Mathematics Letters, eISSN: 2948-3735, 2 (1): 94-107, 2023

COMPARISON OF FEATURES DETECTIONS AND CLASSIFIERS FOR BREAST CANCER CLASSIFICATION

Nur Manisah Suhaimi[1], Wan Nadhirah Wan Abdul Majid[2] and Siti Salmah Yasiran[3*]

Pusat Pengajian Sains Matematik, Kolej Pengajian Pengkomputeran, Informatik Dan Media, Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia.
[1]manisahnur09@gmail.com, [2]wnadhirah43@gmail.com, [3*]salmahyasiran@uitm.edu.my
(indicates the corresponding author)*

ABSTRACT

Breast cancer is one of the most popular diseases which is very harmful especially for women. The method to detect the breast cancer tissue is very important as to aid the radiologist in detecting and diagnosing breast cancer. Computer aided diagnosis (CAD) can be used as one of the systems to help improve the mammographic diagnosis as it can lower false-negative results. A set of 111 abnormal mammographic images from the mini-Mammogram Image Analysis Society (MIAS) database was obtained and used for implementation. Median filter and histogram stretching method are used to pre-process the images. Hybrid Otsu and Mathematical Morphology are used in segmented phases. Then, the features are detected by using both Oriented FAST and rotated BRIEF (ORB) and Binary Robust Invariant Scalable Keypoints (BRISK) as feature detection and feature extraction. A comparison of ORB and BRISK features performance are made as different classifiers, which are Fine KNN and Support Vector Machine (SVM) are also used to classify the breast tissue. The highest accuracy for both ORB and BRISK features is achieved by using SVM classifiers which is 94.5% and 99.4% for testing dataset. However, Fine KNN classifiers are faster in terms of the prediction speed despite having lower accuracy than SVM.

Keywords*: Breast cancer, Binary Robust Invariant Scalable Keypoints (BRISK), Oriented FAST and rotated BRIEF (ORB), Fine K-Nearest Neighbor (KNN), Support Vector Machine (SVM).*

Received for review: 15-03-2023; Accepted: 09-04-2023; Published: 14-04-2023

1. Introduction

Cancer is one of the diseases that is very harmful for humans. This disease can occur in a part or multiple part of the body that is caused by uncontrolled cell replication mechanisms that can lead to death. There are many types of cancer and one of them is breast cancer which is the most common cause of cancer death among women from all over the world. As reported by Ferlay et al. (2021) the cancer statistics in 2020 from International Agency for Research on Cancer (IARC), breast cancer has become the most diagnosed cancer worldwide with 2.26 million cases surpassing lung cancer which has 2.21 million cases and prostate cancer with 1.41 million cases. According to the World Health Organization (WHO), benign tumors are not carcinogenic because their cells have a similar appearance to normal cells, develop slowly and do not invade neighbouring tissues or spread to other parts of the body. Tumors that are malignant are cancerous. Malignant cells can eventually expand beyond the primary tumor to other regions of the body if left untreated. A malignant tumor that has formed from cells inside the breast is referred to as “breast cancer”. A woman that has a first-degree relative that has been diagnosed with cancer has double the risk of getting breast cancer.

A lot of effort has been done to find treatment for this disease. Various studies from a lot of organizations around the world have shown that mammographic screening can reduce the rate of breast cancer (Patil et al., 2020). Mammographic screening remains the most common method that is used to diagnose this disease even though there are a lot of other methods. Therefore, it is very important that we know the accurate reading of this mammogram to find the hidden probability features of breast cancer. The breast cancer features are very complex, so using the computers are very crucial to aid radiologist as they need to examine a big stack of screening mammograms and can be easily distracted. According to Dromain et al. (2013), instead of neglected instances, a lesion that have benign appearance or any unusual discovery that are observed only on one view from previous mammography can be misinterpreted as perceived abnormality and be the most common cause of missed breast cancer. We use computer-aided diagnosis (CAD) in the detection of breast cancer since the CAD systems have been created to improve mammographic diagnosis of breast cancer at screening by lowering the number of interpretations that result in false-negative results.

To get better results, we need to perform feature detection and extraction so that we could transform raw data into numerical features that can be processed while maintaining the information in the original dataset. As there are a lot of features that can be used, we would only like to focus on two features which are Oriented FAST and rotated BRIEF (ORB) and Binary Robust Invariant Scalable Keypoints (BRISK) features.

According to Truong & Kim (2016), BRISK is a point-feature detector and descriptors that can have minimal computing complexity. This is because an intense comparison has been collected from every keypoint neighbourhood in sampling to assembly bit-sharing descriptor and combined it with scale-space FAST-based detection. Saliency criteria are used in BRISK to get the points of interest by identifying across the image and scale dimension. It is stated that BRISK can perform adaptively with a good quality comparable to state-of-the-art algorithms but also significantly at a lower computational cost (Leutenegger et al., 2011). This special characteristic of BRISK can be advantageous for a variety of tasks, especially those that need strict real time limitations or low computing capacity.

Meanwhile, the ORB is a combination of Features from Accelerated and Segments Test (FAST) and Binary Robust Independent Elementary Features (BRIEF). As stated by Rublee et al. (2011) ORB uses the well-known FAST features as a keypoint detector and BRIEF features as the descriptor. Both modified features are fair as they are both low in cost and execute valuable output. Roy et al. (2021) also agreed that the combination of FAST and BRIEF are inexpensive and can carry out worthy output. As ORB performs its task of feature detection, it is a hundred times faster than SIFT features (Roy et al., 2021). According to Rubee et al. (2011), the fact that ORB is not subjected to the license limitations imposed by SIFT and SURF is an additional advantage.

Nowadays, the use of mammography images is very common as one of the ways to detect breast cancer as it has been proven to be the most reliable radiological screening approach (Deshmukh & Bhosle, 2017). In this modern era, we also know that the computational power of computers is getting a lot more intelligent in machine learning classification schemes. For accurate diagnosis, effective classifications of the mammogram are needed so that we can get the right treatment for breast cancer (Deshmukh & Bhosle, 2017). In a study conducted by Utomo et al. (2020) they have utilized BRISK with some hybrid methods, and they have implemented a deep learning approach. Meanwhile, the ORB method is only used for comparison purposes only with other features extraction methods. In addition to that, the classification is only limited to two classes only which are benign and malignant. For this reason, we would like to compare the performance between BRISK and ORB features detection and extraction without any hybrid method by using the machine learning approach. Then, we would like to classify the two classes (benign and malignant) and multiclass types (fatty, dense and glandular) of breast tissue by using KNN classifier.

Referring to Houssein et al. (2021), machine learning is regarded as a subfield of artificial intelligence that employs logical, statistical, and mathematical methods to enable a computer to learn from data without programming by linking the learning issue from sample data to the general concept of inference. Meanwhile, deep learning is the study of artificial neural networks that is related to machine learning algorithms containing numerous layers of nonlinear processing units to extract and transform features (Ongsulee, 2017). Deep learning could give higher accuracy from machine learning and it performs very well on unstructured data but nevertheless it actually requires a significant amount of training data as well as expensive hardware and software (Jakhar & Kaur, 2020). The same point is also stated by Bharati et al. (2020) where deep learning techniques for classifying breast cancer need a lot of labelled images, which are tough to obtain. Many images are needed for a big dataset of breast cancer images, and competent experts are needed to accurately classify the images. Poor results will be provided by using deep learning by using small datasets. Thus, creating datasets takes time and is challenging. So, we have decided to use machine learning as it is reported by Bharati et al. (2020) that it takes less training time and requires less memory.

The objectives of this study are to compare between BRISK and ORB performance in breast cancer classification and to compare the accuracy and efficiency for KNN and SVM classifier.

2. Methodology and Implementation

The proposed CADx can be divided into five main stages. The data will be divided into 60% training and 40% testing (Salama et al., 2018). All of the processes in Figure 1 are using coding and built-in features in MATLAB.

2.1 Pre-processing

The data that will be used in this project is taken from MIAS. The image of the mammography has been minimized to a 200-micron pixel edge in a fixed size of 1024 by 1024 pixels. To remove noise and enhance the Region of Interest (ROI) image, image enhancement techniques will be involved which are median Filter and Histogram stretching (Salleh et al., 2017).

2.2 Segmentation

2.2.1 Otsu and Mathematical Morphology

The data will be divided into 60% training and 40% testing (Salama et al., 2018). After the pre-processing stage, we applied the Modified Otsu method and Mathematical Morphology proposed by Yasiran et al. (2020). Before moving on to the next stage, the breast tissue images from each of the category will be segmented.

2.3 Features Detection and Extraction

2.3.1 Binary Robust Invariant Scalable Key points (BRISK)

The Binary Robust Invariant Scalable Key points (BRISK) algorithm is a feature point detection and description algorithm with scale invariance and rotation invariance. BRISK feature is often used to compare the performance with the other features (Yasiran & Suhaila, 2022). According to Leutenegger et al. (2011), the purpose of BRISK is to overcome the weaknesses of SURF descriptor. The advantages of BRISK features are it has less storage space requirements and has quicker matching speed (Liu et al., 2018). There are three main modules in the BRISK algorithm: scale-space keypoint detection, keypoint description and descriptor matching.

2.3.1.1 Scale-space key point detection

Adaptive and Generic Accelerated Segment Test (AGAST) is the stimulator of methodology in the keypoint detection of BRISK algorithm. Firstly, the scale-space pyramid is created that consists of n octaves c_i and n intra-octaves d_i for $I=\{0,\ 1,\ \ldots,\ n-1\}$ and characteristically n = 4 where octaves are constructed from the novel image by gradually half-sampling the image. To spot the potential region of interest, a detector named FAST 9-16 is utilized on every octaves and intra-octaves with the condition that the threshold value, T must be the same. A non-maxima suppression in scale-space which is commonly used to lessen the redundancy are the subjects of the points that belongs to the region.

2.3.1.2 Key point description

The entire performance of the algorithm depends on the description of the keypoint and also has a massive impact on the matching efficiency. The sampling pattern to represent the feature points is based on the fixed neighbourhood of the keypoint. The BRISK descriptor is composed by the binary bit string. The BRISK sampling pattern consists of small blue circles symbolize the sampling positions and the smoothed area is denoted by the red dashed circles. The interest point is the centre block of 40 by 40 pixels. Four concentric circles are constructed with N = 60 points.

2.3.1.3 Descriptor matching

According to Leutenegger et al. (2011), the similarities between the two feature points will be examined by the BRISK descriptor. The similarity of descriptor is described by calculating Hamming distance since the BRISK techniques use the binary bit string of 0 and 1. The Hamming distance is using the bitwise XOR operation which has the two values to participate. If the value of the matching bits is the same, then the result will be 0 and the value will be set to 1 if the corresponding bits are not matched. The higher the numbers of value 1, the higher the dissimilarity between the two descriptors. The mathematical equation of the BRISK descriptor is represented as Equation (1).

$$p = p_1 p_2 \cdots p_i \cdots p_N \text{ and } q = q_1 q_2 \cdots q_i \cdots q_N \tag{1}$$

where p and q are two BRISK descriptors respectively. The value of p_i and q_i is either 0 or 1. The Hamming Distance equation is as in Equation (2).

$$\mathscr{H} = (p, q) = \sum_{i=1}^{N} p_i \oplus q_i = \sum_{i=1}^{n} I(p_i, q_i)$$
$$\text{where} \quad I(p,q) = \begin{cases} 0 & if \ p = q \\ 1 & if \ p \neq q \end{cases} \tag{2}$$

The term $I(p,q)$ denotes the bit inequality. Then the degree of the two BRISK descriptors is calculated.

2.3.2 Oriented FAST and Rotated BRIEF (ORB)

ORB is a combination of Features from Accelerated and Segments Test (FAST) and Binary Robust Invariant Scalable Keypoints (BRISK). As stated by Rublee et al. (2011) ORB uses the well-known FAST features as a keypoint detector and BRISK features as the descriptor. For the feature point extraction, Oriented FAST (oFAST) is used to increase its uniform rotation (Chuanping et al., 2020). The first step in this algorithm is to set up a coordinates system, A-XY where A is the origin of one corner in the image. Next, compute the gradient $G(x,y)$ in the x and y directions. Implement the Gaussian function on the product of two gradient, G_{x^2}, G_{y^2} and G_{xy} to create the elements below;

$$P = h(G_{x^2})$$
$$Q = h(G_{y^2})$$
$$R = h(G_{xy})$$

Then, calculate the response value, as in Equation (3)

$$V = \det(N) - \tau(trace(N))^2 \tag{3}$$

where $det(N) = PR - Q^2$ and $trace(N) = P + R$. Next, for the feature point descriptor which is BRIEF descriptor, the $M*M$ around the feature point will be selected. Then pick any points (x,y) that is closed to the feature point and then differentiate the pixel values of the two points. If the value of x is less than y, then 1 will be allocated and if x is greater than y, 0 will be allocated. To get the string of binary codes, the neighbourhood for pixel value differentiation are selected according to the number of pairs of points which is basically 256.

2.4 Classification

2.4.1 *k*-Nearest Neighbor (KNN)

KNN is an uncomplicated classifier that functioned well on recognition issues (Maung et al., 2018). Harefa et al. (2016) also stated that among all the machine learning algorithms that are used for classification, KNN is the fair one. It stores obtainable cases and classifies a new case based on the similarity measure. Referring to Than et al. (2018), the value of k will be chosen by the user where k is positive integer and commonly small. According to George and John (2019), below are the phases that included in KNN;

Phase 1: A new formerly unseen instance of something to classify is given. The sets of memorized training samples are looked into by the classifier to discover the K samples that have the closest features. The K value must be positive and small.
Phase 2: The class labels for the example of determined K nearest neighbour will be developed by the classifier.
Phase 3: Then, the labels of those examples will be integrated to make a conjecture for the label of a new object.

In this project, we have conducted parameter analysis which involved the two parameters; number of neighbours and k-NN distance metrics in the process of training the samples which are Euclidean, City Block, Cosine, Correlation, Jaccard, Hamming, Spearman and Chebyshev. The analysis is focusing on obtaining the best result in term of accuracy and efficiency between these two parameters. The highest accuracy of Distance formula will be selected for the 40 percent testing data.

2.4.2 Support Vector Machine (SVM)

Support vector machine (SVM) is a discriminant technique where it classifies unknown individuals into a certain group. According to Awad et al. (2015), SVM has been widely utilized as a classifier, for regression, new determination of tasks and feature reduction. SVM is also renowned for their strong generalization capability, robustness and exceptional worldwide finest solutions. In other word, Zhang (2012) stated that SVM is constructed using a small number of samples in the data present in the training samples to achieve the best classification results. The choice of the kernel function has a significant impact on SVM performance as different kernel functions and coefficients will produce varied results (Zhang, 2012). To achieve a high classification accuracy, we use Radial Basis Function (RBF) Kernel SVM as the complexity of the algorithm will be lessened. The advantage of using RBF kernel is that it reduces computational complexity regardless of the high-dimensional space of the samples. In this project, the parameter value of the selected kernel scale is 1 and 2 box constraints are utilized. The selection of both parameters is based on the findings from Yasiran & Suhaila (2022). They stated that the performance of CADx will decrease as the scale factor rises because the kernel matrix computation becomes wasteful.

2.5 Performance Evaluation

2.5.1 Accuracy

A confusion matrix is a tool to determine the accuracy of the classification algorithm. In this project, we utilized the wrapper method which used a classifier to assess the features. According to Visa et al. (2011), the predicted and actual classification are demonstrated in a confusion matrix of size m x m, where m is the total number of classes.

Predicted / Actual Class	Positive Class	Negative Class
Positive Class	TP	FP
Negative Class	FN	TN

Table 1. Confusion matrix for accuracy metrics (Awad et al., 2015)

The accuracy of the performance will be calculated as in Equation (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{4}$$

The definition of the accuracy will be represented as below,

- $True\ Positive\ (TP)$ when the CAD system predicts a positive case and the actual class is positive.
- $False\ Positive\ (FP)$ when the CAD system predicts a positive case and the actual class is negative.
- $False\ Negative\ (FN)$ when the CAD system predicts a negative case and the actual class is positive.
- $True\ Negative\ (TN)$ when the CAD system predicts a negative case and the actual class is negative.

The disagreement score that related with the confusion matrix is defined in Equation (5).

$$DA = \begin{cases} 0, & if\ \ FP = FN = 0; \\ \dfrac{|FP - FN|}{max\{FP, FN\}}, & otherwise. \end{cases} \tag{5}$$

Based on the equation above, when the value of *FP* and *FN* is equal, the disagreement score is 0 and otherwise.

2.5.2 Efficiency

The time the CAD takes to complete the classification process is used to measure efficiency. The time taken to process the training and testing data will be recorded and compared. The running time is the overall amount of time needed for computations, such as data splitting, data preparation, and model evaluation. Training time is the amount of time needed for a machine to train on a dataset.

The trained model then will be applied to the 40 percent testing data. In this project, the segmentation phase until classification phase are conducted in parallel environment for both training and testing. Since the hardware used is in 64-bits, there will be four workers for the parallel environment configurations. Hence, we choose to utilized k-NN classifier with two number of neighbours and Hamming distance metric. The SVM classifier with the parameter value of the selected kernel scale is 1 and 2 box constraints are also utilized for the BRISK and ORB features detection and extraction for the 40 percent testing data.

3. Results and Discussion

3.1 Parameter analysis for Fine KNN classifiers

For this research, 60% of the datasets are used for training meanwhile 40% are used for testing. We have conducted parameter analysis for Fine KNN classifiers by using different numbers of neighbors and distance metric.

Table 2 shows the performance of the training dataset for Fine KNN by using ORB as feature detection and extraction. Different numbers of neighbors and distance metric are applied to compare the performance of this classifier.

Table 2. Performance accuracy of Fine KNN classifier for training dataset.

Number of Neighbors	**Distance Matrix**	**Accuracy (%)**	**Result Efficiency (PREDICTIONS SPEED (OBS/SEC), TIME)**	
			OBS/SEC	**SECS**
2	Euclidean	87.4%	~700	71.916
	City Block	88.8%	~830	68.578

	Chebyshev	85.6%	~780	71.267
	Cosine	87.2%	~820	68.136
	Correlation	87.8%	~740	94.285
	Spearman	87.0%	~800	84.229
	Jaccard	89.7%	~550	126.22
	Hamming	89.7%	~260	191.11
3	Euclidean	80.5%	~860	64.706
	City Block	82.8%	~890	63.723
	Chebyshev	78.3%	~1000	54.731
	Cosine	80.4%	~910	63.725
	Correlation	80.3%	~850	64.318
	Spearman	80.0%	~860	65.79
	Jaccard	85.0%	~630	91.286
	Hamming	85.0%	~730	79.698
4	Euclidean	72.5%	~850	73.674
	City Block	75.0%	~40	65.826
	Chebyshev	69.3%	~1000	55.378
	Cosine	72.3%	~880	66.803
	Correlation	72.2%	~870	64.325
	Spearman	71.8%	~860	66.819
	Jaccard	79.0%	~600	94.328
	Hamming	79.0%	~730	79.384
5	Euclidean	67.3%	~850	66.91
	City Block	70.4%	~810	69.852
	Chebyshev	62.9%	~1000	54.902
	Cosine	67.3%	~790	67.454
	Correlation	66.9%	~870	65.366
	Spearman	66.7%	~850	67.395
	Jaccard	73.9%	~560	98.04
	Hamming	73.9%	~540	90.946

From Table 2, we have found that the highest accuracy is obtained when the number of neighbors is equal to 2 using Hamming and Jaccard as the distance metric with the same accuracy which is 89.7%. Hamming and Jaccard gives the highest accuracy in comparison with other distance metrics. However, for most neighbors, Hamming has higher prediction speed and takes less time to train the dataset. Meanwhile, Chebyshev remains as the lowest accuracy between other distance metrics despite needing less time taken to train the data. Based on the parameter analysis by using ORB as feature detection and extraction, the highest accuracy achieved by using Fine KNN with Hamming as distance metric and number of neighbors equal to 2.

3.2 Results of Features Detection and Features Extraction for ORB

ORB features result by using Fine KNN and SVM classifiers are recorded and compared.
Figure 1 shows the accuracy comparison of ORB features for Fine KNN and SVM classifiers.

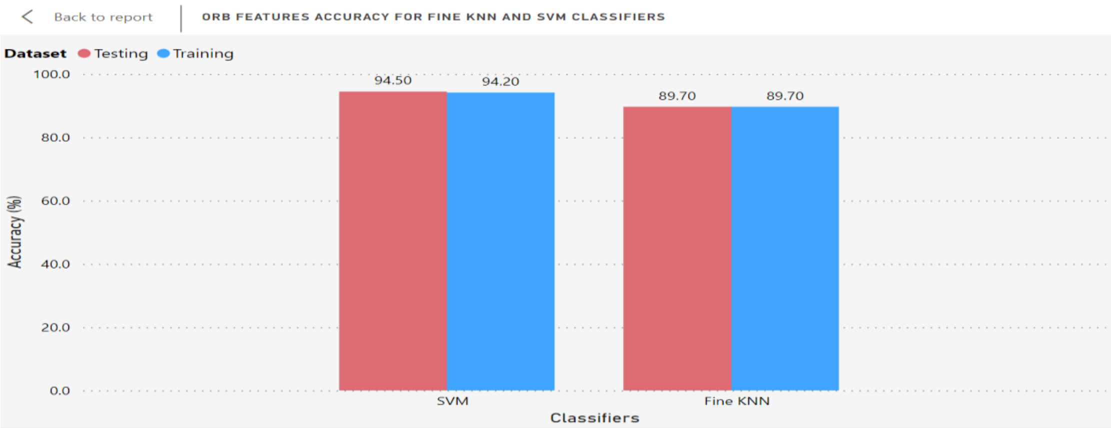
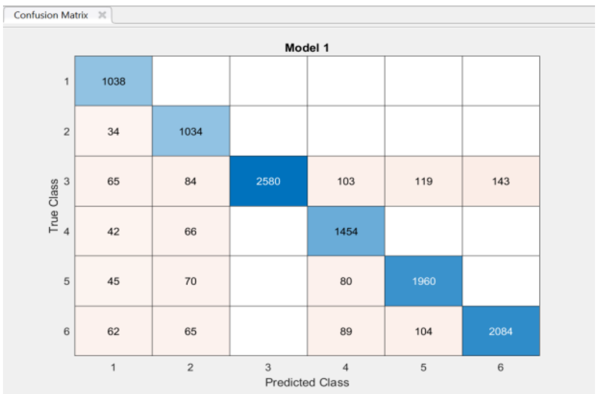


Figure 1. ORB features Accuracy for Fine KNN and SVM classifiers.

From Figure 1, the accuracy for ORB features by using Fine KNN classifiers are the same for both training and testing dataset which is 89.7%. For SVM classifiers, the accuracy for training dataset is 94.2% and it increased to 94.5% for testing dataset. This shows that SVM is more accurate in classifying breast cancer since SVM gives higher accuracy for training and testing dataset compared to Fine KNN classifiers. Figure 2 shows the confusion matrix of ORB features for Fine KNN and SVM classifiers.

Fine KNN	SVM



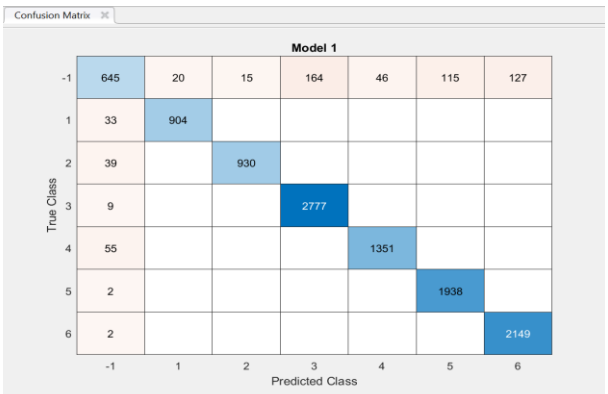


Figure 2. ORB features confusion matrix for Fine KNN and SVM classifiers.

Figure 2 shows the confusion matrix chart for ORB features using the Fine KNN and SVM classifiers. For Fine KNN classifiers, all observations for class 3 are correctly classified but for other classes there exist misclassified observations that are predicted in certain classes but actually belong to other classes. Class 2 has the most misclassified observation which is 285.

From the SVM confusion matrix, we can see that all classes have certain false positive observations that are misclassified but most of the observations are correctly classified. Class '3' has the highest false positive observations which is 164 that are predicted to be in class '3' but actually belong to class '-1'. In this case we have inserted the value '-1' to the data as an "error" to test whether the system can achieve the highest accuracy or not. The reason for inserting the "error" is to avoid the perfect accuracy of 100%. According to Alvarez et al. (2015), error measures are made to have a strong and clear summary potential of the method. It is found that, even though we have inserted the "error", the CADx still gives the highest accuracy of 94.5%.

Figure 3 shows the prediction speed comparison of ORB features for Fine KNN and SVM classifiers.

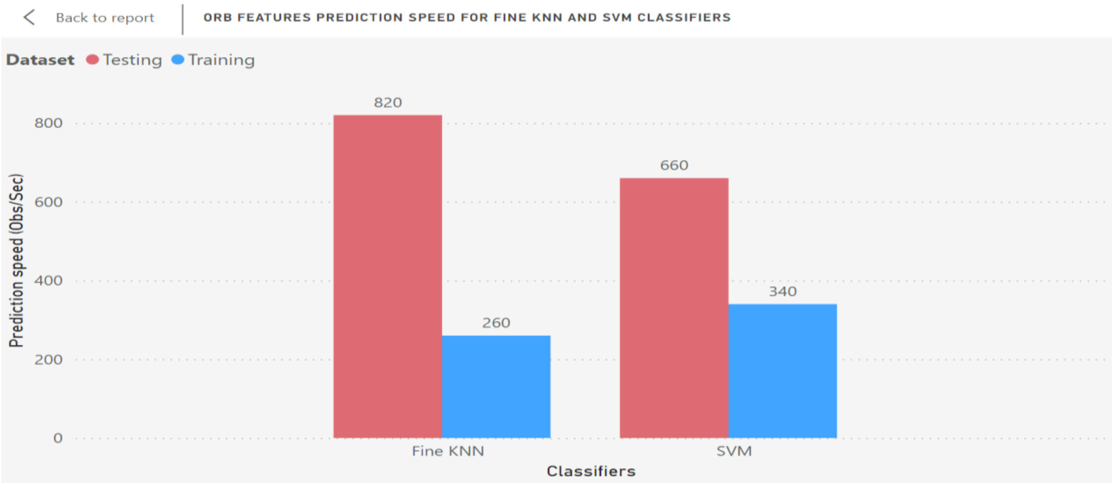


Figure3. ORB features Prediction speed for Fine KNN and SVM classifiers.

From Figure 3, the prediction speed for training dataset by using Fine KNN classifiers is ~260 obs/sec and ~340 obs/sec for SVM classifier. However, we can see that the prediction speed for both classifiers increases for the testing dataset which is ~820 obs/sec and ~660 obs/sec respectively. The prediction speed for SVM is faster than Fine KNN for training dataset but slower for testing dataset. Meanwhile, Fine KNN prediction speed is lower for the training dataset but becomes faster for the testing dataset compared to SVM.

3.3 Results of Features Detection and Features Extraction for BRISK

BRISK features result by using Fine KNN and SVM classifiers are recorded and compared. Figure 4 shows the accuracy comparison of ORB features for Fine KNN and SVM classifiers.

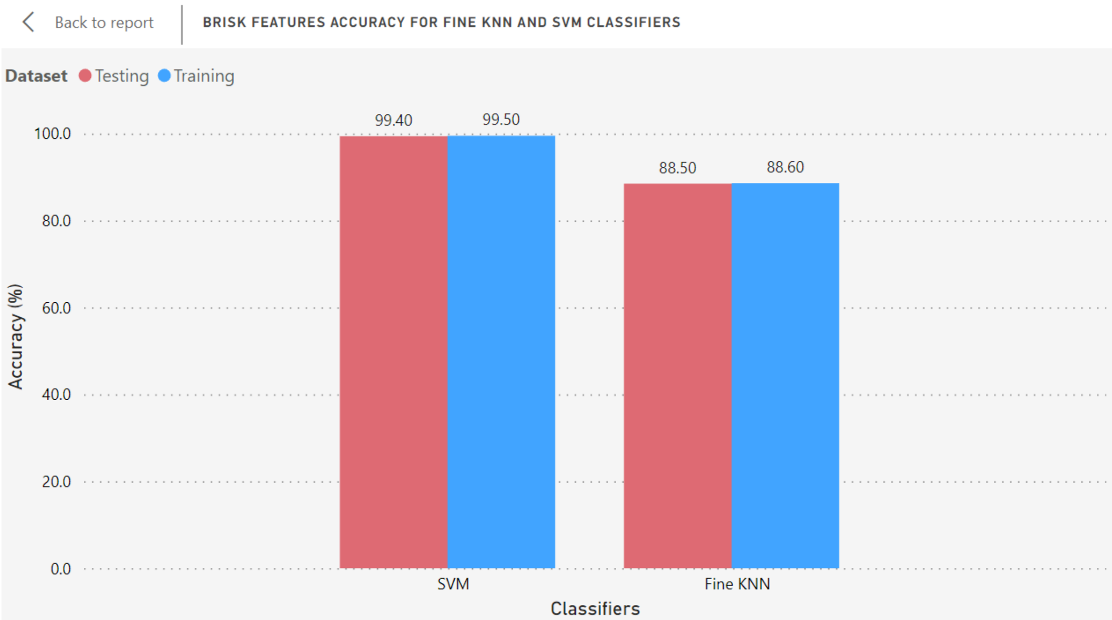


Figure 4. BRISK features Accuracy for Fine KNN and SVM classifiers.

From Figure 4, the accuracy for BRISK features by using SVM classifiers are higher than Fine KNN. SVM classifiers give 99.5% accuracy for training dataset and 99.4% accuracy for testing dataset. Meanwhile, KNN classifier accuracy is 88.6% for training dataset and 88.5% for testing dataset. We can see that the accuracy for BRISK features by using Fine KNN and SVM classifier slightly decreased for the testing dataset compared to the training dataset. Figure 5 shows the confusion matrix of BRISK features for Fine KNN and SVM classifiers.

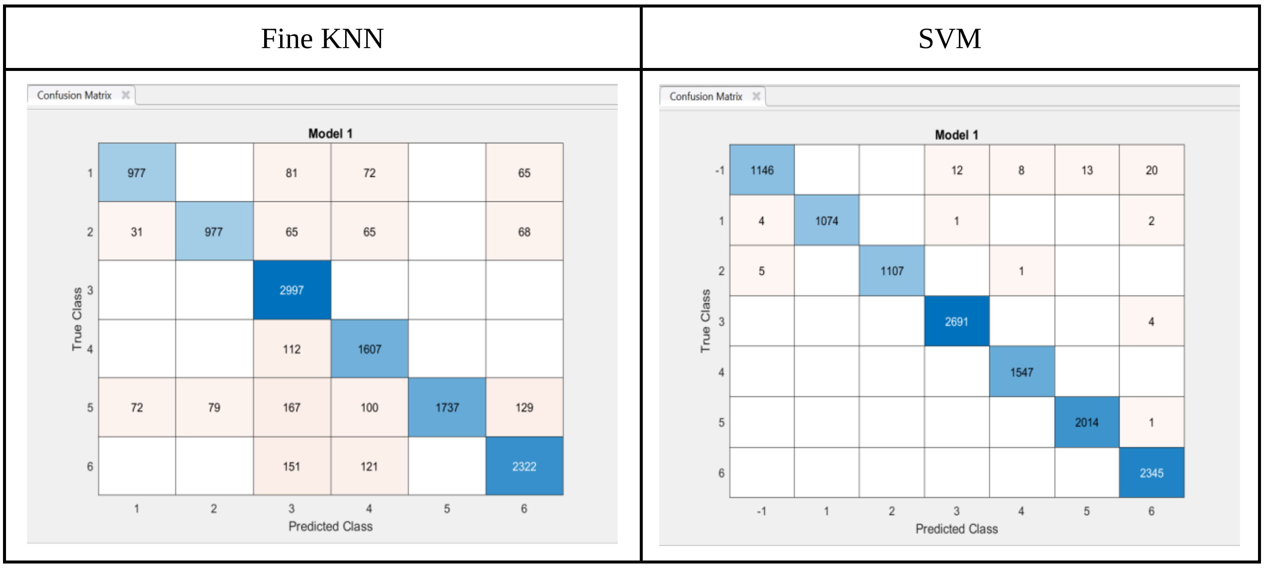


Figure 5. BRISK features confusion matrix for Fine KNN and SVM classifiers.

Figure 5 shows the confusion matrix for BRISK features using the Fine KNN and SVM classifier. For KNN classifiers, there are many misclassified observations from this confusion matrix. Class '3' has the most misclassified observations in comparison with other classes, meanwhile all observations from class '5' are correctly classified. The total of misclassified observations for these classifiers is 1378.

For the SVM classifiers confusion matrix, we can see that only a few of the observations are incorrectly classified. Class '1' and '2' have no misclassified observations meanwhile, for class '-1' and '4' only 9 observations that are misclassified. The total misclassification for SVM classifiers is 71 which is smaller than Fine KNN.

Figure 6 shows the prediction speed comparison of BRISK features for Fine KNN and SVM classifiers.

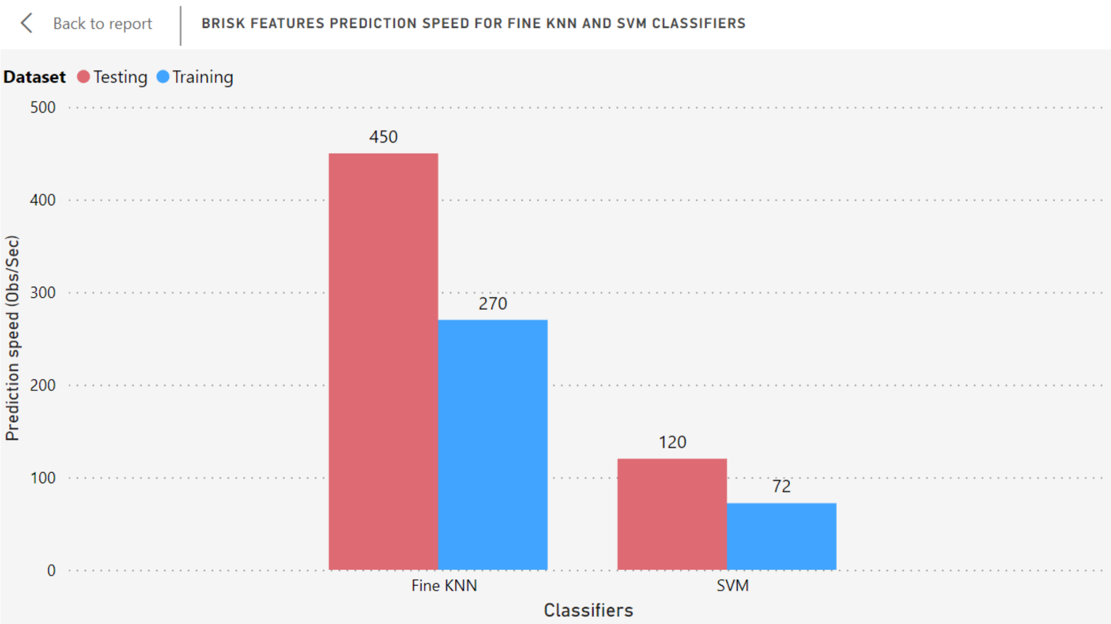


Figure 6. BRISK features Prediction speed for Fine KNN and SVM classifiers.

From Figure 6, we can clearly see that the prediction speed for Fine KNN classifiers is faster compared to SVM. Fine KNN classifiers prediction speed for training dataset is ~270 obs/sec while SVM is only ~72 obs/sec. For testing dataset, the prediction speed using Fine KNN is ~450 obs/sec and ~270 obs/sec by using SVM classifiers.

4. Conclusion and Recommendations

This study utilizes the ORB and BRISK features as both feature detection and feature extraction. Different classifiers which are Fine KNN and SVM are used in this study. Parameter analysis for Fine KNN is experimented to show the combination result using different distance metric and number of neighbors. As a result, Hamming as distance metric and number of neighbors equal to 2 are selected as it gives the highest accuracy for the classifier. The accuracy values for breast cancer classification using ORB features with selected Fine KNN classifiers are 89.7% meanwhile for BRISK features are 88.5%. By using Fine KNN, the accuracy for ORB is higher than BRISK. On the other hand, by using SVM classifiers, the accuracy values for ORB features are 94.5% and 99.4% for BRISK. BRISK features have higher accuracy using SVM classifiers. In addition, for this configuration, the SVM classifier was superior to others as the higher accuracy for both ORB and BRISK features are achieved by using SVM classifiers. However, SVM classifiers take longer training time and lower prediction speed in comparison with though the accuracy is lower.
For the recommendation, further comparison between the accuracy and prediction speed can be studied by using different features such as Speed Up Robust Features (SURF), Fast Retina Keypoints (FREAK), Binary Robust Independent Elementary Features (BRIEF) with combinations of different classifiers such as Decision Tree (DT), Naive Bayes (NB) and Neural Network (NN).

Acknowledgement
The authors would like to thank the School of Mathematics, College of Computing, Informatics and Media Universiti Teknologi MARA for supporting this project.

References

Abudawood, T., Al-Qunaieer, F., & Alrshoud, S. (2018, April). An efficient abnormality classification for mammogram images. In 2018 21st Saudi Computer Society National Computer Conference (NCC) (pp. 1-6). IEEE. https://doi.org/10.1109/NCG.2018.8593208

Al-Masni, M. A., Al-Antari, M. A., Park, J. M., Gi, G., Kim, T. Y., Rivera, P., Valarezo, E., Choi, M. T., Han, S. M., Kim, T. S. (2018). Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. Computer methods and programs in biomedicine, 157, 85-94. https://doi.org/10.1016/j.cmpb.2018.01.017

Alqudah, A. M., Algharib, H. M., Algharib, A. M., & Algharib, H. M. (2019). Computer aided diagnosis system for automatic two stages classification of breast mass in digital mammogram images. Biomedical Engineering: Applications, Basis and Communications, 31(01), 1950007. https://doi.org/10.4015/S1016237219500078

Awad, M., & Khanna, R. (2015). Support vector machines for classification. In Efficient learning machines (pp. 39–66). Springer.
Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). Computer vision and image understanding, 110(3), 346-359. https://doi.org/10.1016/j.cviu.2007.09.014

Bharati, S., Podder, P., & Mondal, M. (2020). Artificial neural network-based breast cancer screening: a comprehensive review. arXiv preprint arXiv:2006.01767. https://doi.org/10.48550/arXiv.2006.01767

Biswas, R., Nath, A., & Roy, S. (2016, September). Mammogram classification using gray-level co-occurrence matrix for diagnosis of breast cancer. In 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE) (pp.161-166). IEEE. https://doi.org/10.1109/ICMETE.2016.85

Chuanping, G., Junjie, Y., Jinchen, L., Ting, Y. & Kexin, S. (2020). An Improved ORB Feature Extraction Algorithm. Journal of Physics: Conference Series, 2-3. https://doi.org/10.1088/1742-6596/1616/1/012026

Deshmukh, J., & Bhosle, U. (2017, March). SURF features based classifiers for mammogram classification. In 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) (pp. 134-139). IEEE. https://doi.org/10.1109/WiSPNET.2017.8299734

Dromain, C., Boyer, B., Ferré, R., Canale, S., Delaloge, S., Balleyguier, C. (2013). Computer-aided diagnosis (CAD) in the detection of breast cancer. European Journal of Radiology, 82(3), 417–423. https://doi.org/10.1016/j.ejrad.2012.03.005

Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. International journal of cancer, 149(4), 778-789, https://doi.org/10.1002/ijc.33588

Ghongade, R. D., & Wakde, D. G. (2017, April). Detection and classification of breast cancer from digital mammograms using RF and RF-ELM algorithm. In 2017 1st International Conference on Electronics, Materials Engineering and Nano- Technology (IEMENTech) (pp.1-6). IEEE. https://doi.org/10.1109/IEMENTECH.2017.8076982

George, J. & John, S. E. (2019). Extreme learning machine-based classification for detecting micro-calcification in mammogram using multi scale features. https://doi.org/10.1109/ICCCI.2019.8821877

Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2009). Vol. 2. Digital Image Processing Using MATLAB. Knoxville: Gatesmark Publishing.

Harefa, J., Alexander & Pratiwi, M. (2016, August). Comparison Classifier: Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) In Digital Mammogram Images. Jurnal Informatika dan Sistem Informasi, 02(02), 37.

Houssein, E. H., Emam, M. M., Ali, A. A., & Suganthan, P. N. (2021). Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. Expert Systems with Applications, 167, 114161. https://doi.org/10.1016/j.eswa.2020.114161

Işık, Ş. (2014). A comparative evaluation of well-known feature detectors and descriptors. International Journal of Applied Mathematics Electronics and Computers, 3(1), 1-6. https://doi.org/10.18100/ijamec.60004

Jakhar, D., & Kaur, I. (2020). Artificial intelligence, machine learning and deep learning: definitions and differences. Clinical and experimental dermatology, 45(1), 131- 132. https://doi.org/10.1111/ced.14029

Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., & Karasfi, B. (2013). Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. Clinical imaging, 37(3), 420-426. https://doi.org/10.1016/j.clinimag.2012.09.024

Lanisa, N. b. (2012, June). DEVELOPMENT OF BREAST CANCER DETECTION SYSTEM USING DIGITAL IMAGE. 23

Leutenegger, S., Chli, M., & Siegwart, R. (2011). BRISK: Binary Robust invariant scalable keypoints. In Proceedings of the IEEE International Conference on Computer Vision. https://doi.org/10.1109/ICCV.2011.6126542

Liu, Y., Zhang, H., Guo, H., & Xiong, N. (2018). A FAST-BRISK Feature Detector with Depth Information. Sensors, 18(11), 3908. MDPI AG. Retrieved from http://dx.doi.org/10.3390/s18113908

López-Cabrera, J. D., Rodríguez, L. A. L., & Pérez-Díaz, M. (2020). Classification of breast cancer from digital mammography using deep learning. Inteligencia Artificial, 23(65), 56-66. https://doi.org/10.4114/intartif.vol23iss65pp56-66

Maung, S. S. & Htay, T. T. (2018). Early Stage Breast Cancer Detection System using GLCM feature extraction and K-Nearest Neighbor (k-NN) on Mammography image. https://doi.org/10.1109/ISCIT.2018.8587920

Martínez-Álvarez, F., Troncoso, A., Asencio-Cortés, G., & Riquelme, J. C. (2015). A survey on data mining techniques applied to electricity-related time series forecasting. Energies, 8(11), 13162-13193. https://doi.org/10.3390/en81112361

McLeod, S. A. (2019, July 19). What does a box plot tell you? Simply psychology: https://www.simplypsychology.org/boxplots.html

Salama, Eltrass & Elkamchouchi, "An improved approach for computer-aided diagnosis of breast cancer in digital mammography", 2018 IEEE international symposium on medical measurements and applications (MeMeA), pp. 1-5, 2018.

Salleh, S., Mahmud, R., Rahman, H., & Yasiran, S. S. (2017). Speed up Robust Features (SURF) with Principal Component Analysis-Support Vector Machine (PCA- SVM) for benign and malignant classifications. Journal of Fundamental and Applied Sciences, 9(5S), 624-643. https://doi.org/10.4314/jfas.v9i5s.44

Salman, M. S., Dey, P. K., Das, P. & Shuvro, R. A. (2016, December). Breast Cancer Detection and Classification Using Mammogram Images. ResearchGate, 3. https://www.researchgate.net/publication/325284976
Ongsulee, P. (2017, November). Artificial intelligence, machine learning and deep learning. In 2017 15th international conference on ICT and knowledge engineering (ICT&KE) (pp. 1-6). IEEE. https://doi.org/10.1109/ICTKE.2017.8259629

Patil, V., Burud, S., Pawar, G., Rayajadhav, T., & Hebbale, S. B. (2020). Breast Cancer Detection using MATLAB Functions. Advancement in Image Processing and Pattern Recognition, 3(2).

Ramadan, S. Z. (2020). Methods used in computer-aided diagnosis for breast cancer detection using mammograms: a review. Journal of healthcare engineering, 2020. https://doi.org/10.1155/2020/9162464

Roy, S. D., Das, S., Kar, D., Schwenker, F., & Sarkar, R. (2021). Computer aided breast cancer detection using ensembling of texture and statistical image features. Sensors, 21(11), 3628. https://doi.org/10.3390/s21113628

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011, November). ORB: An efficient alternative to SIFT or SURF. In 2011 International conference on computer vision (pp.2564-2571). Ieee. https://doi.org/10.1109/ICCV.2011.6126544

Sindhuja, K., Sivakami, R., Sona, C., Sujatha, C. (2019). Tumour Classification and Ellipsification for Breast Cancer in Mammogram Images using Image Processing in MATLAB. SSRG International Journal of Electronics and Communication Engineering (SSRG – IJECE).

Utomo, A., Juniawan, E. F., Lioe, V., & Santika, D. D. (2021). Local Features Based Deep Learning for Mammographic Image Classification: In Comparison to CNN Models. Procedia Computer Science, 179, 169-176. https://doi.org/10.1016/j.procs.2020.12.022

Visa, S., Ramsay, B., Ralescu, A., & van der Knaap, E. (2011). Confusion matrix-based feature selection. CEUR Workshop Proceedings, 710.

Yasiran, S. S., Salleh, S., & Mahmud, R. (2020). Otsu and Mathematical Morphology for Breast Cancer Classification. ASM Science Journal. https://doi.org/10.32802/asmscj.2020.sm26(4.9)

Yasiran, S. S., & Halim, S. A. (2022). Performance of Support Vector Machine (SVM) using Binary Robust Invariant Scalable Keypoints (BRISK) with different RBF kernel scale in classifying type of breast tissue. Mathematics Letters, eISNN: 2948-3735, 1 (2): 1-13, 2022

Zhang, Y. (2012). Support Vector Machine Classification Algorithm and Its Application. In: Liu, C., Wang, L., Yang, A. (eds) Information Computing and Applications. ICICA 2012. Communications in Computer and Information Science, vol 308. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34041-3_27