International Journal of Science Academic Research
Vol. 07, Issue 03, pp.433-440, March 2026
Available online at http://www.scienceijsar.com
ISSN: 2582-6425

Research Article

AI-DRIVEN MEDICAL IMAGES FOR PREDICTION AND CLASSIFICATION BREAST CANCER BY USING CONVOLUTIONAL NEURAL NETWORKS

[1]Ayman Tagelsir Fageer, [2]Hisham Mohammed Makki, [3]Sara Hashim Hassan, [4]Ahmed Khalifa Mohammed Abdelrahman and [1,*]Anas. A. Nicola

[1, 2]Department of Computer Science, Faculty of Computer Sciences and Information Technology, Port Sudan Ahlia, Sudan
[3, 4]Department of Management Information System, Faculty of Computer Sciences and Information Technology, Port Sudan Ahlia, Sudan
[1]Faculty of Telecommunication engineering and space Technology, Future University, Khartoum, Sudan

Received 15th January 2026; **Accepted** 18th February 2026; **Published online** 20th March 2026

Abstract

Breast cancer is a common disease that many women face throughout their lives. Therefore, early diagnosis is the most effective and reliable tool for effectively treating cancer. Computer-aided diagnosis is one of the software technologies designed to assist doctors in detecting or diagnose cancer. In fact, to reduce mortality via using medical images analysis with less time. Recently, medical images analysis used convolutional neural networks to evaluate a vast amount of data for detecting cancer cells images classification. However, breast cancer tumors could be determined by analyzing microscopic images. Analysis taken a complex time consuming. So, there are many risks like misdiagnosis absent from pathologists. On this study our goals refer to assign and reduce the time-consuming for classification process. Fore to increase the accuracy when they attempt using data augmentation for to evaluate the performance of the classifiers, by using various performance metrics. Furthermore, in this study, the pre-trained convolution neural network (CNNs) models are applied to extract the features that Support Vector Machine (SVM), to classify and transfer learning from pre-trained CNN models for feature extraction and classification, this approach was applied to image-based breast cancer classification using histopathological images from the public dataset, and the experimental result shown that the **resent50**achieved high accuracy and **97.2%**has been compared to the **inceptionV3**with accuracy of**96.6%.**Finally, on this study we focusing to recommends expanding to research other diseases such as Pneumonia, Brain Cancer, Covid-19, etc. Therefore, to incorporate other classifier, such as BayesNet, Decision Tree, etc., for to achieve higher classification accuracy by using different type of others models.

Keywords: Deep learning, Computer neural network CNN, Radiation exposure, Image classification, Risk factors.

INTRODUCTION

Breast cancer remains one of the most prevalent in life-threatening diseases, affecting women on the wide area in the world [1]. Although it predominantly affects women. Thus, the diseases also occur in men at the same time, that is can makesa critical public health concern [2]. According to the World Health Organization (WHO)[3], breast cancer accounts for approximately 25% of all cancer cases in women, with millions of new diagnoses reported each year[4].In fact, the process of the discovers threating should be clarified in the Early detection is the crucial for improving survival rates. When peoples delays diagnosis significantly increases the risk of metastasis, and that can makes the process of the treatment more complex and less effective[5].The traditional diagnostic approaches, including mammography[6], and histopathological analysis[7], require specialized expertise for often time-consuming, it can be leads the potential human errors, in the especial area of different regions with had a shortage of trained pathologists[8]. Artificial intelligence (AI) and deep learning, computer-aided diagnosis (CAD) systems[9], has been emerged for to promising tools for to enhancement the process of the breast cancer detection and classification(Arun Kumar and Sasikala 2023). Convolutional Neural Networks (CNNs) [10], a subset of deep learning architectures, also demonstrated remarkable performance in medical image analysis automatically for to extracts the hierarchical features from theraw data.

Unlike conventional machine learning methods that rely on handcrafted feature extraction, CNNs learn to identify complex patterns within histopathological images, improving classification accuracy and reducing reliance on manual feature engineering [11]. The CNN architectures consist of multiple layers [12], including convolutional, pooling, and fully connected layers, which enable efficient feature extraction and classification [13]. By leveraging transfer learning, pre-trained CNN models such as ResNet-50(Categorization and Koonce) and Inception-V3[14], can be fine-tuned to classify breast cancer images with high accuracy. These models have been trained on large-scale datasets like ImageNet, allowing them to generalize effectively to medical imaging tasks. Studies have shown that CNN-based approaches can outperform traditional classification methods, achieving accuracies exceeding 97% in distinguishing benign and malignant tumors. Despite the success of CNNs in medical image classification, there are several challenges remain, including limited availability of labeled datasets[15], high computational requirements[16], and the risk of over fitting [17] Data augmentation techniques, for to increase the size of the training set and provides more illustrative samples for to large-capacity users [18]. Such as image rotation, flipping, and color normalization, help mitigate these challenges by artificially expanding training datasets and improving model generalization. Additionally, integrating CNNs with classifiers like Support Vector Machines (SVM) further enhances classification performance by leveraging complementary feature representations.

Corresponding Author:** ***Anas. A. Nicola,
Faculty of Telecommunication engineering and space Technology, Future University, Khartoum, Sudan.

This paper aims to address these challenges by implementing a CNN-based breast cancer classification framework using pre-trained deep learning models. The key objectives of this study are:

- To reduce the time required for breast cancer classification by automating feature extraction using CNNs.
- To improve classification accuracy through data augmentation and transfer learning.
- To evaluate the performance of different CNN models using various metrics, including accuracy, precision, recall, and F1-score.

By enhancing diagnostic accuracy and efficiency, this study contributes to the ongoing advancements in AI-driven medical imaging, paving the way for more reliable and Accessible breast cancer detection systems. Future extensions of this research may explore applications in detecting other diseases, such as pneumonia, brain tumors, and COVID-19, using similar deep learning techniques rather than requiring human engineering [19]. There are different types of layers included in the CNN architecture and the ones are known as input layers, hidden layers, and output layers. Features may be obtained by the learning process of images with each other, which may be done on the hidden layers. [20]. The workflow and the architecture of CNN are shown in Figure. 1 and 2, respectively.

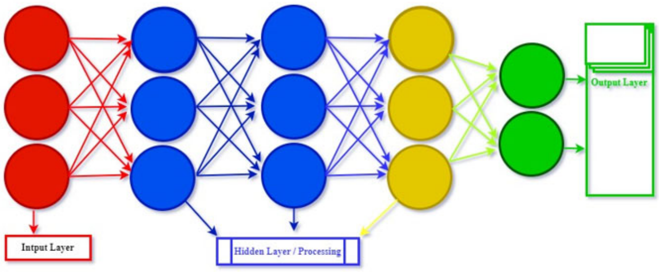

Figure 1. CNN layers architecture

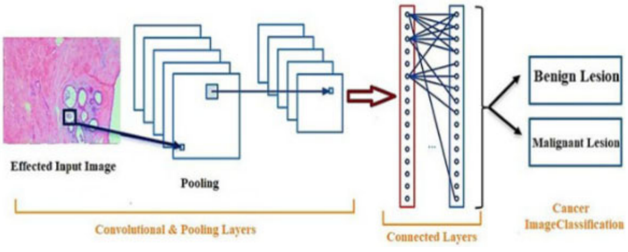

Figure 2. CNN workflow for breast cancer classification

Figure 2, CNN workflow for breast cancer classification Image features can be learned from the connected layers to classify breast cancer images. In addition, features may be extracted automatically from images using pre-trained deep neural networks (ResNet, Inception-V3Net and ShuffleNet),as performed in the research performance.

Related Works

Deep learning has significantly advanced breast cancer classification, with several studies exploring different architectures and methodologies. Paper [21], demonstrated the effectiveness of CNNs with transfer learning for breast cancer classification. Their study showed that a shallow custom CNN achieved 100% accuracy, outperforming pre-trained models like ResNet and InceptionV3, which reached around 92%. However, challenges such as dataset limitations, computational resource dependency, and generalization across imaging modalities remained unresolved. Additionally, most studies relied on specific datasets, limiting real-world applicability.
On this research introduced amodelby [22], build a custom CNN model (BCCNN) alongside pre-trained networks such as Xception, InceptionV3, and ResNet50 for classifying breast cancer into eight categories. Using MRI images and data augmentation with GANs, their model achieved 98.28% F1-scor, demonstrating that high-resolution images significantly enhance classification accuracy. However, the study highlighted the need for more diverse datasets and improved generalization across imaging techniques. To address classification accuracy and model robustness [23], proposed a meta-learning ensemble approach that combined multiple CNN models, including ResNet50, DenseNet121, and InceptionV3. Their ensemble approach improved accuracy to 90% on the BUSI dataset by integrating transfer learning and data augmentation. Nevertheless, dataset imbalance and overfitting issues persisted, necessitating further advancements in model training techniques.

Further refining deep learning for breast cancer detection [24], introduced LMHNet, a deep neural network optimized using the Levenberg–Marquardt algorithm. Their model utilized attention mechanisms for feature refinement, achieving 99% accuracy in binary classification and 88% accuracy in multi-class classification across eight breast cancer subtypes. While the study demonstrated improved convergence and classification performance, the challenge of dataset generalization remained unresolved. [25] explored deep learning for molecular subtyping of breast cancer using Dynamic Contrast-Enhanced MRI (DCE-MRI). Their CNN-based deep feature extraction method improved classification accuracy by leveraging multi-phase MRI analysis. However, issues related to segmentation errors, dataset limitations, and imaging variability highlighted that need for further improvements in feature extraction and generalization. Despite significant advancements, existing studies still face challenges related to dataset diversity, generalization across imaging techniques, and computational efficiency. Our paper aims to bridge these gaps by leveraging a CNN-based approach with transfer learning and enhanced data augmentation techniques to improve breast cancer classification accuracy. Unlike previous studies, our model integrates multiple imaging modalities, optimizes computational efficiency, and enhances generalization through advanced feature extraction and preprocessing techniques. Creating more resilient architectures that are less susceptible to minor input variations [26]. This approach ensures a more robust, scalable, and clinically applicable breast cancer classification system, addressing key limitations in existing methodologies.

METHODOLOGY

The proposed method utilizes Convolutional Neural Networks (CNNs) with transfer learning and Support Vector Machines. The overall workflow is illustrated in Figure 2, providing an overview of the process from data preprocessing to classification. The dataset used in this study consists of histopathological breast tumor images from the Break His dataset, categorized into benign and malignant tumors.

Preprocessing steps includecolor normalization and resizing, ensuring compatibility with the input requirements of the pre-trained CNN models (ResNet-50 and Inception-V3). Figures 2 and 3 depict sample benign and malignant tumor images, respectively. To enhance model performance and reduce overfitting, data augmentation techniques such asimage flipping, rotation, and translation are applied exclusively to the CNN-based approach. This augmentation process is illustrated in Figure 6. Feature extraction is performed usingpre-trained CNN models, where deep features are obtained from the 'fc1000' layer before classification. The extracted features are then used to train anSVM classifier, leveraging its ability to enhance classification performance. A schematic representation of thetransfer learning workflow is presented in Figure 7

The dataset is randomly split into70% training and 30% testing for each magnification factor. Additionally,25% of the training data is reserved for cross-validation, ensuring optimal model selection. After determining the best model parameters, they are incorporated into the final training dataset. The same testing dataset is used across all experiments to maintain consistency. Performance evaluation is conducted using multiple metrics, including accuracy, sensitivity, specificity, precision, F1-score, and AUC. Confusion matrices summarizing classification results are provided in Table 1. A comparative analysis of different models is further illustrated in Figures 6–10, highlighting the effectiveness of our approach.

Our contribution

This study presents a deep learning-based approach for breast cancer classification using Convolutional Neural Networks (CNNs) and Support Vector Machines (SVM). Therefore, our methodology follows a structured pipeline involving dataset selection, preprocessing, data augmentation, feature extraction, classification, and evaluation using multiple performance metrics.

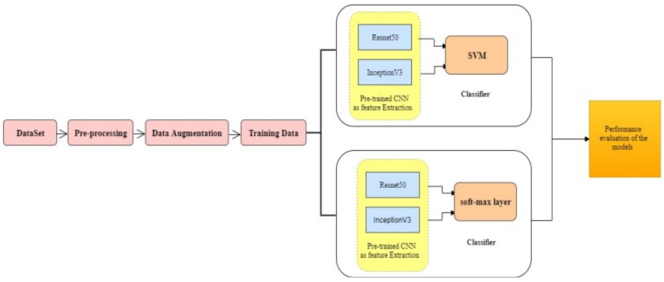

Figure 3. Dataflow Diagram of the Proposed Method

The diagram illustrates the complete workflow of the proposed breast cancer classification method. It begins with the dataset, followed by preprocessing steps such as color normalization and resizing. Then, data augmentation techniques (e.g., flipping, rotation, translation) are applied to enhance the dataset and reduce over fitting. The training data is then fed into two classification pipelines. In both, pre-trained CNN models (ResNet-50 and Inception-V3) are used for feature extraction. The extracted deep features are passed to two different classifiers:

- One pipeline uses a Support Vector Machine (SVM).
- The other uses a Soft-max layer for classification.

Finally, the results from both models are subjected to performance evaluation, using metrics such as accuracy, precision, recall, F1-score, and AUC.

Dataset:

The breast cancer histopathological image classification (BreakHis Dataset) is collecting form online from the laboratory of vision, robotics, and imaging of the federal university of Parana, brazil and composed of (957) microscopic images of breast tumor tissue collected from 82 patients using the magnifying factor (100X,).it contains (260) benign and (697) malignant samples (700X460 pixels, 3-channel RGB, 8-bit depth in each channel, PNG format).

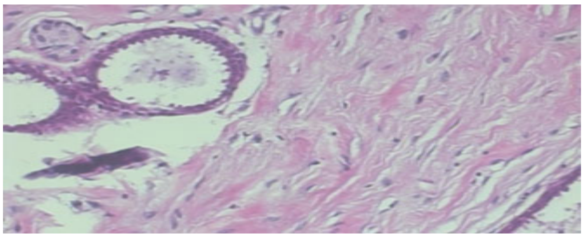

Figure 4. Benign Tumor(Not Cancer)

This is a microscopic Figure4of benign breast tissue from the Break His dataset, stained with Hematoxylin and Eosin (H&E). The pink areas show connective tissue, while the purple spots are cell nuclei. The image displays normal glandular structures and fibrous tissue, with no signs of malignancy.

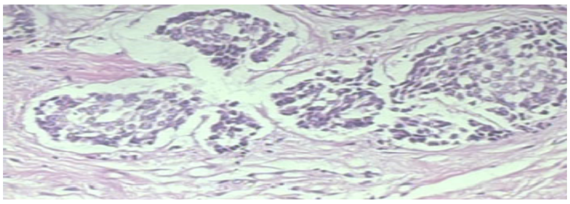

Figure 5. Malignant Tumor (Cancer)

Figure 5 a 100x magnified histopathological image of malignant breast tissue from the BreakHis dataset. It shows densely packed, dark-stained nuclei, irregular glandular structures, and disorganized cells, all indicating breast cancer.

Data Pre-Processing

Before we could use the data, we had to convert the image into (RGB) using a function in Mat lab(gray2rgb),Data pre-processing should be performed within each resampling iteration to avoid information leakage [27], also we had to size normalize it to make it consistent with the requirements of different networks. Resizing of input images was accomplished via data rescaling and cropping to ensure that image sizes fit each pre-trained deep neural network. The networks used in this thesis are pre-trained deep neural network ResNet50, Inception-V3, and they all require different size input images. Specifically, ResNet50 uses 224x224x3 images; Inception-V3 uses 299x299x3 images.

Data Augmentation

Defined as a strategy to artificially increase the variety of input instances for training phase, without really collecting new instances [28], for to enhance the performance of

Convolutional Neural Networks (CNNs), it is essential to provide a large and diverse dataset. However, in many real-world scenarios, especially in medical imaging, the availability of labeled data is limited. To address this, data augmentation techniques are applied to artificially expand the training dataset.

Figure [6] shows examples of original histopathological images (top row) and their augmented versions (bottom row). The augmentation techniques used include rotation (e.g., 90°, 180°, 270°) and reflection (horizontal and vertical flips). These transformations preserve the semantic content of the image while presenting it in different orientations, allowing the model to become invariant to position and direction.This process describes:

- Increases the size of the training dataset,
- Reduces the risk of over fitting,
- Improves generalization capability,
- And helps in addressing class imbalance.

These can be a simple yet effective transformations that play a crucial role in improving model robustness, especially in domains where data collection is expensive or time-consuming see Figure 6.

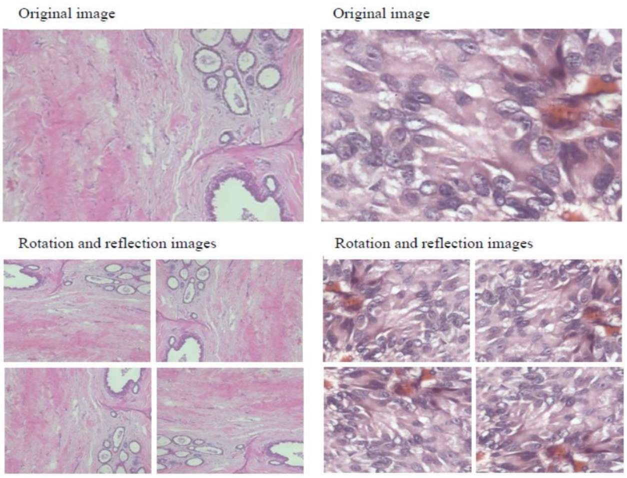

Figure 6. Illustration of Data Augmentation

CNN Pre-trained Models

On this area in the research,they study employed two CNN pre-trained convolution neural networksdifferentiates between healthy tissues and cancerous samples [29]. These networks were ResNet-50 and Inception-V3. Pre-trained CNN will be used to extract learned image features and then use those features to train a classifier.

Resnet-50: The concept architecture of ResNet- [30],is a convolutional neural network that is trained on more than a million images from the Image Net database. The network is consisting of 177 layers and can classify images into 1000 object categories. The network has an image input size of 224-by-224.

Inception-v3: The Inception-V3Net model is a deep CNN that is trained directly on a low-configuration computer. It is fairly difficult to train, and it takes a longer time, up to several days. This problem is solved via transfer learning its [31], trained on more than a million images from the ImageNet database and can classify images into thousand object categories, the Inception-V3Net network has 48 deep layers and can classify images into 1000 object categories. The Inception-V3Net network has 89MB size and accepts inputs of size is 299-by-299.

Train SVM Classifier Using Pre-trained CNN Features

Extract the learned image features from a pre-trained CNN for feature extraction, the layer before the classification layer (called ‘fc1000’), was used to extract features by using the activation method. These features were then used to train and test the SVM classifier.

Transfer Learning:

Due to the complexity and variability of medical images especially histopathological image straining deep learning models from scratch can be challenging and computationally expensive. Transfer learning (TL) offers an efficient solution by allowing a model trained on a large, general dataset (e.g., ImageNet) to be adapted for a specific medical imaging task Transfer learning techniques have been used in classification, regression, and clustering problems [32].

Transfer learning leverages previously learned features such as edges, shapes, and textures from a source task and applies them to a target task with limited data. This approach significantly reduces the need for large labeled datasets and training time, while improving classification performance. Figure [5] illustrates the typical workflow of transfer learning:

1. Load Pretrained Network: A deep network trained on a large dataset (e.g., 1 million images across thousands of classes) are loaded. Early layers capture generic features like edges and colors.
2. Replace Final Layers: The final classification layers are replaced to suit the target task, which often has fewer classes. These new layers will learn task-specific features.
3. Train Network: The modified network is trained using the new dataset. Since only part of the network is retrained, the process is faster and requires fewer images (hundreds rather than millions).
4. Predict and Assess: The trained network is evaluated on unseen data to assess its accuracy and generalization. If necessary, further fine-tuning can be applied to improve performance.

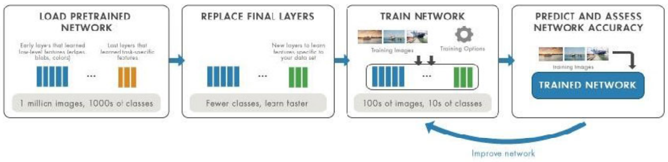

Figure 7. Transfer Learning Workflow

Evaluation and Metrics

Confusion Matrixes: A confusion matrix is a fundamental evaluation tool in classification tasks, providing a detailed breakdown of the model's performance by comparing actual and predicted class labels and been more accurate and eliminated a lot of confusion[33],it allows the identification of classification errors and helps compute several important performance metrics such as accuracy, sensitivity, specificity, and precision. This matrix is particularly useful in imbalanced datasets were relying solely on accuracy may give misleading results.Table1illustrating the relationship between actual and predicted classifications. It includes the counts of true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP), which are used to evaluate classification performance metrics.

Table 1. Confusion Matrixes Details

	Predicted positive	Predicted negative	Total
Actual positive	TP	FN	AP
Actual negative	FP	TN	AN
Total	PP	PN	N

The basic terms to be interpreted from a confusion matrix are:

- True Positive (TP): Observation is positive and predicted to be positive.
- False Negative (FN): Observation is positive but predicted negative.
- True Negative (TN): Observation is negative and predicted to be negative.
- False Positive (FP): Observation is negative but predicted positive.

The proportion of actual positives samples that correctly is given by Sensitivity measures is calculated as:

Sensitivity (Recall): Sensitivity (Recall) measures the model's ability to correctly identify malignant breast cancer cases. It is especially important in medical diagnosis to minimize false negatives. Using CNNs with transfer learning and SVM classifiers, combined with data augmentation, helped improve sensitivity. It is calculated as:

Recall is defined as the proportion of true positives among the actual positives. It measures the model's ability to correctly identify all relevant positive instances (cancerous cells). Recall is critical in medical applications, as a higher recall ensures that most actual cancerous cases are correctly identified, reducing the risk of missed diagnoses (false negatives). Using CNNs with transfer learning and SVM classifiers, combined with data augmentation, helped improve sensitivity. It is calculated as. (Eswar, Dinesh et al.)

$$\text{Recall (Sensitivity)} = \frac{(\text{TP})}{(\text{TP} + \text{FN})}$$

High sensitivity in our results indicates the model effectively detects most cancerous samples, supporting accurate diagnosis.

Specificity: Specificity evaluates the model's ability to correctly identify non-cancerous (benign) cases by minimizing false positives. This is crucial to avoid unnecessary stress and treatment for healthy patients. The combination of CNN-based feature extraction, transfer learning, and SVM classification helped achieve reliable specificity. The ratio of correctly detected negative observations to all the negative cases this metric represents the correctly classified rate of opposite disease classes [34], it is calculated as:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

High specificity in our experiments confirms the model's accuracy in distinguishing benign tissue from malignant.

Positive predictive value or Precision: Precision (Positive Predictive Value) measures how many of the cases predicted as malignant are truly cancerous. In other words, positive predictive value answers the question, 'if I have a positive test, what is the probability that I actually have the disease [35]. This is important to reduce false alarms and ensure the reliability of positive diagnoses. By using CNNs with transfer learning and SVM classifiers, we improved precision alongside other metrics. It is calculated as:

$$\text{Precision(PPV)} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

High precision in our results reflects the model's ability to make accurate positive predictions with minimal false positives.

F1-score: F1 score. This metric is the most used member of the parametric family of the F-measures, named after the parameter value β=1. F1 score is defined as the harmonic mean of precision and recall [36]. It is particularly useful in medical diagnosis where both false positives and false negatives carry significant consequences. The F1-score is calculated as:

$$\text{F1\_ Score} = \frac{2\ (\text{Precision. recall})}{(\text{Precision} + \text{recall})}$$

Using CNN-based feature extraction with transfer learning and SVM classification, our approach achieved strong F1-scores, reflecting robust and reliable classification performance across both benign and malignant classes.

Accuracy: The accuracy metric is defined as the percentage of correctly classified images out of the total number of images in the test set. more frequently and still achieve high accuracy (Eswar, Dinesh et al.)Accuracy reflects the overall effectiveness of the classification model in correctly identifying both benign and malignant cases. It measures the proportion of true results (both true positives and true negatives) among the total number of predictions. Accuracy is calculated as:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

By integrating CNNs with transferlearning and SVM classifiers, our approach achieved high accuracy, demonstrating its reliability in breast cancer diagnosis.

The Area Under the Curve(AUC): AUC is one of several important pharmacokinetic terms that are used to describe and quantify aspects of the profile of an administered drug (and/or its metabolites, which may or may not be pharmacologically active themselves). [37], is used to evaluate the classifier's ability to distinguish between benign and malignant tumors. Unlike accuracy, AUC provides a more comprehensive measure of model performance, especially in imbalanced datasets. The AUC value ranges from 0 to 1, with higher values indicating better discrimination capability. A larger AUC implies stronger ranking performance, making it a reliable metric for assessing the effectiveness of our CNN-SVM classification framework.

Table 2. Comparison of results with approaches in metric measurements

	Sensitivity	Specificity	Precision	F1-Score	Accuracy
Resnet50+ SVM	98.7%	91.8%	91.2%	94.8%	94.9%
TransferLearning+Resnet50	98.6%	94.1%	97.8%	98.1%	97.2%
InceptionV3+ SVM	91.2%	93.4%	93.6%	92.4%	92.3%
TransferLearning +InceptionV3	96.9%	95.7%	98.4%	97.6%	96.6%

RESULT AND DISCUSSION

The performance measurements of all the approaches (InceptionV3+SVM, ResNet-50+SVM, transfer learning from Resnet50, and transfer learning from InceptionV3):

Table 2.Presents a comparative analysis of four classification approachesResNet-50+SVM, TransferLearning with ResNet-50, Inception V3+SVM, and TransferLearning with InceptionV3using key performance metrics. Among the models, TransferLearning with ResNet-50 achieved the highest F1-Score (98.1%) and Precision (97.8%), with an overall Accuracy of 97.2%, indicating excellent balance between sensitivity and specificity. Similarly, TransferLearning with InceptionV3 also demonstrated strong performance, with an F1-score of 97.6% and accuracy of 96.6%. In comparison, ResNet-50+SVM achieved a slightly lower accuracy of 94.9%, while InceptionV3+SVM yielded the lowest accuracy (92.3%) among the tested approaches. These results highlight the effectiveness of integrating transfer learning and SVMs for robust histopathological image classification.

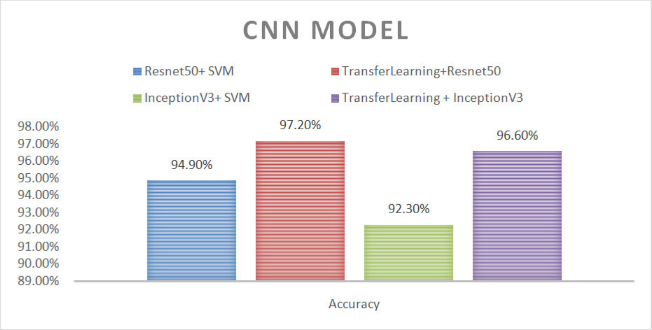

Figure 8. Depicted accuracy for CNN models

Figure 8. Illustrates the accuracy of various CNN models. Among the models evaluated, the TransferLearning + ResNet50 model achieved the highest accuracy at 97.20%. This suggests that transfer learning from pre-trained CNN models outperforms combinations of pre-trained CNNs with SVM classifiers.

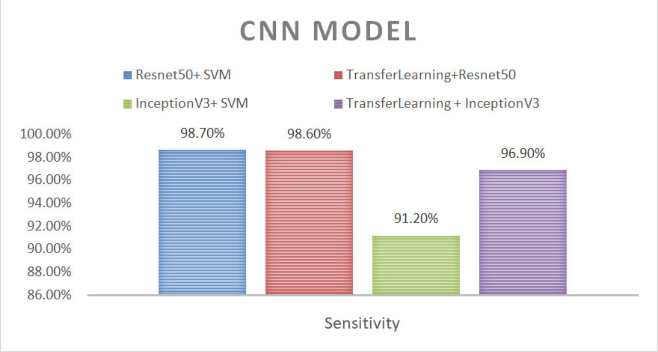

Figure 9. Depicted sensitivity for CNN models

Moreover, TransferLearning + ResNet50 also surpassed the TransferLearning + InceptionV3 model (96.60%) in performance. This may be attributed to the symmetry in the dataset, specifically the large representation of cancer-positive cases. Accuracy, as a metric, reflects the model's ability to correctly identify both true positives and true negatives, which is crucial in medical diagnosis tasks. Figure 9. Presents the sensitivity results of different CNN models. The ResNet50 + SVM and TransferLearning + ResNet50 models demonstrated the highest sensitivity, achieving 98.70% and 98.60%, respectively. These results indicate that ResNet50-based architectures are more effective at correctly identifying true positive cases compared to the InceptionV3-based models. Notably, InceptionV3 + SVM showed the lowest sensitivity at 91.20%, whereas TransferLearning + InceptionV3 reached 96.90%, highlighting that applying transfer learning significantly enhances the model's ability to detect positive cases.

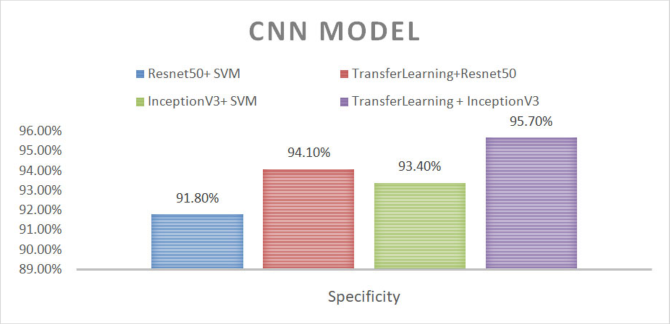

Figure 10. Depicted specificity for CNN models

Figure 10. Illustrates the specificity of the various CNN models. Among the evaluated models, TransferLearning + InceptionV3 achieved the highest specificity at 95.70%, indicating its superior ability to correctly identify true negative cases (i.e., non-cancerous instances). This performance surpasses both pre-trained CNN + SVM combinations and the TransferLearning + ResNet50 model (94.10%). The ResNet50 + SVM and InceptionV3 + SVM models yielded lower specificity values of 91.80% and 93.40%, respectively, emphasizing that transferlearning not only improves sensitivity but also enhances the model's capability to avoid false positives.

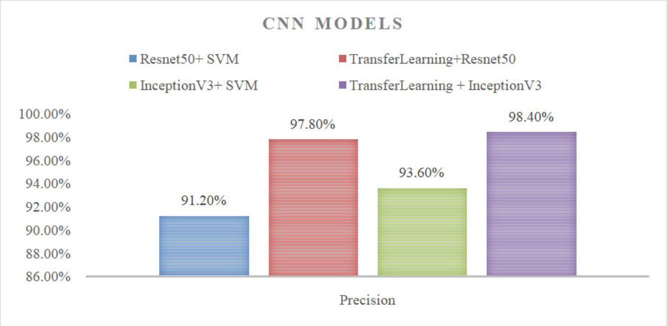

Figure 11. Depicted precision for CNN models

Precision for all models are depicted in Figure9, we can observe that the TransferLearning+Resnet50 and InceptionV3 are obtained the highest precision in comparison with pre-trained CNN+SVM models this because precision assesses competence at identifying actual positives.

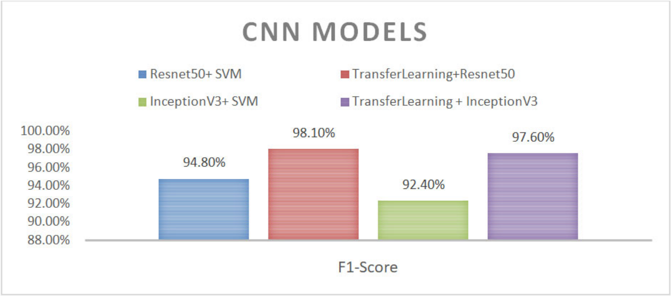

Figure12. Depicted f1-score for CNN models

Figure 12. Presents the F1-scores of the evaluated CNN models. Among all, the TransferLearning+ResNet50 model achieved the highest F1-score of 98.10%. This superior performance can be attributed to the presence of class imbalance in the dataset, where a significant portion of samples belongs to the positive class. In such scenarios, the F1-scorebeing the harmonic mean of precision and recalloffers a more balanced assessment of the model's effectiveness. The high F1-score of TransferLearning+ResNet50 indicates its strong capability to maintain both high precision and high recall, even under skewed class distributions.

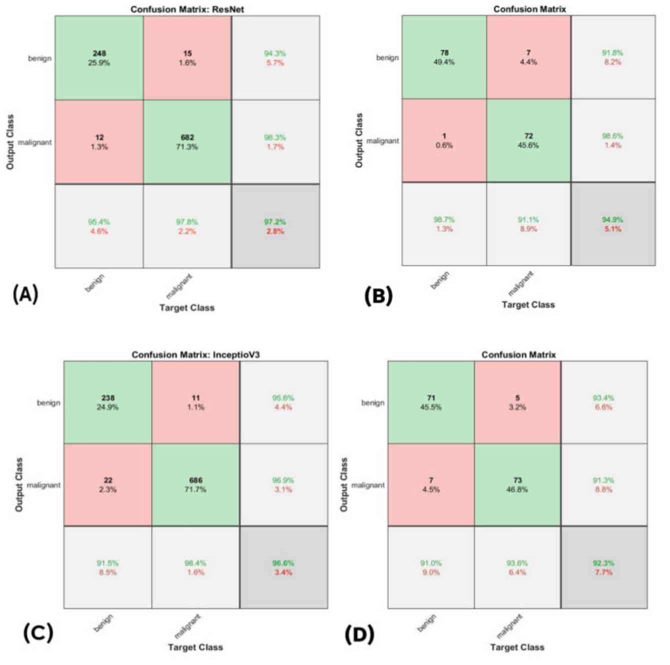

Figure 13: Confusion Matrixes (A)TransferLearning+Resnet50(B) TransferLearning+ InceptionV3, (C) Resnet50+SVM and (D) InceptionV3+SVM

Figure 13.Shows confusion matrices for four models. (A) Transfer Learning with ResNet50 achieved high accuracy (97.2%) with low misclassification (2.8%). (B) Transfer Learning with InceptionV3 had the best malignant class accuracy (98.6%) and the fewest errors. (C) ResNet50 + SVM also performed well (96.6% accuracy), while (D) InceptionV3 + SVM showed slightly higher misclassification (7.7%). Overall, Transfer Learning models outperformed SVM-based ones.

Conclusion and Future work

This paper presents a powerful and efficient framework for breast cancer classification using convolutional neural networks (CNNs) with transfer learning, further strengthened by Support Vector Machine (SVM) classifiers. Through extensive experimentation on the BreakHis histopathological image dataset, the model leveraging Transfer Learning with ResNet50 achieved the highest accuracy (97.2%) and F1-score (98.1%), demonstrating its robustness in differentiating between benign and malignant tumors. The high sensitivity and specificity metrics confirm its clinical reliability. Importantly, the integration of data augmentation techniques improved the model's generalization and addressed data scarcity a common challenge in medical imaging. Looking beyond theoretical validation, this model holds significant potential for real-world deployment, especially in hospitals and diagnostic centers with limited access to experienced pathologists. For instance, it could be embedded in digital pathology platforms to assist with rapid screening or integrated into mobile diagnostic applications for remote or underserved areas. Its fast, automated, and accurate diagnostic capability can greatly reduce diagnostic delays and improve treatment planning. As healthcare moves toward AI-assisted systems, the proposed model offers a practical, scalable, and cost-effective solution to support early breast cancer detection and reduce mortality rates globally. Future work should focus on expanding the dataset diversity, exploring additional diseases like pneumonia and brain cancer, and incorporating explainable AI tools to improve transparency, making the system more adaptable, interpretable, and ready for clinical deployment.

REFERENCES

1. Aghaei, M. (2023). "The Breast Cancer, a Global Health Crisis." International Journal of New Findings in Health and Educational Sciences (IJHES)1(2): 79-87.
2. Karagoz, A., H. Tombuloglu, M. Alsaeed, G. Tombuloglu, A. A. AlRubaish, A. Mahmoud, S. Smajlović, S. Ćordić, A. A. Rabaan and E. Alsuhaimi (2023). "Monkeypox (mpox) virus: Classification, origin, transmission, genome organization, antiviral drugs, and molecular diagnosis." Journal of infection and public health 16(4): 531-541.
3. Arnold, M., E. Morgan, H. Rumgay, A. Mafra, D. Singh, M. Laversanne, J. Vignat, J. R. Gralow, F. Cardoso and S. Siesling (2022). "Current and future burden of breast cancer: Global statistics for 2020 and 2040." The Breast66: 15-23.
4. Pallua, J., A. Brunner, B. Zelger, M. Schirmer and J. Haybaeck (2020). "The future of pathology is digital." Pathology-research and practice 216(9): 153040.
5. Crosby, D., S. Bhatia, K. M. Brindle, L. M. Coussens, C. Dive, M. Emberton, S. Esener, R. C. Fitzgerald, S. S. Gambhir and P. Kuhn (2022). "Early detection of cancer." Science 375(6586): eaay9040.
6. Gao, Y. e., J. Lin, Y. Zhou and R. Lin (2023). "The application of traditional machine learning and deep learning techniques in mammography: a review." Frontiers in Oncology 13: 1213045.
7. Mezei, T., M. Kolcsár, A. Joó and S. Gurzu (2024). "Image Analysis in Histopathology and Cytopathology: From Early Days to Current Perspectives."Journal of Imaging 10(10): 252.
8. Cheng, J. Y., J. T. Abel, U. G. Balis, D. S. McClintock and L. Pantanowitz (2021). "Challenges in the development, deployment, and regulation of artificial intelligence in

anatomic pathology." The American Journal of Pathology 191(10): 1684-1692.

9. Guo, Z., J. Xie, Y. Wan, M. Zhang, L. Qiao, J. Yu, S. Chen, B. Li and Y. Yao (2022). "A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis." Open Life Sciences 17(1): 1600-1611.
10. O'shea, K. and R. Nash (2015). "An introduction to convolutional neural networks." arXiv preprint arXiv:1511.08458.
11. Kattenborn, T., J. Leitloff, F. Schiefer and S. Hinz (2021). "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing." ISPRS journal of photogrammetry and remote sensing173: 24-49.
12. Khan, A., A. Sohail, U. Zahoora and A. S. Qureshi (2020). "A survey of the recent architectures of deep convolutional neural networks." Artificial intelligence review 53: 5455-5516.
13. Alzubaidi, L., J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie and L. Farhan (2021). "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." Journal of big Data 8: 1-74.
14. Xia, X., C. Xu and B. Nan (2017). Inception-v3 for flower classification. 2017 2nd international conference on image, vision and computing (ICIVC), IEEE.
15. Salehi, A. W., S. Khan, G. Gupta, B. I. Alabduallah, A. Almjally, H. Alsolai, T. Siddiqui and A. Mellit (2023). "A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope." Sustainability 15(7): 5930.
16. Abdou, M. A. (2022). "Literature review: Efficient deep neural networks techniques for medical image analysis." Neural Computing and Applications34(8): 5791-5812.
17. Li, Q., W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen (2014). Medical image classification with convolutional neural network. 2014 13th international conference on control automation robotics & vision (ICARCV), IEEE.
18. Oza, P., P. Sharma, S. Patel, F. Adedoyin and A. Bruno (2022). "Image augmentation techniques for mammogram analysis." Journal of Imaging 8(5): 141.
19. Aggarwal, R., V. Sounderajah, G. Martin, D. S. Ting, A. Karthikesalingam, D. King, H. Ashrafian and A. Darzi (2021). "Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis." NPJ digital medicine 4(1): 65.
20. Aljuaid, H., N. Alturki, N. Alsubaie, L. Cavallaro and A. Liotta (2022). "Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning." Computer Methods and Programs in Biomedicine 223: 106951.
21. Masud, M., A. E. Eldin Rashed and M. S. Hossain (2022). "Convolutional neural network-based models for diagnosis of breast cancer." Neural Computing and Applications 34(14): 11383-11394.
22. Abunasser, B. S., M. R. J. Al-Hiealy, I. S. Zaqout and S. S. Abu-Naser (2023). "Convolution neural network for breast cancer detection and classification using deep learning." Asian Pacific journal of cancer prevention: APJCP 24(2): 531.
23. Ali, M. D., A. Saleem, H. Elahi, M. A. Khan, M. I. Khan, M. M. Yaqoob, U. Farooq Khattak and A. Al-Rasheed (2023). "Breast cancer classification through meta-learning ensemble technique using convolution neural networks." Diagnostics 13(13): 2242.
24. Koshy, S. S. and L. J. Anbarasi (2024). "LMHistNet: Levenberg–Marquardt based deep neural network for classification of breast cancer histopathological images." IEEE Access.
25. Hasan, A. M., N. K. Al-Waely, H. K. Aljobouri, H. A. Jalab, R. W. Ibrahim and F. Meziane (2025). "Corrigendum to “Molecular subtypes classification of breast cancer in DCE-MRI using deep features”[Expert Syst. Appl. 236 (2024) 121371]." Expert Systems with Applications: 126763.
26. Kalusivalingam, A. K., D. Chopra, P. Nair, V. Gupta and V. Sharma (2024). "Enhancing Cancer Detection in MRI Scans Using Transfer Learning and Data Augmentation with Convolutional Neural Networks." European Advanced AI Journal 13(3).
27. Pfob, A., S.-C. Lu and C. Sidey-Gibbons (2022). "Machine learning in medicine: a practical introduction to techniques for data pre-processing, hyperparameter tuning, and model comparison."BMC medical research methodology 22(1): 282.
28. Catak, F. O., J. Ahmed, K. Sahinbas and Z. H. Khand (2021). "Data augmentation based malware detection using convolutional neural networks." Peerj computer science7: e346.
29. Gupta, K. and N. Chawla (2020). "Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained CNN." Procedia Computer Science167: 878-889.
30. Chiet, C. C., K. W. How, P. Y. Han and Y. H. Yen (2024). "A lung cancer detection with pre-trained CNN models." Journal of Informatics and Web Engineering 3(1): 41-54.
31. Mujahid, M., F. Rustam, R. Álvarez, J. Luis Vidal Mazón, I.d.l.T. Díez and I. Ashraf (2022). "Pneumonia classification from X-ray images with inception-V3 and convolutional neural network." Diagnostics12(5): 1280.
32. Tsalera, E., A. Papadakis and M. Samarakou (2021). "Comparison of pre-trained CNNs for audio classification using transfer learning." Journal of Sensor and Actuator Networks 10(4): 72.
33. Sathyanarayanan, S. and B. R. Tantri (2024). "Confusion matrix-based performance evaluation metrics." African Journal of Biomedical Research: 4023-4031.
34. Hosseini, A., M. A. Eshraghi, T. Taami, H. Sadeghsalehi, Z. Hoseinzadeh, M. Ghaderzadeh and M. Rafiee (2023). "A mobile application based on efficient lightweight CNN model for classification of B-ALL cancer from non-cancerous cells: A design and implementation study." Informatics in Medicine Unlocked39: 101244.
35. Leal Filho, W., P. C. C. Ribeiro, J. Mazutti, A. Lange Salvia, C. Bonato Marcolin, J. M. Lima Silva Borsatto, A. Sharifi, J. Sierra, J. Luetz and R. Pretorius (2024). "Using artificial intelligence to implement the UN sustainable development goals at higher education institutions." International Journal of Sustainable Development & World Ecology31(6): 726-745.
36. Chicco, D. and G. Jurman (2020). "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation." BMC genomics21: 1-13.
37. Turner, J. R. (2020). "Area under the curve (AUC)." Encyclopedia of Behavioral Medicine: 146-146.

****