Manasa Sandeep, 2025, 13:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

A Review on Deep Learning Based Breast Cancer Classification for Histopathology Images

Assistant professor Manasa Sandeep ,Dr. C Nandini,Bhargavi S.R,Disha.A,Fida.K.S,Harshitha.B.K
Dayananda Sagar Academy of Technology and Management

Abstract- Breast cancer is one of the most common and life-threatening cancers in women worldwide. The clinical gold standard for diagnosis is still histopathological examination, but it is time-consuming, subject to human expertise, and susceptible to human error. This project presents a deep learning system based on DenseNet201 architecture for automated and enhanced accuracy of breast cancer detection from histopathology images. The system is developed using the BreaKHis dataset, employing state-of-the-art preprocessing and data augmentation methods for enhancing robustness. Performance metrics such as accuracy, precision, recall, and AUC-ROC results reflect the system's performance as a sound diagnostic tool in clinical settings. Histopathological diagnosis, while critical, entails a number of challenges including inter-observer variability, workload burden on pathologists, and risk of delayed treatment decisions. Convolutional Neural Networks (CNNs), specifically DenseNet201, have proven to be useful tools for extracting complex visual patterns in medical images. In this research, transfer learning, reuse of features, and a well- designed classification pipeline are utilized to separate benign from malignant samples successfully. The application of artificial intelligence to pathology is not just a means of improving diagnostic correctness but also of broadening access to healthcare through making sound diagnostics available in low-resource environments. By providing a speedy and reproducible second opinion, the model described here is an advance toward real-time, AI-augmented cancer diagnosis that can revolutionize conventional clinical practice.

Keywords- Breast Cancer Detection,Histopathology Images,Deep Learning,DenseNet201,Convolutional Neural Networks (CNNs),Transfer Learning.

I. INTRODUCTION

Diagnosis of breast cancer is the key to minimizing mortality and enhancing patient outcomes. With millions of new cases every year, early and correct detection is still of paramount importance. Traditional diagnostic tools—mammography, ultrasound, and biopsy—are dependent on skilled interpretation, susceptible to human error, and commonly lead to delays in diagnosis. Of these, histopathological examination of images most definitive yet also the most time-consuming. Artificial intelligence, especially deep learning, has indicated promising potential in automatic diagnosis. Convolutional Neural Networks (CNNs) have proven to be highly effective in medical image analysis by learning hierarchical features at deep layers. DenseNet201, a top-performing CNN structure, provides fast feature propagation, solves the vanishing gradient issue, and minimizes the number of parameters, which makes it appropriate for medical image tasks.

This project is aimed at constructing a DenseNet201-based model to classify breast histopathology images as benign or malignant. Through the application of preprocessing methods, transfer learning, and a solid evaluation approach, the model should offer a trustworthy diagnostic device that improves clinical work processes. The general goal is to be part of the current initiative to incorporate AI into actual medical diagnostics in order to speed up, make consistent, and enable cancer detection atscale.

II. LITERATURE SURVEY

Mehmet Gül (2025) proposed a twofold methodology using a tailored 20-layer CNN and a new Local Binary Pattern method called quad-star LBP (QS-LBP) for histopathology image classification toward breast cancer diagnosis. The research utilized both the BreaKHis dataset and a large proprietary dataset with 98.27% accuracy and 97% AUC using the CNN, whereas the QS-LBP with Random Forest provided 94.58% accuracy. The developed hybrid system performed better than traditional approaches and focused on the role of textural enhancement in early detection. In spite of the complicated architecture, the model demonstrated the ability for scalable and correct classification.

Komal S. Gandle and D. B. Kshirsagar (2024) performed a systematic review of methods for detecting breast cancer using histopathological images, with emphasis on segmentation, feature extraction, and classification techniques. The research highlighted the success of deep learning, especially CNNs and SVMs, in enhancing accuracy of diagnosis. It delineated that diagnosis at the early stages is the determinant of survival, whereas issues of human errors and misclassification remain. The article espouses computer-aided CAD systems with the use of sophisticated imaging methods such as liquid biopsy and molecular imaging to achieve improved results, noting that correct classification relies significantly on preprocessing and dataset quality.

Ahmed Omrane Meddas et al. (2024) presented comparative evaluation of Inception V3 and DenseNet201 architectures for binary classification of breast cancer based on histopathology images. The two models were trained under similar conditions on the BreaKHis dataset to compare fairness in performance. DenseNet201 surpassed Inception V3 on accuracy, precision, recall, and F1-score measures, showing improved feature propagation and utilization of dense connections. The research concluded that architectural decisions play a major role in determining model reliability, particularly when trained with standardized protocols and datasets.

Farjana Parvin et al. (2023) suggested an ensemble solution based on AlexNet, ResNet-50, and Inception-v1 ensemble through majority voting for breast cancer histopathological image classification. Employing the BreaKHis dataset with four magnification levels, the ensemble model achieves up to 99.98% accuracy, which is better than that of individual CNNs. The approach showed that ensemble models with complementary strength lead to improved generalization and classification performance. Although the ensemble causes additional computational complexity, it greatly enhances reliability and robustness in binary classification.

Hossena Djouima et al. (2022) used DenseNet201 with a DCGAN-based augmentation pipeline to solve class imbalance in the BreaKHis dataset. Global Average Pooling was used to extract early DenseNet layer features and feed them into a softmax classifier. Accuracy achieved was 96% (40x), 95% (100x), 88% (200x), and 92% (400x) over magnification levels. The approach demonstrated the capability of GAN-based data synthesis to minimize bias and enhance classifier performance, although the accuracy fell at large magnifications, possibly a result of feature redundancy.

Sajiv G. and G. Ramkumar (2023) created a deep learning framework through Artificial Neural Networks (ANN) for classifying breast cancer histopathological images. Using a dataset provided on Kaggle, the ANN model yielded 91.7% accuracy, suggesting that it could serve as an economical replacement for CNNs. The system was centered around feature extraction and classification but with a focus on preprocessing methods. Although performance was slightly less than in the case of deeper CNNs, the interpretability and simplicity of the ANN model make it a good choice to deploy in low-resource settings.

Rajendra Babu Chikkala et al. (2025) proposed a Bidirectional Recurrent Neural Network (BRNN) architecture that adopted ResNet50-based transfer learning, GRUs, and residual feature fusion mechanism. The network attained an accuracy of

97.25% on the multi-class BreaKHis data with eight class labels. It successfully employed spatial and sequential dependencies with attention mechanisms and Adagrad optimization. The paper showed that integrating temporal and spatial feature learning enhances interpretability and performance, particularly in the case of multi-class, albeit model complexity is still an issue for deployment.

Lang Wang et al. (2023) introduced LGViT, a Local-Global Vision Transformer framework combining CNN stems and a new self-attention mechanism for breast cancer image classification. The model used Local-Global Multi-head Self- attention and a Ghost Feed-forward Network to extract local texture as well as global context. Evaluated on the PatchCamelyon dataset, LGViT surpassed some CNN-based benchmarks. The architecture reduced computational overhead while improving feature learning, although transformer- based models are still sensitive to small data set tuning.

Feng He et al. (2022) presented MICNet, a multi-instance classification network proposed with VGG11 and visual explanation mechanisms for histopathological breast cancer diagnosis. It utilized weighted average pooling, mirror padding, and overlap cropping to preserve spatial detail with interpretability support. The model was more accurate and explainable compared to traditional CNNs on BreakHis and Camelyon16 datasets. This method gave clinicians explainable reasoning for predictions but consumed a lot of computational resources because of multiple instance learning.

Eshika Jain and Amanveer Singh (2024) employed a VGG16-based classifier for histopathological images of the BreakHis dataset. The model had an accuracy of 82.94% and an AUC- ROC of 0.87337 when using data augmentation and transfer learning.

Although good performance was obtained for classification, the large number of false negatives demonstrated scope for improvement. The research demonstrated that traditional CNNs such as VGG16 are still competitive with proper improvements but could fall behind newer models in precision and recall.

III. METHODOLOGY

The methodology followed involves the following key steps:

Dataset: The BreaKHis dataset is utilized, which includes thousands of histopathological images with four magnifications (40x, 100x, 200x, 400x) labeled as benign or malignant.

Preprocessing: Images are resized to 224 pixels by 224 pixels and normalized. Data augmentation methods such as flipping, rotation, and zoom are used to improve variability and counteract class imbalance.

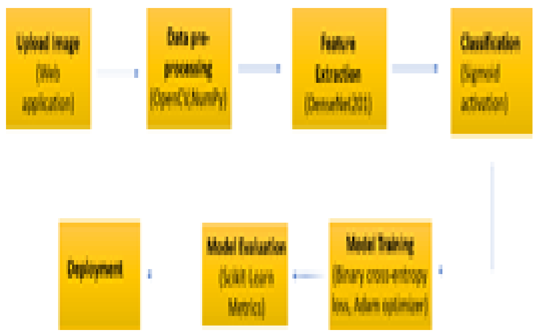
Model Architecture: The base model is DenseNet201, pretrained on ImageNet. Features are extracted from several dense blocks and then fed through a global average pooling layer. A fully connected layer with sigmoid activation does the binary classification.

Training Strategy: Binary Cross-Entropy loss and Adam optimizer are utilized to train the model. Early stopping on validation loss is used to prevent overfitting. Transfer learning is employed by freezing early layers and fine-tuning deep ones.

Evaluation Metrics: The performance of the model is measured based on accuracy, precision, recall, F1-score, and AUC-ROC. They give an overall idea about the capability of the classifier to separate benign and malignant samples.

Deployment: A Flask web application lets users upload histopathology images and obtain real-time predictions and confidence scores.

IV.SYSTEM ARCHITECTURE



V. FUTURE ADVANCEMENTS

This project now only deals with binary classification; however, being extended to multi-class classification to handle subtypes like ductal, lobular, and mucinous carcinoma would be more clinically useful. Adding metadata (like patient history or biomarker profiles) could add further personalization and context to the diagnostic output.

Explainable AI will become increasingly important for medical adoption. Incorporating methods such as Grad-CAM or attention maps may facilitate visualizing the area of concentration of the model, which will instill confidence among clinicians.

Further, deployment into hospital information systems and cloud access may allow for large-scale, real-time diagnostics, particularly in underserved areas.

VI. CONCLUSION

The paper introduces a DenseNet201-derived deep learning system for automated histopathology image-based breast cancer classification. The system is accurate and consistent, reflecting its utility as a decision-support tool in the clinic. By automating tissue classification and reducing diagnostic latency, the system marks a next step in AI-aided pathology, with significant implications for scalable, accurate, and interpretable breast cancer diagnosis.

REFERENCES

1. Gül, Mehmet, A Novel Local Binary Patterns- Based Approach and Proposed CNN Model to Diagnose Breast Cancer by Analyzing Histopathology Images (March 7, 2025).
2. Gandle, Komal S. and Kshirsagar, D. B., Breast Cancer Categories, Analysis, Detection: Systematic Review for Histopathological Images (2024).
3. Meddas, Ahmed Omrane and Jabri, Dalel and Belkhiat, Djamel Eddine Chouaib, Breast Cancer Classification on Histopathological Images Using Inception V3 and DenseNet 201: A Comparative Study (2024).
4. Parvin, Farjana and Hasan, Md. Al Mehedi and Ahmed, Boshir and Mamun, Md. Al and Parvej, S.
5. M. Kausar, Breast Cancer Histopathological Image Classification Using an Ensemble of Deep Convolutional Neural Networks (2023).
6. [Djouima, Hossena and Zitouni, Athmane and Sbaa, Salim and Megherbi, Ahmed Chaouki, Classification of Breast Cancer Histopathological Images using DensNet201 (2022).
7. Sajiv, G. and Ramkumar, G., Deep Learning based Breast Cancer Classification Using Artificial Neural Network on Histopathological Images (2023).
8. Wang, Lang and Liu, Juan and Jiang, Peng and Cao, Dehua and Pang, Baochuan, LGViT: Local- Global Vision Transformer for Breast Cancer Histopathological Image Classification (2023).
9. He, Feng and Zhu, Yuemin and Wang, Weibo and Nanding, Abiyasi and Kuai,

Zixiang and Li, Xiaomei and Liu, Zhengjun, Multi-Instance

10. Classification of Histopathological Breast Cancer Images with Visual Explanation (2022).
11. [9] Jain, Eshika and Singh, Amanveer, Revolutionizing Breast Cancer Diagnosis: VGG16's Breakthrough in Histopathological Image Classification.