Deep Learning-Based Breast Cancer Classification: A Comparative Study of CNN Architectures

Springboard

Capstone -3

Pramod Acharya

Introduction

1. Breast cancer is a prevalent form of cancer in women, emphasizing the importance of early detection and accurate diagnosis for effective treatment.

2. Deep learning techniques, particularly convolutional neural networks (CNNs), are utilized in this project to develop a classifier for distinguishing between benign and malignant histology images of breast cancer.

3. Throughout history, efforts to understand, diagnose, and treat breast cancer have evolved, with advancements in medical knowledge and technology leading to more sophisticated diagnostic and treatment options.

4. Artificial intelligence (AI) and machine learning, especially deep learning, offer promising avenues for improving breast cancer detection and diagnosis through computer-aided diagnosis (CAD) systems.

5. Despite advancements, challenges such as limited dataset availability, interpretability concerns, and regulatory considerations persist, but ongoing research aims to overcome these obstacles and improve outcomes for breast cancer patients globally.

Data Collection

1. The IDC_regular dataset from Kaggle, containing breast cancer specimen patches scanned at 40x magnification, is utilized in this project.

2. With over 277,000 patches available, this dataset offers a substantial amount of data for analysis and can be downloaded from the provided link.

Exploratory Data Analysis (EDA)

1. Exploratory Data Analysis (EDA) was conducted to understand feature distribution, utilizing techniques such as pairplot visualization and feature engineering to assess dataset composition.

2. Visual representation of 18 randomly selected images from each class aided in assessing dataset diversity, providing insights crucial for subsequent data preprocessing and model development stages.

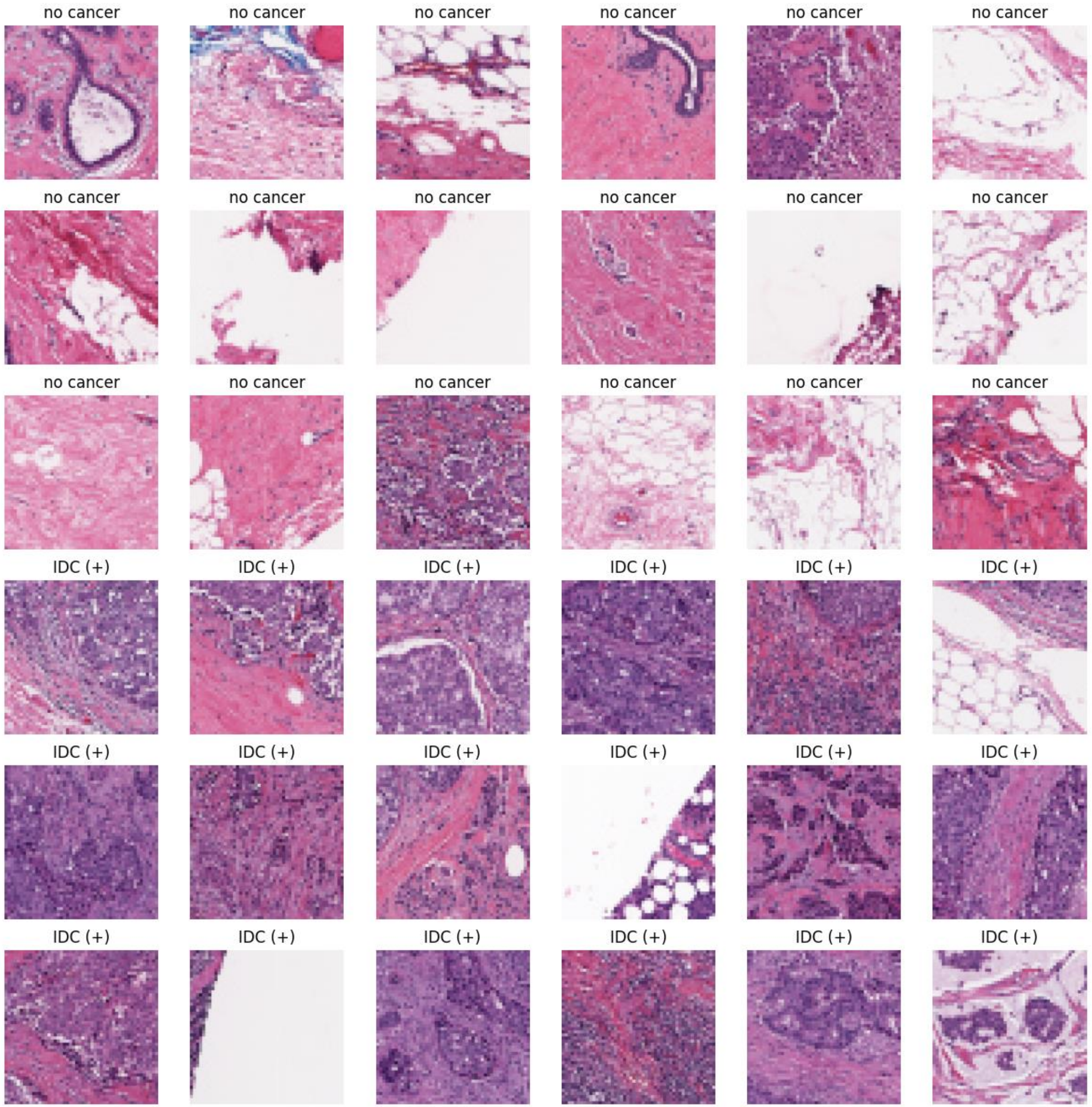
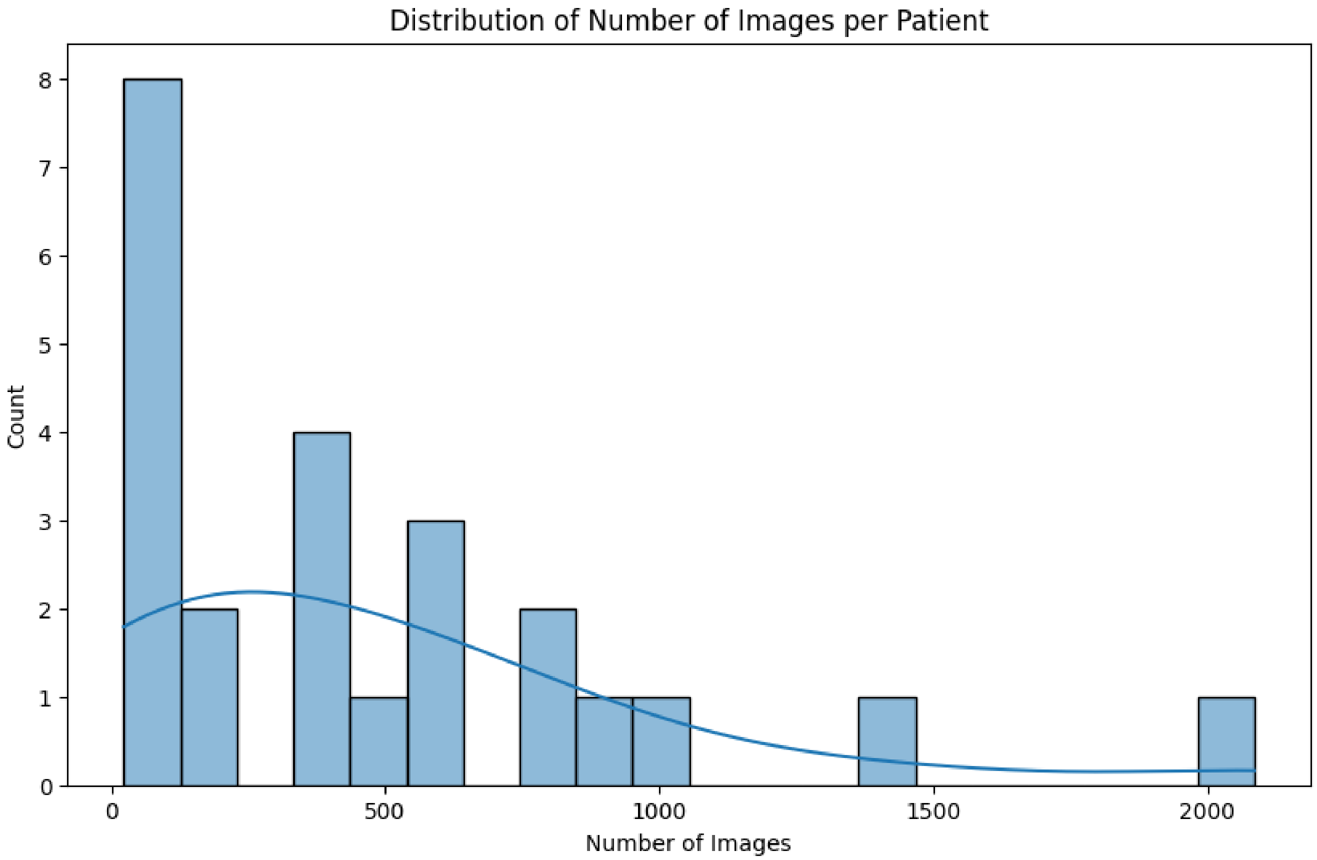


Fig: Samples showing the breast cancer images



Distribution of Number of Images per Patient
Count
Number of Images
0
1
2
3
4
5
6
7
8
0
500
1000
1500
2000

Pre-processing

1. Image preprocessing involves loading, resizing, and normalizing images to 50x50 pixels, ensuring consistency and reducing computational complexity.

2. Dimensionality reduction techniques like PCA and t-SNE aid in visualizing data structure, facilitating pattern identification relevant to breast cancer classification.

3. Features and labels are extracted from preprocessed images, ensuring balanced representation of IDC-negative and IDC-positive cases in the dataset.

4. A CNN architecture, comprising convolutional, batch normalization, activation, max-pooling, and dropout layers, is designed for effective feature extraction and classification, trained using Adam optimizer with binary cross-entropy loss for accurate discrimination between benign and malignant tumors.

Model Architecture

1. Multiple convolutional layers followed by pooling layers for hierarchical feature extraction.

2. Fully connected layers with dropout regularization to prevent overfitting, leading to a softmax output layer for classification.

Tuned Hyperparameters

- Learning Rate: 0.001
- Batch Size: 32
- Number of Convolutional Layers: 4
- Kernel Size: 3x3
- Dropout Rate: 0.5

Modeling

1. Preprocessing involves dataset splitting, label encoding, and defining CNN architecture.
2. CNN architecture includes convolutional, batch normalization, max-pooling, and dropout layers.
3. Model summary provides insights into layer configuration and trainable parameters.
4. Model is compiled with appropriate optimization and loss functions and trained using early stopping.

Training Process

1. Performance metrics are monitored across epochs with training history stored for visualization.
2. Training and validation accuracy and loss are plotted for assessing learning dynamics and generalization capabilities.
3. CNN architecture demonstrates robust performance with high accuracy on both training and testing data.

Model Evaluation

1. Training and validation accuracy reached approximately 99.81% and 95.61%, respectively, after 75 epochs.

2. Loss gradually decreased, indicating improved model performance over epochs.

3. Plots visualizing accuracy and loss demonstrated favorable trends, signifying effective learning and generalization.

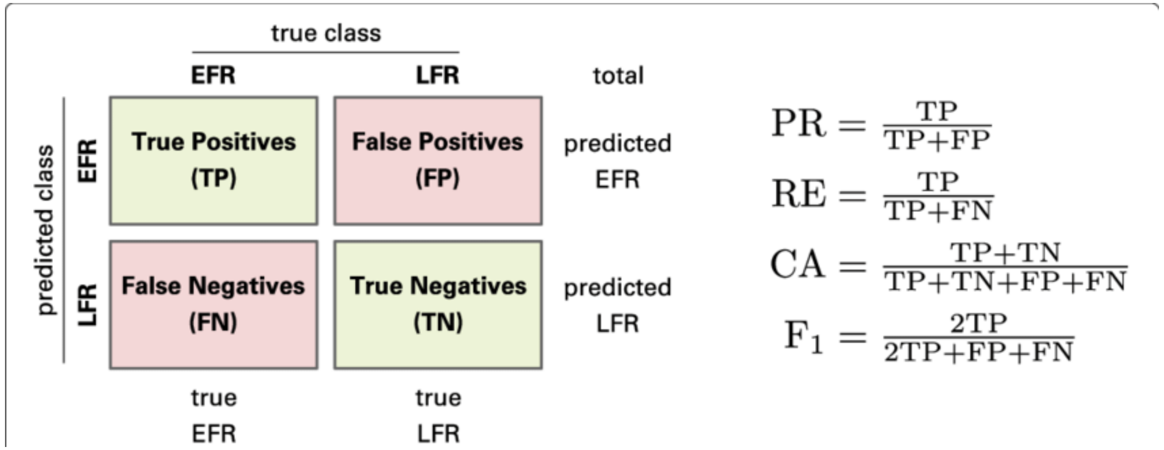
Metrics Calculation

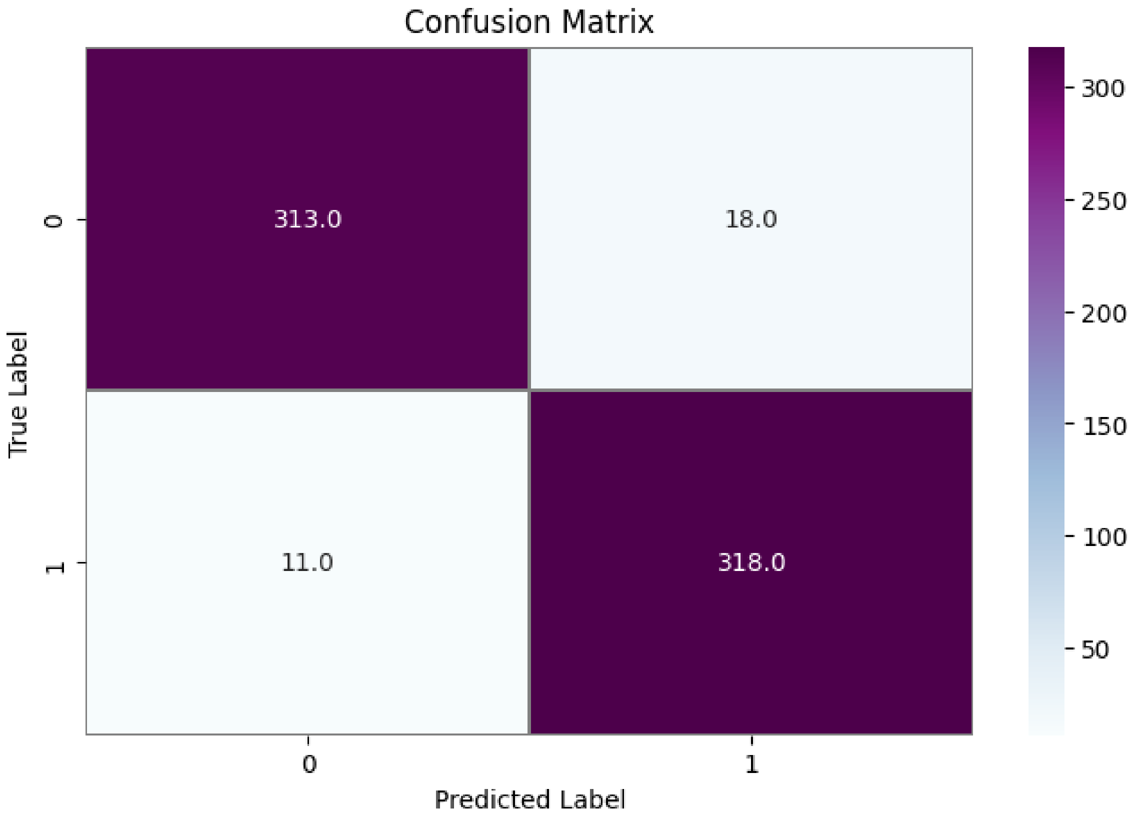
- Confusion matrix provided insights into classification performance, showing distribution of true positive, true negative, false positive, and false negative predictions.

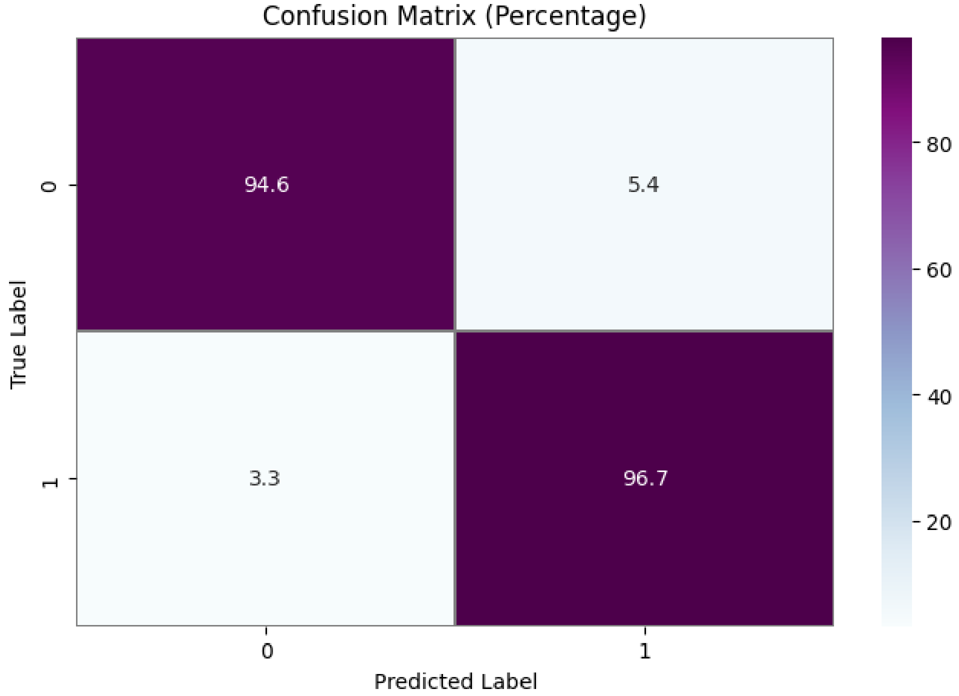
- Metrics including recall, precision, and F1 score were computed, all exceeding 95%, indicating robust classification capability.

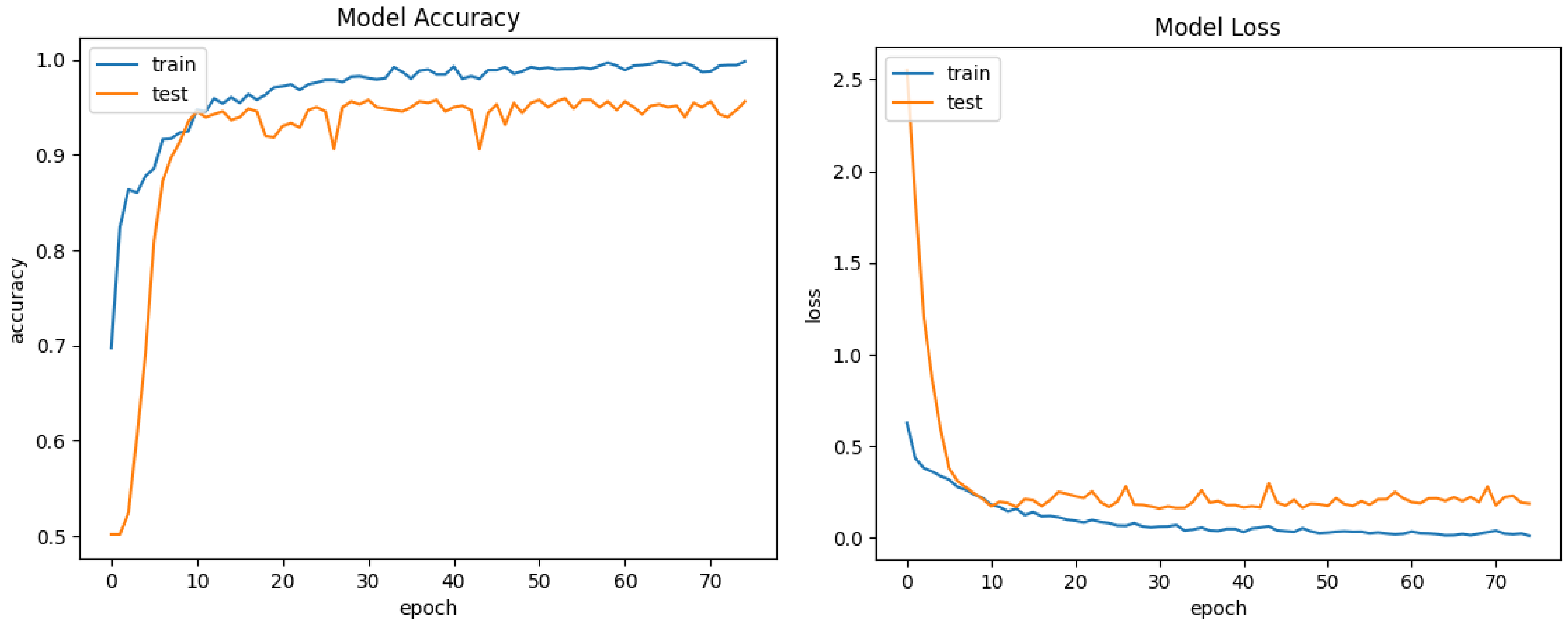
Individual Image Predictions:

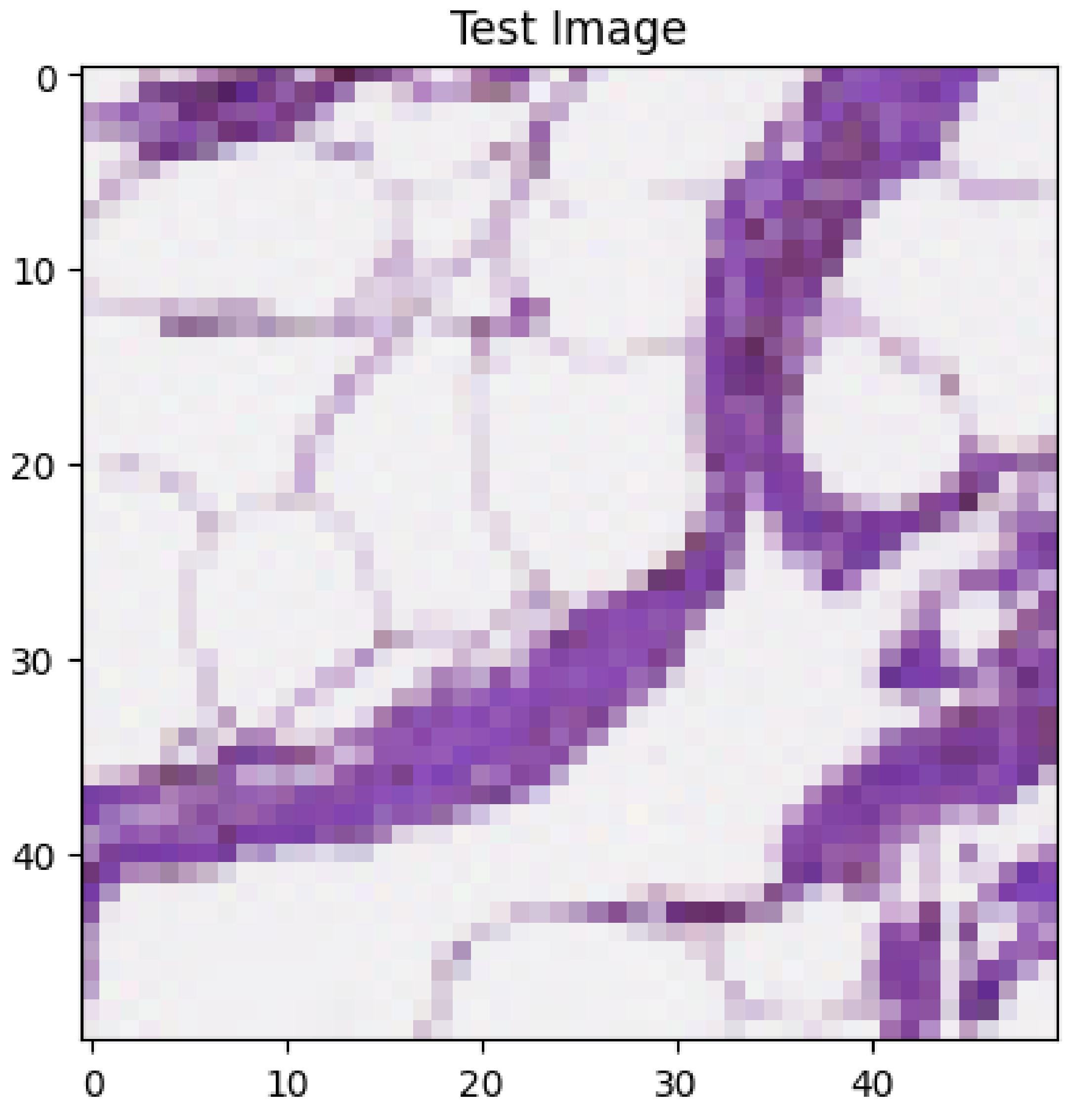
- Predictions on individual test set images showed agreement between predicted and true values, reaffirming model's accuracy on unseen data.


true class
EFR
LFR
total
predicted class
EFR
LFR
True Positives (TP)
False Positives (FP)
predicted EFR
False Negatives (FN)
True Negatives (TN)
predicted LFR
true EFR
true LFR
$\mathrm{PR} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$
$\mathrm{RE} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$
$\mathrm{CA} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}}$
$\mathrm{F}_1 = \frac{2\mathrm{TP}}{2\mathrm{TP+FP+FN}}$


Confusion Matrix
313.0
18.0
11.0
318.0
True Label
Predicted Label
0
1
0
1
300
250
200
150
100
50


Confusion Matrix (Percentage)
94.6
5.4
3.3
96.7
True Label
Predicted Label
0
1
0
1
80
60
40
20


Model Accuracy
train
test
accuracy
epoch
1.0
0.9
0.8
0.7
0.6
0.5
0
10
20
30
40
50
60
70
Model Loss
train
test
loss
epoch
2.5
2.0
1.5
1.0
0.5
0.0
0
10
20
30
40
50
60
70



Test Result:

•The CNN model predicted the test image with index 100 as class 0.

•The true label of the test image with index 100 is also class 0.

Conclusion

1. Utilized deep learning techniques to develop a robust classifier for benign and malignant breast cancer images.

2. Achieved high accuracy (99.81% training, 95.61% validation) through meticulous modeling and evaluation.

3. Confirmed model reliability with comprehensive metrics, including recall, precision, and F1 score (>95%).

4. Demonstrates suitability for real-world applications, ensuring accurate classification and improved patient outcomes.

Thanks!