Abstract—Breast cancer is one of the most feared cancer among woman due to the high statistics of female affected by it around the world. The screening tools that are used for breast cancer detection are mammogram, ultrasound, MRI, CT-scan, PET and thermogram. Although mammogram considered as the gold screening tool among all, this procedure are invasive, expensive, time consuming and cause physical discomfort to the patient. Therefore, to overcome this problem, high interest have been given to thermogram, an ionizing free and non-contact procedure. However, the absence of clear edges and low contrast in thermogram may cause difficulty to analyze the image. This paper will discuss pre-processing method to detect breast and body boundaries to eliminate unwanted area and to extract colors which represent the temperature emitted from the breast area. This pre-processing procedure is a procedure before the data extracted for feature extraction and image classification process.

I. INTRODUCTION

Breast cancer referred to the cells in the breast tissues that grow, change and multiply rapidly without control resulting to the formation of lump or mass as shown in Figure 1.1 [1]. These masses are called tumor and can be either cancerous (malignant) or non-cancerous (benign). Malignant tumor is an invasive type of tumor which invade neighboring tissues and spread to the other part of the body. Sometimes it also metastases, a condition where the tumor spread by move away from its original organ to another organ or bones. Benign cancer however didn’t spread to other parts of the body and always in the condition where it is easy to be removed [2]. Breast cancer usually start in the ducts, structures that carry milk from lobules toward the nipple or lobules, the milk glands.

Figure 1.1 Breast cancer

Breast cancer is one of the most common cancers affecting woman and the most common source of death among middle aged women. Based on the World Health Statistics 2011 by Global Health Observatory (GHO), the mortality among female population all over the world cause by malignant neoplasm is about 11.81% and breast cancer are the highest with 15.80% compare to other types of cancer [3]. Low-income countries show the lowest survival rate with below 40% while middle-income country has around 60% survival rate and high-income countries has survival rates ranging from 80% and above.

In Malaysia, cancer become one of the top ten causes of hospitalization and one of the top five causes of death in public and private hospital. National Cancer Registry (NCR) report in 2007 show breast cancer as the highest type of cancer affecting Malaysian population (18.1%) compare to the other type of cancer such as head and neck, (13.2%), colorectal (12.3%), trachea, bronchus and lung (10.2%), and cervix (4.6%) [4]. Hence, to reduce the number of mortality due to breast cancer, early detection is crucial. Research has shown that 85% chance of cure if the tumor detected earlier and only 10% chance if detected late [5].

There are several imaging techniques use for detection and screening of breast cancer such as mammography, ultrasound, Magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET) and thermography [4]. Mammography is consider as the gold standard screening tool for breast cancer. This process uses system with low dose x-ray, high contrast and high-resolution film [2]. Mammography detect any anatomical changes or structural distinction between tumor and breast tissue. However, mammography have several limitation. Firstly, the tumor must achieve certain size or thickness to be detected in mammogram. This cause the treatment to be delayed and reduce the chance of cure for the patient. Besides that, the density of the breast also affected the sensitivity of mammogram to differentiate between normal tissue and tumor. Mammogram also exposed patient with X-ray radiation which may destroyed the tissue or cause tissue mutation. Not to mention, mammogram procedure are expensive, time consuming, cause physical discomfort and ineffective for surgically implanted breast or post-operative breast scar [2].

Due to the drawback of mammogram, a high interest have been given to thermography as breast cancer screening tool. Thermography did not use ionizing radiation, venous access or other invasive procedure [6]. Besides that, it is a quick, painless, economic, risk free, and patient friendly imaging method. Therefore, it is suitable for all sizes and density of breast, pregnant or nursing women, implanted or non-
implanted breast, and even post-operative patients. Thermography detected physiological changes by measuring the infrared radiation emitted by the body. Human body usually emitted narrow band wavelength of infrared radiation of 8µm – 12µm [5]. Breast normally produced symmetry temperature distribution. Cancer cells however cause asymmetry pattern in breast thermal image. The changes is because cancer cells released more heat by means of higher metabolic activity compare to surrounding tissue. Additionally, cancer cells increase local blood flow by producing chemicals to develop new blood vessels (angiogenesis) for food supply. This cause an increase of temperature to the area.

II. LITERATURE REVIEW

Even though thermography gain some popularity as breast cancer screening tool, thermal image also seem to have several limitation. Low contrast, low signal to noise ratio, absence of clear edges and no definite shape cause difficulty to analyze and segmented thermal image [7].

A survey done by Borchartt et al. reported [8], several studies have been using the combination of edge detection technique with Hough Transform (HT) to extract the region of interest (ROI). HT was used to detect the parabolic shape of inframammary fold, edge between the breast and the lower part of body in thermal image. This paper also discuss on several other segmentation technique such as automatic algorithm based on Level Set Method (LSM), semi-automated algorithm to divide breast area into four quadrants, manual segmentation and wavelet-based denoising method.

S.S. Suganthi et al. [9] comparing the performance of Gaussian filter and anisotropic diffusion filter to identify the lower breast boundaries and inframammary fold. Gaussian filter sigma are directly proportional with the smoothness of the image. High value of sigma will reduce the ability to detect any visible edges. On the other hand, anisotropic filter that depend on two parameters (edge strength threshold (K) and number of iterations (T)), show preserve, smooth and sharp edges.

M. EtehadTavakol [10] use Canny edge detector. It consist of 5 major steps for edge detection; smoothing, gradient finding, non-maximum suppression, double thresholding, and edge tracking by hysteresis. The Canny edge detector use to eliminate inner edge and extract outer boundaries. The segmentation than follow by the extraction of two lower boundary.

Sheeja V. Francis [11] pre-processing technique start by converting thermal image into gray scale image. Enhancing image contrast, manual cropping to obtain breast area, Sobel edge detector and several morphology steps use to obtain their region of interest.

Based on the review done, combination of Canny edge detector and HT algorithm are the most popular technique use in segmentation for breast thermal image. This technique however increase the image processing time. More than 96% processing time are due to the execution of HT algorithm. Hence, another algorithm put into consideration such as Gaussian filter use in [7] or Sobel edge detector use in [10]. This paper will discuss on the use of Sobel edge detector. Another similarity shared by previous paper are their first pre-processing method which is converting RGB image to grayscale image. Equation (1) shows one of the algorithm that can be used in the conversion. This equation also shows some of the data might be lost as each component are reduce by weighted sum value. Therefore, this paper will discuss on pre-processing thermograms on each color plane instead of converting the image into grayscale image.

\[ 0.2989 * R + 0.5870 * G + 0.1140 * B \]  

III. METHODOLOGY

In this work, 50 images of thermogram were chosen to be used as the data sets. The images was obtain from an online database [11]. The images are taken in University Hospital of UFPE (Brazil) and stored in the database with several other information such as age, exam date, patient family history, and patient preparation before the exam. Ethical Committee of UFPE approved the acquisition procedure and the storage of images registered at Brazilian Ministry of Health.

The images resolution are around 600 x 480 pixels with size range from 640 to 670 kbytes. The patient will go through specific protocol for the image acquisition based on two category represent the body heat transfer behaviour; static and dynamic. Two hours before image capturing, patient are highly recommended to not smoking, consumed caffeine or alcohol, doing any physical activities, and applying cream or ointment on breast area. At the time to capture the image, patient body temperature will be recorded and any accessories need to be put aside. For static acquisition, the first step is patient will be seated on the bench for 10 minutes with her arms on the armrest to stabilize body temperature with the surrounding temperature. During this 10 minutes, atmospheric temperature and relative humidity is observed. The picture then capture with patient standing 1.0 meter from the infrared camera and hands over her head. For dynamic acquisition, cooling fan will be directed to the breast with arms above her head. The fan then turn off when the temperature around the chest central region drop to less than 30.5°C and finally the image of the breast is taken [11].

Breast thermogram is a ‘truecolor’ image. It consist of an M-by-N-by-3 data array of RGB values for each pixel. The color of each pixel is determined by the combination of the red, green, and blue intensities stored in each color plane at the pixel's location. Instead of converting RGB thermal image to grayscale image, the analysis is going to be done on each color plane. This is to reduce any information lost during the process. The first step was to obtain the object boundaries, inframammary fold edges and nipple point. This step is to eliminate noises and unwanted area below the breast. Besides that, the edges and nipple point will be used as reference line for future work in feature extraction and image classification. In this paper, we are using Sobel edge
detector to find the edges. Figure 3.1 show the process taken to obtain the edge image.

![Figure 3.1 Edge detection process.](image)

Sobel edge detection or Sobel filter convolved with image using two 3x3 kernels to calculate approximations of the derivatives. One kernel is to compute horizontal changes and the other is for vertical changes. Equation (2) and Equation (3) show the convolution operation towards original image define as A. \( G_x \) represent the gradient approximation for horizontal changes while \( G_y \) represent the gradient approximation for vertical changes. Equation (4) show the formula use to combine both gradient approximation to obtain gradient magnitude for each pixel in the original image.

\[
G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \\
(2)
\]

\[
G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \\
(3)
\]

\[
G = \sqrt{G_x^2 + G_y^2} \\
(4)
\]

Second step was to detect different temperature range in thermal images. In thermal pseudo color image, several ranges of temperature will be represent by several colors. Therefore, extracting each temperature range contour and show the location by merging it with the output of edge detection will increase the possible features that can be extracted for more accurate classification.

IV. RESULT AND DISCUSSION

The RGB image separate to its individual color plane. Figure 4.1 show the images of each color plane that extracted from the original image.

![Figure 4.1 Color Planes of Image](image)

Each of the color plane then process using Sobel edge detector. Sobel method are using Sobel approximation to the derivatives to detect edges. Figure 4.2 show several result of edge detection. 34 out of 50 thermal images show a promising result. The other 16 images contain weak edges at inframammary fold due to the low contrast between breast area and surrounding area. Future work will focus on improving the segmentation technique to obtain edges on these images.

![Figure 4.2 Edge Detected](image)

The second step was to extract individual color range. Each color shown in thermal images represented by certain range of number for each plane. By knowing this range, color can be extracted from the original thermal image. Figure 4.3 show the extraction of yellow color from the original image. This color represent certain range of temperature. By combining the color extraction result with the edge detection result shown in Figure 4.2, the location can be seen clearly. By extracting color in thermal image, the exact contour to represent each temperature can also be obtain.

![Figure 4.3 Color Extraction](image)
V. CONCLUSION

Thermal images usually capture from shoulder to waist. Detecting inframammary fold edge and body boundaries allowing researcher to focus on processing breast region as other information outside that region consider as noise. Multiple color shown in thermogram without exact edges should also be process to gain data. Therefore, edge detection and color extraction are both important aspects to be consider in image pre-processing of thermal images. This early process allow researcher to gain more information from the image. This information will then be used in feature extraction and image classification. Multiple type of data will give opportunity for researcher to test and train classifier for more accurate result.

ACKNOWLEDGMENT

This research received funding from the Ministry of Higher Education, Government of Malaysia under the University Malaya Research Grant (UMRG), Advanced Engineering and Technology (AET), Innovative Technology Research Cluster, University of Malaya (UMRG-RP020A-13AET). Thermal image obtained from an online database which were taken and stored by University Hospital of UFPE (Brazil) with patients consent and registered at Brazilian Ministry of Health [9].

REFERENCES