COMPARISON OF FEATURES DETECTION METHODS IN CLASSIFYING TYPES OF BREAST TISSUE

Nurul Izzati Tajul Azhar[1], Ungku Puteri Nursyahadah Ungku Sharin[2] and Siti Salmah Yasiran[3*]

Department of Mathematics, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor

[1]2019268304@isiswa.uitm.edu.my, [2] 2019229734@isiswa.uitm.edu.my,

[3*] sitisalmah@tmsk.uitm.edu.my

(indicates the corresponding author)*

ABSTRACT

Breast cancer is now the leading cause of cancer death in women. Computer-Aided diagnosis (CADx) has proven effective in assisting medical experts in making an early diagnosis, increasing the probability of recovery. Machine learning is used in Computer-Aided diagnostics to predict whether a patient has benign or malignant cancer. Feature selection algorithms can be utilised to produce more accurate results. Our problem statement is to fill the gap by classify the multiclass type of classification for the breast tissue using SURF as a feature detection approach that distinguishes from other researchers and attempt to compare the performance for both FAST and SURF feature detection. We discovered that Speed Up Robust Feature (SURF) required more computational time than Features from Accelerated and Segments Test (FAST) to complete the feature detection phase. Thus, this study compared and measured the performance of SURF and FAST descriptors for dense, fatty, and glandular tissue. In this report, the digital mammogram image is pre-processed first, this is followed by segmented the pre-processed images. Then, the feature detected, and feature extracted phase will be carried out. Next, the classification phase is conducted using Fine K-nearest Neighbours (KNN). Finally, the model is evaluated. As a result, the SURF method shows the best accuracy for training, while FAST gives the better accuracy for testing. Thus, using the appropriate threshold can help distinguish between tissue and background photos. Both the segmentation phase and the extracted features phase are important to the effective performance of the CADx.

Keywords: *Computer-Aided diagnosis (CADx), Speed Up Robust Feature (SURF), Features from Accelerated and Segments Test (FAST), Fine K-nearest Neighbours (KNN).*

Received for review: 25-08-2022; Accepted: 08-09-2022; Published: 01-11-2022

1. Introduction

According to the World Health Organization (WHO), if the disease is detected earlier, breast cancer treatment can be extremely effective. Breast cancer treatment typically consists of surgical removal, radiation therapy, and medication such as hormonal therapy, chemotherapy and targeted biological therapy. The function is to treat microscopic cancer that has spread from the breast tumour through the blood. This type of treatment, which can prevent cancer growth and spread, helps us in many ways (World Health Organization, 2021). On top of that, Cancer Research Malaysia explains that mammograms have historically been the most effective method of screening for breast cancer. Leading medical organisations recommend

that women over the age of 50 get a mammogram every two years. Women with higher risk factors, on the other hand, may need to begin screening earlier or take a different approach entirely (Cancer Research Malaysia, 2017).

According to Firmino et al., (2016), a Computer-Aided Detection and Diagnosis (CAD) system is a type of computer system that aids in the detection and diagnosis of diseases. The purpose of CAD systems is to enhance radiologists' accuracy while reducing time spent on image interpretation. Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) systems are the two types of CAD systems. CADe are systems that help detect lesions in medical images, while CADx systems characterise lesions, such as differentiating benign and malignant tumours (Firmino et al., 2016). The CAD component is divided into five major phases: data collection and pre-processing, segmentation, feature detection and extraction, feature selection, and classification. According to Cheng et al., (2006), the features detection and extraction phase is a critical step because the performance of CADx is highly dependent on the optimization of features extraction and features selection.

Murtaza et al., (2021) have found an automated detection model for digital histology images using a machine learning approach. They used ten feature extraction methods, including the BRISK descriptors, in their study. An algorithm is developed for the feature selection phase to find the fewest number of redundant features. They did not, however, use the segmentation phase in their study. They used seven different types of classifiers during the classification phase. The developed model's accuracy is found to be 93%.

According to Zhu et al., (2018), Speed Up Robust Feature (SURF) is a new algorithm introduced in 2008 by Hebert Bay. This operator is invariant to rotation, scaling, and brightness changes because it describes the local texture features of key points in different directions and scales of the grayscale image. It also has a good transformation matrix and noise stability, outperforms other methods in terms of uniqueness and robustness, and greatly improves computational efficiency (Zhu et al., 2018). Meanwhile, Features from Accelerated and Segments Test (FAST) is an algorithm proposed by Rosten and Drummond for identifying interest points in images. An interest point in an image is a pixel with a well-defined position that can be reliably detected. Interest points have a high local information content and should ideally be repeatable across images. Image matching, object recognition, tracking, and other applications benefit from interest point detection. The concept of interest point detection or corner detection is not new (Viswanathan, 2011).

Liu et al., (2018) have found many new feature description algorithms, such as Scale-invariant Feature Transform (SIFT), Speeded Up Robust Feature (SURF), Binary Robust Independent Elementary Features (BRIEF), and Binary Robust Invariant Scalable Keypoints (BRISK), have been proposed in recent years under the premise of satisfying the invariance of rotation, scale transformation, and noise. The BRISK algorithm is a scale and rotation invariant feature point detection and description algorithm. It generates the binary feature descriptor by constructing the feature descriptor of the local image using the grey scale relationship of random point pairs in the local image's neighbourhood. When compared to the traditional algorithm, BRISK's matching speed is faster and its storage memory is smaller, but its robustness is reduced (Liu et al., 2018).

According to Truong & Kim (2016), BRISK is a point-feature detector and descriptor that Autonomous Systems Lab has recently developed. BRISK's low computational complexity was achieved through the use of a novel scale-space FAST-based detector associated with the assembly of a bit-string descriptor from intensity comparisons obtained through the dedicated sampling of each keypoint neighbourhood. A saliency criterion is used in BRISK to identify points of interest across image and scale dimensions. Keypoints are detected in the image pyramid's octave layers as well as the layers in between to improve the computation efficiency. Quadratic function fitting is used to obtain the location and scale of each keypoint in the continuous domain (Truong & Kim, 2016).

Yasiran et al., (2020) proposed CAD models that only focus on the classification of benign and malignant tissue. As a result, to complete the feature detection phase, we discovered that SURF required more computational time than FAST. Therefore, the objectives of our research are to compare and measure in term of accuracy and efficiency between SURF and FAST descriptors in classifying breast tissue such as fatty, dense, or glandular. As mentioned earlier, we have used two methods for feature detection: SURF and FAST, while for feature extraction, we used the BRISK method. In addition, we also applied the k-nearest neighbours (KNN) approach for data classification in this paper. According to Joby (2021), the KNN algorithm is a that estimates the likelihood that a data point will join one of two groups based on the data points closest to it. The KNN algorithm is a supervised machine learning method that is used to address issues like classification and regression. Unlike artificial neural network classification, k-nearest neighbours' classification is straightforward to interpret and apply. It is perfect for scenarios with well-defined or non-linear data points (Joby, 2021). In this report, we pre-processed the mammogram image, divide the data for training and testing, then the image is segmented using im2bw, went through the feature detection and extraction phase, next is classified the type of breast tissue whether it is dense, fatty or glandular using KNN approach. Lastly, we evaluate and record the model performance.

2. Methodology

The suggested CADx classify three types of breast tissue is shown in Figure 1.

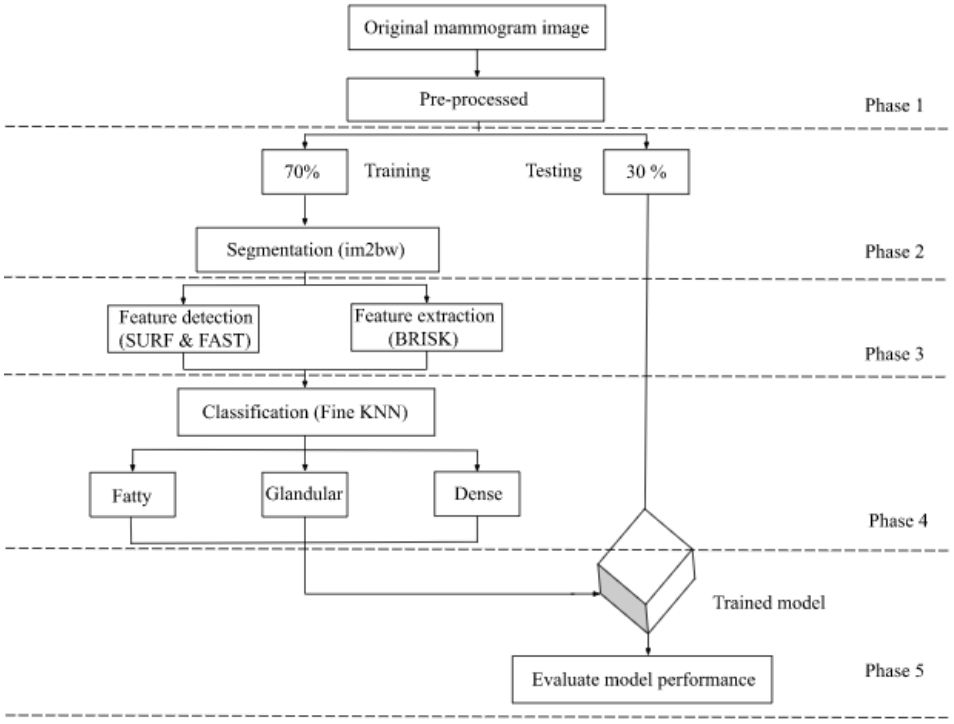

Figure 1. The proposed CADx.

Figure 1 shows the proposed CADx in which there are five main phases included in this research. The first phase is to obtain raw data which is the mammogram images collected from the mini Mammogram Image Analysis Society (MIAS) database (Suckling J et

al.,1994). There are 207 normal images and 115 abnormal images with diverse classes of abnormalities, including 69 benign images and 46 malignant images. The remaining phases are pre-processed, segmentation, feature detection and extraction, feature selection and classification.

2.1 Pre-Processed

Each image's Region of Interest (ROI) is cropped through Adobe Photoshop Software. After that, its size will be standardised to 256 × 256 pixels. The cropping process includes cutting the undesired parts of an image. As a result, we removed practically all of the background information. We utilised the Histogram stretching technique approach, to reduce the effect of over-brightness or over-darkness in images while enhancing the image features. The noise removal step was required prior to this enhancement because it would otherwise enhance noise (Zaıane et al., 2002). The Histogram stretching technique improves image contrast by expanding the dynamic range of grey levels. The median filter is chosen because it smooths the images and minimises noise.

2.2 Segmentation

For the segmentation phase, image binarization with the function “im2bw’ is used.

2.2.1 Binarization of Image

Image binarization is possible with the function "im2bw." Function “im2bw” uses a global threshold to convert a grayscale image to a binary image. The threshold approach is a simple and effective image segmentation method. It divides the image's grey level into numerous portions using one or more thresholds. A threshold is a number used to differentiate between an object's pixel and the backdrop. If a pixel has a gradation higher than or equal to the threshold, it belongs to the object; otherwise it belongs to the background. The threshold approach is particularly successful at distinguishing between an object and a backdrop that is clearly distinct (Huidan & Yuming, 2009). As long as we choose the appropriate threshold, we can separate the item from the backdrop. The threshold used in this study is from 0.1 to 0.9 with a difference of 0.1.

2.3 Feature Detection and Extraction

Selected features should indicate the distinction between benign and malignant cases. In this study, we are interested in comparing the features obtained by using FAST and SURF descriptors. This works efficiently comparing the feature detection which is FAST and SURF combined with BRISK methods. The feature extraction used in this study is BRISK descriptors.

2.3.1 Features from Accelerated Segment Test (FAST)

Features from accelerated segment test (FAST) is a corner detection approach that can be used to extract feature points and afterwards track and map objects. FAST is an algorithm developed by Rosten and Drummond for detecting interest points in images (Rosten & Drummond, 2006). An interest point in an image is a pixel with a well-defined position that can be recognised reliably. Interest points contain a high level of local information value and should preferably be repeated across images. The FAST detector measures the self-dissimilarity of the nucleus using the proportion of the circular arc length L rather than the proportion of the circular area. Fast detector used a circular mask of 16 pixels around the

corner candidate p. When its intensity value I_x is compared to the intensity value of the nucleus I_p using threshold T, each pixel points on the circle, $x\ \varepsilon\ \{\ x_1\ ,\ x_2\ ,\ .\ .\ .\ ,\ x_{16}\}$ occupies one of the three states S_x, as shown in (1).

$$S_x = \begin{cases} d, I_x \leq I_p - T\ (darker) \\ b, I_x \geq I_p + T(brighter) \\ s, I_p - T < I_x < I_p + T\ (similar) \end{cases} \quad (1)$$

If the pixel is to be detected as an interest point, N contiguous pixels out of the 16 must be either above or below I_p by the value T. Depending on the states, the vector P is subdivided into three subsets which is P_d, P_b and P_s . Figure 2 shows the 16 pixels stored for every pixel p surrounding as a vector and repeating the process for all pixels.

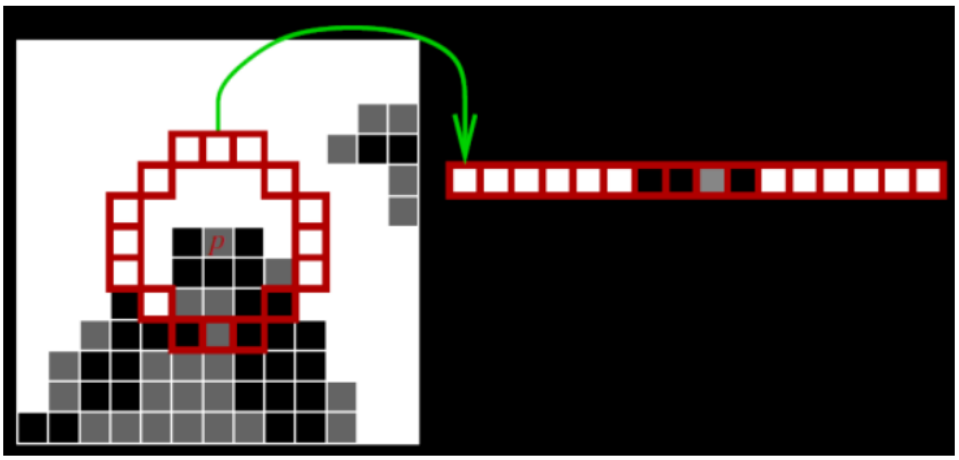

Figure 2. The 16 values that surround the pixel are recorded as vectors.

Figure 2, P is stored as a vector form which contains all the training data. As a result, FAST considers p to be a corner if a segment has at least 12 contiguous pixels that are all brighter or darker than pixel p. Instead, p is a non-corner pixel (Biadgie & Sohn, 2014). The maximum angle of the detected corner is defined by the arc length parameter L. Keeping L as large as possible improves the corner detector's repeatability. The amount of corner responses is controlled by the threshold value T. A large T generates few but strong corners, whereas a tiny T produces corners with smoother slopes.

Using the variable K_p, the ID3 algorithm (decision tree classifier) queries each subset for information regarding the true class. Entropy minimization is the foundation of the ID3 algorithm. Queries as necessary to identify the true class (whether it is an interest point or not) from the 16 pixels. Or to put it differently, choose the pixel x that contains the most details about the pixel p. Mathematically, the entropy for the set P can be expressed as,

$$H(P) = (c + \bar{c})\ log_2(c + \bar{c}) - c log_2 c - \bar{c} log_2 \bar{c}$$
where $c = |\{p|K_p\ is\ true\}|$ (number of corners)
and $\bar{c} = |\{p|K_p\ is\ false\}|$ (number of non-corners) (2)

Non-maximal suppression is applied after detecting the interest point adjacent to one another which can remove the adjacent corners. The process can be summarised mathematically as,

$$V = max \begin{cases} \sum (pixel\ values\ -\ p)\ if\ (value\ -\ p)\ >\ T \\ \sum (p\ -\ pixel\ values)\ if\ (p\ -\ value)\ >\ T \end{cases} \tag{3}$$

where T is the threshold for detection, p is the centre pixel and *pixel values* correspond to the circle containing N consecutive pixels.

2.3.2 Speed Up Robust Feature (SURF)

SURF is a fast interest point detector and descriptor with great repeatability, which means it can dependably detect the same interest point under varied viewing conditions (John & George, 2019). This method employs local image pixels based on interest points and takes the SURF descriptor into consideration. The SURF descriptor is generated in two processes. The first step is to set a reproducible orientation based on data from a circular region around the keypoint. The SURF descriptor is then extracted from a square region aligned to the specified orientation.

2.3.2.1 Integral Images

The integral image of a point $I(x, y)$ can be used to represent an integral image. The (x, y) is the coordinate pixel and I is the grayscale image. The point's mathematical representation of the origin of the rectangle sum in the original image is as follows:

$$U(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i, j) \tag{4}$$

where the whole integrated original image is represented by $U(x, y)$. The pixel value of the original image is represented by the term $I(i, i)$ The Hessian Detector Filters are used in the following phase to identify extreme points using the Hessian Matrix (Viola & Jones, 2004), which is expressed mathematically as;

$$M^{L}(x, y, \sigma) = \begin{bmatrix} \Gamma_{xx}^{L}(x, y, \sigma) & \Gamma_{xx}^{L}(x, y, \sigma) \\ \Gamma_{yx}^{L}(x, y, \sigma) & \Gamma_{yy}^{L}(x, y, \sigma) \end{bmatrix} \tag{5}$$

where L is the scale space and M^{L} is the Hessian Matrix. The image point is then convolved using the two-dimensional Gaussian function, $g^{*}(\sigma) = \frac{1}{2\pi\sigma^{2}} e^{-(x^{2}+y^{2})/2\sigma^{2}}$, followed by the equations shown below;

$$\Gamma_{xx}^{L}(x,y,\sigma)=\frac{\partial^2 g^*(x,y,\sigma)}{\partial x^2}$$
$$\Gamma_{yy}^{L}(x,y,\sigma)=\frac{\partial^2 g^*(x,y,\sigma)}{\partial y^2} \quad (6)$$
$$\Gamma_{xy}^{L}(x,y,\sigma)=\frac{\partial^2 g^*(x,y,\sigma)}{\partial x\partial y}$$

In this case, the terms $\Gamma_{xx}^{L}(x,y,\sigma)$, $\Gamma_{yy}^{L}(x,y,\sigma)$ and $\Gamma_{xy}^{L}(x,y,\sigma)$ are the second order box filters at scale L. If both Γ_{xx}^{L} and Γ_{yy}^{L} are positive while Γ_{xy}^{L} is negative, the local maxima of this filter will be located here. Sigma is a name for the parameter σ in the $g^*(\sigma)$. As a consequence of the previous step, an approximation of a Hessian Matrix was determined, and it may be expressed as;

$$\left|M_{approx}\right|=\mathrm{P}_{xx}.\mathrm{P}_{yy}-\left(\kappa\mathrm{P}_{xy}\right)^2 \quad (7)$$

where $\left|M_{approx}\right|$ is the determinant of the Hessian Matrix's approximation value. This determinant computes each integral image's location. P_{xx}, P_{yy} and P_{xy} are the terms that result from convolution with Γ_{xx}^{L}, Γ_{yy}^{L} and Γ_{xy}^{L} respectively. The weighting parameter κ, also known as the energy conservation between the $g^*(\sigma)$ and the estimated $g^*(\sigma)$ is referred to by the term. We'll build a pyramid scale space to detect the feature points. In this stage, the orientation information for the feature point is also generated. The Haar wavelet is counted in the x and y directions, written as d_x and d_y to produce the features point. As a result, $\sum d_x$ and $\sum d_y$ are used to express each value in the subregion.

2.3.2.2 Orientation Assignment

In being rotationally invariant, we determine a reproducible orientation for the interest points. Haar wavelet is calculated responding in x and y direction which is a circle neighbourhood of radius 6*s* around the interest point, where *s* is the scale at which the interest point was identified. Furthermore, the sampling step is scale dependent and has been chosen to be *s*, and the wavelet responses are computed at that current scale *s*. Furthermore, the wavelets are enormous at large scales. As a result, integral images are used for fast filtering yet again. The sum of vertical and horizontal wavelet responses in a scanning area is then calculated, and the scanning orientation with the highest sum value is chosen as the primary orientation of the feature descriptor.

2.3.2.3 Descriptor Components

The first step in extracting the descriptor is to generate a square region centred on the interest point and oriented along the orientation specified in the Orientation Assignment (Bay et al., 2006). The size of the window in $20s$ which each of the region is divided into 4×4 square sub-regions on a regular basis. We compute a few simple features at 5×5 regularly spaced

sample points for each sub-region. To keep things simple, we refer to d_x as the horizontal Haar wavelet response and d_y as the vertical Haar wavelet response (filter size $2s$). To improve resistance to geometric deformations and localization errors, the d_x and d_y responses are first weighted with a Gaussian centred at the keypoint.

The wavelet responses d_x and d_y are then averaged over each subregion to initialise a set of entries in the feature vector. We additionally extract the total of the absolute values of the responses, $|d_x|$ and $|d_y|$, to bring in information about the polarity of the intensity changes. As a result, for its underlying intensity structure $v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$, each sub-region has a four-dimensional descriptor vector v. This yields a descriptor vector of length 64 for each of the 4×4 sub-regions.

2.3.3 Binary Robust Invariant Scalable Keypoints (BRISK)

The BRISK method's feature extraction is based on the Accelerated Segment Test (AGAST). AGAST is used to detect the keypoint, which is a speed increase over Features from Accelerated Segment Test (FAST) with comparable performance in charge of maintaining identified features. As scale invariance is a key indicator of high-quality keypoints, the BRISK technique extends the FAST algorithm to the image plane as well as scale-space and delivers the best continuous scale space solution (Leutenegger et al., 2011). The BRISK method is divided into three major phases: keypoint detection, keypoint description, and descriptor matching.

2.3.3.1 Keypoint Detection

The Adaptive and Generic Accelerated Segment Test (AGAST) motivates BRISK keypoint detection (Mair et al., 2010). First, a scale space pyramid is created. This is followed by the extraction of stable sub-pixel extreme points using AGAST. The binary feature descriptors are then produced by greyscale correlation of arbitrary point pairs in the local image neighbourhood. Finally, the Hamming distance approach will be used to do the feature matching phase.

2.3.3.2 Keypoint Description

The BRISK descriptor is created as a binary string by concatenating the results of basic brightness comparison tests given a set of keypoints which consist of sub-pixel refined image locations and associated floating-point scale values. Within the block, four concentric circles with uniform distribution and the same spacing within β pixels are constructed, and $N(N = 60)$ points with uniform distribution and the same spacing are obtained on the four concentric circles. A point pair set formed by all pairs of sample points is defined as A:

$$A = \{(p_i, p_j) \epsilon R^2 \times R^2 \mid i < N \wedge j < i \wedge i, j \epsilon N\} \quad (8)$$

(p_i, p_j) is a sampling-point pair of set A. $I(p_i, p_j)$ and $I(p_i, \sigma_j)$ are the grey values smoothed by pixels p_i and p_j, respectively. The following is the local gradient between the two pixel spots.

$$g(p_i, p_j) = (p_i - p_j) \cdot \frac{I(p_j, \sigma_j) - I(p_i, \sigma_i)}{\left\| p_j - p_i \right\|^2} \tag{9}$$

where $1 \leq i \leq N, 1 \leq j \leq N$

The set of short-distance sampling points is defined as Q, while another subset of L long-distance sampling points is defined as $\mathcal{L}$, based on the distance between pixel pairs.

$$Q = \{(p_i, p_j) \in A \mid \|p_i - p_j\| < \delta_{max}\} \subseteq A \tag{10}$$

$$\mathcal{L} = \{(p_i, p_j) \in A \mid \|p_i - p_j\| > \delta_{min}\} \subseteq A \tag{11}$$

The overall characteristic pattern direction of the keypoint k by iterating through the point pairs in $\mathcal{L}$:

$$g = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \frac{1}{L} \cdot \sum_{(p_i, p_j) \in \mathcal{L}} g(p_i, p_j) \tag{12}$$

The long-distance pairings are employed for this computation since it is assumed that local gradients annihilate each other and are hence unnecessary in determining the global gradient (Liu et al., 2018).

2.3.3.3 Descriptor Matching

The descriptors are matched by comparing the commonalities of the descriptors of the two feature points. Since the BRISK approach represents the extracted feature points using a binary bit string of 1 and 0, the similarity of the descriptors is represented by computing the Hamming distance of the descriptor. If their corresponding bits are the same, the result is zero; otherwise, the result is 1. If α and β are both BRISK descriptors, then:

$$\alpha = \alpha_1 \alpha_2 \ldots \alpha_i \ldots \alpha_N \tag{13}$$

$$\beta = \beta_1 \beta_1 \ldots \beta_i \ldots \beta_N \tag{14}$$

The value of α_i and β_i is either 1 or 0. The Hamming Distance equation is given by Equation (13).

$$H(\alpha,\beta) = \sum_{i=1}^{N} \alpha_i \oplus \beta_i = \sum_{i=1}^{n} b(\alpha_i, \beta_i) \tag{15}$$

where

$$b(\alpha,\beta) = \begin{cases} 1 & \alpha \neq \beta \\ 0 & \alpha = \beta \end{cases} \tag{16}$$

The terms $b(\alpha_i, \beta_i)$ refer to the inequality bit. The Hamming distance is calculated to evaluate the degree to which two BRISK descriptors match. The XOR operation is used in the calculation. The lower the degree of descriptor matching, the greater the value of Hamming distance.

2.4 Classification

2.4.1 Fine K-Nearest Neighbour (KNN)

Fine K Nearest Neighbour (Fine KNN) classifiers are utilised in this study to classify benign and malignant tissue. The reason for this selection is that KNN is a simple algorithm that stores the available data and classifies a new case based on the similarity measure (John & George, 2019). Fine KNN has a high and unique ability to minimise feature size. It is used to select and extract processes at the same time. Fine KNN can be efficient for very large datasets by adding incremental changes and providing a more accurate approximation of a probability density function. Moreover, Fine KNN takes one neighbour (K = 1) to distinguish the sample data compared to coarse KNN which uses 100 neighbours (Xu et al., 2013). Fine KNN produces consistent results across all thresholds.

2.5 Performance Evaluation

The suggested CAD system's performance are evaluated in terms of accuracy and efficiency.

2.5.1 Accuracy

According to Zebari et al., (2021), the Receiver Operating Characteristic (ROC) curve is used to evaluate accuracy performance. A receiver operating characteristic (ROC) graph is used to show the comparative trade-off between costs (false positives) and benefits (true positives).

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

2.5.2 Efficiency

The time taken and prediction speed for the CAD to complete the classification process is measured and the efficiency will be compared for both FAST and SURF descriptors.

3. Results and Discussion

In this phase, the analysis of the result from the classification types of breast tissue from two different methods is recorded to compare the accuracy and the efficiency through the time taken and prediction speed for several observations processed per second. For this research, 70% of the data will be used for training and 30% of the data used for testing. Table 1 shows the accuracy performance of training datasets of Fine KNN using FAST as feature detection and BRISK as feature extraction.

Table 1. Accuracy Performance of FAST and BRISK

Threshold	Accuracy (%)	Time (s)	Prediction speed (obs/sec)
0.1	99.7	1.2659	~6000
0.2	96.3	4.8311	~12000
0.3	95.9	1.1837	~33000
0.4	97.8	1.0378	~36000
0.5	94.8	1.1912	~36000
0.6	94.6	1.1269	~30000
0.7	93.9	1.7064	~30000
0.8	96.7	1.1020	~32000
0.9	96.0	0.9818	~42000

Table 2 also shows the accuracy performance of training datasets of Fine KNN using SURF as feature detection and BRISK as feature extraction.

Table 2. Accuracy Performance of SURF and BRISK

Threshold	Accuracy (%)	Time (s)	Prediction speed (obs/sec)
0.1	97.7	1.5771	~7300
0.2	99.3	8.5999	~13000
0.3	96.0	1.5442	~16000
0.4	97.8	1.7679	~23000

0.5	93.8	2.1469	~21000
0.6	94.7	2.6968	~17000
0.7	93.2	2.2241	~13000
0.8	96.5	1.8801	~20000
0.9	95.1	1.2614	~18000

The number of thresholds is experimented from 0.1 until 0.9 for both methods. From the training process, the accuracy performance of FAST and BRISK has the highest accuracy than the performance of SURF and BRISK. The highest accuracy in Table 1 is 99.7% at threshold 0.1 while in Table 2 the accuracy is at the highest for threshold 0.2 with 99.3%. Figure 3 shows the graph of the comparison of the time taken between FAST and SURF.

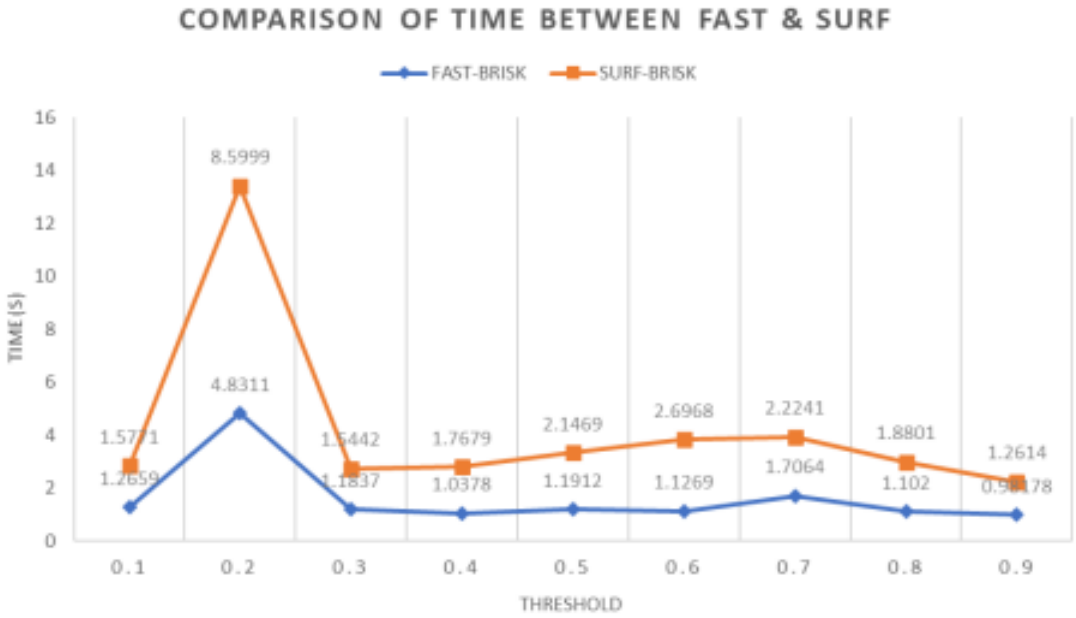

Figure 3. Comparison of time taken between FAST and SURF for training data.

Figure 4 represents the graph of the comparison of the accuracy between FAST and SURF.

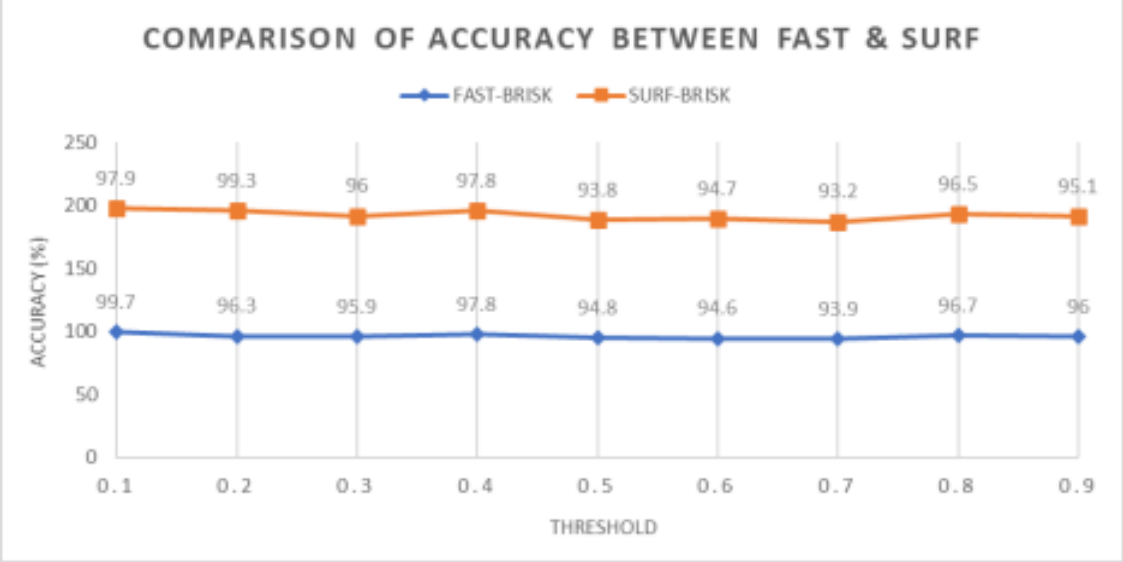

Figure 4. Comparison of accuracy between FAST and SURF for training data.

Figure 3 and Figure 4 above show the comparison between two different methods which are FAST and SURF towards accuracy and time taken. In Figure 3, the time taken for both FAST and SURF in the second threshold took the highest time to compare to the other threshold. For SURF, the huge time taken in threshold 0.2 with the highest accuracy shows that the method is less efficient than FAST. FAST yields more accuracy and has a high-speed performance compared to SURF which is not fast in a real-time system. Less time is taken for FAST than SURF because FAST requires less time to detect the keypoints. After the training data phase is completed, 30% of the datasets are used for testing to make accurate predictions. Table 3 shows the result from both previous methods with a testing dataset.

Table 3. Accuracy performance from testing data

Feature Detection	**Feature Extraction**	**Threshold**	**Accuracy (%)**	**Time(s)**	**Prediction speed (obs/sec)**
FAST	BRISK	0.1	96.6	2.5197	~12000
SURF	BRISK	0.2	95.6	1.2041	~24000

The selected threshold used for testing data is based on the highest accuracy achieved in training. Time and prediction speed are required for the CAD to complete the classification process and the efficiency will be compared for both FAST and SURF descriptors. The results show that the first method which is FAST and BRISK has the highest accuracy which is 96.6% compared to SURF and BRISK with 95.6%.

4. Conclusions and Recommendations

This study utilises the FAST and SURF descriptors for features detection and BRISK as a features extraction stage to compare the efficiency between the two descriptors. For the proposed descriptors, accuracy values for breast tissue classification are 96.6% and 95.6% respectively. The best accuracy was achieved by using FAST as feature detection and BRISK as a feature extraction using threshold 0.1 with 96.6%. Mammogram classification using FAST as feature detection yields more accuracy compared to SURF. FAST is more efficient than SURF as the prediction speed is much faster with ~12000 observations processed per second. To differentiate between tissue and background photos, the appropriate threshold should be used. Both the segmentation phase and the determined features extraction phase play essential keys for the whole CADx to perform accurately. For the recommendation, different types of feature extraction may be further developed. In this study, there is only one feature extraction used which is BRISK. BRISK is frequently used as a feature extraction which encouraged to use of other descriptors to get a better result. The process may be carried out using a variety of descriptors, including Scale Invariant Features Transform (SIFT), Fast Retina Keypoints (FREAK), and Binary Robust Independent Elementary Features (BRIEF). As previously stated, the features detection and extraction phase are important in decreasing computing costs and improving CADx performance.

Acknowledgement

The authors would like to thank to Department of Mathematics, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA for supporting this project.

References

Abudawood, T., Al-Qunaieer, F., & Alrshoud, S. (2018). An Efficient Abnormality Classification for Mammogram Images. 21st Saudi Computer Society National Computer Conference, NCC 2018, 1–6. https://doi.org/10.1109/NCG.2018.8593208

Bay, H., Tuytelaars, T., & Gool, L. Van. (2006). LNCS 3951 - SURF: Speeded Up Robust Features. Computer Vision–ECCV 2006, 404– 417. http://link.springer.com/chapter/10.1007/11744023_32

Biadgie, Y., & Sohn, K. A. (2014). Feature detector using adaptive accelerated segment test. ICISA 2014 - 2014 5th International Conference on Information Science and Applications, 1–4. https://doi.org/10.1109/ICISA.2014.6847403

Cancers. (2017, September 8). Cancer Research Malaysia. Retrieved 19 June 2022, from https://www.cancerresearch.my/our-work/breast-cancer/

Chekanov, M., Shipitko, O., & Skoryukina, N. (2022). Study of Keypoints Detectors and Descriptors Performance on X-Ray Images Compared to the Visible Light Spectrum

Images. IEEE Access, 10, 38964–38972. https://doi.org/10.1109/ACCESS.2022.3159650

Cheng, H.-D., Shi, X. J., Min, R., Hu, L. M., Cai, X. P., & Du, H. N. (2006). Approaches for automated detection and classification of masses in mammograms. Pattern Recognition, 39(4), 646–668.

Fanizzi, A., Losurdo, L., Basile, T. M. A., Bellotti, R., Bottigli, U., Delogu, P., Diacono, D., Didonna, V., Fausto, A., Lombardi, A., Lorusso, V., Massafra, R., Tangaro, S., & Forgia, D. La. (2019). Fully automated support system for diagnosis of breast cancer in contrast enhanced spectral mammography images. Journal of Clinical Medicine, 8(6). https://doi.org/10.3390/jcm8060891

Firmino, M., Angelo, G., Morais, H., Dantas, M. R., & Valentim, R. (2016). Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. *BioMedical Engineering OnLine*, *15*(1). https://doi.org/10.1186/s12938-015-0120-7

Gayathri, S., Gopi, V. P., & Palanisamy, P. (2020). Automated classification of diabetic retinopathy through reliable feature selection. Physical and Engineering Sciences in Medicine, 43(3), 927–945. https://doi.org/10.1007/s13246-020-00890-3

Huidan, C., & Yuming, S. (2009). Application of MATLAB image processing technology in sewage monitoring system. ICEMI 2009 - Proceedings of 9th International Conference on Electronic Measurement and Instruments, 3993–3995. https://doi.org/10.1109/ICEMI.2009.5274144

Joby, A. (2021, July 19). What is K-nearest neighbour? An ML Algorithm to Classify Data. Learn Hub. Retrieved January 21, 2022, from https://learn.g2.com/k-nearest neighbor

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

Kaur, P., Singh, G., & Kaur, P. (2019). Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. *Informatics in Medicine Unlocked*, *16*, 100239. https://doi.org/10.1016/j.imu.2019.100239

Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011). BRISK: Binary Robust invariant scalable keypoints. Proceedings of the IEEE International Conference on Computer Vision, 2548– 2555. https://doi.org/10.1109/ICCV.2011.6126542

Liu, Y., Zhang, H., Guo, H., & Xiong, N. N. (2018). A FAST-BRISK feature detector with depth information. Sensors (Switzerland), 18(11). https://doi.org/10.3390/s18113908

Mair, E., Hager, G. D., Burschka, D., Suppa, M., & Hirzinger, G. (2010). Adaptive and generic corner detection based on the accelerated segment test. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6312 LNCS(PART 2), 183–196. https://doi.org/10.1007/978-3-642- 15552-9_14

Mai Thanh Nhat Truong, & Kim, S. (2016, November). A review on image feature detection and description. Retrieved July 5, 2022, from https://koreascience.or.kr/article/CFKO201629368424723.pdf

Matos, C. E., Souza, J. C., Diniz, J. O., Junior, G. B., de Paiva, A. C., de Almeida, J. D., da Rocha, S. V., & Silva, A. C. (2018). Diagnosis of breast tissue in mammography images based local feature descriptors. *Multimedia Tools and Applications*, *78*(10), 12961–12986. https://doi.org/10.1007/s11042-018-6390-x

Murtaza, G., Wahab, A. W. A., Raza, G., & Shuib, L. (2021). Breast Cancer Detection via Global and Local Features using Digital Histology Images. Sukkur IBA Journal of Computing and Mathematical Sciences, 5(1), 1–36.

Nedra, A., Shoaib, M., & Gattoufi, S. (2018). Detection and classification of the breast abnormalities in Digital Mammograms via Linear Support Vector Machine. Middle East Conference on Biomedical Engineering, MECBME, 2018-March, 141–146. https://doi.org/10.1109/MECBME.2018.8402422

Rosten, E., & Drummond, T. (2006). Erickson - 2008 - Methods of treatment using anti-erbb antibody-maytansinoid conjugates.pdf. 430–443.

Raj, A. (2020, June 9). *Feature description & Extraction*. K. R. I. S. S. https://medium.com/k-r-i s-s/feature-description-extraction-78bbb19ccc8b#51c1

Salleh, S., Mahmud, R., Rahman, H., & Yasiran, S. S. (2018). 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. https://doi.org/10.4314/jfas.v9i5s.44

Suckling J, et al. (1994) The Mammographic Image Analysis Society Digital Mammogram Database, Exerpta Medica. International Congress Series 1069 pp. 375–378

Truong, M. T. N., & S. Kim. (Ed). (2016). A Review on Image Feature Detection and Description [Review of *A Review on Image Feature Detection and Description*]. *Proceedings of the 2016 Fall Academic Conference*, *23*(2), 677– 678. https://www.koreascience.or.kr/article/CFKO201629368424723.pdf

Viola, P., & Jones, M. (2004). Robust Real-Time Face Detection Intro to Face Detection. *International Journal of Computer Vision*, *57*(2), 137–154.

Viswanathan, D. (2011). *Features from Accelerated Segment Test (FAST)*. Www.semanticscholar.org. https://www.semanticscholar.org/paper/Features-from Accelerated-Segment-Test-(-FAST-)-Viswanathan/cd267a4b04d835dbecf01d47fc69ed3a38c23055

WHO. (2021, March 26). Breast cancer. Www.who.int. Retrieved July 5, 2022 from https://www.who.int/news-room/factsheets/detail/breast-cancer

Xu, Y., Zhu, Q., Fan, Z., Qiu, M., Chen, Y., & Liu, H. (2013). Coarse to fine K nearest neighbor classifier. *Pattern Recognition Letters*, *34*(9), 980– 986. https://doi.org/10.1016/j.patrec.2013.01.028

Yasiran, S. S., Salleh, S., & Mahmud, R. (n.d.). Standard Deviation Based Otsu for Breast Cancer Classification. ASM Science Journal. Retrieved January 21, 2022, from https://www.akademisains.gov.my/asmsj/article/standard-deviation-based-otsu-forbreast cancer-classification/

Zaıane, O., Antonie, M., & Coman, A. (2002). Mammography classification by an association rule-based classifier. Mdm/Kdd, 62– 69. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.88.2127&rep=rep1&type=pdf#page=68

Zebari, D. A., Ibrahim, D. A., Zeebaree, D. Q., & Mohammed, M. A. (2021). Applied Sciences Breast Cancer Detection Using Mammogram Images with Improved Multi-Fractal Dimension Approach and Feature Fusion

Zhu, Z., Zhang, G., & Li, H. (2018). *SURF feature extraction algorithm based on visual saliency improvement* [Review of *SURF feature extraction algorithm based on visual saliency improvement*]. *5*(3), 13.