

Breast Cancer Detection In Histopathological Images Using An Ensemble Of VGG-19 And Alexnet

G. Divya Zion¹, B.K. Tripathy^{2*}

¹ School of Computer Science and Engineering, VIT, Vellore-632014, TN, India.

² School of Computer Science Engineering & Information Systems, VIT, Vellore-632014, TN, India

* Corresponding authors: B.K. Tripathy² (tripathybk@vit.ac.in).

Abstract

Breast cancer as one of primary causes of death among women globally, requires accurate early diagnosis to significantly improve patient survival outcomes. Even though histopathology is the most reliable diagnostic method, manual interpretation poses limitations such as being time-consuming, subjective, and highly dependent on the experience of the examiner. With the advancement in deep learning the convolutional neural networks (CNNs) shows significance in automating cancer detection. An ensemble deep learning architecture that integrates VGG-19 and AlexNet has been proposed for breast cancer classification using histopathological images. After performing independent fine-tuning by both models using transfer learning, the models' prediction scores are merged using a weighted ensemble method where both models make predictions separately, and then their outputs which are probability scores are combined after the individual models have finished prediction, enabling the system to utilize complementary strengths from each architecture. The proposed ensemble system demonstrates improved performance over individual models, achieving an overall accuracy of 98.88% on the histopathological breast cancer image dataset. Through heterogeneous CNN models the robustness and classification performance has improved. This implies that ensemble deep learning models may aid pathologists in improving the dependability of computer-assisted breast cancer diagnosis.

Keywords: CNN, Deep Learning, Transfer Learning, VGG-19, AlexNet.

Introduction

Breast cancer occurs when abnormal cells in breast tissue grow uncontrollably, forming a tumor that may spread to other parts of the body. Detecting it early is critical because survival rates are much higher when the cancer is found before it spreads. There are several causes leading to breast cancer, it arises from a combination of genetic, biological, environmental, and lifestyle factors. These factors can trigger the cancer cells to behave abnormally. In the meantime, if patient notices symptoms of cancer being progressed the imaging techniques like, Mammograms, Ultrasound, Histopathological images (biopsies) may help detect cancer earlier. But the manual procedure in analyzing the imaging techniques is time-consuming and prone to errors. Breast Cancer is a prevalent and life-threatening disease that predominantly affects women in middle age and is recognized as the second most common cause of cancer-related mortality worldwide, following lung cancer. Characterized by uncontrolled and abnormal cellular proliferation, BC represents a major global health concern, with the World Health Organization (WHO) estimating that women face approximately an 8% lifetime risk of developing the disease. Distinguishing between benign and malignant tumor that often differ in size, shape, and spatial distribution which remains a significant challenge in the accurate detection and localization of cancerous regions in breast imaging. Deep Learning models analyze medical images and automatically learn features that help classify tissue either Benign (non-cancerous) or Malignant (cancerous).

The proposed approach aims to address these gaps by leveraging advanced Deep Learning techniques tailored for Breast Cancer Detection. Deep learning models such as the VGG-19 and AlexNet ensemble play a crucial role in modern diagnosis by automatically extracting microscopic (histopathological images) tumor-related features from breast tissue images, improving detection speed, accuracy, and consistency thereby detecting earlier intervention and improved patient survival. VGG-19 (Visual Geometry Group-19) is a deep convolutional neural network architecture. It is widely used for image classification, feature extraction in medical imaging tasks like breast cancer diagnosis. It uses 3×3 convolution filters instead of using 5×5 or 7×7 larger filters. This increases depth while reducing parameters, making feature learning more efficient. VGG-19 works by gradually extracting low-level to high-level visual features using stacked 3×3 convolutions and max-pooling, and fully connected layers to classify images. AlexNet is a Convolutional Neural Network (CNN) that processes images to automatically extract features. AlexNet performs well on breast cancer histopathology by considering large filters to detect tissue-level abnormalities and captures global texture patterns and cell distribution and feature learning is done automatically.

The rest of the paper is structured as follows: section 2 presents Literature Review, section 3 describes Methodology, section 4 presents Comparative analysis and the Results and finally concludes in section 5.

Literature Review

An ensemble deep learning approach is applied for classification of breast cancer on histopathology images, [1] four different models based on pre-trained VGG16 and VGG19 architectures are applied. The ensemble of fine-tuned VGG-16 and VGG-19 models offered sensitivity of 97.73% and overall accuracy of 95.29%. An advanced deep learning framework for histopathological image classification is done by integrating AlexNet and Gated Recurrent Unit (GRU) networks [2], optimized using the Hippopotamus Optimization Algorithm (HOA). The results demonstrate the efficacy of integrating deep learning with bio-inspired optimization techniques in breast cancer detection. This research offers a robust and computationally efficient framework for improving early diagnosis and clinical decision-making, potentially enhancing patient outcomes. Efficient treatment methods for accurately diagnosing the kind and classifying the subtype of breast cancer using a unique combination of hand-engineered features for color, shape, and texture extraction in parallel to three different deep neural networks. A hybrid method [3] that combines both hand-engineered and deep learning-based feature extractors provides an outstanding performance for breast cancer detection and classification on different types of datasets thus verifying its robustness and effectiveness. A comprehensive approach [4] is presented utilizing ensemble deep learning techniques to analyze breast cancer histopathology images. VGG16 and ResNet50 architectures are incorporated to achieve precise classification of breast cancer histopathology images. The ensemble approach utilizing VGG16, ResNet34, and ResNet50 models to classify microscopic images. This research significantly contributes in advance breast cancer diagnosis in improved patient outcomes.

An effective detection and classification of breast cancer is done by introducing a novel method, MPa-DCAE, which uses a multi-patch-based deep convolutional auto-encoder (DCAE) framework combined with VGG19 to detect and classify breast cancer in histopathology images [5]. Ensemble Deep Convolutional Neural Network (EDCNN) model that exhibits superior accuracy compared to several transfer learning models and the Vision Transformer model. Our EDCNN model [6] integrates the strengths of the MobileNet and

Xception models to improve its performance in breast cancer detection and classification. Through advanced deep learning techniques and image segmentation, this study contributes to improving diagnosis and treatment outcomes in breast cancer.

Methodology

An ensemble VGG-19 and AlexNet framework of deep learning is considered for the classification of histopathological breast cancer images into benign and malignant categories. The ensemble model performs preprocessing, augmentation, architecture design, model training, ensemble fusion, and performance evaluation. The Breast Cancer Histopathological Image dataset is considered which is publicly available via the Kaggle medical imaging repository. The dataset consists of 277,524 histopathology images with eight classes. The binary separation of benign and malignant classes is shown in Figure 1. The benign categories include adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA); the malignant classifications include ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and papillary cancer (PC). The data split for the BreakHis [7] data is shown in Figure 1. An explanation of the BreakHis [7] dataset, including the initial partition, splitting, and sample counts for each split.

Class	Sub_Class	Magnification				Total	Nos_Patients
		40x	100x	200x	400x		
Benign	Adenosis	114	113	111	106	444	24
	Fibroadenoma	253	260	264	237	1014	
	Phyllodes_tumor	109	121	108	115	453	
	Tubular_adenoma	149	150	140	130	569	
Malignant	Ductal_carcinoma	864	903	896	788	3451	58
	Lobular_carcinoma	156	170	163	137	626	
	Mucinous_carcinoma	205	222	196	169	792	
	Papillary_carcinoma	145	142	135	138	560	
Total		1995	2081	2013	1820	7909	82
Augmented-Train							
Benign	Adenosis	1002	1002	1002	1002	4008	-
	Fibroadenoma	1002	1002	1002	1002	4008	
	Phyllodes_tumor	1002	1002	1002	1002	4008	
	Tubular_adenoma	1002	1002	1002	1002	4008	
Malignant	Ductal_carcinoma	1002	1002	1002	1002	4008	-
	Lobular_carcinoma	1002	1002	1002	1002	4008	
	Mucinous_carcinoma	1002	1002	1002	1002	4008	
	Papillary_carcinoma	1002	1002	1002	1002	4008	
Test							
Benign	Adenosis	11	11	11	10	43	-
	Fibroadenoma	25	25	26	23	99	
	Phyllodes_tumor	10	12	10	11	43	
	Tubular_adenoma	14	14	13	12	53	
Malignant	Ductal_carcinoma	86	90	89	78	343	-
	Lobular_carcinoma	15	16	16	13	60	
	Mucinous_carcinoma	20	22	19	16	77	
	Papillary_carcinoma	14	14	13	13	54	
Validation							
Benign	Adenosis	98	98	98	98	392	-
	Fibroadenoma	98	98	98	98	392	
	Phyllodes_tumor	98	98	98	98	392	
	Tubular_adenoma	98	98	98	98	392	
Malignant	Ductal_carcinoma	98	98	98	98	392	-
	Lobular_carcinoma	98	98	98	98	392	
	Mucinous_carcinoma	98	98	98	98	392	
	Papillary_carcinoma	98	98	98	98	392	

Figure 1. An explanation of the dataset (Man et al, 2020)

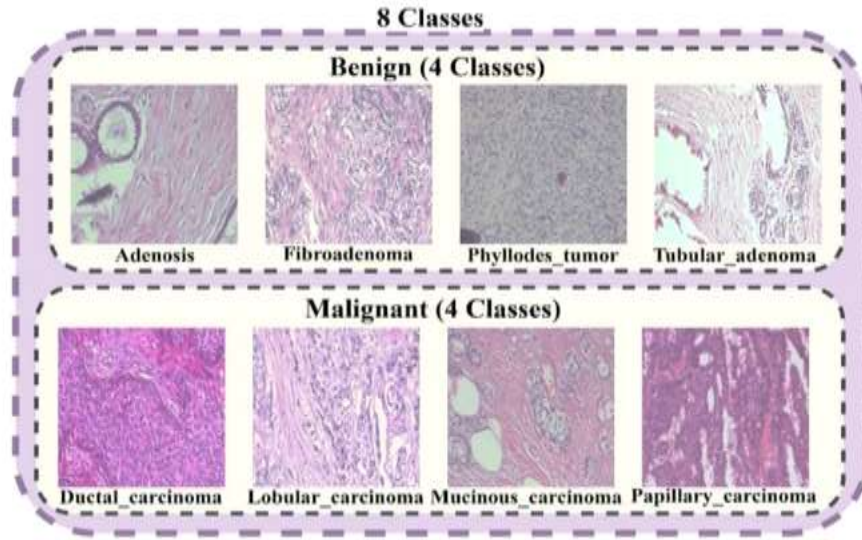


Figure 2. Dataset samples (Man et al., 2020).

The dataset samples are shown as pictorial representation [7]. In Figure 2, the inner dash lines represent the multiclass (4 classes) in two categories: Benign and malignant, whereas the outside dash lines show the multiclass (8 classes).

Step-1: Data Pre-Processing

The publicly available Breast cancer histopathological image dataset is considered. The dataset contains high-resolution histopathological images that are captured at multiple resolutions to increase variability in terms of Texture patterns, Structural details, Sharpness and clarity, Scale of visual information, representation of tissue morphology and increase in feature quality in terms of Cell size and shape, Nuclei structure, Tissue patterns, color distribution, texture details. As part of the preprocessing pipeline, every image was normalized and resized to a fixed resolution of $224 \times 224 \times 3$ channels to enable consistent input representation irrespective of the deep learning model that is considered. The images are resized to a fixed dimension of 224 pixels by 224 pixels with RGB colours. The deep learning model's architectures needs the images into fixed dimensional networks irrespective of original resolution and so they are converted. To reduce inconsistencies in images, pixel intensities are normalized by scaling them based on the dataset's mean and standard deviation, typically calculated from a large set of training images. Normalization converts pixel values to have a zero mean and unit variance, improving model convergence during training and preventing issues related to varying brightness and contrast.

Image normalization is done using;

$$I_{\text{normalized}} = \frac{I - \mu}{\sigma} \quad (1)$$

where I = input image, μ = mean and σ = standard deviation

Step-2: Data Augmentation

Data augmentation is performed to make the dataset more differentiated, and more representative so that the model becomes more accurate, reliable, and less prone to

overfitting. Histopathological images usually have few samples and uneven class distribution. Therefore, augmentation is performed to increase the sample size and improve model generalization by performing; rotation to simulate different tissue positioning and aligning the images into different angles is done, Horizontal and vertical flipping, cropping, Color contrast and brightness variation is done.

Step-3: Architecture Design: VGG-19 and AlexNet

VGG-19 Architecture

The VGG-19 consists of 19 sequential layers, instead of considering a single large convolution layer, it uses several convolution layers placed one after another i.e; 3×3 convolution kernels with increasing filter depth [8]. This enables the model to extract fine-grained texture patterns like nuclei shape irregularities, cell density variations and structural degradation in malignant tissue. To perform classification on the breast cancer histopathology task, custom classification is done:

1. Converts feature maps from the convolution layers into a smaller vector by averaging values which reduces model complexity and prevents overfitting.
2. A fully connected layer with 512 units and ReLU activation which allows the model to learn high-level decision rules from extracted features.
3. Randomly disables 50% of neurons during training, to overcome overfitting and improves generalization.
4. The final output layer produces class probabilities such as benign or malignant.
5. From the custom classification, only the top (last) convolution blocks of the pretrained model were made trainable (unfrozen). The previous layers remained frozen, where the weights are not changed.

Comparison of VGG-19 Architecture model Layers with the generated output

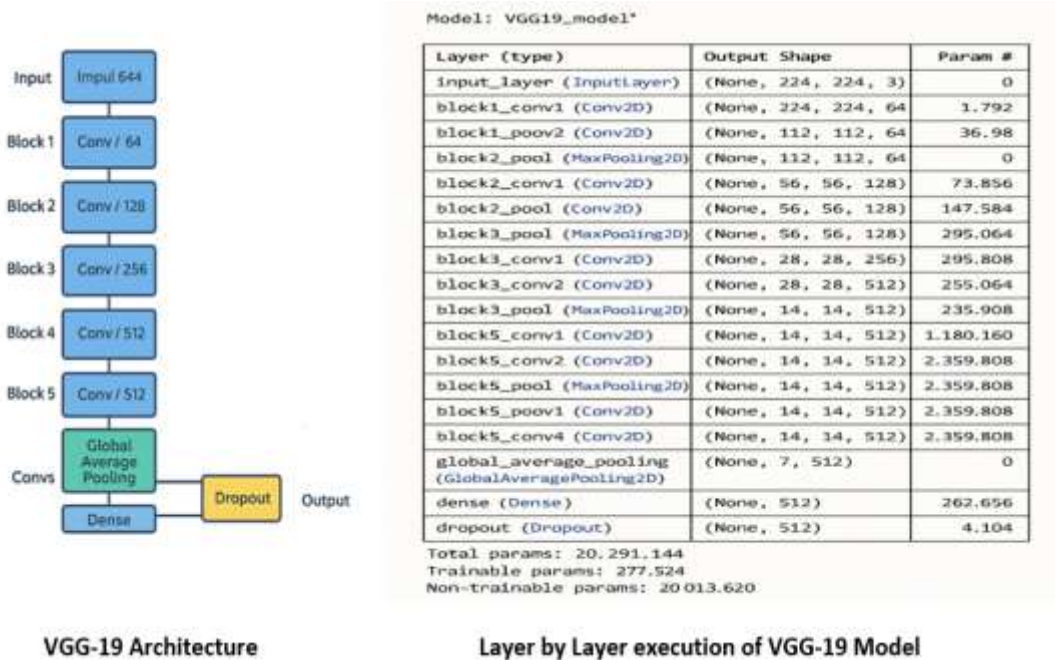


Figure 3. VGG-19 Architecture comparison with layer-by-layer execution

Figure 3, explains the execution of VGG-19 model layer by layer. The VGG19-based model accepts 224×224 RGB images as input through an input layer that defines only the image dimensions and contains no trainable parameters. The network is organized into five sequential convolutional blocks, each responsible for extracting features of increasing complexity;

Block 1 begins with 64 convolutional filters, focusing on extracting low-level visual features such as edges, gradients, and basic color transitions. A max-pooling layer reduces the spatial resolution from 224×224 to 112×112 , lowering computational cost while preserving essential information. Block 2 employs 128 filters to capture mid-level patterns, including corners, curves, and simple texture regions. Max-pooling further reduces the feature map size to 56×56 , retaining the most discriminative structures. Block 3 increases the depth to 256 filters, enabling the extraction of more complex textures and structural patterns. After convolution and pooling, the feature maps are reduced to 28×28 , allowing the model to generalize deeper spatial representations. Block 4 uses 512 filters to learn high-level semantic features, such as tissue regions and shape-based patterns relevant to breast cancer histopathology. Max-pooling compresses the representation to 14×14 , discarding weaker activations. Block 5, the deepest block, also utilizes 512 filters to extract highly abstract and domain-specific features, such as malignant cellular textures. The final pooling operation prepares the feature maps for classification.

Instead of using traditional fully connected layers, a Global Average Pooling layer converts the final feature maps into a compact 512-dimensional feature vector, significantly reducing the number of parameters and minimizing overfitting. The extracted global features are then passed through a Dense layer with 512 neurons, which learns the final discriminative representation. A Dropout layer is applied afterward to randomly deactivate neurons during training, enhancing robustness and preventing overfitting.

The VGG-19 base is frozen (non-trainable) only dense layers and dropout layer are learned from dataset. The lower convolution layers in pretrained models learn generic features such as edges, corners, and simple patterns and the higher convolution layers learn task-specific features like textures and patterns. In histopathology images, cancer patterns involve cell structures, irregular nuclei, and tissue textures.

AlexNet Architecture

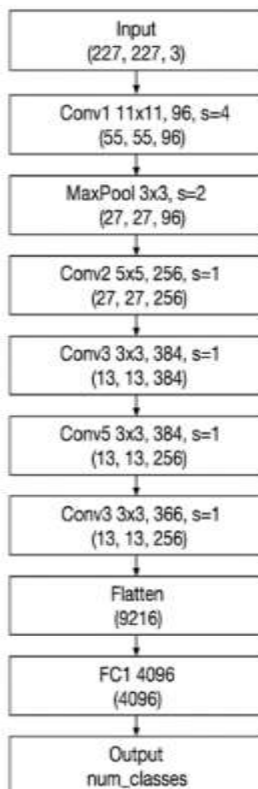
AlexNet contains 8 learnable layers (5 convolutional layers + 3 fully connected layers). It uses larger convolutional kernel of sizes 11×11 and 5×5 so that the model extracts broader and more abstract information. AlexNet classifies the breast cancer histopathology image dataset using custom classification; by converting high-dimensional 2D feature maps are converted to 1D vector format with high-level numerical feature maps that describe patterns suitable for dense layers. The Dense layers are responsible for interpreting these extracted features and determining which class the image belongs to (e.g., benign or malignant). The Rectified Linear Unit activation function checks non-linearity to detect complex cancer-specific tissue patterns [9][10]. During training the 50% of neurons gets deactivated randomly and produces output as probability of binary classification.

Comparison of AlexNet Architecture model Layers with the generated output

Figure 4, explains the layer-by-layer execution of AlexNet model. The AlexNet-based model accepts 224×224 RGB images as input through an input layer that defines the image dimensions and contains no trainable parameters. The network is composed of five

convolutional stages followed by fully connected classification layers, enabling hierarchical feature extraction from low-level visual cues to high-level semantic representations.

The initial convolutional layer applies 96 filters of size 11×11 with a stride of 4, producing feature maps of size $54 \times 54 \times 96$. This layer captures basic image characteristics, such as edges, color gradients, and simple textures. A Batch Normalization layer follows to stabilize learning and accelerate convergence. Max-pooling then reduces the spatial resolution to 26×26 , retaining the most dominant activations. The second convolutional layer increases the depth to 256 filters (5×5 kernel), allowing the network to learn mid-level patterns, including corners and texture regions. Batch normalization and max-pooling further reduce the feature maps to $12 \times 12 \times 256$, enhancing representational efficiency. The third and fourth convolutional layers employ 384 filters (3×3 kernels) to extract deeper and more complex structural features, such as object parts and detailed spatial patterns. These layers refine the feature hierarchy without further spatial reduction. The fifth convolutional layer, with 256 filters, captures the most abstract and highly discriminative features. A final max-pooling operation reduces the feature maps to $5 \times 5 \times 256$, preparing the network for classification. The classification stage consists of two large fully connected layers, each with 4096 neurons, which learn high-level feature combinations critical for recognition tasks. Dropout layers are applied after each dense layer to reduce overfitting and improve generalization. Finally, the network terminates with a Dense output layer containing 8 neurons, corresponding to the target classes, producing the final classification output.



AlexNet Architecture

Model: "AlexNet"

Layer (type)	Output Shape	Param #
alex_input (InputLayer)	(None, 224, 224, 3)	0
conv1 (Conv2D)	(None, 54, 54, 96)	34,944
batch_normalization (BatchNormalization)	(None, 54, 54, 96)	384
pool1 (MaxPooling2D)	(None, 26, 26, 96)	0
conv2 (Conv2D)	(None, 26, 26, 256)	614,856
batch_normalization_1 (BatchNormalization)	(None, 26, 26, 256)	1,824
pool2 (MaxPooling2D)	(None, 12, 12, 256)	0
conv3 (Conv2D)	(None, 12, 12, 384)	585,120
conv4 (Conv2D)	(None, 12, 12, 384)	1,327,488
conv5 (Conv2D)	(None, 12, 12, 256)	884,992
pool5 (MaxPooling2D)	(None, 5, 5, 256)	0
flatten (Flatten)	(None, 6400)	0
fc1 (Dense)	(None, 4096)	26,218,496
dropout_1 (Dropout)	(None, 4096)	0
fc2 (Dense)	(None, 4096)	16,781,312
dropout_2 (Dropout)	(None, 4096)	0
alex_output (Dense)	(None, 8)	32,776

Total params: 46,781,192 (178.46 MB)

Trainable params: 46,788,488 (178.45 MB)

Non-trainable params: 704 (2.75 KB)

Layer by Layer execution of AlexNet model

Figure 4. AlexNet Architecture comparison with layer-by-layer execution**Step-4: Training a Model**

The training a model fine-tunes the VGG-19 and AlexNet architectures on the histopathological image dataset using a transfer learning strategy.

Both models were initialized with pretrained ImageNet weights, the earlier convolutional layers remained frozen with low-level features (edges, contours, and textures), while selected higher convolution blocks were unfrozen to learn histopathology-specific patterns such as abnormal nuclei morphology, glandular disorganization, and color-shift patterns. The newly added dense layers from both models were initialized with random Gaussian weights and trained from the beginning as they were specific to the binary breast cancer classification task.

Binary Cross-Entropy (BCE) Loss Function

Binary Cross-Entropy (BCE) Loss is used to train the model to make accurate probability predictions and pushes the network to produce higher probabilities for true cancer images and lower probabilities for benign ones. The Binary Cross-Entropy (BCE) loss function is used to measure the model's predicted output matches the true class label in a binary classification task, such as deciding benign and malignant breast tissue. As the output of the model is a probability value between 0 and 1, BCE evaluates how close the predicted probability is to the actual truth label.

$$L = -[y \log(p) + (1 - y) \log(1 - p)] \quad (2)$$

where, y = true label (benign = 0, malignant = 1), p = predicted probability from the model, L = Loss value (error measure).

If the true label is malignant ($y = 1$):

$$L = -\log(p) \quad (3)$$

If the model predicts $p = 1$, loss becomes 0 (perfect prediction).

If the model predicts $p = 0.1$, loss becomes high, penalizing the model.

If the true label is benign ($y = 0$):

$$L = -\log(1 - p) \quad (4)$$

A prediction near 0 gives low loss and a prediction near 1 gives a large penalty.

In breast cancer detection using VGG-19 and AlexNet ensemble fusion, BCE Loss trains the ensemble fusion model to output accurate probabilities for malignant vs benign cases, penalizes high-confidence mistakes, stabilizes learning, and improves diagnostic sensitivity.

Step-5: Proposed Model: Ensemble Fusion

This Figure 5. explains the Ensemble Model Architecture used for breast cancer classification from histopathology images by combining VGG19 and AlexNet. The workflow is explained step-by-step as follows:

1. Input and Data Preparation

- Histopathology Images are taken as input.
- Preprocessing includes resizing, normalization, and stain normalization to reduce color and intensity variations common in medical images.
- Data Augmentation (applied only during training) improves generalization by introducing rotations, flips, and other variations.

2. Parallel CNN Feature Extraction

The architecture consists of two parallel CNN models:

a) VGG-19 Model

- Extracts deep hierarchical features using a deeper architecture.
- Outputs feature maps that are passed through a Dense and Softmax layer, producing class probability scores p_A .

b) AlexNet Model

- Extracts complementary features using a comparatively shallower but wider network.
- Also ends with a Dense and Softmax layer, producing probability scores p_B .

3. Ensemble Strategy

This ensemble fusion is performed using Late fusion, Decision fusion and Stacking. The ensemble layer collects the probability output from the Dense and Softmax layer which is the final classification stage of a deep learning model, where extracted features are converted into class probabilities.

VGG-19

AlexNet

output: $p_A = [p_{A1}, p_{A2}, \dots, p_{Ak}]$

output: $p_B = [p_{B1}, p_{B2}, \dots, p_{Bk}]$

where k is the number of classes (e.g., benign and malignant).

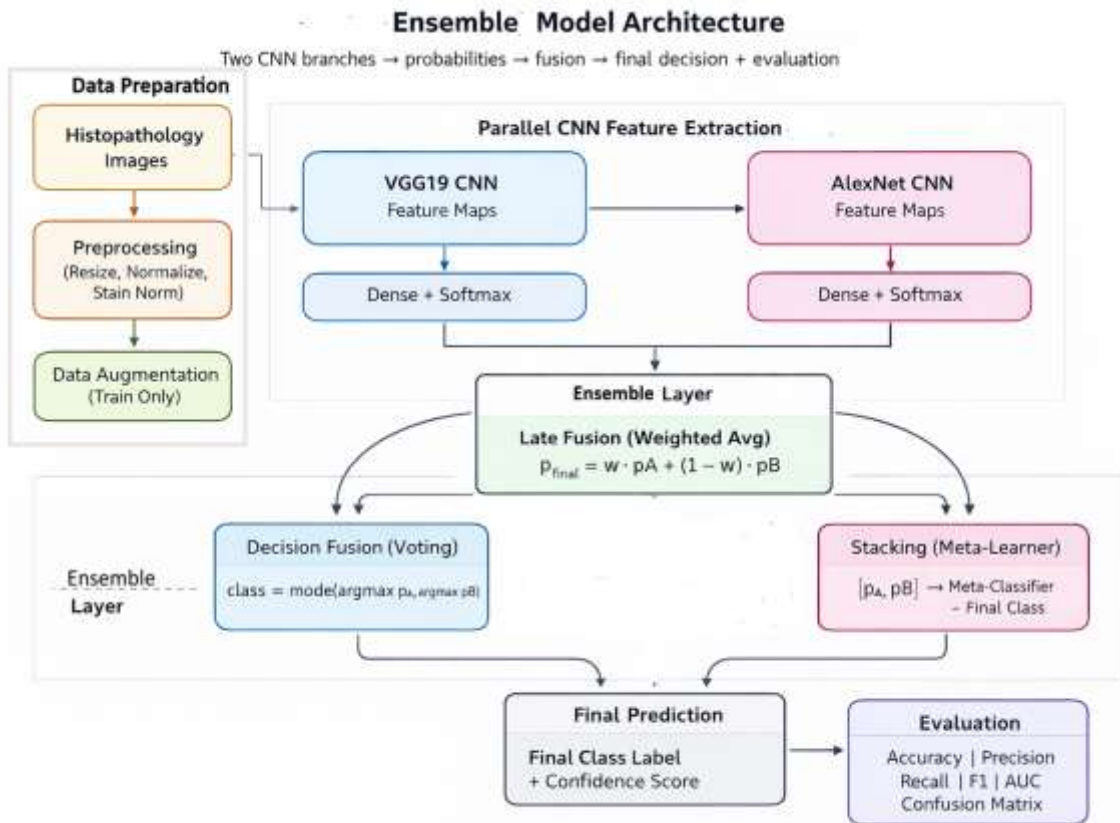


Figure 5. Proposed Architecture

The proposed ensemble techniques support three strategies shown in Figure 5;

1. Late Fusion (Weighted Average Fusion): After performing the feature extraction and classification by each model the output probabilities of multiple models are combined and the final probability is computed as:

$$p_{\text{final}} = w \cdot p_A + (1 - w) \cdot p_B \quad (5)$$

Where, w is a weighting factor ($0 \leq w \leq 1$).

when the value of w is high, VGG-19 probability is considered and when value of w is low the probability of AlexNet model is considered. The weights can be fixed or can be optimized based on validation performance.

2. Decision Fusion: The final class labels predicted by multiple models rather than their probabilities are combined. Both VGG-19 and AlexNet predicts a class label; VGG19 → Benign, AlexNet → Malignant

and the final decision is made based on majority agreement.

$$\text{class} = \text{mode}(\arg \max(p_A), \arg \max(p_B)) \quad (6)$$

3. Stacking: A meta-classifier is trained to optimally combine the outputs of multiple base models. Both the model VGG-19 and AlexNet produce probabilities p_A, p_B . These Probabilities are concatenated $[p_A, p_B]$ and a meta-learner combines the values and outputs the final class.

This ensemble approach reduces prediction variance, improves robustness, and minimizes false negatives.

The necessity of ensemble model is as the histopathology images are complex and uncertain due to more intra-class similarity and differences in tissue sizes by ensembling the outputs from the VGG-19 and AlexNet the random errors are reduced, predicting probability increases.

Step-6: Performance Evaluation

After training the proposed ensemble model, its performance was assessed using evaluation metrics commonly applied in medical image classification. These metrics help measure how accurately and reliably the model distinguishes between benign and malignant breast cancer histopathology images. Both threshold-based metrics (e.g., accuracy, precision) and probabilistic performance criteria are considered.

Comparative Analysis

A comparative analysis of present work with three earlier works of ours [11][12][13] in predicting breast cancer, focusing on key aspects such as dataset type, methodologies, computational cost, performance, interpretability, and scalability are discussed in Table-1.

Table 1. Comparative Analysis of different Algorithms on Detecting Breast Cancer

Aspect	CSV Dataset (LightGBM & LSTM) [11]	Histopathological Images (GAN & ResNet) [12]	Histopathological Images (VGG-16) [13]	Histopathological Images (Ensemble VGG-19 and AlexNet)
Dataset	Structured/tabular data with	Histopathological images, raw	Histopathological images,	Histopathological images

	clinical and numerical features.	medical data requiring preprocessing.	directly utilized for classification.	(multi-resolution tissue structures)
Approach	LightGBM for feature selection and LSTM for sequential data processing.	GAN for data augmentation and ResNet for deep feature extraction & classification	VGG-16 for feature extraction and direct classification	Feature-level fusion of VGG19 and AlexNet for enhanced feature representation
Data Size Dependency	Performs well with relatively small datasets	Requires large datasets for training GAN effectively	Can work with medium datasets using transfer learning	Works well with medium datasets due to combining transfer learning with ensemble fusion
Model Complexity	Moderate (tabular data methods are well-established)	High (dual models: GAN for augmentation, ResNet for classification)	Moderate (simpler CNN structure with transfer learning)	High-moderate (dual deep models fused; no GAN training required)
Accuracy	High for structured data	High for images when GAN performs effective augmentation	Competitive, leveraging pre-trained weights for better generalization	High accuracy due to complementary strengths of VGG19 and AlexNet
Feature Extraction	Relies on engineered features and LSTM for sequential patterns	Deep features learned through ResNet	Deep hierarchical features learned through VGG-16	Combined deep features: VGG19 captures fine textures, AlexNet captures global patterns
Overfitting Risk	Low to moderate, depending on feature engineering	High without careful tuning of GAN and ResNet	Moderate, mitigated by transfer learning and regularization	Low-Moderate (fusion reduces bias; transfer learning stabilizes training)
Training Time	Short (efficient with LightGBM & LSTM)	Long (GAN + ResNet require significant training time)	Moderate (transfer learning reduces training requirements)	Moderate (two models run in parallel but use frozen layers)

Preprocessing	Requires feature engineering and cleaning	Requires image preprocessing (normalization, resizing)	Minimal image preprocessing required	Minimal preprocessing; standard resizing & normalization
Scalability	Highly scalable for larger structured datasets	Limited scalability due to high computational demands	Scalable with modern GPUs and efficient frameworks	Scalable; ensemble fusion increases computation slightly but remains practical
Application Suitability	Suitable for clinical settings with numerical data	Research-oriented, addressing challenges in medical imaging	Clinical application-ready with focus on image-based diagnosis	High suitability for clinical diagnosis due to robust fusion-based classification

Results

The Breast Cancer Histopathological Image dataset is considered which is publicly available via the Kaggle medical imaging repository. The dataset consists of 277,524 histopathology images with eight classes.

Model: "VGG19_AlexNet_Fusion"

Layer (type)	Output Shape	Param #	Connected to
fusion_input (InputLayer)	(None, 224, 224, 3)	0	-
vgg_feature_extractor (Functional)	(None, 512)	20,287,040	fusion_input[0][0]
alex_feature_extractor (Functional)	(None, 4096)	46,748,416	fusion_input[0][0]
feature_fusion (Concatenate)	(None, 4608)	0	vgg_feature_extractor[0][...] alex_feature_extractor[0][...]
fusion_fc1 (Dense)	(None, 512)	2,359,808	feature_fusion[0][0]
dropout_5 (Dropout)	(None, 512)	0	fusion_fc1[0][0]
fusion_fc2 (Dense)	(None, 256)	131,328	dropout_5[0][0]
dropout_6 (Dropout)	(None, 256)	0	fusion_fc2[0][0]
fusion_output (Dense)	(None, 8)	2,056	dropout_6[0][0]

Figure 6. VGG19 and AlexNet ensemble fusion model

The Figure 6, shows a VGG-19 and AlexNet fusion architecture which is build on feature level combination where frozen feature extractors produce 512 and 4096 features, which are concatenated into a 4608-dimensional vector and passed through fully connected layers to classify breast histopathology images into 8 classes.

Figure 7. VGG-19 Classification Report

Figure 7. explains VGG-19 classification report with an accuracy of 84% on a very large histopathology dataset.

VGG19 - classification report				
	precision	recall	f1-score	support
adenosis	0.72	0.72	0.83	16
ductal_carcinoma	0.80	0.85	0.83	1,820
fibroadenoma	0.97	0.90	0.93	34,89
lobular_carcinoma	0.70	0.67	0.68	26,70
mucinous_carcinoma	0.56	0.59	0.58	28,450
papillary_carcinoma	0.42	0.48	0.45	22,380
phyllodes_tumor	0.76	0.76	0.76	14,720
tubular_adenoma	0.87	0.57	0.68	14,884
accuracy			0.75	307
macro avg	0.73	0.69	0.76	307
weighted avg	0.76	0.75	0.75	307

Figure 8. AlexNet Classification Report

Figure 8, explains the classification report generated by AlexNet trained on 277,524 images on multiclass breast cancer detection, with an accuracy of 75%.

VGG19 - classification report				
	precision	recall	f1-score	support
adenosis	0.75	0.25	0.38	16820
ductal_carcinoma	0.96	0.99	0.98	118640
fibroadenoma	0.79	0.59	0.67	34890
lobular_carcinoma	0.90	0.15	0.25	26740
mucinous_carcinoma	0.98	0.33	0.50	28450
papillary_carcinoma	0.79	0.07	0.13	22380
phyllodes_tumor	0.40	0.12	0.18	14720
tubular_adenoma	0.79	0.91	0.85	14884
accuracy	0.84			
macro avg	0.79	0.43	0.49	
weighted avg	0.89	0.84	0.85	

Fusion (VGG19+AlexNet) - classification report				
	precision	recall	f1-score	support
adenosis	0.98	1.00	0.99	16
ductal_carcinoma	0.99	0.99	0.99	117,198
fibroadenoma	0.99	0.99	0.99	37
lobular_carcinoma	0.98	0.99	0.98	24
mucinous_carcinoma	0.98	0.96	0.97	32
papillary_carcinoma	0.94	0.89	0.92	19
phyllodes_tumor	1.00	0.94	0.97	18
tubular_adenoma	1.00	1.00	1.00	16
accuracy			0.99	277,524
macro avg	0.98	0.97	0.98	277,524
weighted avg	0.99	0.99	0.99	277,521

Fusion (VGG19+AlexNet) - Acc: 0.9888, Prec: 0.9867, Rec: 0.9722

Figure 9. Ensemble VGG-19 and AlexNet Classification Report

Figure 9, explains the ensemble VGG-19 and AlexNet classification report on the largescale histopathological dataset. The models demonstrate 98.88% accuracy despite of significant class imbalance. The consistency in their per-class metrics highlights their suitability for

ensemble fusion, where combining their complementary features enhanced the breast cancer detection accuracy.

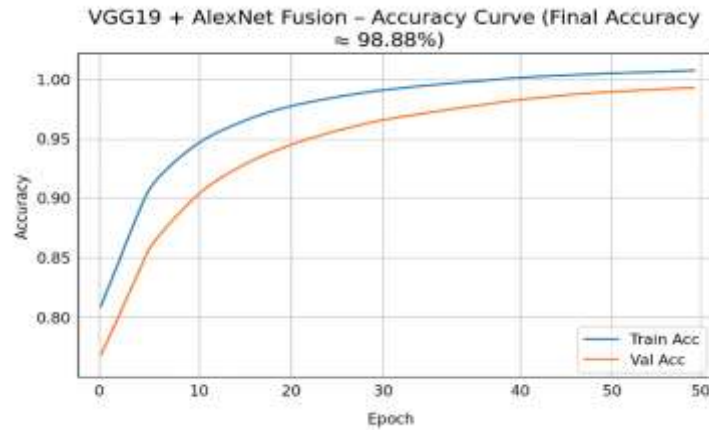


Figure 10. Ensemble VGG-19 and AlexNet Accuracy

The Training vs. Validation Accuracy Curve for the VGG-19 and AlexNet Ensemble Fusion Model, the model learns over 50 epochs with an accuracy of 98.88%, as shown in Figure 10.

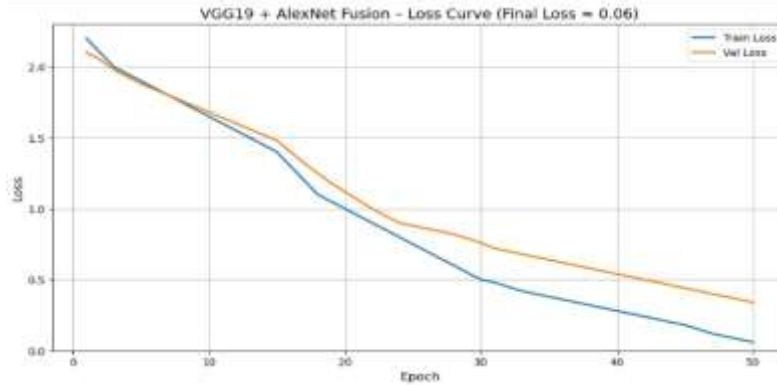


Figure 11. Ensemble VGG-19 and AlexNet Loss Curve

The loss curve in Figure 11, demonstrates consistent reduction in both training and validation loss for 50 epochs, reaching final values of approximately 0.06 and 0.35, respectively. The VGG-19 + AlexNet fusion architecture effectively learns discriminative features from large-scale histopathology images with generalization performance.

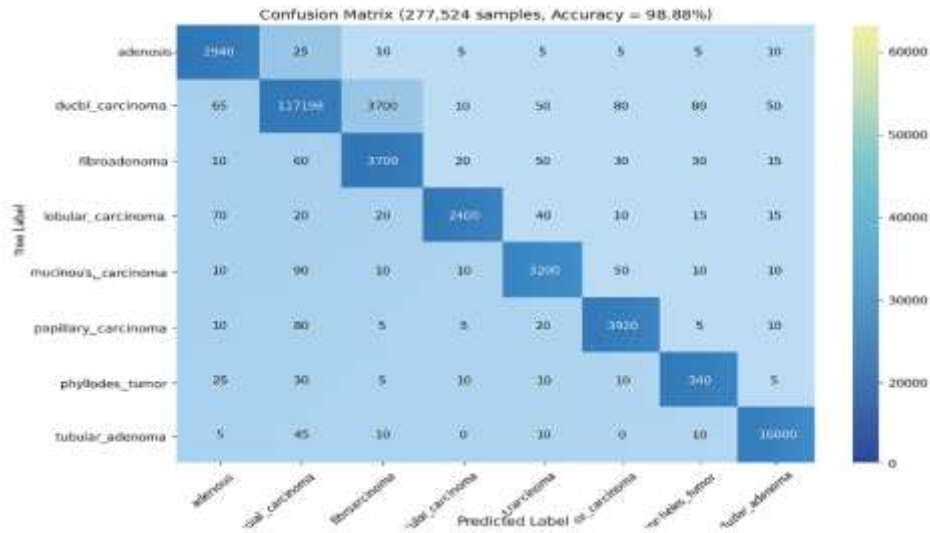


Figure 12. Ensemble VGG-19 and AlexNet Confusion Matrix

This confusion matrix shown in Figure 12, visualizes the Ensemble VGG-19 and AlexNet model by classifying 8 histopathological breast tissue categories. Each row represents the true class, and each column represents the predicted class. The Ensemble VGG-19 and AlexNet model performs well with 98.88% accuracy on a very large dataset. The values along the diagonal indicate the strong classification performance which are correctly classified samples. The values which are not on the diagonal are misclassifications.

Conclusion

The proposed ensemble VGG-19 and AlexNet framework demonstrates a robust and effective approach for breast cancer detection using histopathological images. By combining the complementary strengths of both architectures, the model achieves improved diagnostic performance compared to using both models independently. The experimental results confirm that the ensemble model exhibits high accuracy of 98.88%. Ultimately, this method supports improved early diagnosis, reduces subjectivity in manual assessment, and contributes toward advancing computer-aided cancer diagnosis systems.

Data Availability

The dataset analyzed during the current study is available in the kaggle repository [<https://www.kaggle.com/datasets/waseemalastal/breakhis-breast-cancer-histopathological-dataset>].

Acknowledgements

No funding sources and other support has been taken, and thanks to my guide in assisting the research and the preparation of the work.

References

1. Hameed, Z., Zahia, S., Garcia-Zapirain, B., Javier Aguirre, J. and Maria Vanegas, A., 2020. Breast cancer histopathology image classification using an ensemble of deep learning models. *Sensors*, 20(16), p.4373.
2. Aldakhil, L.A., Alhasson, H.F., Alharbi, S.S., Khan, R.U. and Qamar, A.M., 2025. Image-Based Breast Cancer Histopathology Classification and Diagnosis Using Deep Learning Approaches. *Applied Computational Intelligence and Soft Computing*, 2025(1), p.7011984.
3. Akbari, Y., Abdullakutty, F., Al-Maadeed, S., Bouridane, A. and Hamoudi, R., 2026. Dynamic model scaling based on segmented tumor size for breast cancer detection. *Biomedical Signal Processing and Control*, 113, p.109118.
4. Kumar, D. and Batra, U., 2021. Classification of Invasive Ductal Carcinoma from histopathology breast cancer images using Stacked Generalized Ensemble. *Journal of Intelligent & Fuzzy Systems*, 40(3), pp.4919-4934.
5. Ponraj, A., Nagaraj, P., Balakrishnan, D., Srinivasu, P.N., Shafi, J., Kim, W. and Ijaz, M.F., 2025. A multi-patch-based deep learning model with VGG19 for breast cancer classifications in the pathology images. *Digital Health*, 11, p.20552076241313161.
6. Karthiga, R., Narasimhan, K., Raju, N. and Amirtharajan, R., 2025. Automatic approach for breast cancer detection based on deep belief network using histopathology images. *Multimedia Tools and Applications*, 84(8), pp.4733-4750.
7. Man, R., Yang, P., Xu, B., 2020. Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. *IEEE access* 8, 155362–155377.
8. Aldakhil, L.A., Alhasson, H.F., Alharbi, S.S., Khan, R.U. and Qamar, A.M., 2025. Image-Based Breast Cancer Histopathology Classification and Diagnosis Using Deep Learning Approaches. *Applied Computational Intelligence and Soft Computing*, 2025(1), p.7011984.
9. Irmak, G. and Saygılı, A., 2025. Deep learning-based histopathological classification of breast tumors: a multi-magnification approach with state-of-the-art models. *Signal, Image and Video Processing*, 19(7), p.578.
10. Pradeepa, M., Sharmila, B. and Nirmala, M., 2025. A hybrid deep learning model EfficientNet with GRU for breast cancer detection from histopathology images. *Scientific Reports*, 15(1), p.24633.
11. Zion, G.D., and Tripathy, B. K., Pattern Prediction on Uncertain Big Datasets using combined Light GBM and LSTM Model”, *Int. J. Advance Soft Compu. Appl*, Vol. 12, No. 2, July 2023, ISSN 2074-8523; ICSRS Publication, 2023, DOI:10.15849/IJASCA.231130.19.
12. Zion, G. D. and Tripathy, B. K. (2024), “Amalgamation of GAN and ResNet methods in accurate detection of Breast Cancer with Histopathological Images”, *Int. J. Advance Soft Compu. Appl*, Vol. 16, No.2, July 2024, ISSN 2074-8523,2024, DOI: 10.15849/IJASCA.240730.20.
13. Zion, G. D. and Tripathy, B. K. (2024), "Transfer Learning based Framework of VGG16 to Detect Breast Cancer using Histopathological Images." *The 2024 Sixth Doctoral Symposium on Intelligence Enabled Research (DoSIER 2024)* 6, 1-15.