A Hybrid Approach with Xception and NasNet for Early Breast Cancer Detection

Yassin Benajiba, Mohamed Chrayah, Yassine Al-Amrani
TIMS Laboratory, Abdelmalek Essaadi University, Tetouan, Morocco

***Abstract*—Breast cancer is the most common cancer in women, accounting for 12.5% of global cancer cases in 2020, and the leading cause of cancer deaths in women worldwide. Early detection is therefore crucial to reducing deaths, and recent studies suggest that deep learning techniques can detect breast cancer more accurately than experienced doctors. Experienced doctors can detect breast cancer with only 79% accuracy, while machine learning techniques can achieve up to 91% accuracy (and sometimes up to 97%). To improve breast cancer classification, we conducted a study using two deep learning models, Xception and NasNet, which we combined to achieve better results in distinguishing between malignant and benign tumours in digital databases and cell images obtained from mammograms. Our hybrid model showed good classification results, with an accuracy of over 96.2% and an AUC of 0.993 (99.3%) for mammography data. Remarkably, these results outperformed all other models we compared them with, Top of Form ResNet101 and VGG, which only achieved accuracies of 87%, 88% and 84.4% respectively. Our results were also the best in the field, surpassing the accuracy of other recent hybrid models such as MOD-RES + NasMobile with 89.50% accuracy and VGG 16 + LR with 92.60% accuracy. By achieving this high accuracy rate, our work can make a significant contribution to reducing breast cancer deaths worldwide by helping doctors to detect the disease early and begin treatment immediately.**

Keywords*—*Breast Cancer; CNN; Hybrid Model: Xception; NasNet

I. Introduction

Breast cancer is a serious threat to women, especially those over the age of 50. It is estimated that breast cancer accounts for 25% of all cancer cases and will affect approximately one in ten women during her lifetime. The International Agency for Research on Cancer (IARC) reports that in 2020 there will be 2.3 million women diagnosed with breast cancer and 685,000 deaths worldwide. By the end of 2020, there will be 7.8 million women alive who have been diagnosed with breast cancer in the last five years, making it the most common cancer in the world, although the incidence of breast cancer varies widely around the world. [1] These statistics highlight the importance of continued research and efforts to improve breast cancer diagnosis and treatment. The use of machine learning in this area is particularly relevant and has shown promise in increasing the accuracy of predictions, making it a valuable tool in the fight against this disease.

However, reducing the mortality rate caused by this type of cancer as well as increasing the chances of recovery are only possible if the tumor has been detected at the early stages of its appearance. And to ensure the early detection of such a tumor, one needs to detect it in the early periods, since scientific researchers have found that early diagnosis greatly increases the chances of survival, before these malignant (cancer) cells multiply [2] in a disordered way until creating a tumor which attacks the neighboring tissues. When a breast cancer is not treated, the tumor cells spread locally and invade the neighboring organs (local extension then regional extension). Early diagnosis can therefore increase the chances of treatment before the doctor has to amputate a woman's entire breast. [3].

In recent years, the intersection of medicine and technology has paved the way for innovative approaches to breast cancer detection. Studies have shown the promising potential of deep learning techniques to improve diagnostic accuracy [4], surpassing the capabilities of even experienced medical professionals. While experienced clinicians typically achieve an accuracy rate of up to 79%, machine learning algorithms have demonstrated remarkable capabilities, with accuracy rates as high as 91% and, in some cases, 97% [5].

Motivated by the imperative to improve breast cancer diagnosis, our research delves into the realm of deep learning models, specifically exploring the efficacy of Xception and NasNet. By synergising these models, we aimed to increase classification accuracy in distinguishing between malignant and benign tumours, using digital databases of cell images obtained from mammography. The culmination of our efforts resulted in a hybrid model that showed promising results, with an accuracy of over 96.2% and an impressive AUC (area under the curve) of 0.993 for mammography data.

The research paper was designed in a way that facilitates the scientific journey for the discerning reader. Adding this introduction is given in Section I, Section II delves into the broader world of related work, where we provide an overview of the current state of research, the contributions of different researchers, the techniques of previous studies, and the results obtained in this field. Section III highlights the methodologies we chose in our research, how we divided the data, and ways to determine the effectiveness of our mixed model. We reveal the impressive results we have reached in Section IV by comparing them to the results of other research that share the same goal and research. Finally, Section VI summarizes our conclusions and the obstacles we encountered, while also pointing out future research paths.

Comparative analysis positioned our hybrid model as a leader in the field, outperforming established models such as ResNet101 and VGG, which achieved accuracy rates of only 87%, 88% and 84.4% respectively. Furthermore, our results outperformed recent hybrid models such as MOD-RES +

NasMobile and VGG 16 + LR, underlining the significance of our contribution.

Beyond academia, the implications of our work have profound potential for real-world impact. By achieving exceptional accuracy rates, our research can serve as a critical tool in the arsenal of healthcare professionals, enabling early detection and prompt initiation of treatment protocols. Ultimately, our efforts aim to reduce the burden of breast cancer on a global scale, offering hope in the quest to reduce mortality and improve patient outcomes.

II. Related Work

This section provides a systematic review of scientific research in the field of medical image classification, with the aim of highlighting the contributions of different researchers and improving understanding of research methods, study techniques and results. It evaluates and compares different studies, highlighting the work of individual researchers and providing context for their findings. It also compares and evaluates these findings with the conclusions of other researchers in the same field, providing an insight into the current state of medical image classification. The results of this review will serve as a valuable resource for researchers and practitioners, providing a comprehensive overview of the field, the methods used by experts, and comparisons of results. Numerous studies have highlighted the importance of medical image classification in the early detection and accurate diagnosis of tumours and diseases, thereby improving medical diagnosis, treatment efficacy and reducing mortality rates.

In a study entitled "Automated Breast Cancer Detection Models Based on Transfer Learning" [6], the researchers segmented the data and applied some techniques to it. The best result they obtained when they applied the hybrid model technique between MOD-RES and NasMobile was 89.5% accuracy. This is the highest accuracy value compared to the rest of the other techniques used in the same research and under the same conditions.

In a study entitled: "Boosting breast cancer detection using convolutional neural network" [7], the researchers carried out a proposed prototyping approach in which they used different convolutional neural network (CNN) structures to automatically detect breast cancer and compared the results with those of machine learning (ML) algorithms… After searching and comparing, they found that their model produced results with an accuracy 87% higher than that of machine learning (ML) algorithms, which had an accuracy of only 78%. Thus, the system proposed in their paper improves accuracy by 9% over the results of machine learning (ML) algorithms.

In their paper "Deep Learning RN-BCNN Model for Breast Cancer BI-RADS Classification" [8], the researchers used a combination of augmentation strategies to prevent overfitting and improve the accuracy of mammogram analysis, including rotation, scaling and displacement. The proposed system was evaluated on the MIAS dataset and achieved 88% accuracy using the pre-trained ResNet101 classification network and 70% accuracy using the Nasnet-Mobile network. The study suggests that pre-trained classification networks are more effective and efficient for medical imaging, especially when dealing with small training datasets.

In a scientific paper on medical image classification entitled "Transfer learning and fine tuning in the classification of radiographic mammary abnormalities on the CBIS-DDSM database" [9], the researchers used NasNet and MobileNet to train the mammary abnormality classifier and then compared models such as : VGG, Resnet and Xception and compared them. They found that VGG16 achieved the best accuracy compared to other models, which was 0.844 in the CBIS-DDSM dataset.

In a paper entitled "Breast cancer histology images classification: Training from scratch or transfer learning?" [10], the researchers compared the use of transfer learning and fully trained networks for breast cancer classification using histopathological images. They analysed three pre-trained networks (VGG16, VGG19 and ResNet50) for their breast cancer classification behaviour independent of magnification, and examined the effect of training and test data size on their performance. They found that pre-trained VGG16 with a logistic regression classifier had the best performance, with an accuracy rate of 92.60% and an AUC of 95.65%.

III. Proposed System

In our research, we have proposed a new model for image classification that combines the strengths of two of the best models available. By using the technique of combining them, as shown in the figure, we have created a new classification model that outperforms the use of either model in isolation.

What is unique about our approach is that the two models are not simply run sequentially, but are integrated in parallel. This means that the outputs of both models are combined in a way that captures the strengths of each model while mitigating its weaknesses. The result of this integration is a model that is more accurate and robust than either model alone. In other words, our proposed model represents a significant improvement over current state-of-the-art image classification methods.

To arrive at this model, we carried out extensive experimentation and analysis to determine the optimal configuration and parameters for the two models, as well as the best way to combine their outputs. The result is a model that is not only more accurate, but also more efficient, processing images faster and using fewer computing resources. Fig. 1 shows schematic diagram of the propose system.

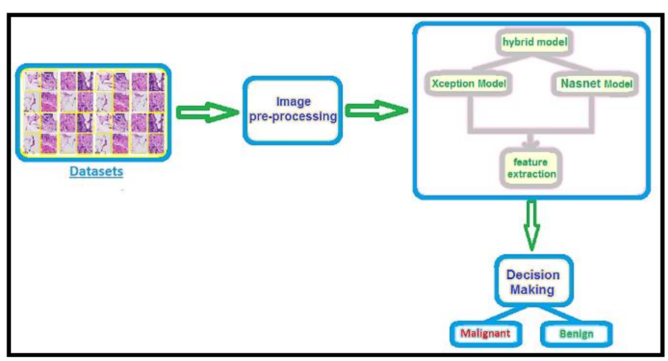

Fig. 1. Schematic diagram of the proposed system.

We believe that our proposed model has the potential to advance the field of image classification and contribute to a wide range of applications, from medical imaging to autonomous driving.

A. Datasets

In this work, datasets taken from the kaggle database were used, and this database consisted of 277,524 loci of 48 × 48 size divided into (198,738 negative IDCs and 78,786 positive IDCs), extracted from The original dataset was from 162 whole slice mammograms, in which cancer (BCa) samples were scanned 40 times, and the database was then extracted, segmented, and classified [11]. Fig. 2 shows example of data images.

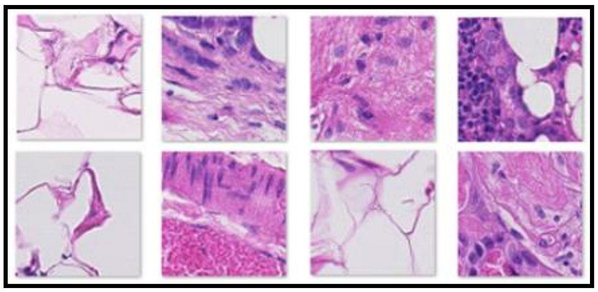

Fig. 2. Example of data images.

The archive is a 3 GB Winrar file, it contains 278 subfolders, and in each of these folders there are two more files: one labeled "0" containing the negatives, and the other labeled. The "1" file contains positive images, the name of each image is: u_xX_yY_classC.png where u is the patient ID, X is the x coordinate of where this patch was cropped, Y is the y coordinate of where this patch was cropped, and C denotes a class where 0 is not idc and 1 is idc.

B. Images Pre-processing

In order to improve the performance and robustness of our proposed model and prevent overfitting, we utilized various techniques to augment and optimize the training data set, including data augmentation, image optimization, rescaling, normalization, and other methods. As the complexity of the model network increased, so did the number of parameters that needed to be learned, making it even more susceptible to overfitting. Therefore, we increased the variety and size of the training data set by implementing several techniques, such as rescaling the pixel values of input images to a range of [0,1] using a factor of 1/255 to improve numerical stability and convergence. We also flipped input images horizontally or vertically and rotated them by 90 degrees to increase diversity and reduce the risk of over fitting. Furthermore, we multiplied the number of training images by a factor of ×2 to zoom in twice to improve the model's ability to handle changes in image scale.

C. Proposed Learning Methods

In this research project, we aim to develop a hybrid model by combining the Xception and NasNet models for parallel and integrated image classification. This approach will lead to improved performance compared to using each model separately. The NasNet model is a deep learning architecture developed by Google's AI researchers and optimised for high accuracy object detection and image classification. It is designed to be computationally efficient, making it ideal for real-time applications such as self-driving cars, robotics and video surveillance. The Xception model, also developed by Google AI researchers in 2016, is a variation of the Inception model designed for advanced accuracy in image classification tasks. It is more efficient than the Inception model, making it suitable for real-time image recognition, natural language processing and object detection applications. The combination of these two models will lead to significant improvements in the early detection of breast cancer by classifying benign and malignant images, as demonstrated by the results of our study.

D. Assessment Metrices

To evaluate the effectiveness of our proposed hybrid model, we compared its performance with that of several unilateral models, as well as with the performance of other hybrid models used by other researchers in image classification for breast cancer detection. We used a number of performance metrics to make this comparison, including the following:

1) Accuracy and loss: When evaluating classification models, an important criterion is their accuracy, which is a measure of how many of their predictions are correct [12] More formally, accuracy refers to the proportion of correct predictions made by the model.

In other words, accuracy is a measure of how well the model is able to correctly classify different types of input. The higher the accuracy, the more reliable the model is likely to be. [13] For this reason, accuracy is often used as a benchmark for comparing different classification models.

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

- True positives (TP): images of benign tumors correctly classified.
- True negatives (TN): images of malignant tumors correctly classified .
- False positives (FP): Images of benign tumors were not correctly classified.
- False negatives (FN): Images of malignant tumors were not correctly classified.

Loss is a technique used to assess how well an algorithm is able to model a given set of data. If the algorithm's predictions are too far from the actual results, the loss function will return a high value. [14] Over time, using an optimisation function, the loss function is able to learn how to minimize the prediction error by adjusting the algorithm's parameters. [15].

In other words, the loss function is a measure of how well the algorithm is able to approximate the correct output for a given input. [16] The optimisation function works by adjusting the parameters of the algorithm to minimise the loss, which in turn improves the accuracy of the model. This process is often repeated many times until the loss is reduced to an acceptable level, at which point the algorithm is considered to have learned the underlying patterns in the data.

2) The ROC Curve (Receiver Operational Characteristic): The ROC curve, short for receiver operating characteristic, is a function used to measure the performance of a binary classifier - a system that classifies items into two groups based on their characteristics. This function is also known as a performance characteristic or sensitivity/specificity curve. [17]. Fig. 3 shows the ROC curve.

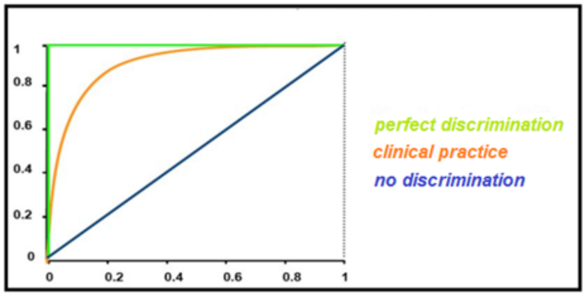

Fig. 3. The ROC curve.

The ROC curve is usually plotted as a curve with the true positive rate (the proportion of true positives that are correctly identified) on the y-axis and the false positive rate (the proportion of true negatives that are incorrectly identified) on the x-axis. [18] This curve provides a visual representation of the performance of the classifier by showing how well it is able to distinguish between the two groups.

In summary, the ROC curve is a tool used to evaluate the performance of a binary classifier. By plotting the true positive and false positive rates, it provides a clear and concise representation of the classifier's ability to distinguish between the two groups.

IV. Results and Discussion

We would like to present the findings of various experiments conducted on our innovative hybrid system, which was applied to a set of images that had been divided into training and testing data with a ratio of 80–20%. In each trial, we modified the parameters of the model and applied it, resulting in a range of different outcomes. We then meticulously analyzed and compared these outcomes with the aim of improving our model and identifying the best possible configuration. Below, we provide a comprehensive summary of the results we obtained from each of these trials using their respective parameter settings:

According to the analysis of the results of the tables, we note that:

For Batch size= 100 and Batch size = 200, the values for accuracy and AUC increased until epoch 15 and then decreased, and for error it was always decreasing. This indicates that optimal performance for these batch sizes was achieved on or around Epoch 12.

For Batch size = 215, the accuracy increased until epoch 15 and then decreased, while the error decreased and then stabilized to a value of 0.113. AUC followed a similar pattern to the other batch sizes, increasing up to Epoch 15 and then decreasing.

For Batch size = 216, the accuracy increased until epoch 16 and then decreased, while the error decreased until epoch 16 and then increased. AUC followed a similar pattern for other batch sizes, increasing until epoch 16 and then decreasing. This indicates that optimal performance for these batch sizes was achieved on or around Epoch 15. Table I shows the summary table of the results obtained by our hybrid model.

TABLE I. Summary Table of the Results Obtained by Our Hybrid Model

Test	Image size	Batch size	Epochs	Accuracy	Loss	Auc
1	(96,96,3)	100	5	0,941	0,138	0,985
2			12	0,948	0,132	0,989
3			15	0,946	0,133	0,987
4		200	5	0,954	0,123	0,985
5			12	0,957	0,114	0,987
6			15	0,958	0,108	0,986
7		215	5	0,956	0,114	0,989
8			12	0,957	0,113	0,992
9			15	0,956	0,113	0,989
10		216	5	0,954	0,124	0,986
11			12	0,958	0,103	0,991
12			15	0,959	0,102	0,993
13			16	0,954	0,104	0,991
14		217	5	0,942	0,134	0,978
15			12	0,954	0,114	0,991
16			15	0,954	0,113	0,99
17		220	5	0,952	0,125	0,988
18			12	0,954	0,117	0,989
19			15	0,955	0,115	0,991

For Batch size = 217 and Batch size = 220, the accuracy decreases and then continues to increase, the error continues to decrease, and the AUC continues to rise until the epoch 15. But its value does not reach the rest of the values of other epochs.

After conducting a thorough analysis and comparing all the results obtained in the parametric study, we have determined that the optimal test for our hybrid model is the one with the following parameters:

Image size = (96 ;96 ;3) ; Batch size = 216 ; Epochs = 15

In order to obtain accurate and reliable results from our analysis, we conducted a comprehensive study by closely observing the evolution of three key metrics: accuracy, loss, and ROC. To ensure that our analysis was as rigorous as possible, we carefully examined the performance of our model from the very first epoch up until the best performing epoch. By doing so, we were able to identify the optimal point at which our model achieved its highest levels of accuracy, lowest levels of loss, and best ROC score. Table II shows the results that we reached during this comparison:

Upon close observation, we have noted that the parameters of Accuracy and Val_Accuracy have consistently displayed a positive and upward trend, with only occasional and minor declines noted at some stages. The general pattern, however, indicates that these performance indicators have achieved the best results possible. Furthermore, Loss and Val_Loss demonstrate a reliable and steady decrease during the implementation stages of the hybrid model utilizing the carefully selected parameters. It is worth noting that after conducting multiple comparisons, we arrived at the conclusion that these chosen parameters were the best for the given task. Thus, it is apparent that the application of the hybrid model with these specific parameters has yielded promising and encouraging results.

TABLE II. ACCURACY, LOSS, VAL_ACCURACY AND VAL_LOSS FOR THE SPECIFIED PARAMETERS

Epoch	Accuracy	Loss	Val_ Accuracy	Val_ Loss
1	0.88366	0.27512	0.92305	0.22194
2	0.93020	0.18116	0.92973	0.17954
3	0.94008	0.15606	0.93808	0.15820
4	0.94816	0.13721	0.95076	0.13379
5	0.95331	0.12474	0.94688	0.14112
6	0.95766	0.11403	0.94814	0.13775
7	0.96111	0.10500	0.95967	0.10887
8	0.96413	0.09785	0.96032	0.10932
9	0.96651	0.09168	0.95764	0.11423
10	0.96950	0.08480	0.95367	0.12746
11	0.97079	0.08029	0.96061	0.10963
12	0.97255	0.07499	0.96011	0.11079
13	0.97403	0.07096	0.96035	0.11164
14	0.97533	0.06706	0.96047	0.11161
15	0.97658	0.06397	0.96285	0.10786

Furthermore, we made sure that the choice of this best performing epoch was the most suitable for the subsequent epochs that followed. In other words, we conducted an extensive analysis to ensure that the performance of the model did not deteriorate after the best performing epoch, as this would indicate that our choice was not the most optimal. In order to conduct a comprehensive analysis of our hybrid model, we closely monitored the accuracy, loss, and ROC curves. To ensure accurate results, we used optimal test parameters and kept all other parameters constant except for three instances when we made changes. By doing this, we gained a comprehensive view of how the model evolved over time and how it performed under different conditions. Fig. 4 shows the evolution of some of the curves selected from among the many curves obtained:

Our study yielded a significant finding related to the performance of our hybrid model. We were able to observe the evolution of the accuracy and ROC curves over time as we adjusted the model's parameters. With each adjustment, we noted a steady improvement in the model's performance, and the accuracy and ROC curves reached their best values at the selected parameters. In addition to this, we also observed a corresponding decrease in the loss curve as we approached the best model for our purposes.

These results provided us with valuable insights into the strength of our hybrid model's selected parameters, which allowed us to fine-tune and improve its performance for future use. Overall, our approach enabled us to effectively assess the effectiveness of our model and make informed decisions about how best to improve it to obtain the best results. By carefully analyzing the performance metrics, we were able to optimize the parameters of our hybrid model and ensure that it delivered accurate results. This outcome is critical for future use, as we can now rely on our hybrid model to provide accurate predictions and insights.

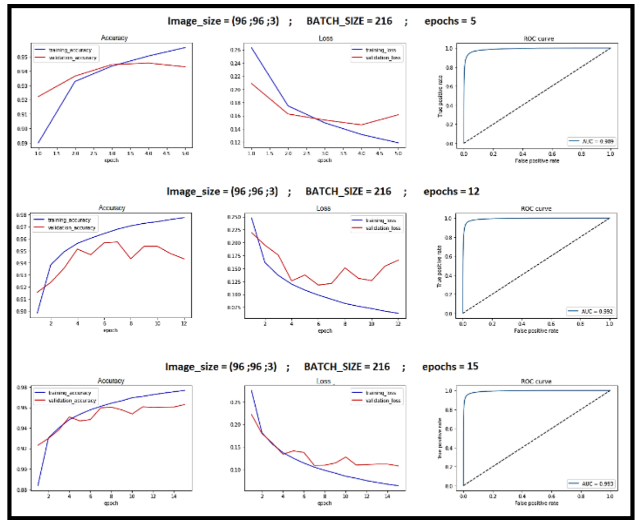

Fig. 4. Accuracy, Loss, and Roc curves by epoch.

V. COMPARED TO OTHER MODELS

The proposed system's effectiveness and dependability are assessed against the latest research on image classification systems for breast cancer screening. In this section, we showcase the results of the proposed system (a hybrid model combining Xception and NasNet deep learning models) and juxtapose them against existing methods (see Table III). The findings in Table III reveal that the proposed system yields considerably more accurate outcomes than the techniques we compared it with. Furthermore, our proposed system surpasses other hybrid models listed in the table with respect to accuracy.

TABLE III. COMPARISON OF OUR MODEL WITH OTHER MODELS IN TERMS OF ACCURACY

Author (year)	Technique	Accuracy
Our hybrid model	Xception + NasNet	96.20%
Madallah Alruwaili and Walaa Gouda (2022) [6]	MOD-RES + NasMobile	89,50%
Saad Awadh Alanazi (2021) [7]	CNN	87%
Shahbaz Siddeeq (2021) [8]	ResNet101	88 %
Lenin G. Falconí, M. Pérez (2020) [9]	VGG	84,4%
Shallu and R. Mehra (2018) [10]	VGG16 + LR	92,60%

In previous studies, researchers used different models to classify data and achieve a high accuracy rate. However, none of those models were able to exceed an accuracy rate of 89.5%, except for the VGG 16 + LR hybrid model, which achieved an accuracy rate of 92.60%. Despite this achievement, our hybrid model was able to surpass the accuracy rate of all the other models, including VGG16 + LR, with an accuracy rate of 96.2%. This makes our hybrid model the best performing model among all the models tested, as it has the highest classification accuracy. Fig. 5 shows this superiority more clearly:

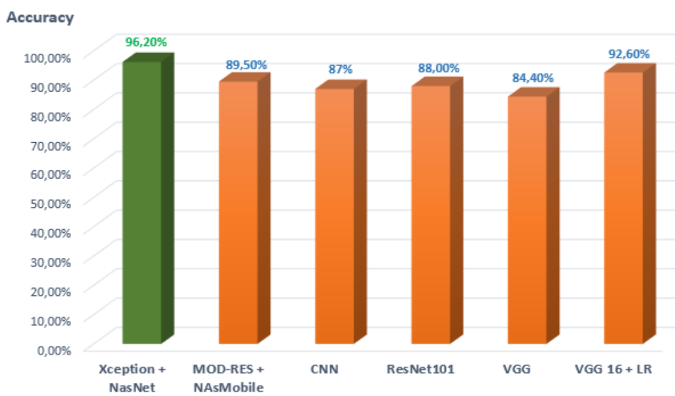

Fig. 5. Accuracy comparison between models.

After analyzing the curve, we found that our hybrid model outperformed all the other models used for classification, with the highest accuracy rate. This is an important finding, as it confirms that our model is superior in identifying and diagnosing breast cancer.

The VGG16 + LR hybrid model ranked second in terms of accuracy, which is consistent with our previous results. The accuracy rates of the other models used in our study did not exceed the accuracy of our hybrid model or the VGG16 + LR hybrid model. However, our model outperformed the VGG16 + LR hybrid model in terms of both accuracy and AUC. Our model achieved an AUC of 99.3%, while the VGG16 + LR hybrid model achieved an AUC of only 84.4%. [9].

This result is significant as it affirms that our model can accurately classify medical images, which will facilitate the process of identifying breast cancer with high accuracy and in the early stages of its existence. This, in turn, will greatly assist doctors in determining the presence of the disease or not, which can lead to early interventions and treatments.

VI. CONCLUSION AND FUTURE DIRECTION

In this study, we utilized a cutting-edge technique by amalgamating two deep learning models, NasNet and Xception, after conducting thorough individual evaluations of each model. We then fine-tuned the parameters to attain the most optimal outcomes. Afterwards, we compared our results with various scientific research papers that focused on classifying breast cancer medical images. Our findings revealed that the model we designed outperformed all other models in terms of accuracy and Area Under the Curve (AUC), with an accuracy of 96.2% and an AUC of more than 99.3%. By achieving this high accuracy rate, our work can make a significant contribution to reducing breast cancer deaths worldwide by helping doctors to detect the disease early and begin treatment immediately.

Although we have succeeded in creating a very precise and accurate model, it takes a long time to apply it to show results. This has led us to look forward to future research focusing on the same model, but with a strong focus on reducing its application time while maintaining its exceptional accuracy and uniqueness. Our plan is to test a number of approaches to determine the most effective ways to achieve this goal. Through extensive research and testing, we aim to simplify the process of implementing our model while ensuring it remains effective and accurate.

REFERENCES

[1] Melina A; Eileen M; Harriet R, “Current and future burden of breast cancer: Global statistics for 2020 and 2040”. December 2022.

[2] Maria S; David P, “Artificial Intelligence in Cancer Research: Trends, Challenges and Future Directions ”. December 2022.

[3] Xin J; Ping M, “Targeting Breast Cancer Metastasis”. September 2015.

[4] Y.Benajiba, M.Chrayah, Y.Al-Amrani, “A Nonlinear Support Vector Machine Analysis Using Kernel Functions for Nature and Medicine”,2021.

[5] Gouda, W.; Selim, M.M.; Elshishtawy, T, “An Approach for Breast Cancer Mass Detection in Mammograms”. January 2012.

[6] Alruwaili, M.; Gouda, W, “Automated Breast Cancer Detection Models Based on Transfer Learning”. January 2022,

[7] Alanazi, S.A.; Kamruzzaman, M.M; Sarker N.I; Alruwaili.M ; Alhwaiti, Y.Alshammari, N.; Siddiqi, M.H, “Boosting breast cancer detection using convolutional neural network”, Apr 2021.

[8] Siddeeq, S.; Li, J.; Ali Bhatti, H.M.; Manzoor, A.; Malhi, U.S, “Deep Learning RN-BCNN Model for Breast Cancer BI-RADS Classification”, January 2021.

[9] Lenin G.F ; Maria P ; Wilbert G.A ; Aura C, “Transfer Learning and Fine Tuning in Breast Mammogram Abnormalities Classification on CBIS-DDSM Database’’, March 2020.

[10]] Shallu. Rajesh M, "Breast cancer histology images classifcation: Training from scratch or transfer learning?", November 2018.

[11] PAUL MOONEY, Breast Histopathology Images Dataset, https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images

[12] Zeljko V, “Classification Model Evaluation Metrics”. July 2021.

[13] Singh L. Alam A, “An efficient hybrid methodology for an early detection of breast cancer in digital mammograms”, May 2022.

[14] Tian Y, Duo Su, Lauria S, Xiaohui L, “Recent advances on loss functions in deep learning for computer vision”. August 2022.

[15] Juan T; Diana M.C.E: Alfonzo R.P; Edgar A.C.U, “Loss Functions and Metrics in Deep Learning. A Review”. July 2023.

[16] Qi W; Yue M; Kun Z; Yingjie T, “A Comprehensive Survey of Loss Functions in Machine Learning”, April 2020.

[17] Narkhede S, “Understanding AUC - ROC Curve”, Jun 2018.

[18] Tom F, “Introduction to ROC analysis”, June 2006.