RESEARCH PAPER

Optimized Deep Learning Model for Early Detection of Breast Cancer in Mammograms Using Vision Transformer

Daphne Sherine. H[1*], G. Revathy[2]

[1,2] Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies, Tamil Nadu, India

[*] **Corresponding Author:** Daphne Sherine. H, Email: salasherine@gmail.com

ABSTRACT

Breast cancer (BC) is a major global health issue and one of the leading causes of cancer-related mortality for women. A timely and precise diagnosis using mammography analysis is essential for increasing survival rates. With the use of Vision Transformers' (ViTs) potent representational capabilities, this research presents an improved deep learning (DL) framework for the early identification and classification of BC in mammograms. The proposed model, Artificial Jellyfish Search-Driven Attention-Enriched Convolutional tuned Pyramid Vision Transformer (AJFS-Att-CPVT), is designed to detect and classify BC by effectively capturing both global and local image features to distinguish between benign and malignant lesions. Mammographic data are collected from multiple publicly available datasets to ensure diversity in breast tissue characteristics and lesion types. The preprocessing pipeline included image normalization and Gaussian blurring to standardize image quality and improve feature visibility. To strengthen model generalization and mitigate overfitting, data augmentation strategies such as random rotation, horizontal flipping, and contrast variation were employed during training. For baseline feature extraction, ResNet50 was applied to the preprocessed images, providing a benchmark for evaluating the performance gains achieved by the proposed AJFS-Att-CPVT model. Model evaluation using Python with key performance metrics, such as a recall of 94.8%, an F1-score of 93.8%, an accuracy of 95.8%, a specificity of 94.20%, an AUC of 93.35%, and a precision of 95.2%, demonstrated higher diagnostic capabilities compared to traditional convolutional approaches. The findings demonstrate how Vision Transformer-based designs have the potential to greatly improve the accuracy and dependability of early BC diagnosis in mammography imaging when refined with bio-inspired algorithms and sophisticated preprocessing.

Keywords: Mammograms, Breast cancer detection, Vision Transformers, Data augmentation, Artificial Jellyfish Search-Driven Attention-Enriched Convolutional tuned Pyramid Vision Transformer (AJFS-Att-CPVT).

How to cite this article: Sherine H D, Revathy G. Optimized Deep Learning Model for Early Detection of Breast Cancer in Mammograms Using Vision Transformer. Int J Drug Deliv Technol. 2026;16(23s): 08-23. DOI: 10.25258/ijddt.16.23s.2

Source of support: Nil.

Conflict of interest: None

1. Introduction

Cancer is defined as uncontrolled cell growth occurring in certain areas of the body [1]. One of the most prevalent illnesses affecting women is BC, which is regarded as a deadly form of the disease worldwide. Additionally, BC is responsible for 9.6% of global deaths. Since BC worsens and has major impacts, it is best to get a diagnosis as soon as possible to make treatment more challenging. Mammography images and their processing are used to diagnose BC, where they play an important role in addition to being one of the techniques used to diagnose disease and treat patients promptly [2]. Compared to alternative imaging modalities, including nuclear medicine, magnetic resonance imaging, and ultrasound, mammograms are more convenient, less expensive, and have a higher spatial resolution. Because of these qualities, mammography is the preferred diagnostic technique for early BC detection. Localization is necessary to ascertain the precise location of the mass in digital mammography; region recognition and classification alone are insufficient in a BC detection system [3]. The most important aspect of treating BC is early detection. Because the initial tumor frequently produces no noticeable symptoms, screening a sizable nonsymptomatic population should be the foundation for early cancer detection. A limited fraction of the population with a higher risk of cancer is frequently the target of these procedures. Instead of looking for morphological abnormalities in the breast tissue, thermography looks for physiological changes while screening for BC. Thermal imaging is especially used to identify changes in blood flow and temperature that indicate the growth of tumors [4].

However, with the introduction of molecular diagnostics, digital imaging, and artificial intelligence (AI), there has been a shift toward more precise and individualized detection methods. Clinicians can detect even the slightest abnormalities using technologies like molecular imaging, 3D mammography, and advanced genetic screening techniques, resulting in earlier intervention and improved patient results [5]. BC can affect the lungs and other bodily regions, as well as spread to lymph nodes. Ductal dysfunction is typically the first indication of BC. However, lobules, breast tissue, and glandular tissues can all be developed. It found that the risk of BC is increased by changes in lifestyle, environment, and hormones. Using a low-dose X-ray, the interior structure of the breast is shown. Medical professionals refer to this process as mammography [6]. A radiographer's manual detection of screening mammography is costly, time-consuming, and has a high false-positive rate. Nevertheless, detection is challenging because of tissue variety and inexperience. To get beyond these obstacles, automated mammography BC detection techniques must be created with specialized computer systems that can help radiologists provide patients with early treatment [7]. Observing a new paradigm in computer vision was introduced by Transformers (ViTs). ViTs employ self-attention techniques to identify both global and local dependencies in images, drawing inspiration from Transformers' successes in natural language processing (NLP). This gives them the ability to simulate long-range interactions between image patches, which leads to a more thorough understanding of the image overall. Because it enables the model to identify subtle spatial correlations that can signal early-stage cancer even in dense breast tissue, this capability is particularly helpful in mammogram analysis [8, 9].

Computer-aided diagnostic (CAD) systems are commonly used by radiologists to treat a large number of patients or to get a second opinion on a diagnosed cancer from an AI-based machine. These new software systems do show promise as diagnostic tools, offering many advantages like segmentation and lesion detection, even for small breast lesions [10]. The BC mortality shares from each human development index (HDI) group are paralleled by increased regional incidence rates, which appear to have a lower share than expected, possibly due to the larger populations in low-, middle-, and middle-HDI countries. Countries in transition experience high mortality rates due to a lack of identification and prompt treatment improvements. HIAs' advanced imaging techniques increase tumor detection, delineation, and stratification. Combining these technologies with traditional methods has dramatically improved early cancer detection and treatment precision, increasing outcomes for patients [11]. Among women globally, BC is presently the second most common type of cancer; new screening tools are urgently needed. The synergistic use of AI techniques such as DL in medical imaging has significantly altered the diagnostic landscape, giving sophisticated solutions that develop the efficiency and quality of BC management. The use of ultrasonic imaging in BC screening received popularity due to its ability to distinguish between solid tumors and cysts and its noninvasive nature [12].

These days, the most used imaging technique for detecting microcalcifications that could be indicators of cancerous changes is mammography. However, the large variations in radiologists' interpretation accuracy make mammography interpretation challenging. While false positive findings might result in unnecessary intrusive treatments, increased costs, and increased patient concern, false negative results can cause a delayed diagnosis, the disease to worsen, and a decreased chance of survival. Clinical professionals also have an added workload because of the massive volume of mammography exams that are performed every day [13]. Two problems, meanwhile, have not been adequately addressed: the massive amount of time radiologists spend examining women who are generally healthy, as well as the similarly considerable percentage of women who, while consistently participating in screening, do not have cancer during screening. Many AI algorithms have been created for software that detects cancer in mammograms. In terms of mammogram analysis, certain software algorithms are performing on par with radiologists, despite the fact that they have not been validated in a fully representative screening sample [14].

1.1 Research Objective:

The research create a better vision transformer-based system; adopts Jellyfish Search-Driven Attention-Enriched Convolutional tuned Pyramid Vision Transformer (AJFS-Att-CPVT) to detect breast cancer and dependably at the earliest stage in mammograms. It incorporates both global and local image characteristics by attention-enriched convolution and AJFS to finely adjust the model parameters. Although using conventional preprocessing and augmentation, the novelty of the framework is that it is based on optimization to create a hybrid design, which enhances the consistency of the diagnosis, convergence and generalization to a variety of mammogram datasets.

1.2 Research Contribution:

The primary value of this research is the creation of an optimization-based hybrid system AJFS-Att-CPVT combining Pyramid Vision Transformers, attention-enriched convolutional tuning, and AJSO to adjust the parameters. In contrast to the existing CNN-ViT hybrids, the model proposes a dynamic dynamically-tuned attention mechanism informed by AJSO, improving lesion localization and converging stability. This combination of standard preprocessing and augmentation in the loop of optimization is strategically designed to enhance the feature visibility and generalization across multiple mammogram datasets, and the overall diagnostic accuracy and reliability are obtained, not just improved performance.

1.3 Research Organization

Section 1 gives an overview of the burden of BC in the world, the importance of early detection by mammography, and proposes ViT and bio-inspired optimization algorithms instead of traditional CNN. Section 2 is a related research of existing on deep learning (DL) in medical imaging, such as CNN-based models such as ResNet50, ViTs, and the issues of generalization and lesion detection in mammograms. Section 3 will describe the research methodology where it presents dataset collection, pre-processing, normalization, contrast enhancement, Gaussian blurring, and data augmentation taking the ResNet50 architecture as the foundation. Section 4 gives the results of the proposed model, which were more successful than those of the baseline ResNet50, despite limitations on the diversity of the datasets and computational costs. Section 5 ends by providing the significance of AJFS-Att-CPVT in early BC detection and areas of future research to integrate multi-modes and deploy to clinical practice.

2. Literature Review

The model proposed a system for early BC detection using mammography images. The proposed methodology plans to use DL to build accuracy of detection [15]. The CNN model is being tested for planning the training and testing data sets. The proposed system was limited based on the input image quality from mammogram images and internet connection, this is a concern that may be assumed to be lacking in rural situations. The method presented a novel framework to enhance BC identification using the Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM) dataset based on a combination of ViT and Graph Neural Networks (GNN) [16]. With an accuracy of 84.2%, the approach performed better than others because the method used the global imaging capabilities of ViT in a way that models structural relationships like with GNN. The model also provided interpretable attention heatmaps showing how the model arrived at a conclusion, which provided the radiologist with useful feedback in clinical situations. Adoption in clinical settings with limited resources may prove difficult due to the comparatively high computing cost of the proposed framework.

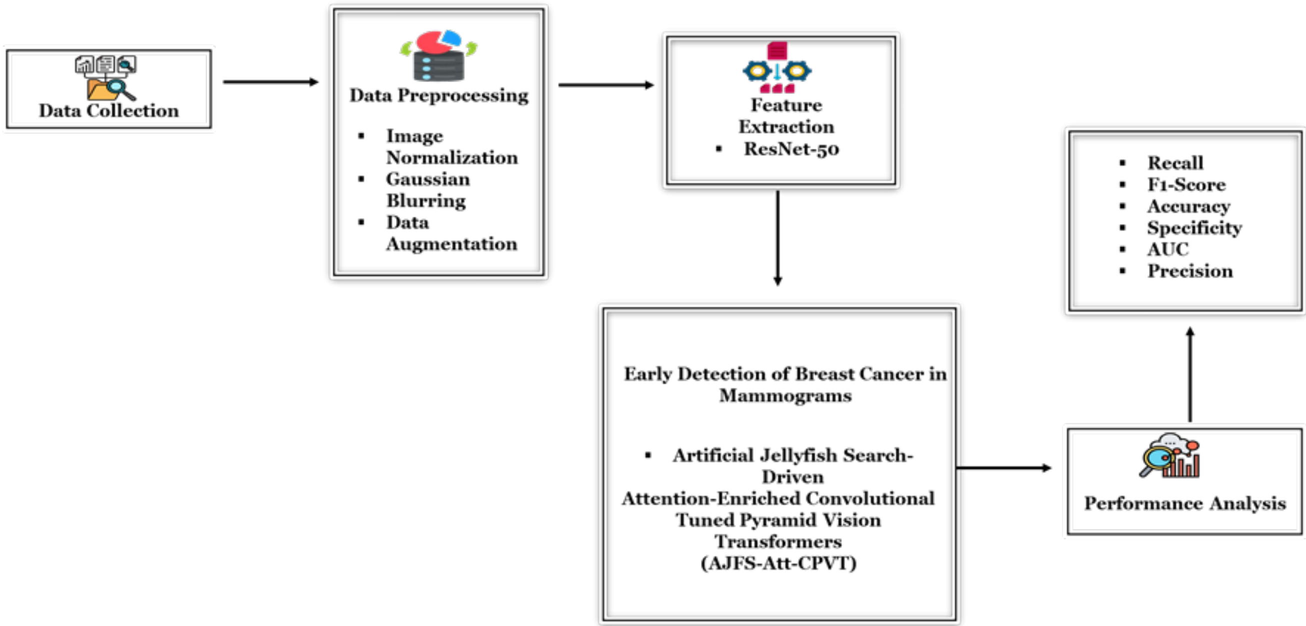
The research analyzed the use of AI-based machine learning methods for diagnosing BC through images [17]; it examined the capabilities of traditional model to evaluate malignant and benign tumors. The models were preprocessed, trained, and validated using publicly available datasets (i.e., the Wisconsin BC Dataset); the datasets included the Annson Cancer Dataset with 63 samples & Mammographic Image Analysis Society (MIAS) database. The research indicated that CNNs surpassed oral classification techniques in BC diagnoses, with excellent precision and recall. A significant concern of bias in datasets can affect the model's generalizability and dependability with different patient populations. Even while the recently proposed strategies for early BC screening have been extremely successful in reaching this objective, they depend on one of the mammography's signs to determine the patient's state [18]. When determining whether breast density is a risk factor or recognizing variations in the shapes and patterns of the results, the inherent, important correlation information among data from associated domains was not completely utilized by modern CAD systems that use single-label categorization. The main disadvantage of relying on several benchmark datasets was that it decreases performance consistency when used with real-world clinical data that hasn't been previously reported.

An end-to-end training approach has been used to create a DL system that can detect BC in mammogram images [19]. The system extracted texture features using an energy layer, performed classification using a TTCNN (transferable texture convolutional neural network), and refined edge detail using a modified contrast enhancement algorithm. Real-time clinical deployment can be hampered by the model's complexity and dependence on several deep feature extraction and fusion stages. The research proposed an advanced hybrid strategy for improved accuracy and the efficiency of BC diagnosis systems [20]. Three DL were used as feature extractors and Term Variance (TV) was used as an additional step to identify relevant features from each CNN model. The model was able to achieve average classification accuracy (CA) of 97.81% on70% of the training data, 98% on 80%, and 90% during training. However, the keylimit of the model was thatit depends on the MIAS dataset, thus limiting its applicability to different real-world mammographic datasets.

The research discussed the application of DL-based methodology to mammography analysis, and the objectives were to enhance the performance of the analysis. The study focused on the implementation of both CNNs and ViTs on a publicly available dataset. Another technique that was used to enhance performance was the use of synthetic data augmentation strategy. The outcomes of the research illustrated the importance of utilizing data preprocessing and augmentation techniques, particularly synthetic augmentation, in obtaining optimal classification performance. One major limitation was with synthetic image enhancement, which limits the ability to adequately represent the diversity of actual clinical mammography data [21]. The suggested method involved creating a DL model to identify BC in digital breast mammograms at different densities and contrasting the research with previous research [22]. 1501 patients who had digital mammograms from February 2007 to May 2015 had their craniocaudal and mediolateral images combined into 3002 views. One significant drawback was a decreased ability to identify breast tissue's high density, which makes it less precise for those specific individuals.

2.1 Research Gap:

The previous researchs on DL in breast cancer detection show significant advancement but have some common weaknesses. Numerous models are based on a small number of studies (e.g., MIAS, CBIS-DDSM) [16, 20], which makes them less applicable to diverse clinical groups. Synthetic data augmentation can be problematic at high computational cost cannot be deployed in low-resource or rural environments [16] and synthetic data augmentation cannot frequently model real mammographic diversity [21]. Moreover, current techniques are biased to datasets, overfit and not effective to thick breast tissues [22]. TTCNNs present complex architectures, and, which make real-time clinical implementation difficult [19]. All these together limit the real-world applicability, which encourages more adaptive, effective, and generalized diagnostic model



development. The proposed AJFS-Att-CPVT framework is capable of overcoming these problems by incorporating a wide range of data to improve generalization, compressing the computation time with AJSO-based optimization, and realistic augmentation to maintain clinical variation. The CNN-ViT nature of it is a hybrid that captures both the local and global features, enhancing its accuracy, strength, and generalization to diverse mammographic images and tissue densities.

Figure 1: Overview of the Proposed Model

3. Methodology:

- To maximize diversity in breast tissue and lesion representation, the AJFS-Att-CPVT framework development was based on mammogram images from public datasets.
- The preprocessing stage consists of image normalization, and Gaussian blurring.
- The datasets were also subjected to data augmentation techniques consisting of random rotation, horizontal flipping, and contrast variation to enhance model generalization.
- Image patches were tokenized and pre-processed by the multi-level attention layers.
- The AJSO manipulated attention parameters to maximize lesion localization and classification.
- Ultimately, the model's performance outperformed standard methods, as demonstrated in Figure 1, which shows the overall methodology for the research.

3.1 Data Collection:

- The breast tumors focused dataset consists of 3,383 mammogram images annotated in a folder structure.
- The data were derived from Roboflow, a machine vision project platform.
- The data is perfect for developing and evaluating DL models to detect breast malignancies using mammograms, as shown in Figure 2.
- The dataset is collected from the Kaggle score (https://www.kaggle.com/datasets/hayder17/breast-cancer-detection).

3.2 Data Preprocessing:

- Preprocessing is also essential to get mammogram images ready to be analyzed with improving the quality and consistency of the images.
- The images were normalized (bringed to the range between [0,1]) and blurred with Gaussian filters to decrease noise and emphasize lesions.
- Random rotation (±15 0), horizontal flipping, and variation contrast were augmented to data as data augmentation methods increased diversity and reduced overfitting.
- The steps resulted in normalized, noise-free, and diverse versions of the inputs, which facilitated the AJFS-Att-CPVT model to successfully model lesion characteristics and better the diagnostic generalization.

3.3 Image Normalization:

- Consistency in the input of mammograms plays a critical role in the ability of the model to identify small pathologic signs in an early case of BC with Vision Transformer-based architectures.

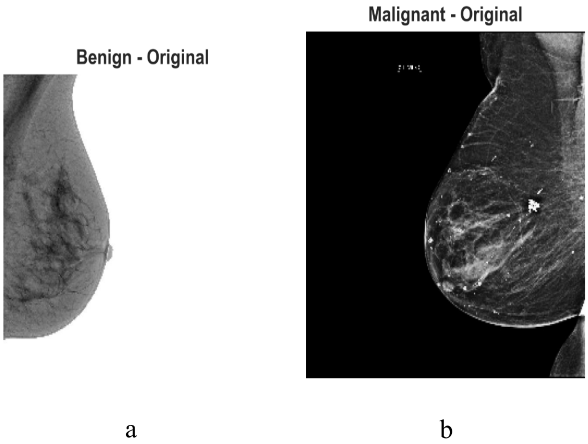


Figure 2: Sample Mammogram Images (a) Benign and (b) Malignant

- To maintain the diagnostically meaningful features, all of the images were down sampled to 224 224 pixels and the ViTs patch-tokenization process was retained.
- All values of the pixels were scaled to the range [0, 1], to decrease the variation related to various imaging devices and purchases, and breast tissue density, as shown in Figure 3.
- This step in the normalization process increases the speed of convergence of the training process, which offers model stability due to centering and scaling pixel intensities, enabling the network to focus on clinically significant features instead of irrelevant pixels.
- This does not only enhance consistency but also sensitivity to subtle mammographic manifestations of cancer especially with regards to the handling of early detection cases involving small micro calcifications, small displacements of internal structure, and ill-defined margins of a mass.
- In addition, the ViT was able to effectively fuse the interactions between local and global features due to consistent values of intensity that resulted in higher accuracy and reliability to classify the lesions when used in early screening of BC.

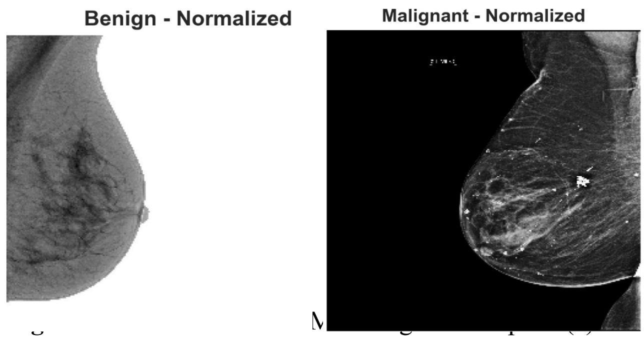
3.4 Gaussian Blurring for Noise Reduction:

- Gaussian blurring is an isotropic smoothing operator used to reduce noise and severe changes in mammographic images, improving the appearance of important characteristics for BC detection.
- The Gaussian kernel G is a circularly symmetric, bell-shaped function that smoothes the image while preserving essential structural information, such as lesions.
- The Gaussian kernel can be computed as in Equation (1).

$$H(Y,X) = \frac{1}{2\pi\sigma^2} f^{-(y^2+x^2)/2\sigma^2}, \quad (1)$$

- Where y and x are Cartesian distances and σ is the standard deviation that controls the level of smoothing.
- This kernel is used as a point spread function in a mammogram pre-processing step for BC detection, which reduces image noise through convolution and preserves the boundaries of lesions.
- Since mammograms consist of discrete pixels, a discrete approximation of the Gaussian kernel is computed before convolution.
- This pre-processing step ensures that image quality is uniform across the datasets and enhances the proposed model's capability to detect and classify benign and malignant BC, as illustrated in Figure 4.

Figure 3: Normalized Mammogram Samples: (a) Benign and (b) Malignant Cases

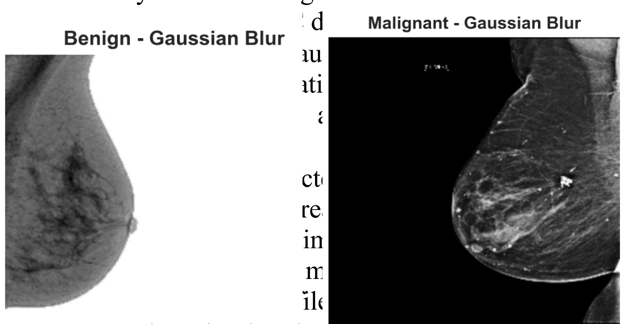


Benign and (b) Malignant Cases

(a) (b)

3.5 Data Augmentation

- In training, the preventive overfitting and model generalization were ensured by applying several clinically safe augmentation methods to



- It was selected only when the augmentation did not affect the clinical interpretability as subtle signs of microcalcifications, speculated masses and architectural distortions would have been lost.

(a) (b)

Random Rotation

- This allows for the variation in breast position occurring during mammography acquisition while allowing the model to maintain orientation invariance.
- The rotated pictures are beneficial, as they keep minor angular adjustments to the patient position from distorting the identification of the lesion.

Horizontal Flipping

- This enables the change in breast position in the process of getting the image during mammography acquisition whilst the model retains its orientation invariance.
- The advantage of the rotated images is that they prevent the occurrence of any minor angular changes in the position of the patient of distortion of the lesion identification.

Contrast Variation

- Adjusting image contrast simulates the change in images of mammogram that may arise because of exposure settings, breast density, and imaging equipment.
- This exposure change conditions the algorithm to be structured and textural based but not absolute intensity values, which strengthens the detection in a variety of datasets..

3.6 Feature Extraction using Residual Network with 50 layers (ResNet50):

- Deep CNNs ResNet-50, a 50-layer CNN, have been extensively applied in medical imaging problems such as BC detection in mammography.
- The architecture works on 224 x 224 images using 3 x 3 convolutional filters.
- Computational efficiency, doubling the number of filters with a halved feature map resolution is done.
- ResNet-50 is unique in its use of skip connections so feature maps of one layer can directly fruit in deeper layers.
- This technique solves the vanishing gradient problem, learns complex hierarchical representations, and trains faster.
- It is particularly beneficial in mammograms because BCs may occur in the form of minute changes in tissue density, thus requiring the low-level feature extraction (edges, textures) and high-level pattern detection (shapes, margins, and microcalcifications).
- Left over block groups containing convolutional and identity layers and batch normalization enhance stability and generalization.
- ResNet-50 has demonstrated a high level of performance in the disparity between benign and malignant tumors of the BC and this gives the good baseline information on the computer-aided diagnosis systems.
- Its ability to gather fine grained image characteristics ensures that it can be used as an architecture to assist radiologists in early detection and diagnostic accuracy.

3.7 To detect and classify BC using AJFS-Att-CPVT:

- The proposed hybrid AJFS-Att-CPVT incorporates the Attention- ECNN, the PVT and the AJSO, to realize the effective and reliable detection of breast cancer in mammograms.
- ECNN is able to extract fine-grained lesion textures, whereas PVT is able to capture the global contextual patterns with the help of multi-head self-attention and feature pyramids. AJSO optimizes its attention weights and network parameters dynamically and improves convergence, stability, and lesion localization.
- The integration in comparison to the traditional CNNVit hybrids is the introduction of adaptive optimization-guided attention, which guarantees to enhance the generalization in various datasets.
- The novelty of the framework is that it is a bio-inspired optimization-based synergy of convolutional precision and transformer context to achieve robust and interpretable early breast cancer diagnosis.

Attention Mechanism

- The model's characteristics are subjected to the attention layer, which automatically determines the value of each mammography region for BC detection and differentiates between benign and malignant cancers.
- The attention mechanism can be thought of as a weighted summation: it first calculates the significance of each feature, normalizes the weights with the softmax function so that they total to one, and then multiplies these weights by the associated features.
- The weighted features are summed to get the final representation for BC classification. The corresponding Equations (2 to 4) are shown below:

$$b_s^l = \frac{\exp(f_s^l)}{\sum_{i=1}^{m} \exp(f_s^i)} \tag{2}$$

$$f_s^l = u_f^S \sigma\left(G_f\left[d_{s-1,} z_{t-1}\right] + V_f d_s + a_f\right) \tag{3}$$

$$\overline{J_s} = \sum_s^S b_s^S\, d_s \tag{4}$$

Here, f_s, $a_f \in Q^S$, $G_f Q^{S\times n}$, and $V_f \in Q^{n\times n}$ are learnable parameters; m is the number of neurons; b_s^l is the attention weight for the l^{th} feature at step s, f_s^l represents the relevance of the feature; u_f^S presents a trainable projection vector used to map the non-linear transformation into a scalar relevance score; $\left[d_{s-1,} z_{t-1}\right]$, indicates the feature space; and $\overline{J_s}$ the attention-enhanced representation that enables the model to focus on diagnostically significant regions in the mammogram for accurate BC detection.

3.8 Enriched Convolutional Neural Networks (ECNN):

- ECNNs are very useful in mammographic analysis because they can automatically learn hierarchical information ranging from low-level edges to high-level lesion characteristics.
- The convolutional operation that extracts features from mammogram patches can be stated as:

$$E_{j,i}^{(s)} = \sigma\left(\sum_{n=1}^{N} \sum_{m=1}^{M} V_{n,m}^{(s)} \cdot Y_{j+n,i_m} + a^s\right) \quad (5)$$

- In Equation (5), Y, is the input image, V^s is the convolution kernel of the kth feature map, a^s is the bias term, and σ represents the activation function (ReLU).
- $E_{j,i}^{(s)}$ is the output feature at location j, i in the s^{th} feature map. N, Mrepresents the feature map. Y_{j+n,i_m} pixel value at location $j + n, i_m$ of the input image.
- This operation allows the network to detect microcalcifications, tissue densities, and mass boundaries in mammograms.
- In Equation (6), to prevent degradation in deeper networks, residual connections are integrated.
- A residual block maps the input yto an output y as:

$$x = E(y, V) + y \quad (6)$$

- Where $E(y, V)$ represents the transformation (convolution, batch normalization, and activation), and the shortcut connection directly adds the input y to the transformed output. x is the output of the residual block.
- This design preserves gradient flow, enabling robust training on high-resolution mammograms.
- Batch Normalization (BN) is applied to stabilize training and improve convergence.
- For a given feature activation y_i, BN transforms it as follows:

$$\widehat{y_i} = \frac{y_i - \mu_A}{\sqrt{\sigma_A^2 + \epsilon}}, x_i = \gamma \widehat{y_i} + \beta \quad (7)$$

- In equation (7),ϵ is a small constant for numerical stability, γ, β, are learnable parameters, and μ_A and σ_A^2 are the batch mean and variance. y_i represents the activation of the i^{th} neuron before normalization.
- $\widehat{y_i}$ denotes normalized activation x_i is the final output after batch normalization.
- Since mammograms contain high-dimensional feature spaces, it is used to reduce redundancy.
- The dimensionality reduction process is expressed as:

$$X_j = V^l y_j \quad (8)$$

- Where y_j is the high-dimensional feature vector, Vis the eigenvector matrix of the covariance matrix, and X_j is the reduced feature representation in Equation (8). Enabling the model to minimize computational complexity while preserving critical lesion features aligning with the objective of improving efficiency and diagnostic accuracy in breast cancer detection.
- V^l indicates the Projection matrix formed from the eigenvectors of the covariance matrix (from PCA or linear transformation).
- This step minimizes computational cost while retaining discriminative lesion features.
- The modified CNN provides a solid platform for mammography classification by combining convolutional operations, residual learning, and BN regularization, resulting in improved lesion detection and minimizing false negatives and false positives.

3.9 Pyramid Vision Transformer

- The PVT is one of the most advanced transformer-based backbones for high-resolution image analysis.
- While alternative backbone networks have evolved, PVT continues to be widely recognized and representative.
- Unlike the standard ViT, which frequently produces low-resolution outputs, PVT uses dense partitioning to generate high-resolution feature maps, a critical property for mammogram analysis, where detecting small abnormalities such as microcalcifications and subtle structural distortions is critical for early BC diagnosis.
- The PVT-small backbone is used in the research because it balances computational efficiency and diagnostic accuracy well.
- The backbone consists of four hierarchical transformer blocks, each including a patch embedding module and transformer encoders.
- Mammograms are divided into fixed-size patches that are linearly embedded, positionally encoded, and processed using a transformer encoder.
- The encoder employs a multi-head self-attention (MHA) method that collects both local lesion-specific characteristics and general breast tissue context.
- For an input sequence Y, the queries (P), keys (L), and values (U) are generated as:

$$P = YV_P \quad (9)$$

$$L = YV_L \quad (10)$$

$$U = YV_U \quad (11)$$

- In Equations (9 to 11), Y_P, Y_L, Y_U are trainable weight matrices. framework to effectively capture both local lesion-specific details and global contextual information, thereby enhancing feature representation and improving the accuracy and reliability of early breast cancer detection.
- The scaled dot-product attention is defined as in Equation (12):

$$VB = softmax\left(\frac{PL^S}{\sqrt{c_l}}\right) \quad (12)$$

- P presents the queries, L^S denotes the Transposed Keys matrix (S usually denotes sequence length).
- c_l denotes the dimension of the key vectors (used for scaling). It plays a vital role in the proposed

framework's objective of enhancing feature representation and attention precision.
- The attention output is shown in Equation (13):

Attention $(P, L, U) = VB \cdot U$ (13)

- This forms the fundamental unit of self-attention.
- VB represents the Attention weights computed from queries and keys and the U is value matrix.
- Extending this, the multi-head attention (MHA) mechanism is expressed in Equation (14):

MHA (P, L, U) = Concat $(head_1, \ldots, head_h)V_A$ (14)

$head_i$ = Attention (P_i, L_i, U_i) (15)

- In Equation (15), P_i, L_i, U_i is the Query, Key, and value matrices for the i^{th} head h is the number of attention heads, and V_A is a projection matrix.The output of each attention head, headi enables the model to capture multi-scale contextual dependencies within mammogram regions, aligning with the study's objective of accurately distinguishing benign and malignant lesions through i^{th} enhanced feature representation and diagnostic reliability.
- The MHA method successfully fuses multi-scale context data, enabling the model to learn tissue distribution in global breast tissue, even as there is limited contextual information about lesion boundaries.
- PVT employs Spatial Reduction Attention (SRA) to remedy the computational challenge of O(n2d) ensuing from the self-attention operation.
- The SRA uses a matrix reduction to reduce the dimension of P and L while retaining structural integrity.
- In addition, the use of a progressive shrinking pyramid reduces the length of sequence at deeper layers, and creates multi-scale feature representation, an intrinsically valuable feature when detecting lesions of various sizes in a mammogram.
- Lastly, a feature pyramid network (FPN) blends outputs from multiple stages to provide better hierarchical feature representation and lesion localization and classification.
- In the proposed system, the already enhanced PVT-based features are augmented further through attention enriched convolutional tuning to provide more reliable BC detection and classification in mammograms.

3.9 Artificial Jellyfish Search Optimization (AJSO):

- The AJSO is an algorithm based on a bio-inspiration of the behavior of jellyfish searching in the water.
- AJSO has the objective of efficiently exploring and exploiting the search space to produce optimal solutions at a global exploration and a local exploitation.
- During the early BC diagnosis of mammograms based on ViTs, attention mechanisms and model parameters are optimized with the help of AJSO, and benign and malignant lesions are reliably identified.
- Chaotic maps and random generation were used for initialization to increase the distribution of candidate solutions, accelerate convergence, and reduce entrapment in local minima.
- The logistic map was determined to perform best, mathematically expressed as follows in Equation (16):

$\vec{Y}_{j+1} = \eta\vec{Y}_j\left(1 - Y_j\right), 0 \leq \vec{Y_0} \leq 1$ (16)

- Here, $\vec{Y_j}$ is the vector of chaotic values for the j^{th} jellyfish, $\vec{Y_0}$ is the initial random vector, and $\eta = 4$. Using the AJFS-Att-CPVT framework to guarantee a varied parameter distribution and speed convergence toward the ideal configuration for precise and trustworthy breast cancer diagnosis.
- The solution with the best fitness, j shows iteration index in the chaotic sequence.
- $\vec{Y}^*$ represents the most promising configuration for improving the ViT's ability to accurately classify lesions in mammographic images.
- Jellyfish positions are updated along the ocean current according to Equation (17):

$\vec{Y}_j(s+1) = \vec{Y_j}(s) + \vec{t} .* (\vec{Y}^* - \beta * t_1 * \mu$ (17)

- $\vec{Y_j}(s)$ is the current position of the j^{th} jellyfish at iteration s. $\vec{Y}^*$ Position of the best solution found so far (global optimum in fitness).
- $\vec{t}$ is time step or movement factor along the ocean current. β denotes scaling factor controlling step size toward the optimal solution. t_1 illustrates random coefficient for stochastic movement.These parameters are optimized to enhance convergence stability and precise lesion localization within the framework for early breast cancer detection.
- μ indicates the attraction factor or influence of ocean current on jellyfish movement. In Equation (18), passive motion perturbs the current location locally:

$\vec{Y}_j(s+1) = \vec{Y}_j(s) + r_3 * \gamma * (V_a - M_a)$ (18)

- V_a target jellyfish position or reference vector in the swarm, M_a current mean position of neighboring jellyfish. In Equation (19), active motion guides jellyfish toward better solutions:

$\vec{Y}_j(s+1) = \vec{Y}_j(s) + \vec{r} * \vec{C}$ (19)

- $\vec{r}$ Random vector controlling stochastic direction of active motion, $\vec{C}$ Active motion coefficient that directs jellyfish toward better solutions in the search space.
- The time control mechanism determines switching among ocean current, passive, and active motions:

$d(s) = \left(1 - \frac{s}{s_{max}}\right) * (2 * r - 1)$ (20)

- Where s is the current assessment, s_{max} is the maximum evaluation, and r is the random number presented in Equation (20). When $d(s) \geq d0$, jellyfish follow the ocean current. To improve lesion localization and classification accuracy in mammography-based breast cancer diagnosis, this technique is improved to adaptively modify attention weights and network parameters.
- Otherwise, they move in the swarm, and select either passive motion or active motion, depending on a secondary random number. This method guarantees efficient tuning of attention weights and parameters of the model in the ViT-based BC detection model, which enhances the detection capability of small and subtle lesions on mammograms.
- AJSO is designed to offer a powerful, adaptable, and effective optimization procedure to guide the search to the optimal solutions without falling into the local minima.
- When ViTs are used in early detection of BC in mammography images by AJSO enhances model generalization and model stabilization, as well as becomes a highly reliable and accurate lesion classification model that can facilitate earlier and more accurate diagnosis. The algorithm 1 depicts hybrid AJFS-Att-CPVT algorithm.

Algorithm 1: AJFS-Att-CPVT for Mammogram Classification

$Input$: $Mammogram\ dataset\ D$
$Output$: $Predicted\ labels\ (Benign/Malignant)$

1. $Preprocessing$:
   $-\ Normalize, Gaussian\ blur, augment\ images\ (rotation, flip, contrast)$

2. $Feature\ Extraction\ (CNN)$:
   $-\ Convolution\ \rightarrow\ Residual\ Blocks\ \rightarrow\ Batch\ Normalization\ \rightarrow\ Dimensionality\ Reduction$

3. $Attention\ Mechanism$:
   $-\ Compute\ feature\ relevance$
   $-\ Compute\ attention\ weights$
   $-\ Generate\ attention\ -\ enhanced\ feature\ representation$

4. $Pyramid\ Vision\ Transformer\ (PVT)$:
   $-\ Patch\ embedding\ \rightarrow\ Multi-Head\ Self-Attention\ \rightarrow\ Fuse\ multi-scale\ features\ via\ FPN$

5. $Artificial\ Jellyfish\ Search\ Optimization\ (AJSO)$:
   $-\ Initialize\ population\ with\ chaotic\ map$
   $-\ Update\ positions\ via\ ocean\ current, passive, or\ active\ motio$
   $-\ Optimize\ attention\ weights\ and\ model\ parameters$

6. $Classification$:
   $-\ Predict\ labels\ (Benign/Malignant)$
   $-\ Evaluate\ Accuracy, Precision, Recall, F1-score, AUC-ROC$

- The hybrid AJFS-Att-CPVT model provides better diagnostic accuracy through the fusion of convolutional precision and transformer contextual learning. Its adaptive optimization, which uses AJSO, enhances convergence and minimizes over-fitting and multi-scale feature fusion makes lesion localization stronger.
- This interaction permits effective generalization of diverse mammogram data, and this guarantees credible and decipherable early breast cancer diagnosis.

4. Result and Discussion

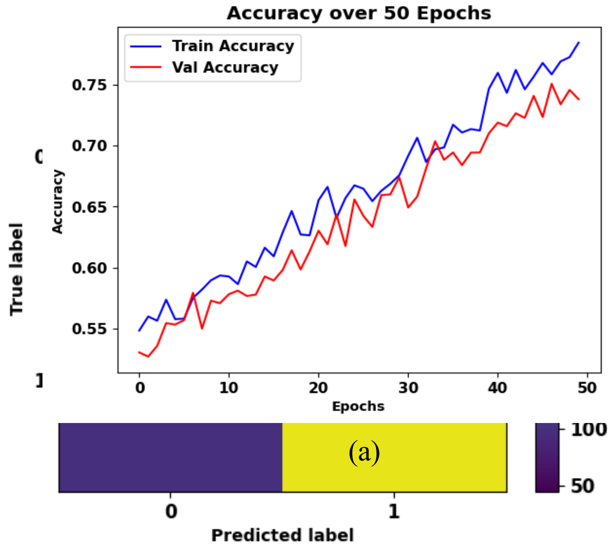
Enhance the generalization, the tests were done with pre-processed mammograms Python and TensorFlow were used to run the experiments in a workstation with an NVIDIA RTX 3090 graphics card, 24 GB of VRAM, and an Intel i9 processor.

Confusion Matrix

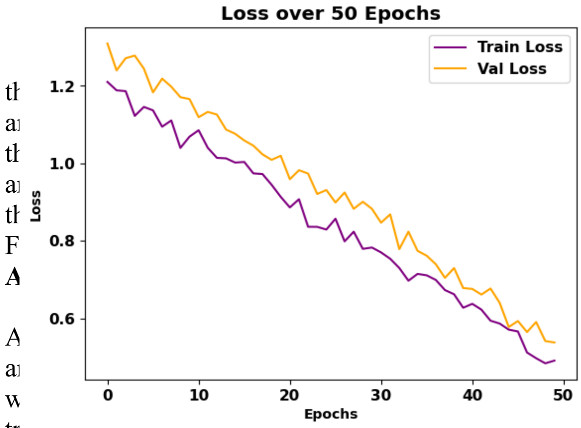
The higher diagnostic capabilities of the suggested model to distinguish between benign and malignant breast tumors are displayed in the confusion matrix in Figure 5. While exhibiting excellent precision in detecting cancer cases and high reliability in identifying benign instances, the algorithm proved resilient when dealing with mammography data, properly recognizing benign and malignant breast lesions in the vast majority of cases. Although there is a small percentage of malignant cases classified wrongly as benign cases, there is good overall equilibrium between sensitivity and specificity; hence, the model can be used to assist clinical decisions. As a result, early diagnosis of BC is improved by reducing false alarms while providing reliable categorization of benign and malignant lesions and making a distinction between lesion types.

Figure 5: Model Performance Confusion Matrix

Receiver Operating Characteristic (ROC) Curve Analysis



(a)



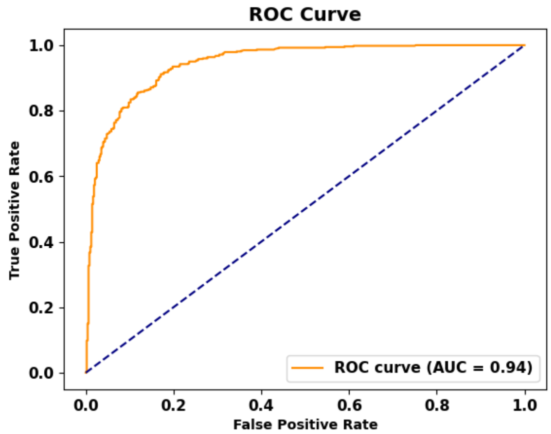
(b)

The proposed AJFS-Att-CPVT model successfully differentiates between benign and malignant breast lesions, as indicated by the ROC curve. The ROC curve shows the connection between the True Positive Rate and the False Positive Rate at various threshold settings. As seen in Figure 6, the model's robustness is demonstrated by the curve's steep climb toward the upper-left corner, which achieves high sensitivity with a low false positive rate, confirming its dependability in the early diagnosis of BC from mammograms.

Figure 6: ROC Curve for BC Classification

Precision–Recall (PR) curve

The PR trade-off of the AJFS-Att-CPVT model in the detection of BC using mammograms is represented by the PR curve. The curve indicates strong reliability in detecting malignant lesions with low false positives as well as detecting false positives within large ranges of



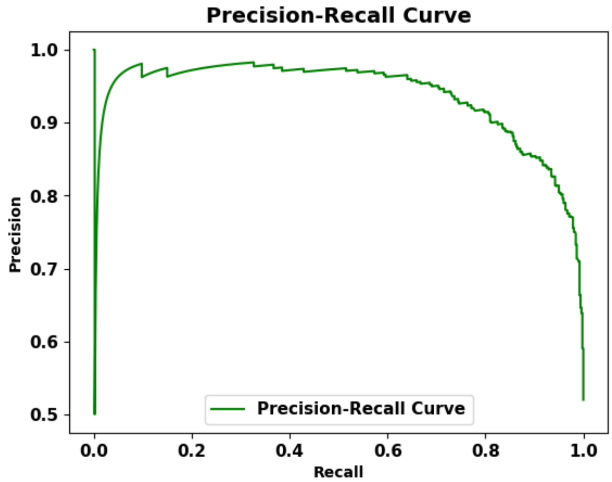
recall values. The fact that the model makes the most confident predictions is manifested in the low recall values and the fact that it tends to decrease gradually to

training and validation accuracies. By the 50th epoch, training accuracy reaches approximately 0.78, while validation accuracy achieves about 0.74, demonstrating strong generalization capability. While the final test accuracy (95.8%) was obtained after fine-tuning the complete hybrid model. This stable upward trend highlights the effectiveness of the hybrid architecture and the optimization process in enhancing classification accuracy for BC detection displayed in Figure 8 (a).

Figure 7: Precision–Recall Curve for BC Classification

Loss curve over 50 epochs

The convergence behavior of the proposed AJFS-Att-CPVT framework over 50 epochs is represented to visualize its performance on both the training and the



validation sets, as seen in the loss plot. The loss curves of 1.2 (training loss) and 1.3 (validation loss) indicated the model's initial (first epoch) difficulty in fitting the training data. However, both curves trended downward in epoch numbering, suggesting two possibilities: learning better features, and optimizing better features.

Figure 8: (a) Training and Validation Accuracy and (b) Loss Curves over 50 Epochs

Evidence of strong generalization (with mild to little overfitting) is clear in the final epoch, where the training loss decreased to about 0.45 and the validation loss to about 0.55. The robustness of the hybrid architecture for detecting BC tasks was confirmed by how close the two curves followed each other, meaning the model effectively minimized classification errors while being stable in the training presented in Figure 8 (b).

Ablation study

This was done to determine the role of each component in the proposed AJFS-Att-CPVT. Findings indicate that the combination of preprocessing, Att-CPVT, and AJSO optimization has a gradual increase in performance. All the modules increase feature extraction, convergence, and generalization, which proves that the overall impact of the hybrid architecture results in maximum accuracy, reliability and diagnostic strength in detecting breast cancer. Table 4 presented the results of ablation study of proposed AJFS-Att-CPVT framework. The addition of every component enhances the performance in all metrics. Data quality is improved with preprocessing, feature extraction and lesion localization is improved with Att-CPVT and parameter tuning and convergence is improved with AJSO. The hybrid model has the highest accuracy (95.8) and balanced specificity, recall, and F1-score, and these results prove the efficiency of the hybrid architecture in providing accurate detection of breast cancer.

Performance Metrics

Standard evaluation measures for BC detection models include metrics. These metrics completely assess the model's ability to classify and differentiate between benign and malignant lesions. To illustrate its improved diagnostic capabilities, the proposed AJFS-Att-CPVT model is compared to existing techniques such as Modified Visual Geometry Group 16-layer network (VGG16) (MVGG16) [23], MVGG16 with data augmentation [23], Hybrid MVGG16 with ImageNet [23], and (Residual Network with 50 layers (ResNet-50 CNN) [24].

Accuracy

Accuracy is the proportion of mammograms that are correctly classified, encompassing both benign and malignant cases. Table 1 and Figure 9 present the BC detection, which reflects the model's overall effectiveness.Yet, due to the tendency to skew mammography datasets with fewer malignant incidents, accuracy is not a sufficient criterion that needs to be taken into consideration, whereas sensitivity and specificity should also be mentioned. The proposed AJFS-Att-CPVT model had the accuracy of 95.8%.

Recall

It is the capacity of the model to pinpoint correctly the actual cancer patients. This attribute is significant to BC detection as a false negative, or an undetected malignant lesion, may lead to a delay in diagnostic and treatment, which may have an impact on the outcome of the patient. Table 1 and Figure 9 shows a recall of 94.8%, the suggested AJFS-Att-CPVT outperformed other models currently in use.

F1- score

To balance the danger of false positives and negatives, the F1-score is calculated as the harmonic mean of accuracy and recall. It offers a unified indicator of the model's ability to consistently identify positive predictions while detecting malignant cases in mammography analysis. The assessment of the proposed AJFS-Att-CPVT achieved a 93.8% compared to existing models illustrated in Table 1 and Figure 9.

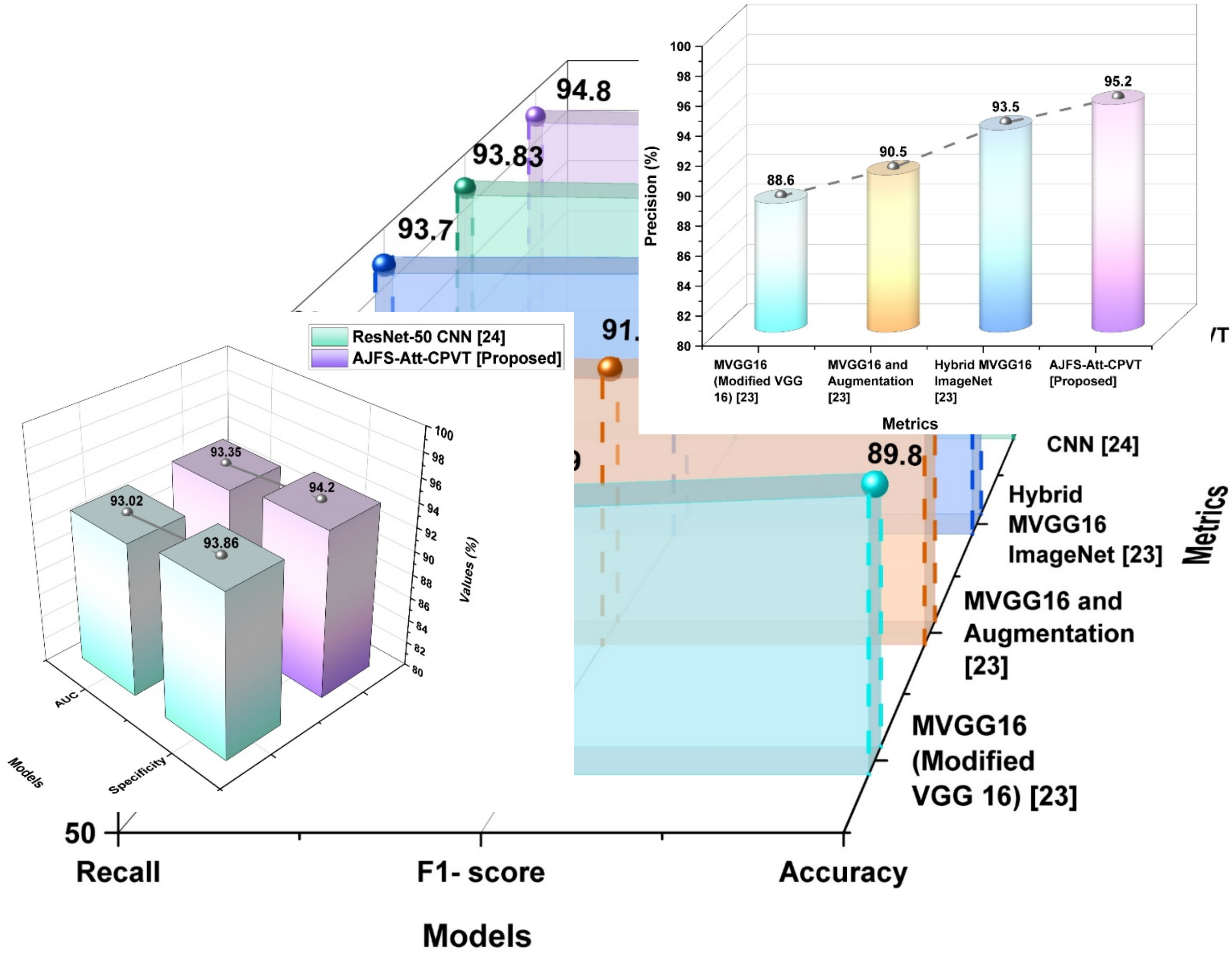


Figure 9: Analysis of the Proposed Method with Existing Methods

Table 1: Comparative Performance of Proposed AJFS-Att-CPVT Model with Existing BC Detection Methods

Models	Recall	F1-score	Accuracy
MVGG16 (Modified VGG 16) [23]	87.2	87.9	89.8
MVGG16 and Augmentation [23]	92.2	91.3	92.8
Hybrid MVGG16 ImageNet [23]	93.7	93.6	94.3
ResNet-50 CNN [24]	93.83	93.03	93
AJFS-Att-CPVT [Proposed]	94.8	93.8	95.8

F1- score

To balance the danger of false positives and negatives, the F1-score is calculated as the harmonic mean of accuracy and recall. It offers a unified indicator of the model's ability to consistently identify positive predictions while detecting malignant cases in mammography analysis. The assessment of the proposed AJFS-Att-CPVT achieved a 93.8% compared to existing models illustrated in Table 1 and Figure 9.

Specificity

The proportion of circumstances that are accurately classified as non-cancerous is known as specificity. In the research objective, excellent specificity is required to avoid misclassifying healthy tissue as malignant, reducing patient concern, and avoiding unneeded diagnostic procedures. Table 2 and Figure 10 show the analysis of the specificity and AUC. The Specificity of 94.20% was achieved by the proposed AJFS-Att-CPVT.

Precision

Out of all instances that are displayed to be malignant, precision is the proportion of correctly predicted malignant cases.

Table 2: Specificity and AUC Comparison between ResNet-50 CNN and the Proposed AJFS-Att-CPVT Model

Model	Specificity	AUC

ResNet-50 CNN [24]	93.86	93.02
AJFS-Att-CPVT [Proposed]	94.2	93.35

Figure 10: Comparative evaluation of specificity and AUC

Table 3: Comparative Performance of Precision for BC Detection Models

Models	Precision
MVGG16 (Modified VGG 16) [23]	88.6
MVGG16 and Augmentation [23]	90.5
Hybrid MVGG16 ImageNet [23]	93.5
AJFS-Att-CPVT [Proposed]	95.2

Table 3 and Figure 11 display the BC detection, with high accuracy, indicating that the model is very dependable when it classifies a lesion as malignant, which reduces needless biopsies and prevents false alarms in clinical settings. The precision achieved was 95.2% for the proposed AJFS-Att-CPVT.

Figure 11: Assessment of precision for BC Detection

Table 4: The Ablation Demonstrates Each Component's Performance Contribution in the Suggested AJFS-Att-CPVT Framework for Breast Cancer Detection

Method	Accuracy (%)	Specificity (%)	AUC (%)	Recall (%)	F1-score (%)
AJSO	93.7	92.6	91.8	92.9	92.3
Att-CPVT	94.2	93.1	92.5	93.4	92.8
Att-CPVT + AJSO (Proposed Model)	**95.8**	**94.2**	**93.4**	**94.8**	**93.8**

4.1 Discussion

Even though the results of the previous DL frameworks on breast cancer detection are encouraging, there are a number of limiting factors that still determine their application in the real world. Synthetic data augmentation cannot generally well represent the natural variability and structural complexity of actual mammographic data [21], hence limiting the extrapolation of models to a clinical setting. These man-made changes are not entirely reflective of changes in the density of breast tissue, morphology of the lesions, or imaging noise which accounts to inconsistencies in performance with various screening situations. As a result, the existing models are inclined to demonstrate low flexibility when implemented to the dissimilar populations of patients and imaging regimes.

The current DL systems are biased in their data set, prone to overfitting, and lack sensitivity in dense or thick breast tissues [22]. These limitations reduce the diagnostic accuracy of various population groups, especially when the features of lesions are hidden by the density of glands. Consequently, early-stage tumors or other minute abnormalities are incorrectly labeled through models and lead to false negatives, which compromises clinical reliability. The mentioned shortcomings underscore the necessity of having algorithms that can strike a balance between sensitivity and specificity without being unstable when applied to heterogeneous imaging datasets.

The Hybrid MVGG16-ImageNet architecture [23] has a limited capacity to be clinically generalizable due to the fact that it uses pretrained weights on Image Net which are not specific to mammography. It also relies on specific 2D and 3D mammographic images and is therefore not applicable in practical screening environments. This framework cannot easily retrieve subtle mammographic characteristics like micro calcifications and minor structural distortions as there is an incompatibility between mammographic features of nature and the properties of medical imaging. Moreover, its pure convolutional structure does not allow capturing global and long-range dependencies, without which the analysis of complex tissue structures is impossible. This lack of mechanisms that operate based on transformers limits the contextual awareness and the lack of interpretability tools limits the confidence that clinicians place on automated results.

ResNet-50 architecture [24] has a relatively lower level of accuracy because it relies on the IN breast dataset that fails to cover all mammographic variations. Its binary classification method does not study multi-class or stage diagnostic complexities, which constrains clinical decision support. In addition, lack of ablation, cross-validation and external dataset assessment undermines reproducibility and scientific rigor. In real-time healthcare systems computational inefficiency also limits deployment, especially in systems with low processing capabilities. These problems demonstrate the need of optimization based, interpretable, and computationally adaptive models.

To overcome these problems, the AJFS-Att-CPVT model uses the capabilities of Vision Transformers (ViTs) with Bio-inspired optimization in the Artificial Jellyfish Search (AJS) algorithm. This hybrid architecture is a hybrid that associates the localized capabilities of CNNs with the contextual capabilities of ViTs, allowing both fine-grained and global representations of an image to be captured at the same time. Adaptive optimization in AJS guarantees a better convergence aspect, stability, and fine-tuning of the parameters. The Pyramid ViT architecture has been proven to be an effective learner of multi-scale contextual relationship, and improved preprocessing, including Gaussian blurring, normalization, and targeted augmentation, can improve image quality and reduce overfitting. All these mechanisms increase strength, diagnostic accuracy, and interpretability in a variety of mammograms.

The AJFS-Att-CPVT framework has a high potential to implement AI-assisted diagnostic systems because it provides better generalization, transparency, and computational efficiency. The modules of the model with attention enrichment allow offering interpretable visualization maps, which will help radiologists make decisions and give them more confidence when making a diagnosis. Its CNNViT hybrid architecture is optimized with AJS to sustain a high detection rate in diverse imaging situations and clinical configurations. Future opportunities are to broaden the framework to multimodal imaging integration, real-time screening on clouds and to adaptive learning environments where the performance of models is optimized by radiologist feedback. These developments make AJFS-Att-CPVT a clinically dependable and cost-effective tool in early breast cancer detection, which is a part of the overall development of intelligent and explainable diagnostic protocols.

5. Conclusion

The present study proposed an adapted deep learning (DL) model known as AJFS-Att-CPVT to identify the early signs of breast cancer (BC) and classify it based on mammographic images. Such a system combines the Artificial Jellyfish Search Optimization (AJSO) with the Attention-Enriched Convolutional Neural Networks (CNNs) and PVT backbone thus enhancing the feature extraction and model optimization. The disadvantages of the classical CNN-based and transformer-based models are the driving force behind the hybrid, because they not only fail to ensure local and global feature learning but also cannot be easily adjusted to other datasets and require a substantial amount of computing capacity. The AJFS-Att-CPVT model differentiates benign lesions and malignant ones through the assistance of the multi-scale representational power of Vision Transformers (ViTs). The attention mechanism combined with the model assists in the highlighting of the areas of the mammograms of significance in the diagnosis and therefore enables the model to understand the details as well as the context of the pictures. In addition to that, the AJSO application opens the doors to the dynamic adjustment of the parameters involved in the model training, which not only affects the stability of convergence, classification accuracy, and robustness positively but also minimizes overfitting. The image was normalized and the quality of the image was also improved in the process of preprocessing. This involved the use of the normalization process, the application of Gaussian blurring, and application of data augmentation techniques such as rotations, flips and adjustment of contrast levels. The images as the inputs were thus guaranteed to be free of noise, well-balanced and also, varied and this was carried out bearing in mind the enhancement of the model generalization as also, counteracting the impact of the variability of the dataset. Overall, the dataset that was used in this work consisted of 3,383 annotated mammogram images that were collated and clustered on the basis of res.

Acknowledgment

None

Conflicts of Interest

The authors have no conflicts of interest to declare.

Data availability

The dataset used in this study is publicly available on Kaggle and can be accessed at: https://www.kaggle.com/datasets/hayder17/breast-cancer-detection

Author's Contribution Statement

Daphne Sherine. H[1]* – Methodology, Analysis, Testing, Writing – original draft, Writing – review and editing, Supervision. G. Revathy[2] – Data Collection, Conceptualization, Investigation, Writing – review and editing.

References

[1]. Mehmood M, Ayub E, Ahmad F, Alruwaili M, Alrowaili ZA. Machine learning enabled early detection of breast cancer by structural analysis of mammograms. Computers, Materials & Continua. 2021;67(1):641-657. https://doi.org/10.32604/cmc.2021.013774

[2]. Almalki YE, Soomro TA, Irfan M, Alduraibi SK, Ali A. Computerized analysis of mammogram images for early detection of breast cancer. Healthcare. 2022;10(5):801. https://doi.org/10.3390/healthcare10050801

[3]. Singh L, Alam A. An efficient hybrid methodology for early detection of breast cancer in digital mammograms. Journal of Ambient Intelligence and Humanized Computing. 2024;15(1):337-360. https://doi.org/10.1007/s12652-022-03895-w

[4]. Khan AA, Arora AS. Thermography as an economical alternative modality to mammography for early detection of breast cancer. Journal of Healthcare Engineering. 2021;2021(1):5543101. https://doi.org/10.1155/2021/5543101

[5]. Alshawwa IA, et al. Advancements in early detection of breast cancer: innovations and future directions. [Journal name not specified]. 2024.

[6]. Jung JI, et al. Thoracic manifestations of breast cancer and its therapy. Radiographics. 2004;24(5):1269-1285. https://doi.org/10.1148/rg.245035062

[7]. Muduli D, Dash R, Majhi B. Automated breast cancer detection in digital mammograms: a moth flame optimization-based ELM approach. Biomedical Signal Processing and Control.2020;59:101912. https://doi.org/10.1016/j.bspc.2020.101912
[8]. Ayana G, et al. Vision-transformer-based transfer learning for mammogram classification. Diagnostics. 2023;13(2):178. https://doi.org/10.3390/diagnostics13020178
[9]. Alruily M, et al. Enhancing breast cancer detection in ultrasound images: an innovative approach using progressive fine-tuning of vision transformer models. International Journal of Intelligent Systems. 2024;2024(1):6528752. https://doi.org/10.1155/int/6528752
[10]. Al-Hejri AM, et al. ETECADx: ensemble self-attention transformer encoder for breast cancer diagnosis using full-field digital X-ray breast images. Diagnostics. 2022;13(1):89. https://doi.org/10.3390/diagnostics13010089
[11]. Sabry M, et al. Enhancing breast cancer diagnosis with multi-resolution vision transformers and robust decision-making. IEEE Access. 2025. https://doi.org/10.1109/ACCESS.2025.3570840
[12]. Annepu V, et al. Advanced breast cancer diagnostics with PolyBreastVit: a combined PolyNet and vision transformer approach. Applied Computational Intelligence and Soft Computing. 2024;2024(1):5574638. https://doi.org/10.1155/2024/5574638
[13]. Lin Q, et al. Artificial intelligence-based diagnosis of breast cancer by mammography microcalcification. Fundamental Research. 2025;5(2):880-889. https://doi.org/10.1016/j.fmre.2023.04.018
[14]. Dembrower K, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. The Lancet Digital Health. 2020;2(9):e468-e474. https://doi.org/10.1016/S2589-7500(20)30185-0
[15]. Croock MS, Khuder SD, Korial AE, Mahmood SS. Early detection of breast cancer using mammography images and software engineering process. Telkomnika. 2020;18(4):1784-1794. https://doi.org/10.12928/telkomnika.v18i4.14718
[16]. Cai Y, Li Z, Wang Y. Enhancing breast cancer detection with vision transformers and graph neural networks. arXiv preprint.2025;arXiv:2507.13372. https://arxiv.org/abs/2507.13372
[17]. Tanveer H, Faheem M, Khan AH, Adam MA. AI-powered diagnosis: a machine learning approach to early detection of breast cancer. International Journal of Engineering Development and Research. 2025;13(2):153-166. https://doi.org/10.18535/ijsrm/v13i04.ec03
[18]. Chougrad H, Zouaki H, Alheyane O. Multi-label transfer learning for the early diagnosis of breast cancer. Neurocomputing.2020;392:168-180. https://doi.org/10.1016/j.neucom.2019.01.112
[19]. Maqsood S, Damaševičius R, Maskeliūnas R. TTCNN: a breast cancer detection and classification approach toward computer-aided diagnosis using digital mammography in early stages. Applied Sciences. 2022;12(7):3273. https://doi.org/10.3390/app12073273
[20]. Elkorany AS, Elsharkawy ZF. Efficient breast cancer mammogram diagnosis using three deep neural networks and term variance. Scientific Reports. 2023;13(1):2663. https://doi.org/10.1038/s41598-023-29875-4
[21]. Prodan M, Paraschiv E, Stanciu A. Applying deep learning methods for mammography analysis and breast cancer detection. Applied Sciences. 2023;13(7):4272. https://doi.org/10.3390/app13074272
[22]. Suh YJ, Jung J, Cho BJ. Automated breast cancer detection in digital mammograms of various densities via deep learning. Journal of Personalized Medicine. 2020;10(4):211. https://doi.org/10.3390/jpm10040211
[23]. Khamparia A, et al. Diagnosis of breast cancer based on modern mammography using hybrid transfer learning. Multidimensional Systems and Signal Processing. 2021;32(2):747-765. https://doi.org/10.1007/s11045-020-00756-7
[24]. Rahman H, Naik Bukht TF, Ahmad R, Almadhor A, Javed A. Efficient breast cancer diagnosis from complex mammographic images using deep convolutional neural network. Computational Intelligence and Neuroscience. 2023;2023(1):7717712. https://doi.org/10.1155/2023/7717712.

Mrs. Daphne sherine.H completed her Master degree from SRM University, kattankulathur, chennai.she is pursuing her PhD in department of computer science and Engineering at vels institute of science technology and advanced studies (VISTAS). Her current research areas are medical image processing,deep learning and machine learning.she presented a paper in the international conference at VIT Chennai and received the best paper award and the paper was published in the book chapter and scope indexed in October 2025.

Dr. G. Revathy completed Master degree from Anna University, Chennai. She earned her Doctor of Philosophy in Computer Science and Engineering with specialization of Cybersecurity from Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai. She has more than 10 years of teaching experience at various institutions. Currently, she is an Assistant professor in the department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS). Her current research interests are Machine learning, Cybersecurity, deep learning, Image processing and wireless networks. She has authored 5 SCI journals and 15 scopus indexed journals and conferences and received the Young Researcher Award in 2023. She is strongly dedicated to advancing data driven technologies and AI based research. She is involved in guiding students and innovating with hands-on learning.