Journal of Carcinogenesis
Covering Genesis to Therapeutics

Journal of Carcinogenesis, ISSN: 1477-3163
2025; Vol 24: Issue 4s, https://carcinogenesis.com/

Multi-Model Learning for Cheast Cancer Detection: Combining Medical Images and Clinical Data using self –Attention with Dilated CNN and Transformers

Dr. S.Sooriya Prabha[1], Dr. R.Saravanan[2], Dhanasekhar D[3], Dr. S.K.Lakshminarayana[4]

[1] Department of Electrical & Electronics Engineering, Mother Theresa Institute of Engineering & Technology, Andhra Pradesh, India

Email ID: sethu25112009@gmail.com

[2]Department of Electrical & Electronics, Engineering Mother Theresa Institute of Engineering & Technology, Andhra Pradesh, India

Email ID: saravaneeetanj@gamil.com

[3] Assistant professor, Department of Electronics and communication engineering, Mother Theresa Institute of Engineering & Technology, Andhra Pradesh, India

Email.ID: degaladhanasekhar@mtieat.org

[4]Associate Professor, Department .of Electronics and Communication Engineering, Mother Theresa Institute of Engineering & Technology, Andhra Pradesh, India

Email.ID: sklnarayana.ece@mtieat.org

ABSTRACT

one of the leading causes of demise for women global continues to be breast most cancers, highlighting the importance of activate and particular detection. Self-Attention with Dilated Convolutional Neural Networks (CNNs) and transformer, a sort of deep mastering set of rules that has proven superb efficacy in clinical photograph and Clinical statistics analysis, offer a strong technique for predicting breast most cancers [2]. This observe investigates the significance of self-attention mechanisms and dilated convolution with CNN fashions for you to detect and classify breast cancer the usage of medical records and mammography pics. Given that the anomalous region in a mammography picture is much smaller than the normal region, CNN models consider all image regions equally. As a result, the anomalous region teaches the models less representational properties. This study uses dilated convolution in unification with a multi-level self-attention enhanced CNN model to classify cancer images. The capacity of the suggested. The dilated CNN component gathers multi-scale spatial information from medical images, while the Transformer layer guarantees long-range relationships between clinical features, providing a thorough picture of the patient's state. We show the effectiveness of this approach on a publicly available dataset comprising clinical and imaging data. The suggested attention-augmented CNN model outperforms conventional CNN techniques in terms of classification accuracy in distinguishing malignant and benign instances by automatically extracting more complicated characteristics from the cancerous zone [9]. In order to lower mortality rates and to enhance the results of breast cancer treatment for patients, the study shows how attention-augmented CNN-based systems might enhance early identification and treatment planning. Techniques like data augmentation, transfer learning, and explainable AI are used to overcome issues including class imbalance, data scarcity, and model interpretability. This work aims to help develop more personalized healthcare solutions by enabling the development of more reliable, data-driven diagnostic tools that could make use of both visible and invisible affected person statistics.

Keywords: *Breast Cancer, Self-Attention, Dilated Convolutional Neural Networks (CNNs), Transformer, Healthcare AI*

How to Cite: Dr. S.Sooriya Prabha, Dr. R.Saravanan, Dhanasekhar D, Dr. S.K.Lakshminarayana, (2025) Multi-Model Learning for Cheast Cancer Detection: Combining Medical Images and Clinical Data using self –Attention with Dilated CNN and Transformers, *Journal of Carcinogenesis*, *Vol.24, No.4s*, 1024-1034

1. INTRODUCTION

Breast cancer is now frequently detected by diagnostic techniques like mammography, biopsies, and ultrasounds [3]. These conventional methods do have certain drawbacks, though, including false-positive and false-negative outcomes. Patients may experience emotional discomfort as a result of needless biopsies caused by false positives, and treatment may be postponed due to false negatives. The possibility of diagnostic errors is further increased by the variability that is frequently introduced by human interpretation of imaging data. These restrictions highlight the need for cutting-edge and trustworthy instruments to help medical practitioners diagnose breast cancer early [4].

Using clinical imaging methods like Xrays, ultrasounds, and mammography, which help radiologists differentiate questionable injuries, is typically part of the demonstration cycle for bosom malignant growth [5]. Consequently, the integration of additional data sources, such as clinical information on the patient's medical history, segment nuances, and genetic factors, has become an indisputable component of developing increasingly accurate demonstration tools [6].

The development of deep learning techniques in recent years has fundamentally altered medical image evaluation due to CNNs' higher performance in responsibilities like picture categorization and detection [7]. Although CNNs are effective at identifying spatial features in medical images, they frequently ignore the rich, non-visual information found in clinical records [1]. The potential of medical images to discriminate is not available to methods that solely use clinical data. Multi-model learning techniques that integrate imaging and clinical data have generated significant interest in an attempt to bridge this gap [8].

Natural language processing (NLP) has been dominated by transformers, which are used in text-speech translation to speech recognition, synthesis, and natural language production. The first transformer was introduced to handle sequential inference tasks in NLP as an example of deep learning systems [17]. Transformers employ stacked self-attention layers to record long-term dependencies of sequential input, while neural networks that repeat (like the long short-term memory network) explicitly employ a series of inference procedures. Transformers are therefore effective because they stack incredibly deep models and efficient since they address a sequential learning problem in a single shot. To solve NLP issues, a number of transformer architectures trained on large-scale designs have gained popularity.Examples include Transformers, BERT, and GPT3, which provide bidirectional encoder representations.

We suggest a unique multi-modal learning architecture that integrates clinical data with medical imaging for breast cancer identification in order to overcome these difficulties. Our method combines Transformers that use self-attention mechanisms to identify long range correlations in clinical features and picture data, and dilated CNNs for improved image feature extraction [10]. By allowing the model to focus on the most informative regions of the image and clinical data, the self-attention method enhances feature fusion and provides a more comprehensive picture of the patient's condition [11]. Furthermore, multi-scale analysis—which is crucial for identifying cancers of various sizes and forms in medical images—is made possible by the usage of dilated CNNs. Figure 1 shows the benign and malignant image of breast.

Figure 1. a. Benign Image b. Malignant Image

2. RELATED WORK

A. Challenges in Diagnosing Breast Cancer

Breast cancer diagnosis that is both accurate and timely is still a challenge despite tremendous progress in medical imaging and machine learning. Numerous factors interact in a complex way in this situation.

Tumor Variability: The size, form, and structure of breast tumors vary greatly. While the margins of some tumors are uneven or hazy, others manifest as well-defined lumps. Using conventional diagnostic techniques to differentiate between benign and malignant growths is frequently made more difficult by this diversity.

Dense Breast Tissue: It can be particularly challenging to detect malignant tumors on mammograms in people with dense breast tissue. Traditional imaging methods' sensitivity and efficacy can be diminished by dense tissue's ability to conceal malignancies.

Large Data Volumes: As imaging technology have advanced quickly, so too has the volume of data produced during brea st cancer screenings. These enormous datasets take a long time to analyze by hand , and it is more likely to error [12].

In order to overcome these obstacles and increase diagnostic efficiency and accuracy, technological breakthroughs are necessary, as demonstrated by these difficulties.

B. AIs Potential to Revolutionize Healthcare

The technology known as artificial intelligence (AI) has revolutionized several industries, including healthcare.To solve difficult problems, artificial intelligence (AI) leverages its ability to analyze vast amounts of data, spot patterns, and ma ke predictions based on that data. Artificial intelligence (AI) systems have demonstrated a tremendous deal of potential in enhancing medical imaging diagnostic skills.

Breast cancer can be can be detected with AI tools by analyzing medical pictures, including ultrasound scans, histopathology slides, mammograms, and clinical data to identifying anomalies that may be signs of the disease. When compared to conventional diagnostic techniques, these technologies provide a number of benefits. For instance, AI systems are able to produce accurate and consistent findings since they are not affected by bias or weariness.

Automated feature extraction from raw data by deep learnig algorithms eliminates the need for human feature selection. In the identification of breast cancer, where minute variations in imaging can reveal malignancy, this is quite helpful [13].

3. METHODOLOGY AND DESIGN

In order to overcome these obstacles, we present in this study an integrated deep learning method that enhances breast cancer diagnosis by fusing self-attention processes with dilated CNNs and Transformers [16].

A. Convolutional Neural Networks (CNNs): An Overview

The deep learning architecture known as Convolutional Neural Networks (CNNs) was created especially for image processing. These models have transformed image-based tasks by attaining remarkable results in classification, segmentation, and object recognition [15].

B. CNN Structure and Function:

CNNs are composed of multiple interconnected layers, each of which performs a specific task in the analysis of inco ming data:

- Convolutional Layers: By using different filters on the input image, these layers extract important elements like edges and textures.
- Pooling Layers: By reducing the feature maps' size, pooling operations allow to save important information while lowering computational costs.
- Activation Functions: Non-linear activation functions such as ReLU (Rectified Linear Unit) allow the model to identify complex patterns by adding non-linearity.
- Fully Connected Layers: These layers create a final classification output by combining the features that were extracted.
- Output Layer: A CNN's last layer is usually a softmax or sigmoid function that offers a probability distribution across all potential classifications [14] (e.g., benign vs. malignant).

C. Strengths of CNNs in Medical Image Analysis:

Where: S [i,j]: Output feature map at position (i,j) , I[i+m,j+n]: Input patch of size F×F , K[m,n]: Kernel weights, b: Bias term, F: Filter size

- Output size (for a single dimension):

$N + 2P - F$

Automatic Feature Learning Unlike traditional machine learning algorithms, CNNs automatically extract relevant feature s from raw pixel data that need manually constructed features.

$S =$

$+ 1 \quad - - - (2)$

O

In medical imaging, when manual feature extraction is difficult due to data complexity, this is especially helpful.

Spatial Hierarchy: Images with spatial hierarchies are well captured by CNNs. CNNs can learn high-level characteristics (like tumors and masses) in deeper layers and low-level features (like edges or textures) in the first layers by stacking numerous convolutional layers.

Reduced Need for Pre-processing: CNNs lessen the need for

Time-consuming pre-processing procedures.

Manual processes such as edge recognition noise reduction, and contrast enhancement are necessary for traditional image analysis techniques. Where: N: Input size, P: Padding, F: Filter size, S: Stride

Pooling Layer is used to down-sample the spatial dimensions of the feature map.

- Operation:
- Max pooling:

[m,, n] = maximum ([m: m + P, n: n + P]) $---$ (3)

- Average pooling:

1

[m, n] =

$$\frac{}{P2}$$

$\sum$ [m + i, n + j]

m,n

$---$ (4)

D. *Problem Statement*Output size:

$N-P$

$S=$

O

$$\frac{}{}$$

+ 1 $---$ (5)

Clinical data and medical imaging are frequently treated as distinct things in traditional breast cancer screening methods, which results in lost possibilities for a more thorough understanding of the patient's situation. It is necessary to develop new methods that can successfully integrate imaging and clinical data in order to provide predictions that are more accurate. Convolutional neural networks (CNNs), in particular, are currently the most successful deep learning models for processing medical pictures. However, they usually have trouble integrating different input modalities, including structured clinical information [16]. CNNs are good at extracting local spatial characteristics from pictures. Since transformer models can handle many forms of input data in parallel and capture long-range dependencies, they present a viable solution and have received a lot of interest in the field of natural language processing (NLP). Transformers are ideally adapted to integrate diverse data, including clinical variables and medical images, because of their self-attention mechanism, which enables them to assess the significance of various input aspects.

E. Equations

In order to extract features from input data, Convolutional Neural Networks (CNNs) use layers that execute mathematical operations [18].

Convolutional filters (kernels) are employed in the Convolutional Layer to extract features from input.

- Operation:

$F-1$

[m, n] = ∑ [m + i, n + j]. [i, j] + b − − − (1)

i=0,=0

Where P: Pool size.

Fully Connected Layer is used to combine features into a vector for classification or regression tasks.

- Operation:

$z = W.x + b$ − − − (6)

Where: z: Output vector, W: Weight matrix, x: Input vector (flattened feature map, b: Bias term

Activation Function is used to introduce non-linearity.

- Common functions:
  - ReLU (Rectified Linear Unit): f(y) = max(0, y)
  - Sigmoid:

1

(y) = 1 + e−y − − − (7)

SELF-ATTENTION MECHANISM

The Self-Attention Mechanism is a core concept in transformer architectures [19]. It allows a model to focus on relevant parts of the input sequence when making predictions or generating outputs.

Step 1: Input Representation

Let the input sequence be represented as a matrix Y ∈ Rn×d, Where: n: Number of tokens in the sequence, d: Embedding dimension of each token.

Step 2: Compute Query, Key, and Value Matrices

Three learnable weight matrices are used to transform the input into Query, Key, and Value **vectors:**

$Q=YW^Q$, $K=YW^K$, $V=YW^V$

Where: W^Q, W^K, $W^V \in R^{d\times dk}$: Weight matrices for queries, keys, and values., Q, K, V, $\in R^{n\times dk}$: Resulting query, key, and value matrices., dk: Dimensionality of the query and key vectors.

Step 3: Scaled Dot-Product Attention

1. Score Calculation: The dot product of Q and K^T are used to compute attention score:

Score (m,n) = Qm· Kn = QK^T

This gives a matrix $S\in R^{n\times n}$, where each element represents the similarity between two tokens.

2. Scaling: To avoid large values causing instability in gradients, the scores are scaled:

$$O' = \left(\frac{QK^T}{\sqrt{d_k}}\right) \quad ---(8)$$

3. Softmax Normalization: The scores are normalized across each row to form probabilities:

$$Attention = softmax(O') = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \quad ---(9)$$

The result is a matrix $A \in R^{n\times n}$, where each row sums to 1.

4. **Weighted Sum of Values: The Value matrix (V) is multiplied by the Attention matrix (A) to produce the output:**

S = A·V where $S \in R^{n\times dk}$ is the self-attention output. Summary of Self-Attention

$$S = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right).V \quad ---(10)$$

$\sqrt{d_k}$

CNN WITH SELF-ATTENTION

1. CNN Part: Convolutional layers using features extracts local features from the input.

Output: A feature map $F \in R^{a \times b \times c}$

Where: a, b: Height and width of the feature map, c: Number of channels.

2. Self-Attention Part: Applies attention to the feature map to capture global relationships across spatial dimensions.

Step-by-Step Explanation

Step 1: CNN Feature Extraction

A standard CNN layer applies convolution, activation, and optionally pooling:

F=CNN(X) Where: $X \in R^{A \times B \times Cin}$: Input image or feature map. $F \in R^{a \times b \times c:}$ Output feature map.

Step 2: Flatten Feature Map for Attention

To apply attention, flatten the spatial dimensions (a×b) into a single sequence of size n = a×b:

Fflat = reshape (F) $\in R^{n \times c}$

Step 3: Compute Query, Key, and Value Matrices

Generate Query, Key, and Value matrices from Fflat using learned weights:

Q = FflatWQ, K = FflatWK, V = FflatWV

Where: W^Q, W^K, $W^V \in R^{c \times dk:}$ Weight matrices, Q, $K,V \in R^{n \times dk}$

Step 4: Compute Scaled Dot-Product Attention

1. Compute similarity scores: $S = QK^T \in R^{n \times n}$

Here, S [m,n] measures the similarity between the i-th and j-th spatial locations.

2. Scale the scores: $S' = S/\sqrt{dk}$
3. Normalize with softmax: $A = \text{softmax}(S') \in R^{n \times n}$

A[i,j] represents the attention weight from the i-th location to the j-th.

4. Weighted sum of values: $S_{att} = A \cdot V \in R^{n \times dk}$

Step 5: Reshape and Merge with CNN Features

Reshape Satt back to the original feature map dimensions (h × w × dk):

Optionally, combine the self-attention output with the original feature map (residual connection):

Fenhanced = F + Oatt_reshaped Final Output

The output after self-attention-enhanced CNN is: Fenhanced $\in$ R h×w×c

This output can be passed to additional CNN layers, fully connected layers, or classification heads, depending on the task.

4. PROPOSED METHODOLOGY

According to this study, the saliency region in mammography pictures is significantly recognized by the self-attention mechanism [23]. Self-attention improves model performance by enabling the model to concentrate on the most pertinent portions of the input data, which helps it identify complex connections and patterns in the image. The model can more accurately distinguish between cancerous and non-cancerous regions because to this focus [20]. This leads to better detection of abnormalities in mammography pictures and more accurate forecasts. Figure 1 shows CNN used to detect the malignant and normal images.

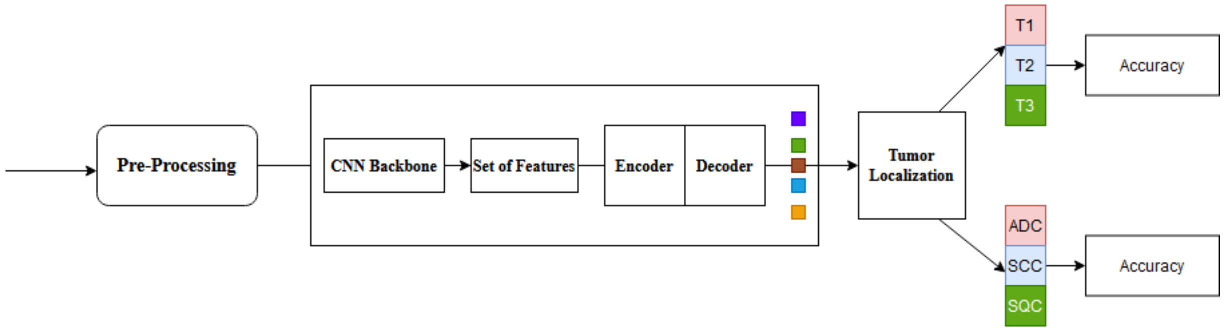

Figure 1: CNN used Cancer image detection

Additionally, self-attention helps the model understand intricate correlations between various features, which improves performance overall. Self-attention also contributes to the model's increased speed and efficiency by lowering its computational complexity [21]. Furthermore, the model can learn even more representational information from the diseased region since the dilated convolution and the self-attention process enhance the effective receptive field. The suggested model, the Self-Attention Augmented CNN Model, is displayed in Figure 2.

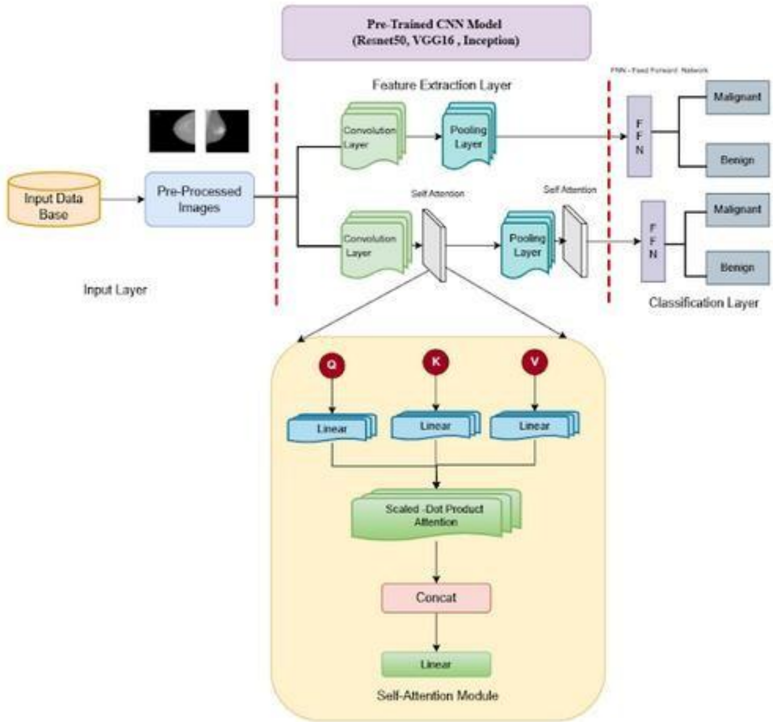

Figure 2. Proposed Model- Self-Attention Dilated CNN Model

The input layer contains image and clinical data. Both medical images and clinical data are normalized to bring them into a comparable scale. Extract features from the medical images [22]. Each convolution layer uses dilated convolutions to capture features at different scales by convolution layer. In activation function typically RELU, applied to introduce non-linearity. Pooling layers are used to reduce the dimension for ease of analysis. The output gives a set of feature maps that represent important visual cues from the image, such as potential tumor regions.

Self- attention plays a major role to detect malignant images and it refines the features extracted from the image by focusing on important regions and learning long-range dependencies. This step ensures that distant spatial features within the image, such as small tumors or distant lymph nodes, are correctly identified as relevant. Both the image features and the clinical data embeddings are passed into the Transformer block. Positional encodings are added to maintain spatial information for the image features in the Transformer. A series of dense layers with activation functions applied to refine the feature representations further are known as feed forward layers. Final stage is called normalization, here layer normalization is applied at multiple points to stabilize training and ensure consistent gradients. After the Transformer has combined the clinical and image data, the output feature vectors are passed through fully connected layers to perform high-level decision-making.

Model	Accuracy	Precision	Recall	F1-Score
CNN	0.9565	0.9156	0.8956	0.899889
CNN with Self-Attention	0.9654	0.9245	0.8979	0.904751

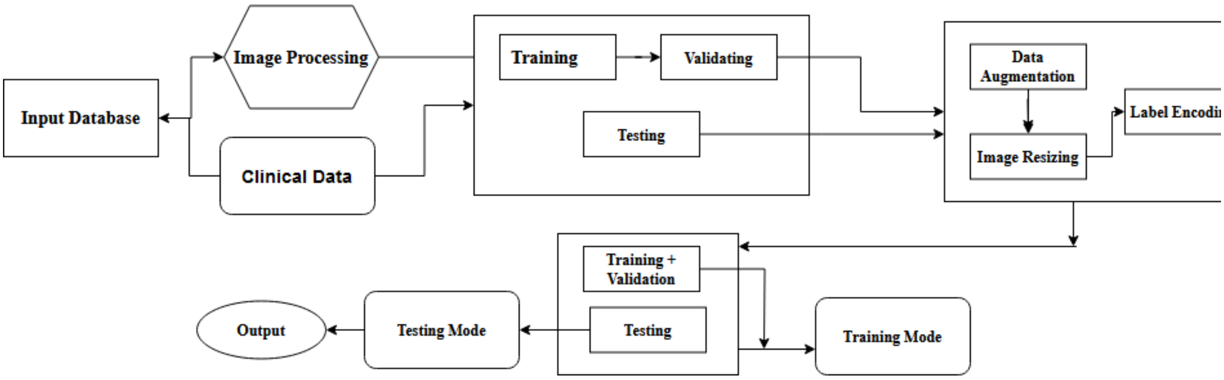

Figure 3: Proposed Model - Self-Attention Dilated CNN Model with Transformer

Self-attention mechanisms employ dilated convolutions, often referred to as atrous convolutions, to improve picture categorization tasks. Expanding the receptive field by the use of a dilated convolution with a 3 x 3 filer in the self-attention module enables the model to learn additional features from the surrounding area of the cancerous spot. Figure 4 depicts the Dilated Convolution.

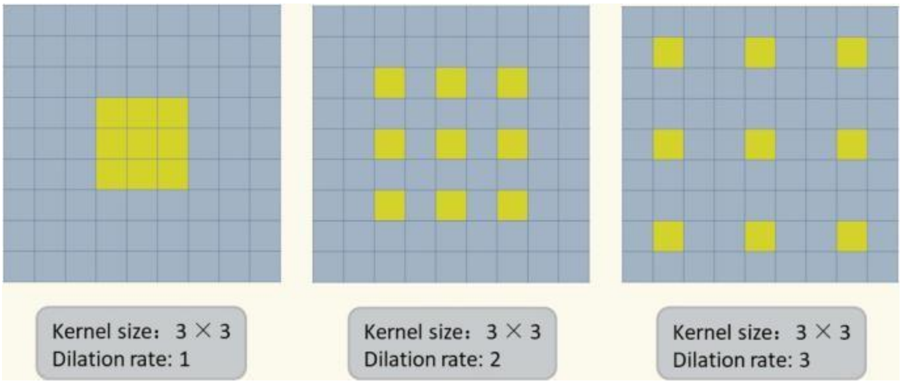

Figure 4: Dilated convolution

A. Results and Discussions

The suggested approach has been tested using the MIAS Mammography Breast Cancer dataset. 1200 benign and 1777 malignant pictures are included in the dataset. The ratio of 70:30 is used for model training. The data imbalance is addressed by increasing the quantity of training data using several picture augmentation techniques, including flipping, rotating, and splitting. For four-fold cross-validation, 200 randomly chosen photos from each category were used to assess the efficacy and resilience of the suggested model. The model is trained using the remaining pictures in each category. Table 1 lists the type of image and the quantity of photos utilized for training and testing.

Image Type	Number of Images
Malignant	1777
Benign	1200
Total	2977

Table 1: Image Type: Malignant & Benign (Testing & Training)

On a desktop computer equipped with an NVIDIA GeForce GTX 1080 Ti, 16 GB of RAM, and an Intel Core i5 or 8th generation processor, the suggested model is implemented using Tensor flow. We train the model using the ADAM optimizer, mini of batch size 16, initialized learning rate 0.01 and 50 epochs. Utilizing the cross-entropy loss function, the suggested model is optimized.

The performance metrics of the CNN model with &without attention module, CNN with transformer model is displayed in Table 2

CNN with Self-attention & Transformer	0.9723	0.9356	0.9154	0.919891

Table 2: Comparison table of CNN, CNN with Self –attention and CNN with self-attention and transformer

Figure 5 shows the Confusion Matrix of CNN with and without Self- Attention Versions.

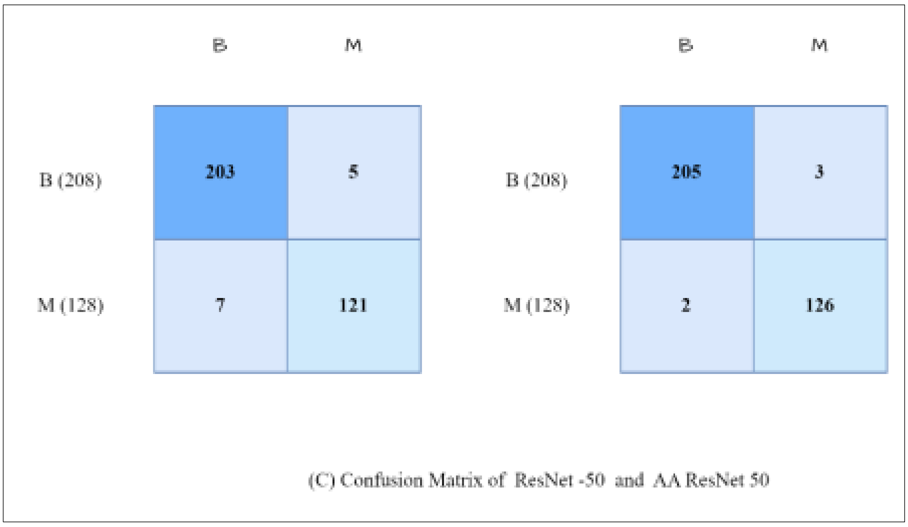

Figure 5: Confusion Matrix of CNN with and without Self- Attention Versions

In figure 6 shows the Quantitative Performance of Pretrained CNN model (Accuracy Vs Epochs for different Modes)

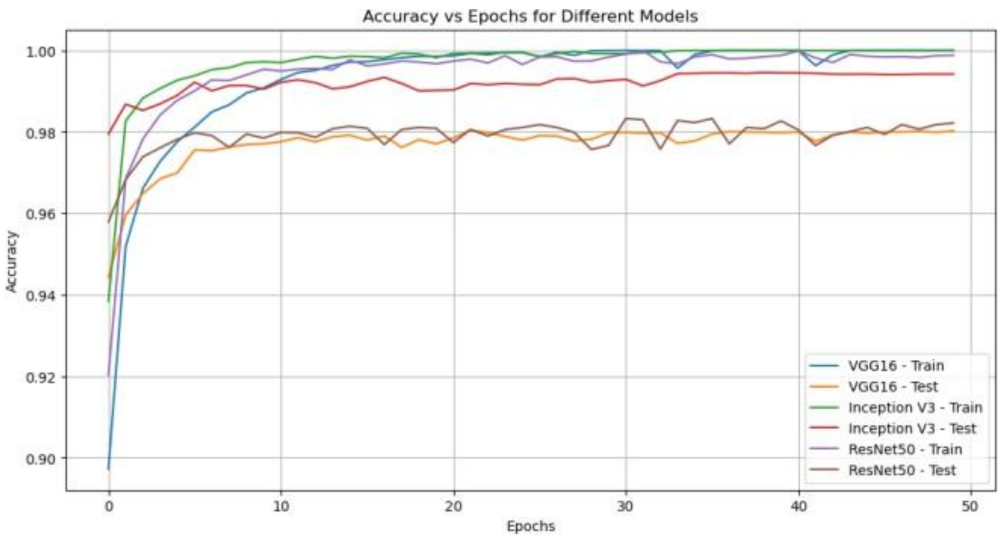

Figure 6: Quantitative Performance of Pretrained CNN model without AA (Accuracy Vs Epochs for different Modes)

The overall performance analysis of all three methods as shown in figure 7.

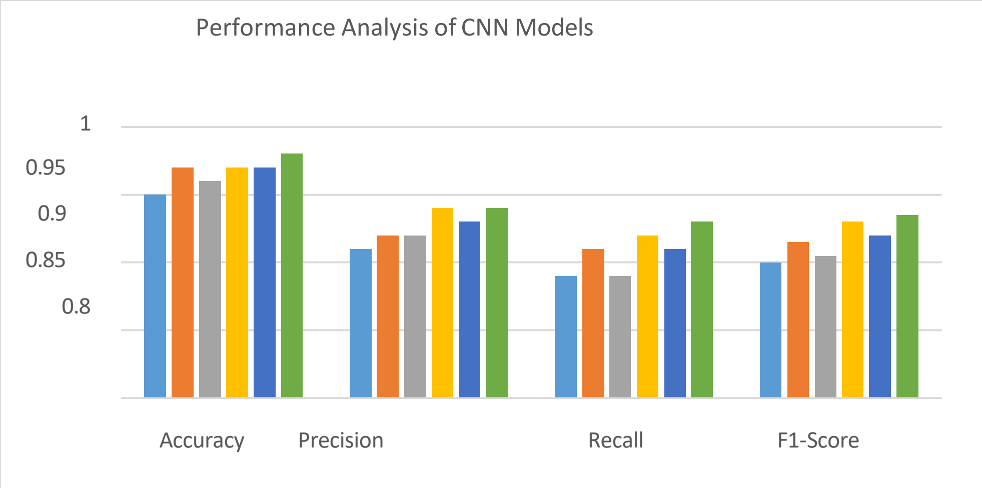

Figure 7. Performance Analysis of CNN Models with & without Self-attention Augmentation

B. Conclusion

A self-attention-augmented CNN model with a transformer in a clinical data set is proposed in this work for breast prediction in mammography pictures. The self-attention mechanism improves feature extraction and classification accuracy by enabling the model to concentrate on pertinent areas of the image. As a result, it may perform better than more conventional CNN models like VGG-16, inception v3, and RESNET 50, which might have trouble identifying long-range dependencies.

Furthermore, self-attention can prioritize important features, lowering the computational cost and improving the model's efficiency. By training the model with unlabeled data, it intends to improve the model in the future through self-supervised learning. Using clinical data, transformer produced precise results.

REFERENCES

[1] M. I. Nazir, A. Akter, M. A. H. Wadud, and M. A. Uddin, "Utilizing customized CNN for brain tumor prediction with explainable AI," Heliyon, vol. 10, no. 10, p. e38997, Oct. 2024. https://doi.org/10.1016/j.heliyon.2024.e38997.

[2] M. Salehi, J. Razmara, S. Lotfi, and F. Mahan, "A one-dimensional probabilistic convolutional neural network for prediction of breast cancer survivability," The Computer Journal, vol. 65, no. 10, pp. 2641–2653, Oct. 2022. https://doi.org/10.1093/comjnl/bxab096.

[3] Z. Jafari and E. Karami, "Breast cancer detection in mammography images: A CNN-based approach with feature selection," Information, vol. 14, no. 7, p. 410, Jul. 2023. https://doi.org/10.3390/info14070410.

[4] N. Zheng, Z. Yao, S. Tao, A. Almadhor, M. S. Alqahtani, R. M. Ghoniem, H. Zhao, and S. Li, "Application of nanotechnology in breast cancer screening under obstetrics and gynecology through the use of CNN and ANFIS," Environmental Research, vol. 234, p. 116414, Oct. 2023. https://doi.org/10.1016/j.envres.2023.116414.

[5] S. Majumdar, P. Pramanik, and R. Sarkar, "Gamma function based ensemble of CNN models for breast cancer detection in histopathology images," Expert Systems with Applications, vol. 213, Part B, p. 119022, Mar. 2023. https://doi.org/10.1016/j.eswa.2022.119022.

[6] N. Faruqui, M. A. Yousuf, F. A. Kateb, M. A. Hamid, and M. M. Monowar, "Healthcare-As-A-Service (HAAS): CNN-Based Cloud Computing Model for Ubiquitous Access to Lung Cancer Diagnosis," Heliyon, vol. 9, no. 11, p. e21520, Oct. 2023. https://doi.org/10.1016/j.heliyon.2023.e21520.

[7] J. G. Melekoodappattu, A. S. Dhas, B. K. Kandathil, et al., "Breast cancer detection in mammogram: Combining modified CNN and texture feature-based approach," Journal of Ambient Intelligence and Humanized Computing, vol. 14, pp. 11397–11406, 2023. https://doi.org/10.1007/s12652-022-03713-3.

[8] S. R. S. Chakravarthy, N. Bharanidharan, and H. Rajaguru, "Multi-deep CNN based experimentations for early diagnosis of breast cancer," IETE Journal of Research, vol. 69, no. 10, pp. 7326–7341, 2022. https://doi.org/10.1080/03772063.2022.2028584.

[9] J. Zuluaga-Gomez, Z. Al Masry, K. Benaggoune, S. Meraghni, and N. Zerhouni, "A CNN-based methodology for breast cancer diagnosis using thermal images," Computer Methods in Biomechanics and Biomedical

Engineering: Imaging & Visualization, vol. 9, no. 2, pp. 131–145, 2021. https://doi.org/10.1080/21681163.2020.1824685.

[10] L. Bouzar-Benlabiod, K. Harrar, L. Yamoun, M. Y. Khodja, and M. A. Akhloufi, "A novel breast cancer detection architecture based on a CNN-CBR system for mammogram classification,"Computers in Biology and Medicine, vol. 163, p.107133,2023. https://doi.org/10.1016/j.compbiomed.2023.107133.

[11] A. Sahu, P. K. Das, and S. Meher, "High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets," Biomedical Signal Processing and Control, vol. 80, Part 1, p. 104292, 2023. https://doi.org/10.1016/j.bspc.2022.104292.

[12] J. N. C. Raimundo, J. P. P. Fontes, L. G. M. Magalhães, and M. A. Guevara Lopez, "An innovative Faster R-CNN-based framework for breast cancer detection in MRI," Journal of Imaging, vol. 9, no. 9, p. 169, 2023. https://doi.org/10.3390/jimaging9090169.

[13] S. Costanzo, A. Flores, and G. Buonanno, "Fast and accurate CNN-based machine learning approach for microwave medical imaging in cancer detection," IEEE Access, vol. 11, pp. 66063-66075, Jun. 30, 2023, doi: 10.1109/ACCESS.2023.3291076.

[14] S. Chakravarthy, N. Bharanidharan, S. B. Khan, V. V. Kumar, T. R. Mahesh, A. Almusharraf, et al., "Multi-class breast cancer classification using CNN features hybridization," Int. J. Comput. Intell. Syst., vol. 17, Art. 191, 2024, doi: 10.1007/s11047-024-09985-x.

[15] S. U. R. Khan, M. Zhao, S. Asif, X. Chen, and Y. Zhu, "GLNET: Global–local CNN's-based informed model for detection of breast cancer categories from histopathological slides," J. Biomed. Eng., vol. 80, pp. 7316–7348, 2024, doi: 10.1007/s11047-024-10021-y.

[16] M. A. Wahed, "Evaluating AI and machine learning models in breast cancer detection: A review of convolutional neural networks (CNN) and global research trends," Front. Genet., vol. 13, sec. Computational Genomics, 2024, Art. 1097207, doi: 10.3389/fgene.2022.1097207.

[17] H. S. Das, A. Das, A. Neog, S. Mallik, K. Bora, and Z. Zhao, "Breast cancer detection: Shallow convolutional neural network against deep convolutional neural networks based approach," Front. Genet., vol. 13, no. 2022, Art. no. 1097207, Jan. 2023. https://doi.org/10.3389/fgene.2022.1097207

[18] V. Sureshkumar, R. S. N. Prasad, S. Balasubramaniam, D. Jagannathan, J. Daniel, and S. Dhanasekaran, "Breast cancer detection and analytics using hybrid CNN and extreme learning machine," J. Pers. Med., vol. 14, no. 8, Art. no. 792, Jul. 2024. https://doi.org/10.3390/jpm14080792

[19] D. Al Kafaf, N. N. Thamir, and A. Al-Azzawi, "Breast cancer prediction: A CNN approach," Multiscience, vol. 6, no. 9, pp. e2024156, 2024. [Online]. Available: https://doi.org/10.31893/multiscience.2024156

[20] A. Pati, M. Parhi, B. K. Pattanayak, B. Sahu, and S. Khasim, "CanDiag: Fog empowered transfer deep learning based approach for cancer diagnosis," Designs, vol. 7, no. 3, pp. 57, 2023. https://doi.org/10.3390/designs7030057

[21] L. Sun, H. Tian, H. Ge, J. Tian, Y. Lin, C. Liang, T. Liu, and Y. Zhao, "Cross-attention multi-branch CNN using DCE-MRI to classify breast cancer molecular subtypes," Front. Oncol., vol. 13, art. 1107850, Mar. 7, 2023. doi: 10.3389/fonc.2023.1107850. https://doi.org/10.3389/fonc.2023.1107850

[22] A. Alshehri and D. AlSaeed, "Breast cancer detection in thermography using convolutional neural networks (CNNs) with deep attention mechanisms," Appl. Sci., vol. 12, no. 24, art. 12922, Dec. 15, 2022. doi: 10.3390/app122412922. https://doi.org/10.3390/app122412922

[23] M. J. Umer, M. Sharif, and S.-H. Wang, "Breast cancer classification and segmentation framework using multiscale CNN and U-shaped dual decoded attention network," Expert Syst., vol. 40, no. 1, p. e13192, Nov. 14, 2022. doi: 10.1111/exsy.13192. https://doi.org/10.1111/exsy.13192