UTM Computing Proceedings
Innovations in Computing Technology and Applications
Volume: 6 | Year: 2024 | ISBN No. : 2976-2278

# BREAST CANCER CLASSIFICATION USING DEEP LEARNING

Muhammad Faris Danial Bin Mohd Yusof, Noor Hidayah Zakaria
Faculty of Computing,
Universiti Teknologi Malaysia
Johor Bahru, Malaysia
mfdanial2@graduate.utm.my, noorhidayah.z@utm.my

***Abstract*— The main goal of this research is to develop and compare the accuracy of classification of breast cancer using ResNet50 V2, VGG19 and Convolutional Neural Networks (C-NN). The reviewed studies shed light on the various ways in predicting breast cancer by making use of machine learning techniques. The contribution of this research is it provides or suggests a solution to classify breast cancer using VGG19, ResNet50 V2 and C-NN in order to improve and enhance the accuracy to predict and detect breast cancer. Using machine learning, we can learn more about predicting breast cancer in the early stage using machine learning. The dataset that is being used for this research is the CBIS-DDSM dataset.**

***Keywords-Breast Cancer, Deep Learning, Convolutional Neural Network,CBIS-DDSM, ResNet50V2, VGG19***

## I. INTRODUCTION

Breast cancer is a disease characterized by the uncontrollable growth of breast cells. Different types of breast cancer arise depending on which abnormal cells become cancerous. The medical industry possesses an immense amount of data, but due to limited technological advancements in the past, drawing definitive conclusions from this data has been challenging. However, with the advent of technology, we now have the potential to anticipate and diagnose breast cancer, which is crucial for early detection and improved patient survival rates. This paper aims to explore how machine learning techniques can contribute to predicting and detecting breast cancer more accurately. While mammograms, the primary diagnostic tool, may not be 100% accurate, machine learning algorithms can leverage large datasets of mammography measurements to improve precision. By improving the breast cancer diagnosis system's accuracy through machine learning, we can make significant advancements in patient outcomes and streamline the therapy process.

Breast cancer is the most common cancer worldwide, affecting women of all ages after puberty. In 2020, the World Health Organization reported 2.3 million new cases and 685,000 deaths globally. By the end of the same year, approximately 7.8 million women had survived breast cancer for five years or more. Early detection is crucial for effective treatment, but the lack of technological advancements in predicting and detecting cancer has been a challenge. Improved survival rates have been observed in countries with early prediction and detection strategies. Mammograms, commonly used for breast cancer detection, have limitations as they may miss small cancers or produce false-positive results. Therefore, leveraging advancements in technology, such as machine learning, can enhance the accuracy of breast cancer diagnosis by utilizing large datasets of mammogram measurements.

Deep learning techniques can assist in predicting breast cancer at the early stage. By analyzing the CBIS-DDSM dataset, machine learning algorithms can identify breast cancer cells. They can detect uncontrollable growth of breast cells on mammogram images and predict breast cancer. This helps patients to get treatment at the early stage.

## II. RELATED WORK

### A. *Classification in Machine Learning for Breast Cancer Diagnosing Technique*

Medical diagnostics were improved using optimization methods and machine learning technologies. This study aims to assess the accuracy of breast cancer detection using these recommended technologies. Even experienced radiologists can make small mistakes when interpreting mammograms and diagnosing breast cancer. Therefore, this research aims to develop a system that can accurately predict breast cancer on mammograms, helping radiologists avoid common errors during the BI-RADS diagnostic process. The dataset (CBIS-DDSM) can be analyzed more thoroughly and quickly. Classification involves organizing structured or unstructured data into classes. Classification predictive modeling draws conclusions from the training input data. The classifier uses various properties to determine the class of a sample.

In a study by Wenwei Zhao et al. (2021), Random Forest (RF) and Support Vector Machines (SVM) were the most commonly used classifiers for breast cancer prediction. Mutra et al. (2021) found the optimal combination of grayscale characteristics for various classifiers, including breast density classification. Since each classifier has its own advantages and disadvantages, it is possible to select the most effective classifier for a specific algorithm by evaluating performance measures such as sensitivity, specificity, and classification accuracy.

There are researchers that have studied and proposed methods for predicting breast cancer using machine learning from mammography images. Various techniques have been proposed to classify mammogram images. In order to identify the best technique for classification, we have compiled a table (Table I) that includes different classification techniques based on the methods employed, data sets used, and results obtained.

REVIEW OF CLASSIFICATION TECHNIQUES

| Publisher | Classifier | Dataset | Accuracy |
|---|---|---|---|
| Vishwanath et al | SVM | Raw Sample Images | 84.84 % |
| Beham et al (2019) | KNN | 149 Mammographic images from the pixel scan hub | 71.14 % |
| Shen et al (2019) | CNN | DDSM | 91.00% |

The first research paper was written by Shen et al, 2019 [1], that discusses how the study demonstrates the effectiveness of an end-to-end deep learning model for accurately classifying screening mammograms. The model can be fine-tuned using additional datasets without region of interest annotations, even with variations in image characteristics. Deep learning algorithms outperform traditional CAD systems, offering competitive and generalizable performance across different mammography platforms. The study also highlights limitations of existing commercial CAD systems and proposes an approach that requires only image-level labels for fine-tuning, allowing scalability and adaptation to evolving mammography systems. Future research can focus on improving patch sampling efficiency and exploring the combination of VGG-based and Resnet-based classifiers. Despite limitations, deep learning shows potential for enhancing breast cancer detection on mammograms and can be applied to other medical imaging problems with limited ROI annotations.

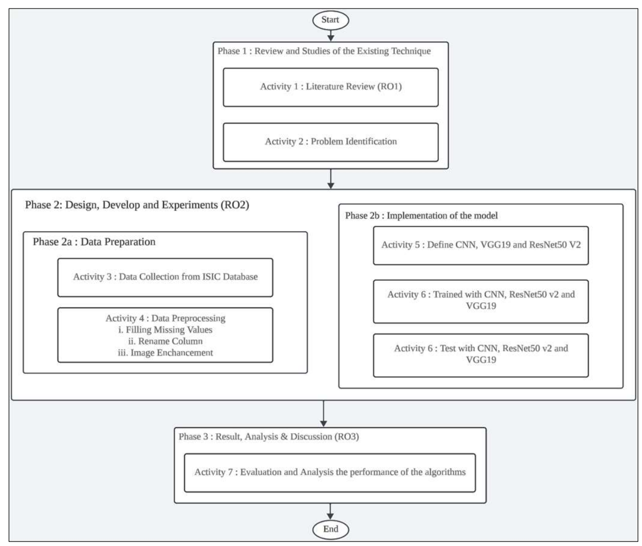
The second research paper by Beham et al, 2019[2], is about a study that introduces a straightforward method for the classification of breast cancer images in mammogram images. The proposed technique involves the extraction of highly discriminant local binary patterns from mammogram images that have been normalized using the wavelet transform. The resulting features are utilized to categorize abnormal cancer cell images using a K-nearest neighbor classifier. To assess the effectiveness of our algorithm, a dedicated mammogram database is established. The experimental findings demonstrate that our algorithms perform well, exhibiting a favorable balance between accuracy and computational efficiency.

The final research paper by Viswanath et al et. al, 2019, discusses the proposed approach aims to automate the classification and segmentation procedures in mammogram analysis, considering normal, benign, and malignant conditions. Supervised predictive models, namely Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), and Random Forest (RF), were tested to determine their accuracy in classifying mammogram image conditions. The evaluation revealed that RF achieved the highest accuracy metrics for both multi-class and binary classifications, using enhanced and raw images, emphasizing its reliability as a classifier for mammogram analysis. The research also emphasizes the importance of image pre-processing in improving classification accuracy. Despite promising results, low classification in specific cases needs to be addressed. Strategies such as incorporating a larger dataset, engineering new features, and modifying pre-processing parameters can be implemented for improvement. The ultimate goal is to develop a fully automatic system integrating image processing and classification techniques, serving as a valuable tool for radiologists in mammographic interpretation.

## III. DISCUSSION

### A. Operational Framework/Research Workflow

There are several phases that are involved in order to classify breast cancer using CNN, ResNet50 V2 and VGG19. The objective is to make sure that the research is done properly. Generally, the first phase centers on the studies of existing techniques and literature review. For the second phase, it is about the design and developing the algorithms and last but not least, the result analysis and discussion.




Flow of the methodology

**Phase 1: Review and Studies of Existing Technique**

This step consisted of gathering and analyzing relevant research and data from a range of sources, including journals, books, papers from SpringerLink, NCBI, IEEE, and ScienceDirect, among others. To provide a better grasp of the study issue in terms of the strengths and weaknesses of a technique employed in prior studies, a thorough evaluation should be conducted. Several breast cancer diagnostic tools have been developed in light of earlier studies. Both traditional techniques and deep learning for breast cancer diagnosis were briefly described in Chapter 2, which helps to justify the research's purpose and aims. After gaining comprehension, the framework for this study is devised and built.

By completing this phase, the first research objective (RO1), which is to study algorithm performance in detecting and classifying breast cancer, is met.

**Phase 2: Design, Develop and Implementation.**

In this research, we follow several key steps to build our model. First, we acquire and preprocess the dataset, using CBIS-DDSM, which contains 55,890 training examples. The images are preprocessed to 299x299 size, and a balanced test data issue is addressed. To improve accuracy, images are resized and cropped, ensuring a precise dataset without missing values. Feature selection is crucial for predictive model development, minimizing input variables and enhancing breast cancer prediction. These steps aim to achieve the research objective (RO2).

**Phase 3: The Result, Analysis and Discussion**

After completing step 2, which includes design, development, and implementation, the classification results from the utilized models are received. Based on this, model comparisons are made. The comparison is assessed based on its precision, sensitivity, and specificity. From this phase 3, the last purpose of this study, RO3, may be accomplished, which is to compare the accuracy of the classification using multiple algorithm models with previous study.

*B. Performance Measurement*

This study evaluated the efficacy of CNN, VGG19 and ResNet50 V2 in identifying breast cancer using performance measurement. The comparison is evaluated according to its accuracy, sensitivity, and specificity.

*C. Confusion Matrix*

The confusion matrix is composed of data on the actual and expected classifications made by the proposed models. Table II displays the confusion matrix based on two class classifiers. True Positive (TP) and True Negative (TN) refers to the number of accurately anticipated positive and negative instances (TN). False Positive (FP) refers to the number of positive instances that were incorrectly anticipated as negative, while False Negative (FN) refers to the number of negative cases that were incorrectly projected as positive (FN). The negative case represented "normal" labels, whereas the affirmative case represented "malignant/benign" labels.

CONFUSION MATRIX

| **Actual** | **Predicted** | |
|---|---|---|
| | ***Positive*** | ***Negative*** |
| Positive | TP | FN |
| Negative | FP | TN |

Consequently, performance assessment approaches in terms of precision, sensitivity, and specificity are determined utilizing evaluation metrics as stated in equations below.

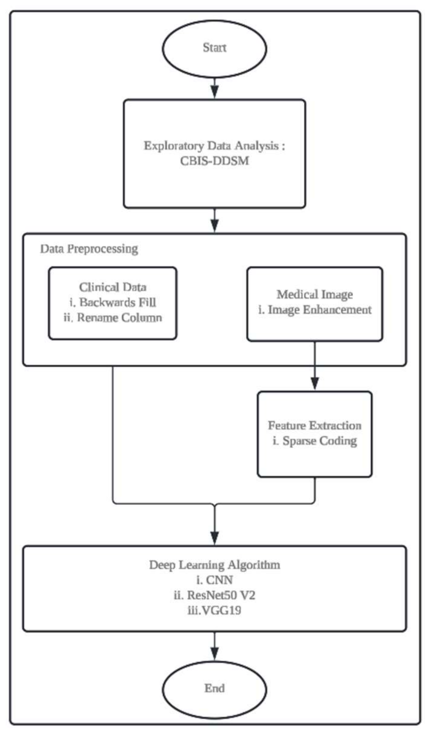
$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \text{ X } 100$$

*D. Data Preparation*

This research consists of two types of datasets collected from CBIS-DDSM, namely images and nominal data. It contains 6774 files of mammography images while the nominal data contains each description and attributes of the mammography images. This dataset comprises the outcomes of several medical examinations conducted on individuals to determine whether or not they have breast cancer. Patients' breast cancer prognosis is included since it is quantified as a decision characteristic, with benign cases labeled as 0 and malignant cases as 1.

*E. Proposed Solution*

This research, consisting of design, development, and experimentation, was initiated. This research began by gathering the dataset from CBIS-DDSM, which comprises breast cancer mammography images. Preprocessing the images is crucial to make sure that the images can be used for classification. In the classification step, the relevant characteristics are provided as inputs into CNN. Classification accuracy, sensitivity, and specificity are used to evaluate the effectiveness of both classifiers in diagnosing breast cancer. Based on test findings, the suggested approach would be able to identify breast cancer in patients.




Proposed Solution

### F. Data Pre-Processing

This research, consisting of design, development, and experimentation, was initiated. This research began by gathering the dataset from CBIS-DDSM, which comprises breast cancer mammography images. Preprocessing the images is crucial to make sure that the images can be used for classification. In the classification step, the relevant characteristics are provided as inputs into CNN. Classification accuracy, sensitivity, and specificity are used to evaluate the effectiveness of both classifiers in diagnosing breast cancer. Based on test findings, the suggested approach would be able to identify breast cancer in patient Data preprocessing involves transforming raw data into a clean dataset, ensuring it is suitable for statistical analysis. This process includes essential steps like data cleaning. Even accurate data can be inconsistently presented, making preprocessing crucial for statistical analysis quality. Key operations in preprocessing include filling data gaps, splitting datasets, and scaling features.

### G. Data Preparation

This research uses two types of datasets from CBIS-DDSM: images and nominal data. The dataset involves outcomes from medical examinations to determine breast cancer presence. Prognosis is quantified with benign cases as 0 and malignant cases as 1. The image dataset includes two types of mammogram images (Mediolateral Oblique and Bilateral Craniocaudal) and two structures (Classification and Masses). It contains 10,239 DICOM format images from 1,566 patients, gathered from DDSM and CBIS-DDSM databases. The dataset is an updated, standardized version with two classes: benign and malignant. It's divided into five tfrecords files, totaling 55,890 training examples, with 14% positive and 86% negative cases. All images are pre-processed to 299x299 pixels by extracting Regions of Interest (ROIs).

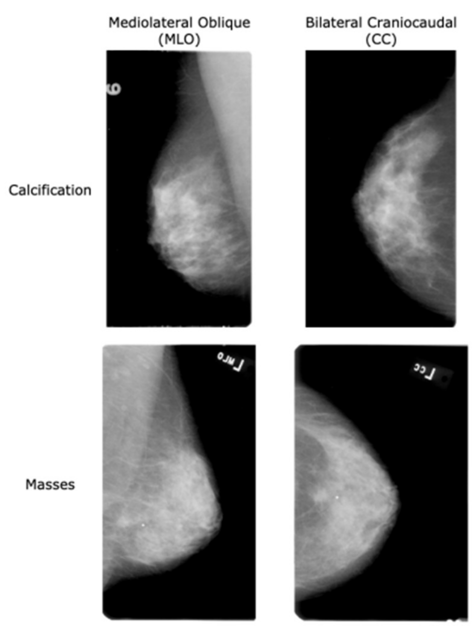



Image views

### H. Checking Missing Value

This research uses two types of datasets from CBIS-DDSM: images and nominal data. The dataset involves outcomes from medical examinations to determine breast cancer presence. Prognosis is quantified with benign cases as 0 and malignant cases as 1. The raw data extracted may be incomplete and have missing values, which can affect the accuracy of model predictions. To address this issue, a data processing check is necessary. Figure 4 indicates missing values in the dataset when using the is null (). sum () function. Several steps can be taken to handle missing values, such as replacing them with mean or mode, dropping the data, or using a backward fill method. In this research, the backward fill method is employed, filling missing values with the next available values in the sequence. This method assumes that upcoming values can provide insights into the missing ones, enhancing the understanding of the data by filling gaps with subsequent values.

MISSING VALUES YPE STYLES

| Attributes | Value |
|---|---|
| patient_id | 0 |
| breast_density | 0 |
| left_or_right_breast | 0 |
| image_view | 0 |
| abnormality_id | 0 |
| abnormality_type | 0 |
| calc_type | 1338 |
| cals_distribution | 1694 |
| assessment | 0 |
| pathology | 0 |
| sublety | 0 |
| image_file_path | 0 |
| cropped_image_file_path | 0 |
| ROI_mask_file_path | 0 |
| mass_shape | 1550 |
| mass_margins | 1589 |

### I. Rename Column

Renaming the column is crucial to achieve a clean dataset. It may lead to syntax error in the future. By removing it, it can lead to consistent and clean dataset that are more compatible with code, thus avoiding potential errors.

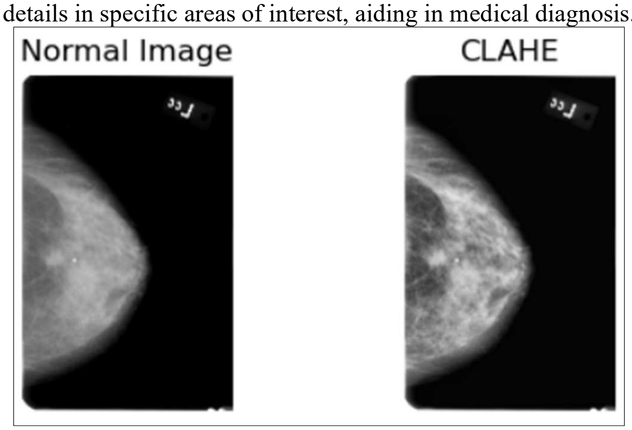
COLUMN NAME

| Column Names | |
|---|---|
| ***Before*** | ***After*** |
| left or right breast | left_or_right_breast |
| ***image view*** | ***image_view*** |

| Column Names | |
|---|---|
| ***Before*** | ***After*** |
| abnormality id | abnormality_id |
| abnormality type | abnormality_type |
| image file path | image_file_path |
| cropped image file path | cropped_image_file_path |
| ROI mask file path | ROI_mask_file_path |
| mass shape | mass_shape |
| mass margins | mass_margins |

### J. *Image Enhancement (CLAHE)*

CLAHE, or Contrast Limited Adaptive Histogram Equalization, is a significant image processing technique in computer vision. It is designed to improve image contrast, particularly in situations with uneven lighting and varying contrast levels. Commonly used in medical imaging like X-rays and CT scans, CLAHE is essential for highlighting important details in specific areas of interest, aiding in medical diagnosis.




CLAHE Enhancement

## IV. RESULT, ANALYSIS

### A. *Baseline Performance for Classification Model*

It is important to have a baseline performance while working on a classification as it can be a benchmark against other research or other algorithm models. The benchmark can be used to compare or determine whether the performance of this thesis had improved or vice versa. Table V shows the results of the previous paper that has been a benchmark for this thesis.

BASELINE PERFORMANCE

| Previous Research | Method | Results | Proposed Model | Results |
|---|---|---|---|---|
| Adam Jaamour et al. (2020) | ResNet 50 | 61% | ResNet 50 V2 | 67% |
| | VGG19 | 64% | VGG19 | 71% |

Adam Jaamour et al. (2020) conducted the research using the same dataset. The research conducted were with multiple CNN architectures which is VGG19 and RestNet50. As shown on the table above, Adam Jaamour et al. (2020) achieved an accuracy of 61% using the ResNet50 algorithm and 64% using VGG19.

### B. *Comparison between Epochs & Hyperparameter Tuning*

#### 1) *Epoch*

An epoch in deep learning signifies one complete pass through the entire training dataset during model training. In each epoch, the model encounters all training samples, and its parameters are adjusted based on the calculated error compared to the true labels. Using multiple epochs is crucial as it allows the model to iteratively refine its parameters by learning from the entire training dataset. This iterative process gradually improves the model's ability to make accurate predictions on new, unseen data. Table VI presents the results of running different numbers of epochs for each model.

EPOCH COMPARISON

| Method | Epoch | | |
|---|---|---|---|
| | 30 | 50 | 70 |
| CNN | 0.59% | 0.64% | 0.59% |
| RESNET50V2 | 0.67% | 0.68% | 0.70% |
| VGG19 | 0.71% | 0.70% | 0.72% |

The data shows that model performance differs with the number of epochs. The CNN model initially improves accuracy from 30 to 50 epochs but sharply declines at 70 epochs. ResNet50V2 consistently improves accuracy with more epochs. On the other hand, VGG19 reaches its highest accuracy at 30 epochs, slightly decreases at 50 epochs, and more significantly drops at 70 epochs.

### C. *Hyperparameter Tuning*

Hyperparameter tuning is the process of finding the best set of predefined parameters for a deep learning model before training. It aims to optimize the model's performance, such as improving accuracy or reducing errors on new data. In this research, HyperOpt, using Bayesian Optimization, is employed for efficient hyperparameter tuning. Table VII displays the different parameter sets tested, including epoch, batch size, and learning rate, which significantly impact the model's performance.

HYPERPARAMETER TUNING

| Parameter | Values |
|---|---|
| Epochs | [30, 50, 70] |
| Batch size | [16, 32, 64] |
| Learning rate | [ 0.0001, 0.01, 0.001] |

#### 1) *VGG19*

As for VGG19, the combination of 30 epoch, batch size of 32, and learning rate of 0.001 achieve the highest accuracy of 0.72%.

VGG RESULTS

| Hyperparameter | Accuracy | Precision | Recall |
|---|---|---|---|
| Epoch: 70<br>Batch Size: 32<br>Learning rate: 0.001 | 0.72% | 0.72% | 0.72% |
| Epoch : 70<br>Batch Size : 1632<br>Learning rate : 0.0001 | 0.69% | 0.69% | 0.68% |
| Epoch : 50<br>Batch Size : 64<br>Learning rate : 0.001 | 0.70% | 0.70% | 0.71% |
| Epoch : 30<br>Batch Size : 16<br>Learning rate : 0.01 | 0.69% | 0.67% | 0.68% |

For ResNet50V2, the best combination of the hyperparameter is epoch of 70, batch size of 32 and learning rate of 0.001 which achieve accuracy of 70%. At some cases, increasing the epoch number has proven to improve the results as it allows the algorithm model to learn and capture the complex patterns.

RESNET50 V2

| Hyperparameter | Accuracy | Precision | Recall |
|---|---|---|---|
| Epoch: 70<br>Batch Size: 32<br>Learning rate: 0.001 | 0.70% | 0.69% | 0.69% |
| Epoch : 30<br>Batch Size : 32<br>Learning rate : 0.0001 | 0.67% | 0.66% | 0.68% |
| Epoch : 50<br>Batch Size : 16<br>Learning rate : 0.001 | 0.68% | 0.68% | 0.68% |
| Epoch : 30<br>Batch Size : 64<br>Learning rate : 0.01 | 0.68% | 0.69% | 0.67% |

## V. DISCUSSION

This research achieved higher accuracy in deep learning algorithms compared to previous studies. The use of an improvised ResNet50 V2 contributed to this improvement, as its architecture enhances training efficiency and accuracy. Additionally, Contrast Limited Adaptive Histogram Equalization (CLAHE) played a crucial role in boosting accuracy by improving image contrast. This preprocessing technique, commonly used in medical images, enhances visibility and performance. Performance comparisons among the three deep learning algorithms are presented in Table X and Figure 6.

ACCURACY COMPARISON

| Method | Accuracy |
|---|---|
| CNN | 0.64% |
| RESTNET50V2 | 0.70% |
| VGG19 | 0.72% |

*A.*

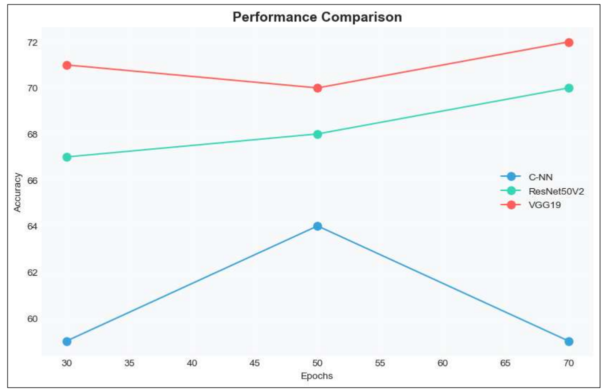



Figure 6: Performance Comparison

Table above and Figure 6 reveal performance trends for three algorithms. CNN shows improvement, reaching 64% accuracy at 50 epochs but drops afterward. ResNet50 V2 consistently improves, reaching 70% accuracy. VGG19 fluctuates, dropping to 70% and then rising to 72% at epoch 70. Overall, VGG19 has the highest performance, but ResNet50 V2 shows consistent improvement throughout the experiment.

## VI. CONCLUSION

The main focus of this research lies in the dataset used for breast cancer data. The objective of this thesis is to assess the accuracy of prediction by employing various algorithm models. The purpose is to enhance the classification of breast cancer and facilitate early detection for prompt treatment. As the preprocessing process is conducted, we hope that it would enhance and increase the accuracy of the model as it would be able to detect the cancer cell thoroughly from the mammography images. As this research has been completed through the pre-processing phase, more discussion and work on the implementation of algorithm models to predict breast cancer in the classifying phase will be conducted in the next phase.

## ACKNOWLEDGMENT

The Universiti Teknologi Malaysia (UTM) supported this work through UTM Fundamental Grant (PY/2023/01586) with Cost Centre No. Q.J130000.3828.23H44.

## REFERENCES

[1] Shen R., Yan K., Tian K., Jiang C., Zhou K. Breast mass detection from the digitized X-ray mammograms based on the combination of deep active learning and self-paced learning. Future Gener. Comput. Syst. 2019;101:668–679. doi: 10.1016/j.future.2019.07.013.

[2] M.P. Beham, R. Tamilselvi, S.M. Roomi, A. Nagaraj, Accurate classification of cancer in mammogram images, in: Innovations in electronics and communication engineering, Springer, Singapore, 2019, pp. 71–77.

[3] Wang, D. Zhang and Y. H. Huang “Breast Cancer Prediction Using Machine Learning” (2018), Vol. 66, NO. 7.

[4] B. Dai, R. -C. Chen, S. -Z. Zhu and W. -W. Zhang, "Using Random Forest Algorithm for Breast Cancer Diagnosis," 2018 International Symposium on Computer, Consumer and Control (IS3C), 2018, pp. 449-452, doi: 10.1109/IS3C.2018.00119.

[5] Yan Wang, Yangqin Feng, Lei Zhang, Zizhou Wang, Qing Lv, Zhang Yi, Deep adversarial domain adaptation for breast cancer screening from mammograms, Medical Image Analysis, Volume 73, 2021, 102147, ISSN 1361-8415, https://doi.org/10.1016/j.media.2021.102147

[6] Heenaye-Mamode Khan, M., Boodoo-Jahangeer, N., Dullull, W., Nathire, S., Gao, X., Sinha, G. R., & Nagwanshi, K. K. (2021). Multi- class classification of breast cancer abnormalities using Deep Convolutional

Neural Network (CNN). PloS one, 16(8), e0256500. https://doi.org/10.1371/journal.pone.0256500

[7] Lee, R., Gimenez, F., Hoogi, A. et al. A curated mammography data set for use in computer-aided detection and diagnosis research. Sci Data 4, 170177 (2017). https://doi.org/10.1038/sdata.2017.177

[8] G. Meenalochini, S. Ramkumar, Survey of machine learning algorithms for breast cancer detection using mammogram images, Materials Today: Proceedings, Volume 37, Part 2, 2021, Pages 2738-2743, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2020.08.543.

[9] M. Thalhammer, C. Karlo, P. Rehak, A. Pilhatsch, H. Mischinger, P. Kohek, A. Beham, 215 POSTER Hyperthermia as a therapeutic option in recurrent breast sarcoma, European Journal of Surgical Oncology (EJSO), Volume 32, Supplement 1, 2006, Page S64, ISSN 0748-7983, https://doi.org/10.1016/S0748-7983(06)70650-3.

[10] Anders CK and Carey LA. ER/PR negative, HER2-negative (triple-negative) breast cancer. UpToDate website. https://www.uptodate.com/contents/er-pr-negative-her2-negative-triple-negative-breast-cancer. Updated June 06, 2019. Accessed July 23, 2019.

[11] Calhoun KE, Allison KH, Kim JN et al. Chapter 62: Phyllodes Tumors. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed. Philadelphia, Pa: Lippincott-Williams & Wilkins; 2014.

[12] Dillon DA, Guidi AJ, Schnitt SJ. Ch. 25: Pathology of invasive breast cancer. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed. Philadelphia, Pa: Lippincott-Williams & Wilkins; 2014.

[13] Esteva FJ and Gutiérrez C. Chapter 64: Nonepithelial Malignancies of the Breast. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed. Philadelphia, Pa: Lippincott-Williams & Wilkins; 2014.

[14] Henry NL, Shah PD, Haider I, Freer PE, Jagsi R, Sabel MS. Chapter 88: Cancer of the Breast. In: Niederhuber JE, Armitage JO, Doroshow JH, Kastan MB, Tepper JE, eds. Abeloff's Clinical Oncology. 6th ed. Philadelphia, Pa: Elsevier; 2020.

[15] Jagsi R, King TA, Lehman C, Morrow M, Harris JR, Burstein HJ. Chapter 79: Malignant Tumors of the Breast. In: DeVita VT, Lawrence TS, Lawrence TS, Rosenberg SA, eds. DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology. 11th ed. Philadelphia, Pa: Lippincott Williams & Wilkins; 2019.

[16] National Cancer Institute. Inflammatory Breast Cancer. 2016. Accessed at https://www.cancer.gov/types/breast/ibc-fact-sheet on August 30, 2021.

[17] National Comprehensive Cancer Network (NCCN). Practice Guidelines in Oncology: Breast Cancer. Version 7.2021. Accessed at https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf on August 30, 2021.

[19] Nora M. Hansen. Chapter 63: Paget's Disease. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed. Philadelphia, Pa: Lippincott-Williams & Wilkins; 2014.