

國立臺灣大學電機資訊學院資訊工程學系

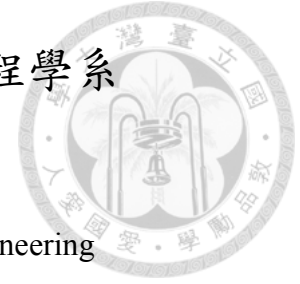
碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis



使用材質資訊於全乳房自動超音波的腫瘤偵測

Tumor Detection Based on Texture Information in  
Automated Breast Ultrasound

林士淳

SHIH-CHUN LIN

指導教授：張瑞峰 博士

Advisor: RUEY-FENG CHANG, Ph.D.

中華民國 106 年 1 月

January, 2017



國立臺灣大學碩士學位論文  
口試委員會審定書

使用材質資訊於全乳房自動超音波的腫瘤偵測

Tumor Detection Based on Texture Information in  
Automated Breast Ultrasound

本論文係林士淳君（學號 R02922082）在國立臺灣大學資訊工程學系完成之碩士學位論文，於民國 104 年 7 月 29 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

張瑞峰

陳啟禎 (指導教授)

羅榮錦

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

趙坤茂

系主任

## 致謝



由於中間去了日本交換學生一年，我的碩士生涯比其他同屆的同學們還要長，謝謝這一路上有著許多人的指導與協助，讓我度過很充實的碩士生涯。

首先要感謝我的指導教授張瑞峰教授，因為我的任性之下跑去實習，也讓我體驗一年交換生活，在研究上給予用心指導，遇到困難與瓶頸時，老師總能適時的給予建議與方向，使我的實驗能夠順利完成，真的很謝謝老師。

接著感謝帶我的學長陳榮泰，在論文研究剛起步時有你的帶領，對於寫程式前的架構環境才能夠順利。實驗的過程以及最後的論文撰寫，因為學長的耐心指導與引領，讓我收起慌張的心情順利完成論文，總是受到你的照顧，辛苦了。鴻豪學長在我論文撰寫的黑暗時期，在數據分析上給予我指導，也給了我不少鼓勵，學長們的一句『快畢業了』，對我來說是最有效的強心針！耀賢學長總會在實驗室待到晚上，在我求救無援時給我許多指導方針。衍緯學長總是在我身旁給我鼓勵，崇銘學長感謝你讓我口試過關，小伶學姊的一句『沒問題的』給予我努力的動力。傑帆學長在我無助時給我許多研究上的方向與諮詢。在碩一時碩二學長們，有你們我的碩一生活既歡樂又難忘，尤其是懿升，謝謝你在我壓力大時陪我聊天也給予我許多研究上的建議。也感謝現在碩二的學弟妹們，讓我最後撰寫論文的生活並不孤單。

感謝同屆的同學們（鄭琮錡、楊名揚、朱于萱、方詩敏、李雅雯）以及崧瑋學長，我們是一同走過來的戰友，謝謝你們的陪伴與共同奮戰，尤其是我的好鄰居琮錡，在我回歸實驗室孤軍奮戰時仍不離不棄給予我過來人的建議，謝謝你。

最後要謝謝我的家人，支持我做的決定，謝謝你們忍受我的固執，鼓勵我勇敢去冒險。感謝女友ありさ，時常忙到太晚無法聯絡，謝謝妳一直以來的應援，之後也請多多指教！想要感謝的人太多了，真的很感謝大家的鼓勵與幫助，我愛你們！

## 摘要



全乳房自動超音波是一種新的乳房腫瘤偵測儀器，與傳統手持式超音波相比，全乳房自動超音波具有幾項優點，包括更高的再現性和更少的操作者依賴性。然而，一組全乳房自動超音波影像包含數百張二維影像，若逐張檢查，醫師需要花費很多時間，因此電腦輔助乳房腫瘤偵測系統被設計來解決檢查全乳房自動超音波影像的耗時問題。先前一個基於分水嶺切割的電腦輔助乳房腫瘤偵測系統能夠達到腫瘤偵測率 100%，其伴隨著一些非腫瘤區域被誤判為腫瘤。由於大部分的腫瘤位於乳腺區域的內部和附近，乳房的解剖學資訊可以被使用於電腦輔助乳房腫瘤偵測系統的偽陽性篩除上。因此，本論文提出一個基於解剖學的偽陽性篩除模型以提升先前電腦輔助乳房腫瘤偵測系統的效能，利用每一個區域中量化的空間和材質特徵設計一個邏輯迴歸模型來從全乳房自動超音波影像取出腺體區域，利用電腦輔助乳房腫瘤偵測系統所偵測出來的腫瘤區域與腺體區域作疊合，再根據疊合率使非腫瘤區域被誤判成腫瘤的個數有所下降。實驗中所用到的測試資料包含 43 個良性以及 39 個惡性腫瘤。根據結果，本系統在腫瘤偵測率為 100% (82/82), 90% (74/82), 80% (66/82) 的情況下，平均每一個全乳房自動超音波掃描會有 3.15, 1.90, 1.42 個非腫瘤區域被誤判成腫瘤，與先前的電腦輔助乳房腫瘤偵測系統相比非腫瘤區域被誤判成腫瘤的情形下降了 67.49%。總結，本論文所提出之基於解剖學的偽陽性篩除模型，在針對全乳房自動超音波掃描作腫瘤偵測時，能夠在不影響腫瘤偵測率下有效地降低非腫瘤區域被誤判成腫瘤的情形。

關鍵詞：全乳房自動超音波、乳癌、電腦輔助乳房腫瘤偵測、偽陽性篩除、多元邏輯斯迴歸、組織分類、分水嶺切割

## Abstract



Automated breast ultrasound (ABUS) is a new screening modality. Compared to the conventional handheld ultrasound, ABUS has several advantages, including higher reproducibility and less operator dependence. However, an ABUS image volume includes hundreds of 2-D slices, and the physicians need a large amount of time to review the volume image. Hence, the computer-aided detection (CAdE) systems were proposed to solve the long-time review issue. A CAdE system based on watershed transform can achieve the sensitivity of 100% with several false positives (FP) in the previous study. Because most tumors are located inside or close to the mammary gland regions, the information of breast anatomy can be used for reducing the FPs of CAdE. Therefore, this study proposed an anatomy-based false positive reduction (FPR) method to decline the FPs/pass with the same sensitivity as the previous CAdE system. A logistic regression model using quantitative spatial and texture features was designed to extract glandular regions from ABUS images. The FPs would be reduced according to the overlapping ratio between glandular regions and suspicious regions detected by CAdE system. The collected database comprised 43 benign and 39 malignant tumors. As a result, the proposed anatomy-based FPR achieved sensitivities of 100% (82/82), 90% (74/82), 80%

(66/82) with FPs/pass of 3.15, 1.90, and 1.42, respectively. A decrease of 67.49% in FPs/pass was achieved in comparison to the previous CADe system. In conclusion, the proposed anatomy-based FPR could reduce FPs with the same tumor detection rate in ABUS images.

*Keywords:* Automated breast ultrasound, breast cancer, computer-aided detection, false positive reduction, multinomial logistic regression, tissue classification, watershed segmentation.

# Table of Contents



口試委員會審定書 .....	i
致謝 .....	ii
摘要 .....	iii
Abstract .....	iv
Table of Contents .....	vi
List of Figures .....	vii
List of Tables .....	ix
Chapter 1 Introduction .....	1
Chapter 2 Materials.....	4
Chapter 3 ABUS Tumor Detection Based on Texture Information .....	9
3.1 CADe.....	13
3.2 Anatomy-based FPR .....	17
3.2.1 Quantitative Feature Extraction.....	18
3.2.1.1 Spatial Features .....	20
3.2.1.2 Texture Features.....	21
3.2.2 Glandular Extraction.....	23
3.2.3 Post Processing .....	28
3.2.4 TIGR Calculation .....	31
Chapter 4 Experimental Results .....	32
Chapter 5 Discussion and Conclusion .....	46
Reference .....	49

## List of Figures



- Fig. 1. ABUS image from a 35-year-old woman with a fibroadenoma. (a) A series of continuous 2-D slices in axial plane. (b) The axial, sagittal and coronal planes of the ABUS image. The red circles indicate the tumor position. .... 8
- Fig. 2. The red lines illustrate the classification classes. (a) mammary glandular tissue (b) subcutaneous fat (c) retromammary fat & pectoralis muscle (d) shadow regions ..... 11
- Fig. 3. The architecture of the proposed CADe system and anatomy-based FPR..... 12
- Fig. 4. Topographic view of watershed transform..... 13
- Fig. 5. The tumor detection result of the CADe system. (a) The ABUS image with a tumor circled. (b) A tumor candidate generated by the CADe system. .... 14
- Fig. 6. The red circles indicate the tumor candidate region chosen by CADe system. Because of low echogenicity, the region of nipple was detected as tumor..... 15
- Fig. 7. The red circles indicate the tumor candidate region chosen by CADe system. Because of low echogenicity, the region of shadow was detected as tumor. .... 16
- Fig. 8. The image showed the location of each region. (a) Original ABUS image (b) Illustration of the glandular tissue, subcutaneous fat, retromammary fat & pectoralis muscle, and shadow in the image. .... 20
- Fig. 9. The distance and direction used in GLCM method. The pixel pair of four directions from the centered pixel( $\odot$ ). Pixel 1 to 4 are the neighboring pixels in the direction of  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  with  $d=1$ , respectively. .... 22
- Fig. 10. The process of building logistic regression model. The left-up side is an original slice of training data. The slice being cut into patches were classified into 4 classes defined by the expert for training of the logistic regression model..... 24
- Fig. 11. The result after applying image classification. (a) The original ABUS image. (b) The classification result of (a) after applying logistic regression model. The image was cut into  $11 \times 11$  patches, and the patches were classified into four classes as shown in the image..... 27
- Fig. 12. The result after applying glandular extraction and post processing. The red circle indicates the tumor region specified in pathology report. The green circle indicates the non-glandular patches which were misclassified into glandular patches. The blue circle indicates the glandular patches which were misclassified into non-glandular patches. (a) The original ABUS image with a tumor circled with



red line. (b) The result after image classification to (a). (c) The result after post processing to (b). .....	30
Fig. 13. The variation graph of the sensitivity with <i>TIGR</i> as a threshold. Using the testing data set in this study, the sensitivity was 100% until the threshold of <i>TIGR</i> was 43%. Red line indicates the <i>TIGR</i> of 43%.....	39
Fig. 14. The free-response receiver operating characteristic (FROC) curve for CADe system before and after the anatomy-based FPR was implemented. ....	39
Fig. 15. True positive case of 2.0 cm invasive ductal carcinoma. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. It was specified in pathology report. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the <i>TIGR</i> of 92.31%.....	41
Fig. 16. True positive case of 1.2 cm fibroadenoma. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. It was specified in pathology report. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the <i>TIGR</i> of 100%. .....	42
Fig. 17. True positive case of 2.0 cm fibroadenoma. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. It was specified in pathology report. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the <i>TIGR</i> of 66.51%. .....	43
Fig. 18. False positive case. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was excluded with the <i>TIGR</i> of 0%.....	44
Fig. 19. False positive case. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the <i>TIGR</i> of 81.01%.....	45

## List of Tables

Table 1. Lesion sizes and BI-RADS categories with histological diagnoses.....	6
Table 2. The description of the adopted quantitative features.....	19
Table 3. The performance of the leave-one-out cross-validation with different patch size. The amount of training data is 1000 patches for each class. That is, totally 4000 patches used as training data. ....	36
Table 4. The performance of the leave-one-out cross-validation with different amount of data. The patch size was $11 \times 11$ .....	36
Table 5. The median value, and p-value of Mann-Whitney U test for various features in glandular extraction. ....	37
Table 6. The number of tumor candidates mistakenly excluded for different sizes of benign and malignant tumors at different sensitivity.....	38
Table 7. The number of tumor candidates mistakenly excluded in different sensitivity rates.....	38
Table 8. The average number of FPs/pass before and after applying the anatomy-based FPR. ....	38

# Chapter 1

## Introduction



In 2016, cancer is a public health problem worldwide, and breast cancer is the second leading cause of death in females aged from 20 to 59 years in the United States [1]. To reduce mortality rates, early detection and treatment have been shown to be important [2]. Mammography [3, 4] and ultrasound (US) [5, 6] are popular screening modalities for breast tumor diagnosis and detection. Conventionally, mammography is the most effective modality used in clinical examinations. Nevertheless, due to the low specificity of mammography, it is nearly impossible to detect breast cancer in adolescent women with dense breasts [7, 8]. US is an adjunct imaging modality owing to the advantages such as observation in real-time, portability and relatively high screening sensitivity in dense breasts. The US examination is radiation-free and makes patients suffer from less emotional pressure.

However, conventional US images are acquired by hand-held probes. The quality of the generated image is operator dependent and poorly reproducible. To overcome the disadvantages of conventional US examination, the automated breast ultrasound (ABUS) imaging systems have been developed to automatically scan whole breasts [9-12].

Nonetheless, a three-dimensional (3-D) ABUS image is composed of more than 300 two-dimensional (2-D) slices. It is time-consuming to review a great amount of images to discover suspicious abnormalities. Thus, computer-aided detection (CAdE) systems of ABUS images have been proposed to assist radiologists to accelerate the reviewing procedure [13-18].

This study was proposed based on previous CAdE system using topographic watershed [19]. The likelihoods of being tumors were evaluated using the quantitative morphology, intensity, and texture features. The regions with low echogenicity were extracted as suspicious regions. Finally, 2-D and 3-D false positive reduction (FPR) model have been applied, which composed of 2-D FPR followed by 3-D FPR. However, this CAdE system did not perform well enough for all types of breast tumors. Although the 2-D/3-D FPR combined various categories of feature sets to reduce non-tumors from the tumor candidates, the information of breast anatomy was not considered. In ABUS image, some regions like nipple and shadow can cause acoustic shadowing. Sometimes those regions were remained with high likelihood as tumor regions. Besides, breast cancer often starts in glandular tissue. It occurs when cell division in glandular tissue goes wrong and begins to clone out of control [20]. Hence, it would be helpful for tumor detection if

the glandular tissue could be extracted. In fact, the ABUS image shows discrete structures being recognized. It is possible to use anatomical information in tumor detection.



Therefore, an anatomy-based FPR was proposed to combine with CAdE system. A logistic regression model using quantitative spatial and texture features was designed to extract glandular regions from ABUS images in this study. The quantitative spatial and texture features were used to train the logistic regression model due to resemblances of the information of breast anatomy. Then, the FPs would be reduced according to the overlapping ratio between glandular regions and suspicious regions detected by CAdE system.

## Chapter 2


### Materials



The ABUS images used in this research were acquired from the Breast Center of National Taiwan University Hospital between July 2012 and December 2014. An ACUSON S2000 Automated Breast Volume Scanner (Siemens Medical Solutions, Mountain View, CA, USA) with a 14L5BV linear array transducer, which was ranged from 5 to 15 MHz.

Depending on the size of each breast, up to five views (anterior-posterior, medial, lateral, superior, and inferior) are taken to cover the whole breast volume. In this study, the common three passes including anterior-posterior, medial and lateral were performed in order to completely scan a breast. Each scanning pass has 318 two-dimensional (2-D) slices with the thickness of 0.5mm, and the spacing are 0.0212, 0.007 and 0.052 cm/pixel for row, column and slice respectively.

In the experiment, 136 passes of ABUS image volumes were acquired from 51 women (age range: 22-71 years, mean:  $48.63 \pm 12.20$  years), which were collected for estimating the performance of the anatomy-based FPR. According to the examination reports, there were 82 abnormal passes with at least one tumor and 54 normal passes. 82

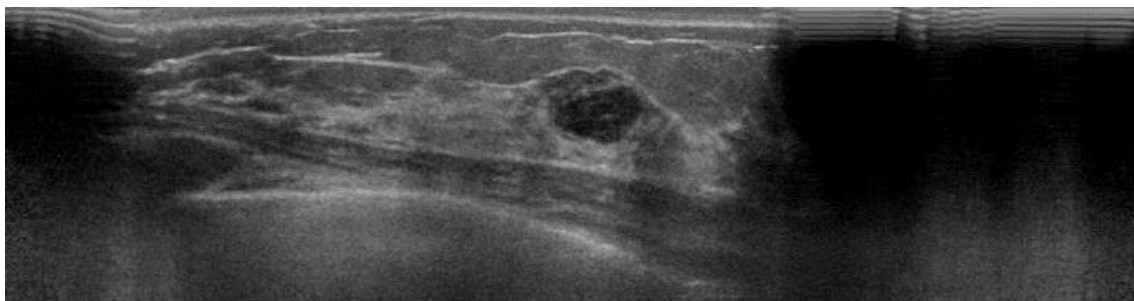
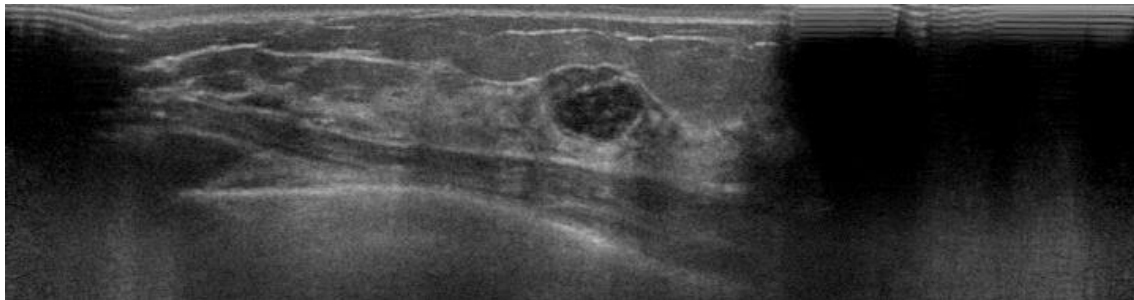
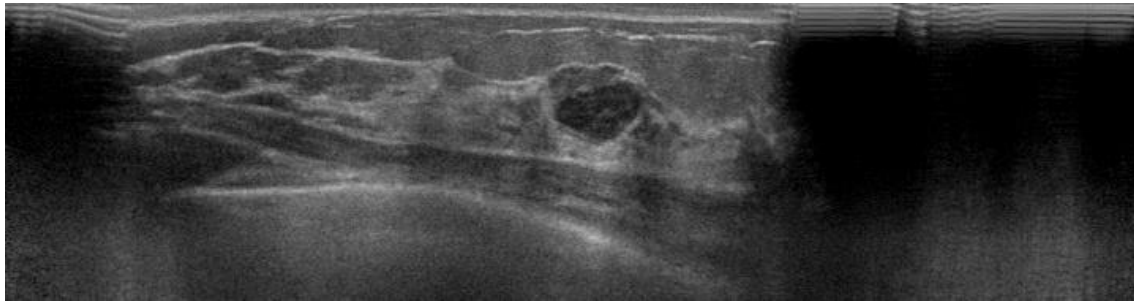
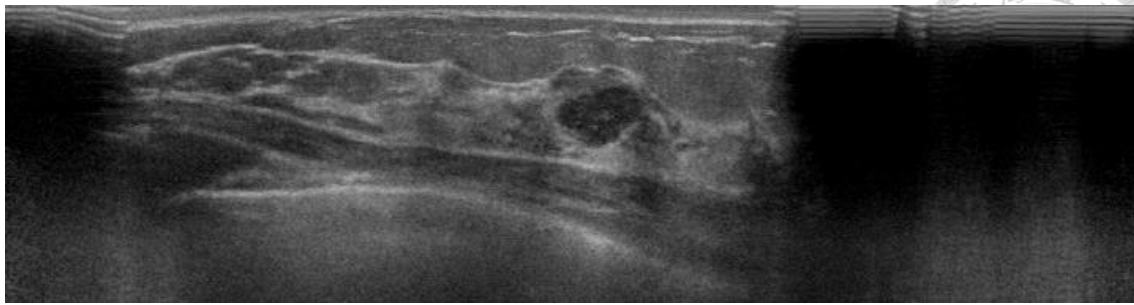


abnormal passes had 82 biopsy-proven lesions (size range: 0.4-4.0 cm, mean:  $1.4341 \pm 0.5911$  cm), including 43 benign lesions and 39 malignant lesions. The lesion sizes and the BI-RADS categories of collected databases are listed in Table 1. The 43 benign lesions were composed of 7 fibroepithelial lesion, 9 fibrocystic changes, 12 papilloma, 14 fibroadenoma and 1 sclerosing adenosis. The 40 malignant lesions were composed of 37 invasive ductal carcinoma (IDC) and 2 mucinous carcinoma. The approval and informed consent for this study were obtained from the institutional review board. Fig. 1(a) shows a series of continuous 2-D slices in an ABUS volume. Fig. 1(b) shows a tumor that was observable in the axial, sagittal and coronal views.

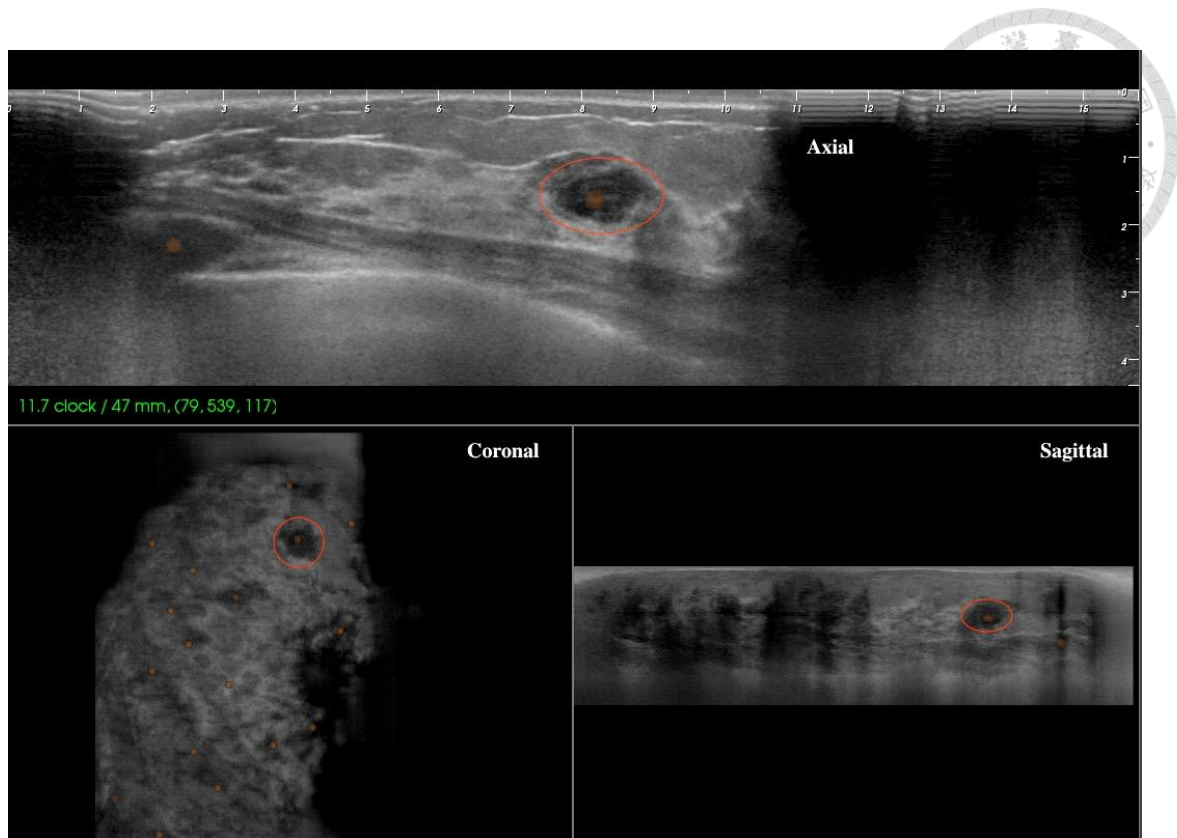
Table 1. Lesion sizes and BI-RADS categories with histological diagnoses.

Characteristic	No. of Cases		
	Benign	Malignant	Total
Lesion Size (cm)			
<1.0	15	0	15
1.0-2.0	23	31	54
2.0-3.0	4	7	11
>3.0	1	1	2
BI-RADS			
2	3	0	3
3	9	0	9
4A	23	6	29
4B	4	7	11
4C	4	26	30





(a)



(b)

Fig. 1. ABUS image from a 35-year-old woman with a fibroadenoma. (a) A series of continuous 2-D slices in axial plane. (b) The axial, sagittal and coronal planes of the ABUS image. The red circles indicate the tumor position.

## Chapter 3

### ABUS Tumor Detection Based on Texture Information



Nowadays, CADe system has been developed for tumor detection in ABUS images.

A previous study applied watershed transform to CADe system by separating different tissues into regions [19]. After doing the watershed transformation, the regions having high likelihood of being tumors would be remained, and the regions like muscle, shadow, skin and rib due to the intensity rank. Then, the 2-D/3-D FP reduction was adopted by using morphology, intensity and texture as features to distinguish between tumors and non-tumors, and the remaining regions were tumor candidates.

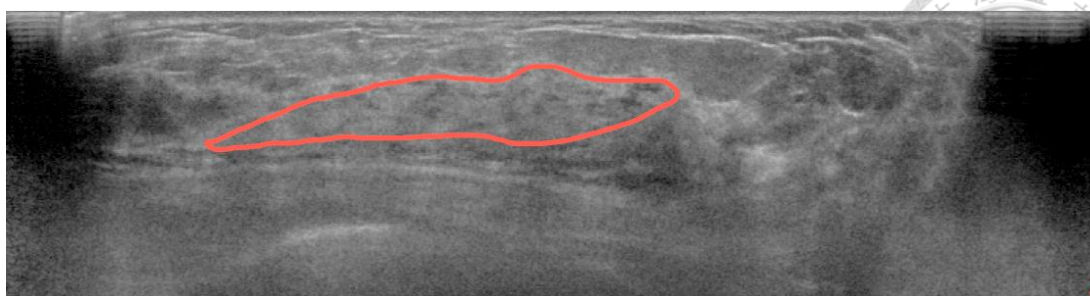
As a matter of fact, most of the breast cancers are carcinomas, a type of cancer starting in the epithelial cells. Moreover, a type of carcinomas often occurs called adenocarcinoma, which is carcinoma that begins to clone from glandular tissue. Therefore, if the overlapping ratio between the tumor candidate and the mammary gland tissue is high, the tumor candidate has high likelihood of being real tumors.

The proposed FPR based on the anatomy information in order to eliminate FPs from the tumor candidates generated by the previous CADe system. In this study, a logistic regression model was developed to classify the breast tissue. ABUS images were

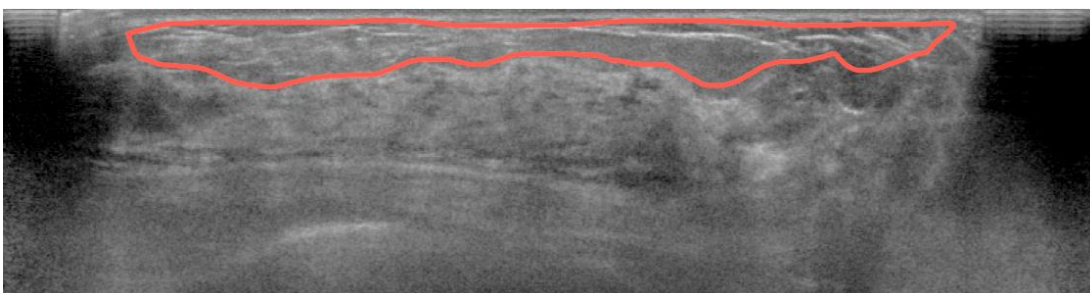
classified into four classes based on texture and spatial features including mammary glandular tissue, subcutaneous fat, retromammary fat & pectoralis muscle, and shadow as shown in Fig. 2. The glandular tissue would be extracted after combining the non-glandular regions. To judge whether the tumor candidates generated by CAdE system were FPs or not, the ratio of the tumor candidate region inside of glandular tissue (*TIGR*) was considered to be meaningful. The architecture of the proposed FPR is shown in Fig.

3.

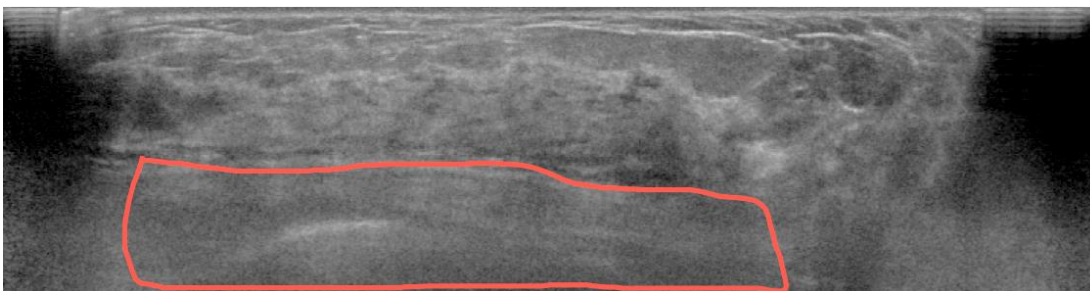




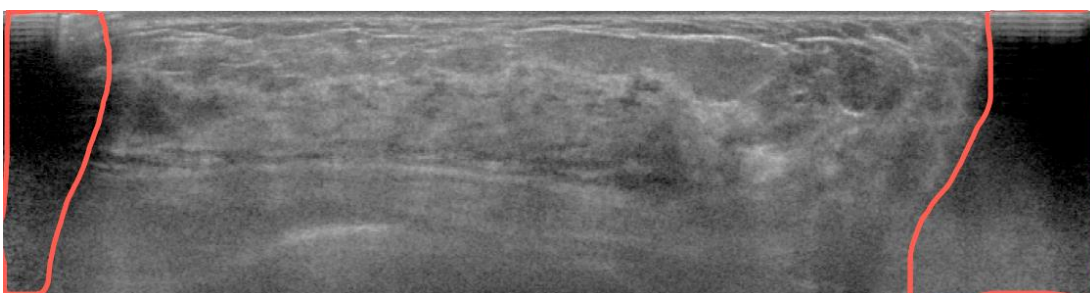
(a)



(b)



(c)



(d)

Fig. 2. The red lines illustrate the classification classes. (a) mammary glandular tissue (b) subcutaneous fat (c) retromammary fat & pectoralis muscle (d) shadow regions

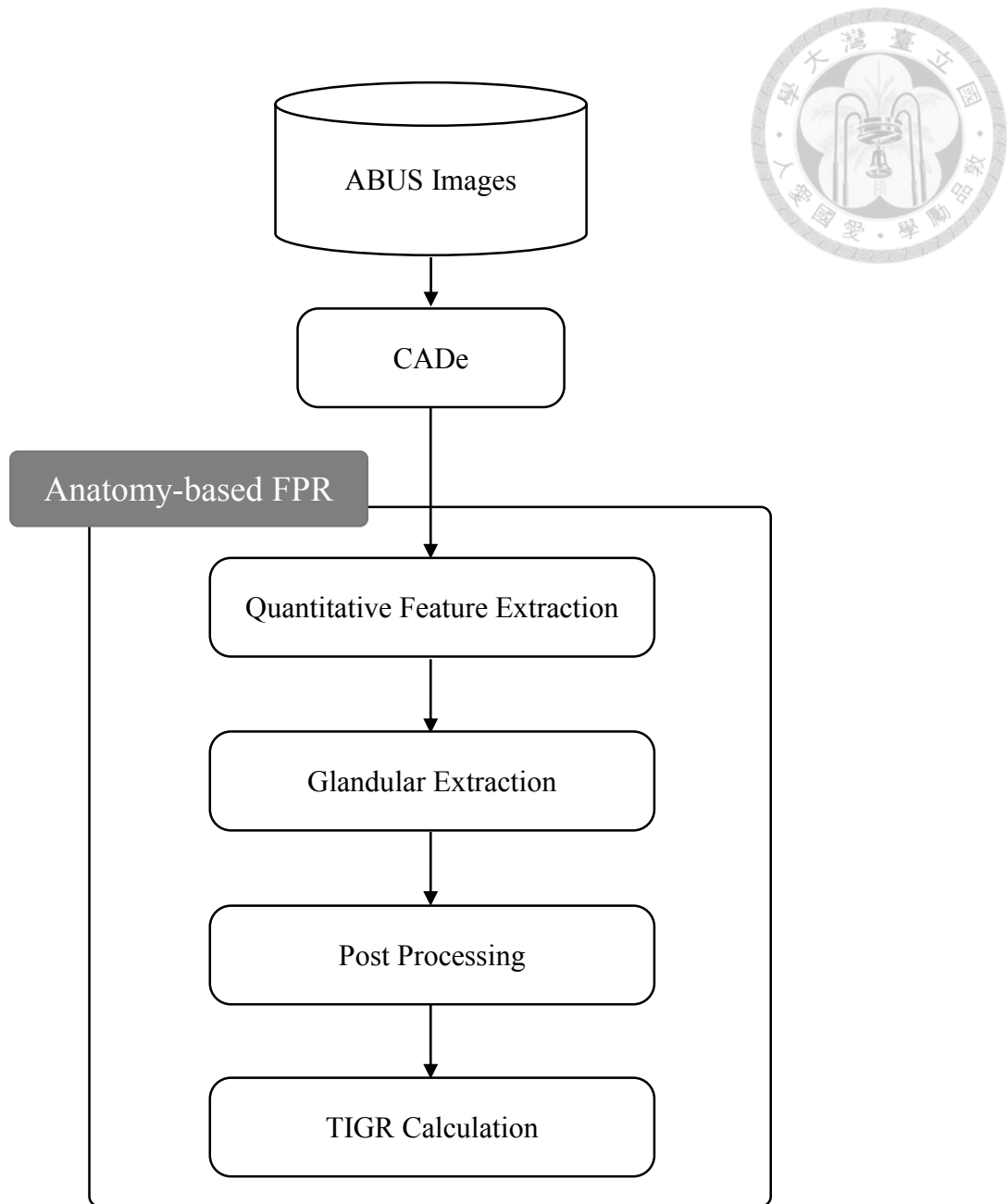


Fig. 3. The architecture of the proposed CADe system and anatomy-based FPR.



### 3.1 CADe

The proposed anatomy-based FPR was implemented on tumor detection results of the CADe system using topographic watershed [19]. The watershed transform gather similar tissues to be homogeneous regions according to their echogenicity as shown in Fig. 4. After the tumor detection, the remaining regions were tumor candidates as shown in Fig. 5.

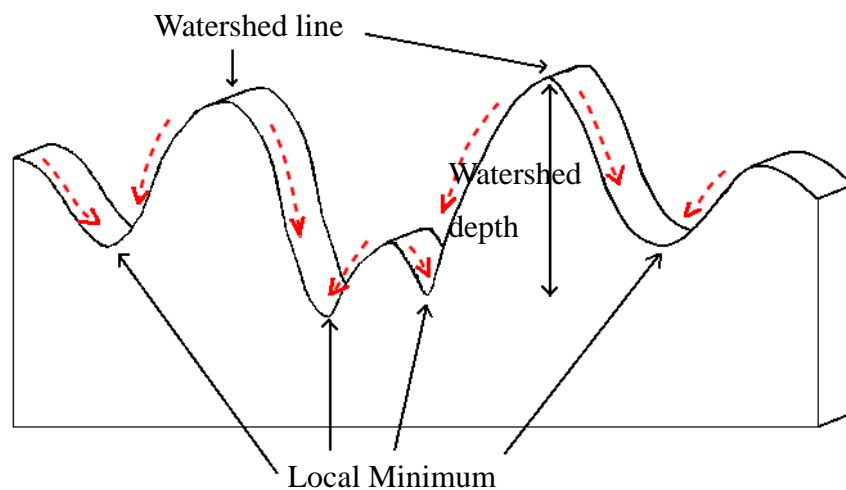
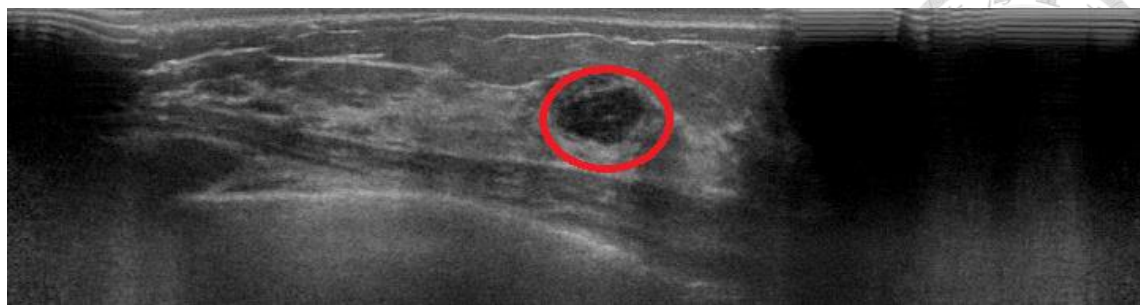


Fig. 4. Topographic view of watershed transform.



(a)



(b)

Fig. 5. The tumor detection result of the CAdE system. (a) The ABUS image with a tumor circled. (b) A tumor candidate generated by the CAdE system.



The CADe system detected potential carcinomas using watershed transform. Although 2-D/3-D FPR had been adopted in CADe, regions with low echogenicity would be judged as tumor candidates, including regions of nipple and shadow as shown in Fig. 6 and Fig. 7 respectively.

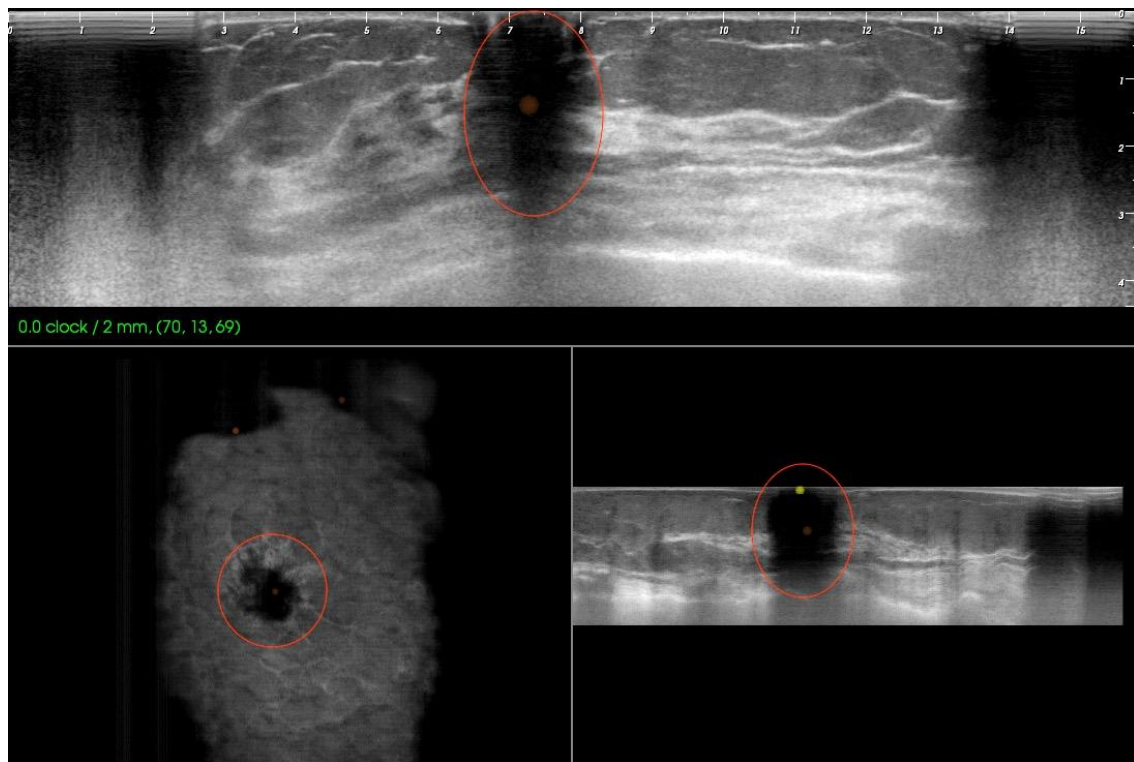


Fig. 6. The red circles indicate the tumor candidate region chosen by CADe system. Because of low echogenicity, the region of nipple was detected as tumor.

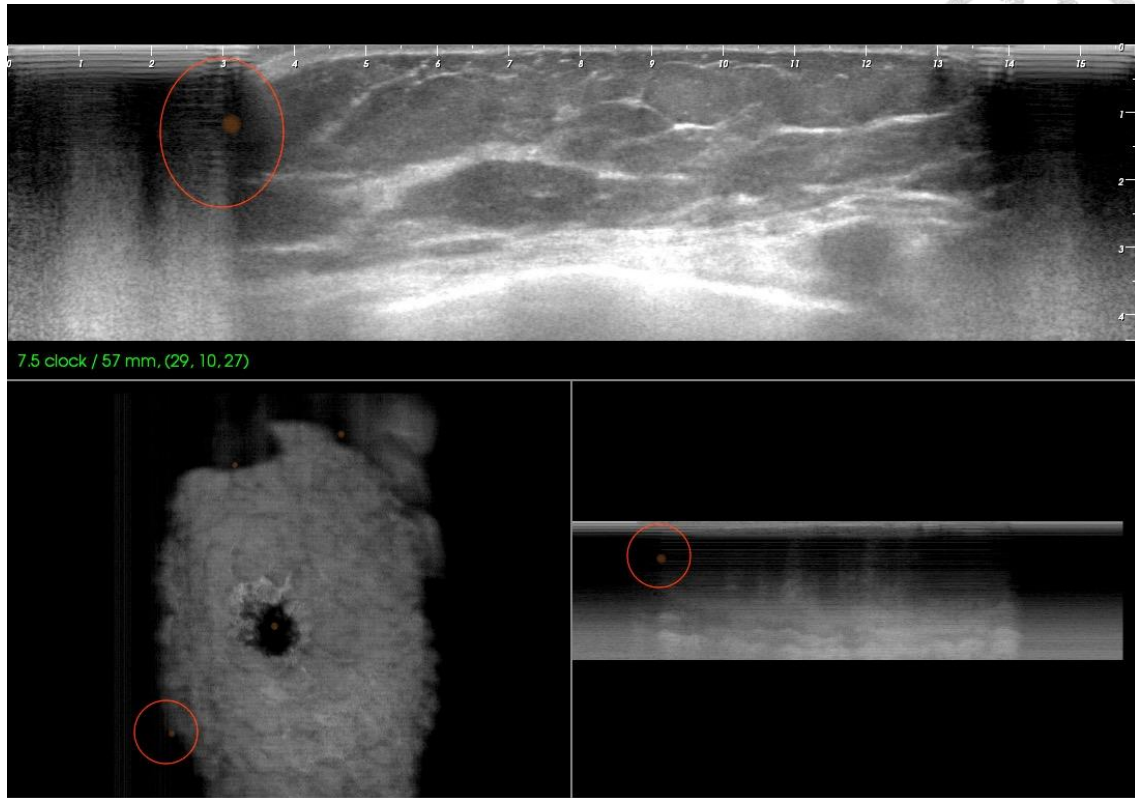


Fig. 7. The red circles indicate the tumor candidate region chosen by CADe system. Because of low echogenicity, the region of shadow was detected as tumor.

### 3.2 Anatomy-based FPR



Most of the diagnosed breast cancers belongs to the adenocarcinoma. Since adenocarcinoma starts to grow in the glands, glandular tissue extraction was expected to be meaningful in tumor detection. As shown in Fig. 6 and Fig. 7, some of the tumor candidates were FPs that need to filter out.

To distinguish FPs from the tumor candidates, anatomy-based FPR was adopted to remove tumor candidates with lower probabilities of being tumor. After applying CADe system, the tumor candidate regions were indicated as shown in Fig. 5. The main concern of the proposed anatomy-based FPR would be focused on the slices that were indicated by the CADe system.

### 3.2.1 Quantitative Feature Extraction

To distinguish non-tumors from tumor candidates, quantitative features were extracted from ABUS images used as input of a classifier, and the patch-based texture analysis was used in this study. ABUS images was divided into several equal size patches, and 20 2-D slices from different ABUS images were randomly chosen to be training cases for classification. After manual division, spatial and texture features were extracted and combined as region characteristics to generate the predicted probabilities in the classification procedure. The features used in this study are listed in Table 2 and explained in the following sections.



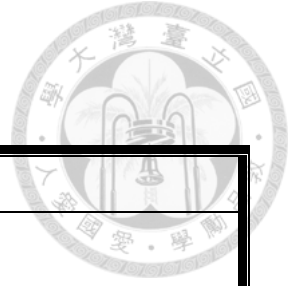
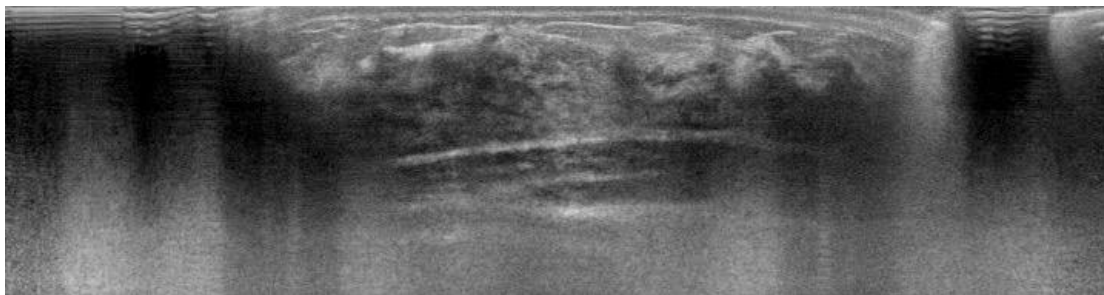


Table 2. The description of the adopted quantitative features

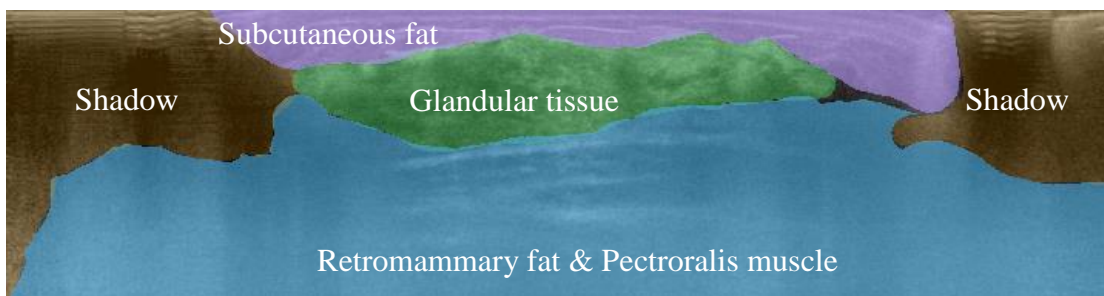
Category	Feature	Description
Spatial	$Lo_{Row}$	Location of row
	$Lo_{Column}$	Location of column
Texture	$En_{Mean}$	Energy mean
	$En_{Std}$	Energy standard deviation
	$Etr_{Mean}$	Entropy mean
	$Etr_{Std}$	Entropy standard deviation
	$Cor_{Mean}$	Correlation mean
	$Cor_{Std}$	Correlation standard deviation
	$Loho_{Mean}$	Local homogeneity mean
	$Loho_{Std}$	Local homogeneity standard deviation
	$In_{Mean}$	Inertia mean
	$In_{Std}$	Inertia standard deviation
	$CS_{Mean}$	Cluster shade mean
	$CS_{Std}$	Cluster shade standard deviation
	$CP_{Mean}$	Cluster prominence mean
	$CP_{Std}$	Cluster prominence standard deviation
	$HC_{Mean}$	Haralick's correlation mean
	$HC_{Std}$	Haralick's correlation standard deviation

### 3.2.1.1 Spatial Features

According to the breast anatomy, as shown in Fig. 8, the center part of the image in axial plane corresponds to the glandular tissue. Besides, the leftmost and rightmost regions correspond to the shadow, as well as the top region (subcutaneous fat) and bottom region (retromammary fat & pectoralis muscle). In other words, the image borders have low likelihoods to be glandular tissue. Thus, the location information was extracted to use for the tissue classification.



(a)



(b)

Fig. 8. The image showed the location of each region. (a) Original ABUS image (b) Illustration of the glandular tissue, subcutaneous fat, retromammary fat & pectoralis muscle, and shadow in the image.

### 3.2.1.2 Texture Features



Texture is one of significant features used to describe the spatial correlations between groups of pixels [21]. In this study, the gray level co-occurrence matrix (GLCM) method was used to interpret the texture pattern of regions. There are two groups of features according to the statistical order: first-order texture features and second-order texture features [22]. The first-order textures were extracted from intensity histogram without pixel locations information. The extended second-order texture features take the position of a pixel and its neighboring into consideration.

The GLCM method is calculated as follows. For an image  $I$ , it was transformed to a co-occurrence  $N \times N$  matrix  $M$  which  $N$  was reduced number of intensity bins. The definition of the matrix is  $= [g(i, j | d, \theta)]$ , the frequencies of two adjacent pixels with the gray value as  $i$  and the other as  $j$  and separated by distance  $d$  and the direction is given by the angle  $\theta$ . In the experiment,  $N=32$ ,  $d=1$  and  $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$  as shown in Fig.

9.

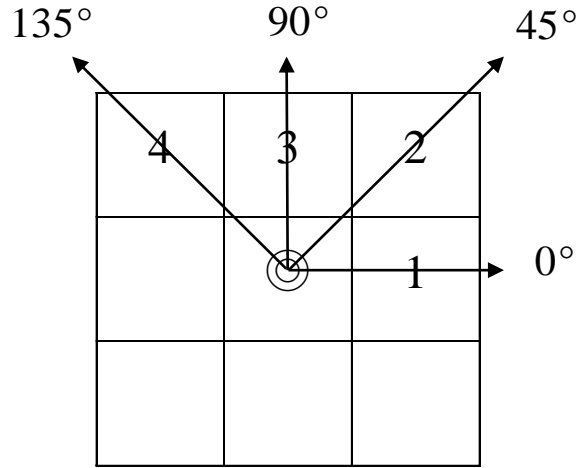


Fig. 9. The distance and direction used in GLCM method. The pixel pair of four directions from the centered pixel( $\odot$ ). Pixel 1 to 4 are the neighboring pixels in the direction of  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  with  $d=1$ , respectively.

There were eight GLCM features, energy, entropy, correlation, local homogeneity, inertia, cluster shade, cluster prominence, and Haralick's correlation. The GLCM texture features [23] are defined as follows:

Energy  $\square$  
$$f_1 = \sum_{i,j} g(i,j|d,\theta)^2 \quad (1)$$

Entropy  $\square$  
$$f_2 = - \sum_{i,j} g(i,j|d,\theta) \log_2(g(i,j|d,\theta)) \quad (2)$$

Correlation: 
$$f_3 = \frac{\sum_{i,j} (i - \mu_x)(j - \mu_y)g(i,j|d,\theta)}{\sigma_x \sigma_y} \quad (3)$$

Local Homogeneity: 
$$f_4 = \sum_{i,j} \frac{1}{1 + (i - j)^2} g(i,j|d,\theta) \quad (4)$$

Inertia: 
$$f_5 = \sum_{i,j} (i - j)^2 g(i,j|d,\theta) \quad (5)$$

Cluster Shade: 
$$f_6 = \sum_{i,j} [(i - \mu_x) + (j - \mu_y)]^3 g(i,j|d,\theta) \quad (6)$$

Cluster Prominence: 
$$f_7 = \sum_{i,j} [(i - \mu_x) + (j - \mu_y)]^4 g(i,j|d,\theta) \quad (7)$$



Haralick's Correlation: 
$$f_8 = \frac{\sum_{i,j} (i \cdot j) g(i, j|d, \theta) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (8)$$

where  $\mu_x, \mu_y, \sigma_x$  and  $\sigma_y$  are mean and standard deviation of the marginal distributions of

$g(i, j|d, \theta)$ .

$$\mu_x = \sum_i i * \sum_j g(i, j|d, \theta), \mu_y = \sum_j j * \sum_i g(i, j|d, \theta) \quad (9)$$

$$\sigma_x^2 = \sum_i (i - \mu_x)^2 * \sum_j g(i, j|d, \theta), \sigma_y^2 = \sum_j (j - \mu_y)^2 * \sum_i g(i, j|d, \theta) \quad (10)$$

### 3.2.2 Glandular Extraction

Image classification is the process to partition an image into several classes. To achieve glandular extraction, tissue classification was applied to classify 2-D ABUS slices into several partitions. Logistic regression [24] is one of the most famous machine learning algorithm for classification. In this study, a multinomial logistic regression [25] was adopted to classify the data into multi-classes.

To make a prediction for glandular extraction in further experiment, the training data was randomly chosen from 20 slices of ABUS images. Those ABUS image slices were cut into 11×11 patches and classified them into 4 classes (glandular tissue, subcutaneous fat, retromammary fat & pectoralis muscle, shadow) manually by the expert to ensure the classification of the training data was correct. The logistic regression model took the

quantitative features from manually classified image patches as input. The training process of logistic regression model is illustrated at Fig. 10.

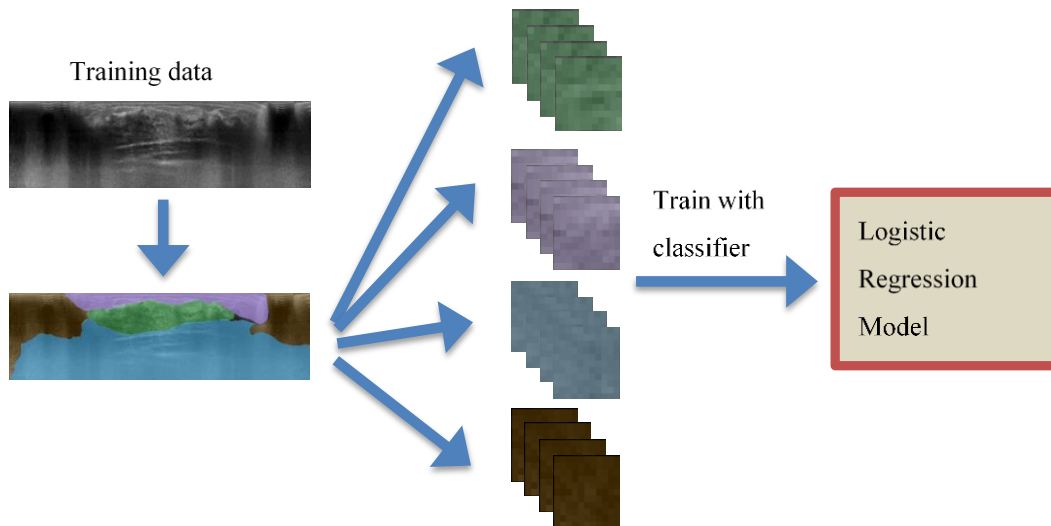


Fig. 10. The process of building logistic regression model. The left-up side is an original slice of training data. The slice being cut into patches were classified into 4 classes defined by the expert for training of the logistic regression model.

The classification performance was evaluated with leave-one-out cross-validation. That is, assume  $n$   $11 \times 11$  patches are in the training data using in logistic regression. One of the patches used as testing data to test the rest of  $n-1$  training data. The procedure is repeated afterwards  $n$  times until every patch being testing data once.

In the multinomial logistic regression, a linear predictor function  $L(k, i)$  is used to predict the probability of each class  $k$  with observation  $i$ . The linear predictor function  $L(k, i)$  is formulated by the following equation:

$$L(k, i) = c_{0,k} + c_{1,k}x_{1,i} + c_{2,k}x_{2,i} + c_{3,k}x_{3,i} + \cdots + c_{F,k}x_{F,i}$$

where  $c_{f,k}$  is the regression coefficient associated with the  $f$ th feature and the  $k$ th class,

$[x_{1,i}, x_{2,i}, \dots, x_{F,i}]$  is the set of features with observation  $i$ , and  $F$  is the number of features.

After that, class  $K$  (the last class) is chosen as a pivot, and the other  $K-1$  classes are separately regressed against class  $K$ . Because all  $K$  of probabilities must sum to one, the probability of class  $K$  is defined as

$$P = \frac{1}{1 + \sum_{k=1}^{K-1} e^{c_k \cdot x_i}}$$

The probabilities of the other  $K-1$  of probabilities would be calculated in the following equation:

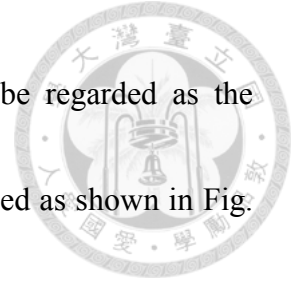
$$Pr(k) = \frac{e^{c_k \cdot x_i}}{1 + \sum_{k=1}^{K-1} e^{c_k \cdot x_i}}$$

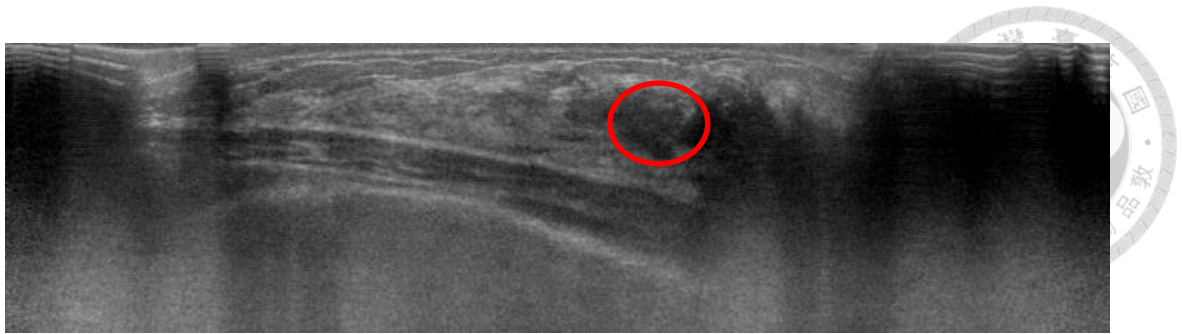
for  $K-1$  possible classes. In the experiment,  $F=18$  and  $K=4$ .

Before classifying testing data, the slices of ABUS images were also cut into  $11 \times 11$  patches and classified the patches into four classes (class 1: glandular tissue, class 2: subcutaneous fat, class 3: retromammary fat & pectoralis muscle, class 4: shadow) using the above-mentioned logistic regression model. After applying logistic regression classifier, each patch would be given a set of four predicted probabilities as the classification result. Each predicted probability  $Pr(k)$  of class  $k$  was between 0 and 1, and

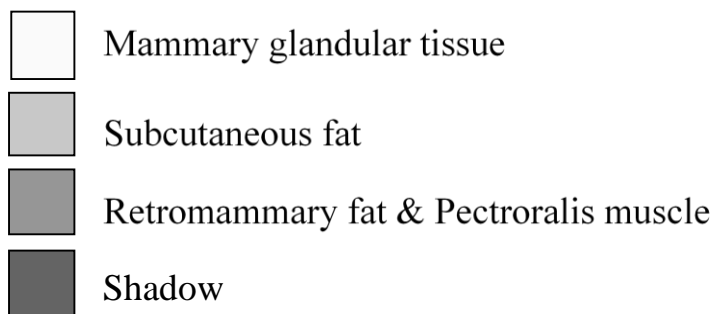
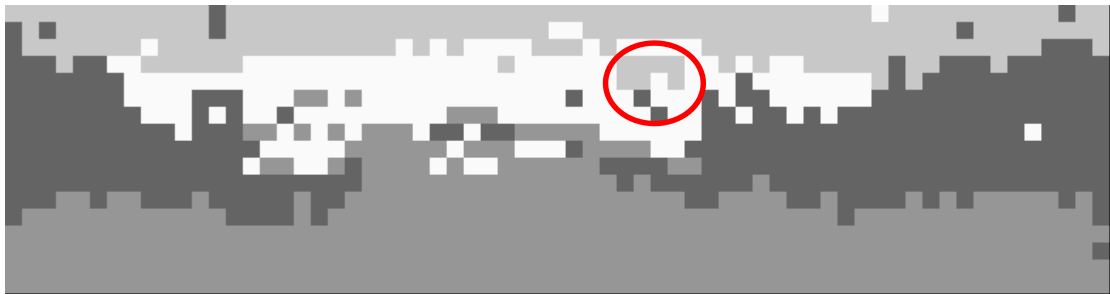
the class  $H$  with the highest predicted probability  $Pr(H)$  would be regarded as the appropriate class for the patch. The classification result was expressed as shown in Fig.

11.





(a)



(b)

Fig. 11. The result after applying image classification. (a) The original ABUS image. (b) The classification result of (a) after applying logistic regression model. The image was cut into  $11 \times 11$  patches, and the patches were classified into four classes as shown in the image.

### 3.2.3 Post Processing



Due to the complicated structure of the ABUS image, it is difficult to distinguish a proper class for the patch. As shown in Fig. 11, it is hard to tell the region indicated by red circle belongs to subcutaneous fat or glandular tissue. The patches from tumor candidate regions might also be classified into non-glandular regions. Therefore, this section would introduce the post processing to assure the validity of further experiments.

First of all, the patch classified into glandular class (class 1) should be reclassified if neighboring patches around it belong to non-glandular class (class 2 to class 4), such as the patches in the green circles in Fig. 12. Second, if the patch had been classified to non-glandular class (class 2 to 4), it should be reclassified to glandular class (class 1) in the following condition:

1.  $Pr(1)$  is the second highest probability.
2. At least two neighboring patches were classified to glandular class (class 1).

If both of the conditions were satisfied, the patch should be reclassified to glandular class as shown in the blue circles of Fig. 12.

After post processing of the patches, there were still some gaps and holes in the image. Some of ABUS images were with less glandular tissue, Furthermore, tumor

regions were possibly being classified into non-glandular class. Therefore, a morphological image process was applied to combine all of the glandular regions together as shown in Fig. 12 (c).



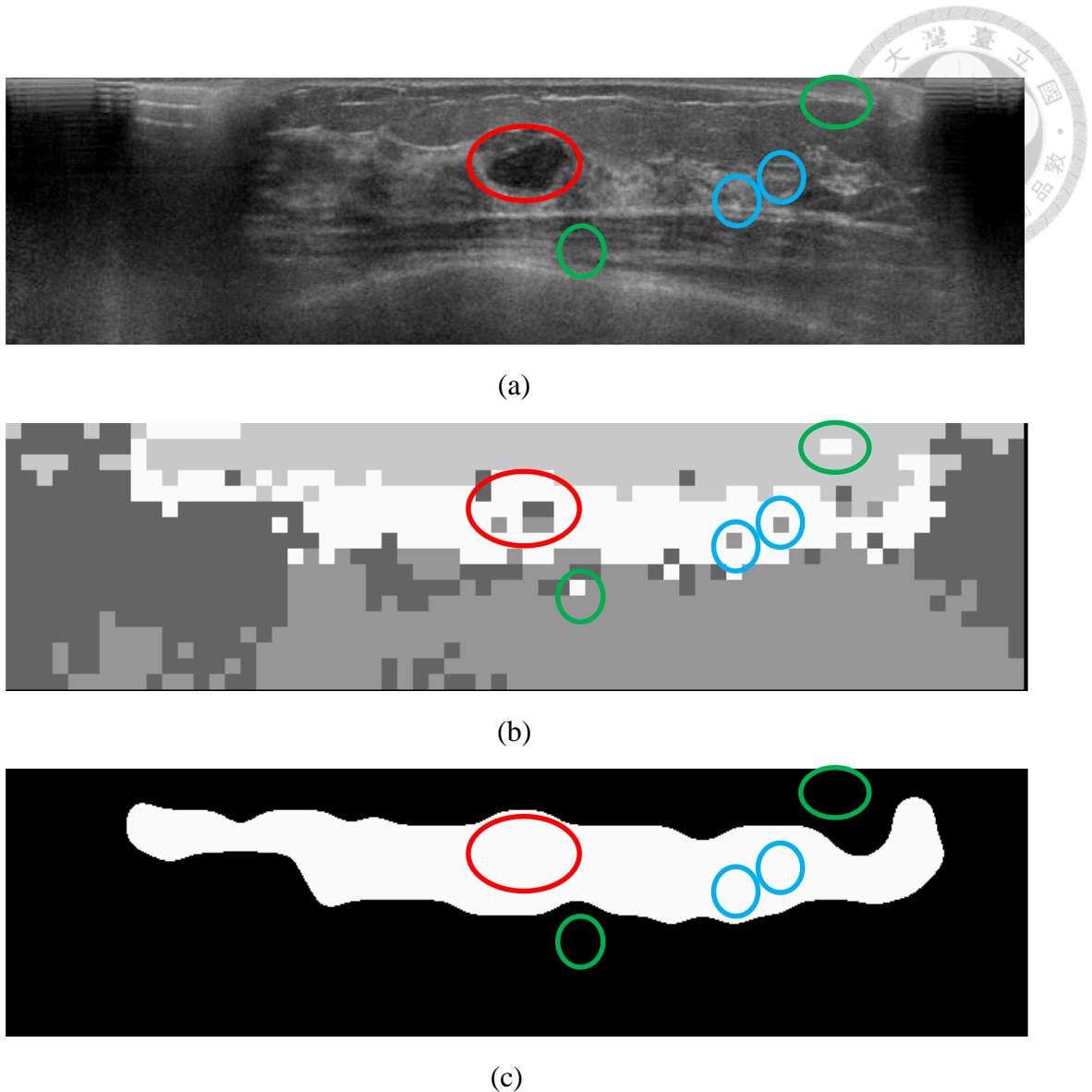


Fig. 12. The result after applying glandular extraction and post processing. The red circle indicates the tumor region specified in pathology report. The green circle indicates the non-glandular patches which were misclassified into glandular patches. The blue circle indicates the glandular patches which were misclassified into non-glandular patches. (a) The original ABUS image with a tumor circled with red line. (b) The result after image classification to (a). (c) The result after post processing to (b).



### 3.2.4 TIGR Calculation

Because breast cancers were usually inside of glandular tissue, the judging criterion

to extract suspicious abnormalities was defined as below:

$$TIGR = \frac{\text{Overlapping area between tumor candidate and glandular tissue}}{\text{Area of tumor candidate}}$$

Based on the *TIGR*, tumor candidates with low *TIGR* can be filter out. More FPs of the

CADe system can be further reduced with the same tumor detection rate.



## Chapter 4

### Experimental Results



The proposed CADe system and anatomy-based FPR for ABUS images was developed with Microsoft Visual Studio 2010 (Microsoft, Seattle, WA, USA) using C++ language and Microsoft Windows 7 operating system (Microsoft, Seattle, WA, USA). The program was running on an Intel® Core™ i7-2600 3.4 GHz processor with 5GB RAM.

A database containing 82 abnormal passes and 54 normal passes was used to measure the performance of the experiments. 82 abnormal passes included 43 benign and 39 malignant lesions. According to the result of CADe system, there were 1472 tumor candidate regions, 82 true positive and 1390 FPs in this experiment.

For image classification, the logistic regression model was developed to classify ABUS images into 4 classes. To determine whether spatial and texture features were useful criteria, the classification results with different groups of quantitative features were analyzed. In the training procedure of logistic regression model, the performance of leave-one-out cross-validation [26] with different feature sets, amount of data and patch size were generated as shown in Table 3 and Table 4, respectively.

According to Table 3, when patch size was  $11 \times 11$ , the accuracy of leave-one-out cross-validation for spatial and texture features corresponded to 69.80% and 68.93%.

With the advantage of the combination (spatial and texture features), the accuracy was improved up to 81.50% and the error rate was decreased to 18.50%. The Cohen's kappa was 0.75. A Cohen's kappa larger than 0.60 indicates good level inter-rater agreement.

Therefore, the spatial and texture features were critical properties for breast tissue classification. Furthermore, compared with training data with patch size  $9 \times 9$ , the accuracy was 82.38%, which approached to the accuracy with patch size  $11 \times 11$  (81.50).

Hence, patch size  $11 \times 11$  was used in the experiment due to its time-efficient running process.

The cross-validation analyzed with different amount of training data are listed in Table 4. Combining spatial and texture features, the corresponding accuracies were 81.50%, 82.67% and 82.00% for different amount of training data (4000, 6000, 8000). The best performance was achieved by 6000 patches of training data. As a result, 6000 patches of training data were used in the experiment, and all of the patches were chosen from the 20 slices of ABUS images.

Student's t test or the Mann-Whitney U test [27] were used to test for various features

in the non-glandular and glandular cases. The Kolmogorov-Smirnov test was applied to test for a normal distribution. Because the distributions of the features were abnormal, the median value was listed and the Mann-Whitney U test was used. The test results of the proposed features are listed in Table 5. It evaluated the distinguishable abilities of the proposed features. A  $p$ -value less than 0.05 indicates the difference between a feature for non-glandulars and glandulars was statistically significant.

For the determination of the anatomy-based FPR, *TIGR* was used as a threshold to observe the details of trade-offs between the sensitivity and FPs/pass. The sensitivity was the true positive rate, which varied with the threshold of *TIGR*. Table 6 shows the distribution of mistakenly excluded tumor candidates at sensitivity from 60% to 90%. The higher sensitivity would lead to less number of missed tumors regardless of the tumor size. The number and percentage of mistakenly excluded tumor candidates in different sensitivity rates are listed in Table 7.

The performance of the proposed anatomy-based FPR, the FPs/pass of CADe system before and after applying FPR in different sensitivity rates are listed in Table 8. The CADe system reduced the FPs to 9.69 at a sensitivity of 100% [19]. The anatomy-based FPR further reduced the FPs to 3.15 at a sensitivity of 100%. FPs/pass was 67.49%

decrease compared with previous CADe system.



In the experiment, according to the variation graph of sensitivity as shown in Fig. 13, the performance kept at the sensitivity of 100% while *TIGR* was less than 43%. That is, all of the tumors specified in the pathology report were successfully retained when the threshold of *TIGR* was 43%. On the contrary, the tumor candidates with *TIGR* less than 43% would be regarded as FP. The details of trade-offs between sensitivity and FPs/pass were illustrated using free response operating characteristics (FROC) [28] curves in Fig. 14.

Table 3. The performance of the leave-one-out cross-validation with different patch size. The amount of training data is 1000 patches for each class. That is, totally 4000 patches used as training data.

Patch size	Feature set	Accuracy (%)	Error (%)	Cohen's kappa $\kappa$
9×9	Spatial	67.55	32.45	0.57
	Texture	68.33	31.68	0.58
	Spatial+Texture	82.38	17.63	0.77*
11×11	Spatial	69.80	30.20	0.60*
	Texture	68.93	31.08	0.59
	Spatial+Texture	81.50	18.50	0.75*

\* Cohen's kappa bigger than 0.60 was good level inter-rater agreement.

Table 4. The performance of the leave-one-out cross-validation with different amount of data. The patch size was 11×11

Amount of training data (patch)	Feature set	Accuracy (%)	Error (%)	Cohen's kappa $\kappa$
4000 (1000 patches for each class)	Spatial+Texture	81.50	18.50	0.75*
6000 (1500 patches for each class)	Spatial+Texture	82.67	17.33	0.77*
8000 (2000 patches for each class)	Spatial+Texture	82.00	18.00	0.76*

\* Cohen's kappa bigger than 0.60 was good level inter-rater agreement.

Table 5. The median value, and p-value of Mann-Whitney U test for various features in glandular extraction.

Feature	Median		<i>p</i> -value
	Non-glandular	Glandular	
<i>Lo<sub>Row</sub></i>	346	379	<0.001*
<i>Lo<sub>Column</sub></i>	104	60	<0.001*
<i>En<sub>Mean</sub></i>	0.0810	0.0492	<0.001*
<i>En<sub>Std</sub></i>	0.0043	0.0040	<0.001*
<i>Etr<sub>Mean</sub></i>	3.9937	4.7228	<0.001*
<i>Etr<sub>Std</sub></i>	0.0664	0.1093	<0.001*
<i>Cor<sub>Mean</sub></i>	0.2390	0.2299	0.1982
<i>Cor<sub>Std</sub></i>	0.0677	0.0389	<0.001*
<i>Loho<sub>Mean</sub></i>	0.6037	0.5582	<0.001*
<i>Loho<sub>Std</sub></i>	0.0346	0.0452	<0.001*
<i>In<sub>Mean</sub></i>	1.4540	1.9517	<0.001*
<i>In<sub>Std</sub></i>	0.2546	0.4568	<0.001*
<i>CS<sub>Mean</sub></i>	0.2152	-0.2933	<0.001*
<i>CS<sub>Std</sub></i>	0.5853	1.4946	<0.001*
<i>CP<sub>Mean</sub></i>	27.7177	159.3316	<0.001*
<i>CP<sub>Std</sub></i>	5.5627	26.3866	<0.001*
<i>HC<sub>Mean</sub></i>	10137.3278	27257.3218	<0.001*
<i>HC<sub>Std</sub></i>	130.6542	459.2121	<0.001*

\* *p*-value less than 0.05 was statistically significant.

Table 6. The number of tumor candidates mistakenly excluded for different sizes of benign and malignant tumors at different sensitivity.

Sensitivity (%)	Malignant (cm)					Benign (cm)				
	<1.0	1.0-2.0	2.0-3.0	>3.0	Total	<1.0	1.0-2.0	2.0-3.0	>3.0	Total
60	0	11	7	1	19	8	9	1	0	18
70	0	4	7	0	11	5	8	1	0	14
80	0	3	5	0	8	4	5	0	0	9
90	0	3	3	0	6	2	1	0	0	3

Table 7. The number of tumor candidates mistakenly excluded in different sensitivity rates.

Sensitivity (%)	Number of tumor candidates mistakenly excluded		
	All	Malignant	Benign
60	37	19 (49%)	18 (42%)
70	25	11 (28%)	14 (33%)
80	17	8 (21%)	9 (21%)
90	9	6 (15%)	3 (7%)

Table 8. The average number of FPs/pass before and after applying the anatomy-based FPR.

Sensitivity (%)	FPs/pass	
	Before FP reduction	After FP reduction
60	1.98	0.43
70	2.45	0.69
80	3.95	1.42
90	6.17	1.90
100	9.69	3.15



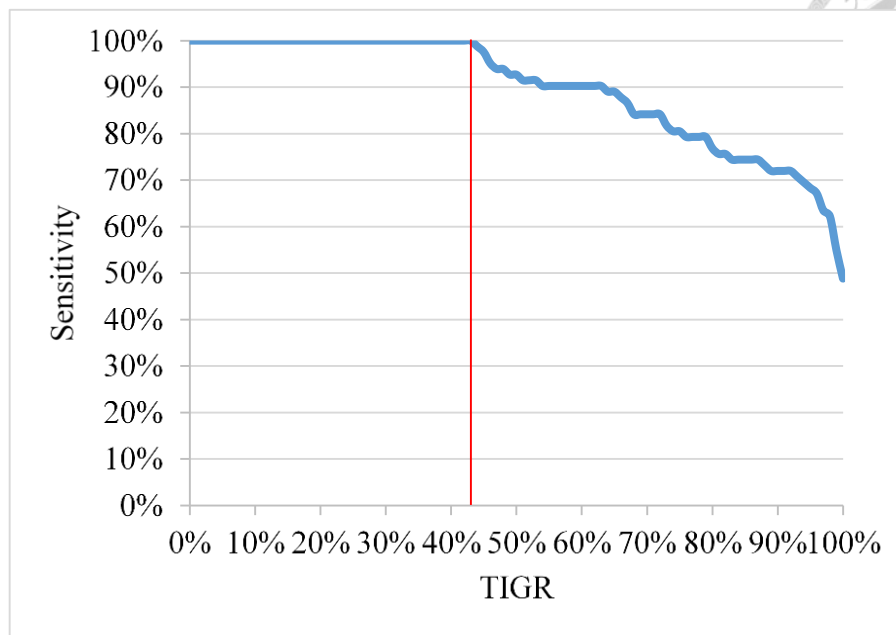


Fig. 13. The variation graph of the sensitivity with *TIGR* as a threshold. Using the testing data set in this study, the sensitivity was 100% until the threshold of *TIGR* was 43%. Red line indicates the *TIGR* of 43%.

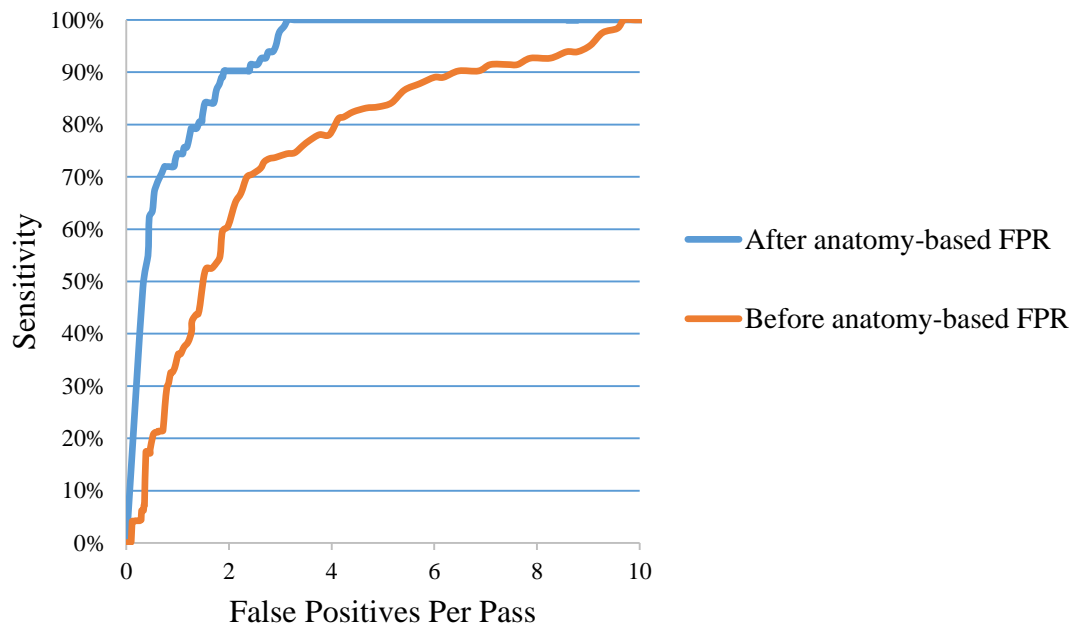


Fig. 14. The free-response receiver operating characteristic (FROC) curve for CADE system before and after the anatomy-based FPR was implemented.

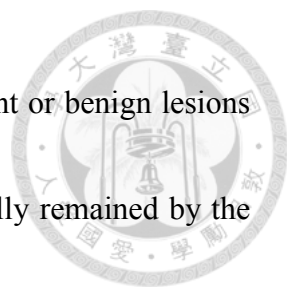
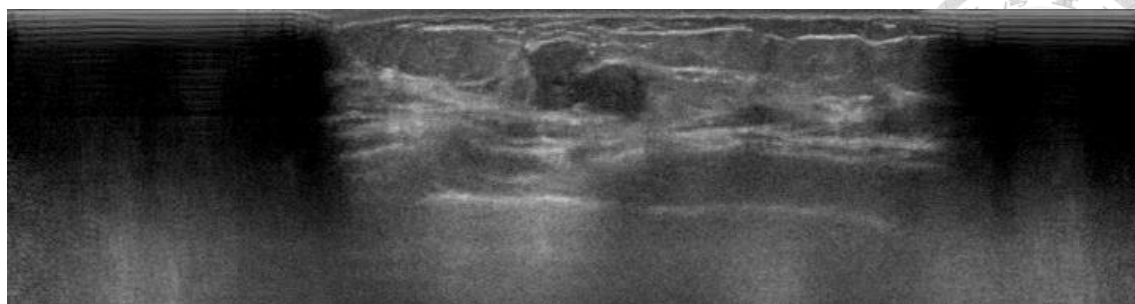
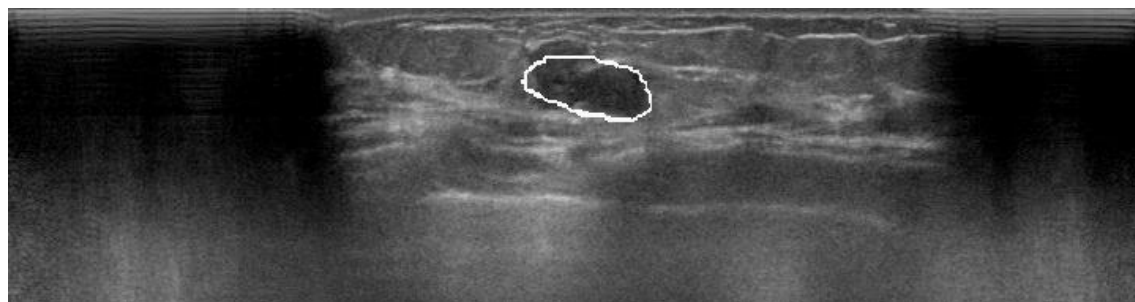


Fig. 15 - Fig. 17 show true positive cases which were malignant or benign lesions specified in the pathology report. All of these cases were successfully remained by the anatomy-based FPR. In Fig. 15, 2.00 cm invasive ductal carcinoma was almost completely inside of the glandular tissue. The *TIGR* was 92.31% and correctly regarded as a positive tumor ( $92.31\% > 43\%$ ). The true positive case of a 1.2 cm fibroadenoma as shown in Fig. 16 was totally overlapping with glandular tissue ( $TIGR = 100\%$ ), so it was correctly regarded as a positive tumor ( $100\% > 43\%$ ). Fig. 17 was also a true positive case of a malignant tumor which was detected as a 2.0 cm invasive ductal carcinoma. The *TIGR* was 66.51% that it was correctly regarded as a positive tumor ( $66.51\% > 43\%$ ).

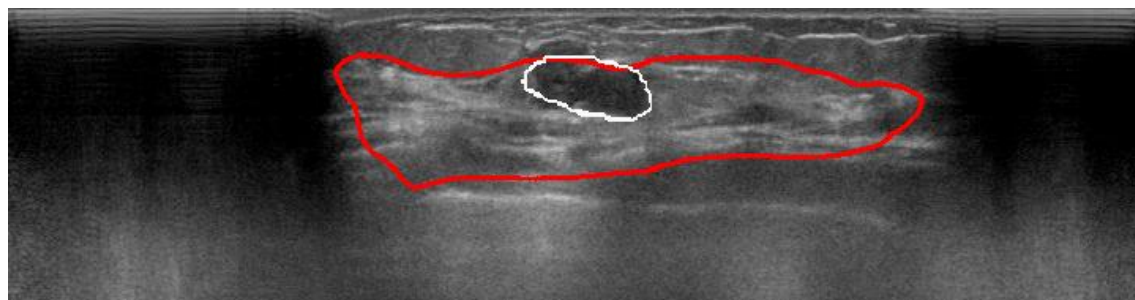
On the other hand, the case with *TIGR* less than 43% would be regarded as FP. Fig. 18 shows a FP. The *TIGR* was 0% that it was correctly regarded as FP. The hypoechoic region as shown in Fig. 19 is also FP, but the *TIGR* of this case was 81.01% that it was mistakenly regarded as a tumor .



(a)

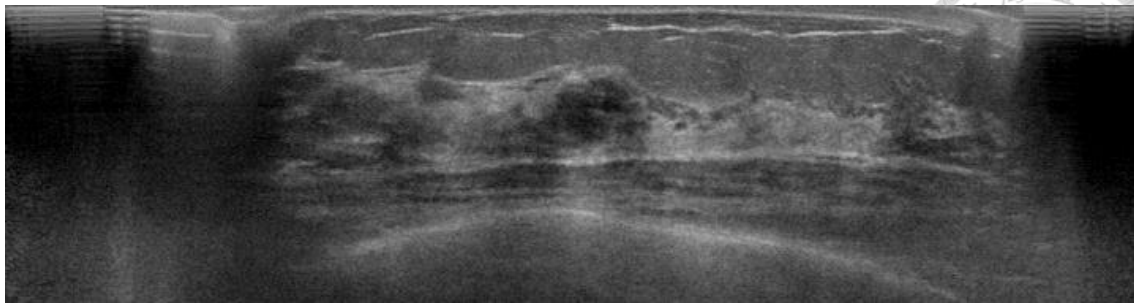


(b)

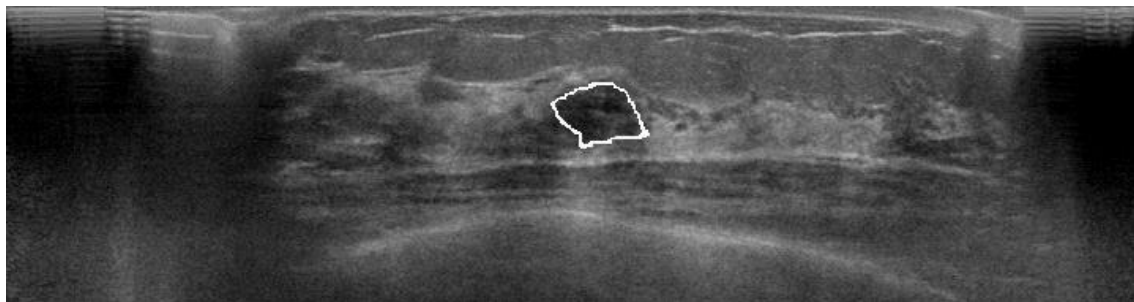


(c)

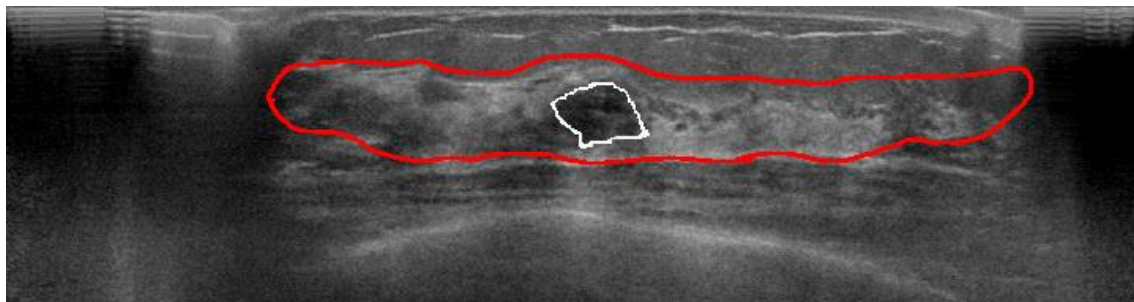
Fig. 15. True positive case of 2.0 cm invasive ductal carcinoma. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. It was specified in pathology report. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the *TIGR* of 92.31%.



(a)

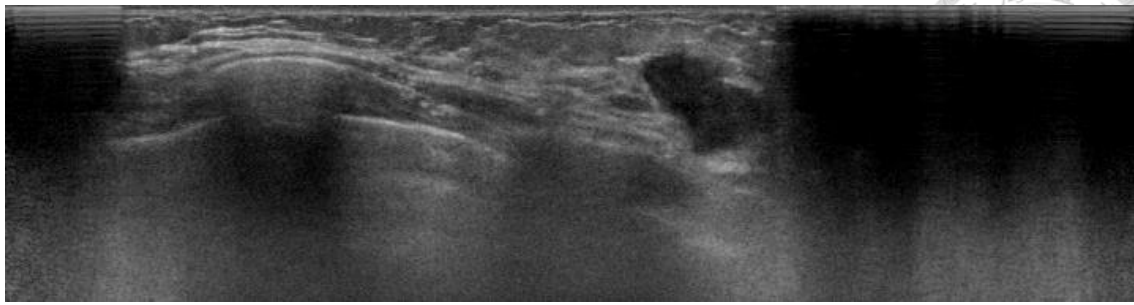


(b)

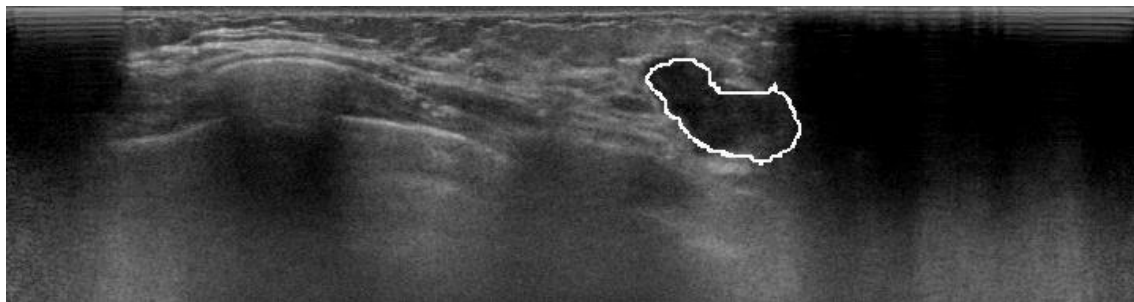


(c)

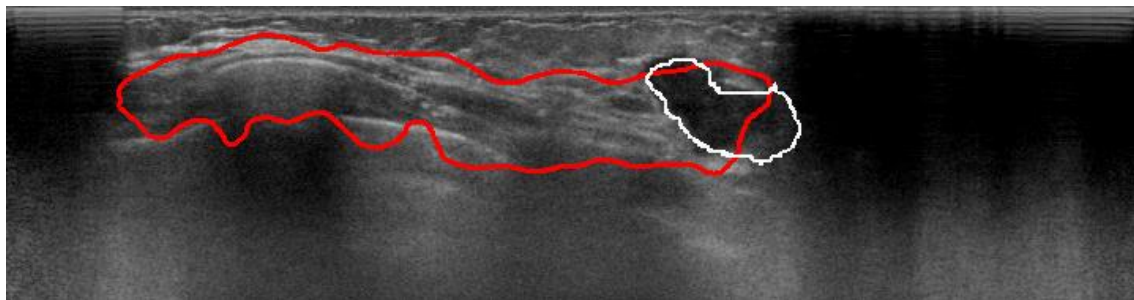
Fig. 16. True positive case of 1.2 cm fibroadenoma. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. It was specified in pathology report. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the *TIGR* of 100%.



(a)

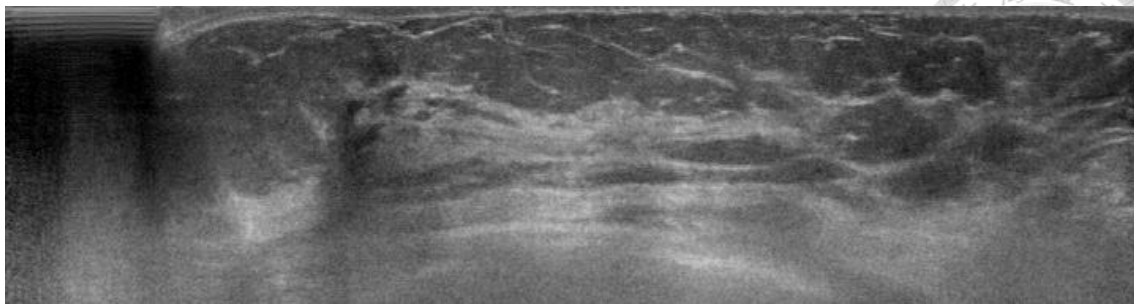


(b)

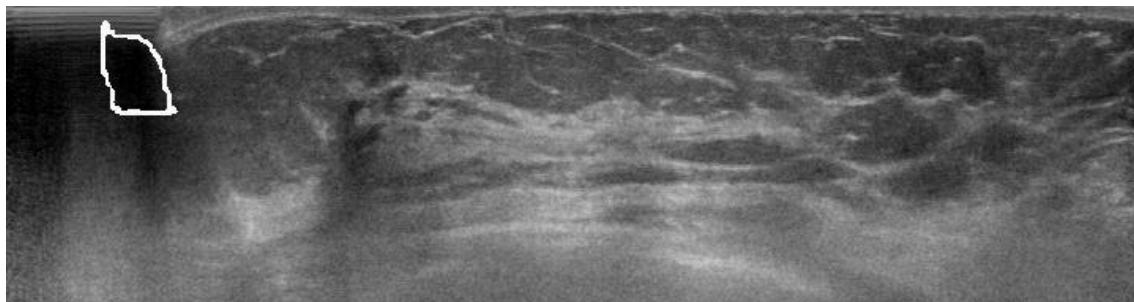


(c)

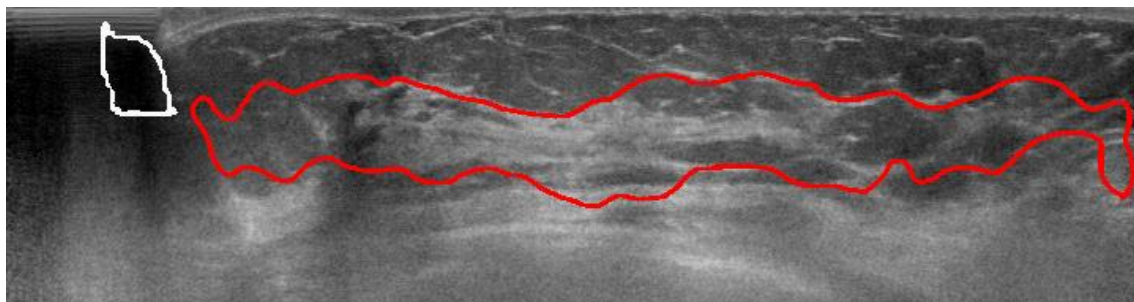
Fig. 17. True positive case of 2.0 cm fibroadenoma. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. It was specified in pathology report. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the *TIGR* of 66.51%.



(a)



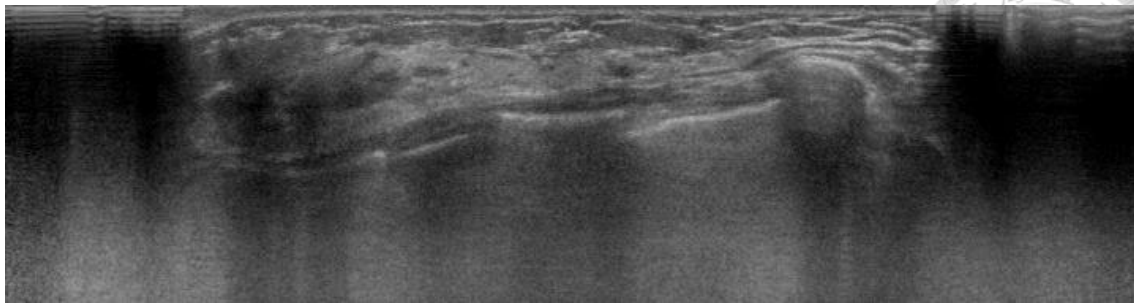
(b)



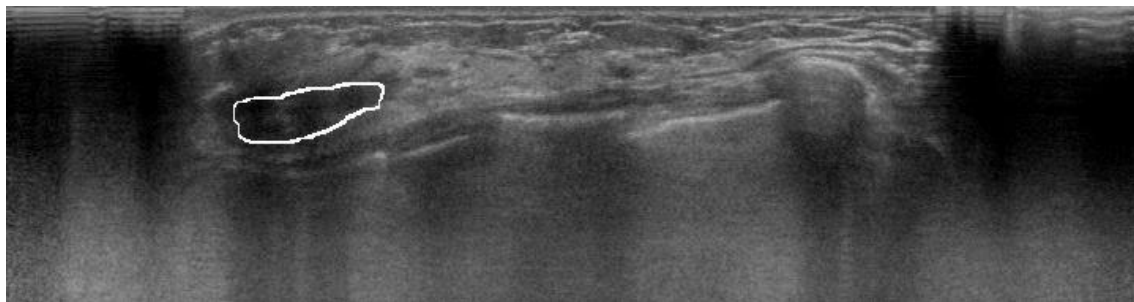
(c)

Fig. 18. False positive case. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CADe system. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was excluded with the *TIGR* of 0%.

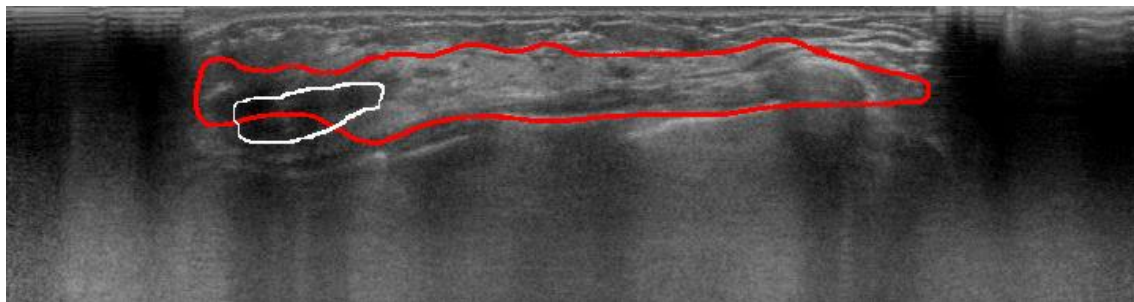




(a)



(b)



(c)

Fig. 19. False positive case. (a) The original ABUS image. (b) The white line indicates the tumor candidate detected by CAdE system. (c) The red line indicates the glandular tissue extracted in this study. The tumor candidate was retained with the *TIGR* of 81.01%.

## Chapter 5

### Discussion and Conclusion



The motivation of this study was to develop the anatomy-based FPR to improve the CAdE system built in the previous study [19]. When the threshold of *TIGR* being set to 43%, the FPR achieved 3.15 FPs/pass at the sensitivity of 100%, which was lower than the result (FPs/pass = 9.69) of the CAdE system before applying the anatomy-based FPR. All of the tumors including malignant and benign specified in the pathology reports were detected without mistakenly excluded.

However, there were still some points to be considered. It should be noted that the proposed FPR was applied to the previous CAdE system based on watershed transform [19]. The watershed transform gathers similar tissues which were hypoechoic to be homogeneous regions. Therefore, sometimes the regions chosen by CAdE systems might be mistaken. As shown in Fig. 19, the FP was detected as tumor since *TIGR* of 81.01% was larger than 43%. Nonetheless, FPs with high *TIGR* were difficult to exclude.

The limitation of the anatomy-based FPR is that tumor candidates near the nipple, shadow and the edge of the glandular tissue were often with low *TIGR*. According to Table 7, at a sensitivity of 90%, the percentage of mistakenly excluded malignant tumor



candidates was larger than that of mistakenly excluded benign tumor candidates. As shown in Fig. 17, the tumor candidate was near the shadow with *TIGR* of 66.51%. Nevertheless, all of the tumor candidates could be successfully retained at a sensitivity of 100%.

Besides, FP with high *TIGR* was difficult to exclude as shown in Fig. 18. To detect all of the tumors specified in pathology report and exclude FPs, it is important to trade-offs between sensitivity and FPs/pass. For a better tissue classification, more quantitative features may need to be calculated. Intensity inhomogeneity correction [29] can be used in ultrasound images before classification to avoid boundary leakage and suppress the shadows.

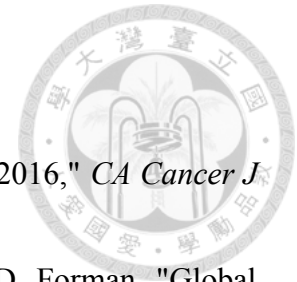
Moreover, according to Fig. 17, the low *TIGR* issue might be related to the accuracy of CAdE system. Although the anatomy-based FPR achieved sensitivity of 100%, the corresponding 3.15 FPs/pass may be further decreased by adding the detecting system in the future.

Besides, this study concentrated on tumors inside of glandular tissues. There are other types of cancers in the breast, too, such as sarcomas, which start in the cells of muscle, fat, or connective tissue. However, the tumor candidates retained were with

higher probability as real tumors. In addition, the database used in this study consisted of both abnormal passes and normal passes. The tumor candidates with high *TIGR* in normal passes would be helpful for radiologists do malignancy evaluation.

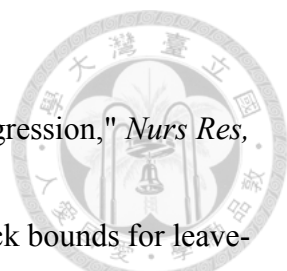


## Reference



- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA Cancer J Clin*, vol. 66, pp. 7-30, Jan-Feb 2016.
- [2] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA Cancer J Clin*, vol. 61, pp. 69-90, Mar-Apr 2011.
- [3] M. A. Roubidoux, M. A. Helvie, N. E. Lai, and C. Paramagul, "Bilateral breast cancer: early detection with mammography," *Radiology*, vol. 196, pp. 427-31, Aug 1995.
- [4] T. E. Wilson, M. A. Helvie, and D. A. August, "Breast cancer in the elderly patient: early detection with mammography," *Radiology*, vol. 190, pp. 203-7, Jan 1994.
- [5] E. A. Sickles, R. A. Filly, and P. W. Callen, "Benign breast lesions: ultrasound detection and diagnosis," *Radiology*, vol. 151, pp. 467-70, May 1984.
- [6] T. M. Kolb, J. Lichy, and J. H. Newhouse, "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, pp. 165-75, Oct 2002.
- [7] J. L. Jesneck, J. Y. Lo, and J. A. Baker, "Breast mass lesions: computer-aided diagnosis models with mammographic and sonographic descriptors," *Radiology*, vol. 244, pp. 390-8, Aug 2007.
- [8] H. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognition*, vol. 43, pp. 299-317, 2010.
- [9] K. M. Kelly, J. Dean, W. S. Comulada, and S. J. Lee, "Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts," *Eur Radiol*, vol. 20, pp. 734-42, Mar 2010.
- [10] E. Wenkel, M. Heckmann, M. Heinrich, S. A. Schwab, M. Uder, R. Schulz-Wendtland, *et al.*, "Automated breast ultrasound: lesion detection and BI-RADS classification--a pilot study," *Rofo*, vol. 180, pp. 804-8, Sep 2008.
- [11] Y.-H. Chou, C.-M. Tiu, J. Chen, and R.-F. Chang, "Automated full-field breast ultrasonography: the past and the present," *Journal of Medical Ultrasound*, vol. 15, pp. 31-44, 2007.
- [12] R.-F. Chang, K.-C. Chang-Chien, H.-J. Chen, D.-R. Chen, E. Takada, and W. K. Moon, "Whole breast computer-aided screening using free-hand ultrasound," in

- International Congress Series*, 2005, pp. 1075-1080.
- [13] R. F. Chang, K. C. Chang-Chien, E. Takada, C. S. Huang, Y. H. Chou, C. M. Kuo, *et al.*, "Rapid image stitching and computer-aided detection for multipass automated breast ultrasound," *Med Phys*, vol. 37, pp. 2063-73, May 2010.
- [14] K. Drukker, M. L. Giger, K. Horsch, M. A. Kupinski, C. J. Vyborny, and E. B. Mendelson, "Computerized lesion detection on breast ultrasound," *Med Phys*, vol. 29, pp. 1438-46, Jul 2002.
- [15] K. Drukker, M. L. Giger, and E. B. Mendelson, "Computerized analysis of shadowing on breast ultrasound for improved lesion detection," *Med Phys*, vol. 30, pp. 1833-42, Jul 2003.
- [16] K. Drukker, M. L. Giger, and C. E. Metz, "Robustness of computerized lesion detection and classification scheme across different breast US platforms," *Radiology*, vol. 237, pp. 834-40, Dec 2005.
- [17] Y. Ikeda, D. Fukuoka, T. Hara, H. Fujita, E. Takada, T. Endo, *et al.*, "Development of a fully automatic scheme for detection of masses in whole breast ultrasound images," *Med Phys*, vol. 34, pp. 4378-88, Nov 2007.
- [18] W. K. Moon, Y. W. Shen, M. S. Bae, C. S. Huang, J. H. Chen, and R. F. Chang, "Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images," *IEEE Trans Med Imaging*, vol. 32, pp. 1191-200, Jul 2013.
- [19] C. M. Lo, R. T. Chen, Y. C. Chang, Y. W. Yang, M. J. Hung, C. S. Huang, *et al.*, "Multi-dimensional tumor detection in automated whole breast ultrasound using topographic watershed," *IEEE Trans Med Imaging*, vol. 33, pp. 1503-11, Jul 2014.
- [20] W. Misdorff, "Tumors of the mammary gland," *Tumors in Domestic Animals, Fourth Edition*, pp. 575-606, 2002.
- [21] W. K. Moon, C. M. Lo, C. S. Huang, J. H. Chen, and R. F. Chang, "Computer-aided diagnosis based on speckle patterns in ultrasound images," *Ultrasound Med Biol*, vol. 38, pp. 1251-61, Jul 2012.
- [22] T. Schindewolf, W. Stolz, R. Albert, W. Abmayr, and H. Harms, "Classification of melanocytic lesions with color and texture analysis using digital image processing," *Anal Quant Cytol Histol*, vol. 15, pp. 1-11, Feb 1993.
- [23] R. M. Haralick and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, pp. 610-621, 1973.
- [24] D. W. Hosmer Jr and S. Lemeshow, *Applied logistic regression*: John Wiley &

- 
- Sons, 2004.
- [25] C. Kwak and A. Clayton-Matthews, "Multinomial logistic regression," *Nurs Res*, vol. 51, pp. 404-10, Nov-Dec 2002.
  - [26] M. Kearns and D. Ron, "Algorithmic stability and sanity-check bounds for leave-one-out cross-validation," *Neural Computation*, vol. 11, pp. 1427-1453, 1999.
  - [27] R. C. Sprinthal, *Basic statistical analysis*, 2011.
  - [28] D. P. Chakraborty, "Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data," *Med Phys*, vol. 16, pp. 561-8, Jul-Aug 1989.
  - [29] C. Y. Lee, Y. H. Chou, C. S. Huang, Y. C. Chang, C. M. Tiu, and C. M. Chen, "Intensity inhomogeneity correction for the breast sonogram: constrained fuzzy cell-based bipartitioning and polynomial surface modeling," *Med Phys*, vol. 37, pp. 5645-54, Nov 2010.
  - [30] T. Tan, B. Platel, R. Mus, L. Tabar, R. M. Mann, and N. Karssemeijer, "Computer-aided detection of cancer in automated 3-D breast ultrasound," *IEEE Trans Med Imaging*, vol. 32, pp. 1698-706, Sep 2013.
  - [31] K. Ganesan, U. R. Acharya, C. K. Chua, L. C. Min, K. T. Abraham, and K. H. Ng, "Computer-aided breast cancer detection using mammograms: a review," *IEEE Rev Biomed Eng*, vol. 6, pp. 77-98, 2013.