

CLASSIFICATION MODEL FOR BREAST CANCER MAMMOGRAMS

SUZANI MOHAMAD SAMURI^{1,3*}, TRY VIANANDA NOVA², BAHBIBI
RAHMATULLAH^{1,3}, WANG SHIR LI^{1,3}, ZIADOON TAREQ AL-QAYSI⁴

¹*Department of Computing, Faculty of Art, Computing and Creative Industry,
Sultan Idris Education University, Malaysia*

²*Department of Informatic, Universitas Ibn Kaldun Bogor, Jawa Barat, Indonesia*

³*Data Intelligence and Knowledge Management Special Interest Group, Faculty of Art,
Computing and Creative Industry, Sultan Idris Education University, Malaysia*

⁴*Computer Science Department, College of Computer Sciences and Mathematics,
Tikrit University, Iraq*

**Corresponding author: suzani@fskik.upsi.edu.my*

(Received: 16th January 2021; Accepted: 13th October 2021; Published on-line: 4th January 2022)

ABSTRACT: Machine learning has been the topic of interest in research related to early detection of breast cancer based on mammogram images. In this study, we compare the performance results from three (3) types of machine learning techniques: 1) Naïve Bayes (NB), 2) Neural Network (NN) and 3) Support Vector Machine (SVM) with 2000 digital mammogram images to choose the best technique that could model the relationship between the features extracted and the state of the breast ('Normal' or 'Cancer'). Grey Level Co-occurrence Matrix (GLCM) which represents the two dimensions of the level variation gray in the image is used in the feature extraction process. Six (6) attributes consist of contrast, variance, standard deviation, kurtosis, mean and smoothness were computed as feature extracted and used as the inputs for the classification process. The data has been randomized and the experiment has been repeated for ten (10) times to check for the consistencies of the performance of all techniques. 70% of the data were used as the training data and another 30% used as testing data. The result after ten (10) experiments show that, Support Vector Machine (SVM) gives the most consistent results in correctly classifying the state of the breast as 'Normal' or 'Cancer', with the accuracy of 99.4%, in training and 98.76% in testing. The SVM classification model has outperformed NN and NB model in the study, and it shows that SVM is a good choice for determining the state of the breast at the early stage.

ABSTRAK: Pembelajaran mesin telah menjadi topik yang diminati dalam penyelidikan yang berkaitan dengan pengesanan awal kanser payudara berdasarkan imej mamogram. Dalam kajian ini, kami membandingkan hasil prestasi dari tiga (3) jenis teknik pembelajaran mesin: 1) Naïve Bayes (NB), 2) Neural Network (NN) dan 3) Support Vector Machine (SVM) dengan 2000 imej digital mammogram hingga teknik terbaik yang dapat memodelkan hubungan antara ciri yang diekstraksi dan keadaan payudara ('Normal' atau 'Cancer') dapat diperoleh. Grey Level Co-occurrence Matrix (GLCM) yang mewakili dua dimensi variasi tahap kelabu pada gambar digunakan dalam proses pengekstrakan ciri. Enam (6) atribut terdiri dari kontras, varians, sisihan piawai, kurtosis, min dan kehalusan dihitung sebagai fitur yang diekstrak dan digunakan sebagai input untuk proses klasifikasi. Eksperimen telah diulang selama sepuluh (10) kali untuk memeriksa kesesuaian prestasi semua teknik. 70% data digunakan sebagai data latihan dan 30% lagi digunakan sebagai data ujian. Hasil setelah sepuluh (10) eksperimen menunjukkan bahawa, Support Vector Machine (SVM) memberikan hasil yang paling konsisten dalam mengklasifikasikan keadaan payudara dengan betul sebagai 'Normal' atau 'Kanser', dengan akurasi 99.4%,

dalam latihan dan 98.76% dalam ujian. Model klasifikasi SVM telah mengungguli model NN dan NB dalam kajian ini, dan ia menunjukkan bahawa SVM adalah pilihan yang baik untuk menentukan keadaan payudara pada peringkat awal.

KEYWORDS: *machine learning; breast cancer detection; mammogram images; data mining; data-driven modelling*

1. INTRODUCTION

Breast cancer screening which is commonly made by mammography is a limited series of x-rays taken of the breast. Radiologist sees for any anomalous signs or patterns on the mammogram that might be breast cancer. These signs normally show up on the mammogram before any lump may be thought at the breast. If there is anything odd on this mammogram, more tests get to be made. These tests may add another mammogram, imaging, or a biopsy. Whether testing mammograms will change the probability of women dying from breast cancer has been argued for decades [1]. It is definitely not an ideal test. Some cancers are failed to be detected by the mammogram. This false-positive mammogram effects lead to more experimentation, which is time-consuming and may have unnecessary apprehension. Since mammographic X-ray images are 2D projections of a 3D object, localization of features identified within the breast volume is not trivial. Furthermore, mammograms represent highly deformed configurations of the breast due to compression, thus the tumor localization process relies on the expertise of the clinicians. Breast cancer-related studies have long been conducted by researchers around the world.

In the last few decades, various information production and machine learning techniques have been produced for breast cancer discovery and categorization, which may be separated into three important phases: Preprocessing, feature extraction, and classification. To facilitate explanation and analysis, the preprocessing of mammography images helps improve the clarity of peripheral regions and level system, and various methods have been reported to help in the process. Several transform-based quality analysis techniques are used to change the image into a new structure using the spatial frequency properties of the component level variations. In fact, there are many studies related to early detection is possible to show whether the state of the breast to be 'Normal' or 'Abnormal'. Among the most common classifier used for classification and early detection of breast cancer is Neural Network (NN). It is a model made to imitate the function of learning that owned the human brain. As stated by [2], the Neural classifier developed able to perform well by successfully classifying cancer and non-cancer (normal) images with an accuracy of 97%. However, this study did not show any comparison between the developed NN classifier and other types of the classifier. With regard to a single technique, [3] mentioned that in his mapping study, found that Decision Trees (DT), Support Vector Machines (SVM) and Artificial Neural Networks (ANN) were those most frequently adopted to build ensemble classifiers.

Therefore, in this paper on cancer detection, the focus will be on the classification part, whereby the use of three classifiers namely Naive Bayes (NB), Neural Network (NN) and Support Vector Machines (SVM), to accurately classify the mammographic images of the breast as Normal or Cancer will be presented. The results of this research will be useful for tracking visible tumors between images, as well as to help radiologists in their tasks to detect subtle abnormalities in a mammogram.

2. METHODOLOGY

This research methodology uses a data mining approach to classify data. In this case, there is an important phase that must be done to get the best research results. There are four phases involved namely sample selection, image enhancement and feature extraction, model development, and evaluation.

2.1 Sample Selection

A total of two thousand (2000) 2d-digital mammogram images were collected from the Central Pertamina Hospital, Indonesia (1000 normal breast images and 1000 cancer images) as shown in Fig. 1. These 2d-digital mammogram images were then converted to 3d-digital mammogram images for better analysis.

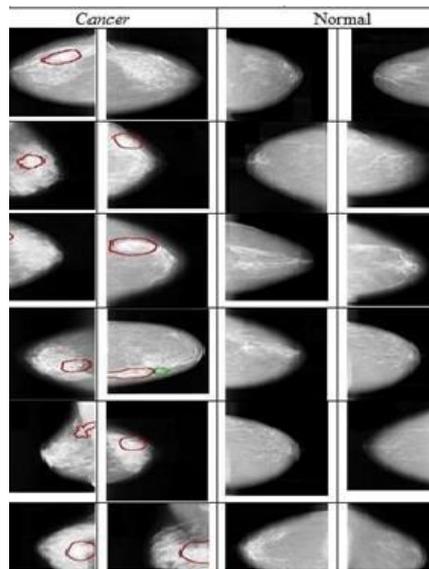


Fig. 1: Some of the mammogram of normal and cancer images.

2.2 Image Enhancement and Feature Extraction

In preprocessing phase, the possibility distribution algorithm [4] is used for image enhancement using logical approach fuzzy with 5 parameters, namely α , β_1 , γ , β_2 and \max as shown in Fig. 2.

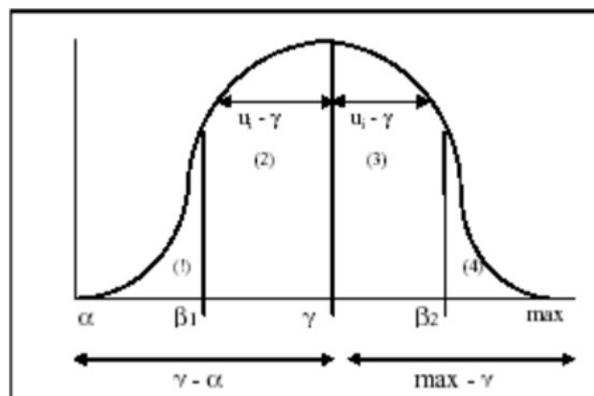


Fig. 2: Possibility distribution function [4].

From the required parameters, α represents the minimum value of the distribution, γ represents the average value of the distribution and \max represents the maximum value of the distribution. The fuzzy transformation function used to get the entire value is defined as follows:

$$\alpha = \min \tag{1}$$

$$\beta_1 = (\alpha + \gamma)/2 \tag{2}$$

$$\gamma = \text{mean} \tag{3}$$

$$\beta_2 = (\max + \gamma)/2 \tag{4}$$

$$\max = \max \tag{5}$$

The purpose of using the possibility distribution algorithm in improving image quality is to lower the level of pixel grayness that has a grayish value between β_1 and β_2 . The way to do this is to provide a new intensity value in pixels between β_1 and γ , γ and β_2 with the opposite board value to the mean γ . The fuzzy rule below is used to perform image contrast enhancement based on Fig. 2:

$$\text{Rule 1: If } \alpha \leq u_i \leq \beta_1 \text{ then } P = 2((u_i - \alpha)/(\gamma - \alpha))^2 \tag{6}$$

$$\text{Rule 2: If } \beta_1 \leq u_i \leq \gamma \text{ then } P = 1 - 2((u_i - \gamma)/(\gamma - \alpha))^2 \tag{7}$$

$$\begin{aligned} \text{Rule 3: If } \gamma \leq u_i \leq \beta_2 \text{ then } P \\ = 1 - 2((u_i - \gamma)/(\max - \gamma))^2 \end{aligned} \tag{8}$$

$$\text{Rule 4: If } \beta_2 \leq u_i \leq \max \text{ then } P = 2((u_i - \gamma)/(\max - \gamma))^2 \tag{9}$$

where $u_i = f(x,y)$ is the intensity of the i - i pixel. Regulations that lower the level of pixel grayness that has a grayish value between β_1 and β_2 are represented by Eqs. (7) and (8). The steps of possibility distribution algorithm can be explained as the following:

Step 1: Initialize parameter data matrix set by image to be improved. From the image, the minimum, maximum and average (mean) grayish values were obtained. The value β_1 is set with the result of $(\min + \text{mean})/2$. Whereas the value of β_2 is set with the result of $(\text{mean} + \max)/2$.

Step 2: Fuzzification for all pixels. In this step, fuzzy rules were applied (Eq. 6, Eq. 7, Eq. 8 and Eq. 9) to get a new grayish value.

Step-3: Modification for all pixels were executed, the fuzzy value of the data obtained from the new grayish value is powered by two.

Step-4: Defuzzification for all pixels. The new contrast data derived from fuzzy data values is calculated and multiplied by the matrix of the initial image data.

In this phase, feature extraction also being done to extract the important variables to be used in the next step which is development of classification model. For feature extraction, GLCM (Grey Level Co-occurrence Matrix) is being used. GLCM is the two dimensions represent the level variation gray at the image. GLCM is one of the popular statistical methods of extracting textural feature from images. According to [5], by extracting the features of an image by GLCM approach, the image compression time can be greatly

reduced in the process of converting RGB to Gray level image when compared to other discrete wavelet transform (DWT) techniques. In this study, measurement of the contrast, variance, standard deviation, kurtosis, mean, and smoothness are computed as features extracted and shall be used in the classification stage.

2.3 Development of Data-driven Classification Model

In this work, Matlab software is used for developing the classification model and analysis. The three (3) machine learning techniques investigated to produce a data-driven classification models, were Neural Network (NN), Support Vector Machine (SVM) and Naïve Bayes (NB). Figure 3 shows an example of the ‘Breast Image Extracted Features – Breast State’ modelling structure. The inputs identified for the classification model were contrast, variance, standard deviation, kurtosis, mean, and smoothness. These features were extracted using GLCM methods explained in the previous section [5]. Whereas the output for this classification model is the state of the breast itself, whether it is ‘Normal’ or ‘Cancer’.

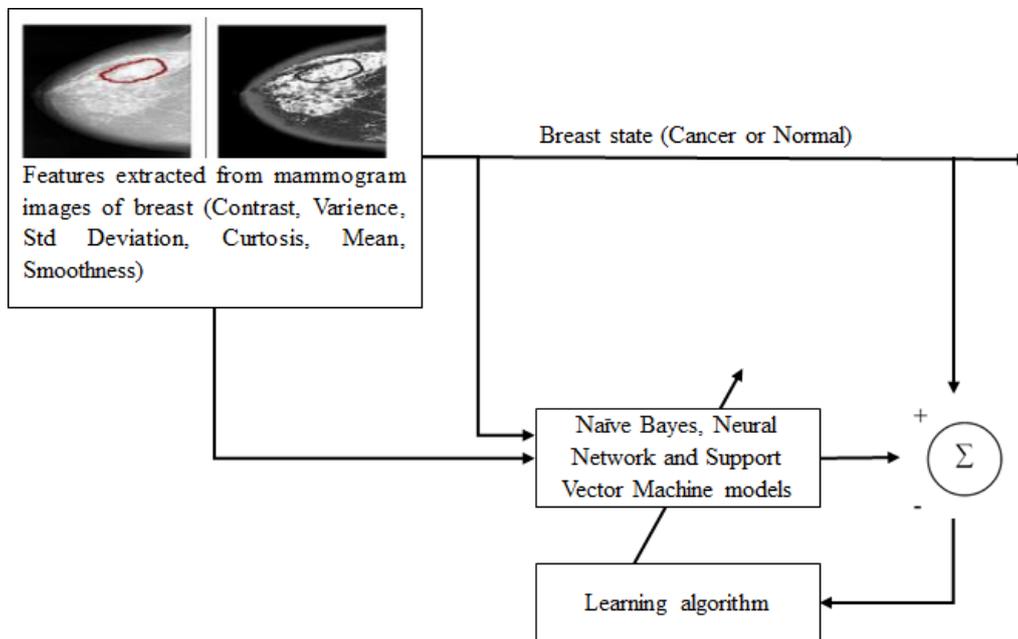


Fig. 3: Model structure for ‘Breast Image Extracted Features – Breast State’. $\varepsilon(t)$ = error between actual and predicted breast category.

2.3.1 Naïve Bayes (NB)

The Naive Bayesian classifier is based on Bayes Theorem assuming independence among predictors. The Naive Bayesian model is easy to build, without the estimation of complex iterative parameters that make it very useful for very large datasets. Despite its simplicity, the Naive Bayesian classification often works very well and is widely used because it often outperforms more sophisticated classification methods.

Bayes Theorem provides a way of calculating the probability of posterior, $P(c | x)$, of $P(c)$, $P(x)$, and $P(x | c)$. The basis of the Naïve Bayes theorem used is the Bayes formula as follows [6]:

$$P(C|X) = \frac{P(x|c)P(c)}{P(x)} \quad (10)$$

The Bayes Naïve classification assumes that the influence of predictor value (x) on a particular class (c) does not depend on the values of other predictors. This assumption is called class conditional independence.

2.3.2 Neural Network (NN)

Neural network is a model created to mimic the learning functions of the human brain. In the Neural Network, neurons are grouped into layers, called neuron layers. Usually, each neuron of a layer is connected to all neurons in the back and front layers (except inputs and outputs). Information sent in a NN, propagated layers – per – layers ranging from inputs to outputs without or through one or more hidden layers. Depending on the algorithm used, information can also be propagated backwards (backpropagation). Figure 4 shows the Neural Network with three neuron layers.

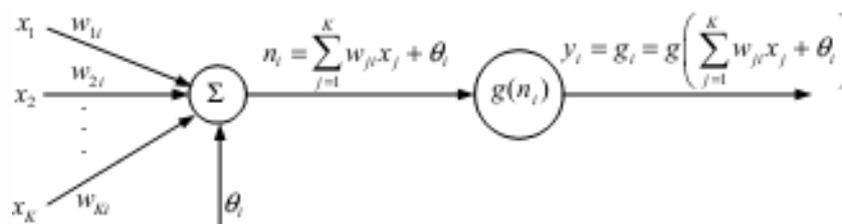


Fig. 4: Algorithm for Neural Network.

2.3.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the prediction techniques, both in the case of classification and regression [7]. SVM has a basic principle of linear classifier that is a classification case that can be linearly separated, but nowadays SVM is also able to solve non-linear problems by adding kernel concepts to high-resolution workspaces. In this space, a hyperplane (hyperplane) will be searched to maximize the distance between data classes.

SVM can actually work on linearly separated data only. When the data is nonlinear then, the way that can be used is to use the kernel on the data. Kernel can be interpreted as a function that remaps data in the original dimension space into more dimensional space as shown in Fig. 5.

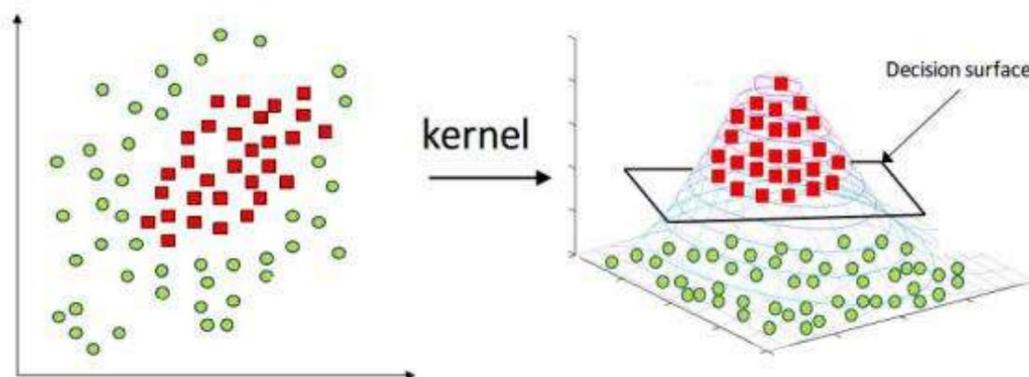


Fig. 5: Representation for Support Vector Machine (SVM).

2.4 Experimental Procedure

As shown in Fig. 3, the features extracted from 2000 mammogram breast images were experimented with different classifiers that are Neural Network (NN), Support Vector

Machine (SVM) and Naïve Bayes (NB). The data has been randomized and the experiment has been repeated for ten (10) times to check for the consistencies of the performance of all techniques. 70% of the data were used as the training data and another 30% used as testing data. Accuracy based on confusion matrix is used as the performance indicator, to show which technique can best model the classification of normal and abnormal state of the breast.

3. FINDINGS

In this section, findings from three (3) different classification models (NB, NN and SVM) will be presented. The inputs identified for all models were the contrast, variance, standard deviation, kurtosis, mean, and smoothness based on the GLCM approach applied to 2000 3d-digital breast mammogram images reported in the previous section. The outputs for all models were the breast state: 'Normal' or 'Cancer', which represented by '1' and '2' in the modelling process respectively. The extracted features with the breast states were then randomized. 70% of the data were used for training and another 30% were used as testing. The objective of this modelling is to mimic the behavior of the expert in identifying the state of the breast whether it is 'Normal' or 'Cancer' by training the model with the extracted features (contrast, variance, standard deviation, kurtosis, mean, and smoothness) from the 3d-digital mammogram images and the identified breast state. The following subsections report the findings on all the training and testing work conducted.

3.1 Naïve Bayes Model Training and Testing Results

The first classification model built in this study is using the NB. The scatterplots for training and testing of NB model are shown in Fig. 6(a) and (b). While the detail results for the accuracy of the model for each experiment in the training and testing are shown in Table 1 and Table 2. In Table 1, it can be seen, that out of 1400 sample data used for training (70% of the total sample data), there are average of 964 correct predictions and average of 436 incorrect predictions for training. Thus, the average accuracy for the training is about 68.83%. Whereas for the testing, out of 600 sample data used for testing (30% of the total sample data), there are average of 420 correct predictions and average of 180 incorrect predictions for the whole 10 experiments. Therefore, the average accuracy for the NB model testing is about 70.07%.

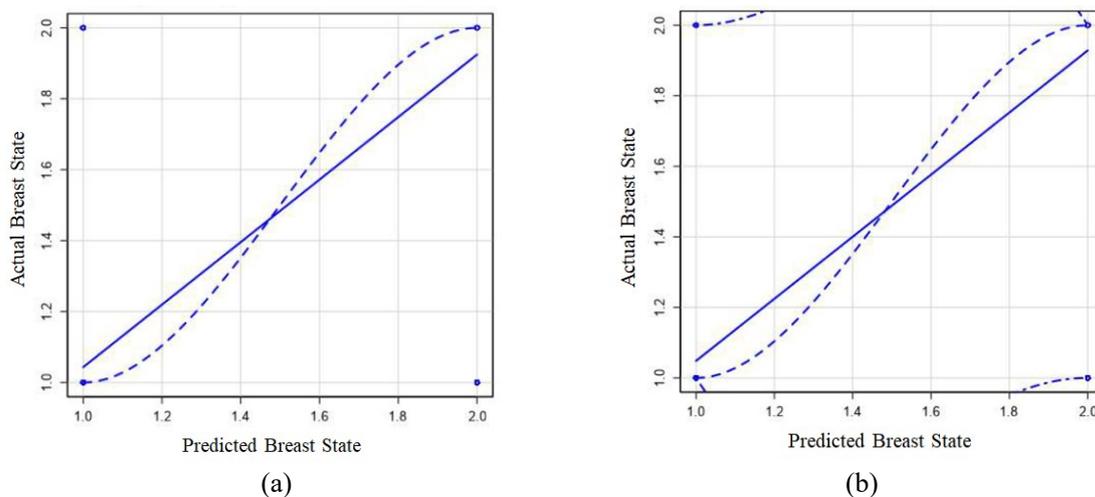


Fig. 6: Scatterplots for (a) Training of the NB Model, (b) Testing of the NB Model.

Table 1: NB model training results for 10 experiments.

Naïve Bayes Model Training Results																					
Actual		Predicted																			
		Exp 1		Exp 2		Exp 3		Exp 4		Exp 5		Exp 6		Exp 7		Exp 8		Exp 9		Exp 10	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1		540	194	540	194	580	165	579	166	520	209	576	168	545	191	536	197	580	166	574	170
2		272	394	271	395	231	424	233	422	291	380	235	421	267	397	276	391	231	423	237	419
Acc		66.69%		66.77%		71.70%		71.51%		64.28%		71.23%		67.31%		66.22%		71.62%		70.95%	

Table 2: NB model testing results for 10 experiments.

Naïve Bayes Model Testing Results																					
Actual		Predicted																			
		Exp 1		Exp 2		Exp 3		Exp 4		Exp 5		Exp 6		Exp 7		Exp 8		Exp 9		Exp 10	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1		228	86	228	86	252	68	253	67	248	71	256	66	223	90	232	83	253	68	258	64
2		120	166	120	166	96	184	95	185	100	181	91	187	125	162	116	169	95	184	90	188
Acc		65.68%		65.75%		72.74%		73.06%		71.44%		73.77%		64.21%		66.82%		72.87%		74.41%	

3.2 Neural Network (NN) Model Training and Testing Results

The next classification model built in this study is using the NN. Figure 7(a) and (b) show the scatterplots for training and testing of NN model, while Table 3 and 4 show the detail results of the NN model accuracy for each experiment in training and testing. As seen in Table 3 and 4, there are average of 1237 correct predictions and average of 163 incorrect predictions for training. Thus, the average accuracy for the training is about 88.38%. Whereas for the testing, out of 600 sample data used for testing (30% of the total sample data), there are average of 530 correct predictions and average of 70 incorrect predictions for the whole 10 experiments. Therefore, the average accuracy for the NN model testing is about 88.27%.

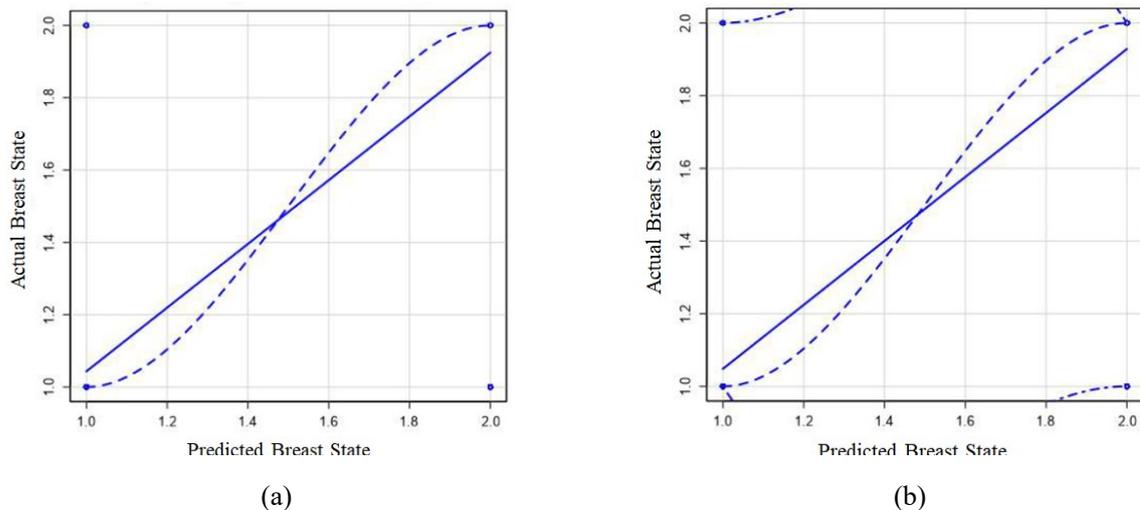


Fig. 7: Scatterplots for (a) Training of the NN Model. (b) Testing of the NN Model.

Table 3: NN model training results for 10 experiments.

Neural Network Model Training Results																					
Actual		Predicted																			
		Exp 1		Exp 2		Exp 3		Exp 4		Exp 5		Exp 6		Exp 7		Exp 8		Exp 9		Exp 10	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1		708	73	717	66	716	67	729	58	711	71	697	81	721	64	718	66	714	68	719	65
2		101	518	93	524	94	523	81	532	99	519	113	509	88	527	93	523	96	522	90	526
Acc		87.55%		88.66%		88.51%		90.04%		87.89%		86.17%		89.11%		88.68%		88.29%		88.90%	

Table 4: NN model testing results for 10 experiments.

Neural Network Model Testing Results																					
Actual		Predicted																			
		Exp 1		Exp 2		Exp 3		Exp 4		Exp 5		Exp 6		Exp 7		Exp 8		Exp 9		Exp 10	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1		314	24	304	31	302	32	299	34	319	20	308	28	308	28	301	33	306	29	300	33
2		33	229	43	222	46	220	48	219	29	232	39	225	38	226	47	219	41	224	47	220
Acc		90.48%		87.60%		87.01%		86.29%		91.91%		88.88%		88.95%		86.73%		88.27%		86.62%	

3.3 Support Vector Machine (SVM) Model Training and Testing Results

The final classification model developed in this study is using the SVM. The scatterplots result of SVM model for training and testing are shown in Fig. 8 (a) and (b), while Table 5 and 6 show the detail results for the accuracy of the model for each experiment during training and testing.

From Table 5, it was found that, the majority of the outputs were correctly predicted by the SVM model, with average of 1392 correct predictions and average of only 8 incorrect predictions for training. Thus, the average accuracy for the training is about 99.40%. Interestingly, Table 6 also shown that the SVM model is able to correctly predict the output of the testing sample data with average of 593 correct predictions as against the incorrect predictions of only 7. This has made the average percentage of the testing accuracy to be 98.76%, slightly lower than its training accuracy.

Table 5: SVM model training results for 10 experiments.

Support Vector Machine Model Training Results																					
Actual		Predicted																			
		Exp 1		Exp 2		Exp 3		Exp 4		Exp 5		Exp 6		Exp 7		Exp 8		Exp 9		Exp 10	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1		799	7	799	7	809	0	809	0	799	7	809	0	799	7	799	7	809	0	809	0
2		10	584	10	584	0	591	0	591	10	584	0	591	10	584	10	584	0	591	0	591
Acc		98.80%		98.80%		100.00%		100.00%		98.80%		100.00%		98.80%		98.80%		100.00%		100.00%	

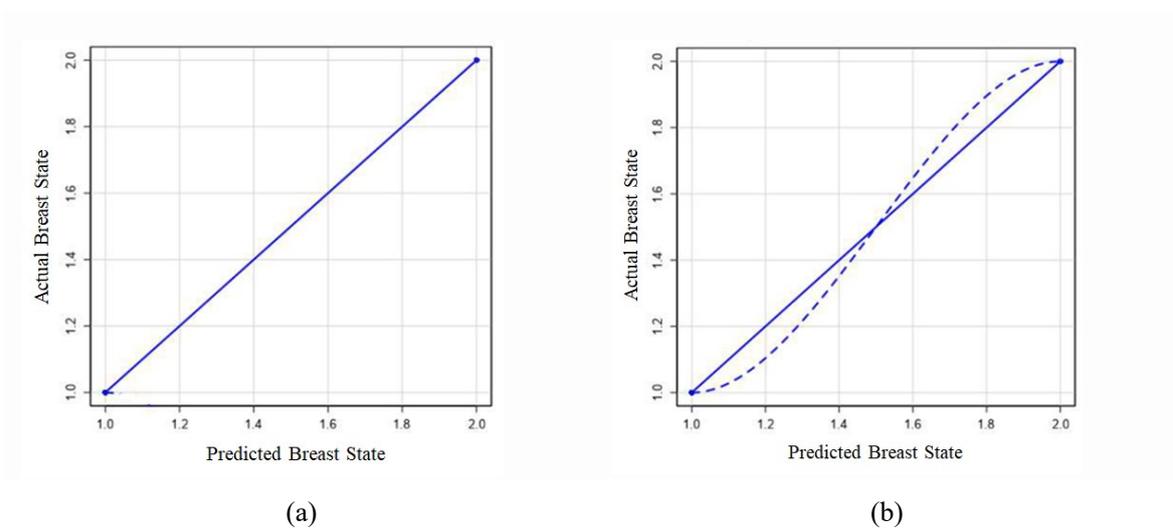


Fig. 8: Scatterplots for (a) Training of the SVM Model, (b) Testing of the SVM Model.

Table 6: SVM model testing results for 10 experiments.

Support Vector Machine Model Testing Results																					
Actual		Predicted																			
		Exp 1		Exp 2		Exp 3		Exp 4		Exp 5		Exp 6		Exp 7		Exp 8		Exp 9		Exp 10	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2		
1		336	8	340	5	347	0	347	0	342	3	342	3	339	6	342	3	342	3	347	0
2		10	246	7	248	0	253	0	253	5	250	5	250	7	248	5	250	5	250	0	253
Acc		97.00%		98.00%		100.00%		100.00%		98.70%		98.70%		98.70%		98.70%		98.70%		100.00%	

4. DISCUSSION

In this study, the 2d-digital mammography images were first transformed into 3d-digital mammography and GLCM approach is applied to these 3d-digital breast images to extract the contrast, variance, standard deviation, kurtosis, mean, and smoothness as features to be input to the models. The supervised learning technique is used to train the models that infer a function from input data with labeled breast state, 'Normal' or 'Cancer'. 2000 digital mammogram images of the breast have been analyzed.

Naive Bayes, Neural Network, and Deep Learning are among the most common machine learning techniques used in cancer detection [2,8-11]. However, in this study, ten (10) experiments have been conducted using the three (3) machine learning techniques namely Naïve Bayes, Neural Network and Support Vector Machine and the accuracy of each model in providing correct prediction for each experiment was recorded as shown in Section 3.

The result after ten (10) experiments show that the Support Vector Machine (SVM) gives the most consistent results in predicting or classifying correctly the state of the breast as 'Normal' or 'Cancer', with an average accuracy of 99.40% in training and 98.76% in testing. The SVM classification model has outperformed NN and NB models in the study, and it shows that SVM is a good choice for determining the state of the breast at the early stage.

As compared to the previously developed NN model reported by [2], the currently developed SVM model outperformed their NN model. Their study also not considering any other machine learning techniques but just focusing on NN model only.

However, these single classifier models can be further experiment and improved. For example, as reported by [3], the majority of research have reported positive results concerning the accuracy of ensemble classifiers as compared to the single classifiers. The following research have supported the given statement [12-21].

Apart from that, the performance index used can also be varied, not just focusing on the accuracy of the model but to consider other aspect like computational time for the training and testing to be done, as well as Precision-Recall curve to capture the exact capability of the classifiers investigated [22-25].

Transfer learning has also seen as the alternative to the conventional supervised learning for breast cancer detection and classification as reported by [3]. Therefore, comparison of model performance developed based on the supervised learning with transfer learning would be interesting to be explored in the near future.

Besides testing procedure that has been implemented in this study, validation with expert (radiologist) is seen crucial. The expert validation is time consuming, but it is important to be done in order to further validate the reliability of the developed models.

5. CONCLUSION

Today, the need for machine learning is rising in the field of breast cancer research. It has been used to screen and differentiate normal breast with cancer. The process of diagnosis can be very complicated. That is why for decades, researchers have developed various machine learning algorithms and architectures that can screen and differentiate cancerous images from normal ones by training the algorithms with various features extracted from the image of the breast. In particular, breast cancer diagnosis is more important than ever because the classification results will affect the treatment and safety of the patients. It requires not only a high prediction of accuracy, but also a high reliability and robustness.

In this study, ten (10) experiments have been conducted using the three (3) machine learning techniques namely Naïve Bayes, Neural Network and Support Vector Machine, using 2000 2d-digital mammogram images that have been converted into 3d-images. The accuracy of each model in providing correct prediction based on the confusion matrix has been reported for each experiment. The SVM classification model has outperformed NN and NB models in the study, and it shows that SVM is a good choice for determining the state of the breast at the early stage. However, further improvements can be done to the classifier models in producing more accurate and reliable results, for example, by exploring the ensemble classifier by combining different algorithm from the same type of classifier or different type of classifier to leverage the strength of individual classifiers.

As part of the future work, these models can also be extended to further classify the stage of the breast cancer (Stage 0, 1, 2, 3 and 4) by using the laboratory results of the patients as well as the auto measured diameter of the detected abnormality extracted from the digital mammogram images.

ACKNOWLEDGEMENT

This research study is funded by Universiti Pendidikan Sultan Idris, Malaysia, under the Special University Research Grant (Fundamental), 2017-0184-109-01.

REFERENCES

- [1] Rubin R. (2017) Do Screening Mammograms Cut Breast Cancer Deaths or Lead to Overtreatment? Probably Both. *Forbes*. Retrieved from <https://www.forbes.com/sites/ritarubin/2017/01/10/do-screening-mammograms-cut-breast-cancer-deaths-or-lead-to-overtreatment-probably-both/>
- [2] Mohamad Samuri, Suzani, Megariani TV. (2019) Intelligent 3D Analysis for Detection and Classification of Breast Cancer. *JITCE (Journal of Information Technology and Computer Engineering)*, 3(2): 96-103. <https://doi.org/https://doi.org/10.25077/jitce.3.02.96-103.2019>
- [3] Hosni M, Abnane I, Idri A, de Gea JMC, Alemán JLF. (2019) Reviewing ensemble classification methods in breast cancer. *Computer methods and programs in biomedicine*, 177: 89-112.
- [4] Dubois D, Prade, H. (2016) Practical Methods for Constructing Possibility Distributions. *Int. J. Intell. Syst.*, 31: 215-239. <https://doi.org/10.1002/int.21782>
- [5] Mohanaiah P, Sathyanarayana P, GuruKumar L. (2013) Image texture feature extraction using GLCM approach. *International journal of scientific and research publications*, 3(5): 1.
- [6] Han J, Kamber M, Pei J. (2012) *Data Mining Concepts and Techniques (3rd ed)*. USA: Elsevier Inc.
- [7] Brereton R, Gavin L. (2010) Support Vector Machines for classification and regression. *The Analyst*. 135: 230-267. [10.1039/b918972f](https://doi.org/10.1039/b918972f).
- [8] Hamidinekoo A, Denton E, Rampun A, Honnor K, Zwigelaar R. (2018) Deep Learning in Mammography and Breast Histology, an Overview and Future Trends, *Medical Image Analysis*. doi:10.1016/j.media.2018.03.006.
- [9] Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, Summers RM, Giger ML. (2019). Deep learning in medical imaging and radiation therapy. *Med Phys*. 46(1):e1-e36. doi: 10.1002/mp.13264. Epub 2018 Nov 20. PMID: 30367497.
- [10] Sechopoulos I, Teuwen J, Mann R. (2020) Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Seminars in Cancer Biology*. <https://doi.org/10.1016/j.semcancer.2020.06.002>.
- [11] Arcadu F, Benmansour F, Maunz A, Willis J, Haskova Z, Prunotto, M. (2019) Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *NPJ digital medicine*, 2(1): 1-9.
- [12] Wang H, Zheng B, Yoon SW, Ko HS. (2018) A support vector machine-based ensemble algorithm for breast cancer diagnosis, *Eur. J. Oper. Res.*, 267: 687–699. doi: 10.1016/j.ejor.2017.12.001.
- [13] Huang, M-W, Chen C-W, Lin W-C, Ke S-W, Tsai C-F. (2017) SVM and SVM ensembles in breast cancer prediction, *PLoS One* 12: e0161501. doi: 10.1371/journal.pone.0161501
- [14] Tsirogiannis GL, Frossyniotis D, Nikita KS, Stafylopatis A. (2004) A meta-classifier approach for medical diagnosis, in: *Methods Appl. Artif. Intell.* pp 154–163. doi: 10.1007/978-3-540-24674-9_17.
- [15] Tan AC, Gilbert D. (2003) Ensemble machine learning on gene expression data for cancer classification. *Appl. Bioinform.* 2 :1–10. doi: 10.1186/1471-2105-9-275.
- [16] Janghel RR, Shukla A, Sharma S, Gnaneswar AV. (2014) Evolutionary ensemble model for breast cancer classification. doi: 10.1007/978-3-319-11897-0_2.
- [17] Joanna J, Piotr J. (2008) GEP-induced expression trees as weak classifiers using gene expression programming to induce, in: *Lect. Notes Comput. Sci.*, pp. 129–141.
- [18] Schaefer G, Nakashima T. (2015) Strategies for addressing class imbalance in ensemble classification of thermography breast cancer features, in: *2015 IEEE Congr. Evol. Comput. CEC 2015 –Proc*, pp. 2362–2367. doi: 10.1109/CEC.2015.7257177.
- [19] Lederman D, Wang XW, Zheng B, Sumkin JH, Tublin M, Gur D. (2011) Fusion of classifiers for REIS-based detection of suspicious breast lesions, in: *Med. Imaging 2011 Image Perception, Obs. Performance, Technol. Assess.*, pp. 1–9. doi:10.1117/12.877368.

- [20] Nguyen TT, Liew AWC, Tran MT, Pham XC, Nguyen MP. (2014) A novel genetic algorithm approach for simultaneous feature and classifier selection in multi classifier system, in: Proc. IEEE Congr. Evol. Comput. CEC 2014, pp. 1698–1705.
doi: 10.1109/CEC.2014.6900377.
- [21] Ghorai S, Mukherjee A, Sengupta S, Dutta PK. (2011) Cancer classification from gene expression data by NPPC ensemble, IEEE/ACM Trans. Comput. Biol. Bioinform. 659–671, doi: 10.1109/TCBB.2010.36.
- [22] Ferri C, Hernández-Orallo J, Modroiu R. (2009) An experimental comparison of performance measures for classification, Pattern Recognit. Lett., 30: 27–38.
doi: 10.1016/j.patrec.2008.08.010.
- [23] Davis J, Goadrich M. (2006) The relationship between precision-recall and ROC curves, in: Proc. 23rd Int. Conf. Mach. Learn.
- [24] Zaidan AA, Ahmad NN, Abdul Karim H, Larbani M, Zaidan BB, Sali A. (2014) On the multi-agent learning neural and Bayesian methods in skin detector and pornography classifier: An automated anti-pornography system. Neurocomputing, 131: 397-418.
<https://doi.org/10.1016/j.neucom.2013.10.003>.
- [25] Alamoodi, A. H.; Garfan, Salem; Zaidan, B. B.; Zaidan, A. A.; Shuwandy, Mocheheb Lazam; Alaa, Mussab; Alsalem, M. A.; Mohammed, Ali; Aleesa, A. M.; Albahri, O. S.; Al-Hussein, Ward Ahmed; Alobaidi, O. R. (2020). A systematic review into the assessment of medical apps: motivations, challenges, recommendations, and methodological aspect. Health and Technology, (), doi:10.1007/s12553-020-00451-4