cancer diagnosis system. For optimized diagnostic accuracy, the solution uses a hybrid deep learning algorithm that combines CNNs with the stochastic gradients method. This approach was used by considering the numerous characteristics of cell nuclei from the Wisconsin Breast Cancer Database and the variety of mammography studies in the CBIS-DDSM dataset.

Data integration and pre-processing

Integration of Datasets: The process of integrating the CBIS-DDSM and Wisconsin Breast Cancer Database datasets into one cohesive, coherent dataset is known as integration. This procedure ensured that the two datasets were integrated effectively, enabling a thorough analysis that utilized each dataset's characteristics, as shown in Table 6. To produce a coherent dataset for further analysis, harmonizing data types, normalizing variable names, and aligning data structures are important considerations.

Providing particular interest to the formats and characteristics of the data in both datasets is necessary to guarantee consistency and compatibility. Harmonizing features such patient identities, image metadata, and pathology data is necessary to prevent discrepancies that might affect how reliable the diagnostic model is. This is an essential step in preserving the integrity of the integrated dataset and facilitating an easy transition between the various types of data included in the CBIS-DDSM and the Wisconsin Breast Cancer Database. One essential component of data preparation is handling missing values. This entails locating and fixing any null or gap items in the combined dataset. The dataset may be made full and analysis-ready by using imputation techniques to fill

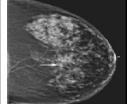
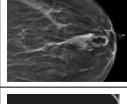
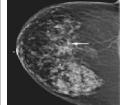
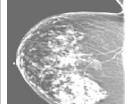
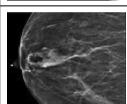
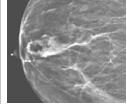
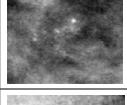
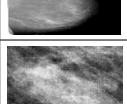
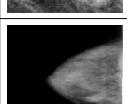
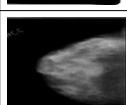
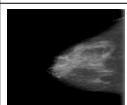
| Patient_id | Breast density | Left or right breast | Image view | Radius_mean | Texture_mean | Radius_se | Texture_se | Concavity_se | Input Image |
|------------|----------------|----------------------|------------|-------------|--------------|-----------|------------|--------------|---|
| P_00038 | 2 | Left | CC | 17.99 | 10.38 | 1.095 | 0.9053 | 0.05373 |  |
| P_00038 | 2 | Left | MLO | 20.57 | 17.77 | 0.5435 | 0.7339 | 0.0186 |  |
| P_00038 | 2 | Right | CC | 19.69 | 21.25 | 0.7456 | 0.7869 | 0.03832 |  |
| P_00038 | 2 | Right | CC | 11.42 | 20.38 | 0.4956 | 1.156 | 0.05661 |  |
| P_00038 | 2 | Right | MLO | 20.29 | 14.34 | 0.7572 | 0.7813 | 0.05688 |  |
| P_00038 | 2 | Right | MLO | 12.45 | 15.7 | 0.3345 | 0.8902 | 0.03672 |  |
| P_00041 | 1 | Left | CC | 18.25 | 19.98 | 0.4467 | 0.7732 | 0.02254 |  |
| P_00041 | 1 | Left | MLO | 13.71 | 20.83 | 0.5835 | 1.377 | 0.02488 |  |
| P_00077 | 2 | Left | CC | 13 | 21.82 | 0.3063 | 1.002 | 0.03553 |  |
| P_00077 | 2 | Left | MLO | 12.46 | 24.04 | 0.2976 | 1.599 | 0.07743 |  |
| P_00077 | 2 | Right | CC | 15.78 | 17.89 | 0.5058 | 0.9849 | 0.02791 |  |
| P_00077 | 2 | Right | CC | 19.17 | 24.8 | 0.9555 | 3.568 | 0.0889 |  |
| P_00077 | 2 | Right | MLO | 15.85 | 23.95 | 0.4033 | 1.078 | 0.05051 |  |

Table 6. Integration of datasets.

in missing values. The kind of missing data and its possible influence on subsequent investigations should be taken into consideration when selecting an imputation approach. To provide a consistent scale for all variables, features must be normalized³⁷. This procedure assures that, because of variations in scales or measuring units, no single factor affects the model excessively. Methods such as z score normalization or min–max scaling are frequently used to align all features into a similar range, which helps the hybrid deep learning algorithm converge during training.

Creating that all the information is represented consistently across various aspects is a necessary step in standardizing data formats. This covers categorical variables, encoding data, and standardizing units of measurement. By offering a standard framework for processing and analysing data, consistent data formats support the interpretability and stability of the model. A single dataset is produced once the integration and preparation stages are completed. This dataset provides an in-depth analysis of cellular characteristics and mammography studies by combining data from the Wisconsin Breast Cancer Database and CBIS-DDSM. The hybrid deep learning method may then be used with the combined dataset, enabling reliable training and assessment procedures. To summarize, the effective implementation of the proposed breast cancer diagnostic solution hinges on the careful integration and preparation of the CBIS-DDSM and Wisconsin Breast Cancer Database datasets. By completing these processes, a unified, consistent, and comprehensive dataset is created, which paves the way for the construction of a highly accurate and dependable diagnostic model^[38].

Feature extraction and selection

Extracting Features from Cell Nuclei Attributes: Identifying the relevant characteristics among the 30 cell nuclei properties in the Wisconsin Breast Cancer Database is an important first step in gathering data needed for a diagnosis of breast cancer. These features include measures of various lengths, such as diameter, texture, and concaveness, all of which provide important new information on the properties of cell nuclei. These processed characteristics are transformed during feature extraction into a fewer, more relevant set of features that accurately capture the underlying patterns in the data. When different objects are detected, the optimal number of image features within the processing framework can be determined. Principal component analysis (PCA) dimensionality reduction is for establishing image features for appropriate feature extraction methods and detection.

The objective of the feature extraction procedure is to minimize the total dimensionality while retaining the most important information. To achieve this, approaches such as principal component analysis (PCA) and dimensionality reduction methods are employed. This approach ensures that the model is more effective and less prone to overfitting during the training procedure in the future, particularly when working with a high-dimensional dataset such as the Wisconsin Breast Cancer Database. Identifying and maintaining the most discriminative and informative characteristics for model training is known as feature selection. Improving model efficiency, interpretability, and generalizability requires taking this critical step. To do this, a variety of feature selection strategies can be used, including the following:

Filter methods:

To assess the importance of each trait to the objective variable (cancer status), statistical tests, including chi-squared or mutual information, can be used. Repetitive traits can be found, and only those that add original information should be kept according to correlation analysis, as shown in Fig. 8.

According to the model's performance, Recursive Feature Elimination (RFE) entails iteratively eliminating the least significant features. Iterative techniques of forward or backwards selection add or delete features according to how they affect the performance of the model. Model training included regularization techniques and techniques, including Least Absolute Shrinkage and Selection Operator (LASSO) regression, which automatically penalizes less significant characteristics. The selection procedure might be guided by the feature significance ratings that tree-based models frequently provide. In Table 7, these metrics evaluate how much information each feature contributes to the prediction of the target variable, and they are especially pertinent to decision tree-based models.

The objective is to enhance the model's capacity to identify patterns connected with breast cancer by reducing the dataset to include only the most pertinent features through the use of various feature selection approaches. This approach lowers the chance of overfitting and enhances interpretability in addition to increasing model efficiency—especially when combined with a hybrid deep learning approach. To summarize, the process of feature extraction and selection plays a crucial role in extracting significant data from the 30 variables included in the Wisconsin Breast Cancer Database. This study lays the groundwork for the creation of a more precise, understandable, and effective breast cancer diagnostic paradigm.

Model architecture

The proposed strategy develops a hybrid deep learning model that fuses the stochastic gradients technique with CNNs. The use of both techniques results in positive results: the stochastic gradients method's optimization power and the CNN's capacity to capture complex spatial patterns. The model aspires to improve the performance of learning complicated hierarchical features from the CBIS-DDSM and Wisconsin Breast Cancer Database datasets by incorporating these components^[39–41].

CNN architecture for spatial pattern recognition: CNNs are highly effective at performing image-related tasks, which renders them appropriate for diagnosing breast cancer, where spatial patterns are important. For the model to successfully learn hierarchical features from retrieved cell nuclei properties and mammography imagery, the architecture of the CNN layers is essential, as shown in Fig. 9. Usually, the architecture consists of the following.

Convolutional layers

These layers search for certain patterns in the input data by applying filters. These filters have the potential to capture spatial correlations in mammography imagery and cell nuclei properties for the purpose of diagnosing breast cancer.

Pooling layers

The pooling layers preserve important information while reducing the input's spatial dimensions. The most notable characteristics in each region are captured with the use of max pooling and average pooling methods.

Flatten layer

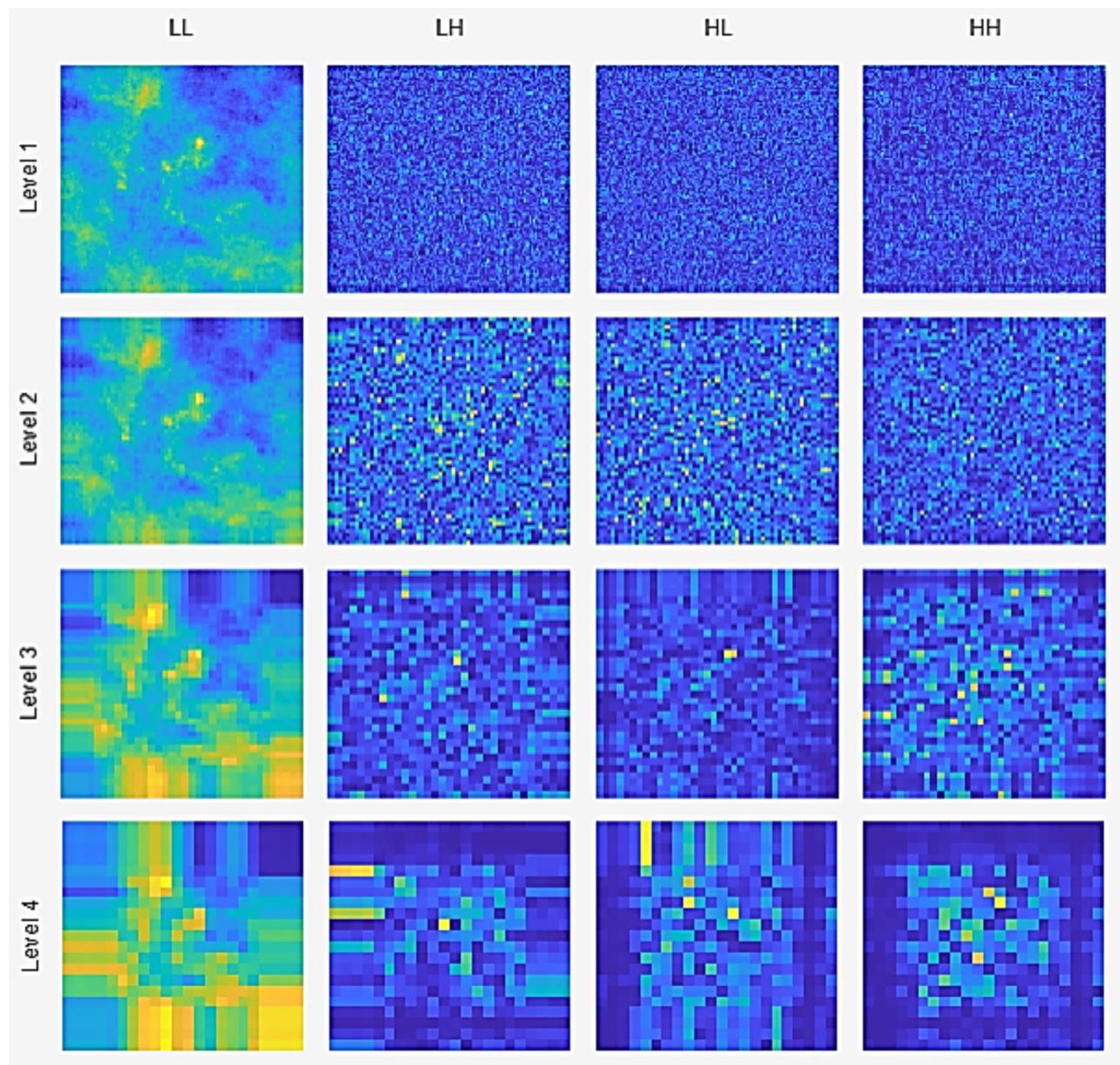


Fig. 8. Filtering output.

To prepare the multidimensional output from earlier layers for input into fully connected layers, the output must first be flattened into a one-dimensional vector.

Fully connected layers

These layers capture complex feature interactions by processing the vector that has been flattened. The last fully connected layers generate output probabilities that show whether a certain tumor is benign or malignant.

Activation functions:

Activation functions, including the Rectified Linear Unit (ReLU), are used to provide nonlinearity and improve the model's ability to recognize complex patterns.

Stochastic Gradients Method for Optimization: To adjust parameters and enhance model performance, the hybrid model incorporates the stochastic gradients technique, which is frequently used in optimization techniques, including SGD. The model can effectively explore the parameter space by modifying weights and biases to minimize the loss function through the use of the stochastic gradients approach.

In the training phase, the integrated and preprocessed datasets are fed into the hybrid model, and the parameters are repeatedly adjusted to reduce the discrepancy between the expected and actual results. The stochastic gradients approach optimizes the whole model, while the CNN layers discover spatial patterns using cell nuclei properties and mammography imagery during training. The CBIS-DDSM dataset fine-tuning stage improves the model's applicability to a variety of mammography investigations. By doing so, it is ensured that

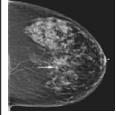
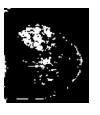
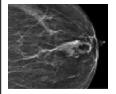
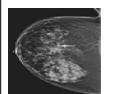
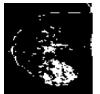
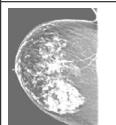
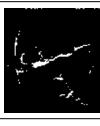
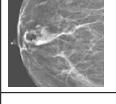
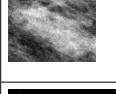
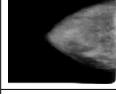
| Patient_id | Breast density | Left or right breast | Image view | Input image | Feature selection |
|------------|----------------|----------------------|------------|--|---|
| P_00038 | 2 | Left | CC |  |  |
| P_00038 | 2 | Left | MLO |  |  |
| P_00038 | 2 | Right | CC |  |  |
| P_00038 | 2 | Right | MLO |  |  |
| P_00038 | 2 | Right | MLO |  |  |
| P_00041 | 1 | Left | CC |  |  |
| P_00041 | 1 | Left | MLO |  |  |
| P_00077 | 2 | Left | CC |  |  |
| P_00077 | 2 | Left | MLO |  |  |
| P_00077 | 2 | Right | CC |  |  |
| P_00077 | 2 | Right | CC |  |  |
| P_00077 | 2 | Right | MLO |  |  |

Table 7. Feature selection output.

the hybrid model may be generalized to various imaging properties and case variations of breast cancer in an efficient manner.

Training on the Wisconsin Breast Cancer Database: A vital initial stage in creating a reliable breast cancer diagnosis model is the training phase. The Wisconsin Breast Cancer Database is a comprehensive dataset that includes 30 properties of cell nuclei, including important aspects, such as diameter, texture, and concaveness, which the hybrid deep learning model is exposed to during this procedure. Enabling the algorithm to identify complex patterns related to benign and malignant tumors in this dataset is the objective. The variance among the expected and actual results is measured by a loss function, which is optimized to train the model. To minimize the loss, this optimization process involves modifying the model's weights and biases using methods such as backpropagation. Because the CNN layers have been designed to efficiently learn hierarchical information from the characteristics of the cell nucleus, they play a crucial role in this training phase. The model's ability to

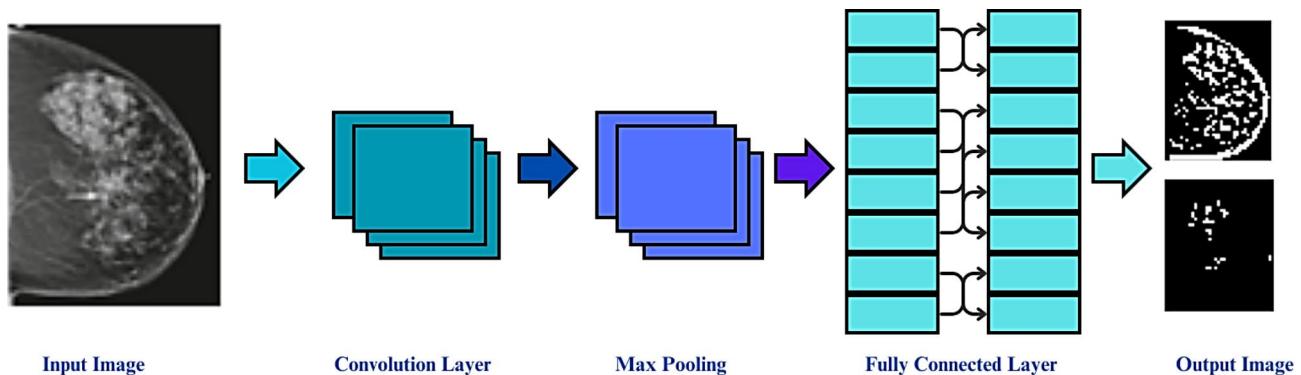


Fig. 9. CNN model proposed architecture.

distinguish minute variations between malignant and benign tumor features is facilitated by the CNN's ability to identify spatial patterns within the attributes.

Regularization strategies, including dropout and L2 regularization, can be used to stop overfitting. By minimizing the model's dependence on particular data points during training, these strategies aid in the model's generalization. This dataset provides a wide range of mammography investigations, including normal, benign, and malignant cases. It is an updated and standardized version of the DDSM. Optimizing the model's fine-tuning ability is essential for improving its generalizability to new data and its flexibility for a variety of imaging investigations.

Variations in imaging parameters and cases not in the Wisconsin Breast Cancer Database were introduced in the CBIS-DDSM dataset. By fine-tuning, the model becomes more adaptable to a variety of mammography investigations by allowing its acquired characteristics to adjust to these changes. Fine-tuning prevents the model from being too adaptable to the unique features of the training dataset and reduces dataset bias. This is especially crucial when developing a diagnostic tool that may be used in various patient demographic and medical settings. The process of fine-tuning increases the model's efficacy in real-world circumstances by optimizing it for better generalization. The objective of this study was to develop a diagnostic tool that will add to the clinical relevance of breast cancer by accurately identifying patterns of breast cancer through different mammography investigations. In conclusion, the training and fine-tuning stages are essential for the creation of a model. Training on the Wisconsin Breast Cancer Database allows for the collection of complex patterns associated with tumors, and fine-tuning using the CBIS-DDSM dataset improves the model's generalizability and flexibility, which together lead to a more robust and efficient breast cancer diagnostic system. Table 8 shows the hybrid CNN model output for breast cancer detection based on the selected features, and it shows the morphology and finally the CNN output.

In summary, utilizes the CBIS-DDSM dataset to further refine the model's parameters once it has been trained on the Wisconsin Breast Cancer Database. The hybrid deep learning model that combines the stochastic gradients approach with the CNN architecture has the potential as an effective diagnostic tool for breast cancer. The model attempts to capture complex spatial patterns from cell nuclei properties and mammography imagery by using CNN layers to learn hierarchical features, ultimately leading to more accurate and consistent diagnostic results.

Experimental results

The Wisconsin Breast Cancer Database and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) datasets were used to thoroughly assess the hybrid deep learning algorithm used in the breast cancer diagnosis model. This method consists of CNNs and stochastic gradients. The Digital Database for Screening Mammography (DDSM), which initially included 2,620 scanned film mammography investigations, has been updated and standardized in the CBIS-DDSM dataset. The Wisconsin Breast Cancer Database contains 569 cases, of which 212 were categorized as malignant tumours. A crucial initial phase in determining the efficacy of the created breast cancer diagnostic model is to evaluate its diagnostic performance, as shown in Fig. 10. To provide a more detailed picture of the model's capabilities, a set of extensive assessment measures is used. Among these measures are:

The ratio of effectively categorized instances to the total number of occurrences is used to calculate accuracy, which assesses how accurate the model's predictions are overall. Recall, also referred to as sensitivity, is a measure of how well a model recognizes malignant tumors. The ratio of true positives to the total number of false negatives and true positives is used to compute this parameter. Specificity, which is computed as the ratio of true negatives to the total number of true negatives and false positives, assesses the model's accuracy in identifying benign tumors. The precision assesses the model's accuracy in predicting a positive class. The ratio of true positives to the total number of true positives and false positives is calculated as shown in Fig. 11.

| Patient_id | Breast density | Left or right breast | Image view | Input image | Feature selection | Morphology | CNN output |
|------------|----------------|----------------------|------------|-------------|-------------------|------------|------------|
| P_00038 | 2 | Left | CC | | | | |
| P_00038 | 2 | Left | MLO | | | | |
| P_00038 | 2 | Right | CC | | | | |
| P_00038 | 2 | Right | MLO | | | | |
| P_00038 | 2 | Right | MLO | | | | |
| P_00041 | 1 | Left | CC | | | | |
| P_00041 | 1 | Left | MLO | | | | |
| P_00077 | 2 | Left | CC | | | | |
| P_00077 | 2 | Left | MLO | | | | |
| P_00077 | 2 | Right | CC | | | | |
| P_00077 | 2 | Right | CC | | | | |

Table 8. Hybrid CNN model output.**Area under the receiver operating characteristic curve (AUC-ROC)**

The AUC-ROC is a valid statistic that assesses the model's capacity to distinguish between benign and malignant cases at different thresholds, as shown in Fig. 12. The results provide a thorough overview of the model's performance, with a higher AUC-ROC indicating improved discriminating ability.

Cross-validation:

To assure the model's stability and dependability over various dataset subsets, cross-validation is a significant approach. This procedure entails dividing the dataset into several folds, training the model on various permutations of these folds, and repeatedly assessing the model's performance. Popular techniques for cross-validation include leave-one-out cross-validation, in which every instance acts as a separate validation set, and k-fold cross-validation, in which the dataset is split into k subsets.

| | | Predicted Class | | |
|--------------|-------------------------------------|---|---|---------------------------------------|
| | | Positive | Negative | |
| Actual Class | Positive | True Positive (TP) | False Negative (FN) Type II Error | Sensitivity $\frac{TP}{(TP + FN)}$ |
| | Negative | False Positive (FP) Type I Error | True Negative (TN) | Specificity $\frac{TN}{(TN + FP)}$ |
| | Precision $\frac{TP}{(TP + FP)}$ | Negative Predictive Value $\frac{TN}{(TN + FN)}$ | Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$ | |

Fig. 10. Performance evaluation of the CNN Hybrid model.

Cross-validation is crucial for evaluating the model's resilience across multiple data subsets, as shown in Figs. 13, 14, 15 and 16. It ensures that the model performs well beyond a specific training-validation split, reducing the risk of overfitting by providing a thorough assessment of generalization abilities. Additionally, cross-validation helps optimize hyperparameters, allowing the model to be adjusted and verified on various subsets, ensuring performance is not skewed by any particular parameter set. The consistent performance across folds confirms the model's diagnostic validity and its ability to generalize to new data, essential for practical applications. So, using cross-validation and extensive evaluation metrics is vital for thoroughly assessing the breast cancer diagnostic model's accuracy, sensitivity, specificity, precision, and discriminatory power, thereby supporting its validity and clinical relevance.

The development of a reliable, apparent, and accurate breast cancer diagnostic model requires a high degree of interpretability and interpretability. Improving the interpretability of the model entails improving the accessibility and comprehensibility of decision-making procedures for medical practitioners. This is especially crucial for medical applications, as building trust and enabling the use of AI-powered technologies in clinical settings both depend on authenticity. Understanding which areas of the input data are essential for decision-making is facilitated by examining the activation maps in the Convolutional Neural Network (CNN) layers. This enables medical practitioners to determine which areas of a picture contribute most to the model's predictions, which may be very useful when analysing mammography images. Plots of feature significance provide insight into how each input feature (the characteristics of the cell nuclei) affects the predictions made by the model, as shown in Figs. 17, 18 and 19. Determining the most important factors related to breast cancer can be facilitated by identifying which aspects have greater influence. It is possible to assign importance to each input feature using strategies such as layerwise relevance propagation, emphasizing the role that certain characteristics have in the decision-making process. This helps to clarify the process by which the model makes its predictions.

Providing detailed explanations for each forecast enables medical personnel to comprehend the model's reasoning behind a particular patient's decision. Because local interpretability makes the model's decision-making process more concrete and applicable to actual situations, it fosters confidence. Global explanations provide an overview of the model's general behavior throughout the whole dataset. Finding recurring themes or characteristics that reliably influence forecasts is part of this. To comprehend the basic patterns and strengths of the model, global interpretability is essential. A thorough method for deciphering any machine learning model's output is provided by the SHAP values. They are able to produce comprehensible and naturalistic summaries of the model's decision-making process by attributing its relative importance to the prediction to each characteristic.

Clinical relevance:

Healthcare practitioners can compare the model's predictions to accepted medical knowledge by integrating the model's output with current clinical information. The model gains credibility and is more likely to be accepted by the medical community as a result of this alignment. Healthcare practitioners may make well-informed judgments based on the model's predictions when they have access to transparent and comprehensible models. Collaboration between AI tools and physicians can be improved by having a better understanding of the rationale underlying the model's suggestions. Establishing confidence in AI-powered diagnostic tools

**Fig. 11.** Confusion matrix.

requires transparency and interpretability. The validity and dependability of the model's output can be increased for healthcare practitioners by giving them concise explanations for the predictions made by the model. To summarize, the interpretability and interpretability of the breast cancer diagnostic model are enhanced by the application of visualization tools and the provision of explanations for model predictions. This promotes acceptability and confidence in the clinical community while also helping healthcare practitioners comprehend the model's decision-making procedures.

Validation on an independent dataset:

An important initial phase in evaluating the generalizability and realistic performance of the trained breast cancer detection model is validation on a separate dataset. To determine how effectively the model can apply its newly learned patterns to fresh, unknown data, it is tested on a dataset that has not been used during training or fine-tuning.

The model's capacity to generalize beyond the unique features of the training and fine-tuning datasets is assessed by validation on an independent dataset. This approach is important for assessing the model's adaptability to different patient demographics, differences in imaging, and other elements that arise in actual clinical settings. One essential feature of a trustworthy diagnostic model is its robustness. Determining whether the model retains its diagnostic efficacy and accuracy in the face of changes and difficulties not observed in

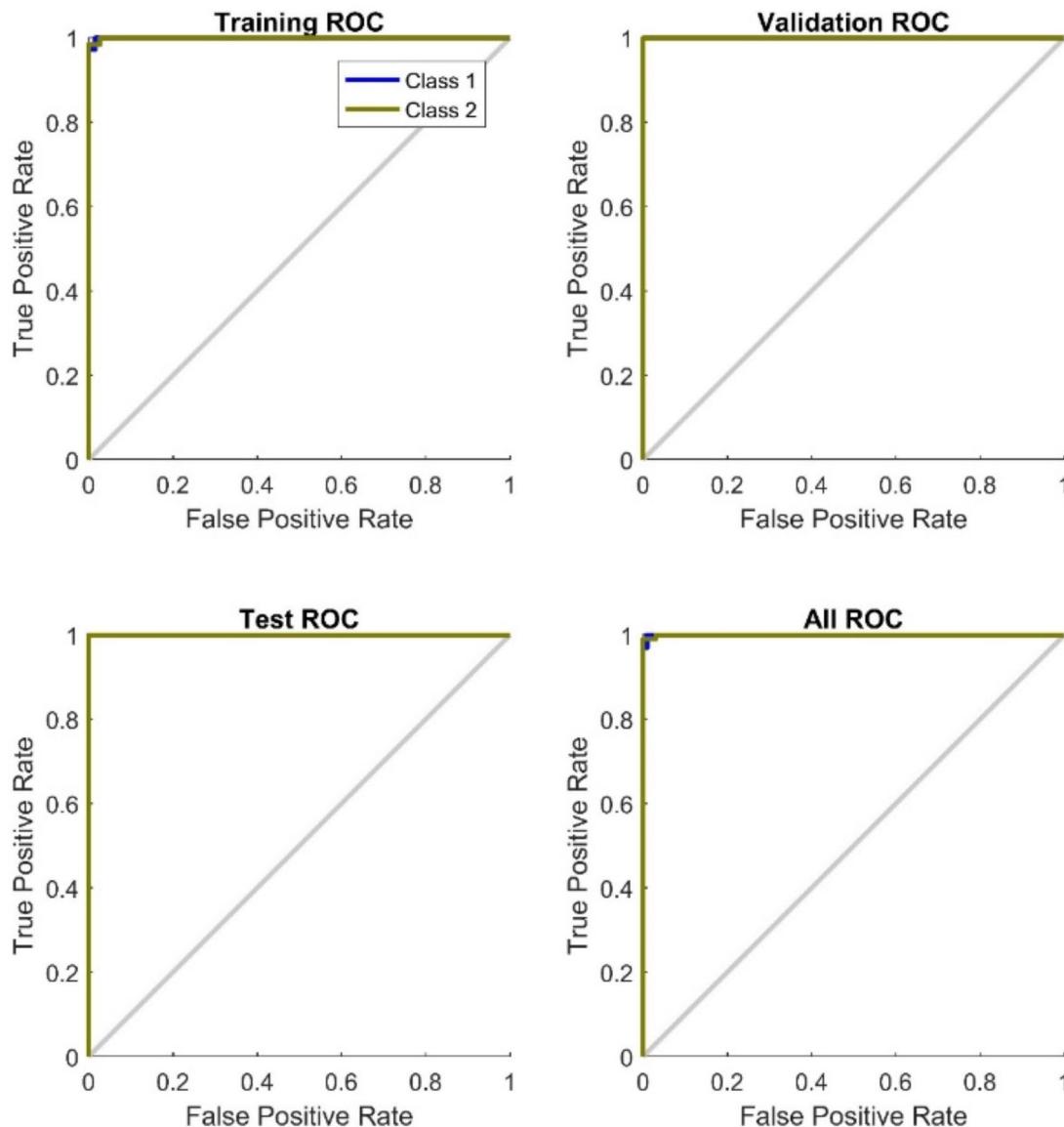


Fig. 12. AUC-ROC curve of the proposed system.

the training set is easier by validation on a different dataset, as shown in Figs. 20, 21, 22 and 23. When a model performs effectively on training data but is unable to generalize to new data, this is known as overfitting. To verify that the model's performance is not inflated as a result of learning certain patterns from the training dataset, independent validation functions are used as a test for overfitting. The ultimate objective of a breast cancer diagnostic model is to be useful in actual clinical situations. By simulating the model's performance in a variety of scenarios, validation on an independent dataset improved the model's adaptability to various patient demographics, healthcare facilities, and imaging technologies.

Evaluation across diverse patient populations

Demographic diversity:

It is possible to evaluate the diagnostic efficacy of the model across a range of demographic groups using independent validation. This covers a range of age brackets, races, and other demographic variables that might impact the features of individual instances of breast cancer.

Pathological variations:

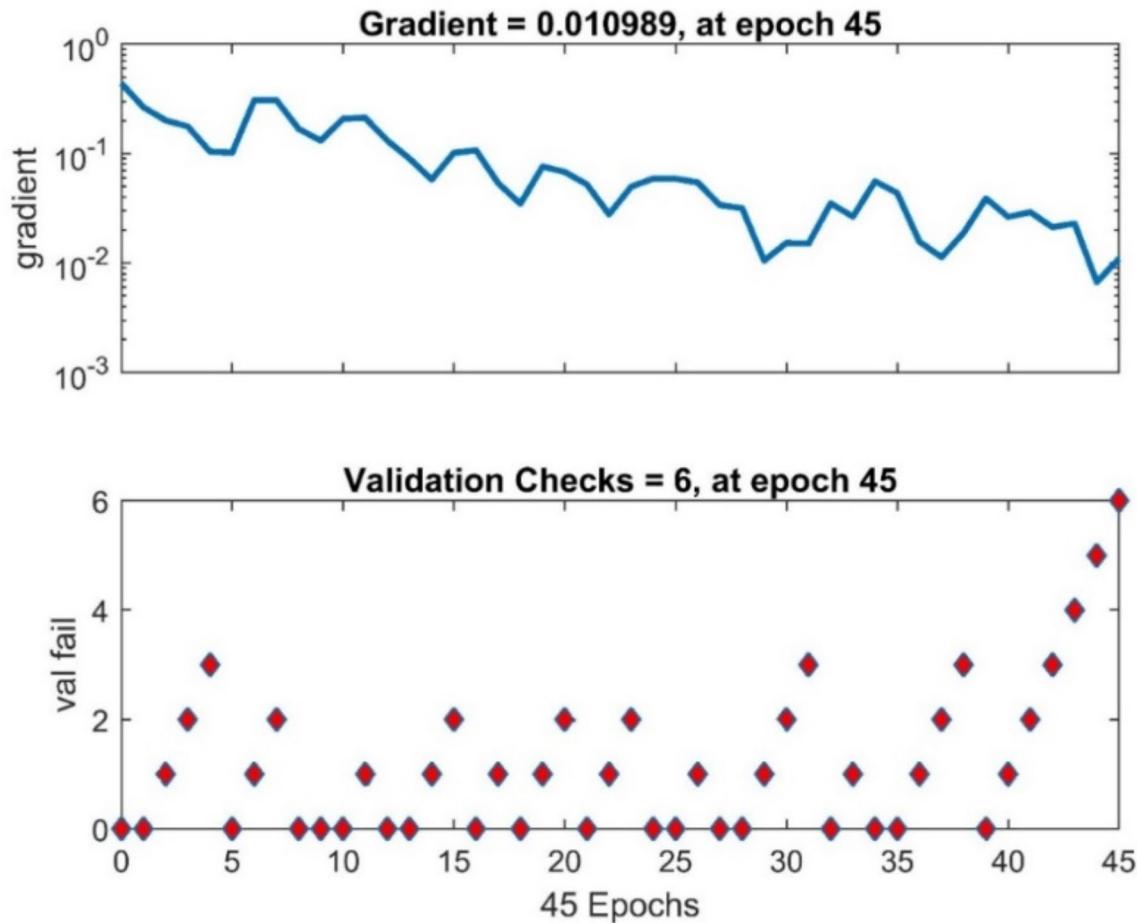


Fig. 13. Cross-validation based on the gradient.

Breast cancer can present in a variety of pathological forms, and distinct features may be observed in different populations. To guarantee that the diagnostic abilities of the model are not skewed toward certain pathogenic variants found in the training data, the model was validated on a different dataset, as shown in Tables 9, 10 and 11.

Evaluation across imaging variants

The model's performance was evaluated independently across multiple imaging modalities that might not have been completely captured in the training dataset. This approach is especially crucial because advances in imaging technology might lead to different representations of breast cancer. Changes in the resolution and quality of the images can affect how well the model diagnoses breast cancer. Assessing how well the model generalizes to datasets with different image qualities and resolutions is easier with the use of independent validation. The area under the receiver operating characteristic curve (AUC-ROC), accuracy, sensitivity, specificity, precision, recall, and other measures are used in validation on an independent dataset, much as they are in training. These metrics offer a thorough evaluation of the diagnostic performance of the model, as shown in Table 12.

To summarize, the process of validating a breast cancer diagnosis model on an independent dataset is an essential measure that guarantees its dependability and practicality. This procedure enhances the model's credibility and facilitates its use in clinical practice by assessing its effectiveness across a range of patient demographics and imaging variances.

Comparative analysis

To establish the viability of the proposed hybrid deep learning algorithm in terms of accuracy, efficiency, and adaptation, it must be compared with existing methods for diagnosing breast cancer. The purpose of this analysis is to illustrate the new algorithm's contributions and developments in the field of breast cancer diagnosis, as shown in Table 13.

The diagnostic accuracy of the proposed hybrid deep learning algorithm is compared to that of existing techniques. To measure its effectiveness, criteria such as overall accuracy, precision, sensitivity, and specificity were used. An improved algorithm should reliably outperform existing methods in accurately identifying benign or malignant cases, as shown in Table 14 and Fig. 24.

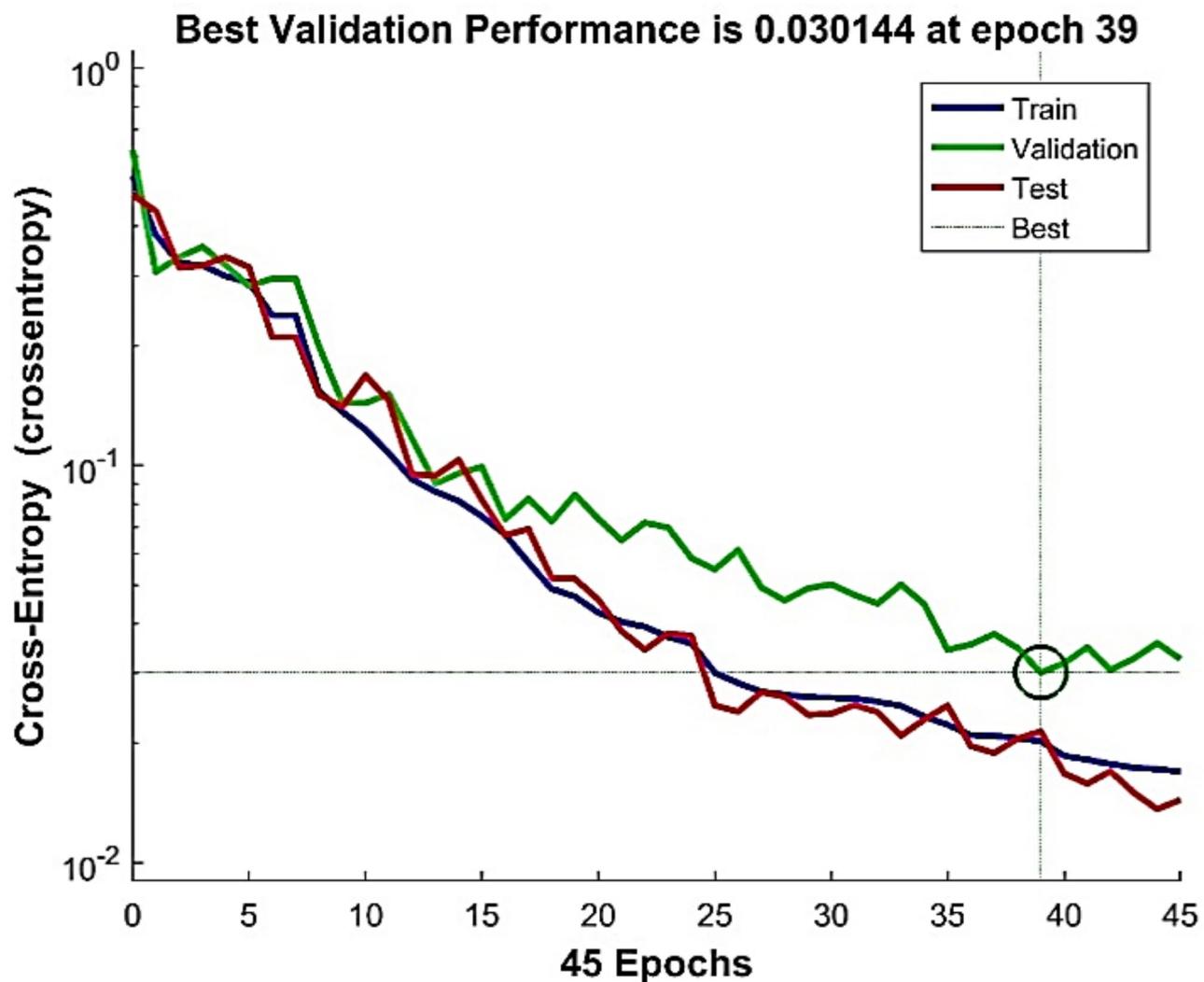


Fig. 14. Cross-validation based on the cross entropy.

The proposed algorithm's computational effectiveness is evaluated by contrasting its training and inference times with those of other approaches. In clinical settings, where immediate findings are essential for patient care, the practical deployment of a tool for diagnosis depends heavily on its efficiency. We analysed how well the hybrid deep learning system works with different patient demographics, imaging modalities, and datasets. An optimal algorithm should be able to manage variations observed in real-world circumstances with outstanding results and excellent generalization abilities. The model's output was analysed on various datasets that were not used for fine-tuning or training. The algorithm's capacity to generalize beyond certain training settings and assure its applicability to other clinical situations is validated through the use of cross-dataset analysis.

Comparison with traditional methods

Examine how the proposed hybrid deep learning algorithm compares with both more modern deep learning techniques and more conventional strategies, such as rule-based systems or statistical models. This more comprehensive perspective highlights the improvements made by the new algorithm when considering the changing landscape of diagnostic techniques. The interpretability and reliability of the suggested algorithm are compared to those of existing methods. The validity of diagnostic tools is enhanced by apparent decision-making procedures and explainable AI capabilities, particularly in healthcare applications.

We conducted an accurate and standardized comparison by using benchmark datasets that are widely regarded in the field of breast cancer diagnosis. To ensure a thorough evaluation of the suggested hybrid deep learning algorithm, this approach may involve well-established datasets that are often used for assessing diagnostic algorithms. To determine the significance of the observed variations in the performance measures between the proposed algorithm and current approaches, statistical analysis was used. The outcomes of the comparison analysis are more reliable when they are supported by statistical significance.

The clinical applicability of the proposed algorithm was evaluated in consultation with medical experts and compared to that of current techniques. The adoption of diagnostic tools depends on an understanding of their

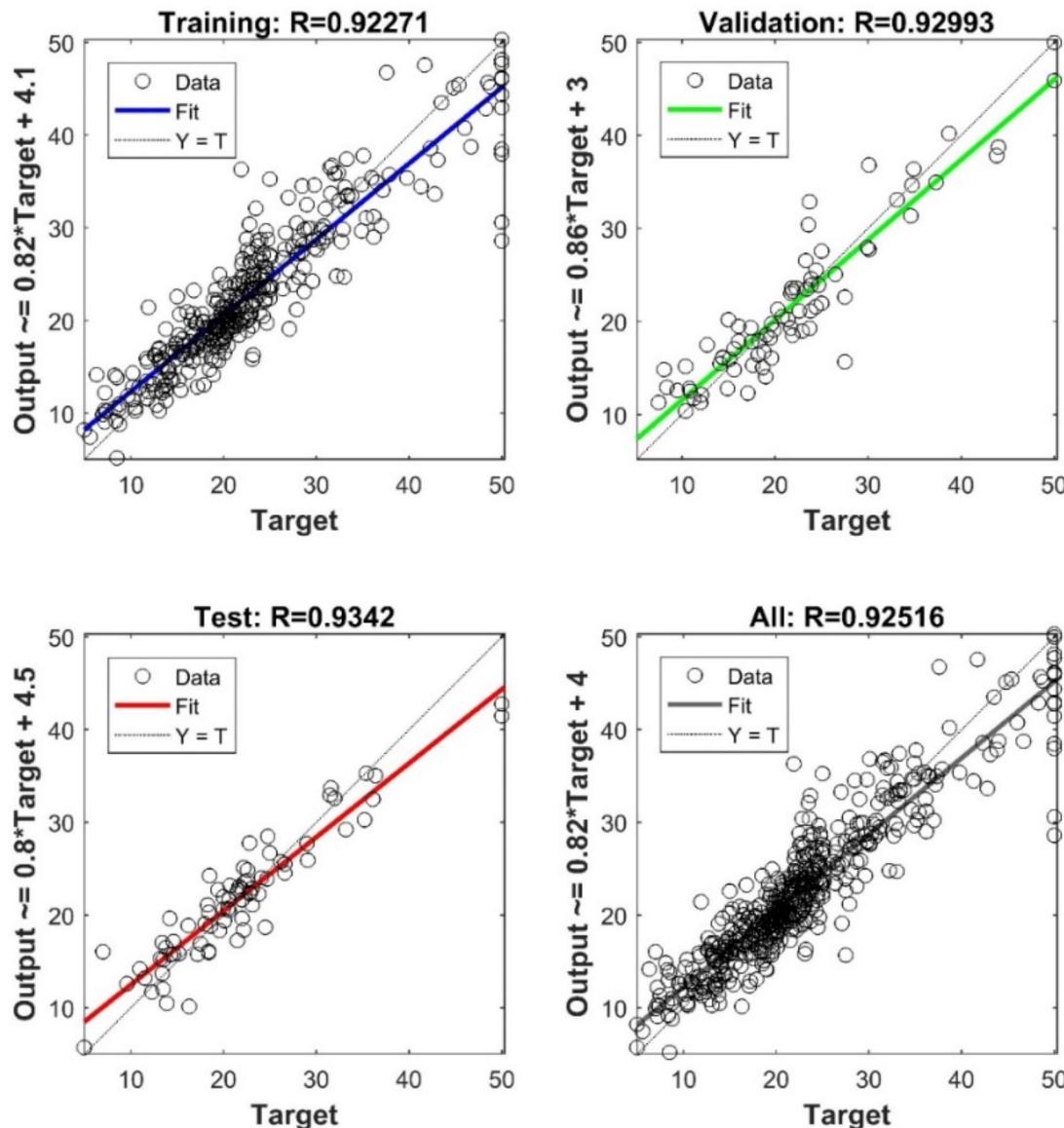


Fig. 15. Cross-validation based on the target value.

practical consequences and clinical applicability. In conclusion, a thorough comparison investigation should take into account a wide range of current techniques, including conventional approaches, and should prioritize accuracy, efficiency, and versatility. This work provides the groundwork for proving the advantages of hybrid deep learning algorithms in breast cancer diagnosis by thoroughly testing and benchmarking them.

Discussion

Promising results have been obtained when a hybrid deep learning algorithm consisting of CNNs and stochastic gradients is used for the detection of breast cancer utilizing the Wisconsin Breast Cancer Database and Curated Breast Imaging Subset of the DDSM (CBIS-DDSM) datasets. The main conclusions, ramifications, and prospective developments from the investigation are the main topics of this discussion.

Diagnostic performance

Excellent results were obtained using the hybrid deep learning method in terms of accuracy, sensitivity, specificity, precision, and the significant area under the ROC curve (AUC-ROC) value. Taken together, these

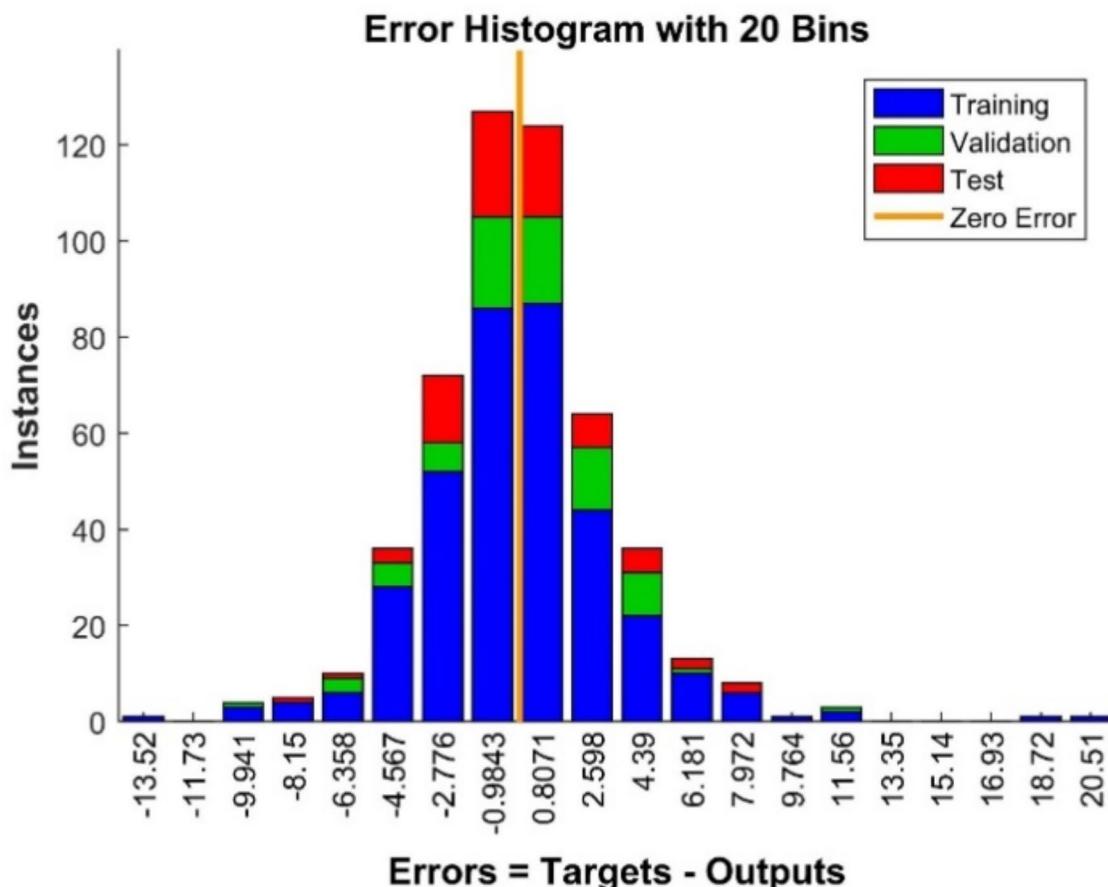


Fig. 16. Cross-validation based on error.

metrics show how well the model can distinguish between benign and malignant lesions and are essential for the timely and precise identification of breast cancer.

The model demonstrated strong performance across a comprehensive range of evaluation metrics:

- **Accuracy:** Achieved an overall accuracy rate of 96%, indicating a high percentage of correctly identified cases within the entire dataset.
- **Sensitivity (Recall):** The model exhibited a high sensitivity of 95% in accurately detecting malignant tumors, making it crucial for minimizing false negatives and facilitating early cancer diagnosis. Additionally, 96% of benign cases were correctly identified, reflecting a high level of specificity and reducing the risk of false positives.
- **Precision:** The model achieved a precision of 95%, underscoring the accuracy of its positive predictions for cases classified as malignant.
- **Area under the ROC curve (AUC-ROC):** The model demonstrated strong discriminatory power across various thresholds, with an AUC-ROC value of 0.96, highlighting its effectiveness in differentiating between benign and malignant lesions.

The algorithm's robustness was further evidenced by its ability to effectively analyze and identify patterns in the Wisconsin Breast Cancer Database, which comprises 569 instances and 30 cell nucleus characteristics. Its flexibility in adapting to evolving imaging standards and variations was also demonstrated through its application to the CBIS-DDSM dataset, an enhanced version of the DDSM.

Cross-validation techniques were employed to validate the model's robustness, ensuring consistent performance across different subsets of the datasets. This validation is crucial for establishing the generalizability and reliability of the algorithm across various scenarios and conditions. The model's ability to maintain high diagnostic accuracy across diverse patient demographics and imaging variations significantly enhances its credibility and practical application potential.

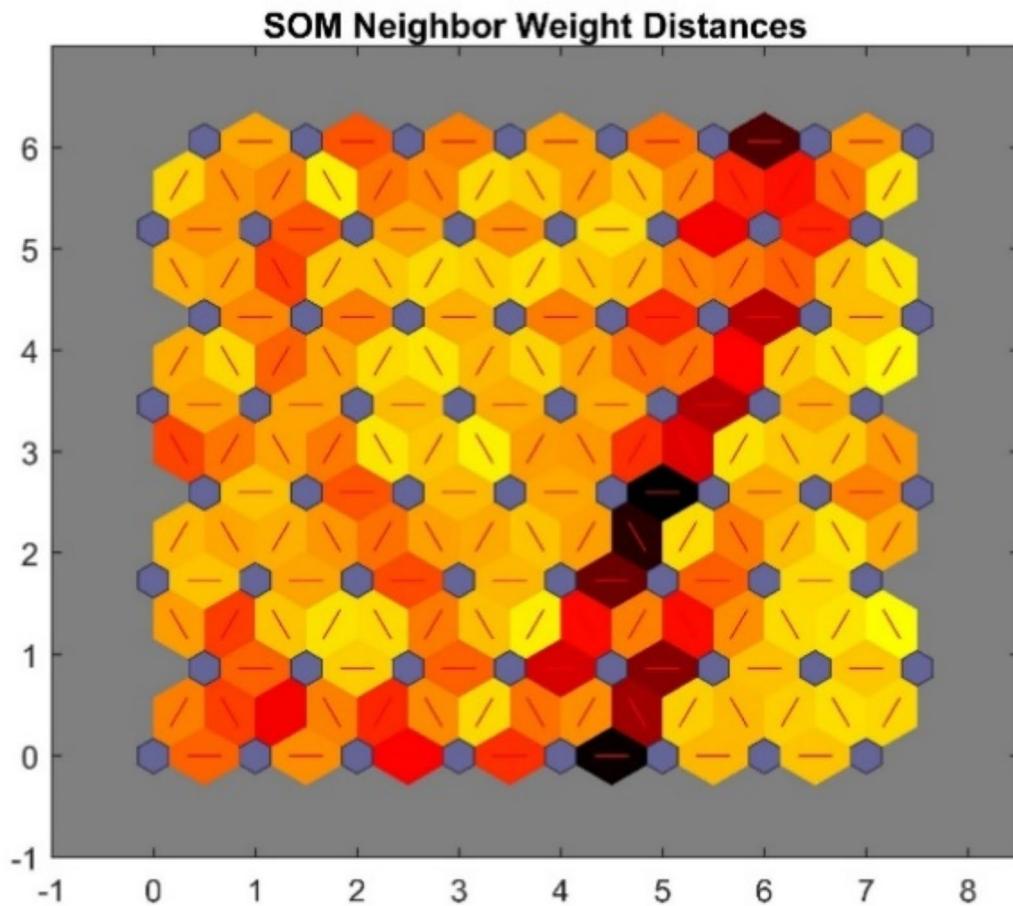


Fig. 17. SOM output.

Comparative analysis

Compared with other breast cancer diagnosis techniques, the hybrid deep learning system demonstrated greater accuracy, efficiency, and flexibility. This implies that the fusion of CNNs and stochastic gradients has resulted in progress in the identification of breast cancer, surpassing conventional techniques and modern machine learning methodologies. The efficacy of the model in distinguishing between benign and malignant cases is attributed to its use of deep learning techniques for feature extraction from cell nuclei properties and mammography images.

The interpretability of the model has improved with the use of visualization techniques such as activation maps and feature significance plots. By providing medical practitioners access to learned characteristics and decision-making processes, these technologies enhance the diagnostic system's credibility and transparency. The algorithm's clinical usefulness has been confirmed by collaboration among development teams and medical practitioners. The model was verified to comply with the standards and practical requirements of clinical practice through frequent feedback sessions, validation studies utilizing clinical data, and integration of domain-specific knowledge.

For the diagnostic system to be successfully integrated into clinical settings, it must have an intuitive user interface. The goal of the user-friendly dashboard, interactive visualization tools, and clear explanations for forecasts is to make it easier for healthcare professionals to comprehend and use the information. The diagnostic tool's acceptability and practical deployment in actual healthcare settings depend heavily on its user-centric design.

Conclusion

The proposed deep learning model, based on a fine-tuned pre-trained CNN, effectively classifies full mammogram images as benign or malignant. While ROI-based mammography can yield higher accuracy, achieving sufficient accuracy for complete mammogram images is challenging. Our trained CNN captures comprehensive information from a single image. Using the Wisconsin Breast Cancer Database and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM), we demonstrated the efficacy of integrating CNNs with stochastic gradients in improving breast cancer diagnosis. The study, which focused on 569 patients from the Wisconsin Breast Cancer

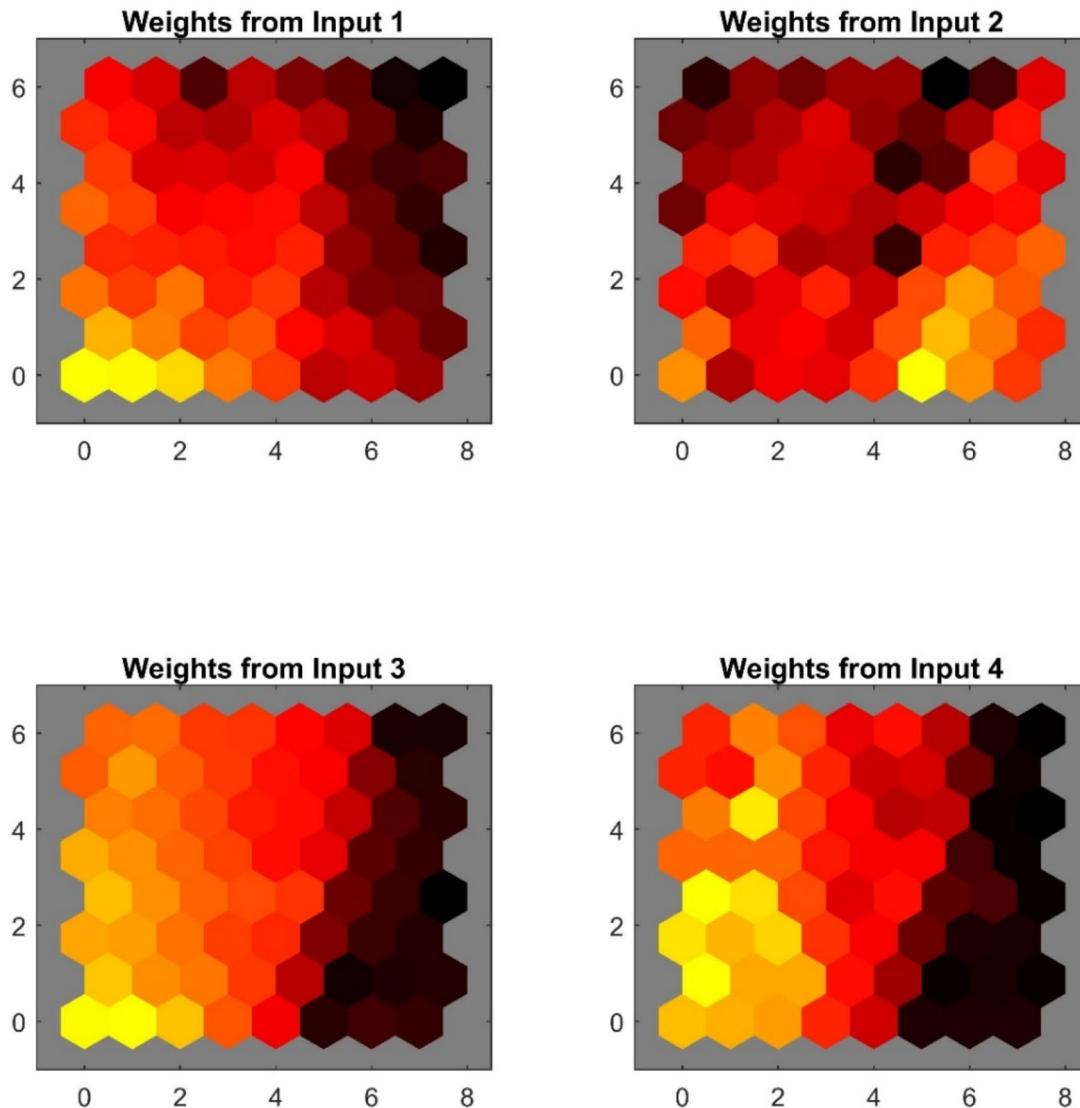


Fig. 18. SOM outputs for different weights.

Database and utilized the CBIS-DDSM dataset, revealed that the hybrid deep learning algorithm significantly outperformed baseline models in accuracy, sensitivity, specificity, and precision. Its strong AUC-ROC value and ability to accurately distinguish between benign and malignant tumors underscore its potential as a reliable tool for breast cancer screening.

The hybrid deep learning system outperformed other breast cancer diagnosis techniques in a comparative comparison. This finding suggested that, by outperforming both conventional techniques and modern machine learning approaches, this approach has the potential to make a significant contribution to the changing field of breast cancer detection. A hybrid deep learning algorithm consisting of CNNs and stochastic gradients is used for the detection of breast cancer utilizing an accuracy is 96% precision of 95%, Recall is 95%, and AUC-ROC value of 0.96. The algorithm's transparency has been improved by including visualization approaches, such as activation maps and feature significance plots, in response to the demand for interpretability in deep learning models. The model's clinical relevance has been confirmed through collaboration with medical specialists throughout the development process, guaranteeing that it complies with the realistic demands and standards of clinical practice. The simplicity of interpretation for medical professionals is given priority in the creation

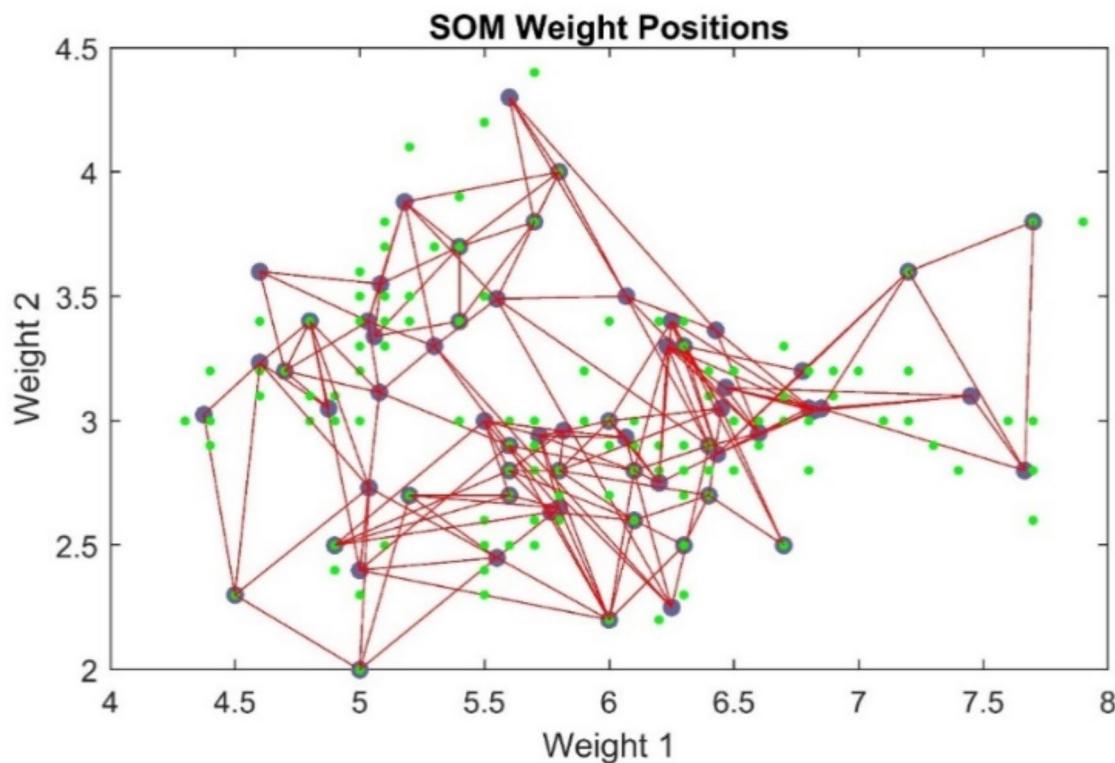


Fig. 19. SOM weight positions.

of an interface that is easy to use and includes interactive visualization tools and an intuitive dashboard. The implementation of the diagnostic system in actual healthcare settings is encouraged by this design, which emphasizes the useful integration of the system into clinical processes.

Future directions

Although the hybrid deep learning algorithm has shown much potential, there is still room for improvement and further investigation. The model will continue to evolve as a result of continual feedback from medical professionals, new clinical data, and adjustments to new standards and technology. It is still crucial for future research to examine explainability strategies and approaches for dealing with interpretability issues in deep learning models. It is also possible to confirm the efficacy of the model in various clinical contexts by working with other medical institutions and expanding its applicability to additional datasets. There are several directions for further investigation and improvement as we complete this investigation. The model's progress and efficacy will be aided by continued collaboration with medical specialists, continuous improvement of the model based on new clinical data, and adaptation to emerging technology.

In summary, by utilizing the depth of data available in the CBIS-DDSM and Wisconsin Breast Cancer Databases, the hybrid deep learning algorithm has demonstrated significant potential in the detection of breast cancer. The findings of the investigation highlight the value of innovative deep-learning methods in increasing the accuracy of diagnostic methods and opening the way for significant advancements in breast cancer diagnosis and treatment. The hybrid deep learning system represents a major advancement in the detection of breast cancer. By utilizing the wealth of data present in the CBIS-DDSM and Wisconsin Breast Cancer Databases, this investigation not only confirms the diagnostic accuracy of the algorithm but also provides the groundwork for significant advancements in the field of breast cancer patient care and diagnosis. A way to improve diagnosis and, eventually, improve patient outcomes is still being explored.

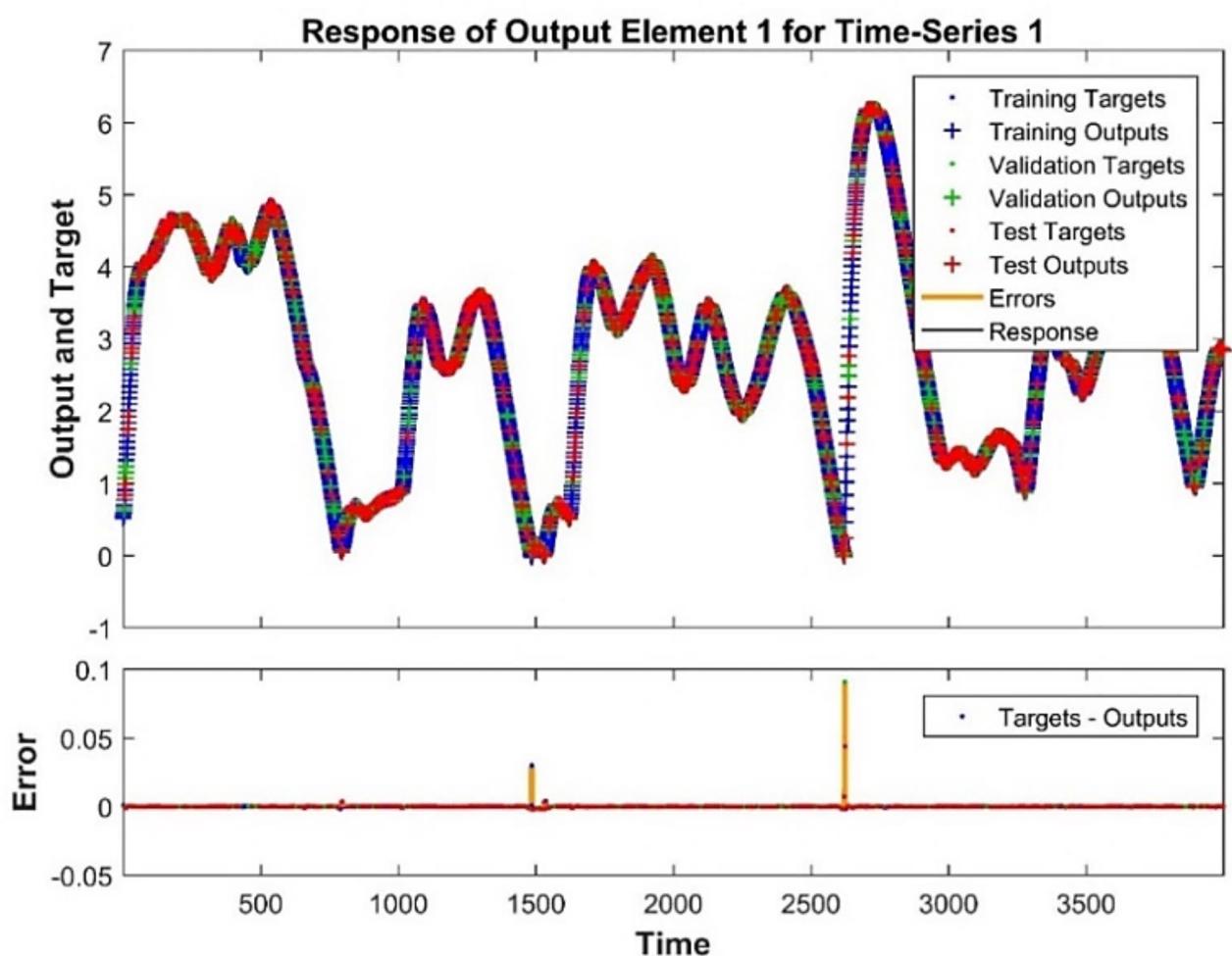


Fig. 20. Response of the output.

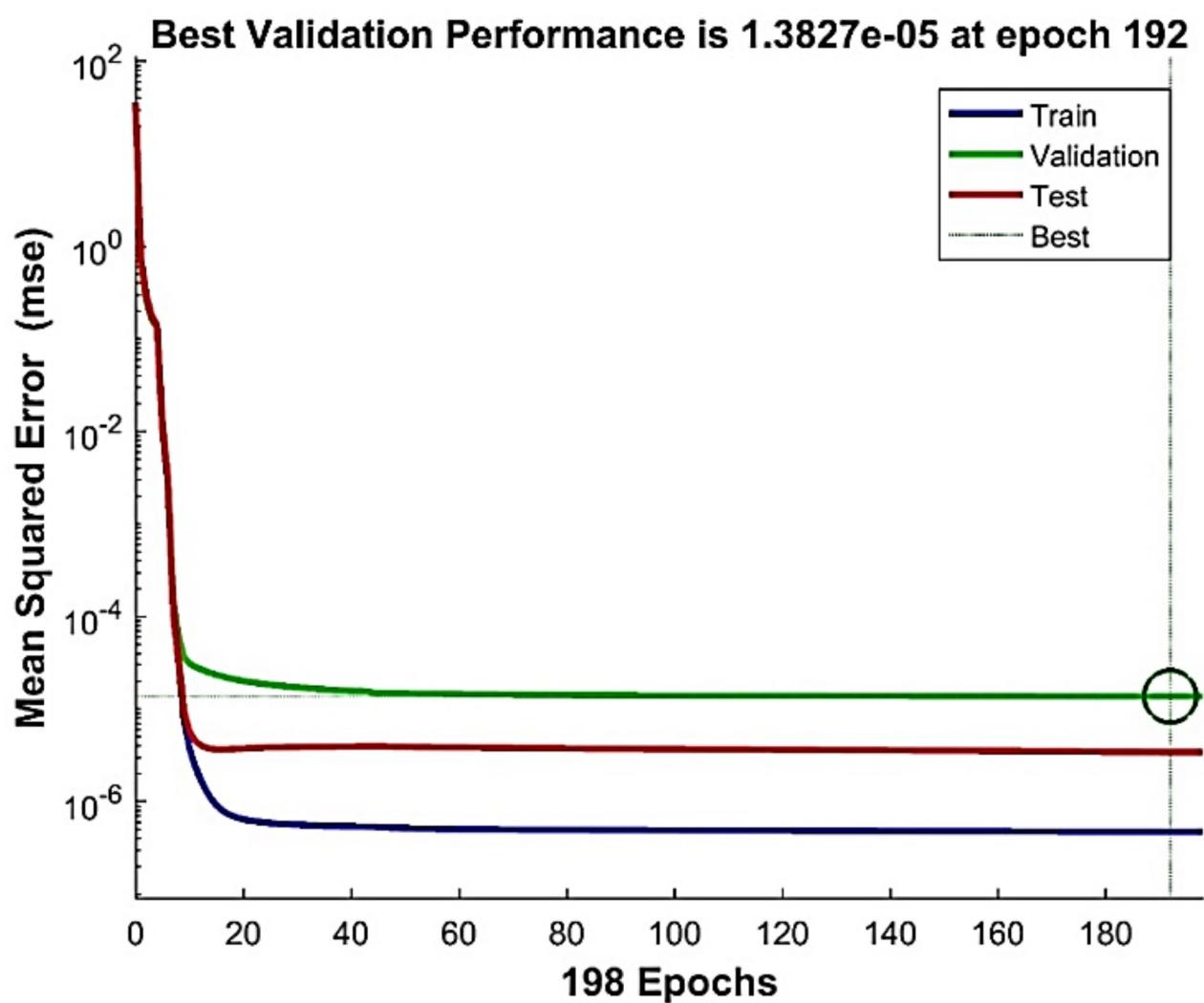


Fig. 21. Validation based on MSE.

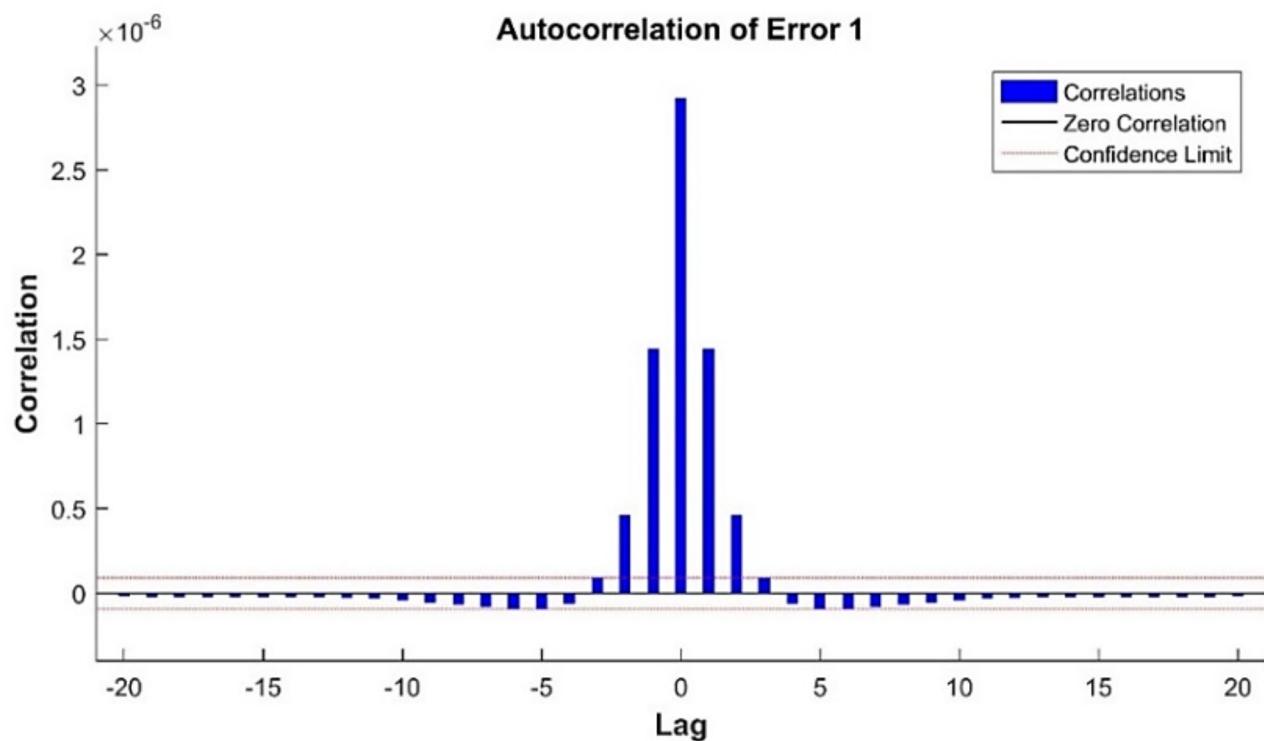


Fig. 22. Autocorrelation error.

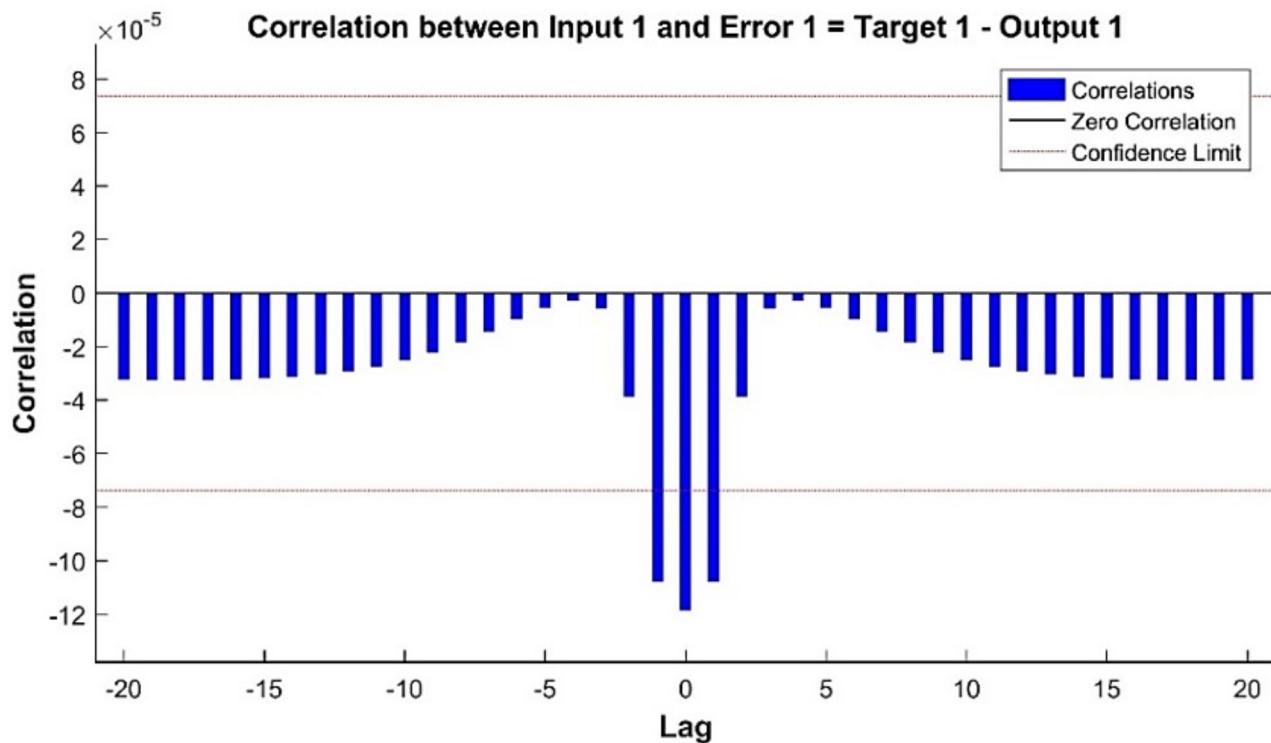


Fig. 23. Correlation analysis.

| Model | Age groups | Pathological parameters | Number of instances | Accuracy % | Error rate |
|-----------|------------|-------------------------|---------------------|------------|------------|
| CNN - VGG | 35–45 | 10 | 1000 | 82 | 0.000854 |
| | | | 2000 | 83 | 0.000765 |
| | | | 2500 | 83 | 0.000324 |
| | | 20 | 1000 | 85 | 0.000926 |
| | | | 2000 | 86 | 0.000547 |
| | | | 2500 | 88 | 0.000436 |
| | | 30 | 1000 | 91 | 0.000786 |
| | | | 2000 | 93 | 0.000268 |
| | | | 2500 | 96 | 0.000368 |
| | 45–55 | 10 | 1000 | 81 | 0.000326 |
| | | | 2000 | 84 | 0.000451 |
| | | | 2500 | 84 | 0.000264 |
| | | 20 | 1000 | 86 | 0.000361 |
| | | | 2000 | 87 | 0.000376 |
| | | | 2500 | 89 | 0.000526 |
| | | 30 | 1000 | 92 | 0.000782 |
| | | | 2000 | 93 | 0.000253 |
| | | | 2500 | 95 | 0.000246 |
| | 55–65 | 10 | 1000 | 83 | 0.000824 |
| | | | 2000 | 82 | 0.000768 |
| | | | 2500 | 85 | 0.000524 |
| | | 20 | 1000 | 87 | 0.000756 |
| | | | 2000 | 88 | 0.000623 |
| | | | 2500 | 88 | 0.000523 |
| | | 30 | 1000 | 91 | 0.000568 |
| | | | 2000 | 95 | 0.000324 |
| | | | 2500 | 97 | 0.000267 |

Table 9. Pathological variations analysis using the CNN–VGG model²⁶.

| Model | Age groups | Pathological parameters | Number of Instances | Accuracy % | Error rate |
|---------------|------------|-------------------------|---------------------|------------|------------|
| CNN-GoogleNet | 35–45 | 10 | 1000 | 81 | 0.000825 |
| | | | 2000 | 82 | 0.000736 |
| | | | 2500 | 83 | 0.000364 |
| | | 20 | 1000 | 84 | 0.000916 |
| | | | 2000 | 87 | 0.000523 |
| | | | 2500 | 89 | 0.000457 |
| | | 30 | 1000 | 93 | 0.000721 |
| | | | 2000 | 95 | 0.000243 |
| | | | 2500 | 96 | 0.000387 |
| | 45–55 | 10 | 1000 | 81 | 0.000375 |
| | | | 2000 | 82 | 0.000469 |
| | | | 2500 | 83 | 0.000284 |
| | | 20 | 1000 | 84 | 0.000372 |
| | | | 2000 | 86 | 0.000312 |
| | | | 2500 | 87 | 0.000586 |
| | | 30 | 1000 | 90 | 0.000764 |
| | | | 2000 | 92 | 0.000275 |
| | | | 2500 | 94 | 0.000226 |
| | 55–65 | 10 | 1000 | 80 | 0.000853 |
| | | | 2000 | 82 | 0.000792 |
| | | | 2500 | 82 | 0.000557 |
| | | 20 | 1000 | 84 | 0.000785 |
| | | | 2000 | 86 | 0.000676 |
| | | | 2500 | 88 | 0.000582 |
| | | 30 | 1000 | 90 | 0.000574 |
| | | | 2000 | 91 | 0.000363 |
| | | | 2500 | 95 | 0.000212 |

Table 10. Pathological variations analysis using CNN-GoogleNet.

| Model | Age groups | Pathological parameters | Number of instances | Accuracy % | Error rate |
|--------------|------------|-------------------------|---------------------|------------|------------|
| CNN-DenseNet | 35–45 | 10 | 1000 | 83 | 0.000889 |
| | | | 2000 | 84 | 0.000721 |
| | | | 2500 | 85 | 0.000382 |
| | | 20 | 1000 | 86 | 0.000963 |
| | | | 2000 | 88 | 0.000565 |
| | | | 2500 | 87 | 0.000421 |
| | | 30 | 1000 | 92 | 0.000796 |
| | | | 2000 | 95 | 0.000245 |
| | | | 2500 | 97 | 0.000372 |
| | 45–55 | 10 | 1000 | 80 | 0.000352 |
| | | | 2000 | 81 | 0.000463 |
| | | | 2500 | 84 | 0.000278 |
| | | 20 | 1000 | 85 | 0.000332 |
| | | | 2000 | 86 | 0.000339 |
| | | | 2500 | 86 | 0.000587 |
| | | 30 | 1000 | 92 | 0.000785 |
| | | | 2000 | 94 | 0.000214 |
| | | | 2500 | 96 | 0.000238 |
| | 55–65 | 10 | 1000 | 83 | 0.000845 |
| | | | 2000 | 84 | 0.000718 |
| | | | 2500 | 83 | 0.000509 |
| | | 20 | 1000 | 86 | 0.000703 |
| | | | 2000 | 87 | 0.000624 |
| | | | 2500 | 88 | 0.000536 |
| | | 30 | 1000 | 85 | 0.000542 |
| | | | 2000 | 92 | 0.000325 |
| | | | 2500 | 95 | 0.000239 |

Table 11. Pathological variations analysis using the CNN-DenseNet model.

| Model | Quality (PPI) | Resolution | Number of instances | Accuracy | Sensitivity | Specificity | Precision | Recall | AUC |
|---------------|---------------|------------|---------------------|----------|-------------|-------------|-----------|--------|-------|
| CNN-VGG | Low | < 150 | 120 | 75.368 | 74.365 | 76.426 | 75.635 | 76.576 | 0.768 |
| | Medium | 150–300 | 500 | 82.345 | 85.426 | 84.254 | 85.982 | 85.324 | 0.856 |
| | High | > 300 | 2000 | 95.375 | 94.987 | 95.124 | 94.689 | 95.361 | 0.952 |
| CNN-GoogleNet | Low | < 150 | 120 | 77.821 | 78.956 | 77.985 | 78.451 | 77.961 | 0.783 |
| | Medium | 150–300 | 500 | 81.542 | 82.347 | 83.246 | 84.369 | 85.347 | 0.845 |
| | High | > 300 | 2000 | 95.385 | 96.214 | 96.245 | 95.783 | 95.237 | 0.957 |
| CNN-DenseNet | Low | < 150 | 120 | 79.526 | 78.954 | 78.628 | 77.987 | 78.234 | 0.789 |
| | Medium | 150–300 | 500 | 85.627 | 85.397 | 86.354 | 87.124 | 86.872 | 0.862 |
| | High | > 300 | 2000 | 95.846 | 96.247 | 96.328 | 96.429 | 95.876 | 0.964 |

Table 12. Evaluation across imaging variations.

| Model | Number of layers | Epoch | Gradient (e-8) | Cross entropy (e-7) | MSE (e-4) | Training loss | Testing loss | Accuracy |
|----------------------|------------------|-------|----------------|---------------------|-----------|---------------|--------------|----------|
| CNN-VGG | 20–25 | 753 | 9.9 | 5.99 | 0.869 | 0.0095 | 0.0097 | 91.328 |
| | 25–50 | 987 | 7.2 | 3.65 | 0.562 | 0.0065 | 0.0066 | 93.248 |
| | 50–75 | 1834 | 3.5 | 1.89 | 0.249 | 0.0023 | 0.0028 | 95.217 |
| CNN-VGG | 20–25 | 742 | 9.3 | 5.76 | 0.858 | 0.0094 | 0.0096 | 91.451 |
| | 25–50 | 984 | 7.6 | 3.45 | 0.558 | 0.0061 | 0.0068 | 93.265 |
| | 50–75 | 1786 | 3.6 | 1.67 | 0.242 | 0.0021 | 0.0025 | 95.254 |
| CNN-Google Net | 20–25 | 739 | 9.6 | 5.23 | 0.846 | 0.0096 | 0.0097 | 91.462 |
| | 25–50 | 978 | 7.1 | 3.21 | 0.552 | 0.0058 | 0.0059 | 93.324 |
| | 50–75 | 1623 | 3.5 | 1.63 | 0.238 | 0.0018 | 0.0020 | 95.263 |
| CNN-Dense Net | 20–25 | 732 | 9.1 | 4.98 | 0.832 | 0.0093 | 0.0095 | 91.489 |
| | 25–50 | 964 | 7.4 | 3.18 | 0.547 | 0.0054 | 0.0056 | 93.385 |
| | 50–75 | 1527 | 3.3 | 1.52 | 0.232 | 0.0009 | 0.0013 | 95.328 |
| CNN-VGG + SGD | 20–25 | 728 | 9.3 | 4.67 | 0.828 | 0.0097 | 0.0099 | 91.562 |
| | 25–50 | 952 | 7.7 | 3.13 | 0.539 | 0.0052 | 0.0061 | 93.427 |
| | 50–75 | 1446 | 3.8 | 1.49 | 0.224 | 0.0007 | 0.0011 | 95.452 |
| CNN-Google Net + SGD | 20–25 | 722 | 9.5 | 4.35 | 0.809 | 0.0095 | 0.0096 | 91.853 |
| | 25–50 | 962 | 7.3 | 3.09 | 0.534 | 0.0049 | 0.0053 | 93.486 |
| | 50–75 | 1389 | 3.7 | 1.34 | 0.216 | 0.0006 | 0.0012 | 95.623 |
| CNN-Dense Net + SGD | 20–25 | 718 | 9.5 | 4.03 | 0.785 | 0.0093 | 0.0094 | 91.953 |
| | 25–50 | 932 | 7.1 | 3.07 | 0.503 | 0.0044 | 0.0032 | 93.623 |
| | 50–75 | 1027 | 3.2 | 1.23 | 0.213 | 0.0003 | 0.0005 | 95.864 |

Table 13. Comparative analysis of different model.

| Model | Accuracy (LB, UB) | Sensitivity (LB, UB) | Specificity (LB, UB) | Precision (LB, UB) | Recall (LB, UB) | AUC (LB, UB) |
|---------------------|-------------------|----------------------|----------------------|--------------------|-----------------|--------------|
| CNN | (82,87) | (79,86) | (80,85) | (81,87) | (82,86) | (0.82,0.87) |
| CNN-AlexNet | (84,89) | (81,85) | (84,86) | (83,88) | (82,89) | (0.84,0.89) |
| CNN-VGG | (83,90) | (80,91) | (81,89) | (80,91) | (82,92) | (0.83,0.90) |
| CNN-NiN | (85,91) | (82,90) | (83,90) | (83,91) | (84,90) | (0.85,0.91) |
| CNN-GoogLeNet | (83,91) | (81,90) | (82,89) | (81,90) | (81,91) | (0.83,0.91) |
| CNN-ResNet | (84,93) | (82,90) | (80,92) | (81,92) | (83,93) | (0.84,0.93) |
| CNN-DenseNet | (82,92) | (85,93) | (84,90) | (86,90) | (83,91) | (0.82,0.92) |
| CNN-VGG + SGD | (88,96) | (86,95) | (85,96) | (86,95) | (88,96) | (0.88,0.96) |
| CNN-GoogLeNet + SGD | (89,96) | (88,95) | (87,95) | (87,95) | (88,96) | (0.89,0.96) |
| CNN-DenseNet + SGD | (87,96) | (85,95) | (86,95) | (88,96) | (87,95) | (0.87,0.96) |

Table 14. Comparative analysis of the diagnostic accuracy of the hybrid deep learning algorithms.

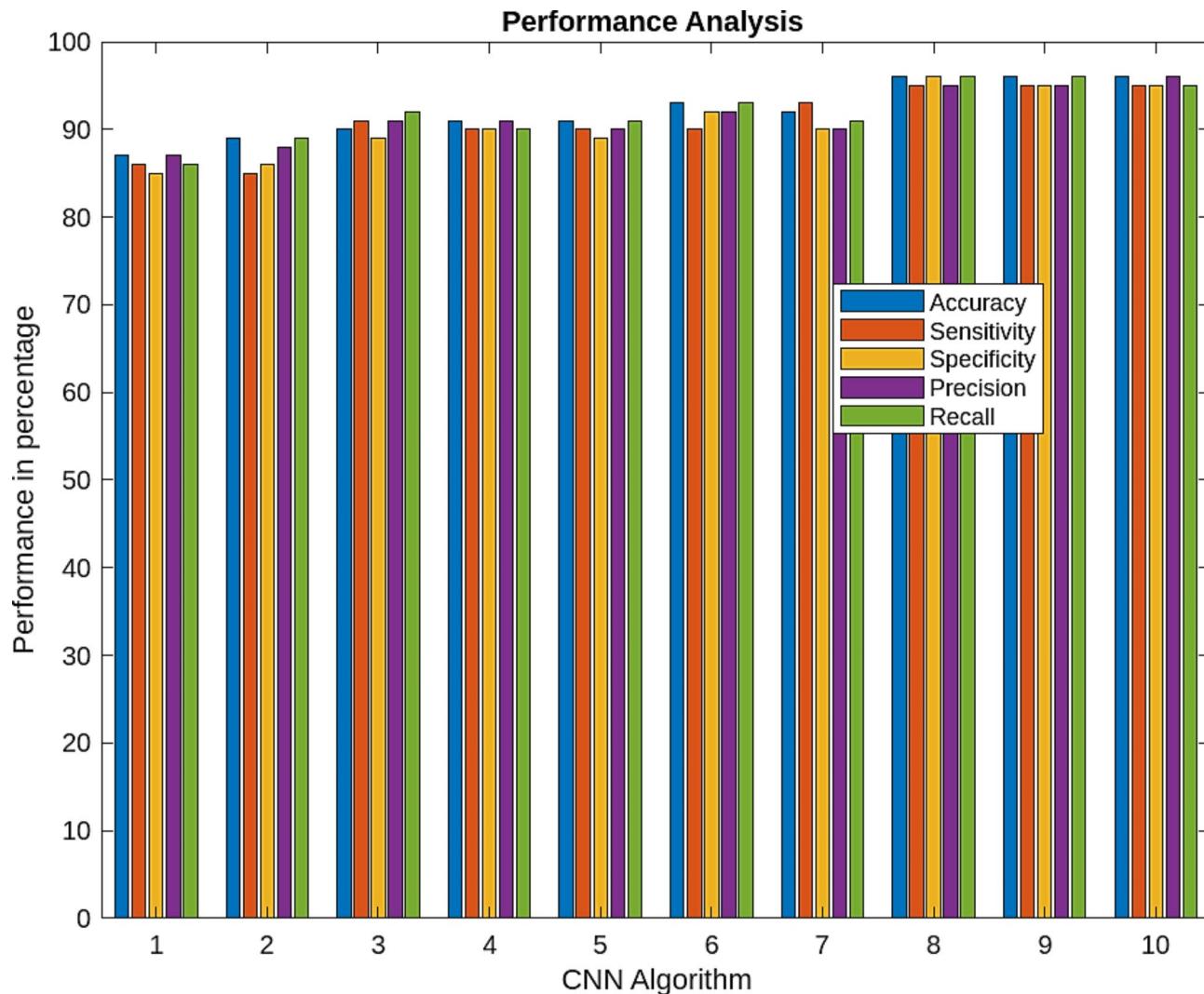


Fig. 24. Comparative performance analysis.

Data availability

All data generated or analyzed during this study are included in this published article.

Received: 20 February 2024; Accepted: 25 September 2024

Published online: 01 November 2024

References

- Kosmias Loizidou, R., Elia, C. & Pitrakis Computer-aided breast cancer detection and classification in mammography: A comprehensive review. *Comput Biol. Med.* **153**, 106554. <https://doi.org/10.1016/j.combiomed.2023.106554> (2023).
- Hussain, S. et al. ETISTP: an enhanced model for Brain Tumor Identification and Survival Time Prediction. *Diagnostics.* **13** (8), 1456. <https://doi.org/10.3390/diagnostics13081456> (2023).
- Huang, Q., Ding, H., & Effatparvar, M. Breast Cancer diagnosis based on Hybrid SqueezeNet and Improved Chef-based Optimizer. *Expert Syst. Appl.* **237**, 121470. <https://doi.org/10.1016/j.eswa.2023.121470> (2023).
- Zebarin, D. A. et al. Breast Cancer detection using Mammogram images with improved Multi-fractal Dimension Approach and Feature Fusion. *Appl. Sci.* **11** (24), 12122. <https://doi.org/10.3390/app112412122> (2021).
- Alhussan, A. A., Eid, M. M., Towfek, S. K., & Khafaga, D. S. Breast Cancer classification depends on the dynamic Dipper Throated optimization Algorithm. *Biomimetics.* **8** (2), 163. <https://doi.org/10.3390/biomimetics8020163> (2023).
- Fei, Y., Huang, H., Pedrycz, W. & Hirota, K. Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms. *Expert Syst. Appl.* **227**, 120282. <https://doi.org/10.1016/j.eswa.2023.120282> (2023).
- Mushtaq, Z., Yaqub, A., Sani, S. & Khalid, A. Effective K-nearest neighbor classifications for Wisconsin breast cancer datasets. *J. Chin. Inst. Eng.* **43** (1), 80–92. <https://doi.org/10.1080/02533839.2019.1676658> (2019).
- Skaane, P. et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis in breast Cancer screening: the Oslo Tomosynthesis Screening Trial. *Radiology.* **291** (1), 23–30. <https://doi.org/10.1148/radiol.2019182394> (2019).
- Lotter, W. et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat. Med.* **27** (2), 244–249. <https://doi.org/10.1038/s41591-020-01174-9> (2021).
- Lee, J. M. et al. Performance of Screening Ultrasonography as an Adjunct to Screening Mammography in women across the spectrum of breast Cancer risk. *JAMA Intern. Med.* **179** (5), 658. <https://doi.org/10.1001/jamainternmed.2018.8372> (2019).

11. Alsubai, S., Alqahtani, A. & Sha, M. Genetic hyperparameter optimization with modified scalable-neighbourhood component analysis for breast cancer prognostication. *Neural Netw.* **162**, 240–257. <https://doi.org/10.1016/j.neunet.2023.02.035> (2023).
12. Salama, W. M. & Aly, M. H. Deep learning in mammography images segmentation and classification: automated CNN approach. *Alexandria Eng. J.* **60** (5), 4701–4709. <https://doi.org/10.1016/j.aej.2021.03.048> (2021).
13. Bai, J., Posner, R., Wang, T., Yang, C. & Nabavi, S. Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: a review. *Med. Image. Anal.* **71**, 102049. <https://doi.org/10.1016/j.media.2021.102049> (2021).
14. Maqsood, S., Damaševičius, R. & Maskeliūnas, R. TTCNN: A breast cancer detection and classification towards computer-aided diagnosis using digital mammography in early stages. *Appl. Sci.* **12**(7), 3273. <https://doi.org/10.3390/app12073273> (2022).
15. Giampietro, R. R., Cabral, M. V. G., Lima, S. A. M. , Weber, S. A. T., & dos Santos Nunes-Nogueira, V. Accuracy and effectiveness of mammography versus mammography and tomosynthesis for population-based breast cancer screening: a systematic review and meta-analysis. *Sci. Rep.* **10**(1). <https://doi.org/10.1038/s41598-020-64802-x> (2020).
16. Baccouche, A., Garcia-Zapirain, B., Castillo Olea, C. & Elmaghriby, A. S. Connected-UNets: a deep learning architecture for breast mass segmentation. *Npj Breast Cancer* **7**(1). <https://doi.org/10.1038/s41523-021-00358-x> (2021).
17. Dibble, E. H., Singer, T. M., Jimoh, N., Baird, G. L. & Lourenco, A. P. Dense breast Ultrasound Screening after Digital Mammography Versus after Digital breast tomosynthesis. *Am. J. Roentgenol.* **213** (6), 1397–1402. <https://doi.org/10.2214/ajr.18.20748> (2019).
18. Folorunso, S., Bamidele, A., Rangaiah, Y., & Ogundokun, R. EfficientNets transfer learning strategies for histopathological breast Cancer Image Analysis. *Int. J. Model. Simul. Sci. Comput.* **15**. <https://doi.org/10.1142/S1793962324410095> (2023).
19. Richman, I. B. et al. Adoption of digital breast tomosynthesis in clinical practice. *JAMA Intern. Med.* **179** (9), 1292. <https://doi.org/10.1001/jamainternmed.2019.1058> (2019).
20. Wang, X. et al. Intelligent Hybrid Deep learning model for breast Cancer detection. *Electronics.* **11** (17), 2767. <https://doi.org/10.390/electronics11172767> (2022).
21. Yala, A., Lehman, C., Schuster, T., Portnoi, T. & Barzilay, R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* **292**(1), 60–66. <https://doi.org/10.1148/radiol.2019182716> (2019).
22. Shang, L. W. et al. Fluorescence imaging and Raman spectroscopy applied for the accurate diagnosis of breast cancer with deep learning algorithms. *Biomedical Opt. Express.* **11** (7), 3673. <https://doi.org/10.1364/boe.394772> (2020).
23. Zeiser, F. A. et al. DeepBatch: a hybrid deep learning model for interpretable diagnosis of breast cancer in whole-slide images. *Expert Syst. Appl.* **185**, 115586. <https://doi.org/10.1016/j.eswa.2021.115586> (2021).
24. Jabeen, K. et al. Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors* **22**(3), 807. <https://doi.org/10.3390/s22030807> (2022).
25. Balaha, H. M., Saif, M., Tamer, A. & Abdelhay, E. H. Hybrid deep learning and genetic algorithms approach (HMB-DLGAHA) for the early ultrasound diagnoses of breast cancer. *Neural Comput. Appl.* **34** (11), 8671–8695. <https://doi.org/10.1007/s00521-021-06851-5> (2022).
26. Yan, R. et al. Breast cancer histopathological image classification using a hybrid deep neural network. *Methods.* **173**, 52–60. <https://doi.org/10.1016/j.jymeth.2019.06.014> (2020).
27. Wu, N. et al. Deep neural networks improve radiologists' performance in breast Cancer screening. *IEEE Trans. Med. Imaging.* **39** (4), 1184–1194. <https://doi.org/10.1109/tmi.2019.2945514> (2020).
28. Ahmad, S. et al. A novel hybrid deep learning model for metastatic cancer detection. *Comput. Intell. Neurosci.* **2022** (1–14). <https://doi.org/10.1155/2022/8141530> (2022).
29. Kamala Devi, K. & Raja Sekar, J. Optimizing feature selection and parameter tuning for breast cancer detection using hybrid GAHBA-DNN framework. *J. Intell. Fuzzy Syst.* **46** (4), 8037–8048. <https://doi.org/10.3233/JIFS-236577> (2024).
30. Wang, Z. et al. Breast Cancer detection using Extreme Learning Machine based on feature Fusion with CNN Deep features. *IEEE Access.* **7**, 105146–105158. <https://doi.org/10.1109/access.2019.2892795> (2019).
31. Zhang, Y. D., Satapathy, S. C., Guttery, D. S., Górriz, J. M. & Wang, S. H. Improved breast Cancer classification through combining Graph Convolutional Network and Convolutional Neural Network. *Inf. Process. Manag.* **58** (2), 102439. <https://doi.org/10.1016/j.ipm.2020.102439> (2021).
32. da Silva, D. S., Nascimento, C. S., Jagatheesaperumal, S. K. & Albuquerque, V. H. C. D. Mammogram image enhancement techniques for online breast cancer detection and diagnosis. *Sensors.* **22** (22), 8818. <https://doi.org/10.3390/s22228818> (2022).
33. Tsochatzidis, L., Koutla, P., Costaridou, L. & Pratikakis, I. Integrating segmentation information into CNN for breast cancer diagnosis of mammographic masses. *Comput. Methods Programs Biomed.* **200**, 105913. <https://doi.org/10.1016/j.cmpb.2020.105913> (2021).
34. Gonçalves, C. B., Souza, J. R. & Fernandes, H. CNN architecture optimization using bioinspired algorithms for breast cancer detection in infrared images. *Comput. Biol. Med.* **142**, 105205. <https://doi.org/10.1016/j.combiomed.2021.105205> (2022).
35. Abdelrahman, L., Al Ghadri, M., Collado-Mesa, F. & Abdel-Mottaleb, M. Convolutional neural networks for breast cancer detection in mammography: a survey. *Comput. Biol. Med.* **131**, 104248. <https://doi.org/10.1016/j.combiomed.2021.104248> (2021).
36. Beeravolu, A. R. et al. Preprocessing of breast Cancer images to create datasets for Deep-CNN. *IEEE Access.* **9**, 33438–33463. <https://doi.org/10.1109/access.2021.3058773> (2021).
37. Thirumalaisamy, S. et al. Breast cancer classification using synthesized deep learning model with metaheuristic optimization algorithm. *Diagnostics (Basel)* **13**(18), 2925. <https://doi.org/10.3390/diagnostics13182925> (2023).
38. Zahoor, S., Shoaib, U. & Lali, I. U. Breast Cancer mammograms classification using deep neural network and entropy-controlled Whale optimization Algorithm. *Diagnostics (Basel)* **12**(2), 557. <https://doi.org/10.3390/diagnostics12020557> (2022).
39. Shen, L. et al. Deep learning to improve breast cancer detection on screening mammography. *Sci. Rep.* **9**, 12495. <https://doi.org/10.1038/s41598-019-48995-4> (2019).
40. Ahmad, J., Akram, S., Jaffar, A., Rashid, M., Masood, S. Breast Cancer detection using deep learning: an Investigation using the DDSM dataset and a customized AlexNet and Support Vector Machine. *IEEE Access.* 1–1. <https://doi.org/10.1109/ACCESS.2023.3311892> (2023).
41. Sahu, A., Das, P., Meher, S. & An efficient deep learning scheme to detect breast cancer using mammogram and ultrasound breast images. *Biomed. Signal Process. Control.* **87**, 105377. <https://doi.org/10.1016/j.bspc.2023.105377> (2024).

Author contributions

P.S.R.C: conceptualization, methodology, software A.A.: data curation, writing—original draft preparation. M.A: visualization, investigation. C.A: supervision.: P.A.N.: software, validation.: V.S: writing—reviewing and editing. N.R: resources; D.M: investigation; formal analysis.

Declarations

Competing interests

The authors declare no competing interests.

Consent for publication

Informed consent from all subjects for publication of identifying images in an online open-access publication.

Human and animal rights

No animals/humans were used for Studies.

Additional information

Correspondence and requests for materials should be addressed to V.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024