https://dx.doi.org/10.4314/dujopas.v9i3b.30
ISSN (Print): 2476-8316
ISSN (Online): 2635-3490
Dutse Journal of Pure and Applied Sciences (DUJOPAS), *Vol. 9 No. 3b September 2023*

Generic Hybrid Model for Breast Cancer Mammography Image Classification Using EfficientNetB2

[1]Oluwasegun Abiodun Abioye , [2]Sadiq Thomas, [2] Chinomso Roselyn Odimba, [1]Awujoola Joel Olalekan

[1]Directorate of Information and Communications Technology,
NDA,
Kaduna

[2]Department of Computer Engineering,
Nile University
Abuja

Email: segunabioye@nda.edu.ng

Abstract
Breast cancer is a global health issue that necessitates precise classification for early detection and effective treatment. In recent years, pre-trained models have shown great potential in the field of medical image classification, including breast cancer classification. These models have been trained on extensive datasets, and they possess the ability to capture intricate features and patterns within medical images, facilitating accurate classification. However, some of the models are non-generic. They can be sensitive to dataset biases, leading to over fitting on specific patterns present in the training data, and they equally struggle to handle data from different distributions. In this work, we proposed a generic hybrid model for image classification. The features were extracted from two datasets: the mammographic image analysis society (MIAS) and the INbreast dataset, respectively, through the pre trained EfficientNetB2 architecture. However, three classifiers were used in the image classification of the extracted features: MGSVM, CUBIC SVM, and XGBOOST. Eight evaluation metrics were selected to assess the performance of the proposed models. These metrics include accuracy, precision, F1-score, AUC, sensitivity, false negative rate (FNR), Kappa score, and time complexity. Experimental results show that the hybrid of EfficientNetB2 and the MGSVM classifier is more generic and efficient for breast cancer diagnosis and classification. It exhibits a strong performance when classifying mammography breast images from both datasets, achieving impressive metrics such as an overall accuracy of 99.47%, a sensitivity rate of 99.31%, precision of 99.44%, F1-score of 99.44%, AUC of 99.44%, a low FNR (False Negative Rate) of 0.007, a kappa score of 0.98, and a manageable time complexity of 231.44 seconds on the MIAS Dataset.

Keywords: breast cancer; mammogram images; feature extraction; deep learning; generic hybrid model

INTRODUCTION
Breast cancer, a prevalent non-communicable disease affecting women, has prompted extensive research into early detection and classification methods. In the year 2022 alone, approximately 287,850 new cases were diagnosed, resulting in 43,250 fatalities (Jabeen et al.,

*Author for Correspondence

2023). Limited expertise in breast cancer nursing due to staff shortages poses a challenge in resource allocation within healthcare systems, leading to delayed detection and treatment (Azamjah et al., 2019). Conventional classification methods for medical conditions such as breast masses involve manual feature extraction followed by machine learning, a laborious and less effective process (Li et al., 2021). Computer-aided detection (CAD) technologies, which alleviate radiologists' workload and enhance accuracy in breast cancer diagnosis, continue to evolve (Hekal et al., 2021). Nonetheless, complex image issues and tumor region variations present challenges in achieving accurate classification. Consequently, artificial intelligence (AI) has made strides in medical image analysis, particularly in breast cancer detection. Convolutional neural networks (CNNs) have demonstrated exceptional performance in cancer detection by learning complex patterns from large datasets (Aljuaid et al., 2022). Transfer learning, a technique in deep learning, has been instrumental in medical image analysis, using pre-trained models for feature extraction followed by fine-tuning on smaller datasets. This approach improves model performance with less data (Wang et al., 2019). Several studies have leveraged deep learning architectures for breast cancer classification, achieving promising results. Ting et al. (2019) used a deep CNN achieving 89.47% sensitivity, 90.50% accuracy, and 90.71% specificity. Abbas et al. (2016) introduced a multi-layer DL model with 92% sensitivity, 84.2% specificity, 91.5% accuracy, and an AUC of 0.91. Togaçar et al. (2020) introduced BreastNet, which achieved 98.80% accuracy. Sha et al. (2020) utilized a CNN-based approach with sensitivity, specificity, and accuracy of 96%, 93%, and 92%, respectively. Wahab et al. (2019) used pre-trained CNNs to classify mitoses, attaining precision, recall, and F-measure of 0.50, 0.80, and 0.621. Lotter et al. (2019) utilized a pre-trained ResNet50, achieving sensitivity, specificity, and AUC values of 96.2%, 90.9%, and 0.94. Khan et al. (2019) proposed a model with 97.525% accuracy. Deniz et al. (2018) fine-tuned AlexNet and VGG16 for 91.37% accuracy. Celik et al. (2020) achieved F-score and accuracy values of 92.38% and 91.57%. Karthiga et al. (2022) combined transfer learning and CNNs for improvements. Almalki et al. (2022) proposed a technique with 92% accuracy. Hikmah et al. (2022) introduced a framework achieving detection accuracy of 88.0% and 80.5%. Hekal et al. (2021) achieved accuracy rates of 91% and 84%.

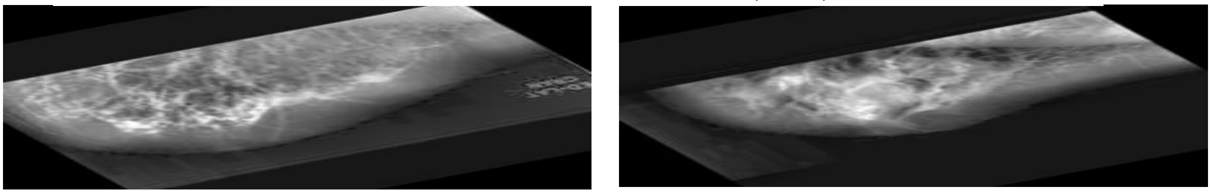
The purpose of this research work is to create a generic hybrid model capable of effectively classifying a Mammography Dataset containing Benign and Malignant Breast Masses. The proposed model utilizes a Convolutional Neural Network (CNN) based on transfer learning, specifically using the fine-tuned EfficientNetB2 for feature extraction. The extracted features are then fed into XGBOOST, CUBIC SVM, and MGSVM classifiers for accurate image classification. The aim is to leverage the strengths of both deep learning and traditional machine learning techniques to achieve enhanced classification performance in mammography image analysis.

MATERIALS AND METHODS

MATERIALS

In the context of medical imaging and breast cancer research, a mammography dataset typically contains a collection of medical images known as mammograms. These images are specifically captured using mammography, a specialized imaging technique that uses low-dose X-rays to visualize the internal structures of the breast. The images are made up of Benign and Malignant breast masses; benign masses are non-cancerous lumps or growths in the breast tissue. These masses can be caused by a variety of factors, including hormonal changes, inflammation, and fibrocystic changes in the breast tissue. While benign masses are typically not harmful and do not require treatment, they can sometimes cause discomfort or pain and may need to be monitored over time. In mammography image analysis, differentiating between benign and malignant masses is critical for accurate diagnosis and treatment of breast cancer. Malignant masses are cancerous lumps or growths in the breast tissue. They can invade surrounding tissues and potentially spread to other parts of the body. In mammography image analysis, identifying and diagnosing malignant masses is crucial for early detection and treatment of breast cancer. Malignant masses can have various characteristics in mammography images, including an irregular or spiculated border, an irregular shape, and a heterogeneous or non-uniform appearance. These masses may also have associated calcifications or areas of distortion in the surrounding breast tissue.

Fig 1. (a) Original Mammogram with a Benign Mass. (b) Original Mammogram with a Malignant Mass. **Laisné, M. (2019)**



a) Malignant Mass of Breast Cancer Mammogram

b) Benign Mass of Breast Cancer Mammogram

MAMMOGRAPHY DATASET

In this study, we used mammographic images from Mendeley Data. The datasets used for this work are namely the Mammographic Image Analysis Society database (MIAS) and the INbreast datasets. The MIAS is a database of digital mammograms that is widely used for research in the field of mammography and breast cancer diagnosis. The INbreast dataset is a publicly available dataset for mammography image analysis. The mammogram images from Mendeley Data were preprocessed using data augmentation and contrast-limited adaptive histogram equalization techniques. Following augmentation, the MIAS and INbreast datasets contained 3816 images and 7632 images, respectively.

METHODS

This research work utilized the EfficientNetB2 network to serve as a feature extractor for capturing high-level and distinctive features from input images. Moreover, our approach involved the integration of three classifiers: XGBOOST, CUBIC SVM, and MGSVM as transfer learning techniques to categorize various images obtained from both the MIAS and INbreast datasets.

FINE TUNED MODEL

To begin, we loaded the pre-trained EfficientNetB2 model, excluding its default classification or fully connected layer. This step involved setting the "include_top" parameter to FALSE during the model loading phase. Following this, the extracted images underwent classification using each of the classifiers, and we evaluated the performance of each classifier individually using relevant metrics. During the preprocessing stage, the input size for the pre-trained EfficientNetB2 model was standardized at 260x260 pixels, ensuring a consistent approach in handling the input data. Notably, the pre-trained model employed for feature extraction in this research underwent weight updates during training. This adaptation enabled the model to align the pre-trained features with the specific task at hand. The objective was to determine the best combination of hyper parameters that maximized the model's performance while mitigating the risk of over fitting. Each classifier underwent fine-tuning with a comprehensive set of parameters. The first among them, the XGBOOST CLASSIFIER, refers to Extreme Gradient Boosting. This widely-used machine learning algorithm excels in gradient boosting frameworks and is recognized for its impressive performance across diverse tasks. Its hyper parameters were adjusted as follows: "n_estimators" was set to 100, "max_depth" to 6, and "random_state" to 42. Next, the CUBIC SVM CLASSIFIER, an alternative version of the Support Vector Machine (SVM), employed cubic polynomials as its kernel function. SVM is a supervised machine learning algorithm employed in classification and regression tasks. The hyper parameters of the CUBIC SVM were fine-tuned: the regularization parameter "C" was set to 10, and the degree of the polynomial used in the kernel function, "Degree," was set to 4. Lastly, the MGSVM CLASSIFIER, signifying Multiclass Gaussian Support Vector Machine, addressed multiclass classification challenges. This SVM variant optimizes the parameters of each binary classifier by maximizing the margin between two classes in the feature space, accounting for the influence of other classes. For this classifier, the hyper parameters were fine-tuned as follows: the regularization parameter "C" was set to 0.1, and the gamma parameter "gamma" was set to 0.001.

RESULT AND DISCUSSION

MODEL EVALUATION RESULT OF EFFICIENTNETB2 ON MGSVM, CUBIC SVM AND XGBOOST USING MIAS DATASET

Table 1: Model Classification Report for the Fine –Tuned EfficientNetB2 with MGSVM, CUBIC SVM and XGBOOST Classifiers on MIAS Dataset

Classifier	Sensitivity Rate %	Precision Rate %	F1 Score	AUC	Recall %	Accuracy %	FNR	Kappa Score	Time(s)
MGSVM	**99.30**	**99.44**	**99.44**	**99.44**	**99.44**	**99.47**	**0.007**	**0.98**	**231.44**
CUBIC SVM	**92.70**	**97.31**	**96.61**	**96.04**	**96.04**	**96.85**	**0.073**	**0.93**	**336.77**
XGBOOST	**90.97**	**95.88**	**95.20**	**94.64**	**94.65**	**95.54**	**0.090**	**0.90**	**2007.85**

Table 1. Classification results of the original dataset after deep feature extraction using fine-tuned EfficientNetB2. The report provides an evaluation of the sensitivity rate, precision, F1-score, AUC, Accuracy, FNR, kappa score, and time complexity metrics for both the “Benign” and “Malignant” classes, along with the support (number of instances) for each class. The classifiers exhibited varying degrees of performance. MGSVM displayed exceptional results, achieving a sensitivity rate of 99.30%, a precision rate of 99.44%, an F1 score of 99.44%, an AUC value of 99.44%, an overall accuracy of 99.47%, and a low false negative rate (FNR) of 0.007%. CUBIC SVM, the second classifier, achieved a sensitivity rate of 92.70%, a precision rate of 97.31%, an F1 score of 96.61%, an AUC value of 96.04%, and an overall accuracy of 96.85%. XGBOOST, the third classifier, attained a sensitivity rate of 90.97%, a precision rate of 95.88%, an F1 score of 95.20%, an AUC value of 94.64%, and an overall accuracy of 95.54%. In conclusion, MGSVM demonstrated superior performance across sensitivity rate, precision rate, F1 score, AUC, accuracy, and Kappa score, making it the most effective classifier for this particular dataset.

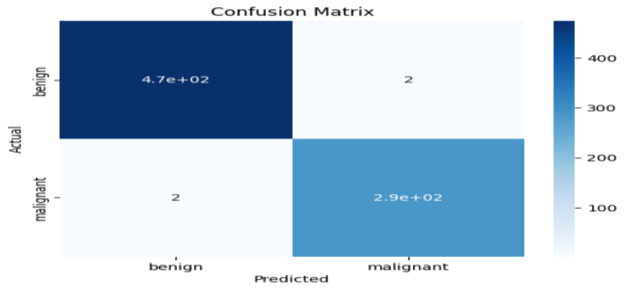


Fig 2 Confusion matrix of MGSVM with Deep Feature Extractions from EfficientNetB2 using MIAS Dataset

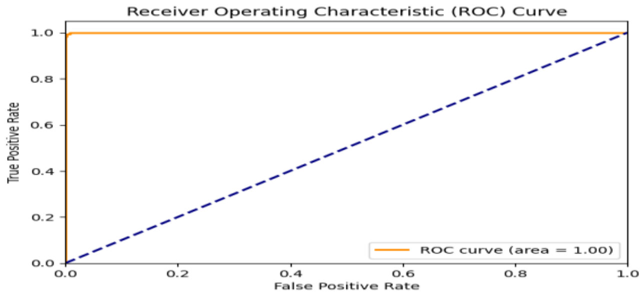


Fig 3 Receiver Operating Character of MGSVM with Deep Feature Extraction from EfficientNetB2 using MIAS Dataset

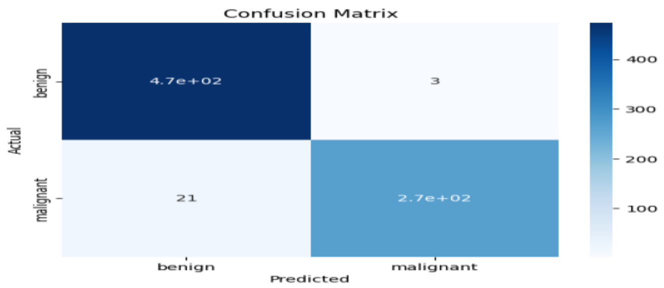


Fig 4 Confusion matrix of CUBIC SVM with Deep Feature Extractions from EfficientNetB2 using MIAS Dataset

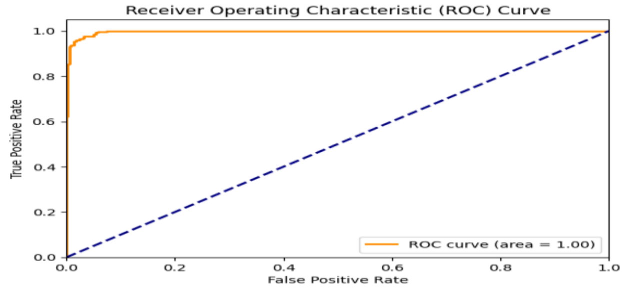


Fig 5 Receiver Operating Character of CUBIC SVM with Deep Feature Extractions from EffcieintNetB2 using MIAS Dataset

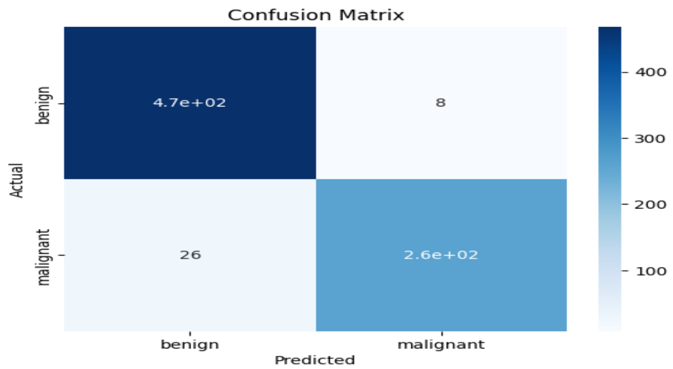


Fig 6 Confusion matrix of XGBOOST with Deep Feature Extractions from EfficientNetB2 usina MIAS Dataset

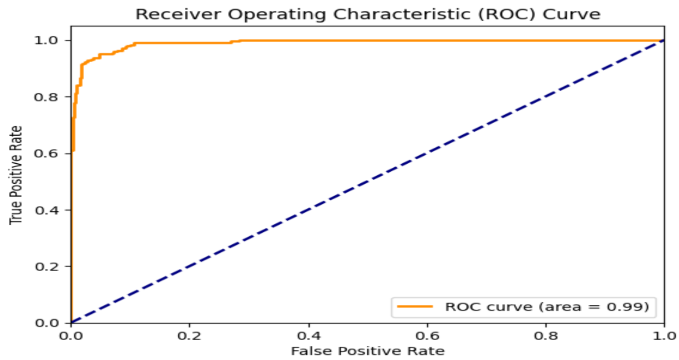


Fig 7 Receiver Operating Character of XGBOOST with Deep Feature Extractions from EfficientNetB2 using MIAS Dataset

MODEL EVALUATION RESULT OF EFFICIENTNETB2 ON MGSVM, CUBIC SVM AND XGBOOST CLASSIFIERS USING INBREAST DATASET

Table 2: Model Classification Report for the Fine –Tuned EfficientNetB2 with MGSVM, CUBIC SVM and XGBOOST Classifiers using INbreast Dataset

Classifier	Sensitivity Rate %	Precision Rate %	F1 Score	AUC	Recall %	Accuracy %	FNR	Kappa Score	Time(s)
MGSVM	99.31	98.51	98.28	98.07	98.07	98.49	0.007	0.965	542.026
CUBIC SVM	99.41	98.08	97.52	97.02	97.02	97.83	0.006	0.950	1424.97
XGBOOST	98.82	97.42	97.01	96.64	96.63	97.38	0.011	0.940	5075.91

Table 2 presents the classification results of the INBREAST dataset after deep feature extraction using fine-tuned EfficientNetB2. The evaluation includes sensitivity rate, precision, F1-score, AUC, accuracy, false negative rate (FNR), kappa score, and time complexity metrics for both the "Benign" and "Malignant" classes, along with the number of instances (support) for each class. The dataset comprises two classes: "Benign" and "Malignant," and the performance of three different classifiers, namely MGSVM, CUBIC SVM, and XGBOOST, was analyzed. Each classifier's effectiveness in correctly classifying instances from both classes was assessed using diverse performance metrics, providing valuable insights into their classification capabilities. The MGSVM classifier displayed strong performance, achieving a sensitivity rate of 99.31%, precision rate of 98.51%, F1 score of 98.28%, AUC score of 98.07%, recall percentage of 98.07%, overall accuracy of 98.49%, and a low FNR of 0.007%. The Kappa score of 0.965 indicated good agreement between predicted and actual classes. The inference and prediction time was 542.026 seconds. The CUBIC SVM classifier demonstrated notable results, achieving a sensitivity rate of 99.41%, precision rate of 98.08%, F1 score of 97.52%, AUC score of 97.02%, recall percentage of 97.02%, overall accuracy of 97.83%, and a low FNR of 0.006%. The Kappa score of 0.950 indicated good agreement between predicted and actual classes. The inference and prediction time was 1424.97 seconds. The XGBOOST classifier achieved a sensitivity rate of 98.82%, precision rate of 97.42%, F1 score of 97.01%, AUC score of 96.64%, recall percentage of 96.63%, overall accuracy of 97.38%, and a FNR of 0.011%. The Kappa score of 0.940 indicated good agreement between predicted and actual classes. The inference and prediction time was 5075.91 seconds.

In summary, all three classifiers demonstrated strong performance in breast cancer image classification. MGSVM excelled in sensitivity rate, CUBIC SVM achieved the highest AUC score, and XGBOOST showcased a balanced accuracy in classifying both benign and malignant instances.

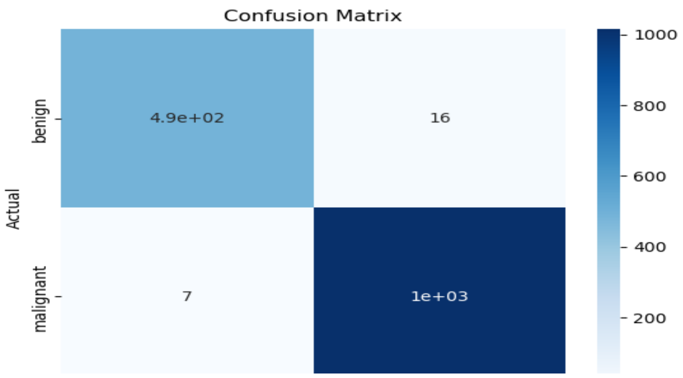


Fig 8 Confusion matrix of MGSVM with Deep Feature Extractions from EfficientNetB2 using INbreast Dataset

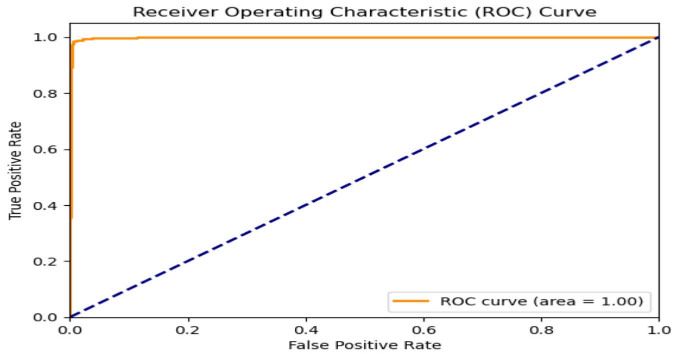


Fig 9 Receiver Operating Character of MGSVM with Deep Features Extraction from EfficientNetB2 using INbreast Dataset

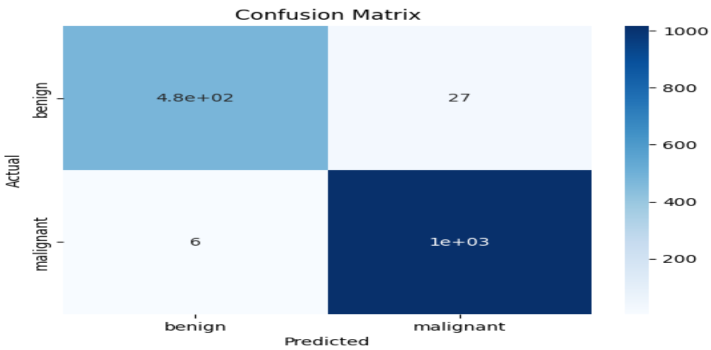


Fig 10 Confusion matrix of CUBIC SVM with Deep Features Extractions from EfficientNetB2 using INbreast Dataset

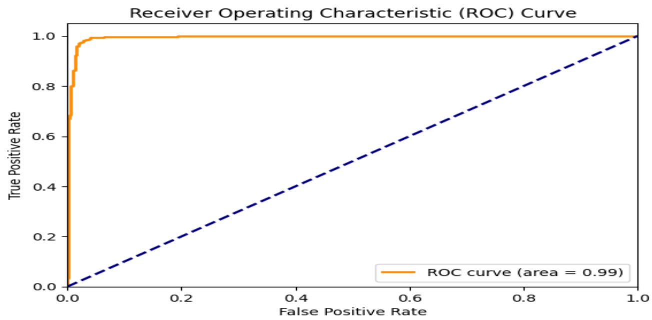


Fig 9 Receiver Operating Character of CUBIC SVM with Deep Features Extraction from EfficientNetB2 using INbreast Dataset

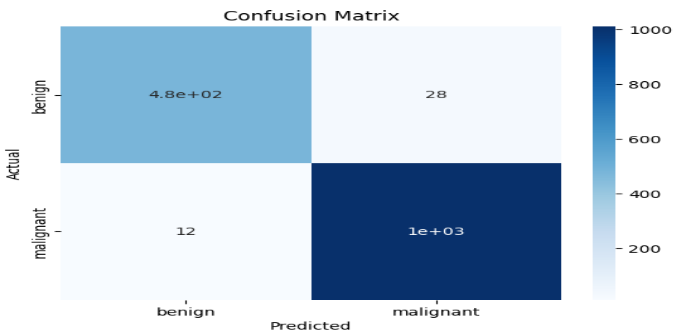


Fig 11 Confusion matrix of XGBOOST with Deep Features Extractions from EfficientNetB2 using INbreast Dataset

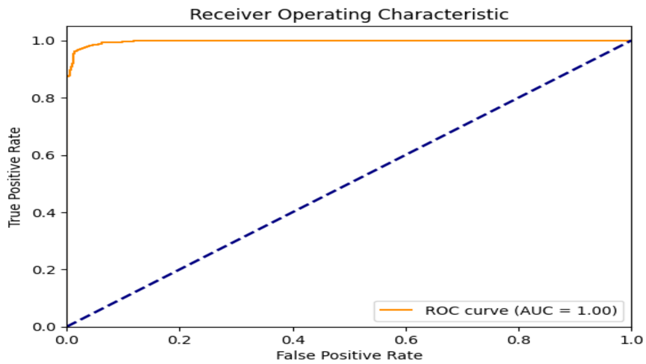


Fig 9 Receiver Operating Character of XGBOOST with Deep Features Extraction from EfficientNetB2 using INbreast Dataset

COMPARING RESULT WITH STATE OF THE ART MODELS

Table 3: Accuracies of State of the Art Models in Breast Cancer Classification

Author	Model	Accuracy
Deniz et al. (2018)	Transfer Learning Based Histopathologic Image Classification for Breast Cancer Detection	91.37%
Khan, Sanaullah, et al (2019)	"A Novel Deep Learning Based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning	97.525%
Togaçar, Mesut, et al (2020)	BreastNet: A Novel Convolutional Neural Network ˘ Model through Histopathological Images for the Diagnosis of Breast Cancer	98.80%
Proposed Model	**EfficientNetB2 & MGSVM Classifier**	**99.47%**

When comparing the proposed model to other state-of-the-art models in the realm of breast cancer image classification, we noted its strong performance across multiple crucial metrics. The model outperformed others in terms of accuracy, achieving an impressive rate of 99.47%, as demonstrated in Table 3. This outcome underscores the proficiency of our proposed model in accurately classifying instances, surpassing the capabilities of other models.

CONCLUSION

Given that both the INbreast and MIAS datasets are independent of each other, the performance of the fine-tuned EfficientNetB2 model with the MGSVM classifier across these datasets indeed indicates a certain level of generalizability. This observation becomes particularly significant because the model has shown consistent high performance on distinct datasets, suggesting its ability to adapt and perform well in diverse settings. The generalizability of a model is a highly desirable trait in machine learning, as it implies that the model has learned meaningful patterns and features that transcend specific dataset characteristics. In this context, the fine-tuned EfficientNetB2 model's success in maintaining strong performance on both the INbreast and MIAS datasets suggests that it has captured essential features relevant to breast cancer classification, which are applicable across different datasets. To establish the model's generalizability definitively, future research could involve testing the model on datasets with varying characteristics, such as different imaging modalities, diverse populations, and potential differences in data quality. Rigorous evaluation across these diverse datasets would provide a more comprehensive understanding of the model's ability to adapt and perform consistently across different contexts. In conclusion, while the fine-tuned EfficientNetB2 model's performance on the INbreast and MIAS datasets is an encouraging sign of its generalizability, it's necessary to continue assessing its performance on additional independent datasets to make more robust claims about its ability to be considered a "generic" model. The journey towards confirming such generalizability involves extensive validation and testing to ensure its reliability and effectiveness across a wide spectrum of real-world scenarios.

REFERENCES

Azamjah, N., Soltan-Zadeh, Y., & Zayeri, F. (2019). Global trend of breast cancer mortality rate: a 25-year study. Asian Pacific Journal of Cancer Prevention: APJCP, 20(7), 2015.

Abbas, Q. (2016). DeepCAD: A computer-aided diagnosis system for mammographic masses using deep invariant features. Computers, 5(4), 28.

Aljuaid, H., Alturki, N., Alsubaie, N., Cavallaro, L., & Liotta, A. (2022). Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. Computer Methods and Programs in Biomedicine, 223, 106951.

Almalki, Y. E., Soomro, T. A., Irfan, M., Alduraibi, S. K., & Ali, A. (2022, April). Computerized Analysis of Mammogram Images for Early Detection of Breast Cancer. In Healthcare (Vol. 10, No. 5, p. 801). MDPI.

Charan, S., Khan, M. J., & Khurshid, K. (2018, March). Breast cancer detection in mammograms using convolutional neural network. In 2018 international conference on computing, mathematics and engineering technologies (iCoMET) (pp. 1-5). IEEE.

Celik, Y., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2020). Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. Pattern Recognition Letters, 133, 232-239.

Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V., & Budak, Ü. (2018). Transfer learning based histopathologic image classification for breast cancer detection. Health Information Science and Systems, 6, 1-7.

Hekal, A. A., Elnakib, A., & Moustafa, H. E. D. (2021). Automated early breast cancer detection and classification system. Signal, Image and Video Processing, 15, 1497-1505.

Hikmah, N. F., Sardjono, T. A., Mertiana, W. D., Firdi, N. P., & Purwitasari, D. (2022). An Image Processing Framework for Breast Cancer Detection Using Multi-View Mammographic Images. EMITTER International Journal of Engineering Technology, 136-152.

Jabeen, K., Khan, M. A., Balili, J., Alhaisoni, M., Almujally, N. A., Alrashidi, H., ... Cha, J. H. (2023). BC2NetRF: breast cancer classification from mammogram images using enhanced deep learning features and equilibrium-jaya controlled regula falsi-based features selection. Diagnostics, 13(7), 1238.

Khan, S., Islam, N., Jan, Z., Din, I. U., & Rodrigues, J. J. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognition Letters, 125, 1-6.

Karthiga, R., Narasimhan, K., & Amirtharajan, R. (2022). Diagnosis of breast cancer for modern mammography using artificial intelligence. Mathematics and Computers in Simulation, 202, 316-330.

Li, H., Niu, J., Li, D., & Zhang, C. (2021). Classification of breast mass in two-view mammograms via deep learning. IET Image Processing, 15(2), 454-467.

Lotter, W., Diab, A. R., Haslam, B., Kim, J. G., Grisot, G., Wu, E., ... Wang, M. (2021). Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. Nature Medicine, 27(2), 244-249.

Laisné, M. (2019), "Breast Cancer", Mendeley Data, V1, doi: 10.17632/jkg5h9h7bz.1.

Medeiros, G. C., Thuler, L. C. S., & Bergmann, A. (2019). Delay in breast cancer diagnosis: a Brazilian cohort study. Public Health, 167, 88-95.

Sha, Z., Hu, L., & Rouyendegh, B. D. (2020). Deep learning and optimization algorithms for automatic breast cancer detection. International Journal of Imaging Systems and Technology, 30(2), 495-506.

Ting, F. F., Tan, Y. J., & Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. Expert Systems with Applications, 120, 103-115.

Toğaçar, M., Özkurt, K. B., Ergen, B., & Cömert, Z. (2020). BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. Physica A: Statistical Mechanics and its Applications, 545, 123592.

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2020). ChestX- ray8:

Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2097-2106).

Wahab, N., Khan, A. and Lee, Y.S. (2019). Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images. Microscopy, 68(3), pp.216-233.