International Journal of Basic and Applied Sciences, 14 (4) (2025) 685-697

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/qnnt4s07
Research paper



Twins^2-Sat: A Taylor Bird Swarm Optimized Twin-Spatially A separable Self-Attention Transformer Model for Highly Accurate Breast Cancer Classification

P. S. Anu Rakhi [1] *, R. S. Rajesh [2]

[1] *Research Scholar, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India*
[2] *Professor, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India*
**Corresponding author E-mail: anurakhi@msuniv.ac.in*

Received: May 21, 2025, Accepted: June 29, 2025, Published: August 27, 2025

Abstract

In light of the growing need for sophisticated diagnostic systems in medical imaging, this work presents a novel framework for the diagnosis of breast cancer that focuses on accurate tissue characterization using texture and attention-based learning. Mammography images of the breast are first preprocessed using an Adaptive Savitzky-Golay Filter (ASGF) to reduce noise and improve important structures. The next step is the precise segmentation of problematic areas using the Markov Random Field (MRF) model, which efficiently defines tumor borders by utilizing spatial dependencies. Following the isolation of the region of interest, texture features are extracted using the Gray-Level Co-occurrence Matrix (GLCM), which uses statistical metrics like energy, contrast, and correlation to capture crucial spatial correlations in pixel intensities. After that, these descriptive characteristics are sent to a TwinS²SAT model, which is an improved Twins-Spatially Separable Self-Attention (SSSA) Transformer that has been tuned using the Taylor Bird Swarm Algorithm (TBSA) to enhance classification performance and learning efficiency. The combination of a lightweight, attention-driven deep architecture with GLCM features allows for a reliable differentiation between benign and malignant tissue patterns. Python software is used to evaluate the proposed model, and the comparison is made with the Self-Attention Transformer (SAT) model. The results show that compared to SAT with 94.99% of accuracy, the TBSA- TwinS²SAT model ranks with excellent accuracy of 97.94%, precision of 98%, and 99% recall, confirming its potential as a useful tool for automated breast cancer diagnosis.

Keywords: *Adaptive Savitzky-Golay Filter (ASGF); Markov Random Field (MRF); GLCM; TwinS²SAT; TBSA.*

1. Introduction

One of the leading causes of death worldwide is cancer, which is defined as the unchecked and destructive proliferation of aberrant body cells [1]. Based on the distinction between benignity and malignancy, there are two main forms of cancer. The benign cancer cell develops very slowly and is frequently non-cancerous [2]. However, malignant tumors are harmful, develop quickly, and use the blood flow to propagate to other parts of the body. The most common cause of mortality for women nowadays is breast cancer, which is also the deadliest type of disease. Breast cancer [3] is hazardous and primarily relies on pathological diagnosis because of its high frequency and widespread nature. Increased cancer survival rates result from accurate breast cancer diagnoses. In order to quickly identify this illness, imaging techniques are being developed [4]. Medical images, which are utilized for screening, diagnosis, and supplementary review to provide field professionals more confidence in their initial conclusion, are essential to diagnose cancer. Imaging modalities include PET/CT, MRI, ultrasound, thermography, histology, and mammography [5]. Every modality has its benefits and drawbacks of its own.

Since medical images are utilized to characterize tumor features in all interval cancer subgroups, they are an invaluable tool for pathological analysis of this illness [6]. Accurately classifying these images is necessary for a correct diagnosis. It has been demonstrated that the prognosis for missing interval cancers is worse than that of actual interval cancers due to their histological tumor features [7]. To extract and compute high-quality image features, more experienced researchers are required, as traditional manual feature extraction and image classification typically need extensive knowledge. Usually, this process is time-consuming and yields unsatisfactory classification outcomes [8]. Since the diagnosis of breast cancer mostly depends on the classification of histopathological imagery, it is imperative in this discipline to create an accurate, automated approach for analyzing these images. Additionally, prior studies have demonstrated the significant clinical implications of histopathological images in the early stages of breast cancer screening [9].

Automated and exact breast cancer classification systems for early detection could be developed using the precise classification of breast images. Thus, accurate histopathological examination is essential for the early identification and diagnosis of breast cancer [10]. Numerous studies on the classification of histological images have been published. Traditional Machine Learning (ML) techniques [11, 12] were first applied to the categorization of breast cancer images decades ago. Deep Learning (DL) [13, 14] is developed to address ML

Copyright © P. S. Anu Rakhi, R. S. Rajesh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

models' accuracy problems. DL has spread widely and been used effectively in numerous applications. Large volumes of data learning are made possible by DL.

Table 1: Summary of Deep Learning Model in Classification of Breast Cancer

Model Used	Merits	Demerits	Classification Tasks	Output Type
EfficientNetV2-ViT [21]	Supports in capturing both local and global features with efficient scaling and reduced training time.	More interpretability is limited due to architectural complexity.	Binary and Multi-class Classification.	Predicted class labels.
Xception+ Transformer [22]	Reduces diagnostic time and variability	Higher computational cost with reduced interpretability.	Binary and Multi-class classification	Predicted class labels.
Ensemble deep learning algorithm [23]	Class-wise weighting improves classification accuracy	Dependent on the quality and balance of class-wise training data	Multi-class classification	Predicted HER2 status category
Dual Decoder Attention ResUNet (DDA-AttResUNet) [24]	Attention mechanism improves focus on tumor regions.	Requires careful tuning for decoder balance and attention weight.	Segmentation of breast tumors in ultrasound images	Segmented tumor region masks
External Attention Transformer [25]	Reduces overfitting with higher classification accuracy.	Necessitates fine-tuning of attention components.	Binary classification	Predicted class label: Benign or Malignant

Swin Transformer and depth-wise separable convolutional networks are combined in this hybrid multi-class model [15] to capture both local and global data. High accuracy, efficient multi-dataset generalization, and decreased computing complexity by patch merging are some of its advantages. It might, however, have issues with memory requirements and poor performance on tiny datasets. Support Vector Machine (SVM) for final classification of HER2-stained breast cancer images is used in conjunction with a ResNet50-based transfer learning model [16]. The model shows resilient performance on heterogeneous entire slide images, low model complexity, and good accuracy. However, its drawbacks include a greater preprocessing overhead because of image tiling and augmentation, as well as a dependence on pre-trained models that might restrict domain-specific adaptability. An improved ResNet-based deep learning model called FastLeakyResNet-CIR is presented in [17] to detect breast cancer from microscopic pictures with accuracy and efficiency, but suffers from potential overfitting on smaller datasets and restricted interpretability because of its deep design. LMHistNet is a hybrid convolutional neural network that combines asymmetric convolutions, Levenberg-Marquardt optimization, and a convolution block attention module provides fast convergence [18], efficient handling of binary and multi-class tasks, and enhanced feature discrimination via batch normalization. However, the model is more intricate and exhibits sensitivity to hyperparameter adjustment at different magnification levels.
Attention-Guided Deep Broad Convolutional Neural Network (CNN) (ADBNet) [19] combines a deep broad block with a modified convolutional block attention module, offering extraordinarily high classification accuracy, good handling of magnification variance, but resilience on synthetic data, which could cause distributional shifts if improperly evaluated. 1D CNN with Long Short-Term Memory (LSTM) networks [20] is used to identify and categorize breast cancer. It is useful for practical medical applications because of its advantages, which include high accuracy, enhanced feature extraction, minimal computing time, and great interpretability. Its drawbacks include a probable inability to handle spatial image data directly and a reliance on the caliber of sequential characteristics that have already been retrieved. HoVer-Transformer [26] uses spatial relationships and anatomical priors to provide an interpretable, ROI-free technique for diagnosing breast cancer. Compared to traditional deep models, it enhances automation and explainability. It is still necessary to address robustness and real-world clinical validation under various imaging situations. The RI-ViT model [27] efficiently extracts both local and global information for the identification of breast cancer from histopathology images by combining residual networks with Vision Transformers. It overcomes the drawbacks of manual diagnosis and improves performance in complicated cases. However, generalizability across various clinical settings is impacted by dependence on the unbalanced dataset. By utilizing Vision Transformers with multi-head external attention, minimizing inductive bias, and capturing global dependencies, the BMEA-ViT model [28] improves the diagnosis of breast cancer. It provides less computational complexity and better generalizability; nevertheless, transformer-based models still take a lot of processing power to train and require large datasets to function at their maximum. The accuracy of breast cancer diagnosis is improved by the multi-resolution ViT framework [29] with ensemble decision-making, which efficiently captures both fine and high-level tumor features. Stacking and post-processing enhance prediction quality and resilience. However, the intricacy of architecture and its dependence on several magnifications raise the computing cost and make real-time implementation more difficult. The ETECADx framework [30] uses both binary and multi-class classification to improve breast cancer diagnosis by combining ensemble CNNs and Vision Transformers, demonstrating increased accuracy. However, scalability in resource-constrained environments is limited by the model's intricacy and reliance on sizable, well-annotated datasets. Breast SwinFedNetX [31] effectively utilizes federated learning and Swin Transformers to provide high diagnosis accuracy while protecting patient privacy. Its real-time deployment and explainable AI (Grad-CAM) improve usability and therapeutic trust. Adoption in healthcare settings with limited resources, however, is constrained by the system's intricacy and dependence on a large infrastructure for secure federated training. Furthermore, for optimal tuning of hyperparameters, Ant Colony Optimization (ACO) [32] efficiently chooses a small subset of important genes from high-dimensional datasets, with improved accuracy in detecting breast cancer. It lowers computational complexity while improving classifier performance. For very large search spaces, ACO encounters difficulties such as sluggish convergence or becoming stuck in local optima. Improved quantum-based Grey Wolf Optimization (GWO) with Support Vector Machine (SVM) [33] enhances the MIAS dataset's sensitivity, specificity, and accuracy in predicting breast cancer. The Teaching–Learning Based Optimization (TLBO) with Salp Swarm Algorithm (SSA) [34] effectively lowers feature space for breast cancer diagnosis, improving classification performance. It successfully combines the learning-based optimization of TLBO with the exploratory capabilities of SSA. The method's reliance on several evaluation steps increases processing overhead. However, scalability in big clinical applications is impacted by the algorithm's complexity and possible sensitivity to parameter adjustment. With good accuracy, the Particle Swarm Optimized (PSO) -CNN technique [35] efficiently automates hyperparameter tweaking and architecture selection for breast cancer diagnosis. Without requiring manual trial-and-error, it improves CNN performance. However, for vast search spaces or complex CNN structures, PSO could need a lot of processing power. To overcome these limitations, the proposed research introduced TBSA for optimal tuning of hyperparameters. Table 1 provides an overview of some of the existing deep learning models in the classification of breast cancer. Addressing the limitations of the existing classification model is crucial for more precise diagnosis. As a consequence, the proposed research comes forth with the following contributions:

- Advanced preprocessing using ASGF enhances the quality of the image by smoothing the noise with preserved edges.
- Robust segmentation with MRF accurately isolates tumor regions and reduces false boundaries.

- Effective extraction of features using GLCM to capture spatial relationships and tissue heterogeneity.
- A novel TwinS^2-SAT model for effective classification of breast cancer with improved learning rate.
- Optimization using TBSA tunes the TwinS^2-SAT model, offering enhanced convergence speed, reducing overfitting, and improving classification accuracy.

2. Description of proposed work

The proposed system for the detection of breast cancer classification initiates by utilizing a dataset of breast cancer images as input. The input dataset of breast cancer is preprocessed by ASGF, which effectively smooths the image by removing noise, preserving edges, and textures. Followed by preprocessing, MRF segmentation is done to accurately outline the Region of Interest (ROI), such as tumor boundaries.

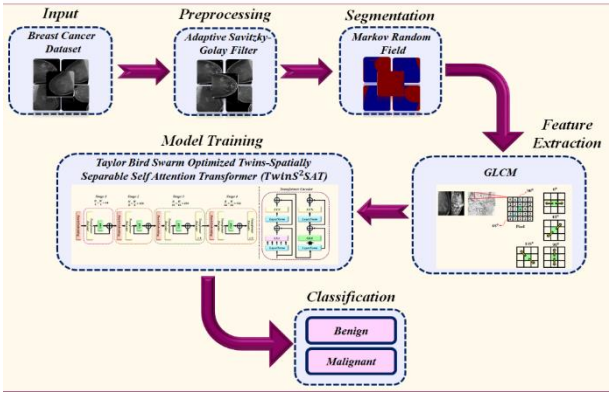


Fig. 1: Proposed System Architecture for Classification of Breast Cancer.

After completion of segmentation, features are extracted using the GLCM method, which examines the spatial correlation between pixel intensities. This GLCM approach determines key features that support differentiating benign and malignant patterns. At the last stage, after extracting important features, model training is performed using a novel deep learning model with the combination of TBSA with TwinS^2-SAT. The proposed TwinS^2-SAT model, with its spatially separable self-attention mechanism, captures local and global dependencies within the breast image data leading to classification. The incorporation of the TBSA technique fine-tunes the developed TwinS^2-SAT model leading to more effective and accurate classification.

3. System methodology

3.1. Preprocessing by adaptive Savitzky-Golay filter

ASGF is an advanced de-noising technique that removes high-frequency noise from breast cancer images while preserving significant structural details. As opposed to the traditional Savitzky–Golay filter, which applies a fixed polynomial smoothing across a sliding window, ASGF dynamically adjusts window size M and polynomial order k in responding to local variations in image intensity to avoid over- or under-smoothing. The expression is provided by

$$\mathrm{d} = (\delta_1, \delta_2, \dots, \delta_{u-1}) \tag{1}$$

Where δ denotes the number of extrema segments as u and the nearby local minima/maxima. The variance surrounding peakpoints between the local extrema is thus determined step-by-step by the following equation:

$$\sigma_{(\mathrm{d})} = \frac{1}{(u-1)\sum_{i=1}^{u-1}\delta_i^2 - \bar{\delta}^2} \tag{2}$$

Real variance is calculated using the expression$\gg \varepsilon_1$based on previous values.This supplies the first segment of the following portion; however, the polynomial degree must change depending on the frame size and maximum distribution, while the window must coincide with the spread peak distribution. Consequently, using the fuzzy relation that the equation provides,

$$\left(\mathrm{d}_{\max} \gg \bar{\mathrm{d}}_S\right) = \frac{1}{1+e^{-(\mathrm{d}_{\max} \gg \bar{\mathrm{d}}_S)}} \in [0,1] \tag{3}$$

The observed peak is defined as$\delta_{\max} = \max(\mathrm{d})$, where the average segment length in the current S portions of the peak distribution are expressed as$\bar{\mathrm{d}}_S$. As an illustration, the greater amplitude peak in the noteworthy segment is indicated when$f\left(\mathrm{d}_{\max}, \bar{\mathrm{d}}_S\right) = 1$. To provide the optimal smoothness while considering the structural complexity of the image, the filter uses adaptive logic to select window sizes and segment-specific polynomial orders. This process is done iteratively to smooth out transitions and correct all over-smoothed areas. A de-noised image is the end product, suitable for segmentation.

3.2. Segmentation by Markov random fields

MRFs are more reliable than clustering methods that ignore voxel proximity for image segmentation because they offer appropriate tools to represent spatial interactions between nearby labels. Examine two random fields, $X = (X_s)_{s\in S}$ and $Y = (Y_s)_{s\in S}$, where s is a voxel index is$s = (s_x, s_y, s_z)$and Sdefines collection of all voxels. The 26 neighborhood system is taken into consideration in this work, and the set of

surrounding voxels $\mathcal{S}$ next to the voxel is represented by $N_{\mathcal{S}}$. A set of tissue c lasses is gven by $\Omega = \{1, \ldots, K\}$, where K represents number of classes that are known in advance. A classification accepts values from the label set Ω, and is a realization of the field $X = x$. A realization $Y = y$ of field Y takes its values in $\mathbb{R}$, the field Y is the observable image that characterizes the statistical properties of tissues. The segmentation challenge involves selecting the optimal tissue classification for each voxel based on observed image intensities. Using the observed image y, the Bayes theorem is used to calculate the segmentation probability x:

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{4}$$

Where $P(x)$ represents the prior probability of segmentation x, $P(y|x)$ represents the conditional probability of the image y given segmentation x, and $P(y)$ represents the prior probability of the image y devoid of segmentation x. The ideal segmentation $\hat{x}$ is obtained by applying the maximum a posteriori (MAP) criterion with previous model knowledge:

$$\hat{x} = \text{argmax} P(y|x)P(x) \tag{5}$$

MRF-based segmentation is capable of handling noise, image artifacts, and partial volume effect by combining spatial relationships and intensity information. For the diagnosis, treatment planning, and progression tracking of breast cancer, this leads to a more precise characterization of breast tissues and tumors.

3.3. Extraction of features by GLCM

Following the preprocessing phase, GLCM is used to detect and extract pixel features, which record the texture information of tissues. Pixel intensities that are too sensitive to angle and distance between pixel pairs are analyzed spatially using the GLCM approach.

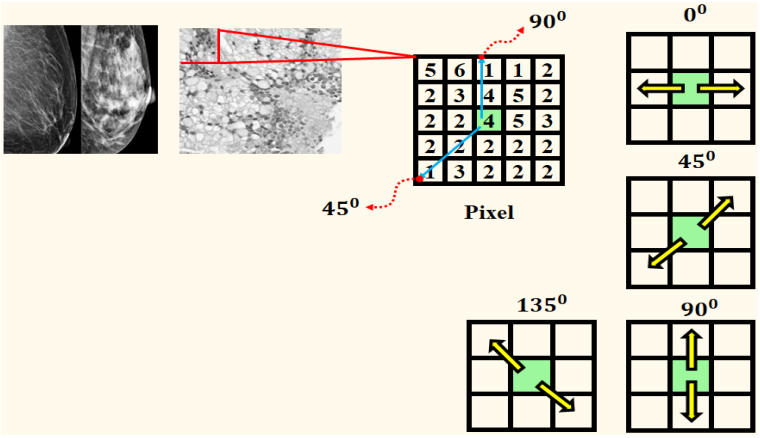


Fig. 2: Illustration of GLCM with Angles in Breast Cancer Image.

Choosing the right distance is crucial because too long a distance produces unnecessary pixel relations, while too short a distance produces homogenous information. Additionally, orientations like 0°, 45°, 90°, and 135° are calculated in order to extract textural information from breast images. To ensure that the co-occurrence matrix adds up to one, each element is partitioned by the total number of collected co-occurrences before GLCM variables are evaluated. By standardizing the matrix, this normalization makes the texture features that are extracted from each image comparable. The normalized GLCM is used to analyze the variables contrast, dissimilarity, homogeneity, energy, and correlation, as illustrated in Fig. 2. These factors improve the proposed work's capacity to accurately categorize tissues based on their texture. The variables are stated as follows:

Contrast is the measurement of intensity between a reference and a nearby pixel at a given distance and angle. A high contrast image is one with a lot of variance. It is stated as

$$\text{Contrast} = \sum_{i,j=0}^{N_{g-1}} p(i,j)(i-j)^2 \tag{6}$$

Dissimilarly is the distance between two pixels in the ROI. Greater intensity fluctuation between neighboring pixels is indicated by larger values. It is established

$$\text{Dissimilarity} = \sum_{i,j=0}^{N_{g-1}} p(i,j)|i-j| \tag{7}$$

In GLCM, homogeneity is the measurement of pixel's proximity distribution, while contrast is a measurement of inverse proportionality. The contrast level drops as homogeneity rises. The statement is provided by

$$\text{Homogeneity} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1+(i-j)^2} \tag{8}$$

Energy is the definition of the root of the angular second moment, and it is written as

$$\text{Energy} = \sum_i \sum_j \{p(i,j)\}^2 \tag{9}$$

In GLCM, correlation a linear dependency of gray level value—describes the local dependency in gray level texture images.

$$\text{Correlation} = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}[(ij)p(i,j)]-\mu_x\mu_y}{\sigma_x\sigma_y} \quad (10)$$

Where the normalized symmetric co-occurrence matrix is denoted by $p(I, j)$. Following feature extraction from the breast image, the data is fed into a TwinS^2-SAT model for classification.

3.4. Classification by TBSA optimized twins2-sat

To classify breast cancer from mammography images, the model must learn both fine-grained local patterns (e.g., cell shapes) and global contextual dependencies. The SSSA mechanism in the TwinS²-SAT model is a novel attention design aimed at balancing computational efficiency and feature learning capabilities, specifically for breast cancer classification. The Spatially Separable Self-Attention module TwinS2-SAT model provides a potent solution by combining two essential elements, including Self-Attention in Local Groups (LSA) for capturing fine-grained patterns and Sub-sampled Global Self-Attention (GSA) for modeling broader contextual dependencies across distant regions shown in Fig. 3. This approach is inspired by the efficiency of separable convolutions, combining local and global perspectives while reducing computational overhead. In the first stage, LSA partitions the feature map into uniform sub-windows (m × n grid) and applies self-attention independently within each window, allowing the model to capture fine-grained, localized features such as cell structures and nuclear abnormalities, while significantly reducing computational complexity compared to full global attention. However, LSA alone cannot model long-range interactions between distant tissue regions. To overcome this limitation, GSA is introduced in the second stage, which extracts representative features from each sub-window to serve as global keys, enabling efficient communication across remote regions. This mechanism preserves essential spatial structure while allowing the model to associate contextual information across the entire image, critical for identifying dispersed or irregular tumor patterns. The combination of LSA and GSA mimics the functionality of depthwise separable convolutions, offering a dual-scale perspective: LSA for capturing micro-level features and GSA for macro-level relationships, all while maintaining lower computational overhead.

3.4.1. Locally-grouped self-attention (LSA)

Inspired by depth-wise convolutions' group-wise architecture for effective inference, the 2D feature maps are uniformly split up into several sub-windows. Self-attention processes are limited to taking place within each sub-window of this structure. Since interactions are restricted to channels within the same attention head, this method is consistent with the multi-head self-attention process. For localized attention processing, the feature maps are specifically divided into $m \times n$ grid of sub-windows. Since there are HW/mnelements in each group, the computation cost of self-attention in this window is $\mathcal{O}\left(\frac{H^2W^2}{m^2n^2}d\right)$, and the total cost is $\mathcal{O}\left(\frac{H^2W^2}{mn}d\right)$. The cost is calculated as $\mathcal{O}(k_1k_2HWd)$ if$k_1 = H/m$ and$k_2 = W/n$. This is much more efficient when $k_1 \ll H$and $k_2 \ll W$ grows linearly with HW if k_1 and k_2 are fixed. While LSA mechanisms successfully process fine-grained patterns within sub-parts (such as cell structures), they fail to account for interactions between distant regions of the tissue if they are not appropriately linked.

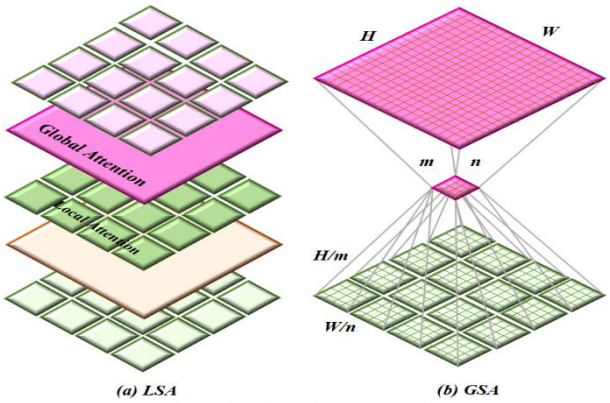


Fig. 3: Schematic View of (A) LSA and (B) GSA.

3.4.2. Global sub-sampled attention (GSA)

The GSA technique is used to get around LSA's location restriction. It is critical to capture both local cellular structures and broad tissue patterns when using mammography image analysis to diagnose breast cancer. Stacking global self-attention layers after local attention blocks is a simple way to facilitate cross-region exchange of data. But this raises the computational complexity to $\mathcal{O}(H^2W^2d)$which is a significant increase. In order to get around, this research uses the most important details from each of the m × n image sub-windows using a single representative feature. By serving as global keys and facilitating effective communication between sub-regions, these representatives significantly reduce computation to $\mathcal{O}(mnHWd) = \mathcal{O}\left(H^2W^2d - k^1k^2\right)$. In order to detect scattered tumor patterns across large biopsy images, the model has to be able to associate distant cancerous and non-cancerous tissue regions. This is made possible via a technique known as GSA. By combining LSA to capture fine-grained cellular morphology inside sub-windows and GSA to relate large-scale tissue architecture, an effect similar to separable convolutions (depth-wise + point-wise), but with lower processing needs, is attained. The ideal sub-window size for standard diagnostic image sizes (e.g., $H = W = 224$) is approximately 15×15. For instance set $k^1 = k^2 = 7$ to ensure consistency in breast cancer image analysis networks across stages of different resolutions. Standard strided convolutions, depth-wise strided convolutions, and average pooling are assessed for the purpose of summarizing local feature representations from high-resolution images. Regular strided convolutions provide the best balance, maintaining crucial morphological and structural patterns required for precise diagnosis, according to empirical evaluation on breast cancer datasets. Spatially Separable Self-Attention (SSSA), the general attention framework, functions as follows:

$$\hat{z}_{ij}^{l} = \mathrm{LSA}\left(\mathrm{LayerNorm}\left(z_{ij}^{l-1}\right)\right) + z_{ij}^{l-1} \tag{11}$$

$$z_{ij}^{l} = \mathrm{FFN}\left(\mathrm{LayerNorm}\left(\hat{z}_{ij}^{l}\right)\right) + \hat{z}_{ij}^{l} \tag{12}$$

$$\hat{z}_{ij}^{l+1} = \mathrm{GSA}\left(\mathrm{LayerNorm}\left(z^{l}\right)\right) + z^{l} \tag{13}$$

$$z_{ij}^{l+1} = \mathrm{FFN}\left(\mathrm{LayerNorm}\left(\hat{z}_{ij}^{l+1}\right)\right) + \hat{z}_{ij}^{l+1} \tag{14}$$

Where $i \in \{1,2,\ldots,m\}, j \in \{1,2,\ldots,n\}$. LSA captures fine-grained local textures, including nuclear and cellular details, within a sub-window. GSA interacts with representative keys from each sub-region$\hat{z}_{ij} \in \mathcal{R}^{k^1 \times k^2 \times C}$ to facilitate information sharing across remote sub-windows. The model effectively learn both local and global properties that are essential for distinguishing between benign and malignant tissue regions owing to this dual methodology. The complete SSSA block combines both LSA and GSA, ensuring both local detail extraction and global context modeling. This mimics the effect of depth-wise separable convolutions, which apply a spatial filter followed by point-wise (1×1) convolution, but does so in the attention domain. It enables the model to effectively learn discriminative patterns across both microscopic and macroscopic scales, which is essential for reliable breast cancer diagnosis. Similar to ordinary self-attention, both LSA and GSA use multi-head attention to ensure thorough identification of features at various scales. Furthermore, Positional Encoding Generator (PEG) is integrated to preserve spatial awareness in high-resolution, variable-length breast cancer images.

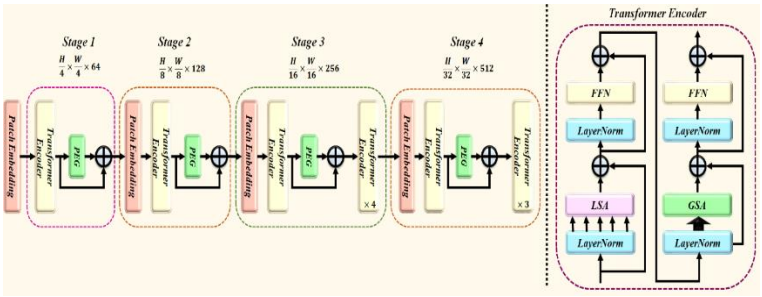


Fig. 4: Architecture of TwinS^2-SAT Model for Breast Cancer Diagnosis.

The TwinS^2-SAT model's architecture for image analysis of breast cancer is depicted in Fig.4. The architecture has a four-stage hierarchical design, with each stage gradually reducing spatial resolution while increasing the depth of feature channels. In order to convert raw pixel data into an initial feature representation, a patch embedding layer first divides the high-resolution breast tissue image into non-overlapping patches. To provide spatial consistency, which is crucial for determining the exact position of diseased characteristics in thick tissue environments, PEG is introduced after the first block in each stage, which is made up of a Transformer Encoder.

Stage 1 reduces the feature map resolution to $\frac{H}{4} \times \frac{W}{4} \times 64$ feature channels; Stage 2 reduces it to $\frac{H}{8} \times \frac{W}{8} \times 128$ channels; Stage 3 reduces it to *H*/ 16×*W*/ 16 H/16×W/16 with 256 channels; and Stage 4 reduces it to $\frac{H}{32} \times \frac{W}{32} \times 512$ channels. GSA and LSA are included into the transformer encoders. By concentrating on local contextual factors, LSA enables the model to identify small nuclear abnormalities and cellular-level alterations, which frequently signify early-stage cancers. Whereas, GSA makes it possible for the model to relate data between remote tissue areas, which is essential for identifying dispersed tumor patterns and preventing the incorrect classification of independent benign anomalies. As the design progresses from precise local textures at higher resolutions to global tissue architecture at lower resolutions, it offers a natural inductive bias for multi-scale feature learning. This multi-resolution ability is well-suited to complicated nature of breast cancer images, which require both micro (cellular) and macro (tissue structural) information for diagnosis. For breast cancer detection tasks, Feed-Forward Neural Networks (FFN) and normalization layers modify the final transformer block outputs, guaranteeing strong feature extraction and excellent classification performance. In order to fine tune the parameters of TwinS^2-SAT model this research introduced TBSA for more precise classification.

3.4.3. Hyperparameter tuning using Taylor bird swarm algorithm

TBSA is utilized in this work to tune hyperparameters of TwinS^2-SAT model for accurate classification in breast cancer diagnosis.

Step 1: Initializing the population and algorithmic parameters is the first stage. Every potential set of hyperparameters for TwinS^2-SAT classifier is represented by the candidate solution$F_{ij}; (1 \le i \le j)$, where j indicates population size, h_{max}indicates maximum iterations, prob indicates likelihood of selecting foraging behavior, and F_t indicates frequency of birds' flight behavior for exploring hyperparameter space.

Step 2: The objective function here is to minimize the validation error of TwinS^2-SAT during training on breast cancer image datasets.

Step 3:Bird positions (hyperparameter sets) are changed in three probability-based phases to explore the hyperparameter space. When a random number$\mathrm{Rand}(0,1) < \mathrm{prob}$The hyperparameters are adjusted in response to foraging behavior; if not, vigilance behavior is used. To improve exploration, the swarm is further separated into producers and scroungers to mimic flight behavior. Following updates, the viability of each candidate hyperparameter set is examined (within predetermined bounds), and the best option at the moment is noted.

Step 4: Each bird, which stands for a set of hyperparameters, adjusts its location by looking for better solutions based on the swarm's global best solution as well as its own past performance (personal best hyperparameters). The updated equation for the hyperparameter optimization process's foraging behavior is given by:

$$F_{i,j}^{h+1} = F_{i,j}^{h} - F_{i,j}^{h}\mathrm{Rand}(0,1)[Z + T] + \mathrm{Rand}(0,1)\left[P_{i,j}Z + Y_j T\right] \tag{15}$$

Algorithm 1

Pseudocode for TBSA Optimized TwinS^2-SAT Model
Input
Initial TBS population $F_{ij}; (1 \leq i \leq j)$
Breast Cancer Dataset
Output
Optimaleper $-$ parameterof TwinS^2 $-$ SATlassifier
Procedure
Begin
Initialize Population of candidate hyperparameter sets $F_{ij}; (1 \leq i \leq j)$
Read Algorithm Parameters
b $\rightarrow$ Population Size, $h_{max} \rightarrow$ Minimum Number of Iterations
prob $\rightarrow$ Probability of Applying Foraging Behavior,
Fl $\rightarrow$ Frequency Factor for Flight Behavior
Evaluate fitness of all candidate hyperparameter sets (using validation error
of TwinS^2 $-$ SAT)
While h $<$ h_{max} do
For each bird k $=$ 1: b
Generate random number rand(0,1)
If rand(0,1) $<$ prob
Apply Foraging behavior to update hyperparameters
Else
Apply Vigilance behavior to update hyperparameters
End if
End For
Else
Split swarm into producers and scroungers
For each bird k $=$ 1: b
If bird k is a producer
Update hyperparameters using equation (17)(explorative flight)
Else
Update hyperparameters using equation (18)(Scrounger Behavior)
End if
End For
End if
Check feasibility of updated hyperparameter sets (ensure within allowed ranges)
Update Personal best and global best solutions based on validation performance
Increment iteration h $=$ h $+$ 1
End While
Return global best hyperparameter set as optimal solution
End

Where $\mathrm{Rand}(0,1)$ is an independent, uniformly distributed number set, Y_j defines the best previous location shared, Z represents the cognitive accelerated coefficients, and T represents the social accelerated coefficients.$F_{l,j}^{h+1}$ and $F_{l,j}^{h}$ indicate the location ofith bird in jth dimension at $(h + 1)$, and h, $P_{l,j}$ refers to the previous best position of ith bird. Z And T These are positive numbers in this case.

Step 5: To simulate competition between configurations, candidate hyperparameter sets (birds) move toward the center of the swarm's which is the most well-known solution. By paying attention to potential areas, this behavior improves hyperparameter fine-tuning. Modeling the hyperparameter optimization vigilance behavior is expressed as

$$F_{i,j}^{h+1} = F_{i,j}^{h} + V_1\left(\mu_j - F_{i,j}^{h}\right) \times \mathrm{Rand}(0,1) + V_2\left(\mathcal{U}_{0j} - F_{i,j}^{h}\right) \times \mathrm{Rand}(-1,1) \quad (16)$$

Where the number of birds is represented as$V$.

Table 2: Hyperparameters of TwinS^2-SAT Model Tuned Using TBSA

Parameter	Range	Parameter	Range
Population Size	$20 - 50$	Maximumterations	$\mathbf{50 - 100}$
Flight Frequency	$0.5 - 1.0$	Foraging Probability	$\mathbf{0.5 - 0.7}$
Acceleration Coefficient (Z)	$0.1 - 0.5$	Acceleration Coefficient (T)	$\mathbf{0.1 - 0.5}$
Learning Rate Search Space	$0.0001 - 0.01$	Batch Size Searchpace	$\mathbf{16 - 128}$
Number of Transformer Encoder Layers	$4 - 12$	Number heads per multi $-$ head attention block	$\mathbf{4 - 12}$
Size of Hidden Layers in Classifier	$\mathbf{256 - 1024}$	**Dropout Rate**	$\mathbf{0.1 - 0.5}$

Step 6: The discovery of new areas in hyperparameter space is simulated in this phase. If existing settings are not ideal or further investigation is needed, birds (candidate hyperparameter sets) fly to new locations. When the birds reach a new point, they look for better combinations of hyperparameters. While some birds are scavengers (following successful solutions), others are producers (actively researching). A mathematical model of flight behavior for tuning is expressed by:

$$F_{i,j}^{h+1} = F_{i,j}^{h} + \mathrm{Randr}(0,1) \times F_{i,j}^{h} \quad (17)$$

$$F_{i,j}^{h+1} = F_{i,j}^{h} + \left(F_{\gamma,j}^{h} - F_{i,j}^{h}\right) \times \mathrm{Fl} \times \mathrm{Rand}(0,1) \quad (18)$$

Gaussian-distributed random number with zero mean and standard deviation is represented as $\mathrm{Rand}(0,1)$ guiding the stochastic exploration of hyperparameter solutions.
Step 7: The TwinS^2-SAT classifier's validation error is used to assess each candidate hyperparameter set's quality. The previously best solution is replaced if a newly calculated hyperparameter configuration produces a lower validation error than the existing best.
Step 8: Until the maximum number of iterations is reached, the population of hyperparameter sets is updated iteratively during the optimization process. The final global best solution is the TwinS^2-SAT model's optimal hyperparameter configuration. The pseudo-code of the proposed TBSA for hyperparameter tuning of TwinS^2-SAT is illustrated in Algorithm 1, and specifications are in Table 2.
The convergence criteria and optimization process of the TBSA for tuning the TwinS²-SAT model are guided by minimizing the validation error during model training. The algorithm begins by initializing a population of candidate hyperparameter sets (birds), with each bird representing a potential solution. During each iteration, birds explore the hyperparameter space using probabilistic behavioural strategies, either foraging or vigilance, based on a predefined probability. The algorithm also simulates social interactions by dividing the swarm into producers and scroungers, enabling both global and local exploration. Each bird updates its position using stochastic equations influenced by its personal best and the global best solution found so far. This process continues iteratively, with the algorithm evaluating the fitness, i.e., validation error of each updated hyperparameter set at every step. Convergence is determined when the maximum number of iterations is reached or when the global best solution stabilizes with minimal improvement over successive iterations. The final output is the hyperparameter configuration that yields the lowest validation error, ensuring optimal performance of the TwinS²-SAT model for breast cancer diagnosis.

4. Results and discussion

In this section, the outcomes of the proposed research are discussed. In this work, Breast Cancer Detection is used, which consists of 3,383 mammography images from Kaggle that are particularly focused on breast cancers, some of which are displayed in Fig. 5. The TwinS^2-SAT classifier has been built and validated using Python software for evaluation. To guarantee an accurate and trustworthy evaluation of model performance, the dataset was split into 80% for training and 20% for testing. For training, 2373 images were used, of which 209 were benign images and 130 were malignant. For testing, 339 images were chosen out of which 1569 are of the benign class and 804 are of the malignant class. These datasets offer greater diversity in imaging protocols, patient demographics, and pathological conditions, thus presenting a more challenging and realistic diagnostic landscape. By evaluating the model on datasets like DDSM, which contains scanned film mammograms with verified biopsy outcomes, or INbreast, which includes full-field digital mammograms with precise lesion annotations, the robustness and adaptability of TwinS²-SAT across varying clinical environments can be validated. Such cross-dataset evaluation not only helps mitigate overfitting to specific data distributions but also enhances confidence in the model's deployment in heterogeneous healthcare systems, ultimately improving its translational value in breast cancer screening and diagnosis. To maximize accuracy and robustness in tumor classification, the assessment between the SAT model and the proposed TwinS^2-SAT after preprocessing is carried out, along with normalization of all images to match the classifier's input specifications.

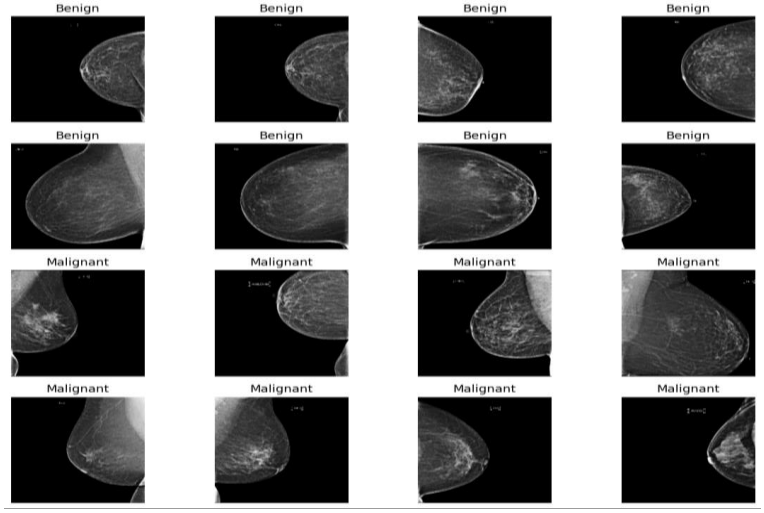


Fig. 5: Radom Images from Dataset.

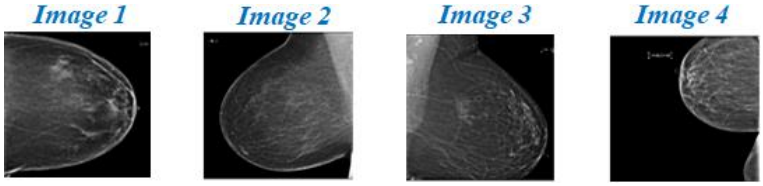


Fig. 6: Input Image Under Test.

A representative set of input mammography images used for the TwinS2-SAT model classification is shown in Fig. 6. Images 1 and 2 represent benign examples, in which there are no notable anomalies or distinct masses that are usually connected to malignancy in breast tissue. On the other hand, malignant cases are shown in Images 3 and 4, which are distinguished by higher tissue density borders or irregular shapes that are frequently seen in breast cancer pathology. These images are used as input samples for the classifier's training and evaluation stages, which helps the model accurately differentiate between benign and malignant tumor patterns.

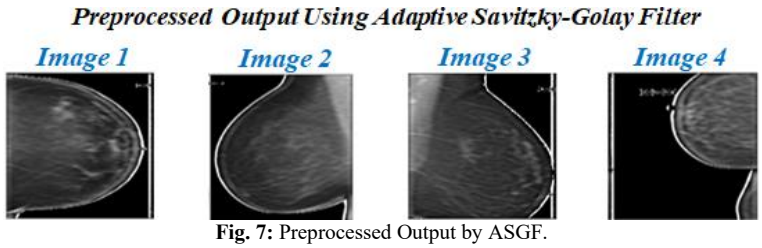


Fig. 7: Preprocessed Output by ASGF.

After applying ASGF, which improves image quality and lowers noise while maintaining crucial structural characteristics of breast tissue, the preprocessed output is shown in Fig. 7. By fitting local polynomials to intensity data, the ASGF smoothes the image and makes small patterns that are essential for tumor detection more visible. These preprocessed outputs have better contrast and distinct tissue boundaries than the raw images, which helps the TwinS^2-SAT model perform better in feature extraction and classification. Fig. 8 displays the pixel intensity histograms following the execution of ASGF, highlighting the effect of filtering on tissue contrast enhancement. Because of their homogenous tissue composition, images 1 and 2 continue to show closely grouped intensity values with some smoothing effects. On the other hand, the peaks in images 3 and 4 are wider and more noticeable, indicating dense tumor areas have better contrast.

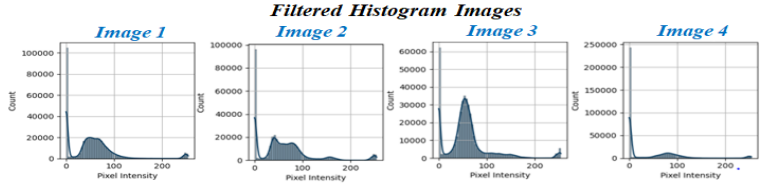


Fig. 8: Filtered Histogram Images.

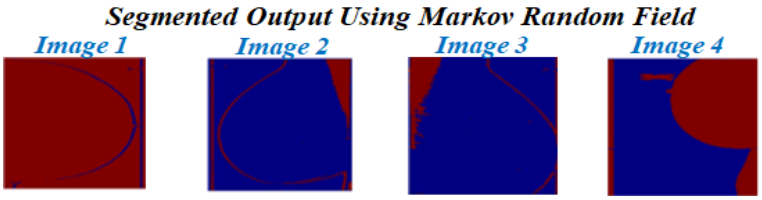


Fig. 9: Segmented Output Images Using MRF.

Following the application of the MRF approach, the segmented outputs of breast tumor mammography images are displayed in Fig.9. To facilitate accurate categorization, the segmentation draws attention to tumorous areas and those that are not. Benign instances are depicted in images 1 and 2, which show mostly homogeneous tissue regions with few spots of tumor emphasized. On the other hand, malignant cases are shown by images 3 and 4, where the marked boundaries show more irregular and obvious tumor regions.

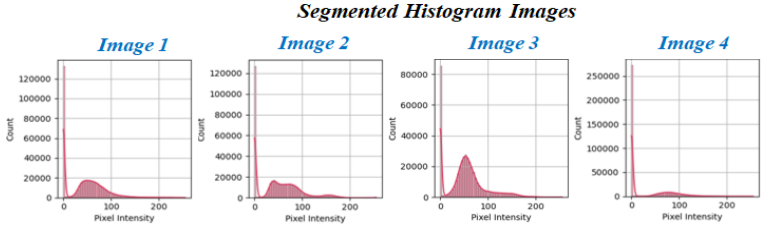


Fig. 10: Segmented Histogram Images.

The black areas in histograms of segmented breast cancer images are probably dense or aberrant tissues, as seen by the strong peaks at low pixel intensities in Fig. 10. Similar patterns with moderate intensity variation are seen in Images 1 and 2, with a wider spread around intensity 40. Image 3 implies extensive segmentation, and the darkest areas in Image 4 may indicate extremely suspicious tissue. These patterns aid in distinguishing between areas that are benign and those that are cancerous.

Table 3: Features Extraction Using GLCM

Features	Image 1	Image 2	Image 3	Image 4
Contrast	4.2100	5.7485	5.6588	5.0535
Correlation	0.8771	0.9319	0.9383	0.9803
Energy	0.9630	0.9109	0.9034	0.7108
Homogeneity	0.9921	0.9892	0.9893	009905
Dissimilarity	0.1830	0.2499	002460	0.2197

GLCM-based texture elements that were taken from the mammography images are displayed in Table 3. Benign images 1 and 2 show smoother, more consistent textures with increased energy and lower contrast. Images of malignancy (3 and 4) show irregular and diverse tumor patterns with increased contrast, higher correlation, and lower energy. The TwinS2-SAT classifier greatly benefits from these feature variations when it comes to differentiating between benign and malignant breast tissues.

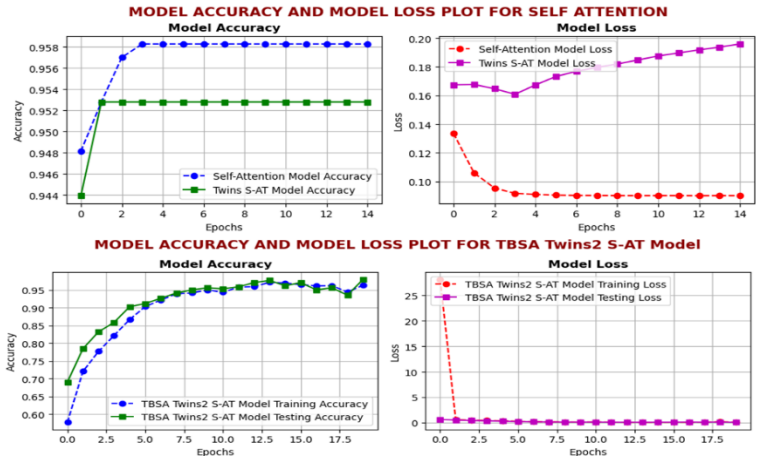


Fig. 11: Comparison of Model Accuracy vs. Loss.

The proposed TBSA-TwinS2-SAT model and the SAT model's training and testing results are contrasted in Fig.11. According to the accuracy plot, the SAT model varies at 94.99%, whereas the proposed TBSA-TwinS^2-SAT model continuously improves across epochs, reaching around 97.94% accuracy. For TBSA- TwinS^2-SAT, the loss plot shows a sharp initial decline that stabilizes at a significantly lower value than the SAT model. These findings clearly show that the TwinS^2-SAT model's classification performance and convergence speed on breast cancer image data are greatly improved by TBSA.

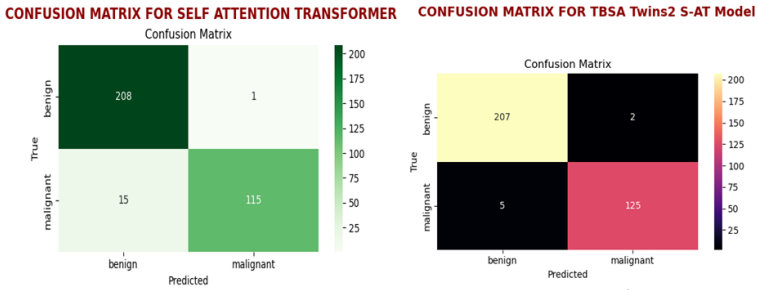


Fig. 12: Confusion Matrix of SAT Model and Proposed TBSA- TwinS^2-SAT Model.

With only one cancer instance correctly classified and 15 cases incorrectly classified as benign, the SAT exhibits poor malignant identification. A significant increase, however, is shown by the TBSA- TwinS^2-SAT model, which properly detects 125 malignant instances with just 5 misclassified as benign while retaining a high accuracy for benign cases. These findings in Fig. 12 demonstrate that the TwinS^2-SAT model's sensitivity and overall classification accuracy for breast cancer diagnosis are greatly increased by TBSA.

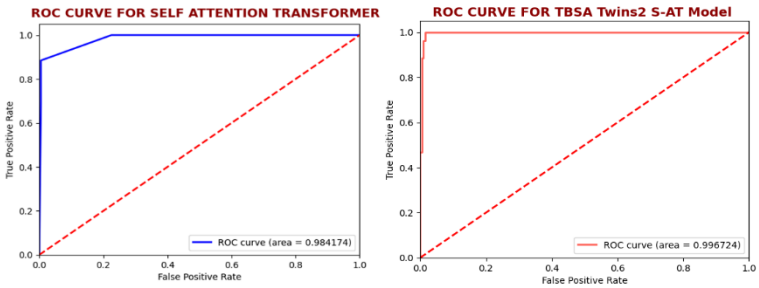


Fig. 13: Assessment of ROC.

The ROC curves for the SAT model and the TBSA- TwinS^2-SAT model proposed are shown in Fig. 13. With an Area Under Curve (AUC) of 0.98 for both benign and malignant classes, the SAT model performs poorly in discrimination, indicating almost random classification. In contrast, the TBSA- TwinS^2-SAT model achieves an AUC of 0.99 for both benign and malignant cases, demonstrating exceptional classification capacity. This notable enhancement shows that the integration of TBSA for fine-tuning of the TwinS^2-SAT framework allows for extremely precise and consistent separation of benign and malignant breast tumor images.

Table 4: Classification Report for Proposed Model vs. SAT

Classification	Precision	Recall	F1-Score	Support
Classification Report for SAT Model				
Benign	0.93	1.00	0.96	209
Malignant	0.99	0.88	0.93	130
Macro Avg	0.96	0.94	0.95	339
Weighted Avg	0.96	0.95	0.95	339
Classification Report for TBSA-TwinS2-SAT Model				
Benign	0.98	0.99	0.98	209
Malignant	0.98	0.96	0.97	130
Macro Avg	0.98	0.98	0.98	339
Weighted Avg	0.98	0.98	0.98	339

The proposed TBSA-TwinS²-SAT model and the current SAT model's classification performance are contrasted in Table 4. The proposed approach shows significant improvement in all measures. The recall increases from 0.88 to 0.96 for the malignant class, indicating improved cancerous case detection. Confirming the improved accuracy, balance, and dependability of the TBSA- TwinS²-SAT model for breast cancer classification, both macro and weighted averages in F1-score, precision, and recall rise from 0.95 (SAT) to 0.98 (proposed model).

Table 5: Comparison of Accuracy

Author	Year	Method	Dataset	Accuracy
Khan et.al [36]	2019	GoogleNet	LRH Hospital	97.52%
Moon et.al [37]	2020	Ensemble Model	BUSI	91.10%
Zuluaga-Gomez et.al [38]	2020	ResNet	DMR-IR	94%
Togacar et.al [39]	2020	CNN	BreakHis	95.03%
Hekal et.al [40]	2021	AlexNet and ResNet-50	CBIS-DDSM ROI	93.2%
Maqsood et.al [41]	2022	TTCNN	DDSM	97.49%
Tavakoli et.al [42]	2023	Block-Based CNN	MIAS	94.68%
Proposed	-	TBSA- TwinS^2-SAT	Breast Cancer	97.94%

A comparison of classification accuracy across different datasets and breast cancer detection techniques is shown in Table 5. With an accuracy of 97.94%, the proposed TBSA- TwinS^2-SAT model outperforms earlier methods like GoogleNet with 97.52%, TTCNN with 97.49%, and conventional CNN and ResNet-based approaches. In comparison to current state-of-the-art techniques, this highlights the proposed model's exceptional performance and generalization ability in breast cancer detection tasks, establishing it as a highly competitive and positive solution.

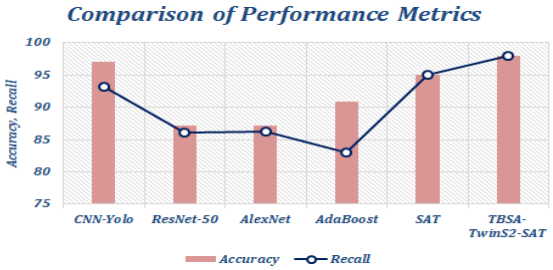


Fig. 14: Comparison of Accuracy and Recall.

The comparative analysis of different models based on their accuracy and recall in breast cancer classification is demonstrated in Fig. 14. The proposed TBSA-TwinS²-SAT model achieves the highest accuracy of 97.94% and recall of 98%, outperforming other state-of-the-art techniques such as CNN-Yolo [43] with 97% accuracy, 93.2% recall, and SAT model with 94.99% accuracy, 95% recall. Traditional models like ResNet-50 [44], AlexNet [45], and AdaBoost [46] show relatively lower performance, emphasizing the superior sensitivity and reliability of the proposed model in correctly identifying cancerous cases.

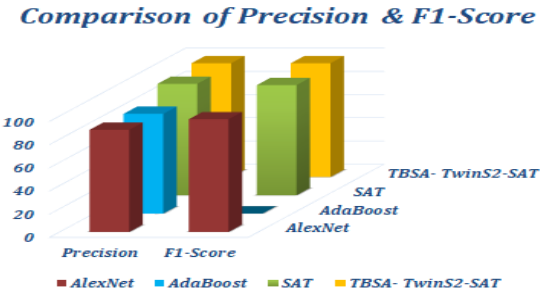


Fig. 15: Comparative Study of Precision & F1-Score.

The precision and F1-score of various models used for the classification of breast cancer are demonstrated in Fig. 15. The proposed TBSA-TwinS²-SAT model achieves the highest values in both metrics, with 98% precision and 98% F1-score, representing strong performance in both identifying positive cases and minimizing false predictions. In comparison, SAT performs well with 96% precision, 95% F1-score, while AlexNet [45] shows a relatively high F1-score of 97.1% but somewhat lower precision 88%. AdaBoost [46] lags expressively with 86% precision and F1-score, indicating potential issues in balanced classification. This analysis reveals that TBSA-TwinS²-SAT demonstrates superior balance between precision and recall.

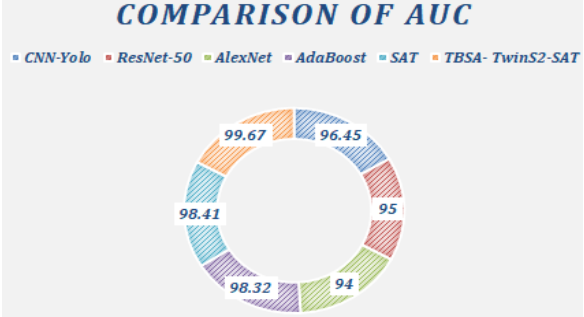


Fig. 16: Comparison of AUC.

The TBSA-TwinS²-SAT model has the greatest AUC of 99.67, according to AUC comparison, demonstrating remarkable discrimination abilities between malignant and non-cancerous cases shown in Fig. 16. It performs better than AdaBoost [46] with 98.32 and SAT 98.41, both of which are reliable but not as good. Models with lower AUCs, such as CNN-Yolo [43], ResNet-50 [44], and AlexNet [45], indicate a lower level of classification confidence. This demonstrates that the most strong and dependable performance in breast cancer detection is shown by TBSA-TwinS²-SAT.

Table 6: Comparison of Training Time

Model	Training time (s)
AdaBoost [46]	60
CNN-Yolo [43]	720
ResNet-50 [44]	540
AlexNet [45]	300
Proposed	180

Table 6 presents a comparative analysis of training times for various models used in breast cancer image classification, based on simulations conducted using a GPU environment. The proposed TwinS²-SAT model demonstrates the fastest training time among deep learning models, completing in just 180 seconds, owing to its efficient spatially separable self-attention mechanism and hyperparameter optimization. In contrast, AlexNet and ResNet-50 require 300 seconds and 540 seconds, respectively, reflecting the increasing complexity and depth of their architectures. CNN-YOLO, known for its high-resolution object detection capabilities, exhibits the longest training duration at 720 seconds, due to its dense layer structure and region proposal computations. On the other hand, AdaBoost, a non-deep learning method, trains significantly faster in only 60 seconds, but lacks the deep feature learning capabilities required for complex medical image analysis.

5. Conclusion

This research proposes an advanced breast cancer detection framework using the TwinS2-SAT classifier with TBSA for fine-tuning. Mammogram images after preprocessing with ASGF to improve contrast and lower noise, the MRF technique is followed for segmenting important areas, and the GLCM approach is used to extract features. The TwinS²-SAT model's hyperparameters were successfully optimized using the TBSA technique, which greatly enhanced classification performance. The spatially separable self-attention mechanism further contributes to reducing the model's computational complexity by decomposing global attention into localized and sub-sampled components, allowing efficient multi-scale feature learning in medical images. The proposed TBSA- TwinS²-SAT model reached a peak accuracy of 97.94%, significantly surpassing the traditional SAT model, which only achieved 94.99% accuracy, according to experimental validation using the Breast Cancer Detection dataset. Furthermore, for both benign and malignant instances, the proposed model achieved a good AUC of 0.99. These findings confirm the proposed method's better diagnostic capacity, sensitivity, and robustness in differentiating between benign and malignant breast cancers. This makes the proposed model potentially suitable for automated breast cancer diagnosis in healthcare environments. Practically, the model's lightweight yet accurate architecture makes it well-suited for integration into clinical workflows, particularly in settings where real-time decision support is critical, such as radiology departments or telemedicine platforms. It could assist radiologists by highlighting suspicious regions or triaging cases for further review, thereby improving diagnostic throughput and consistency. The computational complexity of this work is carefully balanced to achieve high diagnostic accuracy while maintaining efficiency, making it suitable for clinical integration. The adopted SSSA mechanism significantly reduces the computational burden typically associated with standard self-attention operations. While the current design integrates spatially separable self-attention to reduce computational load, future efforts could explore model pruning and quantization to compress the architecture without compromising diagnostic accuracy. Additionally, implementing dynamic attention mechanisms that adaptively allocate resources based on image complexity could further minimize inference cost.

References

[1] Hirra I, Ahmad M, Hussain A, Usman Ashraf M, Saeed IA, Qadri SF, Alghamdi AM & Alfakeeh AS (2021), Breast cancer classification from histopathological images using patch-based deep learning modeling, *IEEE Access,* 9, 24273-24287, https://doi.org/10.1109/ACCESS.2021.3056516.

[2] Saber A, Sakr M, Abo-Seida OM, Keshk A & Chen H (2021), A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. *IEEE Access,* 9, 71194-71209, https://doi.org/10.1109/ACCESS.2021.3079204.

[3] Liu M, Hu L, Tang Y, Wang C, He Y, Zeng C, Lin K, He Z & Huo W (2022), A deep learning method for breast cancer classification in the pathology images, *IEEE Journal of Biomedical and Health Informatics,* 26, 10, 5025-5032, https://doi.org/10.1109/JBHI.2022.3187765

[4] Naseem U, Rashid J, Ali L, Kim J, Haq QEU, Awan MJ & Imran M (2022), An automatic detection of breast cancer diagnosis and prognosis based on machine learning using ensemble of classifiers, *IEEE Access,* 10, 78242-78252, https://doi.org/10.1109/ACCESS.2022.3174599.

[5] Rehman MU, Shafique A, Ghadi YY, Boulila W, Jan SU, Gadekallu TR, Driss M & Ahmad J (2022), A novel chaos-based privacy-preserving deep learning model for cancer diagnosis, *IEEE Transactions on Network Science and Engineering,* 9, 6, 4322-4337, https://doi.org/10.1109/TNSE.2022.3199235.

[6] Fan M, Yuan C, Huang G, Xu M, Wang S, Gao X & Li L (2022), A framework for deep multitask learning with multiparametric magnetic resonance imaging for the joint prediction of histological characteristics in breast cancer, *IEEE Journal of Biomedical and Health Informatics*, 26, 8, 3884-3895, https://doi.org/10.1109/JBHI.2022.3179014.

[7] Gamal A, Sharafeldeen A, Alnaghy E, Alghandour R, Alghamdi NS, Ali KM & Shamaa S (2024), A novel machine learning approach for predicting neoadjuvant chemotherapy response in breast cancer: integration of multimodal radiomics with clinical and molecular subtype markers, *IEEE Access,* https://doi.org/10.1109/ACCESS.2024.3432459.

[8] Jabeen K, Khan MA, Alhaisoni M, Tariq U, Zhang YD, Hamza A, Mickus A & Damaševičius R (2022), Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion, *Sensors,* 22, 3, 807, https://doi.org/10.3390/s22030807.

[9] Jadoon EK, Khan FG, Shah S, Khan A & Elaffendi M (2023), Deep learning-based multi-modal ensemble classification approach for human breast cancer prognosis, *IEEE Access,* 11, 85760-85769, https://doi.org/10.1109/ACCESS.2023.3304242.

[10] Jebarani PE, Umadevi N, Dang H & Pomplun M (2021), A novel hybrid K-means and GMM machine learning model for breast cancer detection", *IEEE Access*, vol. 9, pp. 146153-146162, https://doi.org/10.1109/ACCESS.2021.3123425

[11] Sheng W, Xia S, Wang Y, Yan L, Ke S, Mellisa E, Gong F, Zheng Y & Tang T (2022), Invasive ductal breast cancer molecular subtype prediction by MRI radiomic and clinical features based on machine learning, *Frontiers in Oncology,* 12, 964605, https://doi.org/10.3389/fonc.2022.964605.

[12] Chikkala RB, Anuradha C, Murty PSC, Rajeswari S, Rajeswaran N, Murugappan M & Chowdhury ME (2025), Enhancing Breast Cancer Diagnosis with Bidirectional Recurrent Neural Networks: A Novel Approach for Histopathological Image Multi-Classification, *IEEE Access,* https://doi.org/10.1109/ACCESS.2025.3542989.

[13] Abdullah KA, Marziali S, Nanaa M, Sánchez LE, Payne NR & Gilbert FJ (2025), Deep learning-based breast cancer diagnosis in breast MRI: systematic review and meta-analysis, *European Radiology*, 1-16, https://doi.org/10.1007/s00330-025-11406-6.

[14] Paul Joshua K, Alex SA, Mageswari M & Jothilakshmi R (2024), Enhanced conditional self-attention generative adversarial network for detecting cotton plant disease in IoT-enabled crop management, *Wireless Networks*, 1-15, https://doi.org/10.1007/s11276-024-03762-w

[15] Sreelekshmi V, Pavithran K & Nair JJ (2024), SwinCNN: An Integrated Swin Trasformer and CNN for improved breast cancer grade classification, *IEEE Access,* https://doi.org/10.1109/ACCESS.2024.3397667.

[16] Rashid TA, Majidpour J, Thinakaran R, Batumalay M, Dewi DA, Hassan BA, Dadgar H & Arabi H (2024), NSGA-II-DL: metaheuristic optimal feature selection with deep learning framework for HER2 classification in breast cancer, *IEEE Access,* 12, 38885-38898, https://doi.org/10.1109/ACCESS.2024.3374890

[17] Zeng R, Qu B, Liu W, Li J, Li H, Bing P, Duan S & Zhu L (2024), FastLeakyResNet-CIR: A novel deep learning framework for breast cancer detection and classification, *IEEE Access*, https://doi.org/10.1109/ACCESS.2024.3401729

[18] Koshy SS & Jani Anbarasi L (2024), LMHistNet: Levenberg–Marquardt based deep neural network for classification of breast cancer histopathological images, *IEEE Access,* https://doi.org/10.1109/ACCESS.2024.3385011.

[19] Rahman M, Deb K, Dhar PK & Shimamura T (2024), ADBNet: An Attention-Guided Deep Broad Convolutional Neural Network for the Classification of Breast Cancer Histopathology Images, *IEEE Access*, https://doi.org/10.1109/ACCESS.2024.3419004.

[20] Rastogi M, Vijarania M, Goel N, Agrawal A, Biamba CN & Iwendi C (2024), Conv1D-LSTM: Autonomous Breast Cancer Detection Using a One-Dimensional Convolutional Neural Network with Long Short-Term Memory, *IEEE Access*, https://doi.org/10.1109/ACCESS.2024.3514662.

[21] Hayat M, Ahmad N, Nasir A & Tariq ZA (2024), Hybrid Deep Learning EfficientNetV2 and Vision Transformer (EffNetV2-ViT) Model for Breast Cancer Histopathological Image Classification, *IEEE Access,* https://doi.org/10.1109/ACCESS.2024.3503413

[22] Zeynali A, Tinati MA & Tazehkand BM (2024), Hybrid CNN-Transformer Architecture with Xception-Based Feature Enhancement for Accurate Breast Cancer Classification, *IEEE Access,* https://doi.org/10.1109/ACCESS.2024.3516535.

[23] Senapati K & Pandey AK (2024), A novel decision level class-wise ensemble method in deep learning for automatic multi-class classification of her2 breast cancer hematoxylin-eosin images, *IEEE Access*, 12, 46093-46103, https://doi.org/10.1109/ACCESS.2024.3382212

[24] Hekal AA, Elnakib A, Moustafa HE, & Amer HM (2024), Breast cancer segmentation from ultrasound images using deep dual-decoder technology with attention network, *IEEE Access*, 12, 10087-10101, https://doi.org/10.1109/ACCESS.2024.3351564.

[25] Vanitha K, Manimaran A, Chokkanathan K, Anitha K, Mahesh TR, Vinoth Kumar V & Vivekananda GN (2024), Attention-based Feature Fusion with External Attention Transformers for Breast Cancer Histopathology Analysis, *IEEE Access*, https://doi.org/10.1109/ACCESS.2024.3443126

[26] Mo Y, Han C, Liu Y, Liu M, Shi Z, Lin J, Zhao B, Huang C, Qiu B, Cui Y & Wu L (2023), Hover-trans: Anatomy-aware hover-transformer for roi-free breast cancer diagnosis in ultrasound images, *IEEE Transactions on Medical Imaging,* 42, 6, 1696-1706, https://doi.org/10.1109/TMI.2023.3236011.

[27] Monjezi E, Akbarizadeh G & Ansari-Asl K (2024), RI-ViT: A Multi-Scale Hybrid Method Based on Vision Transformer for Breast Cancer Detection in Histopathological Images. *IEEE Access*, 12, 186074 – 186086, https://doi.org/10.1109/ACCESS.2024.3514322

[28] Maurya R, Pandey NN & Mahapatra S (2025), BMEA-ViT: Breast Cancer Classification using Lightweight Customised Vision Transformer Architecture with Multi-Head External Attention. *IEEE Access*, 13, 44317 – 44329, https://doi.org/10.1109/ACCESS.2025.3547862

[29] Sabry M, Balaha HM, Ali KM, Soliman TH, Gondim D, Ghazal M, Alghamdi NS & El-Baz A (2025), Enhancing Breast Cancer Diagnosis with Multi-Resolution Vision Transformers and Robust Decision-Making. *IEEE Access*, 13, 89704 – 89722, https://doi.org/10.1109/ACCESS.2025.3570840.

[30] Al-Hejri AM, Al-Tam RM, Fazea M, Sable AH, Lee S & Al-Antari MA (2022), ETECADx: Ensemble self-attention transformer encoder for breast cancer diagnosis using full-field digital X-ray breast images. *Diagnostics*, 13, 1, 89, https://doi.org/10.3390/diagnostics13010089.

[31] Ahmed MR, Rahman H, Limon ZH, Siddiqui MI, Khan MA, Pranta AS, Haque R, Swapno SM, Cho YI & Abdallah MS (2025), Hierarchical Swin Transformer Ensemble with Explainable AI for Robust and Decentralized Breast Cancer Diagnosis. *Bioengineering*, 12, 6, 651, https://doi.org/10.3390/bioengineering12060651.

[32] Hamim M, El Moudden I, Pant MD, Moutachaouik H, & Hain M (2021), A hybrid gene selection strategy based on fisher and ant colony optimization algorithm for breast cancer classification, 148-163. https://doi.org/10.3991/ijoe.v17i02.19889.

[33] Bilal A, Imran A, Baig TI, Liu X, Abouel Nasr E & Long H (2024), Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization. *Scientific Reports,* 14, 1, 0714, https://doi.org/10.1038/s41598-024-61322-w

[34] Thawkar S (2021), A hybrid model using teaching–learning-based optimization and Salp swarm algorithm for feature selection and classification in digital mammography. *Journal of ambient intelligence and humanized computing,* 12, 8793-8808, https://doi.org/10.1007/s12652-020-02662-z

[35] Aguerchi K, Jabrane Y, Habba M & El Hassani AH (2024), A CNN hyperparameters optimization based on particle swarm optimization for mammography breast cancer classification. *Journal of Imaging,* 10, 2, 30, https://doi.org/10.3390/jimaging10020030

[36] Khan S, Islam N, Jan Z, Din IU & Rodrigues JJC (2019), A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, *Pattern Recognition Letters,* 125 1-6, https://doi.org/10.1016/j.patrec.2019.03.022

[37] Moon WK, Lee YW, Ke HH, Lee SH, Huang CS & Chang RF (2020), Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks, *Computer methods and programs in biomedicine*, 190, 105361, https://doi.org/10.1016/j.cmpb.2020.105361.

[38] Zuluaga-Gomez J, Masry ZA, Benaggoune K, Meraghni S & Zerhouni N (2021), A CNN-based methodology for breast cancer diagnosis using thermal images, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization,* 9, 2, 131-145, https://doi.org/10.1080/21681163.2020.1824685.

[39] Toğaçar M, Özkurt KB, Ergen B & Cömert Z (2020), BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer, *Physica A: Statistical Mechanics and its Applications,* 545, 123592, https://doi.org/10.1016/j.physa.2019.123592.

[40] Hekal AA, Elnakib A & Moustafa HE (2021), Automated early breast cancer detection and classification system, *Signal, Image and Video Processing,* 15, 7, 1497-1505, https://doi.org/10.1007/s11760-021-01882-w

[41] Maqsood S, Damaševičius R & Maskeliūnas R (2022), TTCNN: A breast cancer detection and classification towards computer-aided diagnosis using digital mammography in early stages, *Applied Sciences,* 12, 7, 3273, https://doi.org/10.3390/app12073273.

[42] Tavakoli N, Karimi M, Norouzi A, Karimi N, Samavi S & Soroushmehr SR (2023), Detection of abnormalities in mammograms using deep features, *Journal of Ambient Intelligence and Humanized Computing,* 1-13, https://doi.org/10.1007/s12652-019-01639-x

[43] Al-Masni MA, Al-Antari MA, Park JM, Gi G, Kim TY, Rivera P, Valarezo E, Choi MT, Han SM & Kim TS (2018), Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system, *Computer methods and programs in biomedicine*, 157, 85-94, https://doi.org/10.1016/j.cmpb.2018.01.017

[44] Ragab DA, Attallah O, Sharkas M, Ren J & Marshall S (2021), A framework for breast cancer classification using multi-DCNNs, *Computers in biology and medicine,* 131, 104245, https://doi.org/10.1016/j.compbiomed.2021.104245

[45] Ragab DA, Sharkas M, Marshall S & Ren J (2019), Breast cancer detection using deep convolutional neural networks and support vector machines, *PeerJ*, 7, e6201, https://doi.org/10.7717/peerj.6201/fig-1.

[46] Zhang H, Wu R., Yuan T, Jiang Z, Huang S, Wu J, Hua J, Niu Z & Ji D (2020), DE-Ada*: A novel model for breast mass classification using cross-modal pathological semantic mining and organic integration of multi-feature fusions, *Information Sciences*, 539, 461-486, https://doi.org/10.1016/j.ins.2020.05.080.