Article
JOURNAL OF UNIVERSITY OF BABYLON
For Pure and Applied Sciences (JUBPAS)
Vol. ; No. | 2025

Deep Learning-Based Approaches for Breast Cancer Detection and Diagnosis

Sura Jasim Mohammed
College of Information Technology, University of Babylon,
sura.phd2018@uobabylon.edu.iq , Babel, Hilla, Iraq.

أساليب التعلم العميق لاكتشاف وتشخيص سرطان الثدي

سرى جاسم محمد
1 قسم البرمجيات، كلية تكنولوجيا المعلومات جامعة بابل، بابل، العراق

Accepted: x/x/2025 **Published:** x/x/2025

ABSTRACT

Background:

Breast cancer is a disease that develops in breast tissue as a result of internal and external factors, such as genetic mutations, lifestyle, environmental exposure, and psychological and social factors. The overall incidence of breast cancer ranges from 5% to 10%, 20% to 30%, respectively.

Materials and Methods:

Breast masses, classified as malignant or benign, were identified using ultrasound images. Pre-processing included resizing images to 150x150 pixels, contrast normalization, data augmentation, and resampling to address class imbalance. A pre-trained EfficientNetB7 model was used for binary classification. The dataset was split into training (80%) and testing (20%). Model performance was validated using the "Multi Cancer Dataset" from Kaggle, which contains 2,191 images.

Results:

The model performed exceptionally well in distinguishing malignant and benign cases. For malignant cases, it achieved 0.98 accuracy, 0.99 sensitivity, and 0.99 F1-score. For benign cases, accuracy was 0.92, with sensitivity and F1-score of 0.87 and 0.89, respectively. The overall accuracy, sensitivity, and F1-score were 0.98. The ROC curve showed an AUC value above 0.9, indicating high sensitivity and a low false-positive rate. On external datasets, the model maintained an accuracy of 0.94.

Conclusion:

The results indicate that pre-trained deep convolutional neural networks can make significant progress in accurately detecting malignant cases while reducing false alarms, enhancing its clinical application and role in improving breast cancer diagnosis using AI techniques.

Key words:

Breast cancer; deep learning; transfer leaning; pre-trained model; convolutional neural networks.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

INTRODUCTION

Breast cancer is one of the most frequent malignant tumors affecting women worldwide. It originates from breast cells due to a combination of internal and external factors, including genetic mutations, lifestyle choices, environmental exposure, and social-psychological influences. Statistically, 5%–10% of breast cancer cases are attributed to genetic mutations and family history, while 20%–30% stem from modifiable factors, such as diet and physical [1]. The disease begins with abnormal changes in breast cells, which can lead to a cancerous tumor capable of invading nearby tissues and spreading throughout the body [2].

Breast cancer typically arises from the cells lining the milk ducts (ductal carcinoma) or the lobules responsible for milk production (lobular carcinoma). Both types can remain localized (in situ) or become invasive, spreading into surrounding tissues. Less common types include triple-negative breast cancer, inflammatory breast cancer, and Paget's disease of the breast [3].

The global prevalence of breast cancer underscores its significance as a major public health issue. It is the leading cause of cancer-related deaths in women, with approximately 2.09 million new cases and 630,000 deaths reported globally in 2018[4]. While the incidence varies regionally, it is particularly high in developed nations, such as the United States, Canada, and parts of Europe. Alarmingly, countries like China have seen a steady rise in breast cancer cases, driven by changes in lifestyle, diet, and population aging [5].

New diagnostic methods and early detection are some of the factors behind this improved prognosis in the patient. Unfortunately, it is usually asymptomatic and can be diagnosed mostly by chance during routine screenings [6].

Advanced tools for breast cancer classification and analysis, with high accuracy in classifying masses, as a precursory experiment on breast masses of an image in ultrasound lately demonstrated, is the deep transfer learning technique. This research supports these results by the utilization of a pre-trained transfer-learning model for medical image classification and breast cancer detection.

MATERIALS AND METHODS

Deep learning (DL) has brought significant improvement in the automatic detection of a mass in a digital mammogram image addressing the shortcomings of the previous work which was considerably manual. Different studies have utilized several proposed approaches that have contributed to the analysis and detection of lesions [7-42]. A summary of related works has been shown in table 1.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

Table 1. Summary of related works in the current study.

Method	Study	Objectives	Methods Used	Key Findings
Enhancing Mass Detection	[7,8]	High-accuracy mass detection	DL + CRF	Sensitivity 98% and Accuracy 90%
	[9]	Mass segmentation	FCN + CRF	Dice Score 0.97
	[10]	Rapid detection models	Faster R-CNN, Mask R-CNN	AUC 0.95, Sensitivity 97%
Improving Segmentation	[11,22]	Mass segmentation	Conditional GAN	Jaccard Index 0.89
	[37]	Segmentation enhancement	Unet3+, DeeplabV3+	Accuracy 90.99% and mean IU 52.57%
Transfer Learning	[18]	Breast density classification	AlexNet with Transfer Learning	AUC 0.9882
	[39]	Detection and classification using YOLO	YOLOv5 with Transfer Learning	mAP 0.835
	[29]	Classification improvement	GAN + ResNet-50	AUC 0.91
Enhancing Classification with CNN	[13]	Mass classification	CNN	Accuracy 90.5%
	[36]	Classification into 8 types	BCCNN + ResNet50	Accuracy 98.28% at 400X magnification
Addressing Data Scarcity	[31]	Tackling data scarcity	Synthetic Data Generation (GAN)	Improved performance with limited data

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

	[30]	Data augmentation	GAN-based DiaGRAM	AUC 0.9
Multi-Class Classification	[40]	Classification into multiple categories	ResNet50 and VGG16	High accuracy up to 98.28%
	[42]	Tissue risk assessment	CNN with customized features	AUC 0.8

Table 1 shows studies that have applied the prominent use of Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), or both, towards breast cancer detection. It reflects a comparable methodology shared between various research attempts. These advanced techniques have proven effective in overcoming issues like data scarcity, leveraging augmentation and synthetic data generation strategies to enhance dataset size and diversity.

DATASET SELECTION

The Breast Cancer Histological Image Classification or BreakHis_400X dataset comprises 1693 micrographs of breast tumor tissues obtained from 82 individual patients at magnifications of 40X, 100X, 200X, and 400X. Presently, it holds 547 samples of benign tissues and 1146 samples of malignant tissues (all 700x460 pixels, 3 RGB channels, 8 bits per channel—PNG). It is publicly available at: https://www.kaggle.com/datasets/scipygaurav/breakhis-400x-breast-cancer-dataset.This database was developed in partnership with the P&D Laboratory - Histopathology and Cytopathology in Paraná, Brazil. The BreaKHis dataset is divided into two categories: benign and malignant tumors. The term “histologically benign” refers to a lesion that does not meet any of the criteria for malignancy, such as significant cellular abnormalities, cell division, basement membrane rupture, metastases, etc. Benign tumors are typically “innocent,” slow growing, and limited. Malignant is the same as cancer: the lesion can infiltrate and destroy nearby structures (locally invasive) before spreading to distant sites (metastasis) and causing death [43].

In this section, we will discuss the most important studies that used the data set for conducting tests on the proposed systems [44-49]:

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Table 2. Summarizing the studies in the current study.

No.	Reference	Method	Aim	Models	Results
1	[44]	CNN-ANN model using the BreakHis_v1_400X dataset. Evaluated using metrics like accuracy, sensitivity, specificity, precision, F1-score, AUC, and ROC curves.	Enhance early breast cancer diagnosis and reduce global mortality by developing an automated detection system using histopathological images.	CNN-ANN	Accuracy: 89.47%, Precision: 86.18%, AUC: 89.46%, F1-score: 89.08%. Model outperformed individual components.
2	[45]	Transfer learning with a fine-tuned DenseNet201 model applied to the BreakHis dataset.	Improve accuracy and reliability in breast cancer diagnostics using deep learning techniques.	DenseNet201	Training Accuracy: 97.00%, Validation Accuracy: 92.00%. DenseNet201 demonstrated high efficacy in tumor classification.
3	[46]	Classification using CNNs with transfer learning on a hybrid dataset (BreakHis + Histo). Pretrained models (DenseNet201, ResNet50, ResNet101, MobileNet-v2).	Address limitations of single datasets and improve diagnostic accuracy through efficient histopathological image classification.	DenseNet201, ResNet50, ResNet101, MobileNet-v2	DenseNet201: Accuracy 91.37%, Sensitivity 100% (at 200x magnification), outperforming other models.
4	[47]	Hybrid deep transfer learning models: Xception with SVM (XSV) and Xception with RF (XRF).	Improve precision of breast cancer classification in histopathological images and	XSV (Xception + SVM), XRF (Xception + RF)	XSV: Accuracy up to 90.17% (BreakHis) and 87.35% (BHID). XRF:

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية جامعة بابل للعلوم الصرفة والتطبيقية

		Evaluated across multiple datasets and magnification levels.	compare hybrid vs. traditional models.		Accuracy up to 88.98% (BreakHis) and 87.29% (BHID). Both outperform traditional classifiers.
5	[48]	Custom CNN model with hyperparameter optimization (GWO and MGTO). Pretrained models like MobileNetV3, EfficientNetB0, VGG16, ResNet50V2 also used for comparison.	Develop a reliable and accurate breast tumor classification model using histopathological images and advanced optimization techniques.	Custom CNN, MobileNetV3, EfficientNetB0, VGG16, ResNet50V2	Custom CNN (optimized with MGTO): Accuracy 93.13% (10 iterations). Pretrained models: Accuracy between 74%-82%. Custom model significantly outperformed pretrained models.
6	[49]	Few-shot learning with a feature fusion strategy to handle limited labeled datasets. Tested on BreakHis and skin lesions datasets.	Overcome challenges of limited labeled medical images to improve tumor classification accuracy using few-shot learning techniques.	Few-shot learning with feature fusion	BreakHis: Accuracy 91.22%. Skin Lesions: Accuracy 71.20%. Achieved superior results with only 10 labeled samples compared to other state-of-the-art methods.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Undoubtedly, when examining the methodologies adopted in the aforementioned studies that target the same dataset, we observe numerous limitations and gaps. These limitations have contributed to the development of our methodology, which leverages the power of transfer learning through pre-trained deep learning algorithms for detecting cancerous tumors, specifically breast cancer, the focus of our research. Table3 below highlights the most notable gaps that can be observed.

Table 3. The most notable gaps in the current study.

No.	Strengths	Limitations or Challenges
1	-Combines CNN and ANN for improved performance. - Utilizes comprehensive evaluation metrics.	- Model complexity increases training challenges. - Lack of testing on external datasets. - No comparison with other strong models.
2	-Excellent performance with DenseNet201. - Transfer learning enhances results.	- Relies on a single dataset (BreakHis). - Risk of overfitting (97% training accuracy). - Lack of analysis on other influencing factors.
3	- Applied transfer learning with multiple pre-trained models. - Strong performance by DenseNet201.	- Reliance on specific magnification (200x). - Combining datasets (BreakHis and Histo) may introduce variability. - Focus on a single model.
4	- Hybrid models enhance classification accuracy. - Comparative analysis with traditional classifiers.	- Training complexity due to hybrid design. - Limited interpretability. - Overemphasis on accuracy metrics without practical implementation focus.
5	- Custom CNN design achieved strong performance. - Optimization using MGTO and GWO.	- Parameter optimization increases training complexity. - Unknown impact on training time. - No external dataset testing.
6	- Efficient learning from a small number of labeled samples. - Few-Shot Learning demonstrates good performance.	- Limited to very few samples, which may affect robustness. - Lack of practical implementation discussion. - Limited comparison with other few-shot methods.

The data preparation phase is a necessary process to prepare the data used in training the model to ensure its quality is improved and the different classes are balanced, which helps to enhance the model's performance and the reliability of its results. The process begins by defining the target classes, which are "malignant" and "benign", and adjusting the dimensions of the images to be a uniform size of 150 x 150 pixels, with the target number of samples for each class set to 3701 to avoid biasing the model towards the largest class. To achieve data enhancement, the

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

"ImageDataGenerator" library is used to apply enhancements such as rotation, panning, zooming, and distortion, in addition to horizontally flipping the images, with the aim of increasing the diversity of the data and enhancing the model's ability to generalize. Table 4 shows the preprocessing stage.

Table 4. The preprocessing stage in the current study.

The samples	preprocessing	
	Before	After
Benign	547	3071
Malignant	1146	3546
Total Samples	1693	6617

Table 4 shows the preprocessing stage of the dataset. The dataset was suffering from an imbalance in samples, which was 547 samples in the “benign” class compared to 1146 samples in the “malignant” class, which in turn affected the biased results and reduced the performance of the model. However, after performing the data augmentation technique, the size of the dataset increased from 1693 to 6617 samples, which ensures diversity and robustness by generating improved variations of the original images. Figure 1 shows the distribution of classes in the dataset after the preprocessing stage.

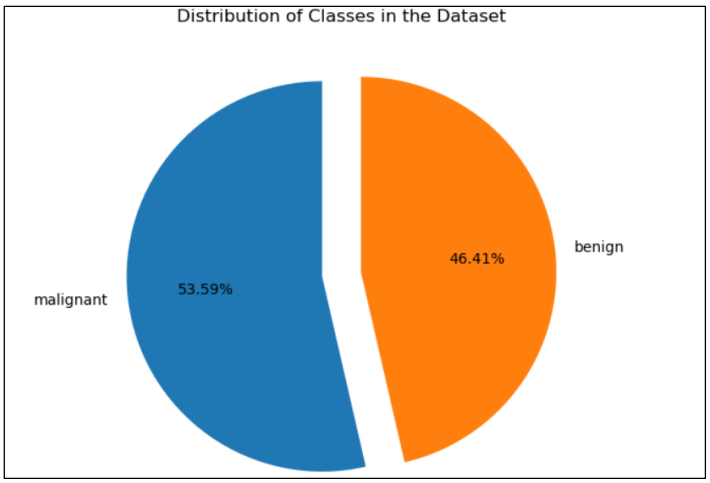

Figure 1. The samples of dataset in the current study.

Our proposed model is built using the pre-trained EfficientNetB7 model. EfficientNetB7 offers high accuracy and computational efficiency compared to traditional models such as VGG16, ResNet50, and MobileNet. It uses a Compound Scaling strategy that improves performance by balancing depth, width, and image resolution, making it more resource-efficient. EfficientNetB7 achieves higher accuracy (84.3% Top-1 on ImageNet) with fewer parameters (66 million),

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

reducing training time and computational consumption [50]. It also supports scaling across different versions (B0 to B7) to meet different performance needs, making it ideal for complex classification applications such as disease diagnosis and computer vision while maintaining execution speed and memory efficiency[51]. The dataset underwent a range of image preprocessing techniques, including image size reduction to 150 × 150 pixels to ensure computational efficiency. Data augmentation was then applied, which includes a combination of techniques such as rotation, zooming, and inversion to improve diversity and increase data size. Finally, data shuffling randomly shuffles the images to ensure the model does not learn from a specific order. At the end of the preprocessing stage, important features are extracted from the post-processed images and prepared for the training and testing stage. The last 20 layers of EfficientNetB7 were tuned to preserve previously learned features and reduce the risk of overfitting. Next, custom layers were added to improve model performance, including a Global Average Pooling (GAP) layer (2D), 50% and 30% dropout layers to reduce overfitting, 128- and 64-node dense layers with ReLU enabled, and a double Softmax output layer. Adam optimization was used to train the model at a relatively small learning rate, typically `1e-4`. This is a much more evolved technique as compared to Stochastic Gradient Descent or RMS prop, in a sense that it considers the advantages or the best practices of the other two optimizers in one single variant. Additionally, the imposed condition of early stopping with a patience of 10 will not only restore best weights, as an added measure to improve performance but also ensure that there is no overfitting. Table 5 shows the architecture of proposed model.

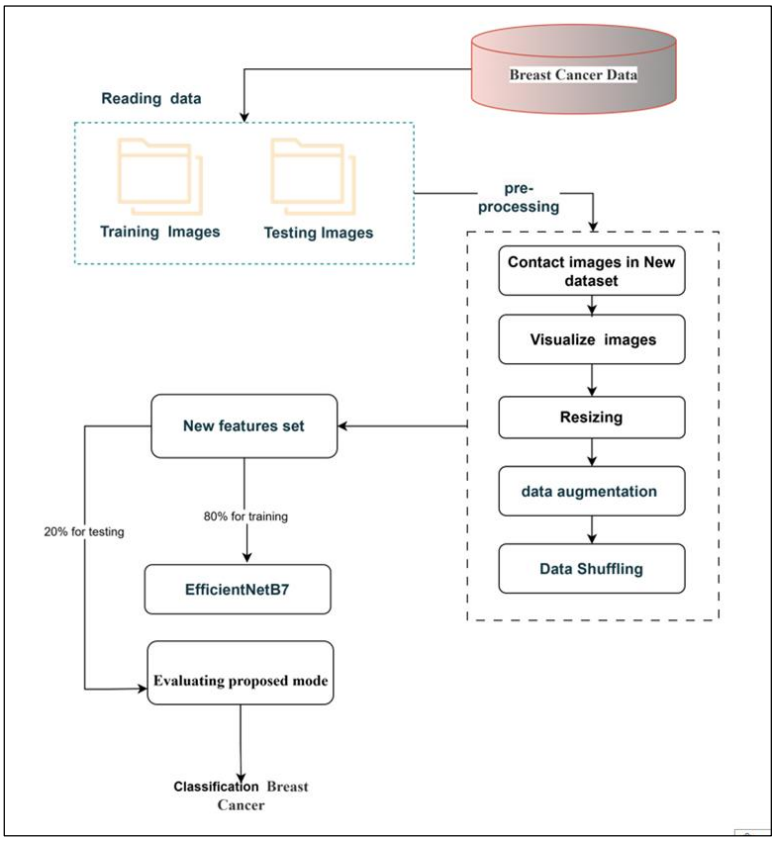

Figure 2. Basic stages of building the proposed model for skin cancer detection in the current study.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

Table 5. The architecture of the proposed model in the current study.

Component	Output	Parameters
EfficientNetB7	(None, 5, 5, 2560)	64,097,687
GlobalAveragePooling	(None, 2560)	0
Dropout 1	(None, 2560)	0
Dense 1	(None, 128)	327,808
Dense 2	(None, 64)	8,256
Dense 3	(None, 2)	130
Dropout 2	(None, 64)	0

The proposed model shown in figure 2 was trained on 80 epochs with 20 epochs and a batch size of 32 shown in Table 6, while the remaining 20% is used for testing. The trained model is evaluated to ensure its effectiveness, and the final output is the classification of breast cancer images into benign or malignant categories.

Table 6. The EfficientNetB7 model training in the current study.

Epoch	Accuracy	Loss	Val Accuracy	Val Loss	Time per Step (s)	Total Time (s)
1/20	0.5791	0.6342	0.9279	0.2085	15	1503
2/20	0.9602	0.1788	0.9647	0.1029	14	1157
3/20	0.9853	0.0624	0.9801	0.0566	14	1176
4/20	0.9918	0.0278	0.9770	0.0733	14	1167
5/20	0.9984	0.0142	0.9816	0.0628	15	1195
6/20	0.9974	0.0137	0.9755	0.0695	15	1196
7/20	0.9948	0.0137	0.9831	0.0561	16	1376
8/20	0.9958	0.0149	0.9877	0.0407	16	1322
9/20	0.9963	0.0122	0.9847	0.0612	16	1329
10/20	0.9983	0.0065	0.9739	0.1003	16	1293

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

11/20	0.9911	0.0257	0.9724	0.0949	16	1343
12/20	0.9954	0.0179	0.9785	0.0782	16	1335
13/20	0.9981	0.0073	0.9831	0.0572	16	1345
14/20	0.9977	0.0072	0.9755	0.0782	17	1427
15/20	0.9972	0.0076	0.9770	0.0703	16	1352
16/20	0.9935	0.0106	0.9770	0.0763	16	1342
17/20	0.9995	0.0036	0.9831	0.0789	16	1299
18/20	0.9999	0.0015	0.9770	0.0928	16	1290

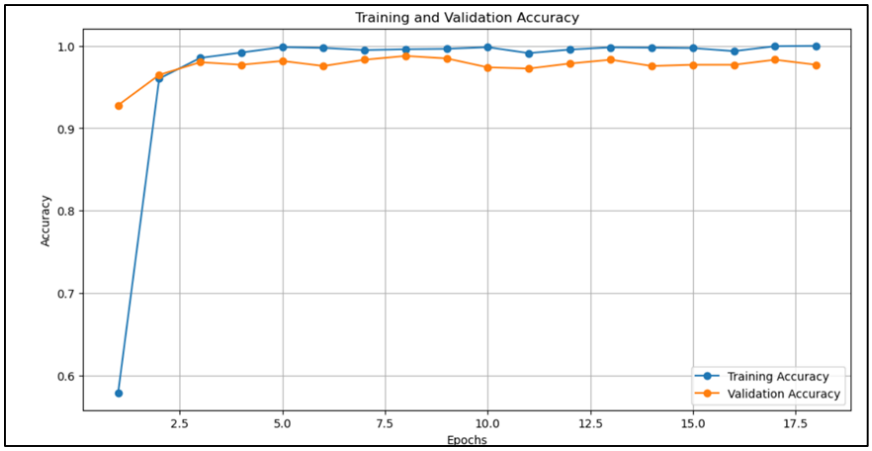

Figure 3. Training phase (accuracy curve) in the current study.

Figure 4. The training phase (loss curve) in the current study.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

Figures 3 and 4 illustrate the performance of the EfficientNetB7 model during the training phase. From the accuracy curve, the model starts with rather low accuracy—in this case around 60%, though by about the fifth epoch both the training and validation data get to stable high accuracy, which is about 97%, that is to say, learning without over-fitting. The loss curve indicates that training loss is depressed with validation loss settling but remaining low in variability, which obviously indicates that the model has good generalization to new data. Overall, the trend of the model is very strong and quite stable showing no clear signs of overfitting.

RESULTS AND DISCUSSION

The model's performance was evaluated based on the classification task metrics as shown in Figure 4, where the binary classification model gave good results.

Classification Report:

	precision	recall	f1-score	support
malignant	0.98	0.99	0.99	723
benign	0.92	0.87	0.89	92
micro avg	0.98	0.98	0.98	815
macro avg	0.95	0.93	0.94	815
weighted avg	0.98	0.98	0.98	815
samples avg	0.98	0.98	0.98	815

Figure 4. Report classification for EfficientNetB7 model in the current study.

The results of the proposed model EfficientNetB7 indicate robust and accurate performance in classifying breast cancer images into "malignant" and "benign" categories. The Confidence Interval (95%): (0.9724, 0.9854)) ensures that the model's true performance falls within this 95% range, demonstrating a high level of accuracy and reliability. Additionally, the T-Statistic: 0.9703 showed no significant difference in performance (Fail to reject H0), meaning that the model's recorded results do not significantly deviate from statistical expectations, enhancing the credibility of the results and confirming the model's stability.

In terms of errors, the model recorded 7 false positives, where malignant tumors were misclassified as benign, which can lead to delayed diagnosis and inaccurate treatment decisions. As for false negatives, the model recorded 12 cases that were misclassified as malignant when they were actually benign.

This can cause unnecessary anxiety and additional testing, but it is less serious than false positives. In terms of classification performance, the model achieved a precision of 0.98 and a recall of 0.99 in the malignant tumor category, meaning that 98% of cases classified as malignant were correctly classified and that it detected 99% of all malignant cases.

This results in an F1-score of 0.99, reflecting an excellent balance between precision and recall. In contrast, in the benign tumor category, the precision was 0.92 and the recall was 0.87, with an F1-score of 0.89, demonstrating good performance despite some errors.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

The ROC curve also demonstrates the model's ability to effectively differentiate between malignant and benign tumors. It features a low false positive rate (FPR), meaning the model rarely classifies benign tumors as malignant, and a high true positive rate (TPR), reflecting a high ability to detect malignant tumors. An AUC > 0.9 also demonstrates excellent performance, enhancing confidence in classification decisions and indicating that the model performs well.

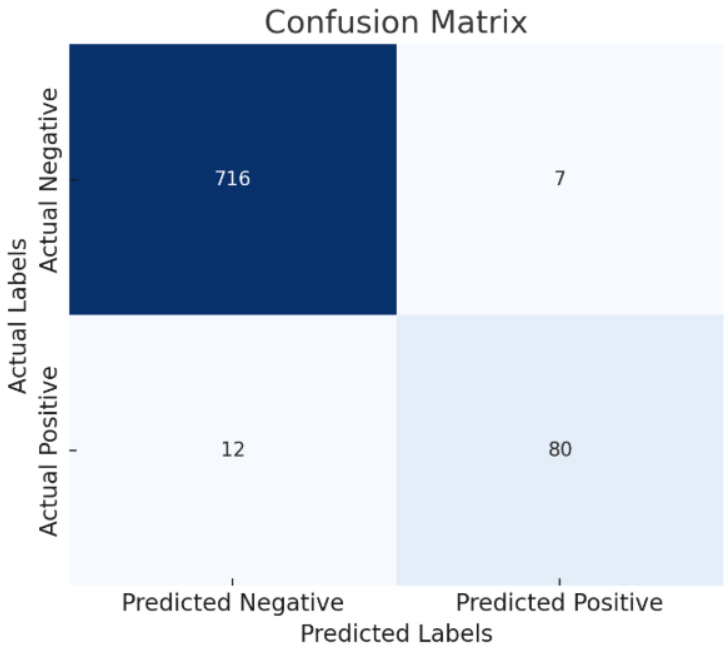

Figure 5. The confusion matrix in the current study.

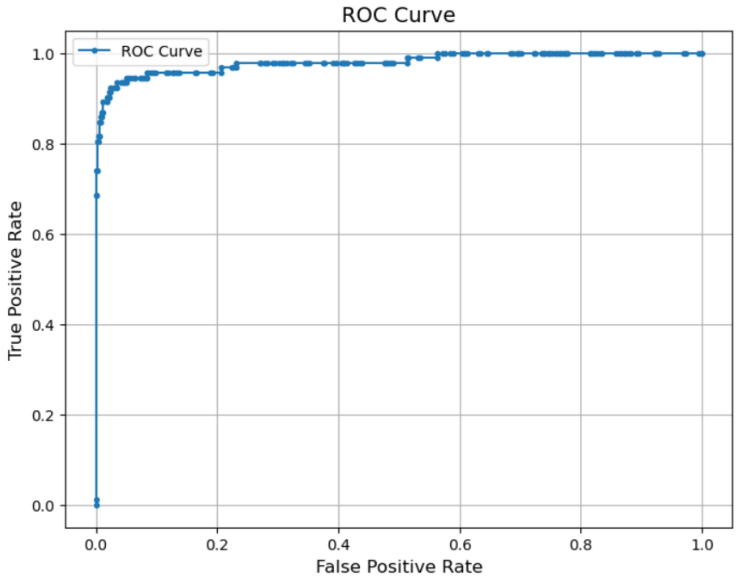

Figure 6. The ROC curve in the current study.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

The results in figure 7 show the prediction probabilities for classes “0” and “1”, where the predicted class is chosen based on the highest probability. The performance indicates a high accuracy of the model, as most probabilities are close to 1 for one class and close to zero for the other, reflecting a strong ability to discriminate between classes.

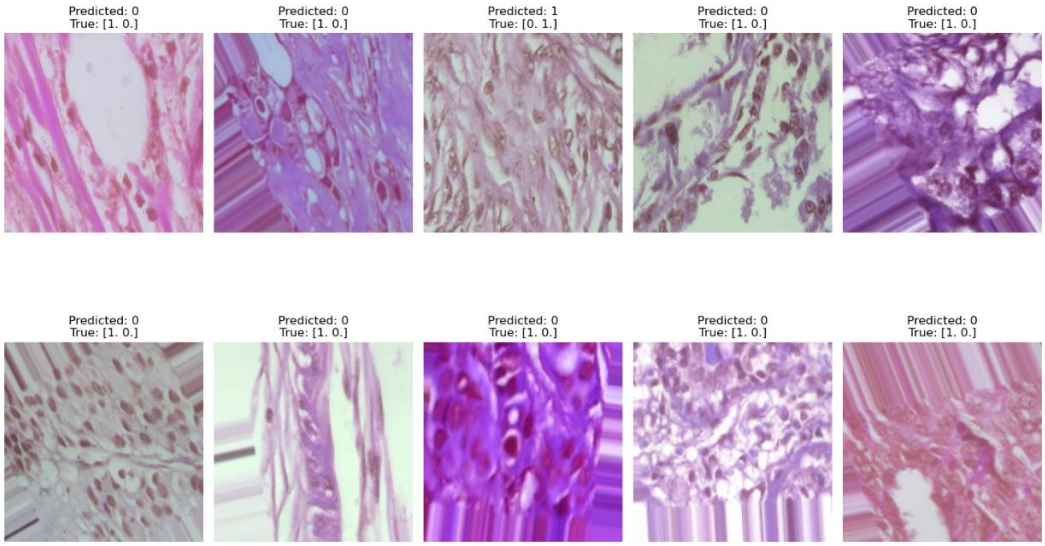

Figure 7. The prediction images in the current study.

Validation with External Data

To validate the model and ensure its ability to accurately classify real-world data, we tested our proposed model using an external dataset. The "Multi Cancer Dataset' is available on Kaggle at https://www.kaggle.com/datasets/obulisainaren/multi-cancer.this dataset contains images related to several types of cancer, among them breast cancer. The dataset used for experimental testing is comprised of 2,191 samples of breast cancer images benign and malignant. In the new dataset, it underwent the pre-processing phase, which resulted in image processing techniques being applied for image resizing to 150 x 150 and data augmentation and resampling due to class imbalance.The trained model was reloaded and tested, achieving a test accuracy of 94%.

```
30/30 [==============================] - 97s 3s/step - loss: 0.3985 - accuracy: 0.9407
Test Loss: 0.39847269654273987
Test Accuracy: 0.9406779408454895


30/30 [==============================] - 97s 3s/step
Classification Report:
              precision    recall  f1-score   support

   malignant       0.94      0.99      0.96       751
      benign       0.93      0.77      0.84       193

   micro avg       0.94      0.94      0.94       944
   macro avg       0.94      0.88      0.90       944
weighted avg       0.94      0.94      0.94       944
 samples avg       0.94      0.94      0.94       944
```

Figure 8. The classification report for External Data in the current study.

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

As indicated in Figure 8 of the classification report, the recall accuracy for the 'malignant' class is very high, which means high sensitivity for detecting malignant cases. The recall for the 'benign' class is low, which indicates that the model has low sensitivity to correctly classify some benign cases. In fine the overall performance of the model is good but class-specific accuracy in benign class could be improved.

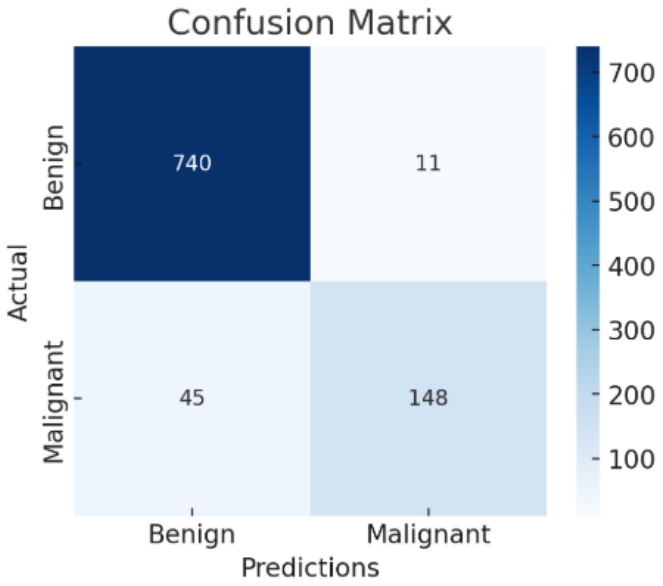

Figure 9. The confusion matrix for External Data in the current study.

Referring to the confusion matrix in figure 9, it' is observed that the model has high performance in classifying malignant cases. This class includes 740 correctly classified cases and 11 misclassified as benign. For the benign class, correctly classified cases are 148 and misclassified are 45 as the malignant class.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

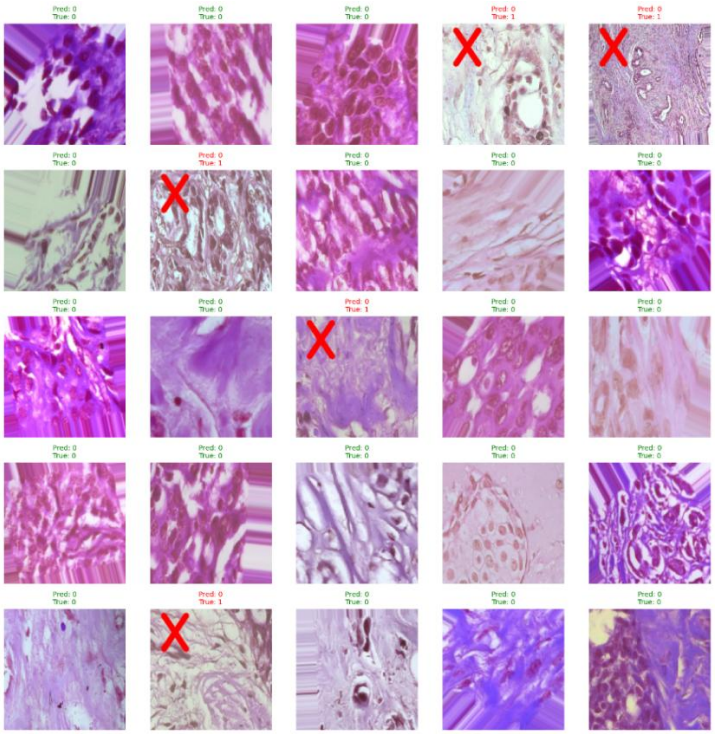

Figure 10. The evaluation for External Data in the current study.

We also subjected our model to predictions on 25 randomly selected images from the test data, shown in figure 10, with the model's predictions and the actual truth for each image shown. The presence of a red X indicates the model's classification errors. Overall, this distribution shows that the model has a low false positive rate.

Comparison with related works

Considering the achieved results, we can notice the superiority of our proposed model in detecting breast cancer compared to related studies that presented their models with the BreakHis_400X dataset. In the following section, we highlight the comparison of the results with related studies.

Table 7. Comparison of the Proposed Model with Related Works in the current study.

No.	Reference	Models	Results
	Our proposed system	EfficientNetB7	With high accuracy (98%), recall (99%), average F1 score (98%), ROC curve greater than 0.9, high sensitivity (TPR) and low false positive rate (FPR).
1	[42]	CNN-ANN	Accuracy: 89.47%, Precision: 86.18%, AUC: 89.46%, F1-score: 89.08%. Model outperformed individual components.
2	[40]	DenseNet201	Training Accuracy: 97.00%, Validation Accuracy: 92.00%. DenseNet201 demonstrated high efficacy in tumor classification.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

3	[44]	DenseNet201, ResNet50, ResNet101, MobileNet-v2	DenseNet201: Accuracy 91.37%, Sensitivity 100% (at 200x magnification), outperforming other models.
4	[45]	XSV (Xception + SVM), XRF (Xception + RF)	XSV: Accuracy up to 90.17% (BreakHis) and 87.35% (BHID). XRF: Accuracy up to 88.98% (BreakHis) and 87.29% (BHID). Both outperform traditional classifiers.
5	[46]	Custom CNN, MobileNetV3, EfficientNetB0, VGG16, ResNet50V2	Custom CNN (optimized with MGTO): Accuracy 93.13% (10 iterations). Pretrained models: Accuracy between 74%-82%. Custom model significantly outperformed pretrained models.
6	[47]	Few-shot learning with feature fusion	BreakHis: Accuracy 91.22%. Skin Lesions: Accuracy 71.20%. Achieved superior results with only 10 labeled samples compared to other state-of-the-art methods.

Table 7, our proposed system, based on EfficientNetB7, demonstrates superior performance in detecting breast cancer compared to related studies using the BreakHis_400X dataset. Achieving a high accuracy of 98%, recall of 99%, F1 score of 98%, and an AUC greater than 0.9, the model exhibits excellent sensitivity (TPR) and a low false positive rate (FPR). Compared to other studies, [42] CNN-ANN model achieved lower accuracy (89.47%) and AUC (89.46%) with limited precision (86.18%).[40] reported a validation accuracy of 92% with DenseNet201, which lags behind our model.[44] achieved a sensitivity of 100% with DenseNet201 at 200x magnification but an overall accuracy of 91.37%. [45] hybrid models (XSV, XRF) attained accuracies of up to 90.17% for BreakHis, while [46] achieved a maximum accuracy of 93.13% using a custom CNN. [47] employed few-shot learning with feature fusion, achieving 91.22% accuracy for BreakHis. These comparisons highlight the efficiency and advanced capabilities of our proposed system, setting it apart as a robust solution for breast cancer detection.

CONCLUSION

This research proposed models high efficiency in breast cancer image classification; it confirms that what it can do appropriately distinguishes between a malignant and benign case. Tests have proven the model's robustness in the face of external world data, and therefore, its reliability increases in clinical settings. The results also reflect that tumour can be developed to classify benign cases. Since it performs very well, the model could be implemented into medical applications to support the diagnostic process.

Acknowledgments:

Thanks to everyone who helped complete this achievement.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية | مجلة جامعة بابل للعلوم الصرفة والتطبيقية

Conflict of interests:

There are non-conflicts of interest.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA Cancer J. Clin., vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.

[2] N. Harbeck et al., "Breast cancer," Nat. Rev. Dis. Primers, vol. 5, no. 1, p. 66, 2019, doi: 10.1038/s41572-019-0111-2.

[3] C. E. DeSantis, J. Ma, A. Goding Sauer, L. A. Newman, and A. Jemal, "Breast cancer statistics, 2017, racial disparity in mortality by state," CA Cancer J. Clin., vol. 67, no. 6, pp. 439–448, 2017, doi: 10.3322/caac.21412.

[4] W. Chen, R. Zheng, P. D. Baade, S. Zhang, H. Zeng, F. Bray, and J. He, "Cancer statistics in China, 2015," CA Cancer J. Clin., vol. 66, no. 2, pp. 115–132, 2016, doi: 10.3322/caac.21338.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Med. Image Comput. Comput. Assist. Interv. – MICCAI 2015, Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[6] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, and C. I. Sánchez, "A survey on deep learning in medical image analysis," Med. Image Anal., vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.

[7] N. Dhungel, G. Carneiro, and A. P. Bradley, "A deep learning approach for the analysis of masses in mammograms with minimal user intervention," Med. Image Anal., vol. 37, pp. 114–128, 2017, doi: 10.1016/j.media.2017.01.009.

[8] N. Dhungel, G. Carneiro, and A. P. Bradley, "Tree RE-Weighted Belief Propagation Using Deep Learning Potentials for Mass Segmentation From Mammograms," in Proc. 12th Int. Symp. Biomed. Imaging, IEEE, 2015, pp. 760–763, doi: 10.1109/ISBI.2015.7163983.

[9] W. Zhu, X. Xiang, T. D. Tran, G. D. Hager, and X. Xie, "Adversarial Deep Structured Nets for Mass Segmentation From Mammograms," in Proc. 15th Int. Symp. Biomed. Imaging, IEEE, 2018, pp. 847–850, doi: 10.1109/ISBI.2018.8363704.

[10] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with deep learning," Sci. Rep., vol. 8, no. 1, p. 4165, 2018, doi: 10.1038/s41598-018-22437-z.

[11] V. K. Singh et al., "Conditional Generative Adversarial and Convolutional Networks for X-ray Breast Mass Segmentation and Shape Classification," in Proc. Med. Image Comput. Comput. Assist. Interv., 2018, pp. 833–840.

[12] V. Agarwal and C. Carson, "Using Deep Convolutional Neural Networks to Predict Semantic Features of Lesions in Mammograms," Stanford Univ., 2015. [Online]. Available: http://cs231n.stanford.edu/reports/2015/pdfs/vibhua_final_report.pdf.

[13] F. Gao, T. Wu, J. Li, B. Zheng, L. Ruan, D. Shang, and B. Patel, "SD-CNN: a shallow-deep CNN for improved breast cancer diagnosis," Comput. Med. Imaging Graph., vol. 70, pp. 53–62, 2018, doi: 10.1016/j.compmedimag.2018.09.004.

[14] Y. B. Hagos, A. G. Mérida, and J. Teuwen, "Improving Breast Cancer Detection Using Symmetry Information with Deep Learning," in Proc. Image Anal. Moving Organ, Breast, Thoracic Images, 2018, pp. 90–97.

[15] J. Teuwen, S. van de Leemput, A. Gubern-Mérida, A. Rodriguez-Ruiz, R. Mann, and B. E. Bejnordi, "Soft Tissue Lesion Detection in Mammography Using Deep Neural Networks for Object Detection," in Proc. 1st Conf. Med. Imaging Deep Learning, 2018.

[16] H. Jung, B. Kim, I. Lee, M. Yoo, J. Lee, S. Ham, J. Kang, "Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network," PLoS One, vol. 13, no. 9, p. e0203355, 2018, doi: 10.1371/journal.pone.0203355.

[17] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "End-to-End Training for Whole Image Breast Cancer Diagnosis Using an All Convolutional Design," Cornell Univ., 2017. [Online]. Available: https://arxiv.org/pdf/1708.09427.pdf.

[18] A. A. Mohamed, W. A. Berg, H. Peng, Y. Luo, R. C. Jankowitz, and S. Wu, "A deep learning method for classifying mammographic breast density categories," Med. Phys., vol. 45, no. 1, pp. 314–321, 2018, doi: 10.1002/mp.12683.

[19] C. K. Ahn, C. Heo, H. Jin, and J. H. Kim, "A Novel Deep Learning-Based Approach to High Accuracy Breast Density Estimation in Digital Mammography," in Proc. Comput.-Aided Diagnosis, 2017, doi: 10.1117/12.2254264.

[20] J. Wang, X. Yang, H. Cai, W. Tan, C. Jin, and L. Li, "Discrimination of breast cancer with microcalcifications on mammography by deep learning," Sci. Rep., vol. 6, no. 1, p. 27327, 2016, doi: 10.1038/srep27327.

[21] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with deep learning," Sci. Rep., vol. 8, no. 1, p. 4165, 2018, doi: 10.1038/s41598-018-22437-z.

[22] V. K. Singh, S. Romani, H. A. Rashwan, F. Akram, N. Pandey, M. M. K. Sarker, and D. Puig, "Conditional Generative Adversarial and Convolutional Networks for X-ray Breast Mass Segmentation and Shape Classification," in Proc. Med. Image Comput. Comput. Assist. Interv., 2018, pp. 833–840.

[23] V. Agarwal and C. Carson, "Using Deep Convolutional Neural Networks to Predict Semantic Features of Lesions in Mammograms," Stanford Univ., 2015. [Online]. Available: http://cs231n.stanford.edu/reports/2015/pdfs/vibhua_final_report.pdf.

[24] Y. B. Hagos, A. G. Mérida, and J. Teuwen, "Improving breast cancer detection using symmetry information with deep learning," in Proc. Image Anal. Moving Organ, Breast, Thoracic Images, 2018, pp. 90–97.

[25] J. Teuwen, S. van de Leemput, A. Gubern-Mérida, A. Rodriguez-Ruiz, R. Mann, and B. E. Bejnordi,"Soft tissue lesion detection in mammography using deep neural networks for object detection," in Proc. 1st Conf. Med. Imaging Deep Learn., 2018.

[26] H. Jung, B. Kim, I. Lee, M. Yoo, J. Lee, S. Ham, and J. Kang, "Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network," PLoS One, vol. 13, no. 9, p. e0203355, 2018. doi: 10.1371/journal.pone.0203355.

[27] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "End-to-end training for whole image breast cancer diagnosis using an all convolutional design," Cornell Univ., 2017. [Online]. Available: https://arxiv.org/pdf/1708.09427.pdf.

[28] C. K. Ahn, C. Heo, H. Jin, and J. H. Kim, "A novel deep learning-based approach to high accuracy breast density estimation in digital mammography," in Proc. Comput.-Aided Diagnosis, 2017. doi: 10.1117/12.2254264.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652 | www.journalofbabylon.com | jub@itnet.uobabylon.edu.iq | info@journalofbabylon.com

[29] N. Wu et al., "Breast density classification with deep convolutional neural networks," in Proc. Int. Conf. Acoust., Speech, Signal Process., 2018, pp. 6682–6886. IEEE. doi: 10.1109/ICASSP.2018.8462671.

[30] J. Xu, C. Li, Y. Zhou, L. Mou, H. Zheng, and S. Wang, "Classifying mammographic breast density by residual learning," Cornell Univ., 2018. [Online]. Available: https://arxiv.org/abs/1809.10241.

[31] M. Kallenberg, K. Petersen, M. Nielsen, A. Y. Ng, P. Diao, C. Igel, and M. Lillholm, "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1322–1331, 2016. doi: 10.1109/TMI.2016.2532122.

[32] G. V. Ionescu, M. Fergie, M. Berks, E. F. Harkness, J. Hulleman, A. R. Brentnall, and S. M. Astley, "Prediction of reader estimates of mammographic density using convolutional neural networks," J. Med. Imaging, vol. 6, no. 3, p. 031405, 2019. doi: 10.1117/1.JMI.6.3.031405.

[33] D. Lévy and A. Jain, "Breast mass classification from mammograms using deep convolutional neural networks," Cornell Univ., 2016. [Online]. Available: https://arxiv.org/pdf/1612.00542.pdf.

[34] R. K. Samala, H. P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter, and K. H. Cha, "Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning using deep neural nets," IEEE Trans. Med. Imaging, vol. 38, no. 3, pp. 686–696, 2019. doi: 10.1109/TMI.2018.2870343.

[35] T. Alam, W. C. Shia, F. R. Hsu, and T. Hassan, "Improving breast cancer detection and diagnosis through semantic segmentation using the Unet3+ deep learning framework," Biomedicines, vol. 11, no. 6, p. 1536, 2023.

[36] B. S. Abunasser, M. R. J. Al-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, "Convolution neural network for breast cancer detection and classification using deep learning," Asian Pac. J. Cancer Prev., vol. 24, no. 2, p. 531, 2023.

[37] A. Egami, Y. Tarumi, A. Okamura, K. Aoyama, H. Kataoka, T. Kokabu, and T. Mori, "A case report of occult breast cancer detected by diagnostic laparoscopy for suspected ovarian cancer," Case Rep. Obstet. Gynecol., vol. 2024, no. 1, p. 8851045, 2024.

[38] J. Wang, L. Qiao, S. Zhou, J. Zhou, J. Wang, J. Li, and J. Shi, "Weakly supervised lesion detection and diagnosis for breast cancers with partially annotated ultrasound images," IEEE Trans. Med. Imaging, 2024.

[39] F. Prinzi, M. Insalaco, A. Orlando, S. Gaglio, and S. Vitabile, "A YOLO-based model for breast cancer detection in mammograms," Cogn. Comput., vol. 16, no. 1, pp. 107–120, 2024.

[40] L. de Munck et al., "Method of primary breast cancer detection and the disease-free interval, adjusting for lead time," JNCI: J. Natl. Cancer Inst., vol. 116, no. 3, pp. 370–378, 2024.

[41] H. Kode and B. D. Barkana, "Deep learning-and expert knowledge-based feature extraction and performance evaluation in breast histopathology images," Cancers 2023, vol. 15, p. 3075, 2023.

[42] R. O. Ogundokun et al., "Hybrid deep learning for breast cancer diagnosis: Evaluating CNN and ANN on BreakHis_v1_400X," in Proc. 2024 Int. Conf. Sci., Eng., Bus. Sustain. Dev. Goals (SEB4SDG), Apr. 2024, pp. 1–6, IEEE.

[43] K. Logu and S. J. J. Thangaraj, "Progressing breast cancer assessment: Precise tumor categorization through DenseNet201-based deep learning," in Proc. 2024 Int. Conf. Trends Quantum Comput. Emerg. Bus. Technol., Mar. 2024, pp. 1–6, IEEE.

[44] E. O. Simonyan, J. A. Badejo, and J. S. Weijin, "Histopathological breast cancer classification using CNN," Mater. Today: Proc., vol. 105, pp. 268–275, 2024.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

[45] M. Gupta et al., "Deep transfer learning hybrid techniques for precision in breast cancer tumor histopathology classification," 2024.

[46] A. Heikal et al., "Fine tuning deep learning models for breast tumor classification," Sci. Rep., vol. 14, no. 1, p. 10753, 2024.

[47] W. Wang et al., "Medical tumor image classification based on few-shot learning," IEEE/ACM Trans. Comput. Biol. Bioinf., 2023.

[48] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," Technologies, vol. 11, no. 2, p. 40, 2023.

[49] A. Kumar, “Ensemble methods in machine learning: Examples,” Vitalflux, 2023. [Online]. Available: https://vitalflux.com/5-common-ensemble-methods-in-machine-learning/

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

الخلاصة

المقدمة:

يعتبر سرطان الثدي من أكثر الأورام الخبيثة شيوعًا التي تصيب النساء في جميع أنحاء العالم. وينشأ نتيجة لعوامل داخلية وخارجية مختلفة، بما في ذلك الطفرات الجينية، وعادات نمط الحياة، والتعرض البيئي، والعوامل النفسية الاجتماعية. ووفقًا للإحصاءات، فإن 5% –10% من حالات سرطان الثدي مرتبطة بالطفرات الجينية والتاريخ العائلي، في حين أن 20% –30% ترجع إلى عوامل قابلة للتعديل مثل النظام الغذائي والنشاط البدني .

طرق العمل:

تضمنت مجموعة البيانات صورًا بالموجات فوق الصوتية لكتل الثدي، حيث كانت الكتل من فئتين: خبيثة وحميدة. طُبقت على الصور خطوة معالجة مسبقة، تتضمن عادةً تغيير حجمها إلى 150 × 150 بكسل، وتطبيع التباين، والتطبيع، وزيادة البيانات، وإعادة أخذ العينات لتصحيح اختلال الفئات، وذلك لضمان جودة البيانات. أجرت الدراسة تصنيفًا ثنائيًا لكتل الثدي باستخدام التعلم العميق بالانتقال استنادًا إلى نموذج EfficientNetB7 المُدرّب مسبقًا. قُسِّمت مجموعة البيانات إلى 80% للتدريب و20% للاختبار. استُخدمت هنا مجموعة بيانات خارجية، "مجموعة بيانات السرطان المتعددة"، متوفرة من Kaggle، وتتألف من 2191 صورة لسرطانات الثدي (الحميدة والخبيثة)، لاختبار قدرة النموذج على التعميم.

النتائج

أظهر النموذج المقترح أداءً ممتازًا في تصنيف الصور إلى فئتين "خبيثة" و"حميدة"، محققًا دقةً بلغت 0.98، وحساسيةً (استرجاعًا) تقارب 0.99، وقيمة F1 مقابلة. بلغت دقة النموذج الحميد 0.92، بينما بلغت الحساسية وقيمة F1 0.87 و0.89 على التوالي. بلغت الدقة والحساسية وقيمة F1 0.98 لمجموعة البيانات بأكملها، مما يؤكد موثوقية النموذج الإجمالية. كان أداء النموذج جيدًا نسبيًا على منحنى خصائص مشغل المستقبل (ROC)، حيث اقتربت المساحة تحت المنحنى من القيمة القصوى 1 (0.9<)، مما يدل على حساسية عالية (معدل إيجابي حقيقي) وانخفاض معدل إيجابي كاذب (FPR). أما على مجموعة البيانات الخارجية، فقد كان أداؤه جيدًا، محققًا دقةً بلغت 0.94.

الاستنتاجات:

تشير النتائج إلى أن الشبكات العصبية التلافيفية العميقة المدربة مسبقًا يمكن أن تحقق تقدمًا كبيرًا في الكشف الدقيق عن الحالات الخبيثة مع تقليل الإنذارات الكاذبة، وتعزيز تطبيقها السريري ودورها في تحسين تشخيص سرطان الثدي باستخدام تقنيات الذكاء الاصطناعي.

الكلمات المفتاحية:

سرطان الثدي، التعلم العميق، التعلم الانتقالي، النموذج المدرب مسبقًا، الشبكات العصبية التلافيفية.