



Design and Development of the Computer Vision Algorithm for a Real-Time Breast Self-Examination

A Thesis Presented to the
Department of Electronics and Communications Engineering
Gokongwei College of Engineering
De La Salle University

In Partial Fulfillment of the Requirements for the Degree of
Master of Science in Electronics and Communications Engineering

Submitted By:
Eman Mohammadi Nejad
MS-ECE
Adviser:
Dr. Elmer P. Dadios

2014




APPROVAL SHEET

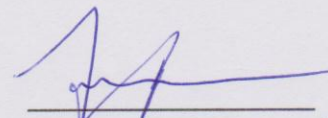


This thesis entitled:

“Design and Development of the Computer Vision Algorithm for a Real-Time Breast Self-Examination”


Prepared and submitted by **Eman Mohammadi Nejad**, in partial fulfillment of the requirements for the degree of **Master of Science in Electronics and Communications Engineering** has been examined and is recommended for acceptance and approval for **ORAL EXAMINATION**.


Dr. Elmer P. Dadios
Thesis Adviser

Approved by the Committee on Oral Examination with a grade of **PASSED** on March 20, 2014.


Dr. Laurence A. Gan Lim
Panel Chair
Assistant Prof. Melvin K. Cabatuan
Member
Assistant Prof. Edwin Sybingco
Member

Accepted in partial fulfillment of the requirements for the degree of Master of Science in Electronics and Communications Engineering.


Dr. Rosemary R. Seva
Dean
Gokongwei College of Engineering



Abstract

Breast cancer is the most common cancer among women worldwide. Early detection of breast cancer is the key to reduce breast cancer mortality. Breast self-examination (BSE) is considered as the most cost-effective approach available for early breast cancer detection. A large fraction of breast cancers are actually found by patients using this technique today. In BSE, the patient should use a proper search strategy to cover the whole breast region in order to detect all possible tumors and abnormalities. However, the majority of women don't perform the correct BSE due to the lack of confidence and knowledge on BSE performance. Therefore, there is a need for an application to educate and evaluate the BSE performance. So, women can perform the BSE using a webcam, computer, and the stated application that has the ability to evaluate the BSE performance. The general objective of this thesis is to design and develop the computer vision algorithm to evaluate the BSE performance in terms of covering the entire breast region. In this research, three individual algorithms were developed and implemented. The first algorithm focuses on detecting and tracking the nipples in frames while a woman performs BSE; the second algorithm focuses on localizing the breast region and blocks of pixels for making palpation on the breast, and the third algorithm focuses on detecting the palpated blocks in the breast region. The palpated blocks are highlighted at the time of BSE performance. Finally, if all areas in the breast region are palpated, the BSE training is considered as a correct performance in terms of covering and palpating the whole breast region. If any abnormality, such as masses, is detected, then this must be reported to a doctor, who will confirm the presence of this abnormality or mass and proceed to do other confirmatory tests.



Acknowledgement

I would like to thank God for the opportunity of working with computer vision. I try to recognize the wonderful algorithms He has created for us to see. Also, I would like to express my sincere gratitude to my advisor, Dr. Elmer P. Dadios, for all his guidance, suggestions and supports throughout this master thesis. I would like to extend my appreciation to Dr. Laurence A. Gan Lim, Dr. Jose Ma. C. Avila, and Melvin Cabatuan for their comments and suggestions about this research.

Besides, I am grateful to all the participants who generously contributed their images and videos while performing BSE to our database.

I also wish to thank the Philippines Commission on Higher Education Research Network (CHED-PHERNET) Program in Sustainability Studies for their funding support of this research.

Lastly, I thank my wife, Atefeh, for all her love and support. I am also forever indebted to my parents who have been supporting me in every aspects of life.



Table of Contents

Approval Sheet.....	i
Abstract.....	ii
Acknowledgement	iii
Table of Contents.....	iv
List of Figures.....	viii
List of Tables	x
Nomenclature.....	xi
Chapter 1	1
Introduction	1
1.1. Background of the Study.....	1
1.2. Statement of the Problem	2
1.3. Significance of the Study.....	3
1.4. Objectives of the Study.....	3
1.5. Scope and Limitations of the Study	4
Chapter 2.....	6
Review of Related Literature	6
2.1. Breast Cancer	6
2.2. Breast Cancer Screening	7
2.2.1. Mammogram	7
2.2.2. Ultrasound.....	7
2.2.3. Magnetic Resonance Imaging (MRI)	8
2.2.4. Clinical Breast Examination (CBE).....	8
2.2.5. Breast Self-Examination (BSE).....	9
2.3. Computer Vision Algorithms for BSE.....	11
2.4. Object Detection Algorithms.....	13



Chapter 3	17
Theoretical Framework	17
3.1. OpenCV library	17
3.2. Qt Framework.....	17
3.3. Pattern Recognition.....	18
3.4. Artificial Neural Network (ANN)	20
3.5. Histogram of Oriented Gradients (HOG)	22
3.5.1. Gradient Computation	23
3.5.2. Orientation Binning	25
3.5.3. Descriptor Blocks	25
3.5.4. Block Normalization	25
3.6. Color Features	26
3.6.1. The RGB Color Model	26
3.6.2. HSV Color Model	27
3.6.3. Color Feature Extraction	27
3.7. Gaussian Smoothing.....	28
3.8. Template Matching	29
3.8.1. Matching Functions	30
3.8.2. Normalized Methods	31
3.9. Farneback Optical Flow	32
3.9.1. Displacement Estimation in Farneback Optical Flow	34
3.10. Histogram Equalization.....	35
3.10.1. Image Histogram	35
3.10.2. Improving the Contrast	36
3.11. Haar-like Feature	37
3.12. AdaBoost Classifier	39
3.13. Cascade	42



Chapter 4.....	44
Procedure.....	44
4.1. System Overview	44
4.2. Nipple Detection and Tracking	48
4.2.1. Breast Detection	48
4.2.2. Nipple Detection.....	51
4.2.3. Training the Neural Network.....	54
4.2.4. Nipple Tracking.....	56
4.3. Breast Region Identification	57
4.4. Detecting the Palpated Blocks	60
Chapter 5.....	68
Results and Evaluation	68
5.1. Datasets	68
5.1.1. Dataset-1.....	68
5.1.2. Dataset-2.....	69
5.1.3. Dataset-3.....	69
5.1.4. Dataset-4	70
5.2. Performance Criteria	71
5.3. Evaluation Results.....	74
5.3.1. Breast Detection	74
5.3.2. Nipple Detection	75
5.3.3. Nipple Tracking.....	76
5.3.4. Palpation Detection.....	77
5.3.5. Evaluation of the Integrated Algorithm.....	79
5.4. Experiments	80
Chapter 6.....	85
Conclusions	85



6.1. Summary	85
6.2. Comments for Future Work	86
Appendix A	88
Medical Doctor's Comments	88
Appendix B	90
Create the Cascade of Haar-Like Classifier	90
B.1. Explanation of the Training Stages	90
B.1.1. Gathering Image Database	90
B.1.2. Arranging Negative Images	90
B.1.3. Crop the Positive Objects	91
B.1.4. Create the Vector of Positive Samples	91
B.1.5. Training	91
References.....	92



List of Figures

Figure 1. The block Diagram of the Multimedia System for BSE performance	4
Figure 2. The Types of Breast Cancer: a) Localized b) Regional c) Distant [21]	6
Figure 3. BSE Tactile Examination: (a) Vertical strip method (b) Palpation pressure levels	11
Figure 4. Hand motion segmentation result [38]	12
Figure 5. Left: R-HOG block. Right: C-HOG block [49]	23
Figure 6. The processing stage for HOG feature extraction [49]	24
Figure 7. The sample of haar-like feature for face detection [70]	38
Figure 8. The samples of Haar-Like Features	39
Figure 9. The shaded area shows the breast region for BSE performance	44
Figure 10. The System of BSE Performance Evaluation	44
Figure 11. The Block Diagram of the Algorithm for BSE Evaluation	46
Figure 12. The Pseudo code of the Integrated Algorithm for BSE Evaluation	47
Figure 13. The Block Diagram of Nipple Detection Algorithm	48
Figure 14. The Haar-Like Features for using in AdaBoost	49
Figure 15. The Form of the Cascade	51
Figure 16. a, b) The structure of the ROI windows c) The sum of the pixels in 5th square must be less than others for left-breast detection d) the sum of the pixels in 2nd square must be less than others for right-breast detection	53
Figure 17. The Structure of the Neural Network for Nipple Detection	55
Figure 18. The Samples of Breast Regions Identification	57
Figure 19. the block diagram of the breast region identification	58
Figure 20. The Detected Breasts with Starting and Ending Points	59
Figure 21. The Flowchart of the palpation detection algorithm	61



Figure 22. a) The Sample of BSE Performance. b) The 2nd and 3rd blocks were highlighted after palpation	64
Figure 23. The Pseudocode of the third objective.....	65
Figure 24. Dividing the Output Matrix of Optical Flow into Four Equal Tiles	66
Figure 25. The Score of Palpation Process with Respect to Time for Each Tile of the Block	66
Figure 26. The Detection of Non-Palpated Tiles in the Block	67
Figure 27. The Sample of Positive Samples in Dataset-1.....	68
Figure 28. The Positive Samples of Dataset-2 for Training and Testing the Neural Network	69
Figure 29. The Samples of Frames in Dataset-3.....	70
Figure 30. The Samples of Blocks of the Videos in Dataset-4. a, b) First and Second Extracted Positive Blocks from Videos. C,d) First and Second Extracted Negative Blocks from Videos.....	71
Figure 31. Cross validation- Holdout method.....	72
Figure 32. The Accuracy of Thresholding Values for Palpation Detection	78
Figure 33. The Examples of the Results for Breast Detection	83
Figure 34. The Examples of the Results for Nipple Detection	83
Figure 35. The Examples of Breast Region Tracking Based on Tracking the Opposite Nipple. a, b) Right Breast Region Tracking in the Actual Tests C, d) Left Breast Region Tracking in the Actual Tests	84
Figure 36. The Examples of Palpated Blocks in the Actual Tests.....	84



List of Tables

Table 1. Algorithm Summary of AdaBoost.....	50
Table 2. Confusion Matrix for a Binary Classification.....	72
Table 3. Training, Test and Validation Set for Breast Detection	74
Table 4. The Resultant Confusion Matrix for Breast Detection	74
Table 5. Performance of Breast Detection Classifier	75
Table 6. Training, Test and Validation Set for Nipple Detection	75
Table 7. The Resultant Confusion Matrix for Nipple Detection	75
Table 8. Performance of Nipple Detection Classifier.....	76
Table 9. The Evaluation Nipple Tracking Using NSDTM.....	76
Table 10. The Evaluation Nipple Tracking Using NCTM	77
Table 11. The Evaluation Nipple Tracking Using NCCTM.....	77
Table 12. The Evaluation of Thresholding Values of NormL2 for Palpation Detection.....	78
Table 13. The Resultant Confusion Matrix for Palpation Detection	79
Table 14. Performance of Palpation Detection Evaluation.....	79
Table 15. The Resultant Confusion Matrix for the Integrated Algorithm	80
Table 16. The Notebook's Characteristics for the Actual Tests.....	82



Nomenclature

ANN: Artificial Neural Network

BSE: Breast Self-Examination

CBE: Clinical Breast Examination

Fps: Frames per second

HOG: Histogram of Oriented Gradients

MRI: Magnetic Resonance Imaging

ROI: Region of Interest

RGB: Red, Green, and Blue Color Model

HSV: Hue, Saturation, and Value Color Model

IDE: Integrated Development Environment

LAHE: Local area histogram equalization

BSD: Berkeley Software Distribution

DSP: Digital Signal Processing



Chapter 1 Introduction

1.1. Background of the Study

Breast cancer is the major cause of death from cancer among females worldwide [1, 2]. In 2010, approximately 1.4 million females were diagnosed with breast cancer worldwide with corresponding 460000 deaths [2]. In the Philippines, the approximate number of breast cancer incidence was 11524, and the approximate number of mortality cases was 4085 among women in 2008 [15]. Studies have proven that in general, mortality rate of breast cancer can be decreased by early detection and treatment of the cancer [3, 4]. Thus, proper education and appropriate breast cancer screening methods which are both effective and accessible to the general population are essential to help solve this universal problem. Breast cancer screening methods are mainly performed through breast imaging technologies, such as mammography, ultrasonic, and visual/palpation-based physical inspections, which includes clinical breast examination (CBE) and breast self-examination (BSE) [5].

BSE is a way for women to check their breasts for detecting lumps and abnormalities by themselves. Some studies propose that majority of women lack the confidence both in their breast examination skill and understanding of results [6]. The aforementioned reasons lead to doubts on the effectiveness and benefits of BSE in breast cancer screening. Despite the controversy, BSE remains important for breast cancer screening [7], [8], [53]. BSE is appealing as a routine screening method because the examination has no financial cost, can be conducted in private, and it is theorized that a woman is highly sensitive to subtle changes in her own breast tissue [9], [10]. This is ideal for breast cancer consciousness and screening in resource-limited areas of developing countries where women cannot



afford annual medical examination (CBE) and mammography [11]. Furthermore, BSE has been shown to detect new breast cancers in high-risk women subjected to regular CBE and mammography [12]. Additionally, mammography is not accurate on dense breast tissues common to younger women ages below 40 years; hence, an alternative method such as BSE is valuable. Moreover, there is a belief that between women who perform BSE, those who develop breast cancer are more likely to find it at an earlier stage and this is expected to lead to earlier treatment and hence decrease their risk of dying from the cancer [9]. Regardless of possible benefits of BSE, only 14–40% of women report performing BSE on a monthly basis and correct way [13], [14]. Therefore, women performing BSE should receive proper training/instructions and have their technique supervised and regularly reviewed. In other words, the development of an application that can guide and evaluate a BSE performance objectively, or at least encourage breast awareness, is very important.

1.2. Statement of the Problem

It has been known that early detection of the breast cancer coupled with proper treatment increases the rate of survival [16] while in [17]; the researchers suggest that the majority of breast cancers could be first detected during BSE. BSE is one of the effective screening methods for detecting the breast cancer in early stages, and it has to be performed in the correct way with covering the entire breast region. Majority of women do not perform BSE due to the lack of confidence and knowledge about the BSE performance. The promotion of knowledge about breast cancer and BSE in females can be done using a multimedia system [18], [19]. The video analysis algorithm, which is developed in this research, is the most important part of the multimedia application in which the BSE performance is evaluated in real time.



1.3. Significance of the Study

Unfortunately, many women are reluctant to perform the BSE due to variation, inconsistencies techniques, and the lack of confidence when performing BSE. Women must be encouraged to perform BSE regularly and correctly, and a computer vision-based algorithm in a multimedia system has the ability to educate females in performance of the BSE, and evaluate the BSE's correctness in real-time. Currently, fully interactive systems do not exist and women depend on their subjective feeling while performing BSE procedures. So, there is a need to develop a computational intelligence system incorporating pattern recognition, computer vision techniques, and provide real-time feedback and objective data to educate women in performing BSE and evaluate the BSE performance in real-time.

1.4. Objectives of the Study

The general objective of the study is to design and develop the computer vision algorithm to evaluate the BSE performance in terms of covering and checking the entire breast region.

The specific objectives of the study are the following:

1. To utilize Adaboost, Haar-Like features, HOG, Neural Network, and template matching for detecting and tracking the nipples in real time.
2. To design and implement an algorithm for localizing and tracking of breast regions and blocks of the pixels for making palpation in a BSE training.
3. To utilize Farneback optical flow to create a method for detecting the palpated blocks of the pixels in the breast regions.
4. To integrate algorithms mentioned in specific objectives 1, 2, and 3 to run in real-time, and monitor the final evaluated results.



1.5. Scope and Limitations of the Study

The BSE is an appropriate way to detect the breast cancer in early stages. But, it is important to keep in mind that BSE is a screening procedure, not a cancer diagnostic tool. Thus as much as possible, BSE should be utilized in combination with breast cancer diagnostic tools, and not as a substitute.

The promotion of knowledge about breast cancer and BSE in females can be done using a multimedia system [18], [19], [80], [81], [82]. As shown in figure.1, the multimedia system consists of two main blocks. The first block is created to educate women in breast cancer and BSE, and the second block is made for guidance and evaluation of the BSE's performance in real-time. The first block involves films, slides, texts, animations, and audio to educate women in BSE and the second block involves video analysis to check the BSE's performance in real-time and audio analysis for users to be connected with the system. This research focuses on designing and developing the algorithm for video analysis to evaluate the BSE performance in terms of covering the entire breast regions. The mentioned block, which is developed in this study, is highlighted in the Fig. 1.

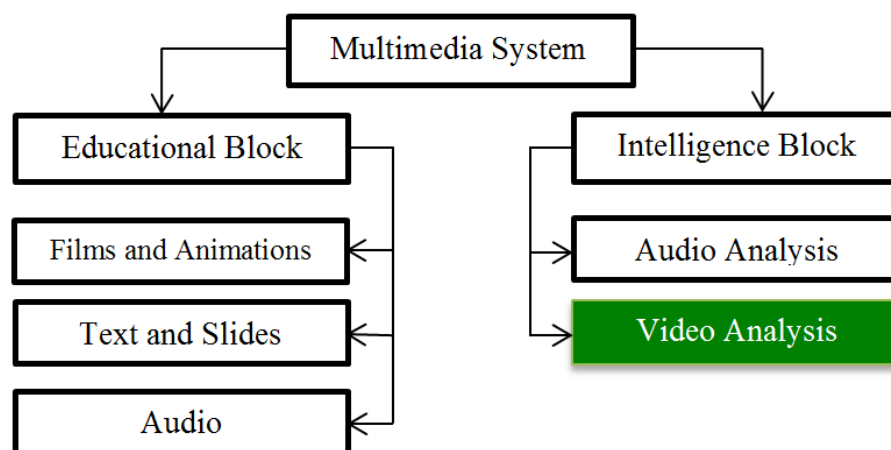


Figure 1. The block Diagram of the Multimedia System for BSE performance



In this thesis, the combination of neural network, integral imaging, cascaded Adaboost algorithm, and HOG was used to solve the problem for the first objective. In the first objective, the cascaded Adaboost is used to detect the left and right breasts. Then, the integral imaging, HOG, color moments, and neural network are used to detect the left and right nipples. The new method for detecting the palpated blocks of the breast region was created using the Farneback optical flow for the third objective.



Chapter 2

Review of Related Literature

2.1. Breast Cancer

Breast cancer is a kind of cancer initiating from breast tissue, most commonly from the inner lining of milk ducts or the lobules that supply the ducts with milk [20]. There are three primary stages of breast cancer: localized, regional and distant as shown in Fig .2. In the localized stage, lumps are small (less than 2 cm), and the cancer has not spread outside the breast. Even if breast cancer is found in several different locations in the breast, it is still considered local. In the regional stage, lumps are medium sized (2 cm to 5 cm), and the cancer has spread to the lymph nodes in the armpit area. In the distant stage, lumps are large (greater than 5 cm), and the cancer has spread to other parts of the body.

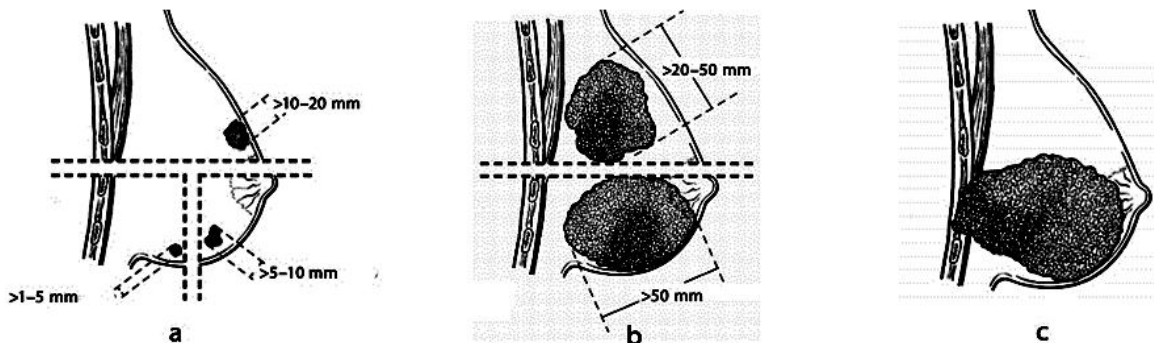


Figure 2. The Types of Breast Cancer: a) Localized b) Regional c) Distant [21]

The five-year relative survival rate from cancer is highly dependent on the stage in which it is detected. For detection in the distant stage, survival rate is as low as 15%. The regional stage survival rate is around 67%. Finally, survival from breast cancer can be as high as 93% for detection in the localized stage [22], [23].



Therefore, early detection of the abnormalities in the breast is the key contributor to the successful treatment of this potentially fatal disease.

2.2. Breast Cancer Screening

There are a number of screening methods to help in the detection of breast cancer, including Mammogram, Ultrasound, MRI, CBE, and BSE.

2.2.1. Mammogram

A mammogram is an x-ray of the breast that can reveal lumps and abnormalities. Mammography is good at finding breast cancer, especially in women ages 50 and older [24]. Generally, the sensitivity of mammography is about 78 percent [24], [25]. This means mammography correctly detects around 78 percent of females who really have breast cancer. Among females over 50, sensitivity is about 83 percent [24], [25]. False negatives are more common in younger women because their breasts are more dense, which makes a mammogram's detection of breast cancer or marks of cancer more challenging [26], [84].

2.2.2. Ultrasound

Ultrasound uses sound waves to make images of the breast. It is non-invasive and often used as a follow-up test after an abnormal finding on a mammogram, breast MRI or clinical breast exam. Studies show ultrasound alone is not a good breast cancer screening tool and has many false positive and false negative results [25], [27]. Ultrasound is better than mammography at discovering tumors within dense breast tissue. Studies have shown mammography joint with ultrasound can detect more breast cancers than mammography alone in females with dense breasts [28]. For women at higher risk of breast cancer, screening ultrasound



does not look to add extra advantage to screening with breast MRI and mammography [29]. However, it may be a useful addition to screening mammography among female at higher risk of breast cancer for whom breast MRI is not yet recommended [28].

2.2.3. Magnetic Resonance Imaging (MRI)

A breast MRI uses magnetic fields to make an image of the breast. It can sometimes detect cancers in dense breasts that are not seen on mammograms. Breast MRI is often used with mammography for screening some women at a high risk of breast cancer. However, it can be expensive and often finds something that looks abnormal, but turns out to be benign (false positive). There is growing evidence that mammography plus breast MRI, compared to mammography alone, can increase detection of breast cancer in Women with a strong family history of breast cancer [30].

2.2.4. Clinical Breast Examination (CBE)

A CBE is a physical exam done by a health care provider as part of the regular medical check-up. CBE can be helpful in finding abnormalities in women under age 40 and for whom mammography is not recommended. In females ages 40 and older, CBE combined with mammography may find more cancers than mammography alone. When used together at yearly check-ups, fewer breast cancers are neglected [24], [31]. Although an important complement to mammography, CBE is not a substitute for mammograms in women 40 and older. Most health organizations mention CBE as a part of regular breast cancer screening. However, there are some weaknesses to its use. One disadvantage is the increased chance of false positive results that require follow-up tests [25], [31]. False positives occur when a CBE detects something that looks or feels like



cancer, but turns out to be benign (not cancer). One large study found the chance of having a false positive consequence after 10 yearly CBE was around 20 percent [32]. Nevertheless, this does not outweigh the life-saving profits of regular screening.

2.2.5. Breast Self-Examination (BSE)

A BSE involves checking the breasts to help detect breast problems or changes. The role of BSE in breast cancer screening is controversial due to the findings of previous studies pointing out its low sensitivity (low detection capability) and specificity (high possibility of a false-positive result) [33]. Furthermore, some studies suggest that most women lack the confidence both in their breast examination skill and interpretation of results [6]. The aforementioned reasons lead to doubts on the effectiveness and benefits of BSE in breast cancer screening [34]. Despite the controversy, BSE remain significant for breast cancer screening [7], [8]. BSE has benefits of being simple, non-hazardous, inexpensive, and a non-invasive procedure. This is ideal for breast cancer consciousness and screening in resource-limited areas of developing countries where women cannot afford annual medical examination (CBE) and mammography [11]. Furthermore, BSE has been shown to detect new breast cancers in high-risk women subjected to regular CBE and mammography [12]. Additionally, mammography is not accurate on dense breast tissues common to younger women ages below 40 years; hence, an alternative method such as BSE is valuable. Finally, it is hypothesized that a woman would be highly sensitive to changes in her own breast tissue.

BSE consists of two stages: visual and tactile examination of the breast [35]. The visual examination searches for abnormal changes in the breast such as dimpling of the skin, 'orange peel' appearance, changes in breast contour,



inversion of the nipple, and any change in the size, shape or symmetry. The tactile examination checks for any new lump in the breast or the armpit area. It is performed with the finger pads of the three middle fingers while circular motion is applied at each site. For BSE to be effective, the total breast area should be covered while incorporating three levels of palpation pressure at each portion- superficial: low, medium and deep. A superficial or gentle palpation disturbs only the portion just beneath the skin surface. A medium (mid-level) palpation examines the middle portion of the breast tissue. Finally, a deep palpation requires enough pressure to reach the breast tissue down to the chest wall. These three palpation pressure levels are necessary as new lumps can be located at any level in the breast. The area thoroughness measures the totality of examination in terms of coverage area – from the armpit region, down to the lower part of the breast tissue, and upward towards the collarbone as shown in Fig. 3.

As suggested by [36], the proficient BSE technique can be assessed with the following eight components:

- BSE should consist of a visual examination
- The middle three fingers should be utilized
- Finger pads should be the surface used for the tactile exam
- Search pattern used should not be random
- Palpation technique should exhibit small circular motion
- All or most of the breast area should be covered during the examination
- Examination of the armpit region should be included
- BSE frequency should be performed monthly, i.e. 12 or more times per

year

The examination can be performed in numerous patterns but the vertical strip method was shown to have the maximal thoroughness in coverage of the breast area [37] as shown in Fig. 3.

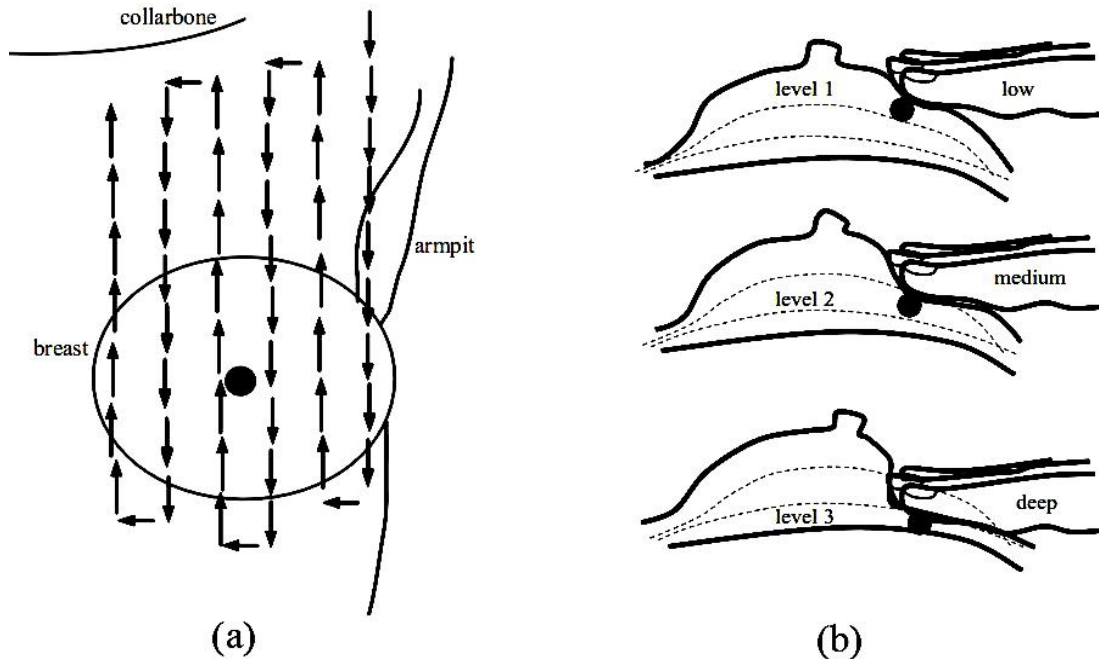


Figure 3. BSE Tactile Examination: (a) Vertical strip method (b) Palpation pressure levels

2.3. Computer Vision Algorithms for BSE

Generally, there is a lack of resources in nipple, and breast region detection for BSE performance. For the hand detection, skin color modeling and classification play important roles in hand detection. A hand is a crucial tool used in BSE that needs to be detected and analyzed during the procedure of BSE. However, the background of a woman's moving hand is her breast that has the same or similar color as the hand. Moreover, color images captured by a web camera are strongly affected by the lighting or brightness conditions. Therefore, it is a challenging task to detect the hand against the breast without using any artificial signs, such as colored nail polish [38].

The hand detection for BSE performance was surveyed by [38]. In that research, a two-dimensional Gaussian skin color model was employed in a



specific way to identify a breast but not a hand. First, an input image is transformed to $YCbCr$ color space, which is less sensitive to the lighting conditions and more tolerant of skin tone. The breast, thus detected by the Gaussian skin model, is used as the baseline or framework for the hand motion. Secondly, motion cues were used to segment the hand motion against the detected baseline as shown in Fig. 4. The proposed method is an appropriate way for hand detection, but there must be a method for breast region identification before performance of hand detection. Other problem of this method is that other movements are considered as the hand. I.e. if the body of the woman is moved, there would be errors in detection the hand.

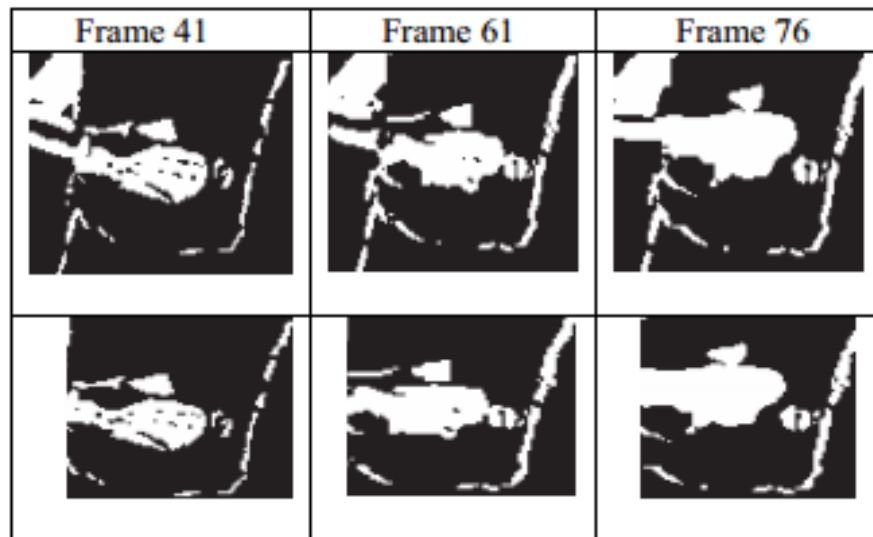


Figure 4. Hand motion segmentation result [38]

A method of estimating the degree of hand pressure was presented in [85]. Pressure is estimated by means of computing image entropy. The method proves to be accurate and the ease of the algorithm produces very fast pressure detection, thus real time pressure estimation can be realized in an interactive multimedia environment. The dynamic model of the breast is close to Hook's rule as verified by the proposed pressure estimation in [85]. In complicated situations, such as



body movement and fingertips hidden under a large breast should be considered in [85]. Aside from that the breast region must be detected and tracked accurately in different luminance conditions in real time before detecting the level of the pressure.

2.4.Object Detection Algorithms

One of the main challenges of computer vision is efficiently detecting and classifying objects in an image or video sequence. Several machine learning approaches have been applied to this problem, demonstrating significant improvements in detection accuracy and speed [95]. The Haar-Like features and Adaboost are considered as a very effective and accurate model for object detection. This technique was used in different researches for recognition a sort of objects.

The Adaboost and Haar-like features were used for hand gesture detection as described in [96]. With the algorithm in [96], real-time performance and high recognition accuracy was obtained. In the paper [96], the higher level implements the linguistic hand gesture recognition using a context-free grammar-based syntactic analysis. Given an input gesture, based on the extracted postures, the composite gestures can be parsed and recognized with a set of primitives and production rules using the Adaboost algorithm [96].

In the paper [97], a new distinctive feature, called joint Haar-like feature, for detecting faces in images was proposed. As described in [97], this was based on co-occurrence of multiple Haar-like features. Feature co-occurrence, which captures the structural similarities within the face class, makes it possible to construct an effective classifier. As described in [97], the joint Haar-like feature can be calculated very fast and has robustness against addition of noise and



change in illumination. In the paper [97], a face detector is learned by stage wise selection of the joint Haar-like features using AdaBoost [97].

The paper [98] describes a vision-based pedestrian detection system for robots, and autonomous vehicles. For that purpose the Haar-like features were used to discriminate pedestrians. Those features were used as input in a learning algorithm, based on AdaBoost, which selects a small number of critical visual features from a larger set and yields an extremely efficient classifier. As indicated in [98], the presented system can run in real-time applications achieving good detection rates.

A fast and robust algorithm for detecting pedestrians in a cluttered scene from a pair of moving cameras was presented in [99]. This is achieved through stereo-based segmentation and neural network-based recognition. The algorithm includes three steps. First, the image is segmented into sub-image object candidates using disparities discontinuity. Second, the sub-image object candidates is merged and divided into sub-images that satisfy pedestrian size and shape constraints. Third, intensity gradients of the candidate sub-images are used as input to a trained neural network for pedestrian recognition [99]. As indicated in [99], the proposed system has high accuracy.

A neural network-based upright frontal face detection system was presented in [100]. As described in [100], a retinal connected neural network examines small windows of an image and decides whether each window contains a face. The system arbitrates between multiple networks to improve performance over a single network. A straightforward procedure for aligning positive face examples for training was proposed in [100]. To collect negative examples, they use a bootstrap algorithm, which adds false detections into the training set as training progresses. This eliminates the difficult task of manually selecting nonface training examples, which must be chosen to span the entire space of nonface



images. Based on the paper [100], simple heuristics, such as using the fact that faces rarely overlap in images, can further improve the accuracy.

The cascade-of-rejecters approach with the Histograms of Oriented Gradients (HOG) features to achieve a fast and accurate human detection system was presented in [101]. The features used in the system are HoGs of variable-size blocks that capture salient features of humans automatically. Using AdaBoost for feature selection, the appropriate set of blocks are identified, from a large set of possible blocks. In the paper [101], the integral image representation and a rejection cascade which significantly speed up the computation was used. For a 320×280 image, the system can process 5 to 30 frames per second depending on the density in which the image is scanned, while maintaining an accuracy level similar to existing methods [101].

While Histograms of Oriented Gradients (HOG) plus Support Vector Machine (SVM) (HOG+SVM) is the most successful human detection algorithm, it is time-consuming [102]. The paper presented in [102] proposes two ways to deal with this problem. One way is to reuse the features in blocks to construct the HOG features for intersecting detection windows [102]. Another way is to utilize sub-cell based interpolation to efficiently compute the HOG features for each block [102]. As indicated in [102], the combination of the two ways results in significant increase in detecting humans—more than five times better. To evaluate the proposed method, a top-view human database have established in [102]. Experimental results on the top-view database have demonstrated the effectiveness and efficiency of the proposed method [102].

A new approach on nipple detection for adult content recognition was presented in [103]. It combines the advantage of Adaboost algorithm that is rapid speed in object detection and the robustness of nipple features for adaptive nipple detection. This method first locates the potential nipple-like region using Adaboost algorithm for fast processing speed. It is followed by a nipple detection



using the information of shape and skin color relation between nipple and non-nipple region. As this method uses the nipple features to conduct the adult image detection, it can achieve more precise detection and avoids other methods that only detect the percentage of exposure skin area to decide whether it is an adult image [103]. The proposed method can be also used for other organ level detection. As described in [103], the experiments show that this method performs well for nipple detection in adult images [103].

The nipple detection algorithm was proposed in [104] using the neural network. Self-Organizing Map (SOM) is an unsupervised neural network using for object detection and recognition in numerous image processing applications [104]. A methodology to achieve an automated detection of nipples by obscene pictures was presented in [104]. The proposed system is composed of a human skin detection using for non-skin removal process based on image processing and a nipple detection using for nipple existence identification based on Self-Organizing Map. The goal is to identify whether the given picture is an obscene picture by detecting the nipple existence in the picture [104]. The proposed system is shown to be effective for a wide range of shades and colors of skin and human configurations. As presented in [104], it is validated for detecting a nipple existence for obscene pictures. In terms of scalability, the nipple detection model can be modified to support distributed processing [104].



Chapter 3

Theoretical Framework

3.1. OpenCV Library

OpenCV is an open source library for designing and developing computer vision applications. It is used in both academic and commercial applications under a BSD license that allows freely using, distributing, and adapting it. It is developed by Intel, and now supported by Willow Garage and Itseez. OpenCV is written in C++ and its primary interface is in C++, but it still retains a less comprehensive though extensive older C interface. Now it has several hundreds of image processing and computer vision algorithms which make developing advanced computer vision applications easy and well-organized [39].

The OpenCV library has more than 2500 adjusted algorithms, which contains a complete set of both typical and advanced computer vision and machine learning algorithms. These algorithms can be used to detect and classify faces, classify objects, classify human movements in videos, track camera movements, track moving items, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to yield a high resolution image of a whole scene, find similar pictures from an image database, remove red eyes from images taken by flash, follow eye movements, and recognize scenery and establish signs to overlay it with augmented reality. The library is used widely in companies, research groups and by governmental groups [39], [40].

3.2. Qt Framework

Qt is a complete Integrated Development Environment (IDE) for C++ application that was first developed by Trolltech, a Norwegian software company



which was developed in 2008 by Nokia. It is offered under the LGPL open source license as well as under a commercial license for the development of copyrighted tasks. It is composed of two distinct components: a cross-platform IDE, called Qt Creator, and a set of Qt class libraries and development tools. Using the Qt Software Development Kit (SDK) to develop C++ applications has several profits [39], [41], [42]:

- It is an open source initiative, developed by the Qt community, which offers contact to the source code of the diverse Qt components.
- It is cross-platform, meaning that user can develop applications that can run on diverse operating systems such as Windows, Linux, Mac OS X.
- It comprises a complete and cross-platform GUI library that tracks an actual object-oriented and event-driven model.
- Qt also includes some cross-platform libraries to progress multimedia, graphics, database, multithreading, web application, and many other interesting building blocks worthwhile for planning advanced applications.

3.3. Pattern Recognition

Pattern recognition is the study of how machines can notice the surroundings, learn to identify/ classify patterns of interest from their background, and make sound and reasonable decisions about the groups of the patterns. Given a pattern, its recognition/classification may consist of one of the following two tasks: 1) supervised classification in which the input pattern is identified as a member of a predefined class, 2) unsupervised classification in which the pattern is assigned to an unknown class. The recognition problem here is being posed as a classification or categorization task, where the classes are either distinct by the system designer (supervised classification) or are learned based on the similarity and match of patterns (unsupervised classification). These applications include data mining



(classifying a “form”, e.g., correlation, or an outlier in millions of multidimensional patterns), document classification (competently searching text documents), financial predicting, organization and retrieval of multimedia databases, and biometrics. The rapidly growing and available computing power, while enabling quicker processing of enormous data sets, has also simplified the use of elaborate and different methods for data analysis and classification. At the same time, demands on automatic pattern recognition systems are growing enormously due to the accessibility of massive databases and stringent performance requirements (speed, accuracy, and cost) [43].

The plan of a pattern recognition system fundamentally involves the following three aspects:

- data acquisition and preprocessing
- data representation
- decision making

The problem area directives the choice of sensor(s), preprocessing method, representation scheme, and the decision making model. It is commonly agreed that a well-defined and adequately constrained recognition problem will lead to a compact pattern representation and a simple choice making strategy. Learning from a set of samples is an imperative and desired attribute of most pattern recognition approaches. The four best known methods for pattern recognition are [43]:

- template matching,
- statistical classification,
- syntactic or structural matching,
- neural networks



3.4. Artificial Neural Network (ANN)

An ANN is an information processing example that is inspired by the way of biological nervous systems, such as the brain, process information. The key element of this model is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements, called neurons, working in union to solve specific problems. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems contains adjustments to the synaptic connections that exist among the neurons [43].

When writing an ANN, this is mimicked by using a "perceptron" as the basic component instead of the neuron. The perceptron can take several weighted inputs and summarize them, and if the combined input exceeds a threshold it will activate and send an output. Which output it sends is determined by the activation function and is often chosen to be between 0 and 1 or 1 and -1. Since the derivative of the activation function is often used in the training of the network, it is suitable if the derivative can be conveyed in terms of the original function value, as few additional calculations are needed to calculate the derivative in this case. The equation for a perceptron can be written as:

$$y = \varphi \left(\sum_{i=1}^n \omega_i x_i + b \right) \quad (3.1)$$

Where y is the output signal, φ is the activation function, n is the number of connections to the perceptron, ω_i is the weight associated with the i th connection and x_i is the value of the i th connection. b signifies the threshold. The threshold b is a neuron with a constant value of -1. By letting the network to adjust the weight associated with b , a dynamic threshold for when the perceptron activates is achieved. This is a very simple design, and its strength can be presented when



several perceptron are combined and work together. The perceptron are often organized in layers, where each layer takes input from the previous, applies weights and then signals to the next layer if suitable. A classifier must be capable to learn from examples and adapt. In an ANN, this is achieved by updating the weights linked with the connections among the layers. There are a number of ways of doing this, and most involve initializing the weights and fed the network an example. The error made by the network at the output is then calculated and feed backwards through a process called "back-propagation". This process is then used to update the weights, and by repeated use of this process, the network can learn to differentiate between several diverse classes. The exact equations involved vary from case to case [43].

One problem with the ANN method is over-fitting of the data, which occurs when the classifier becomes too good at distinguishing the training examples, at the expense of not being able to identify a general input. This can be avoided by cross-validation, where the network is trained on one set of data, and then assessed on a separate one. When the error starts increasing in the validation set, the network might be over-fitted. If previous networks are saved, the network can then be rolled back to the one which gave the smallest error [44].

ANNs have been employed for pattern classification and recognition in different applications [43]. Multi-Layer Perceptron Neural Network is among the best neural networks in operating classification problems. These networks consist of 3 layers in most cases, input, hidden and output layer. Generally, ANN has a number of activation functions, training methods, and variations. The activation function of hidden and output layers has to be adapted in accord with the problem [43].

The main features of neural networks are that they have the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data. The most common family of neural



networks for pattern classification tasks is the feed-forward network, which includes multilayer perceptron and Radial-Basis Function (RBF) networks [45]. Another popular network is the Self-Organizing Map (SOM) [46], which is mostly used for data clustering and feature mapping. The learning process includes updating network architecture and connection weights so that a network can professionally perform a specific classification/clustering task. The increasing reputation of neural network models to solve pattern recognition problems has been mainly due to their seemingly low dependence on domain-specific knowledge and obtainability of effective learning algorithms for experts to use. ANNs provide a new suite of nonlinear algorithms for feature extraction (using hidden layers) and classification (e.g., multilayer perceptron). In addition, current feature extraction and classification algorithms can also be mapped on neural network architectures for effective implementation [43].

3.5. Histogram of Oriented Gradients (HOG)

Histogram of oriented gradients (HOG) is used as feature descriptors for the purpose of object detection, where the incidences of gradient orientation in localized parts of an image play significant roles. This technique is similar to scale-invariant feature transform (SIFT) but differs in that it operates on a dense grid of uniformly spaced cells and uses local contrast normalization on overlapping blocks for enhanced precision. HOG feature descriptors are first described by [47] where great achievement has been gotten on pedestrian detection in both images and videos, as well as on a variety of common animals and vehicles in motionless pictures. The basic clue behind HOG is that the appearances and figures of local objects within an image can be well defined by the distribution of intensity gradients as the votes for dominant edge directions. Such feature descriptor is attained by first dividing the image into small



connecting areas of equal size, called cells, then collecting a histogram of gradient directions for the pixels within each cell, and at last bonding all these histograms. In order to advance the detection correctness against diverse illumination and shadowing, local contrast normalization is applied by computing a measure of the intensity across a bigger region of the image, called a block, and using the resultant value to normalize all cells within the block. Thus, the HOG feature descriptor holds some significant benefits over other techniques. For one thing, this method results in important invariance to geometric and photometric transformations since it operates on localized cells so that those changes would only appear in larger spatial areas. For another, according [47], HOG features are more tolerant to the separate body movement of pedestrians as long as they have an upright position. So, this descriptor is particularly appropriate for human detection. In fact, there are four variants of HOG block scheme as shown in Fig. 5: Rectangular HOG (R-HOG), Circular HOG (C-HOG), Bar HOG and Center-Surround[48], among which R-HOG is the most popular.

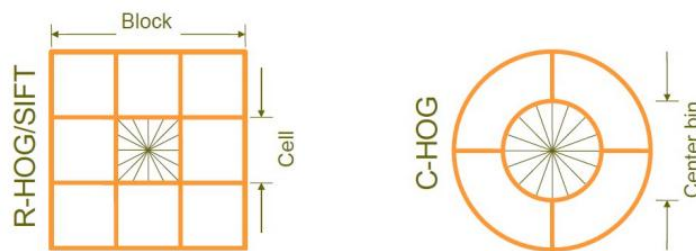


Figure 5. Left: R-HOG block. Right: C-HOG block [49]

3.5.1. Gradient Computation

The first stage in HOG feature extraction is the calculation of image gradients. This is prepared by applying the 1D centered, point discrete derivative mask in



both the horizontal and vertical directions, which in specific are filter kernels of the following form:

$$[1; 0; 1] \text{ and } [1; 0; 1]^T \quad (3.2)$$

In fact, there are numerous complex masks, such as Sobel, Prewitt, Canny or diagonal masks, but these masks generally result in poorer performance [47]. The magnitude and orientation at each pixel $I(x; y)$ is calculated by

$$\begin{cases} G_{mag}(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \\ \theta(x, y) = \arctan\left(\frac{G_y(x, y)}{G_x(x, y)} + \pi/2\right) \end{cases} \quad (3.3)$$

where $G_x(x; y)$ and $G_y(x; y)$ are the gradient values at each pixel in horizontal and vertical direction, respectively. For color images, the channel with the largest magnitude gives the pixel's dominant magnitude and orientation. It should be noted that $\pi/2$ is necessary since the \arctan operator results in a range between $\pi/2$ and $-\pi/2$, but for unsigned orientation scheme which gives better performance, it ranges from 0 to π . The processing stage for HOG feature extraction is shown in Fig. 6.

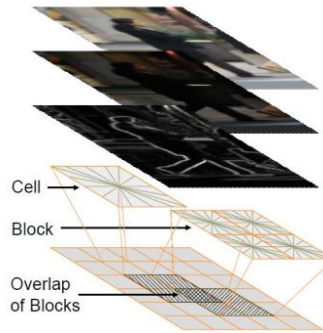


Figure 6. The processing stage for HOG feature extraction [49]



3.5.2. Orientation Binning

In this process, the histograms for each cell are formed. The cells are pixel areas that are either rectangular or radial in figure, and the histogram bins are expanded from 0 to 180 (or from 0 to 360 in the event of signed orientation), so every histogram bin has a spread of 20. Every pixel in the cell casts a weighted voting into one of the 9 histogram bins that its orientation belongs to. As for the weight of votes, it can either be the gradient magnitude itself, or some function of the magnitude, for example, the square root or square of the gradient magnitude, or some clipped version of the magnitude. Generally, the gradient magnitude is directly used.

3.5.3. Descriptor Blocks

To achieve the robustness against diverse illumination and contrast, the gradient strengths should be locally normalized. This leads to combination the cells into larger pixel areas called blocks. These blocks overlap with neighboring blocks, so that each cell can contribute its orientation distribution more than once as show in Fig. 6. Totally 4 block geometries exist, with one most normally used: Rectangular HOG blocks. R-HOG blocks are usually square grids, and the optimal parameters are found to be 22 cell blocks of 88 pixel cells or 33 cell blocks of 66 pixel cells. Furthermore, a Gaussian spatial window is applied to all block before histogram voting so that the weight of every pixel around the edge of the block can be considerably suppressed.

3.5.4. Block Normalization

There are three different ways recommended to normalize the blocks. Let v indicates the non-normalized feature vector that saves all cell histograms from a



given block, $\|v\|_k$ denote its k -norm for $k = 1, 2$ and eps denote certain small constant. Then the normalization structures have the following forms:

$$\text{L2-norm: } \hat{v} = \frac{v}{\sqrt{\|v\|_2^2 + eps^2}} \quad (3, 4)$$

$$\text{L1-norm: } \hat{v} = \frac{v}{(\|v\|_1 + eps)} \quad (3, 5)$$

$$\text{L1-sqrt: } \hat{v} = \sqrt{\frac{v}{(\|v\|_1 + eps)}} \quad (3, 6)$$

All the normalization schemes offer much better performance than non-normalized case. The final HOG feature descriptor is then the vector comprising the elements of the normalized cell histograms from all of the block areas.

3.6. Color Features

3.6.1. The RGB Color Model

The RGB color model is an additive color model in which the main colors red, green, and blue light are added together in different ways to reproduce a broad array of colors. The name comes from the initials of the three colors Red, Green, and Blue. Each of the three beams is called a component of that color, and each can have subjective intensity, from completely off to completely on, in the mixture [50]. Zero intensity for each component gives the darkest color (no light, considered the black), and full intensity of each gives a white. A color in the RGB color model is defined by indicating how much of each of the red, green, and blue is comprised in each component which can vary from zero to a defined maximum value which depends of the application. In computing, the component values are



often stored as integer numbers in the range 0 to 255. The chief purpose of the RGB color model is for sensing, representation, and display of pictures in electronic systems, such as televisions and computers [50].

3.6.2. HSV Color Model

HSV color model (Hue, Saturation, and Value) is a non-linear transformation of the RGB color model, and the colors are the mixture of the three values: the Hue (H), Saturation or color quantity (S), and itself value (V) [51].

Hue: The form of color (e.g. red, green, or yellow). These are represented as a degree of angle whose possible values range from 0° to 360° (although for some applications are normalized from 0 to 100%). Saturation: Is represented as a distance from the axis of the black-white glow. The possible values range from 0 to 100%. Value: Represents the height in the black-white axis. The possible values range from 0 to 100%. 0 is always black. Depending on the saturation, 100 could be white or a more or less saturated color. The transformation of an input image in RGB color model is achieved using the following expressions:

$$H = \arccos \frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{[(R-G)^2+(R-B)(G-B)]}} \quad (3, 7)$$

$$S = 1 - 3 \frac{\min(R,G,B)}{R+G+B} \quad (3, 8)$$

$$V = 1/3 (R + G + B) \quad (3, 9)$$

3.6.3. Color Feature Extraction

The color features can be considered by the color moments and most color distribution information is focused in the lower-order moments, only the first moment (mean), the second moment (variance) and the third moment (skewness) are taken as features [51]. These features are defined as follows:



$$\mu = \frac{1}{N} \sum_i \sum_j p_{ij} \quad (3, 10)$$

$$\sigma = \left[\frac{1}{N} \sum_i \sum_j (p_{ij} - \mu)^2 \right]^{\frac{1}{2}} \quad (3, 11)$$

$$S = \left[\frac{1}{N} \sum_i \sum_j (p_{ij} - \mu)^3 \right]^{\frac{1}{3}} \quad (3, 12)$$

where N is the total number of the pixels in the image, and p_{ij} is the image pixel in location (i, j) .

3.7. Gaussian Smoothing

A Gaussian smoothing is the effect of blurring a frame by a Gaussian function. It is a broadly used effect in graphics applications, usually to diminish image noise and decrease detail. The visual result of this smoothing method is a smooth blur approaching that of viewing the image through a translucent screen, definitely different from the effect made by an out-of-focus lens or the shadow of an object under standard illumination. Gaussian smoothing is also used as a pre-processing phase in computer vision algorithms in order to modify image constructions at diverse scales implementation [54], [55], [56].

Mathematically, applying a Gaussian Filter to a frame is the similar as convolving the frame with a Gaussian function. Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the result of decreasing the image's high-frequency constituents; therefore, a Gaussian blur is a low pass filter [54], [55].

The Gaussian blur is a kind of image-blurring filters that uses a Gaussian function for computing the transformation to apply to each pixel in the frame. The equation of a Gaussian function in one dimension is [54], [55]:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (3, 13)$$



In two dimensions, the Gaussian Filter is the product of two such Gaussians, one in every dimension [54], [55]:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3, 14)$$

where x , is the distance from the starting point in the horizontal axis, y is the distance from the starting point in the vertical axis, and σ is the standard deviation of the Gaussian distribution. When applied in two dimensions, this formula product a surface whose contours are concentric circles with a Gaussian distribution from the center point. Values from this distribution are used to form a convolution matrix which is applied to the original frame. Each pixel's novel value is set to a weighted average of that pixel's neighborhood. The original pixel's value receives the weightiest weight and neighboring pixels receive slighter weights as their distance to the original pixel growths. This results in a blur that keeps borders and edges better than other smoothing approaches [56].

3.8. Template Matching

Template matching function in OpenCV is not based on histograms; rather, the function matches an actual image patch against an input image by “sliding” the patch over the input image using one of the matching methods. Template Matching includes systematically moving a pattern of an object on a frame to be recognized and tracked and recording the positions of any matches. Similarity measure is computed between them to determine the maximum match. If there is a demand on sub pixel accuracy then interpolation of the matching method output values are to be used. This method is applied when there is a slight rotation and scaling [57].

The first problem that comes with this method is its unreliability. Due to the variability of both the original objects and images, exact matches cannot be



estimated, so some moderation of the recognition condition is important. Variability takes the following forms [57], [58]:

- The objects may have slightly different sizes and forms, and in general three dimension sizes and shapes will have to be considered.
- They may have surfaces that have varying roughness, brightness and color.
- They may have gross defects in size, form, roughness, and color.
- They may appear at different distances and orientations in both 2D and 3D.
- They might be viewed in different lighting conditions, and lighting conditions may vary.
- There will be digital quantization effects, due to the particular spatial, grey-scale or color resolution that is used.
- There might be added digital noise.
- There might be lens distortion or blurring.

Due to the mentioned reasons, the nipple must be detected and captured in the algorithm. Then, the captured nipple can be used in nipple tracking algorithm using template matching. In this way, the current nipple is used for tracking instead of using a copied model of nipple.

3.8.1. Matching Functions

1. Square difference matching method:

These methods match the squared variance, so a flawless match will be 0 and wicked matches will be large [57]:

$$F_{sqdiff}(x, y) = \sum_{x', y'} [T(x', y') - i(x + x', y + y')]^2 \quad (3, 15)$$

2. Correlation matching methods [57], [58], [59]:

These methods multiplicatively match the template against the image, so a flawless match will be big and wicked matches will be slight or 0.



$$F_{ccorr}(x, y) = \sum_{x', y'} [T(x', y') \cdot i(x + x', y + y')]^2 \quad (3, 16)$$

3. Correlation coefficient matching methods [57]:

These methods match a template relative to its mean against the image relative to its mean, so a flawless match will be 1 and a perfect mismatch will be -1; a value of 0 simply means that there is no correlation.

$$F_{ccoeff}(x, y) = \sum_{x', y'} [T'(x', y') \cdot i'(x + x', y + y')]^2 \quad (3, 17)$$

$$T'(x', y') = T(x', y') - \frac{1}{(w.h) \sum_{x'', y''} T(x'', y'')} \quad (3, 18)$$

$$i'(x + x', y + y') = i(x + x', y + y') - \frac{1}{(w.h) \sum_{x'', y''} i(x + x'', y + y'')} \quad (3, 19)$$

3.8.2. Normalized Methods

For each of the three methods just described, there are also normalized forms first developed by Galton as described by [60]. The normalized methods are beneficial because, they can help decline the effects of lighting differences between the template and the image. In each case, the normalization coefficient is the same:

$$N(x, y) = \sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} i(x + x', y + y')^2} \quad (3, 20)$$

The values for method that give the normalized computations are listed here:

1. Normed SQDIFF:

$$F_{sqdiff_normed}(x, y) = \frac{F_{sqdiff}(x, y)}{N(x, y)} \quad (3, 21)$$

2. Normed CCORE:

$$F_{ccor_normed}(x, y) = \frac{F_{ccor}(x, y)}{N(x, y)} \quad (3, 22)$$

3. Normed CCOEF:

$$F_{ccoeff_normed}(x, y) = \frac{F_{ccoeff}(x, y)}{N(x, y)} \quad (3, 23)$$



As typical, more accurate matches are obtained (at the charge of more computations) as it moves from simpler measures (square difference) to the more sophisticated ones (correlation coefficient). It's best to do some test trials of all these settings and then choose the one that best trades off accuracy for speed in the application [57].

A good match should have good matches nearby, because slight misalignments of the template shouldn't vary the results too much for real matches. Looking for the best matching can be done by slightly smoothing the result image before seeking the maximum (for correlation or correlation-coefficient) or minimum (for square-difference) matching methods [57].

3.9. Farneback Optical Flow

The idea of optical flow was presented to the field of computer science in the early 1980's [61] and since then it is continuously being advanced and improved. Optical flow represents motion in a form of a vector field that allows transforming one image from the video sequence into the next one, by moving the blocks of the first image in the direction indicated by the components of the vector field. Furthermore, the length of each vector describes the speed observed at the given block. In other words, optical flow is a set of translations describing the motion. In many books and papers the results obtained with this technique are often illustrated by very idealistic motion fields [61], [62], [63], and [64]. There are many methods to the calculation of optical flow but the goal is always the same: for every pixel the velocity vector is calculated [64]. In optical flow, a pixel $F(x, y, t)$ is in the first frame, and it moves by distance (d_x, d_y) in next frame taken after dt time. So, since those pixels are the same and intensity does not change, this formula is described [64]:

$$F(x, y, t) = F(x + dx, y + dy, t + dt) \quad (3, 24)$$



Then, the Taylor series approximation is computed of right-hand side; common terms are removed, and divided by dt to achieve the following equation:

$$f_x u + f_y v + f_t = 0 \quad (3, 25)$$

where:

$$f_x = \frac{\partial f}{\partial x}; f_y = \frac{\partial f}{\partial y}; u = \frac{dx}{dt}; v = \frac{dy}{dt} \quad (3, 26)$$

The above formula contains the information about the speed and the direction in which the pixel is moving. Originally, optical flow algorithms were using the gradient-based estimation. Unfortunately, results that they provide are satisfactory only for certain kinds of motion [62]. Therefore, the palpation detection procedure calculates optical flow with the tensor-based algorithm, proposed by Farnebäck in [63]. The Farneback is a dense optical flow algorithm. The word dense means the motion is calculated for every pixel in the image. This is usually costly, but Farneback's method is linear which is easy to solve. Farneback method uses Polynomial Expansion to approximate the neighbors of a pixel. The Expansion could be seen as a quadratic equation with Matrices and Vectors as variable and coefficients. This dense optical flow analysis yields a motion field from two successive video frames [63], [65].

In summary, Farnebäck proposes an algorithm for estimating dense optical flow based on modeling the neighborhoods of every pixel by quadratic polynomials. The notion is to signify the image signal in the neighborhood of each pixel by a 3D surface, and determine optical flow by finding where the surface has moved in the next frame. The optimization is not done on a pixel-level, but rather on a neighborhood-level, so that the optimum displacement is found both for the pixel and its neighbors [63], [65].



Motion estimation algorithms always contain a trade-off among speed and precision. The method presented by [63] and [65] is primarily intended to be precise but some compromises have been prepared in order to keep it fast as well.

The [65] presented a novel two-frame motion estimation algorithm and evaluation of the mentioned method in compare with other optical flow algorithms. The stated solution, in that paper, is to estimate the motion, or displacement, field from only two frames and try to compensate for the background motion. The [65] presents a novel method to estimate displacement. It is related to the orientation tensor methods in the first processing step, and a signal transform called polynomial expansion.

The notion of polynomial expansion is to approximate several neighborhood of each pixel with a polynomial. The stated polynomial in [65] is:

$$f(x) \sim x^T A x + b^T x + c \quad (3, 27)$$

where A is a symmetric matrix, b is a vector, and c is a scalar. The coefficients are estimated from a weighted least squares fit to the signal values in the neighborhood.

The weighting has two components called certainty and applicability. These terms are the same as in normalized convolution [66], which polynomial expansion is based on. While this may sound computationally very demanding it turns out that it can be executed efficiently by a hierarchical scheme of separable convolutions [67].

3.9.1. Displacement Estimation in Farneback Optical Flow

Based on the [65], since the result of polynomial expansion is that each neighborhood is approximated by a polynomial, the polynomial action on ideal translation was analyzed. Consider the exact quadratic polynomial:



$$f_1(x) = x^T A_1 x + b_1^T x + C_1 \quad (3, 28)$$

A new signal f_2 is made by a global displacement by d [65]:

$$\begin{aligned} f_2(x) &= f_1(x - d) = (x - d)^T A_1 (x - d) + b_1^T (x - d) + C_1 \\ &= x^T A_2 x + b_2^T x + C_2 \end{aligned} \quad (3, 29)$$

Equating the coefficients in the quadratic polynomials yields [65]:

$$A_2 = A_1, \quad b_2 = b_1 - 2A_1 d, \quad C_2 = d^T A_1 d - b_1^T d + C_1 \quad (3, 30)$$

Based on the [65], this observation holds for any signal dimensionality:

$$2A_1 d = -(b_2 - b_1), \quad d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \quad (3, 31)$$

3.10. Histogram Equalization

This section of the chapter describes a method of image processing that allows images to have better contrast. This is achieved via the histogram of the image, using a method that permits the areas with low contrast to gain higher contrast by spreading out the most frequent intensity values.

3.10.1. Image Histogram

Generally, a histogram is the approximation of the likelihood distribution of a specific type of data. A histogram in an image, is a kind of histogram that offers a graphical representation of the tonal distribution of the pixel values in a digital image. The image's histogram can be used for analyzing the frequency of appearance of the different pixel levels contained in the image [68].

A perfect histogram is that which covers all the possible values in the pixel scale or gray scale used. This kind of histogram proposes that the frame has perfect contrast and those details and particulars in the frame may be detected in the easiest way [68].



3.10.2. Improving the Contrast

As described earlier, generally, the histogram equalization spreads out intensity values along the total range of values in order to gain higher contrast. This method is particularly useful when an image is represented by close contrast values, such as images in which both the background and foreground are bright at the same time, or else both are dark at the same time [69].

There are a number of diverse kinds of histogram equalization algorithms. Here is a list of different histogram equalization techniques [68]:

- ***Histogram expansion:***

Advantages: Simple and improve contrasts of an image.

Disadvantages: If there are gray values (0-255) that are physically far apart from each other in the image, then this method fails.

- ***Local area histogram equalization:***

Advantages: Offers a perfect enhancement of image contrast.

Disadvantages: Computationally very slow, needs a high number of processes per pixel.

- ***Cumulative histogram equalization:***

Advantages: Has good performance in histogram equalization.

Disadvantages: Requires a few more operations because it is necessary to create the cumulative histogram.

- ***Par sectioning:***

Advantages: Easy to perform.

Disadvantages: Better suited to hardware performances.

- ***Odd sectioning:***

Advantages: Offers good image contrast.

Disadvantages: Has problems with histograms which cover almost the full gray scale.



These methods were studied and compared in order to determine which one offers the finest equalization and is also finest suited to image processing implementation. Based on the literature, [68], cumulative histogram equalization is proposed for implementation in the DSP. This algorithm was carefully chosen due to its good performance and easy performance in the C language.

The cumulative distribution function is [68]:

$$cdf(x) = \sum_{t=1}^x h(t) \quad (3, 32)$$

where x is a pixel or gray value and h is the frame's histogram. The cumulative distribution function for each gray tone is calculated.

The general histogram equalization formula is [68], [69]:

$$eh(i) = round \left(\frac{cdf(1) - cdf_{min}}{U.V - cdf_{min}} * (L - 1) \right) \quad (3, 33)$$

where the minimum value of the cumulative distribution function is cdf_{min} ; U and V are the frame's number of columns and rows, and L is the number of gray or pixel levels.

3.11. Haar-like Feature

Haar-like features [70], [71] are reminiscent of Haar basis functions which can be used to describe the local gray scale distribution. [72] Introduce an extended set of Haar-like features, which include edge features, line features and center-surrounded features.

Haar-like features is used in object recognition. These kinds of features get the name due to their similarity with Haar wavelets [70]. The initial type of this feature set is discussed in [73] which is based on Haar wavelets. Later, this idea of using Haar wavelets was adapted in [70] and was further developed into the so called Haar-like features. In general, Haar-like feature sums up the intensity



values of the pixels in each rectangular region all inside a subsection of an image and computes the difference between the sums. The difference can be then used to identify subsections of the image. For example, among all human faces the region of eyes is darker than the region of cheeks. As a result, a Haar-like feature for human faces can be obtained by placing two adjacent rectangles over the eye and the cheek region as shown in Fig. 7.

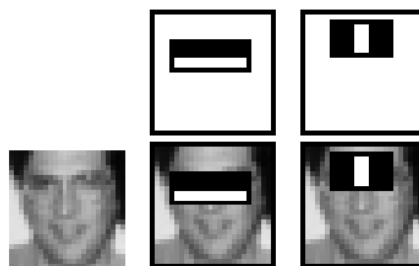


Figure 7. The sample of haar-like feature for face detection [70]

According to [70], a bounding box of proper size slides over the input image, and for each subsection of the image, Haar-like feature is calculated. Each time, the difference is compared to a learned threshold for object / non-object decision. However, this kind of classifier works only slightly better than random guessing, so an ensemble scheme is used to group a large number of Haar-like features in a classifier cascade to jointly create a strong classifier which can define an object with adequate accuracy. The major advantage of Haar-like features is the fast calculation speed, which makes it rather competitive in real-time object recognition. The fast computation of Haar-like features is attained by using summed area tables [74] also called integral images. For improved performance of object recognition in more orientations, tilted and rotated Haar-like features are also proposed, in [72]. Several samples of Haar-like features are shown in Fig. 8.

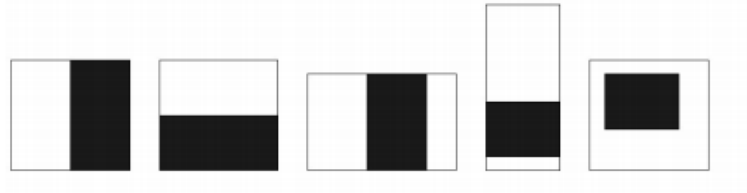


Figure 8. The samples of Haar-Like Features

3.12. AdaBoost Classifier

One of the most popular techniques in machine learning is AdaBoost. It is widely used for object recognition and classification, due to its outstanding performance and the ease to use. Moreover, its capabilities to automatically select the most relevant feature descriptors from large feature sets are also often exploited.

AdaBoost, short for Adaptive Boosting, is a technique which can be used to improve the performance of many learning algorithms [75]. Generally, AdaBoost sequentially applies a given learning algorithm with respect to a set of training samples and adds each prediction to an ensemble. When being added to the ensemble, the prediction is typically weighted according to its accuracy. After this, the dataset is also reweighted: samples that are misclassified gain weights and samples that are correctly classified lose weights. Thereby each successive classifier is forced to focus on those samples that are misclassified by previous ones in the sequence [75]. AdaBoost is chosen in this thesis work since its basic idea is quite simple but still very successful, with performances comparable to more complex methods such as Support Vector Machines.

AdaBoost is equivalent to this model and it uses the exponential loss function below for binary case [76]:

$$L(y, f(x)) = \exp(-yf(x)) \quad (3, 34)$$



For binary AdaBoost, training samples are input as feature vectors $\{x_i\}$ with their desired outputs $\{y_i\} \in \{-1, 1\}$ where $i = 1, 2, \dots, N$, the basic functions in the forward stage wise additive model are the weak learners $T^{(m)}(x) \in \{-1, 1\}$. Using the exponential loss function, the problem becomes solving [76]:

$$(a^{(m)}, T^{(m)}) = \underset{a, T}{\operatorname{argmin}} \sum_{i=1}^N \exp[-y_i (f^{(m-1)}(x_i) + aT(x_i))] \quad (3, 35)$$

for the weak learner $T^{(m)}$ and its corresponding weight coefficient $a^{(m)}$ to be added at each step. This can be expressed as [76]:

$$(a^{(m)}, T^{(m)}) = \underset{a, T}{\operatorname{argmin}} \sum_{i=1}^N w_i^{(m)} \exp(-ay_i T(x_i)) \quad (3, 36)$$

where:

$$w_i^{(m)} = \exp(-y_i f^{(m-1)}(x_i)) \quad (3, 37)$$

Since $w_i^{(m)}$ is independent on a and $T(x_i)$, it can be regarded as a weight factor that is applied to each training samples. This weight depends on $f^{(m-1)}(x_i)$ so the weight changes during each iteration m . It can be observed that [76]:

$$\begin{cases} \text{IF } y_i = T(x_i), & \text{Then } y_i \cdot T(x_i) = 1; \\ \text{IF } y_i \neq T(x_i), & \text{Then } y_i \cdot T(x_i) = -1; \end{cases} \quad (3, 38)$$

therefore:

$$e^{-a} \sum_{y_i=T(x_i)} w_i^{(m)} + e^a \sum_{y_i \neq T(x_i)} w_i^{(m)} \quad (3, 39)$$

which in turn can be rewritten as:

$$(e^a - e^{-a}) \sum_{i=1}^N w_i^{(m)} I(y_i \neq T(x_i)) + e^{-a} \sum_{i=1}^N w_i^{(m)} \quad (3, 40)$$

Apply gradient descent method to last equation and solve for a , by taking partial derivative respect to a and set the resulting equation is to 0, one obtain a as:



$$a^{(m)} = \frac{1}{2} \log \frac{1 - \text{err}^{(m)}}{\text{err}^{(m)}} \quad (3, 41)$$

where $\text{err}^{(m)}$ is the minimized weighted error rate:

$$\text{err}^{(m)} = \frac{\sum_{i=1}^N w_i^{(m)} \mathbb{I}(y_i \neq T^{(m)}(x_i))}{\sum_{i=1}^N w_i^{(m)}}. \quad (3, 42)$$

as a result, the approximation can be updated as:

$$f^{(m)}(x) = f^{(m-1)}(x) + a^{(m)} T^{(m)}(x) \quad (3, 43)$$

so the weight for the next iteration can be accordingly updated as:

$$w_i^{(m+1)} = \exp(-y_i f^{(m)}(x_i)) = w_i^{(m)} \cdot e^{-a^{(m)} y_i T^{(m)}(x_i)} \quad (3, 44)$$

considering the fact that:

$$-y_i T^{(m)}(x_i) = 2 \cdot I(y_i \neq T(x_i)) - 1 \quad (3, 45)$$

the updating scheme of sample weights becomes:

$$w_i^{(m+1)} = w_i^{(m)} \cdot e^{\beta^{(m)} I(y_i \neq T(x_i))} \cdot e^{-a^{(m)}} \quad (3, 46)$$

where $\beta^{(m)} = 2a^{(m)}$. The multiplication factor $e^{-a^{(m)}}$ is applied to all weights, so it can be ignored.

When AdaBoost is asked to classify a previously unknown sample, each classifier in the ensemble contributes its own weight $\beta^{(m)}$ to either one of the two classes it predicts, and in the end, the class with the higher value is chosen as the final prediction.

During each boosting round, the weights of wrongly classified samples are increased. In this way, the weak learner for the next boosting round will be forced



to pay attention on those misclassified. When combining the predictions after each boosting round, the training error rate is thus decreased to some extent.

It should be noted that the weight for each classifier in the ensemble should be a positive value, that is,

$$\beta^{(m)} = \log \frac{1 - \text{err}^{(m)}}{\text{err}^{(m)}} > 0. \quad (3, 47)$$

The solution to this inequality is:

$$0 < \text{err}^{(m)} < 0.5, \quad (3, 48)$$

So each weak learner must have an accuracy greater than 50%, otherwise the weight distribution for the training dataset would not be updated or to be updated towards the wrong direction thus causing AdaBoost out of work. This is also the reason why original AdaBoost algorithm can easily fail to work when facing multi-class classification problems that are more complicated than binary cases.

3.13. Cascade

One of major contributions of Viola and Jones in the paper [70] is to combine strong classifiers in a cascade construction. The evolution is in the face recognition field and is termed the cascade of classifiers. In the paper [70], the construction of a cascade declines computation time while increasing performance. In this thesis, we adopt their approach for breast detection problem.

The inputs of the model are sub window images, which are test patterns needed to classify as positive or negative samples. Layers of the cascade examine the test data one after another. The examination process is stopped in two ways; either it is rejected by one of the cascade layer or the final layer accepts it. Then, this links to a decision of negative or positive correspondingly.



In practice, in an image, most of sub-windows are negative. As such, cascade rejects most of the negative in the initial steps and will not continue on evaluating the rest of the steps. This declines the computational time sharply. Cascade is made with a number of classification layers. AdaBoost create the layers through training process. Each subsequent stage is trained by all of the positive samples and only the false positive samples of the previous cascade. Thus, the number of positive samples of the training set is fixed while the negative samples decrease in every step to deal with more expensive examples. There are two significant evaluation parameters for cascade of classifier: the detection rate and the false positive rate. If the number of cascade layer be N , then the final false detection rate is calculated as;

$$F = \prod_{x=1}^N f_x \quad (3, 49)$$

Similarly the detection rate is described as;

$$D = \prod_{x=1}^N d_x \quad (3, 50)$$

The algorithm is directly taken from Viola and Jones study [70].



Chapter 4 Procedure

4.1. System Overview

Generally, in BSE, women should use a proper search strategy to cover the whole breast region for detecting all possible abnormalities. The breast region is including top to bottom from the collarbone to the bra line; side to side from center of chest area to armpit area; and from the surface of the breast tissue to deep along the chest wall. The shaded areas in the Fig. 9 show the location of breast region.

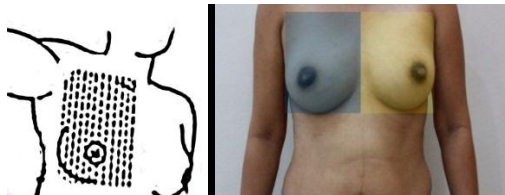


Figure 9. The shaded area shows the breast region for BSE performance

As shown in Fig. 10, the woman performs BSE utilizing a computer and a webcam. The video frames are captured by webcam while the woman performs BSE. In the processing unit, the computer vision algorithm focuses on evaluation of the BSE performance in terms of covering the entire breast region in real time by checking the palpated areas in the left and right breast regions. The woman who performs BSE can follow the BSE procedure in real-time on the monitor.

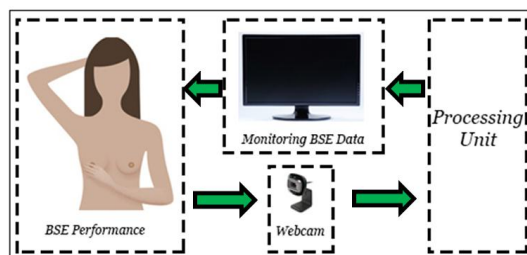


Figure 10. The System of BSE Performance Evaluation



In this research, the left and right breasts are detected using the Haar-Like features and Adaboost learning algorithm. Then, nipples are detected by analyzing the regions that were detected by Adaboost algorithm. For nipple detection, the integral imaging, color features, HOG features, and neural network are used. Then, after verifying the correction of the nipple detection, by the woman who performs BSE, the distance between the nipples is computed by counting the pixels between them. This process must be done due to the breast region tracking by tracking the opposite nipple. Because of having occlusion on the nipples while performing BSE, the breast regions are tracked based on the opposite nipples. I.e. for the left breast region, the right nipple is tracked and for the right breast region the left nipple is tracked. The nipples are tracked using the template matching technique. The breast regions are identified and clarified using a matrix based on the opposite nipple tracking. Then, motion detecting algorithm is applied for detecting the palpated blocks in the layout matrix. This process is done using the Farneback Optical flow algorithm. The palpated blocks are highlighted after palpation in real-time for the user to understand which blocks were palpated and which blocks must be palpated in the rest of the BSE performance. The percentage score of the palpated blocks are computed and monitored in real-time.

The block diagram of the BSE evaluation algorithm in terms of covering and checking the entire breast regions is shown in Fig.11. The pseudo code of the integrated algorithm for BSE evaluation is shown in Fig. 12.

In this chapter, the explanation about the dataset for developing and testing the algorithms is written in the section 4.2, the nipple detection algorithm is described in the section 4.3, the breast region identification algorithm is explained in the section 4.4, and the motion detection algorithm is interpreted in the section 4.5.

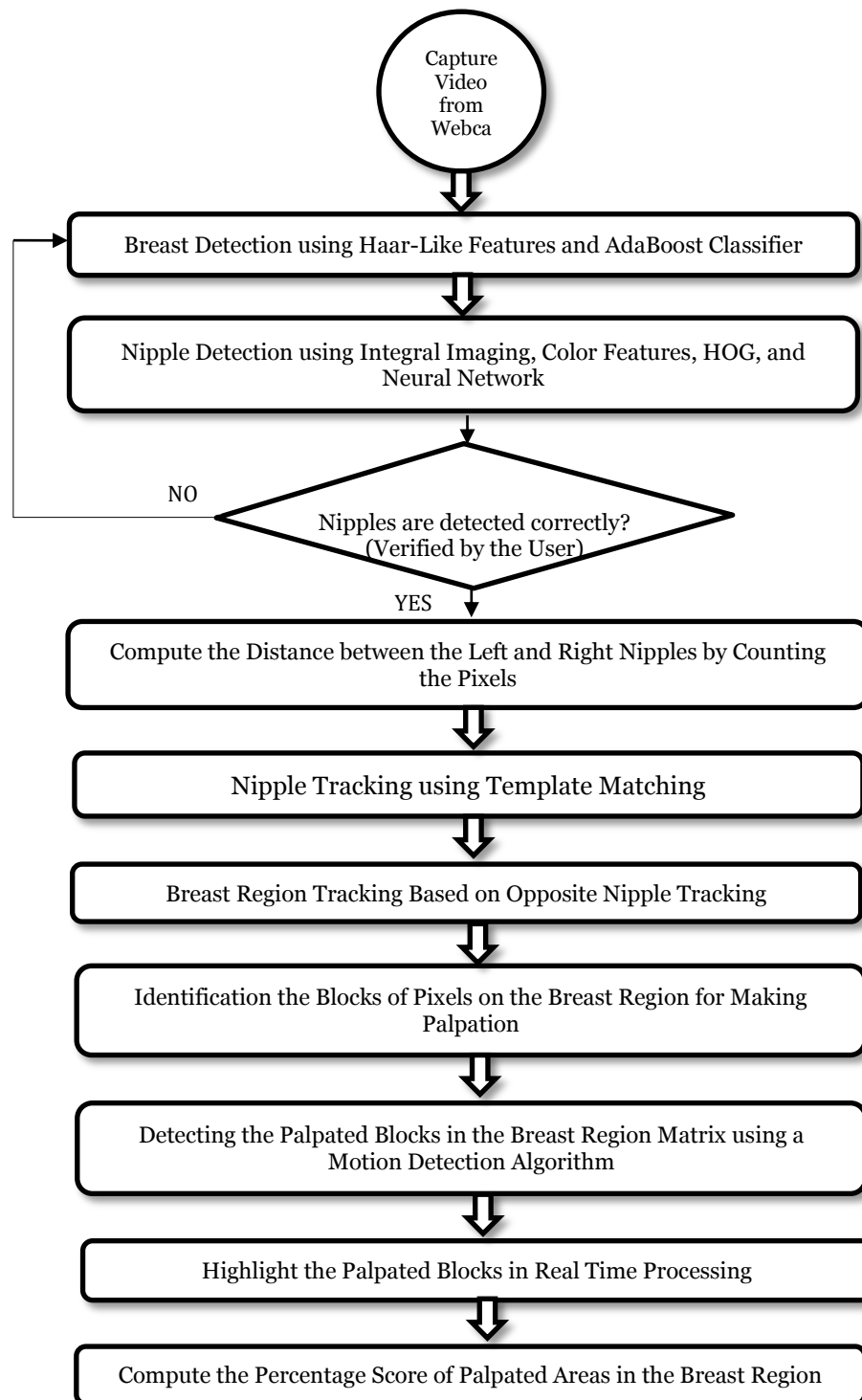


Figure 11. The Block Diagram of the Algorithm for BSE Evaluation



INPUT: Image Sequences from Webcam; (X_i) where i is the number of frames
OUTPUT: Evaluation the BSE Performance

```
1. Capture Video from Webcam
2. For( $X_i$ )
3. Breast detection using Adaboost and Haar-like features  $\leftarrow (B)$ 
4. Feature extraction for nipple detection (integral imaging, HOG, color features)
5. Nipple detection using the extracted features in 4, and neural network
6. If the breasts are detected correctly(must be verified by user)
    6.1. IF NO
    6.1.1. Return to 2 for analyzing the next frame
    ENDIF
    6.2. IF YES
    6.2.1. Capture and save the left and right nipples for tracking  $\leftarrow (L,R)$ 
    6.2.2. Compute the distance between the nipples by counting the pixels  $\leftarrow (D)$ 
7. FOR ( $X_i$ )
7.1. Tracking the opposite nipple using  $(L,R)$  in 6.2.1, and template matching
7.2. Compute the Size of the breast region matrix from  $(B)$  in 3, and  $(D)$  in 6.2.2.
7.3. Locate the breast region matrix on the left and right breast
7.4. Palpation detection on the blocks of the matrix using Farneback Optical Flow as follows (the Matrix has 6 Blocks):
7.4.1. Resize the blocks to 100 by 100 pixels
7.4.2. Apply Farneback Optical Flow on each block sequentially
7.4.3. Divide the output matrix of optical flow into four small tiles
7.4.4. Apply NormL2 on each tile
7.5. IF(NORML2>100000 in 200 frame)
7.6. Consider the tile as a palpated tile
7.7. IF(four tiles in one block were palpated)
7.7.1. Highlight the palpated block
7.7.2. Compute the percentage score of the palpated areas on breast region
    ENDIF
7.7.3. Show the percentage score and highlighted blocks in real-time
    ENDIF
    ENDFOR
  ENDFOR
ENDFOR
```

Figure 12. The Pseudo code of the Integrated Algorithm for BSE Evaluation



4.2. Nipple Detection and Tracking

In nipple detection algorithm, the left and right breasts are detected using the Haar-like features and cascaded Adaboost classifier. Then, the features and pixel intensities inside the breast regions are analyzed for detecting the nipples. The integral imaging, color features, HOG, and neural network are used for nipple detection. The block diagram of the nipple detection and tracking algorithm is shown in Fig. 13.

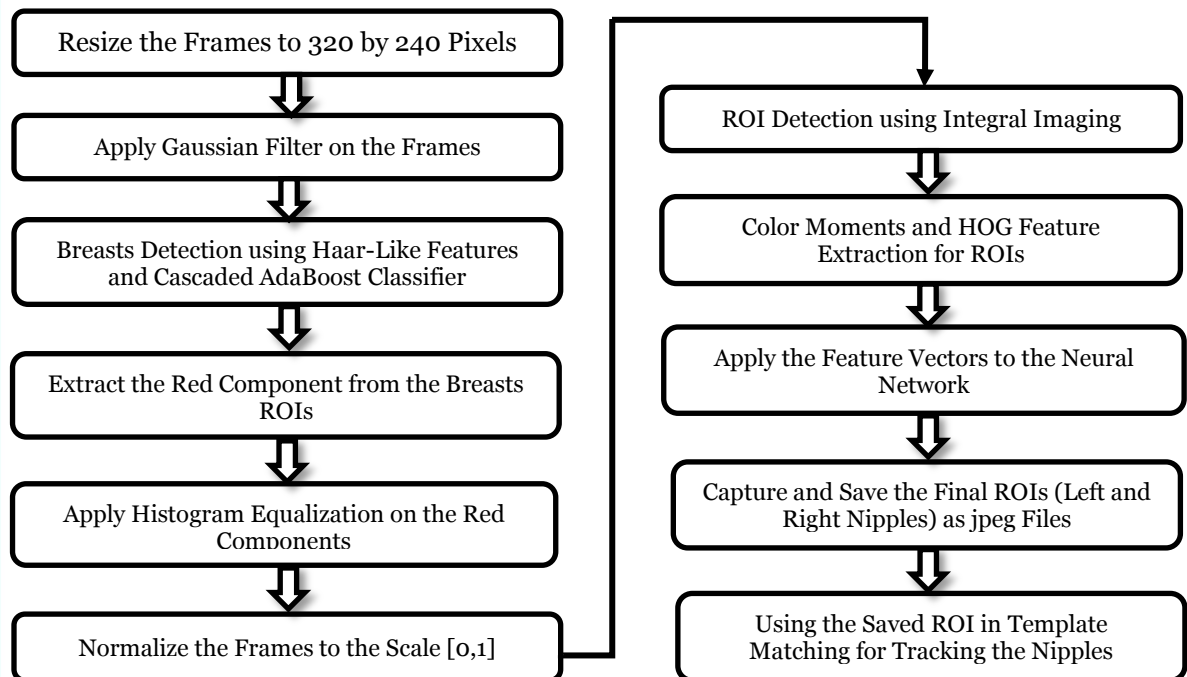


Figure 13. The Block Diagram of Nipple Detection Algorithm

4.2.1. Breast Detection

For a known training data set and set of features, every machine learning algorithm can be used in learning a classification. Boosting is a kind of machine learning technique, which performs supervised learning. The idea explains that a



strong classifier can be created by linearly linking a number of weak classifiers. Weak Classifiers are acquired by response of the features from training data and have the detection rate, which is a little better than the random guessing results. To increase the detection rate, it is essential to employ a machine learning approach. AdaBoost, short for Adaptive Boosting, is the most popular boosting algorithm, which was introduced by [75].

It has two significant modifications over the previous methods:

- Using a weighted sample to focus learning on most challenging samples of training data.
- Using a weighted vote for integrating and combining classifiers.

It is named AdaBoost because it is adjusted adaptively to the errors of weak hypothesis and returned by weak learn [75]. As in Viola and Jones paper [70], a variant of AdaBoost handles the two tasks: selecting the features and training the classifier. The weights of training set define the probability of being selected for a feature and they are continuously updated in each iteration. If a training pattern is exactly classified, then its chance of being used again in the next round is decreased. While the weights of properly labeled examples are decreased, the wrongly classified examples weights are increased. This is named reweighing method. The summary of the variant of AdaBoost, which is directly derived from [70], is described in table 1. The samples of haar-Like features that were used in this algorithm are shown in fig 14.

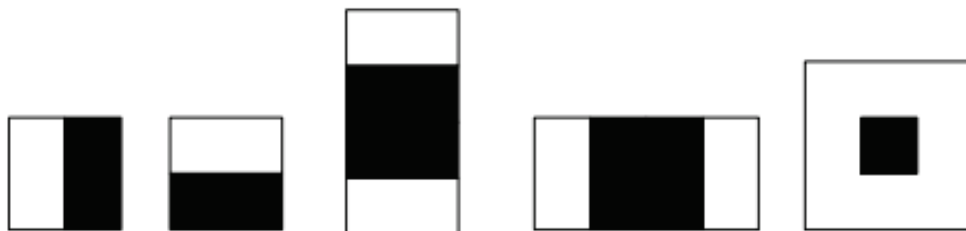


Figure 14. The Haar-Like Features for using in AdaBoost



Table 1. Algorithm Summary of AdaBoost

- Given example images $(x_1, y_1) \dots (x_n, y_n)$ where $y_i = 0, 1$ shows negative and positive samples respectively.
 - Initialize weights $w_{1,i} = \frac{1}{2neg}, \frac{1}{2pos}$ for $y_i = 0, 1$ respectively, where neg and pos are the number of negatives and positives respectively.
- For $t = 1, \dots, T$:
1. Normalize the weights: $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$; so that w_t is a probability distribution.
 2. For each feature, j , train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to $w_t, \epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$.
 3. Choose the classifier, h_t , with the lowest error ϵ_t .
 4. Update the weights: $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$
 where $e_i = 0$ if example x_i is classified correctly; $e_i = 1$ otherwise, and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$
 - The final strong classifier is: $h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{otherwise} \end{cases}$;
 where $\alpha_t = \log \frac{1}{\beta_t}$

The parameter α_t is an important measure of the classifier. By rewriting the equation the relationship between the error and the α_t can be written as follow:

$$\alpha_t = \frac{1}{2} \ln \frac{1-\epsilon_t}{\epsilon_t} \quad (4, 1)$$

$$\text{So, } \alpha_t \geq 0 \text{ if } \epsilon_t \leq \frac{1}{2} \quad (4, 2)$$

This rule can be true since it is dealing with the binary case that has only the 0 and 1 labels. If the classifier is trained for a feature that has $\alpha_t > \frac{1}{2}$, similar to random guessing, then the feature is considered as a good candidate for one of the single feature classifier in the weak learning algorithm. In this study, the



number of negative samples is 8000, and the number of positive samples is 1000. The size of the window for finding the features is 24 by 24 pixels.

Cascade of Classifiers is used to increase the accuracy and speed of the algorithm. As shown in Fig. 15, a cascade consists of layers. As a substitute of applying all the features on a window, the features are grouped into different stages of classifiers for applying one-by-one. If a window fails the first stage, it would be discarded, and the remaining features are not applied on it. If it passes, the second stage of features would be applied and this process is continued up to the final stage. The window which passes all stages is a breast region.

In this study, the complete breast detection cascade has 7 stages. The Gaussian Filter was applied on all the samples before starting to train. The reason of using Gaussian is to reduce the noise, and smooth the images. Therefore, for applying each frame into the Cascade, that frame should be smoothed using the Gaussian Filter. Gaussian kernel size (height and width) both can differ but they must be positive and odd. Here, the size of the Gaussian kernel is 3 by 3 pixels. In the OpenCV, The Gaussian Sigma is computed based on the kernel size. The sigma was considered to be 0.3 based on the kernel size.

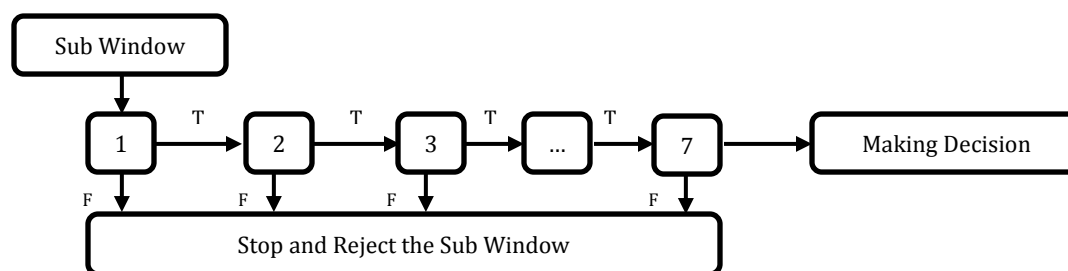


Figure 15. The Form of the Cascade

4.2.2. Nipple Detection

The extracted ROIs (Left and Right detected Breasts in the previous section) are first converted into RGB components, but only the red component is used for the nipple detection. The reason for using the red level is that the color of nipples



is close to red, and the result would be more accurate while analyzing pixel intensities for detecting the nipples. Then, the histogram equalization would be applied on the frames. The reason for using the histogram equalization is to enhance and adjust the contrast in the detected breast frames. As described in [68], cumulative histogram equalization is proposed for implementation in the DSP. This algorithm was carefully chosen due to its good performance and easy performance in the C language. Then, the image normalization would be done on the frames. This process should be employed due to decreasing the effect of luminance conditions on the algorithm. At the next step the integral imaging is performed on the frames. The nipple detection algorithm is based on integral image features. The Integral Image was used as a fast and effective way of calculating the sum of pixel intensities in a given image. The integral image features are accomplished by means of convolution between a kernel and an image. In this study, the kernel size is 60 by 90 pixels, and the breasts windows which were detected in the previous section are resized to 100 by 100 pixels. The form of the kernel is shown in Fig 16. As shown in Fig 16, the kernel window has six smaller windows inside it. The sums of the pixel intensities, in the smaller windows, are computed for detecting the nipples. The rules for detecting the region of interests (ROI) for the nipples are described below:

- If the sum of the pixel intensities in the fifth window is smaller than other windows, the fifth window is considered as an ROI for detecting the right nipple.
- If the sum of the pixel intensities in the second window is smaller than other windows, the second window is considered as an ROI for detecting the left nipple.

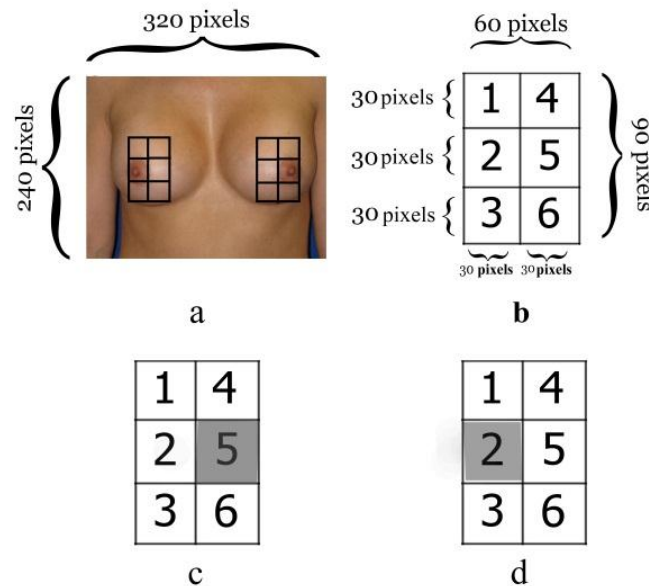


Figure 16. a, b) The structure of the ROI windows c) The sum of the pixels in 5th square must be less than others for left-breast detection d) the sum of the pixels in 2nd square must be less than others for right-breast detection

Using this method, the nipples are detected. But, there might be several false positive detections. The color features, HOG features, and artificial neural network are used to eliminate the false positives. The color features are the mean and standard deviation of the pixels in the ROIs that were detected as nipples. The color features are derived from the Hue layer. The output of the HOG features is a vector with 9 integers. These integers are derived from binning the orientation of the gradients. Basically, there are 360 degrees for the orientation of the gradients. But, that degree is reduced to 9 groups in order to increase the speed of computation, and each orientation (from 0 to 360) would be a vote for each group. Finally, the detected small windows are resized to 10 by 10 pixels for using the pixel intensities in the vector that would be applied in the neural network. So, the vector for using in the neural network has 111 elements (100



elements are from the pixel intensities, 9 elements are from the HOG features, and 2 elements are from the color features).

4.2.3. Training the Neural Network

Artificial Neural Networks (ANNs) have been employed for pattern classification and recognition in different applications. Multi-Layer Perceptron Neural Network is among the best neural networks in manipulating classification problems. These networks consist of 3 layers in most cases, input, hidden and output layer. Generally, ANN has a number of activation functions, training methods, and variations. The activation function of hidden and output layers has to be adapted in accord with the problem. In this research, the activation function is sigmoid. The optimization methods to minimize error are different but Levenberg-Marquardt has shown to have better results in most cases. Therefore, in this study, the Levenberg-Marquardt algorithm was used to train the neural network. Training continues until the network continues modifying on the validation set.

As said previously, the vectors with 111 elements are used for training and testing the neural network. 100 elements of the vector are derived from the pixel values of the ROI from the red component, after resizing them to 10 by 10. 9 elements are derived from the HOG features, and 2 elements are derived from the color features. In this study, The ANN had 20 neurons in the hidden layer, 111 neurons in the input layer, and 2 neurons in the output layer. All the elements are normalized for transferring to the range $[0, \dots, 1]$, before training and testing the ANN. The structure and feature vectors of ANN used in this study is shown in Fig. 17.

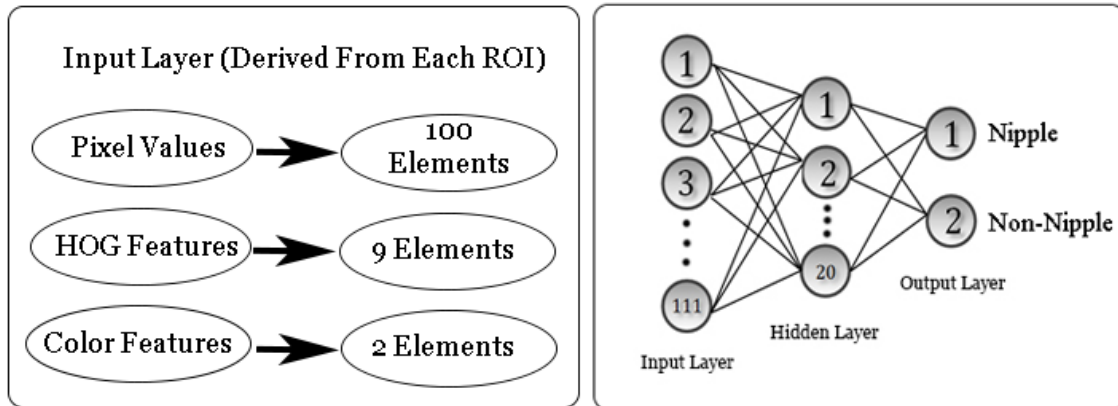


Figure 17. The Structure of the Neural Network for Nipple Detection

250 positive ROIs and 5600 negative ROIs were used for training the neural network. Positive samples are those windows including a nipple, and negative samples are those small windows without a nipple. There are two neurons are located in the output of neural network. one of them is for classifying the nipples and another one is for classifying a bounding window which is not containing a nipple. After preprocessing and feature extraction, the vectors of features were used as the input for training the neural network. 80% of the data was used for training, 10% of the data for test and the remaining 10% for validation. The test set provides a completely independent measure of network accuracy. Since the ANN starts with random initial weights, the results differ slightly every time the ANN is run. The stopping criterion for the training is based on the movement of validation. After performing 326 epochs, the validation is constant in three sequences at the epochs 324, 325, and 326. During iterative training of the ANN, an epoch is a single pass through the entire training set, followed by testing of the verification set. The training stopped based on the constant error values of validation at the epochs 324, 325, and 326. The best performance as measured in terms of mean squared error was 0.0001108 after 326 epochs with a 99.1 % correct classification accuracy.



4.2.4. Nipple Tracking

After detecting the left and right breasts and nipples, the nipples must be tracked in real time. For checking and palpation on the left breast region, the right nipple is tracked, and for the checking and palpation on the right breast region, the left nipple is tracked. The template matching is used for tracking the nipples. The references of left and right nipples for the template matching are derived from detecting the nipples in the previous section. That means, the nipples are saved as two jpeg files for using as the references in template matching for tracking the nipples.

For the nipple tracking, the normed square difference matching method, normed correlation matching method, and normed correlation coefficient matching method were applied and tested to recognize the most accurate model. The testing process is described in the chapter.5. The normed correlation Coefficient matching method is considered as the most accurate method for tracking the nipples. This method match a template relative to its mean against the image relative to its mean, so a flawless match will be 1 and a perfect mismatch will be -1; a value of 0 simply means that there is no correlation. The formulation of the Correlation coefficient matching method is described as follow:

$$F_{coeff}(x, y) = \sum_{x', y'} [T'(x', y') \cdot i'(x + x', y + y')]^2 \quad (4, 3)$$

$$T'(x', y') = T(x', y') - \frac{1}{(w \cdot h) \sum_{x'', y''} T(x'', y'')} \quad (4, 4)$$

$$i'(x + x', y + y') = i(x + x', y + y') - \frac{1}{(w \cdot h) \sum_{x'', y''} i(x + x'', y + y'')} \quad (4, 5)$$

Normed CCOEF:



$$N(x, y) = \sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} i(x + x', y + x')^2} \quad (4, 6)$$

$$F_{ccoeff_normed}(x, y) = \frac{F_{ccoeff}(x, y)}{N(x, y)} \quad (4, 7)$$

where the $F(x, y)$ is the result, $i(x, y)$ is the input image, and $T(x', y')$ is the template (the nipples that were captured).

4.3. Breast Region Identification

The BSE should be performed with checking the entire breast region for finding probable lumps and abnormalities. So, the Breast region or area thoroughness must be properly detected for the evaluation of BSE performance. Each breast region is identified by a matrix with the size 2 by 3 blocks. Here, each small window in the matrix is called a Block. The blocks are shown on the screen and women should perform palpation and checking the breasts by following the numbers of the blocks. The fig. 18 shows the samples of left and right breast regions identification.



Figure 18. The Samples of Breast Regions Identification

The block diagram of the algorithm is shown in Fig. 19. As described in Fig. 19, the breasts and nipples must be properly detected and captured in the first stage. Then, the distance between the nipples is calculated by counting the pixels. The distance between the nipples is needed for locating the matrix on the breast



regions while the woman performs BSE. Then, the nipple tracking is applied. For locating the matrix on each breast regions, the opposite nipple is tracked. The matrix is located on each breast region based on the distance between the nipples.

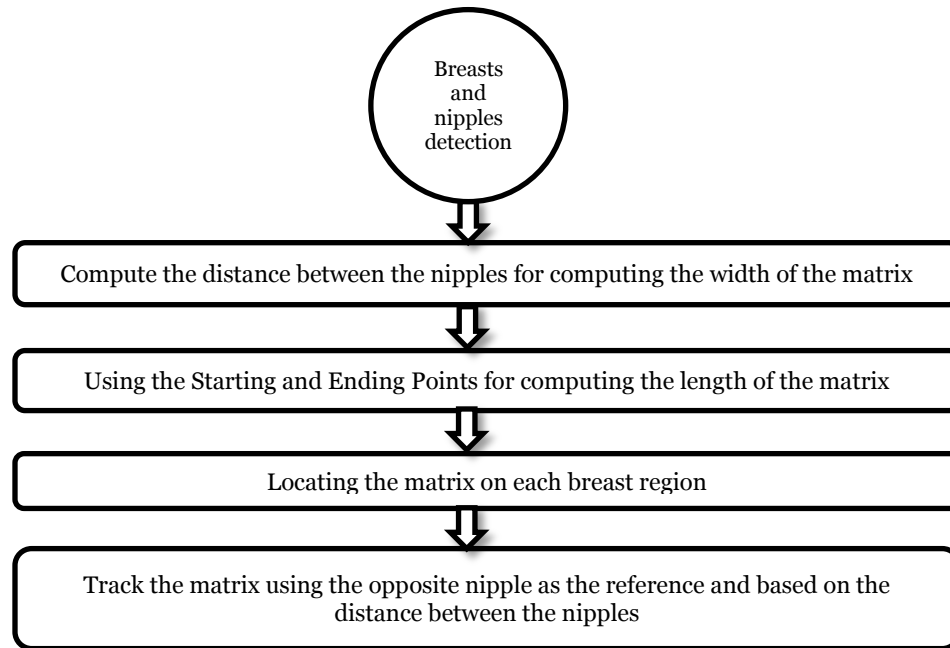


Figure 19. the block diagram of the breast region identification

In this algorithm, the point at the top and left side of the square of breast detection, using Cascaded AdaBoost in the previous chapter, is called the Starting Point, and the point at the low and left side of the square is called the Ending Point. As shown in Fig. 20, first, the red squares were detected using the cascaded Adaboost and Haar-Like features, and at the next stage, the starting and ending points were identified. In the Fig.20, the green circles show the starting points for the left and right breasts, and the yellow circles show the ending points for the left and right breasts.

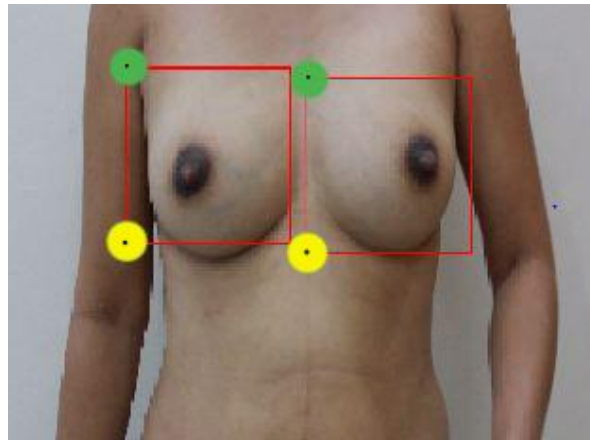


Figure 20. The Detected Breasts with Starting and Ending Points

The width of the matrix, that identifies the breast region for making palpation, is based on the distance between the nipples. The distance between the nipples is divided by two to compute the width of the matrix on breast region. The width of the matrix is started from the Starting Point. The length of the matrix is based on the distance between the Starting and Ending Points. These distances are computed at the first stage of the algorithm. Since the motion in Z axes is very small at the time of BSE performance (there is no zoom in or zoom out), the algorithm does not need to compute the distances in real time. Instead, they are computed in the first frame and the BSE performance is started after verifying the correction of breasts and nipples detection by the user (woman who performs BSE).

As said previously, in real time, each matrix is located on left and right breast regions based on the opposite nipples. For locating the matrix on the left breast region, the right nipple is used as the reference, and for locating the matrix on the right breast region, the left nipple is used as the reference. Since the BSE performance is done using the left hand for the left breast and right hand for the right breast, normally, the occlusion and covering the opposite nipples are not occurred in the BSE performance. However, for the sake of robustness, the simple



method for the time of occlusion is added here. The rules for the mentioned method are described below:

- If the distance between the matrices (the upper and left point of the matrix) of two frames is more than 20 pixels; that is considered as a fake motion and the new position of the matrix would be based on the data from the previous frame.
- If the distance between the matrices (the upper and left point of the matrix) of two frames is less than 20 pixels; that is considered as a real motion and the position of the matrix would be based on the data from the new frame.

4.4. Detecting the Palpated Blocks

The blocks which are detected at the previous section must be palpated for finding abnormalities. This section explains how to detect the palpated blocks in real time.

After identifying the matrix and enumerated blocks on the breast region, the palpation can be started. There are six blocks in each matrix for the left and right breast regions. The palpation detection for each block is derived from the motion detection using the farneback optical flow algorithm. The flowchart of this objective is shown in Fig. 21.

The algorithm, analyzes each block separately and sequentially. Initially, the first block is analyzed for palpation detection. If the palpation is done correctly on the first block, the second block would be checked for palpation detection, and this process sequentially would be performed on the next blocks. In other words, if palpation is done on the second block without completing palpation on the first block, the palpation on the second block is not detected. This method is used to increase the accuracy of palpation detection algorithm on the breast regions.

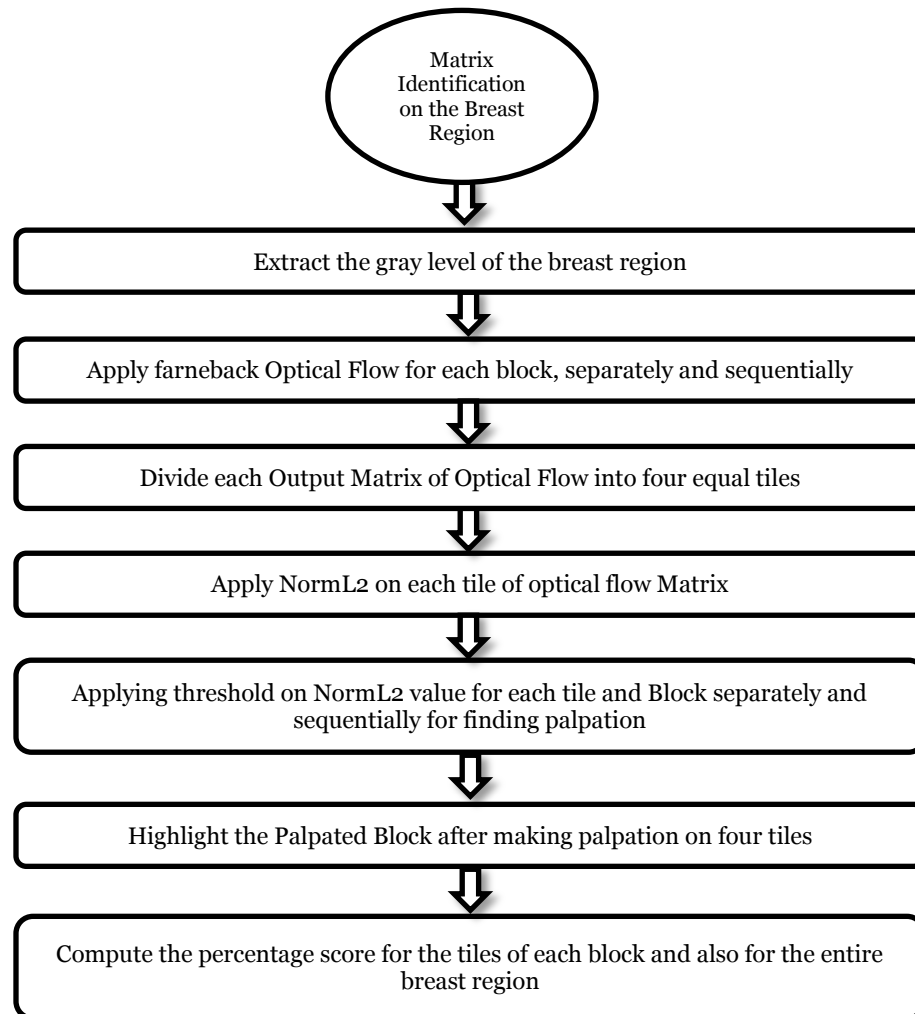


Figure 21. The Flowchart of the palpation detection algorithm

First, the video sequence is treated as a stack of consecutive frames. In a spatiotemporal image volume like this, the vertical axis corresponds to the moments of time. A moving point depicted in the image sequence will create a line inside this three-dimensional space. The orientation of the line can be expressed with the orientation tensor T . For N -dimensional signals T takes the form of an $N \times N$ real symmetric matrix. In order to approximate T we first



assume that the signal can be locally (within some neighborhood) modeled by the second degree polynomial:

$$f(x) \sim x^T A x + b^T x + c \quad (4, 8)$$

where A is a symmetric matrix, b is a vector, and c is a scalar. The coefficients are estimated from a weighted least squares fit to the signal values in the neighborhood. The parameters A , b , and c are computed from image data by the Normalized Convolution [89], an algorithm that assigns the signal a certainty component conveying the level of confidence in the reliability of each measure (typically, in this case the weighting function is Gaussian). The procedure is equivalent to calculating a weighted least squares approximation of the signal but can be efficiently executed with the assistance of convolution filters. Finally T is expressed as [63]:

$$T = A A^T + \gamma b b^T \quad (4, 9)$$

where the value of parameter γ should be decreased when the size of the neighborhood becomes bigger (as the quadratic part should become more significant than the linear component). The faster version of the algorithm does not estimate the velocity from the tensors point by point but assumes that the velocity field over a region can be parameterized according to the affine motion model and uses all the tensors in the region to compute the parameters.

Having extracted the sequence of blocks containing palpation, the optical flow is computed. Each block is resized to 100 by 100 pixels before applying the optical flow. For the Farnebäck algorithm, three parameters were adjusted to detect the palpation: the averaging window size " w ", the size of the pixel neighborhood considered when finding polynomial expansion in each pixel " s ", and the standard deviation " σ " of the Gaussian used to smooth derivatives in the polynomial expansion. The remaining parameters were set to their default OpenCV values.



The displacement is calculated from the center block and the block before. The averaging over neighborhoods is done using a 70 by 70 Gaussian weighting function with standard deviation 6. The polynomial expansion is done with a 19 by 19 Gaussian applicability with standard deviation 3.5. In order to reduce the errors near the borders, the polynomial expansions have been computed with certainty set to zero on the border. Additionally pixels close to the borders have been given a reduced weight because the expansion coefficients still can be assumed to be less reliable there. The constant and affine motion models have been used with a single iteration.

Finally a motion vector is determined for each small neighborhood. Initially, it was planned to exploit the data describing the directions of the movements but this information seemed to be too noisy and unstable in practice. Thus, it was decided to use only the information about the velocity, i.e. it was considered only the length of each motion vector.

The output matrix of optical flow including the velocity vectors is divided into four equal tiles. Then, the Norm-L2 of the motion vectors in each tile is computed and classified as palpated or non-palpated block using a threshold and time. The Norm-L2 is calculated as follows:

$$|x| = \sqrt{\sum_{k=1}^n |x_k|^2} \quad (4, 10)$$

Four threshold values were analyzed to choose the most efficient and accurate value for classifying palpation on blocks. The thresholding values are 50000, 75000, 100000, and 125000 respectively. After having several experiments the 100000 was chosen as the best value to detect palpation. The detail of evaluation results is described in the section 5.3.4. The approximate time for the palpation is 10 seconds based on the suggestions of the pathologists. Since the minimum fps is 20 in this research, there must be minimum 200 frames for each tile for the palpation consideration. Each block is highlighted after performing palpation on



four tiles in the block. This process is employed to monitor BSE evaluation data to the user (woman who performs BSE). So, the user continues palpation until all the blocks are highlighted. This procedure is done for the left and right breasts separately. The score of the palpation percentage is computed for each block, and monitored in real time. This score shows how many present of each block was palpated. The sample of highlighted blocks is shown in Fig. 22. As shown in Fig. 22, the second and third blocks were highlighted, and the rest of the blocks must be palpated for finding abnormalities.

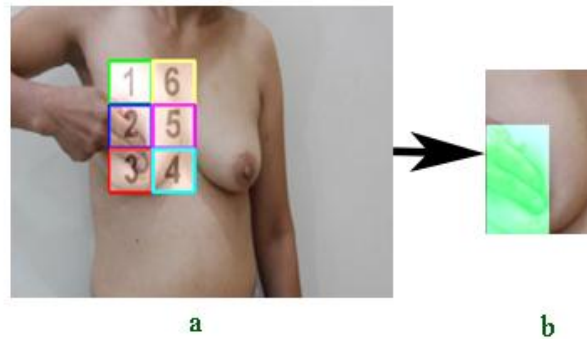


Figure 22. a) The Sample of BSE Performance. b) The 2nd and 3rd blocks were highlighted after palpation

The Pseudo code of this objective is shown in Fig. 23. As said previously, in the integrated algorithm, the woman who performs BSE must check and palpate the blocks on the breast region sequentially. I.e. the palpation must be done from the first block to the last block, one by one. To increase the accuracy of the model, it is decided to not analyze all the blocks at the same time. Instead, the blocks are analyzed separately, and if the palpation is done correctly on one block, the next block is started to be analyzed for palpation detection. Aside from that, a filter is used to reduce the noise to classify the palpated blocks. In the filter, if the palpation between two frames was detected based on the threshold on NormL2 value, the forward non-palpated frames are counted until the next palpated detection. There are two rules in that filter:



- If the number of non-palpated detections, after the palpation detection between two frames, is less than 20 sequentially, the non-palpated detections are considered as noise and counted as the palpation detections.
- If the number of non-palpated frames, after a palpated frame, is more than 20 sequentially, the palpation detection is considered as the noise, and not counted as the palpation detection.

Input $f(x)$: Image Sequences with Breast Region Matrix and Blocks of the Pixels

Output: Palpation Detection

```
1. For( $f(x)$ )
2.   Extract the gray level of the breast region matrix
3.   Resize the blocks to 100 by 100 pixels
4.   Apply optical flow on each block sequentially
5.   Divide each output matrix of optical flow into four equal tiles
6.   Apply NormL2 on each tile
7.   IF(NormL2>100000)
8.     Count the frames considering the filter
9.   ENDIF
10.  IF(Number of Frames=200)
11.    Consider that tile as a palpated tile
12.    Compute the percentage score for each block
13.  IF (NormL2<100000)
14.    Palpation is not done with respect to the filter
15.  ENDIF
16.  IF(four tiles in each block were palpated)
17.    That block is considered as a palpated block
18.    Highlight the palpated block
19.    Compute the percentage score for tiles and blocks in the breast region
20.    Monitor the evaluated data (highlighting and percentage score) on the screen
21.  ENDIF
22. ENDFOR
```

Figure 23. The Pseudocode of the third objective

In order to increase the accuracy of the palpation detection algorithm, the output matrix of the optical flow is divided into four equal tiles as shown in Fig. 24. Since, sometimes, the palpation is done on the borders of the blocks; there is a need to detect the palpation for each tile separately. Generally, it is not possible to divide the blocks before applying the optical flow. Because, the size of the blocks



would be very small and the farneback algorithm with the stated parameters is not able to detect palpation, based on the experiments. After dividing the output matrix of optical flow into four equal tiles, the NormL2 is applied in each tile to compute the integer for each tile. As said previously, using the threshold on the NormL2 value, the filter to reduce the noise, and the time, the palpation on each tile and block is detected. The Fig. 25 shows the graphs about the percentage of palpated tiles with respect to time.

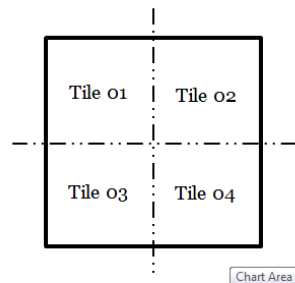


Figure 24. Dividing the Output Matrix of Optical Flow into Four Equal Tiles

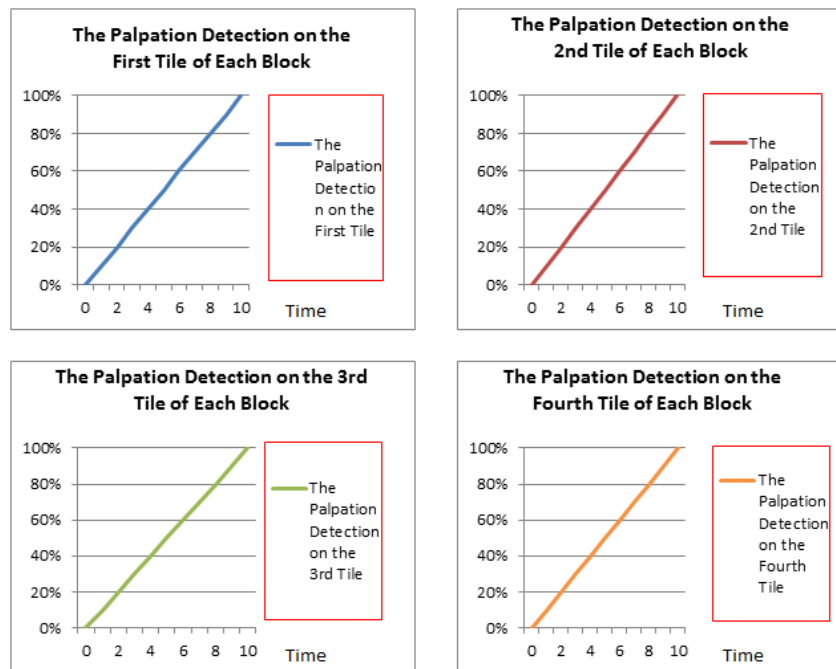


Figure 25. The Score of Palpation Process with Respect to Time for Each Tile of the Block



As shown in Fig. 25, if the palpation is done correctly with respect to time and the threshold on the NormL2 value, the percentage score for palpating the tiles is gradually increasing until it reaches the 100%. This process is done for each tile in each block separately. However, if the palpation is done on four tiles simultaneously, the percentage scores of four tiles would be increased simultaneously. But, if at the time of palpation on each block, the next block is palpated, the palpation on the next block is ignored and not detected. i.e. the palpation process must be performed for the blocks separately and sequentially.

The graphs of the example in which the palpation is not performed correctly are shown in Fig. 26. As shown in Fig. 26, in the first tile of the block, the palpation was performed for four seconds with respect to threshold and filter. So, 40% of the first tile was palpated. 20% of the second tile, 60% of the third tile, and 10% of the fourth tile were palpated, based on the time, threshold, and filter.

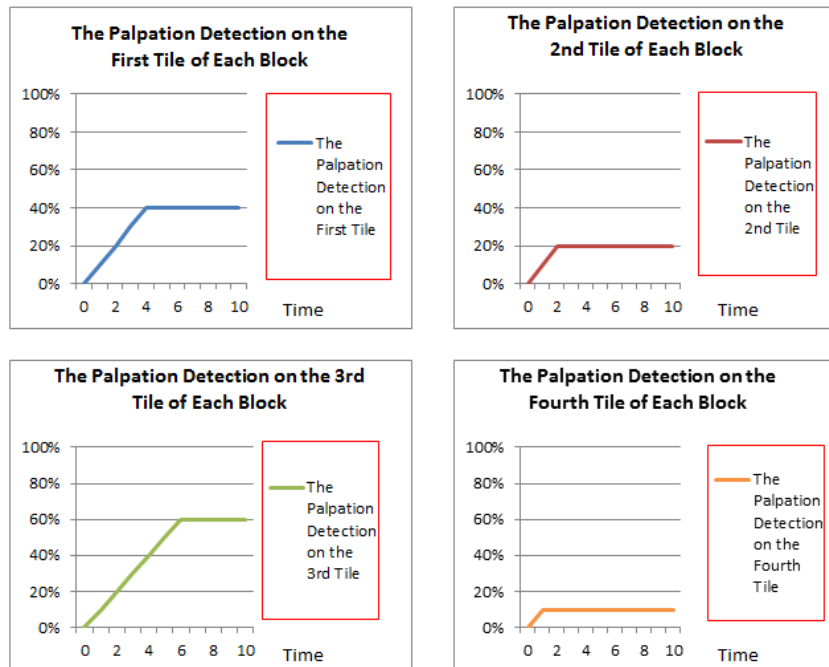


Figure 26. The Detection of Non-Palpated Tiles in the Block



Chapter 5

Results and Evaluation

5.1. Datasets

There are mainly 4 datasets used in the experiments. For simplicity, these datasets are called Dataset-1, Dataset-2, Dataset-3, and Dataset-4.

5.1.1. Dataset-1

The Dataset-1 consists of positive images including the breasts [87], and negative images without any breast [88] for training and evaluation of breast detection using the Adaboost and Haar-like features. The number of positive samples in this dataset is 1000 and the number of negative samples is 8000. The positive samples were selected with different forms of breasts, and luminance conditions. Since, in the actual BSE performance, the women must sit in front of the webcam and work with the computer, cellphone, or tablet, the distance between the woman and the webcam would be in the range of 40 centimeters to 80 centimeters. Thus, in this dataset, the positive samples were captured and selected with the mentioned distance between the webcam and women. The positive samples of images in dataset-1 are shown in Fig. 27.



Figure 27. The Sample of Positive Samples in Dataset-1



5.1.2. Dataset-2

The Dataset-2 consists of positive and negative samples for training and testing the neural network for the nipple detection. The positive samples are those sub-windows containing the nipple and the negative samples are those sub-windows without any nipple. In the Dataset-2, the number of positive sub-windows is 250, and the number of negative sub-windows is 5600. The size of all samples is 10 by 10 pixels. The positive samples of sub-windows in Dataset-2 are shown in Fig. 28.



Figure 28. The Positive Samples of Dataset-2 for Training and Testing the Neural Network

5.1.3. Dataset-3

The Dataset-3 consists of 20 videos for the evaluation of nipple tracking algorithm. The length of the videos is 10 seconds including 200 frames. For the evaluation, 40 frames are extracted from each video (1 frame from each 5 frames). Thus, the total number of frames is 800 (20 videos, and 40 frames from each video) for the evaluation of nipple tracking. The videos were captured in different luminance conditions. The samples of frames of these videos are show in Fig. 29.



Figure 29. The Samples of Frames in Dataset-3

5.1.4. Dataset-4

Dataset-4 consists of positive and negative videos for the evaluation of palpation detection algorithm. Since, the appropriate time for the palpation on each block was considered 10 seconds, the length of the videos in the dataset-4 is 10 seconds. The positive videos are those with correct palpation on the block, and the negative videos are those without any palpation on the block. The number of positive videos is 20 and the number of negative videos is 30. The sample videos were captured in different luminance conditions. The videos were captured by the researcher from women volunteers. The samples of the positive and negative blocks of these videos are shown in Fig. 30. Since palpation detection algorithm detects the motion between two frames, the samples of two sequence blocks from the frames are shown in Fig. 30. In the Fig.30, the “a” and “b” show the blocks from the positive videos, and “c” and “d” show the blocks from the negative videos.

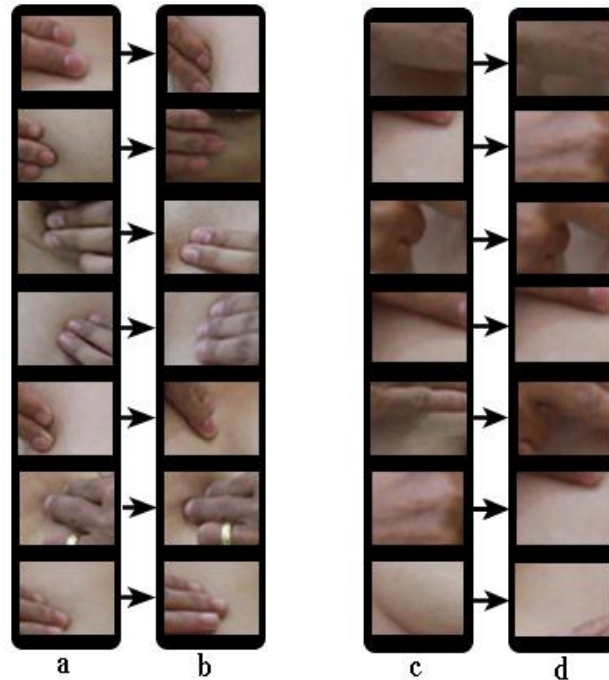


Figure 30. The Samples of Blocks of the Videos in Dataset-4. a, b) First and Second Extracted Positive Blocks from Videos. C,d) First and Second Extracted Negative Blocks from Videos

5.2. Performance Criteria

Since we have a limited number of data sets, cross validation is an appropriate method to evaluate the performance of the Adaboost and neural network models. The holdout method is the easiest method of cross validation. The data set is divided into two sets, which are called the training set and the testing set. The learning algorithms decide the classifier using the training set, and then the classifier computes the values for the data in the testing set. It is vital that the classifier should not use the test dataset while training process. The errors it makes are accumulated as before to give the mean absolute test set error, which is used to evaluate the model (Fig. 31). The holdout method is usually preferable to the other cross validation methods since the number of repetitions to evaluate the learning algorithm is zero.

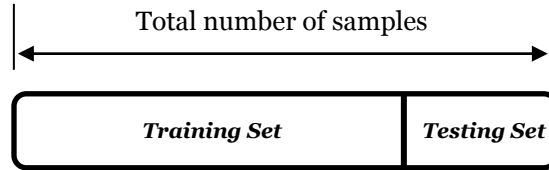


Figure 31. Cross validation- Holdout method

The performance evaluation for nipple and breast region detection is carried out over the image sub-windows. For this purpose the comparison of the ground truth data with the result of the Adaboost Algorithm and neural network for each image block is required. The ground truth data were labeled as true manually while more than half of the pixels of sub-windows belong to that class of object.

The performance equations are the key concept to evaluation of the algorithm and defined by labeled ground truth data and the result made by the AdaBoost algorithm and neural network. The most popular one is Confusion Matrix. The binary case of the confusion matrix is presented in Table 2. It contains information about actual and predicted classifications done by a classification system. Performance of such systems is normally evaluated using the data in the matrix. In this thesis, the confusion matrix is also used for the evaluation of the palpation detection algorithm based on the positive and negative videos which were described in the previous section.

Table 2. Confusion Matrix for a Binary Classification

		<i>Predicted</i>	
		Negatives	Positives
<i>Actual</i>	Negatives	<i>A</i>	<i>B</i>
	Positives	<i>C</i>	<i>D</i>

$$Accuracy (AC) = \frac{A + D}{A + B + C + D} \quad (5, 1)$$



$$\text{Recall or True Positive Rate (TP)} = \frac{D}{C + D} \quad (5, 2)$$

$$\text{False Positive Rate (FP)} = \frac{B}{A + B} \quad (5, 3)$$

$$\text{True Negative Rate (TN)} = \frac{A}{A + B} \quad (5, 4)$$

$$\text{False Negative Rate (FN)} = \frac{C}{C + D} \quad (5, 5)$$

$$\text{Precision (P)} = \frac{D}{B + D} \quad (5, 6)$$

Many measures for evaluating the tracking performances have been proposed [93], typically with the comparison against ground truth, considering the target presence and position.

The three basic types of errors in tracking are:

- Deviation: the track's location deviated from the ground truth.
- False positive: tracker identifies a target which is not a target.
- False negative: tracker misses to detect and locate the target.

A reasonable choice for overlap of target and object is the PASCAL criterion [90]:

$$\frac{|T^i \cap GT^i|}{|T^i \cup GT^i|} \geq 0.5 \quad (5, 7)$$

where T^i signifies the tracked bounding box in frame i , and GT^i represents the ground truth bounding box in frame i [91]. In many works this PASCAL overlap measure is adopted without threshold, similar to the similarity metric without threshold called Dice in [94]. As explained in [93], it is better to use it with threshold as it makes it easier to evaluate large sets of sequences.



For ntp , nfp , nfn indicating the number of true positives, false positives and false negatives in a video:

$$recision = \frac{ntp}{ntp+nfp} \quad (5, 8)$$

$$recall = \frac{ntp}{ntp+nfn} \quad (5, 9)$$

The F – score combines the precision and recall to compute the accuracy of the nipple tracking algorithm [92]:

$$F = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (5, 10)$$

5.3. Evaluation Results

There are two binary classes in this research. The first binary class is for breast detection using the Adaboost and the second binary class in for nipple detection using the neural network.

5.3.1. Breast Detection

The data set is divided as in the table 3 using holdout method, and the resultant confusion matrix for the breast detection is shown in table 4.

Table 3. Training, Test and Validation Set for Breast Detection

	<i>Training</i>	<i>Validation</i>	<i>Test</i>
<i>Positive Images</i>	700	50	250
<i>Negative Images</i>	8000	500	1500

Table 4. The Resultant Confusion Matrix for Breast Detection

		<i>Predicted</i>	
		Negatives	Positives
<i>Actual</i>	Negatives	1478	22
	Positives	9	241



The evaluation parameters for the breast detection are calculated and presented in table 5. As shown in table 5, the accuracy of the breast detection is 98.2%.

Table 5. Performance of Breast Detection Classifier

<i>Accuracy (AC)</i>	<i>0.982</i>
<i>True Positive Rate or Recall (TP)</i>	<i>0.964</i>
<i>False Positive Rate (FP)</i>	<i>0.014</i>
<i>True Negative Rate (TN)</i>	<i>0.985</i>
<i>False Negative Rate (FN)</i>	<i>0.036</i>
<i>Precision (P)</i>	<i>0.916</i>

5.3.2. Nipple Detection

The data set is divided as in the table 6 using holdout method, and the resultant confusion matrix for the nipple detection is shown in table 7.

Table 6. Training, Test and Validation Set for Nipple Detection

	<i>Training</i>	<i>Validation</i>	<i>Test</i>
<i>Positive Images</i>	150	30	70
<i>Negative Images</i>	5450	50	100

Table 7. The Resultant Confusion Matrix for Nipple Detection

		<i>Predicted</i>	
		Negatives	Positives
<i>Actual</i>	Negatives	93	7
	Positives	3	67

The evaluation parameters for the nipple detection are calculated and presented in table 8. As shown in table 8; the nipple detection algorithm has 94.1% accuracy, and 90.5% precision.



Table 8. Performance of Nipple Detection Classifier

<i>Accuracy (AC)</i>	<i>0.941</i>
<i>True Positive Rate or Recall (TP)</i>	<i>0.957</i>
<i>False Positive Rate (FP)</i>	<i>0.07</i>
<i>True Negative Rate (TN)</i>	<i>0.93</i>
<i>False Negative Rate (FN)</i>	<i>0.04</i>
<i>Precision (P)</i>	<i>0.905</i>

5.3.3. Nipple Tracking

The nipples are tracked using the template matching method. Three types of template matching are evaluated in this section using the database-3. The accuracy of the nipple tracking algorithm was computed using the F-Score method.

The evaluation matrix of nipple tracking using the normed square difference (NSDTM) is presented in table 9. Based on the presented data in table 9, the accuracy of the nipple tracking using the NSDTM is 94.9%.

Table 9. The Evaluation Nipple Tracking Using NSDTM

Number of True Positives	723
Number of False Positives	77
Precision	0.903
Recall	1
F. Score	0.949

The evaluation matrix of nipple tracking using the normed correlation matching (NCTM) is presented in table 10. The NCTM method for nipple tracking has 95.8% accuracy.



Table 10. The Evaluation Nipple Tracking Using NCTM

Number of True Positives	736
Number of False Positives	64
Precision	0.92
Recall	1
F. Score	0.958

The evaluation matrix of nipple tracking using the normed correlation coefficient method (NCCTM) is presented in table 11. Based on the stated data in table 11, the NCCTM for nipple tracking has 97% accuracy.

Table 11. The Evaluation Nipple Tracking Using NCCTM

Number of True Positives	755
Number of False Positives	45
Precision	0.943
Recall	1
F. Score	0.970

Based on the evaluation of the nipple tracking methods, the normed correlation coefficient method was selected to use in the algorithm. As described in the evaluation performance, the mentioned tracking method has the highest accuracy among other methods.

5.3.4. Palpation Detection

The palpation detection algorithm was evaluated using the videos in dataset-4. For this evaluation, just one block in the image sequences was analyzed. There are 20 positive and 30 negative videos in this evaluation. Positive videos are those in which the palpation is done correctly, and negative videos are those in which palpation is not done correctly or at all. The experimental results for selecting the best threshold for the NormL2 values are shown in table 12. Based on the



evaluation, the best value for the thresholding is 100000. The resultant confusion matrix for the evaluation of palpation detection is shown in table 13. The evaluation parameters of palpation detection are calculated and presented in table 14. As described in table 14, based on the data in the confusion matrix shown in table 13, the palpation detection algorithm has 92% accuracy, and 86.3% precision. As shown in Fig. 32, based on the stated data in table 12, the accuracy of the thresholds 50000, 75000, 100000, and 125000 are 72%, 78%, 92%, and 86% respectively. So, in this research, the 100000 was selected to be used as the threshold for palpation detection.

Table 12. The Evaluation of Thresholding Values of NormL2 for Palpation Detection

Thresholding Values	Positive videos		Negative Videos	
	Correct Palpation Detection	Incorrect Palpation Detection	Correct Palpation Detection	Incorrect Palpation Detection
50000	20	0	16	14
75000	19	1	20	10
100000	19	1	27	3
125000	15	5	28	2

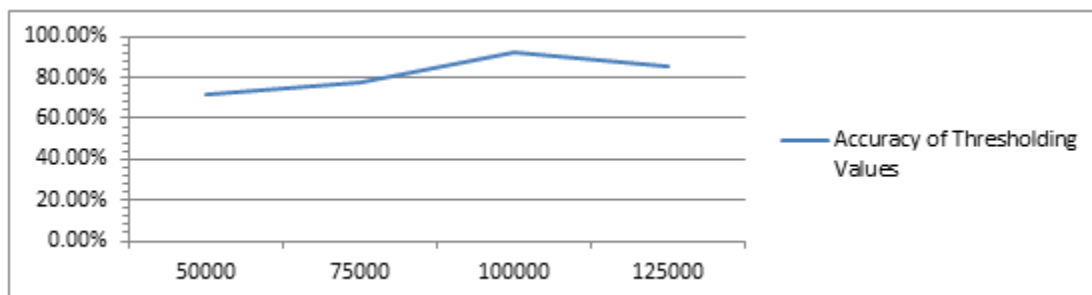


Figure 32. The Accuracy of Thresholding Values for Palpation Detection



Table 13. The Resultant Confusion Matrix for Palpation Detection

		<i>Detected</i>	
		Negatives	Positives
<i>Actual</i>	Negatives	27	3
	Positives	1	19

Table 14. Performance of Palpation Detection Evaluation

<i>Accuracy (AC)</i>	<i>0.92</i>
<i>True Positive Rate or Recall (TP)</i>	<i>0.95</i>
<i>False Positive Rate (FP)</i>	<i>0.1</i>
<i>True Negative Rate (TN)</i>	<i>0.9</i>
<i>False Negative Rate (FN)</i>	<i>0.05</i>
<i>Precision (P)</i>	<i>0.863</i>

5.3.5. Evaluation of the Integrated Algorithm

The integrated algorithm was created by combining of the individual algorithms (nipple detection and tracking, breast region tracking, and palpation detection). The input of the integrated algorithm is the image sequences that are taken by a webcam and the output is the highlighted blocks to show the palpated blocks and also an integer that shows the percentage score of palpated areas on the breast region.

The integrated algorithm was evaluated by five actual BSE trainings. So, there are five tests on the left breasts and five tests on the right breasts. The resultant confusion matrix for the evaluation of integrated algorithm is shown in table 15. The evaluation parameters of integrated algorithm are calculated. As shown in table 15, evaluation was performed by analyzing the palpated blocks in breast regions. As said previously, the matrix on the breast region has 6 blocks. During the BSE performance, all the blocks must be palpated for finding abnormalities



and lumps in the breast. After making palpation on each block, that block is highlighted to inform about palpated and non-palpated blocks.

Table 15. The Resultant Confusion Matrix for the Integrated Algorithm

		Detected					
	The Block Numbers on the Breast Region	1	2	3	4	5	6
Actual Palpation	1	9	0	0	0	0	1
	2	0	8	0	0	2	0
	3	0	0	9	1	0	0
	4	0	0	0	10	0	0
	5	0	0	0	0	10	0
	6	0	0	0	0	0	10

Based on the data in the confusion matrix (table 15), the accuracy of the integrated algorithm is 93.3%.

5.4. Experiments

In the actual BSE performance, the woman must sit in front of a webcam and perform the BSE. Both breasts must be visible on the screen before starting the BSE performance. The application would evaluate the BSE training at the real-time. So, the woman can see which areas of the breast region were palpated and which areas must be palpated at the rest of the BSE performance.

The first objective of this research is the nipple detection and tracking algorithm. In this algorithm, the breasts are detected using the Haar-Like features and cascaded adaboost. Then, the nipples are detected by analyzing the pixels and features inside the detected breast squares. The integral imaging is used to find ROIs including nipples. Then, the color moments, HOG, and pixel intensities are used to make a feature vector for each nipple's ROI. Then, by applying that feature vector in the neural network the nipple and non-nipple ROIs are



classified. Then, the extracted nipples are saved and used in template matching for nipple tracking. The result of the nipple detection is 94.1 % and the result of the nipple tracking has 97% accuracy.

The second objective of the research is related to breast region identification. A matrix with six blocks is located on each breast region. The size of the matrix on the breast region is computed by calculating the distance between the nipples and analyzing the breast squares that were detected by cascaded Adaboost. Then, the matrix on the breast region is tracked by tracking the opposite nipple. The accuracy of this objective is 98.2 %.

The third objective of the research was related to palpation detection on the blocks of breast regions. The Farneback optical flow is applied on each block sequentially. Then, the output matrix of optical flow is divided to four tiles. Then, the NormL2 is applied on each tile. If the value of NormL2 is bigger than 100000, that is considered as a palpation. If the palpation is done for 10 seconds on a tile with respect to the filter for reducing the noise, that is considered as a complete palpated tile. The percentage score is calculated for each block and for the entire breast region respectively. The accuracy of the third objective is 92%.

The integrated algorithm was created by combining the individual algorithms. So, the input of this algorithm is the image sequences that are captured by a webcam, and its output are the evaluated data that are monitored on the screen.

Initially, the integrated algorithm was tested on the manikin. In using the manikin, the picture of the breasts was printed on the sticker papers and labeled on the manikin. Then, the BSE was performed on the manikin. The manikin was used to test the different levels of thresholding values for palpation detection algorithm. As said previously, after having several experiments, the best value for the threshold was selected as 100000. I.e. all the NormL2 values, bigger than 100000, are considered as palpation on a tile with respect to time and filter. The manikin also used as a prototype for testing the integrated algorithm, before



testing the algorithm on actual models. While testing on the manikin, the algorithm was shown to have a high accuracy in evaluating the BSE performance in terms of checking the entire breast region.

The actual BSE test using the integrated algorithm was performed by 2 women in different luminance conditions. The results of the BSE performances were confirmed by Dr. Jose Ma. C. Avila. Dr. Avila's comments about the current progress and future works are presented in Appendix-1. The experimental results show that the integrated algorithm is robust in guidance and evaluation of BSE performance in different luminance conditions.

The actual tests were performed by a notebook and a webcam (HD-A4Tech) in real-time with 22 fps. The notebook's characteristics are described in table 16.

Table 16. The Notebook's Characteristics for the Actual Tests

<i>Operating System</i>	<i>Windows 7</i>
<i>System Type</i>	<i>64 bit</i>
<i>Processor</i>	<i>P 6000 @ 1.87GHz 1.87GHz</i>
<i>RAM</i>	<i>2.00 GB</i>
<i>IDE</i>	<i>Qt + OpenCV Library</i>

The examples of the results from the individual algorithms are shown in this section. Fig. 33 shows the results of breast detection using the adaboost algorithm. Fig. 34 shows the results of nipple detection using the neural network based on the breast detection using the adaboost algorithm. The results of the breast region tracking based on the tracking of the opposite nipple are shown in Fig. 35. The examples of the two consecutive blocks which were considered as the correct palpation are shown in Fig. 36.



Figure 33. The Examples of the Results for Breast Detection



Figure 34. The Examples of the Results for Nipple Detection

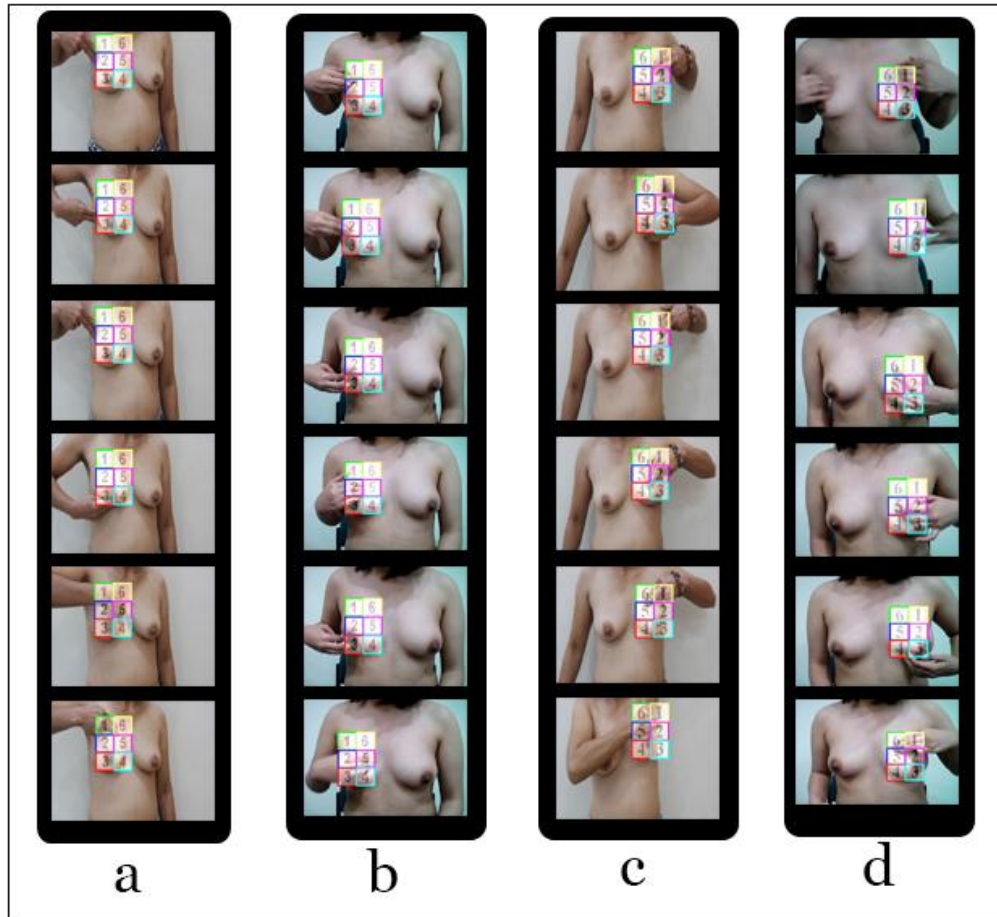


Figure 35. The Examples of Breast Region Tracking Based on Tracking the Opposite Nipple. a, b) Right Breast Region Tracking in the Actual Tests C, d) Left Breast Region Tracking in the Actual Tests

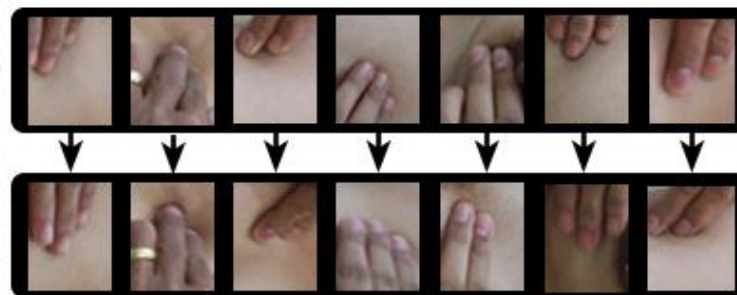


Figure 36. The Examples of Palpated Blocks in the Actual Tests



Chapter 6 Conclusions

6.1. Summary

Breast cancer is the most important cause of death among women worldwide. Early detection of breast cancer is the key of survival. The breast self-examination is one of the efficient methods for detecting the breast cancer in early stages. However, the majority of women do not perform the BSE at all or in the correct way. Therefore, there was a need for the application to educate and evaluate the BSE performance in real time.

The major goal of this research was to design and develop a computer vision algorithm for real-time evaluation of BSE in terms of covering and checking the entire breast regions. Women can perform BSE using the stated application, a webcam, and a computer to evaluate their BSE performance in real-time. If any abnormality, such as masses, is detected, then this must be reported to a doctor, who will confirm the presence of this abnormality or mass and proceed to do other confirmatory tests. Unfortunately, there is lack of literatures in the field of computer vision and BSE. So, the created algorithm, in this study, can be considered as the first step in this field. The algorithm has three main stages: Nipple detection and tracking, Breast region identification with blocks, and detecting the palpated blocks.

The Cascaded Adaboost Classifier, integral imaging, color moments, HOG features, and ANN are used for detecting the nipples. The template matching method is used to track the nipples in real-time. The second stage is focusing to identify the breast regions. A matrix is used to locate the breast region by tracking the opposite nipple. Due to the occlusion on the nipple at the time of BSE performance, the opposite nipple is used as the reference for tracking. The matrix



has 6 enumerated blocks for making palpation, and the palpation must be done on the stated blocks. The third stage focused to detect the palpated blocks in breast region. The motion detection method, farneback optical flow, is employed to detect the palpated blocks. After making palpation, each block is highlighted, and the score of covering the breast region is monitored on the screen. What women like about this application is that they can perform the BSE alone without any supervision. With this application they can comfortably check their own breasts for finding abnormalities.

Experimental results have shown that the BSE evaluation algorithm in this thesis provide robust performance. The nipple detection algorithm has the 94.1% accuracy, the nipple tracking algorithm has 97% accuracy, the breast region detection has 98.2% accuracy, and the algorithm for detecting the palpated blocks has 92% accuracy.

6.2. Comments for Future Work

Although an automated BSE evaluation have been developed in this thesis, there are a number of areas in which future research, based on the results presented here, can be explored.

- 1) As explained before, the nipple detection algorithm is the most important algorithm in this study. By enlarging the dataset for training the neural network and Adaboost, the accuracy of the nipple detection algorithm might be enhanced.
- 2) For future work, how to enhance system performance against partial occlusion should be considered. In this thesis work, the simple algorithm is used for the time of occlusion on nipples. In this study, the BSE for the left breast must be done with the left hand, and the BSE for the right breast must be done with



the right hand. But after solving the problem of occlusion, the BSE can be performed with no limitation of using left or right hand.

- 3) In this study, a matrix with 6 blocks is located to identify the breast region. While this appears to be valid, Dr. Avila's suggestion is that it would have been better if the grid be changed to a quadrant system, with the nipple as the center. This is the natural way doctors examine the breast and report findings. For example: the left breast with have four quadrants – the left upper outer, left upper inner, the left lower outer, and the left lower inner quadrants. The right breast would have something similar. Another possibility would be to further subdivide the breast like a clock face, with the again the nipple as the center of the clock. The breast is then divided into twelve areas like a clock. Areas involved by abnormality would then be described as “at around 3 o'clock” or “between 8 and 9 o'clock”. Everybody will understand this and can easily be transmitted from doctor to doctor with a clear understanding of location of the mass or abnormality.
- 4) In this study, the BSE is evaluated in terms of covering and checking the entire breast region. But, detection the level of the pressure on each block of the breast region is an important problem that should be solved in the future. Because a mass or abnormality might be existed in any depth of the breast.
- 5) Although, the accuracy of the palpation detection model is good, but it might be increased to higher accuracy by detecting the fingers as the preprocess step for applying optical flow. I.e. first, the fingers can be detected, then, the motion detection algorithm can be applied for finding the palpated blocks on the breast region.



Appendix A

Medical Doctor's Comments

The objective of the study is to teach the ordinary Filipino how to conduct breast self-examination by means of an original video, and application that has been produced. If any abnormality, such as masses, is detected, then this is reported to a doctor, who will confirm the presence of this abnormality or mass and proceed to do other confirmatory tests. I think that this objective must be stated very clearly.

The application tries to locate the presence of an abnormality by means of a six area grid of the breast.

While this appears to be valid, my suggestion is that it would have been better if the grid be changed to a quadrant system, with the nipple as the center. This is the natural way doctors examine the breast and report findings. For example: the left breast with have four quadrants – the left upper outer, left upper inner, the left lower outer, and the left lower inner quadrants. The right breast would have something similar.

Another possibility would be to further subdivide the breast like a clock face, with the again the nipple as the center of the clock. The breast is then divided into twelve areas like a clock. Areas involved by abnormality would then be described as “at around 3 o’clock” or “between 8 and 9 o’clock”. Everybody will understand this and can easily be transmitted from doctor to doctor with a clear understanding of location of the mass or abnormality.

Pendulous breasts which are common in older patients would be a problem with the present system because certain areas will not be represented equally in



the six area grid, particularly the lower areas. My suggestion would be to shoot the video with the patient lying down and supporting the breast somehow such that the grid areas would be almost the same throughout. This can be done through an uplifting bra or some contraption. This is something similar to that of a mammogram wherein the breast is squeezed into a space to sort of flatten it out, making examination of the entire breast easier.

Of course, these recommendations can be adopted in a second study phase, where a more precise localization of the breast lesion can be made the objective.

Jose Ma. C. Avila, M.D.
Pathologist



Appendix B

Create the Cascade of Haar-Like Classifier

Training Steps to Create the Haar-like Classifier has 4 stages as follows:

- 1) Gathering of positive and negative training samples
- 2) Marking positive samples (cropping the objects)
- 3) Creating a vector file based on positive marked samples using createsamples.exe
- 4) Training the classifier by means of haartraining.exe

B.1. Explanation of the Training Stages

B.1.1. Gathering Image Database

The positive images are those images that contain the object, and negatives are those which do not contain the object. Having more number of positive and negative images will normally cause a more precise classifier. In this thesis the number of positive samples is 1000, and the number of negative samples is 8000.

B.1.2. Arranging Negative Images

The negative samples must be copied to a specified folder, and then, the create_list.bat should be run. Running this batch file, the text file is created, and the lines of the text file would be as follow:

Img01.jpg

Img02.jpg

Img03.jpg

...

This negative data file is needed later, while training the classifier.



B.1.3. Crop the Positive Objects

In this stage, the data file (vector file) that contains the names of positive images as well as the location of the objects in each image must be created. The process of cropping the positive objects can be done using “objectmaker.exe”. The text file is created using “objectmaker.exe”. The data file contains information about the positive samples and location of objects in positive samples as follows:

```
rawdata\image200.bmp 1 24 42 24 54
rawdata\image201.bmp 2 25 45 80 99 30 45 90 92
rawdata\image202.bmp 2 20 34 90 90 45 88 90 82
```

The first number in each line expresses the number of objects in the given image. After the first number, each set of four numbers defines the location of the object in the given image.

B.1.4. Create the Vector of Positive Samples

In order to create the data file for positive samples, the “samples_creation.bat” must be run. This file reads the created info file in the previous section and makes the vector of positive images. For running the “samples_creation.bat”, the following DLL files must be copied to the folder of execution.

```
cv097.dll
cxcore097.dll
highgui097.dll
libguide40.dll
```

After running this file, the “facevector.vec” is created in the given folder.

B.1.5. Training

For the training the “haartraining.exe” must be run. Before training the “haartraining.bat” can be modified. The following data can be adjusted before training:

- Path and for storing the cascade of classifiers
- Path which points the location of vector file
- Path which points to background file
- Number of positive samples \leq no. positive bmp files
- Number of negative samples
- Number of proposed stages for training
- Quantity of memory assigned in MB
- Sample size

Then, the XML file should be created by combining all created stages (classifiers). This XML file would be the final file as “cascade of Haar-like classifiers”. The XML file can be made by running the “convert.bat” in the given folder.



References

- [1] Zhang J. Kesteloot HE, "Differences in breast cancer mortality worldwide: unsolved problems," *European Journal of Cancer Prevention*, vol. 15, no. 5, pp. 416–423, Oct 2006.
- [2] H. R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin Ferlay. J, "Cancer Incidence and Mortality Worldwide," GLOBOCAN, Lyon, France, Retrieved from <http://globocan.iarc.fr/> IARC Cancerbase No, 10, 2010.
- [3] Maira Caleffi, Ute-Susann Albert, Tony H. H. Chen, Stephen W. Duffy, Dido Franceschi, Lennarth Nystrom Smith. Robert A, "Breast Cancer in Limited-Resource Countries: Early Detection and Access to Care," *The Breast Journal*, vol. 12, no. 1, pp. 16-26, Jan-Feb 2006.
- [4] Abbie O. Beacham and Paul B. Jacobsen Michael A. Andrykowski, "Prospective, Longitudinal Study of Leisure-Time Exercise in Women with Early-Stage Breast Cancer," *American Association for Cancer Research*, vol. 16, no. 3, pp. 430-438, Mar 2007.
- [5] Katrina Armstrong, Constance D. Lehman, Suzanne W. Fletcher Joann G. Elmore, "Screening for Breast Cancer," *The Journal of the American Medical Association*, vol. 293, no. 10, pp. 1245-1256, Mar 2005.
- [6] D.M. Parkin, C. Ngelangel, D. Esteban, L. Gibson, M. Munson, M.G. Reyes and A. Laudico Pisani. P, "Outcome of screening by clinical examination of the breast in a trial in the Philippines," *Int. J. Cancer*, vol. 118, no. 1, pp. 149-154, Jan 2006.
- [7] A. Dueck, R. Gray, N. Wasif, M. Giurescu, R. Lorans, V. Pizzitola, and B. Pockaj Ma. I, "Clinical and Self Breast Examination Remain Important in the Era of Modern Screening," *Annals of Surgical Oncology*, vol. 19, no. 5, pp. 1484-1490, May 2012.
- [8] J.G. Elmore, J.P. Yi-Frazier, L.M. Reisch, N.V. Oster, and D.L. Miglioretti Roth. M.Y, "Self-detection remains a key method of breast cancer detection for U.S. women," *J Womens Health (Larchmt)*, vol. 20, no. 8, pp. 1135-1139, Aug 2011.
- [9] EA Paul Hackshaw AK, "Breast self-examination and death from breast cancer: a meta-analysis," *British Journal of Cancer*, vol. 88, no. 7, pp. 1047-1053., Apr 2003.
- [10] Lian-Hua Huang, Yeu-Sheng Hsieh, Ue-Lin Chung, Chiun-Sheng Huang, Herng-Dar Bih Yang Rea-Jeng, "Motivations and reasons for women attending a Breast Self-Examination training program: A qualitative study," *BMC Women's Health*, vol. 10, no. 23, pp. 10-23, Jul 2010.
- [11] P. Ramakant, E.R. Sanchez Forgach, J.C. Rendo, J. M. Chaparro, C. Sanchez, B.M. Margaritoni Agarwal. G, "Breast Cancer Care in Developing Countries," *World Journal Surgery*, vol. 33, no. 10, pp. 2069-2076, Oct 2009.
- [12] Broadwater G, Rabiner S, Owens E, Yoon S, Ghate S, Scott V, Walsh R, Baker J, Soo MS, Ibarra-Drendall C, Stouder A, Robertson S, Barron A, Seewaldt V Wilke LG, "Breast self-examination: defining a cohort still in need," *American Society of Breast Surgeons*, , vol. 198, no. 4, pp. 575-579, Oct 2009.
- [13] E., Rimer, B., Brogan, D., and Elon, L Frank, "US women physicians' personal and clinical breast cancerscreening practices," *Journal of Women's Health and Gender-Based Medicine*, vol. 9, no. 7, pp. 791-801, Sep 2000.
- [14] Mark Connerb, Rebecca Lawtonb, Wendy Baileyb, Jessica Litmanb, and Victoria Molyneauxb Andrew Prestwicha, "Individual and collaborative implementation intentions and the promotion of breast self-examination. ," *Taylor & Francis*, vol. 20, no. 6, pp. 743-



- 760, Feb 2007.
- [15] (2008, Aug) GLOBOCAN, International Agency for Research on Cancer. [Online]. HYPERLINK "<http://globocan.iarc.fr/factsheet.asp>" <http://globocan.iarc.fr/factsheet.asp>
- [16] I., and Miller, A. B Jatoui, "Why is breast-cancer mortality declining? ," *The Lancet Oncology*, vol. 4, no. 4, pp. 251-254, Apr 2003.
- [17] S.A. Amin, R.N.G. Naguib, A. Todman, H. Al-Omishy Oikonomou. A, "Breast Self Examination Training Through the Use of Multimedia: A Prototype Multimedia Application," in *IEEE EMBS*, Cancun, Mexico, 2003, pp. 1295-1298.
- [18] S.A. Amin, R.N.G. Naguib, A.G. Todman, H. Al-Omishy Oikonomou. A, "IRiS: An Interactive Reality Systemfor Breast Self-Examination Training," in *IEEE EMBS*, San Francisco CA, USA, 2004, pp. 5162-5165.
- [19] Sariego. J., "Breast cancer in the young patient.," *The American surgeon*, vol. 76, no. 12, pp. 1397-1401, Dec 2010.
- [20] American Cancer Society, "Breast Cancer Staging," *AJCC*, vol. 7, 2009.
- [21] Camara O, Kavallaris A, Krauspe S, Malarski N, Gajda M, Kroll T, Jörke C, Hammer U, Altendorf-Hofmann A, Rabenstein C, Pachmann U, Runnebaum I, Höffken K Pachmann K, "Monitoring the response of circulating epithelial tumor cells to adjuvant chemotherapy in breast cancer allows detection of patients at risk of early relapse," *J Clin Oncol*, vol. 26, no. 8, pp. 1208-1215, Mar 2008.
- [22] Ignatiadis M, Apostolaki S, Perraki M, Kalbakis K, Agelaki S, Stathopoulos EN, Chlouverakis G, Lianidou E, Kakolyris S, Georgoulas V, Mavroudis D. Xenidis N, "Cytokeratin-19 mRNA-positive circulating tumor cells after adjuvant chemotherapy in patients with early breast cancer," *J Clin Oncol*, vol. 27, no. 13, pp. 2177-2184, May 2009.
- [23] Lehman CD, Seger DJ, Buist DS, White E. Oestreicher N, "The incremental contribution of clinical breast examination to invasive cancer detection in a mammography screening program," *AJR Am J Roentgenol*, vol. 184, no. 2, pp. 428-432, Feb 2005.
- [24] Lichy J, Newhouse JH. Kolb TM, "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations," *Radiology*, vol. 225, no. 1, pp. 165-175, Oct 2002.
- [25] Helen M Zorbas, "Breast cancer screening," *Med J Aust*, vol. 178, no. 12, pp. 598-599., Jun 2003.
- [26] Thaler K, Chapman A, Kaminski-Hartenthaler A, Berzaczy D, Van Noord MG, Helbich TH Gartlehner G, "Mammography in combination with breast ultrasonography versus mammography for breast cancer screening in women at average risk," *Cochrane Database Syst Rev*, vol. pub2, p. CD009632, Apr 2013.
- [27] Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, Böhm-Vélez M, Mahoney MC, Evans WP 3rd, Larsen LH, Morton MJ, Mendelson EB, Farria DM, Cormack JB, Marques HS, Adams A, Yeh NM, Gabrielli G Berg WA1 and ACRIN 6666 Investigators, "Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk," *JAMA*, vol. 307, no. 13, pp. 1394-1404, Apr 2012.
- [28] Weigel S, Schrading S, Arand B, Bieling H, König R, Tombach B, Leutner C, Rieber-Brambs A, Nordhoff D, Heindel W, Reiser M, Schild HH Kuhl C1, "Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk



- of breast cancer: the EVA trial," *J Clin Oncol*, vol. 28, no. 9, pp. 1450-1457, Mar 2010.
- [29] Schrading S, Leutner CC, Morakkabati-Spitz N, Wardelmann E, Fimmers R, Kuhn W, Schild HH Kuhl CK1, "Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer," *J Clin Oncol*, vol. 23, no. 33, pp. 8469-8476, Nov 2005.
- [30] Majpruz V, Brown P, Thriault M, Shumak R, Mai V Chiarelli AM, "The contribution of clinical breast examination to the accuracy of breast screening," *J Natl Cancer Inst*, vol. 101, no. 18, pp. 1236-1243, Jun 2009.
- [31] Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW Elmore JG, "Ten-year risk of false positive screening mammograms and clinical breast examinations," *N Engl J Med*, vol. 338, no. 16, pp. 1089-1096, Apr 1998.
- [32] Gao DL, Ray RM, Wang WW, Allison CJ, Chen FL, Porter P, Hu YW, Zhao GL, Pan LD, Li W, Wu C, Coriaty Z, Evans I, Lin MG, Stalsberg H, Self SG Thomas DB, "Randomized Trial of Breast Self-Examination in Shanghai: Final Results," *Natl Cancer Inst*, vol. 94, no. 19, pp. 1445-1457, Oct 2002.
- [33] V. Cokkinides, O.W. Brawley Smith. R.A., "Cancer Screening in the United States, A Review of Current American Cancer Society Guidelines and Current," *CA Cancer J. Clin*, vol. 62, no. 2, pp. 129-142, Jan 2012.
- [34] Mojtabaii. A Khatib OMN, *Guidelines for the early detection and screening of breast cancer*, World Health Organization ed., Atord Modjtabei Oussama M.N. Khatib, Ed. Cairo, Egypt: EMRO Technical Publications Series ; 30, 2006.
- [35] A B Miller, C J Baines, and P N Corey B J Harvey, "Effect of breast self-examination techniques on the risk of death from breast cancer," *CMAJ*, vol. 157, no. 9, pp. 1205-1212., Nov 1997.
- [36] Pilgrim CA, Pennypacker HS Saunders KJ, "Increased proficiency of search in breast self-examination," *Cancer*, vol. 58, no. 11, pp. 2531-7., Dec 1986.
- [37] Naguib RN, Todman AG, Amin SA, Al-Omishy H, Oikonomou A, Tucker N Hu Y, "HAND MOTION SEGMENTATION AGAINST SKIN COLOUR BACKGROUND IN BREAST AWARENESS APPLICATIONS ," in *Conf Proc IEEE Eng Med Biol Soc*, CA-USA, 2004, pp. 3221-3224.
- [38] Robert Laganière, *OpenCV 2 Computer Vision Application Programming Cookbook*, Roger D'souza, Kavita Iyer, Neha Shetty Usha Iyer, Ed. Birmingham, UK: Packt Publishing Ltd, 2011.
- [39] (2012) OpenCV documentation. [Online]. HYPERLINK
"http://docs.opencv.org/modules/core/doc/intro.html"
<http://docs.opencv.org/modules/core/doc/intro.html>
- [40] Jasmin Blanchette and Mark Summerfield, "A Brief History of Qt," in *C++ GUI Programming with Qt 4.:* Prentice-Hall, June 2006, pp. xv-xvii.
- [41] Mark Summerfield, "Advanced Qt Programming ," in *Creating Great Software with C++ and Qt 4.:* Addison-Wesley, 2010.
- [42] Debnath Bhattacharyya, Tai-hoon Kim Jayanta Kumar Basu, "Use of Artificial Neural Network in Pattern Recognition," *International Journal of Software Engineering and Its Applications*, vol. 4, no. 2, pp. 23-34, Apr 2010.
- [43] S Haykin, *Neural Networks and Learning machines*. Delhi, India: Pearson Education, 2009.



- [44] Jianchang Mao, K. M. Mohiuddin Anil K. Jain, "Artificial Neural Networks: A Tutorial," *IEEE Computational Science & Engineering*, vol. 29, no. 3, pp. 31-44, Mar 1996.
- [45] Kohonen Teuvo, *Self-organizing maps*, T. Kohonen, M.R. Schroeder Thomas Huang, Ed. New York, USA: Springer, 2001.
- [46] Dadal. N, "Histogram of Oriented Gradients for Human Detection," in *Computer Vision and Pattern Recognition, IEEE*, San Diego, USA, Jun, 2005, pp. 886-893.
- [47] Dadal. N, "Finding People in Images and Videos," French National Institute for Research in Computer Science and Control (IN-RIA), PhD thesis Jul 2006.
- [48] Yixiao Yun, "Analysis and Classification of Objects Poses," CHALMERS UNIVERSITY OF TECHNOLOGY., Sweden, Master Thesis 2011.
- [49] Sabine Süsstrunk, Robert Buckley, and Steve Swen, "Standard RGB Color Spaces," *Color and Imaging Conference, 7th Color and Imaging Conference Final Program and Proceedings*, vol. 8, pp. 127-134, Jan 1999.
- [50] Gualberto Aguilar Torres, Gabriel Sánchez , Toscano Medina, Pérez Meana, Escamilla Hernadez Jorge Alberto Marcial Basilio, "EXPLICIT CONTENT IMAGE DETECTION," *SIPIJ*, vol. 1, no. 2, pp. 47-58, Dec 2010.
- [51] YUNG-KUAN CHAN and CHIN-CHEN CHANG, "A Color Image Retrieval Method Based on Color Moment and Color Variance of Adjacent Pixels," *Int. J. Patt. Recogn. Artif. Intell*, vol. 16, no. 1, pp. 113-125, Feb 2002.
- [52] J.G. Elmore, J.P. Yi-Frazier, L.M. Reisch, N.V. Oster, and D.L. Miglioretti Roth.M.Y, "Self-detection remains a key method of breast cancer detection for U.S. women," *J. Womens Health (Larchmt)*, vol. 20, no. 8, pp. 1135-1139, Aug 2010.
- [53] L. G. & Stockman, G. Shapiro, *Computer Vision*. NJ, USA: Prentice Hall, 2001, pp. 137-150.
- [54] Mark Nixon and Alberto S Aguado, *Feature Extraction and Image Processing for Computer Vision.*: Academic Press, 2008.
- [55] R.A. Haddad and A.N. Akansu, "A Class of Fast Gaussian Binomial Filters for Speech and Image Processing ," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 39, no. 3, pp. 723-727, Mar 1991.
- [56] Gary Bradski and Adrian Kaehler, "Template Matching," in *Learning OpenCV*, Mike Loukides, Ed. CA, USA: O'Reilly Media, 2008, ch. 7, pp. 214-219.
- [57] T. Mahalakshmi, R. Muthaiah and P. Swaminathan, "Review Article: An Overview of Template Matching Technique in Image Processing," *Research Journal of Applied Sciences*, vol. 4, no. 24, pp. 5469-5473, Dec 2012.
- [58] Coates, R. J., Uhler, R. J., Brogan, D. J., Gammon, M. D., Malone, K. E., Swanson, C. A., Flagg, E.W., and Brinton, L. A, "Patterns and predictors of the breast cancer detection methods in women under 45 years of age (United States)," *Cancer Causes and Control*, vol. 12, no. 5, pp. 431-442, Jun 2001.
- [59] J. P. Lewis, "Fast Template Matching," in *Canadian Image Processing and Pattern Recognition Society*, Quebec City, 1995, pp. 120-123.
- [60] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *American Statistician*, vol. 42, no. 1, pp. 59-66, 1988.
- [61] Brian G. Schunck Berthold K. P. Horn, "Determining Optical Flow," *ARTIFICIAL INTELLIGENCE*, vol. 17, pp. 185-204, 1981.



- [62] Maciej Smiatacz, "Liveness Measurements Using Optical Flow for Biometric Person Authentication," *Metrology and Measurement Systems*, vol. 19, no. 2, pp. 257–268, May 2012.
- [63] G Farnebeck, "Fast and Accurate Motion Estimation using Orientation Tensors and Parametric Motion Models," *15th International Conference on Pattern Recognition*, vol. 1, pp. 135-139, 2000.
- [64] Peter O'Donovan, "Optical Flow: Techniques and Applications," *The University of Saskatchewan*, pp. 1-24, Apr 2005.
- [65] Gunnar Farneback, "Two-Frame Motion Estimation Based on Polynomial Expansion," *Springer: Image Analysis, Lecture Notes in Computer Science*, vol. 2749, pp. 363-370, 2003.
- [66] H., Westin, C.F Knutsson, "Normalized and Differential Convolution: Methods for Interpolation and Filtering of Incomplete and Uncertain Data," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York City, USA, 1993, pp. 515-523.
- [67] G Farneback, "Polynomial Expansion for Orientation and Motion Estimation," Linköping University, Sweden, PhD thesis 91-7373-475-6, 2002.
- [68] Robert Krutsch and David Tenorio, "Histogram Equalization," Microcontroller Solutions Group Guadalajara, Tutorial 2011.
- [69] R. C. Gonzalez and R. E. Woods, *Digital Image Processing.*, 2008.
- [70] Michael Jones Paul Viola, "Rapid object detection using a boosted cascade of simple features," *Computer Vision and Pattern Recognition*, vol. 1, pp. 511-518, 2001.
- [71] Michael Jones Paul Viola, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [72] R., Maydt, J. Lienhart, "An extended set of Haar-like features for rapid object detection," *Image Processing, IEEE*, vol. 1, pp. 900-903, 2002.
- [73] M. Oren, and T. Poggio C. Papageorgiou, "A general frame-work for Object Detection," in *ICCV, IEEE Computer Society Washington*, 1998, p. 555.
- [74] Franklin C. Crow, "Summed-area tables for texture mapping," *SIGGRAPH*, vol. 18, no. 3, pp. 207-212, Jul 1984.
- [75] Robert E Schapire Yoav Freund, "A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119-139, Aug 1997.
- [76] Kegl Balazs. (2009) Introduction to AdaBoost.
- [77] Allen NE, Spencer EA, Travis RC Key TJ, "The effect of diet on risk of cancer," *Lancet*, vol. 360, no. 9336, pp. 861-868, Sep 2002.
- [78] Zujewski JA, Kamin L, Giordano S, Domchek S, Anderson WF, Bartlett JM, Gelmon K, Nahleh Z, Bergh J, Cutuli B, Pruneri G, McCaskill-Stevens W, Gralow J, Hortobagyi G, Cardoso F Korde LA, "Multidisciplinary Meeting on Male Breast Cancer: Summary and Research Recommendations," *Journal of Clinical Oncology*, vol. 28, no. 12, pp. 2114-2122, Mar 2010.
- [79] Susan B. Leight, Roy S. Nutter, Jr., Christy B. Schmidt, Douglas J Victor W. Leight, "Instrumented breast model," Grant US6669483 B1, Dec 30, 2003.
- [80] S.A. Amin, R.N.G. Naguib, A.Todman, H. Al-Omishy Oikonomou. A, "A prototype



- multimedia application for breast self examination training," in *IEMBS*, 2002, pp. 2623 - 2624 vol.3.
- [81] S. Amin, R.N.G. Naguib, A. Todman, H. Al-Omishy Oikonomou. A, "Breast self examination training through the use of multimedia: a benchmark multimedia development methodology for biomedical applications," in *EMBSS, IEEE*, Istanbul, Turkey, 2001, pp. 4020-4023.
- [82] S. Amin, R.N.G. Naguib, A. Todman, H. Al-Omishy. Oikonomou. A, "Breast self examination training through the use of multimedia: a comparison between multimedia development approaches," in *EMBS, IEEE*, Houston, TX, USA, 2002, pp. 2623-2624.
- [83] C. Qicai, R.N.G. Naguib, A. Oikonomou Shuyue. C, "Hand Pressure Estimation by Image Entropy for a Real-Time Breast Self Examination Multimedia System," in *Medicine and Biology 27th Annual Conference, IEEE*, Shanghai, China, 2005, pp. 1732-1735.
- [84] Richman AR, Brewer NT Salz T, "Meta-analyses of the effect of false-positive mammograms on generic and specific psychosocial outcomes," *Psychooncology*, vol. 19, no. 10, pp. 1026-1034, Oct 2010.
- [85] C. Qicai, R.N.G. Naguib, A. Oikonomou Shuyue. C, "Hand Pressure Detection Among Image Sequence In Breast Self-examination Multimedia System," in *IFTA*, Chengdu, 2009, pp. 127-130.
- [86] Computational Vision. [Online]. HYPERLINK "%20http://www.vision.caltech.edu/html-files/archive.html" <http://www.vision.caltech.edu/html-files/archive.html>
- [87] Real Self. [Online]. HYPERLINK "http://www.realself.com/breast-implants/before-and-after-photos" \l "tags=9951" <http://www.realself.com/breast-implants/before-and-after-photos#tags=9951>
- [88] Image Processing Place. [Online]. HYPERLINK "http://www.imageprocessingplace.com/root_files_V3/image_databases.htm" http://www.imageprocessingplace.com/root_files_V3/image_databases.htm
- [89] G Farneback, "Spatial Domain Methods for Orientation and Velocity Estimation," Linkopings University, Sweden, PhD Thesis 755, Mar 1999.
- [90] L. J. V. Gool, C. K. I. Williams, J. M. Winn, and A. Zisserman M. Everingham, "The pascal visual object classes," *IJCV*, vol. 88, no. 2, pp. 303-338, 2010.
- [91] S Oron, A. Bar-Hillel, D. Levi, and S. Avidan, "Locally orderless tracking," in *CVPR*, Providence, RI, Jun 2012, pp. 1940-1947.
- [92] J. Kwon and K. M. Lee, "Tracking of a non-rigid object via patch-based dynamic appearance modeling and adaptivebasin hopping monte carlo sampling," in *CVPR*, Miami Beach, Florida, USA, 2009, pp. 1208-1215.
- [93] A Smeulders et al., "Visual Tracking: An Experimental Survey," *Pattern Analysis and Machine Intelligence*, vol. PP, no. 99, p. 1, Nov 2013.
- [94] F. Bremond, M. Thonnat, and V. Valentin A. T. Nghiem, "Etiseo, performance evaluation for video surveillance systems," in *AVSS*, London, Sep 2007, pp. 476-481.
- [95] Nascimento, J.C; Marques, J.S., "Performance evaluation of object detection algorithms for video surveillance," *IEEE Transactions on Multimedia*, vol. 8, no. 4, pp. 761 - 774, Aug 2006.
- [96] Qing Chen, Georganas, Nicolas D. ; Petriu, E.M., "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features," in *Instrumentation and Measurement Technology*



- Conference Proceedings, Warsaw, May 2007, pp. 1-6.
- [97] Mita, T.; Kaneko, T. ; Hori, O., "Joint Haar-like features for face detection," in Computer Vision, 2005. ICCV 2005, Beijing, 2005, pp. 1619 - 1626.
 - [98] Goncalo Monteiro Paulo Peixoto , Urbano Nunes, "VISION-BASED PEDESTRIAN DETECTION," Robotica, pp. 45-50, Apr 2006.
 - [99] Liang Zhao, Thorpe, C.E., "Stereo- and neural network-based pedestrian detection," Intelligent Transportation Systems, IEEE Transactions on, vol. 1, no. 3, pp. 148 - 154, Sep 2000.
 - [100] Rowley, H.A.; Baluja, S. ; Kanade, T., "Neural network-based face detection," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 20, no. 1, pp. 23 - 38, Jan 1998.
 - [101] Qiang Zhu; Yeh, M.-C. ; Kwang-Ting Cheng ; Avidan, S., "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," in Computer Vision and Pattern Recognition, 2006, pp. 1491 - 1498.
 - [102] Yanwei Panga, Yuan Yuan, Xuelong Lib, Jing Panc, "Efficient HOG human detection," Signal Processing, vol. 91, no. 4, pp. 773-781, Apr 2011.
 - [103] Yue Wang, Jun Li, HeeLin Wang, and ZuJun Hou, "Automatic Nipple Detection Using Shape and Statistical Skin Color Information," Springer-Verlag Berlin Heidelberg, pp. 644-649, 2010.
 - [104] Piyabute Fuangkhan; Thitipong Tanprasert, "Nipple detection for obscene pictures," SSIP'05 Proceedings of the 5th WSEAS international conference on Signal, speech and image processing, vol. 32, no. 7, pp. 242-247 , 2005.