私 立 東 海 大 學

資 訊 工 程 研 究 所

碩士論文

指導教授：黃育仁 博士

乳房保留手術中之腫瘤邊緣評估

Intra-operative tumor margin evaluation

in breast-conserving surgery

研究生：陳怡君

中華民國 一零八 年 六 月 十七 日

Intra-operative tumor margin evaluation in breast-conserving surgery with deep learning

by

Yi-Chun Chen

Advisor

Prof. Yu-Len Huang

Submitted to the Department of Computer Science of

Tunghai University

in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING

June 17 2019

東海大學碩士學位論文考試審定書

東海大學資訊工程學系 研究所

研究生 陳 怡 君 所提之論文

乳房保留手術中之腫瘤邊緣評估

經本委員會審查，符合碩士學位論文標準。

學位考試委員會

召 集 人 吳士駿 簽章

委 員 張孟仁

黃育仁

指 導 教 授 黃育仁 簽章

中華民國 108 年 6 月 17 日

摘要

乳癌是現今婦女最常罹患的癌症，但隨著醫學研究的發展，若能早期發現並及早接受治療，可有效地提高乳癌的治癒率。乳房保留手術(BCT)再加上術後放射治療，是治療早期乳癌的優先選擇。乳房保留手術能保有乳房的外觀，並降低復發的機率，然而，對於任何惡性腫瘤，陽性邊緣可能導致 BCT 後局部復發的風險增加，為了減少陽性邊緣的數量，將為外科醫生提供關於陽性切除邊緣存在的及時資訊。本論文提出了在乳房保留手術的術中腫瘤邊緣評估，首先利用門檻值取出感興趣的區域，並以傳統的影像切割方法，多重門檻值、K-means 和區域成長法以及兩個深度學習網路來進行腫瘤區域切割，接著評估其周圍正常組織的邊緣寬度。本論文所提出的方法，可使外科醫生獲得更多的資訊，希望在進行乳房保留手術時獲得乾淨的邊緣。本研究總共使用 30 個病例進行評估，最後將實驗結果與醫師手繪的腫瘤區域以及病理報告進行比較。實驗結果顯示，比起傳統的影像切割方法，深度學習網路能繪製出更符合病理報告的結果。在深度學習技術的輔助下，本研究有潛力成為術中輔助量測系統。

關鍵字： 乳癌、乳房保留手術、腫瘤邊緣評估、乳房 X 光攝影、深度學習、影像切割

Abstract

Breast cancer is the most commonly diagnosed cancer in women. Breast-conserving therapy (BCT) followed by irradiation is the treatment of choice for early-stage breast cancer. Breast retention surgery preserves the appearance of the breast and reduces the chance of recurrence. A positive margin may result in an increased risk of local recurrences after BCT for any malignant tumor. In order to reduce the number of positive margins would offer surgeon real-time intra-operative information on the presence of positive resection margins. This thesis aims to design an intra-operative tumor margin evaluation scheme by using specimen mammography in breast-conserving surgery. The proposed method first utilizes image thresholding to extract regions of interest and then to segment cancer tissue using various segmentation methods, i.e. multi-thresholding, K-means and regional growth methods and two deep learning networks. Finally, the margin width of normal tissues surrounding it is evaluated as the result. With this work, surgeons would acquire more information to get clean margins when performing breast conserving surgeries. This study evaluated total of 30 cases, the results were compared with the manually determined contours and pathology report. The experimental results reveal that deep learning techniques can draw results that are more consistent with pathology reports than traditional segmentation methods. With the aid of deep learning techniques, the proposed scheme would be a potential procedure in the intra-operative measurement system.

Keywords—Breast cancer, breast-conserving therapy, tumor margin evaluation, specimen mammography, deep learning, image segmentation

Acknowledgements

The author wishes to express her deepest appreciation to her thesis advisor, Professor Yu-Len Huang, for his inspiration and guidance throughout the course of this work. The author is particularly grateful to Dr. Dar-Ren Chen and Dr. Hwa-Koon Wu; the ideas in this theis and the cognition on surgical margin width have benefited enormously from discussion with them.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

Breast cancer is the most commonly diagnosed cancer in women. Fortunately, early detection and treatment could prevent the disease from worsening and reduce patient's mortality significantly. Surgery is one of the most important treatments for breast cancer. There are two kinds of surgery, one is mastectomy, which occurs when the tumor is too large (particularly in a small breast) or more than one area of cancer in the breast. The other one is breast-conserving therapy (BCT), in which only cancerous tissue plus a rim of normal tissue can be cleaned without removing the breast [1].

BCT is the best choice for the treatment of early stage invasive breast cancer. The surgery can not only remove the tumor but also preserve the shape of the breast. However, a positive margin may result in an increased risk of local recurrences after BCT for any malignant tumor. Until now, the definition of a positive margin has been the subject of frequent debate [2]. In reality, surgeon removes the tumor is done by rough estimation of the boundary. Surgeon could not accurately determine the margin width until the pathologist makes a microscopic assessment. Pathologist’s report might require a week or more to completed. If it shows the margins are not wide enough, the patient must undergo a second operation to remove the remaining malignant tissue. The operation would cause second physical and mental injury to patient.

For removing the tumor while minimizing the risk of leaving residual disease, many intra-operative methods have been proposed for tumor margin assessment, such as optical coherence tomography (OCT) [3], MarginProbe system [4], spectroscopy [5], molecular fluorescence imaging [6], and so on. OCT is a high-resolution imaging technique involving 2mm real-time microscopic images below the tissue surface. Within a short period of time after the invention in the 1990s, it became an important clinical imaging modality in several fields of

biomedical science. OCT has found clinical applications in ophthalmology [7], cardiology [8], gastroenterology [9], and oncology. OCT was reported to have a sensitivity of 82% and a specificity of 100% in breast cancer. MarginProbe system which made by Israel-based Dune Medical Devices, allows a surgeon to examine cancerous tissue removed from a breast to ensure that there are no malignant cells on its outer boundaries. A pen-like probe reads electromagnetic waves from the tissue and indicates on an attached console whether or not its edges are healthy, but with 75.2% sensitivity and 46.4% specificity. Multi-modal spectral histopathology (MSH), a multimodal imaging technique combining tissue auto-fluorescence and Raman spectroscopy was used to detect microscopic residual tumor at the surface of the excised breast tissue. The sensitivity and specificity of the MSH as 91% and 83%, respectively. However, these methods have not widely been accepted as part of standard of care due to high equipment cost. The molecular fluorescence imaging required contrast medium injection for patient few days before the BCT, but some of the patients feel nausea, vomiting, discomfort, or causing allergy after tracer injection, this it is not suitable for everyone.

Mammography is one of the most common screening tools to diagnose breast tumor, which is using low-energy X-rays to examine the human breast. The advantage of using mammography is that detection of micro-calcifications more precisely and becoming cheaper to produce as the technology becomes more widespread. In order to reduce the number of positive margins and offer surgeon real-time intra-operative information on the presence of positive resection margins, this study proposed multiple method computer-aided system by using the specimen mammography during BCT. Specimen mammography [10] is routinely used to evaluate the surgical margin. The specimen is transported to a room near the operating room where mammography device was placed after the surgeon has resected the specimen. Mammogram is captured immediately and stored in a hospital diagnostic imaging system. The surgeon could quickly review the digital mammogram to assess the integrity of the resection. Specimen mammography with digital system have an edge on low equipment cost and less time

consuming.

The rest of this thesis is organized as follows. Material and methods of this study can be found in Chapter 2. The flowchart (see Fig. 1) was given to explain the proposed procedure, a detailed description of each step is recorded in followed sections. Chapter 3 presents the experimental results. Computing time, similarity measures and average difference were used to compare in different situation. Finally, Chapter 4 concludes the thesis.

CHAPTER 2

MATERIAL AND METHODS

2.1. Data acquisition

Two full field digital mammography (FFDM) systems were included in the study, i.e. GE Senographe Essential and Hologic Selenia Dimentions system. After wide excision of the tumor, location stiches were made on 12 (0º), 3(90º), 6 (180º) and 9 (270º) o'clock direction and clipped on the stiches in order to be easily identified on specimen mammogram. There were 1, 2, 3 and 4 clips on each stich. This study included 30 patients who received BCT. Each specimen mammogram is high resolution (over 1000 × 1000) and varies in size, has a corresponding ground truth image which is manually annotated by experienced surgeons. All obtained images were stored on the hard disk and transferred to a personal computer using a DICOM connection for image analysis. The pathologic report of margin distance was considered as the ground truth in this study. The format of pathology report descriptions are shown in Table 2.1.

Table 2.1. Pathology report descriptions

PATHOLOGY REPORT
The specimen submitted consists of a suture oriented breast measuring X cm in size totally, in fresh state. On cut, there is an irregular firm tumor measuring X cm in size. The tumor is grayish-white in color and contains minute chalky white streaks. It is located at X cm beneath superficial margin and X cm above the pectoral fascia. The 3 o'clock, 6 o'clock, 9 o'clock, and 12 o'clock margins measure X cm, X cm, X cm, and X cm, respectively. There is no gross/gross evidence of invasion of the underlying pectoral fascia and muscles. The parenchyma in the remainder of the breast is not remarkable**.**

2.2. Flow-chart of the proposed method

At first, the proposed method measured the pixel density and extracted the region of interest (ROI) by detecting the specimen boundary in mammography. Then the tumor boundary was detected by the proposed contouring methods from the ROI. A distance evaluation step was applied in a final stage to obtain the result. Flow-chart of the proposed method is shown in Fig. 2.1. The flowing sections described the procedures within the proposed method in detail.

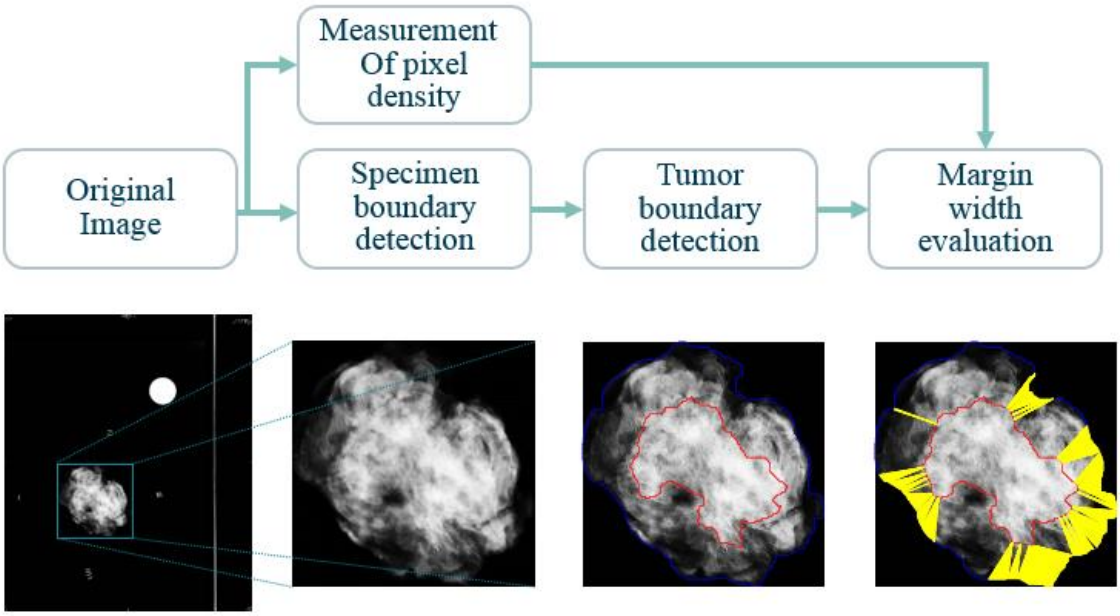
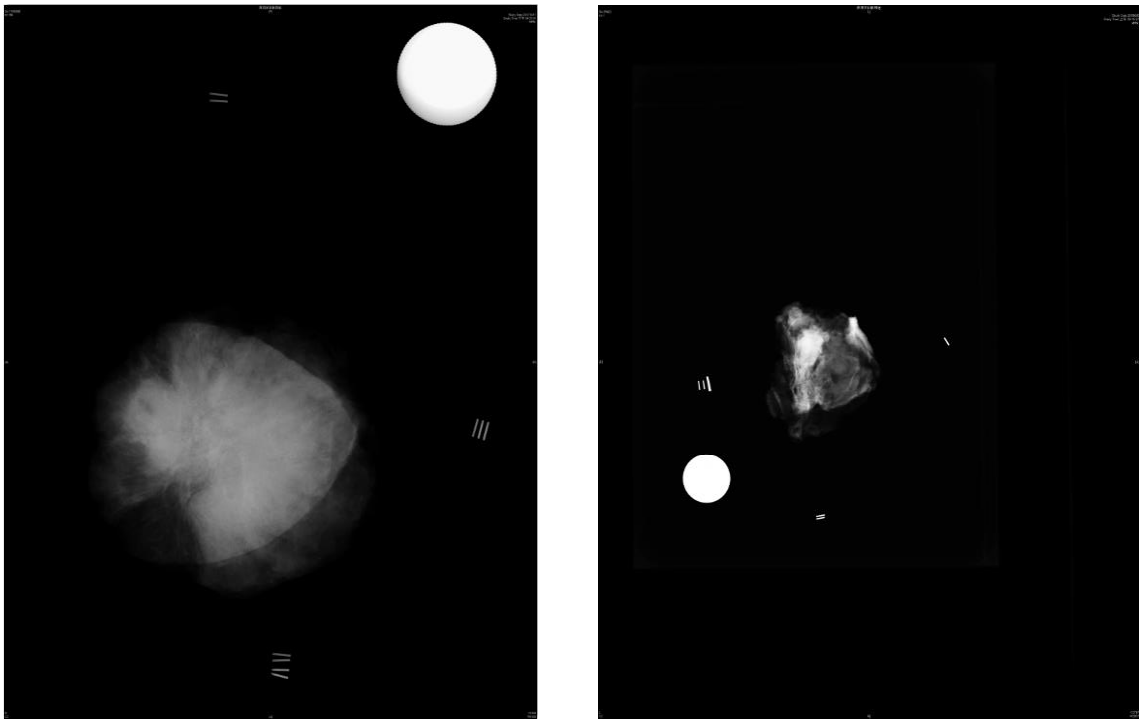


Figure 2.1. Flow-chart of the proposed method

2.3. Measurement of pixel density

In the dataset, each tumor has a different proportion in the image. That is the operator will zoom in or out the specimen mammography during the filming stage in order to observe the specimen clearly. In order to accurately measure the distance between the tumor and the tissue, a standard one-dollar coin (20mm diameter) was placed in the specimen mammogram as a measuring scale. Pixel resolution was converted to millimeter by according the radius of the coin.



(a) (b)

Figure 2.2. (a) The case with large pixel density (239 pixels per 1mm), (b) the case with small pixel density (104 pixels per 1mm)

2.4. Specimen boundary detection and ROI extraction

Image pre-processing which removes the noise and enhances the quality of the image is very important step for image segmentation. The major problem with the precise segmentation of the specimen boundary is that the existence of noises which might affect the segmentation results. In order to suppress the noise in the background, an automatic thresholding method [11] was performed to the specimen mammogram (Fig. 2.3(a)). However, the images also contain artifacts in the form of labels, wedges, markers and some patient information in the background region. In the proposed method, the connected component algorithm was utilized to extract the largest component (specimen), which means artifacts were discarded (Fig. 2.3(b)). In order to extract the tumor region completely, fill-hole operator was utilized to eliminate the black cavity. The morphological operators [11], i.e. opening, closing and erosion, were used to smooth the boundary (Fig. 2.3(c)). The obtained specimen boundary was utilized as extract ROI for the following tumor boundary detection step (Fig. 2.3(d)).

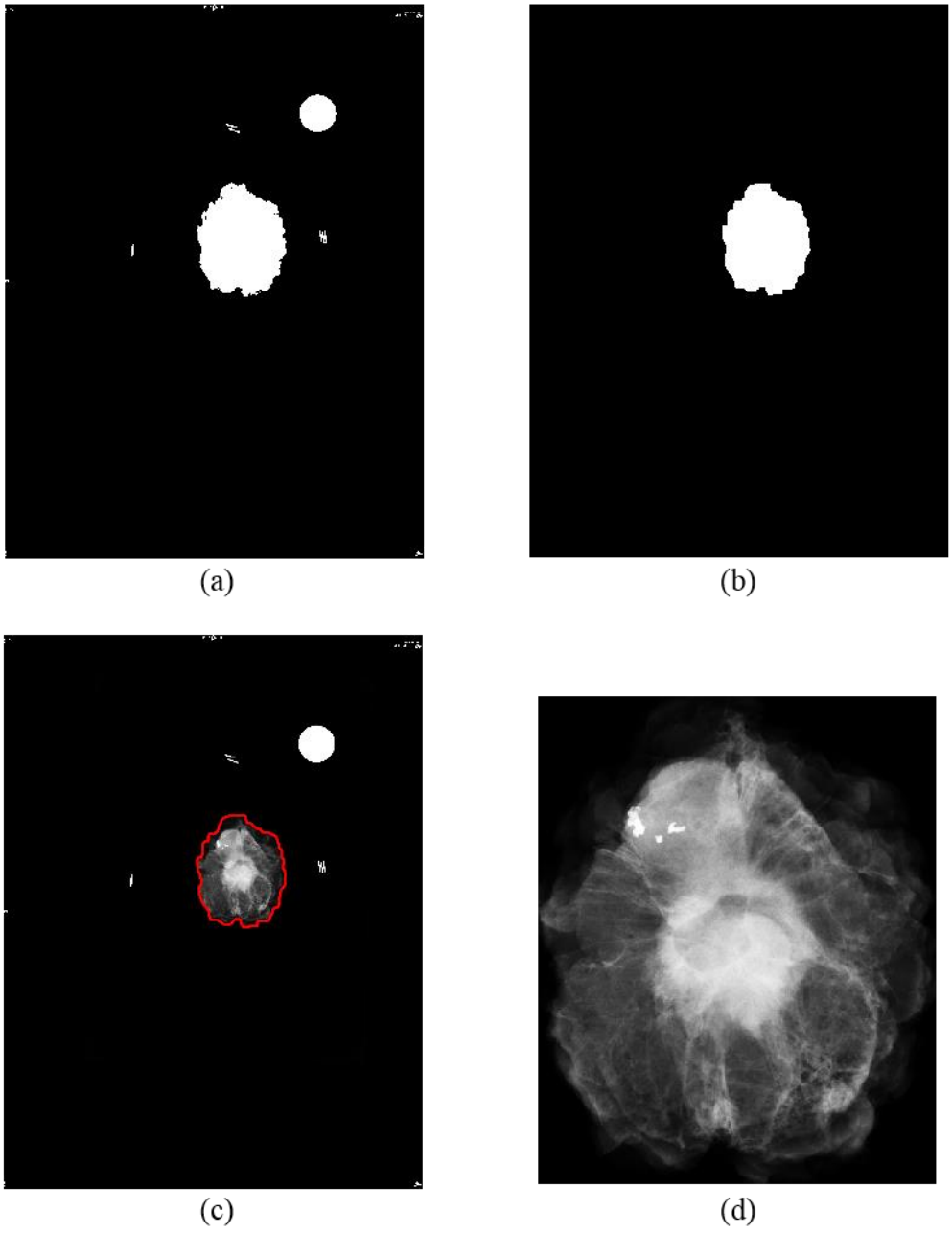


Figure 2.3. Result of specimen boundary detection and ROI extraction: (a) binary image, (b) specimen region, (c) extracted specimen boundary (red) and (d) extracted ROI

Morphological image processing is a collection of non-linear operations [12] related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suitable for the processing of binary images. Morphology with two types of sets of pixels are used in image processing: objects and structure elements (SE). Typically, objects are defined as sets of foreground pixels. The SE is positioned at all possible locations in the image and it is compared with the corresponding neighborhood of pixels. Some operations test whether the element "fits" within the neighborhood, while others test whether it "hits" or intersects the neighborhood. Morphological operation forms a new image with the same size as original image which only contain zero (background value) initially, then slide SE over original image at each increase, if the operation test is successful, mark the region of original image as one (foreground value) on the new image.

Erosion and dilation are two fundamental operations of the morphological processing [11]. Generally, erosion operator is used to shrink the components of a set. The expression is written as following, with A and B as sets in Z^2, the erosion of A by B, denoted $A \ominus B$, is defined as

$$A \ominus B = \{z \mid (B)_z \subseteq A\}, \tag{1}$$

where A is a set of foreground pixels, B is a structure element, and the z's are foreground values. On the contrary, dilation expands the components of a set. The expression is written as following, with A and B as sets in Z^2, the dilation of A by B, denoted $A \oplus B$, is defined as

$$A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\}. \tag{2}$$

This equation is based on reflecting B about its origin and shifting this reflection by z. The dilation of A by B then is the set of all displacements, z, such that $\hat{B}$ overlap at least one element of A.

Opening and closing are two other important morphological operations, which are composed of the erosion and dilation. Opening generally smooth the contour of an object, breaks narrow bridges, and eliminates thin protrusions. The opening of set A by structure element B, denoted $A \circ B$, is defined as

$$A \circ B = (A \ominus B) \oplus B. \tag{3}$$

Thus the opening A by B is the erosion of A by B, followed by a dilation of the result by B. Similarly, closing also tends to smooth section of contours, but the difference is that closing fuse narrow breaks, eliminates small holes, and fill the gaps in the contour. The closing of set A by structure element B, denoted $A \cdot B$, is defined as

$$A \cdot B = (A \oplus B) \ominus B. \tag{4}$$

The closing of A by B is the dilation of A by B, followed by erosion of the result by B. Figure 2.4 demonstrates the opening and closing procedure.

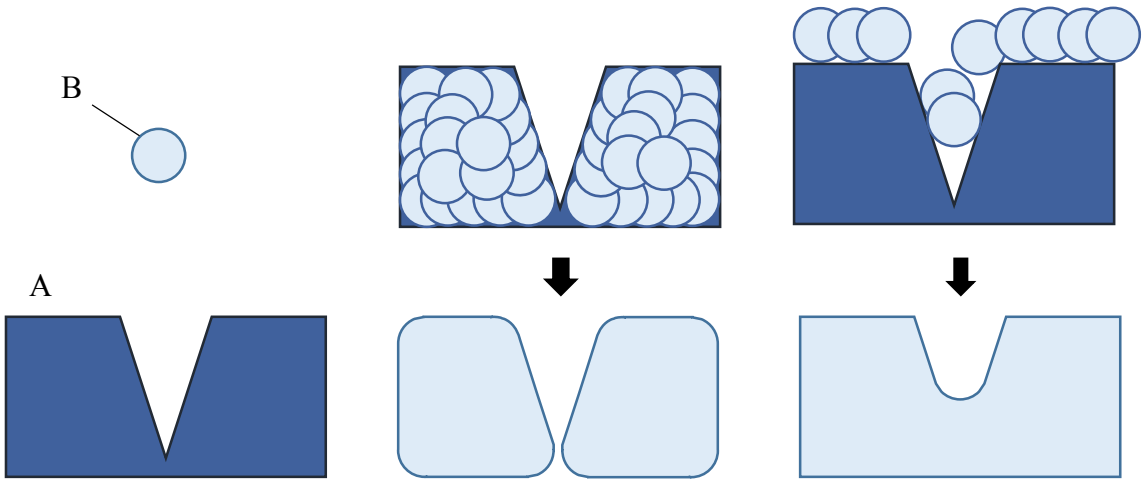


Figure 2.4. (a) Object A and structure element B, (b) opening of A by B and (c) closing of A by B

2.5. Tumor boundary detection

The varying quality of specimen mammography makes tumor boundary detection becoming a difficult task. In order to overcome the conditions where specimen mammography has varied contrast, this study performed five contouring methods to sketch tumor boundary, i.e. multi-thresholding, K-means clustering [13], region-growing [14], U-net [15] and SegNet [16].

2.5.1. Multi-thresholding

The threshold technique is the simplest one in segmenting methods, which partitioning images directly into regions based on intensity value of every pixel. Generally, the pixel intensities in an 8-bits grey-scale image ranges between 0 and 255. A specimen mammogram contains at least three regions, i.e. background, normal tissue, and cancerous tissue (see Fig. 2.6(b) as an example). This study utilized the multi-thresholding classifies a point (x, y) as belonging to the background if $f(\mathrm{x,y}) \leq T_1$, to normal tissue class if $T_1 < f(x,y) \leq T_2$, and to cancerous tissue class if $f(x,y) > T_2$. That is, the segmented image is given by

$$g(x,y) = \begin{cases} c1, & \text{if } f(x,y) \leq T_1 \\ c2, & \text{if } T_1 < f(x,y) \leq T_2 \\ c3, & \text{if } f(x,y) > T_2 \end{cases}, \tag{5}$$

where c1, c2, and c3 are three distinct intensity values. T_1 and T_2 are predefined thresholds, which is given by

$$T_i = \frac{255}{\mathrm{NC}} \times i, \tag{6}$$

where NC is the number of classes, and i = 1, 2..., NC-1. The pixels were combined into homogenous regions according to the intensity levels of the regions.

However, the fixed thresholds were not suitable for this study. Tissue density change from a mammogram to another one, either the mammograms of the same patients with different ages, either mammograms of different patients, hence the number of classes changes. Therefore, this study performed dynamic algorithm [17] to solve the problem. Below more details about the

steps of the proposed dynamic algorithm are given:

Step 1: The number of classes is initialized by three (NC=3) and the multi-threshold is applied for partitioning the ROI into three classes, the highest class is selected and calculate the sum of pixels as Sum(NC).

Step 2: Increment the number of classes (NC=NC+1) and the multi-threshold is reapplied for partitioning ROI to K classes. Select the new highest class and calculate the sum of pixels as Sum(NC+1).

Step 3: The sum of pixels when the number of classes equals (NC) is compared with the sum of pixels when the number of classes equals (NC+1).

If the difference between Sum (NC) and Sum (NC+1) was less than 5%, the number of classes equal (NC) was taken and finished the algorithm, else Step 2 was repeated. The criterion for stopping this algorithm is that the classes containing the tumor has become stable, i.e. the classes almost obtained the same regions. The morphological operators opening and closing were utilized to exclude undesired regions and extract the region of tumor. Figure 2.5 shows the dynamic algorithm apply on multi-threshold. Figure 2.6 shows the example of tumor detection using the multi-thresholding.

Figure 2.5. Dynamic algorithm applies on multi-thresholding: When NC=3, the pixel intensity was divided into three parts with two thresholds (T_1 and T_2); When NC=4, the pixel intensity was divided into four parts with three thresholds (T_1, T_2 and T_3)

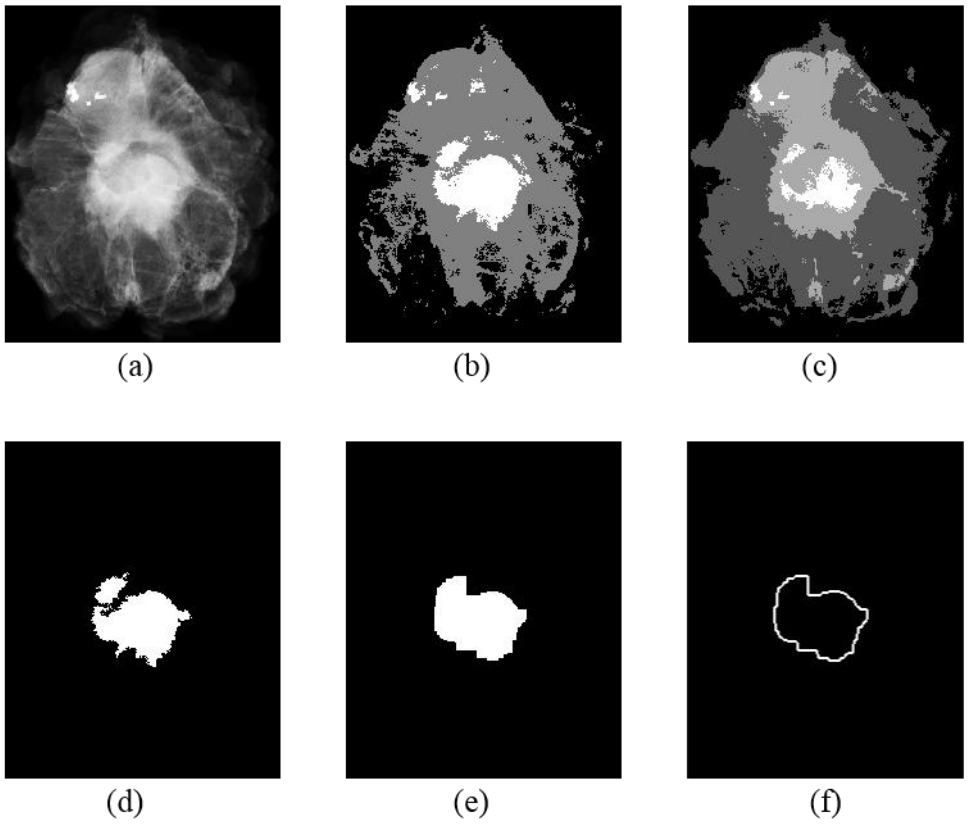


Figure 2.6. Example of tumor detection using the multi-thresholding method: (a) ROI result after specimen boundary detection, (b) ROI partitioned into three classes, (c) ROI partitioned into four classes (stop criteria would be checked), (d) extraction of binary tumor image from (b), (e) extracted tumor area and (f) extracted tumor boundary

2.5.2. K-means clustering

Clustering is a method to divide a set of data into a specific number of groups. K-means clustering is one of the popular clustering techniques [13]. The objective of k-means clustering is to partition the set of observations into k disjoint cluster sets $C = \{C_1, C_2, \dots, C_k\}$,the objective function is defined as

$$\arg\min_{C}\left(\sum_{c=1}^{k}\sum_{i=1}^{n}\|x_i - \mu_c\|^2\right), \tag{7}$$

where μ_c is the mean of points in *x*, also is the centroid of *x*'s cluster. Typically, the Euclidean norm is used, so the term $\|x_i - \mu_c\|$ is familiar Euclidean distance from a sample in C_i to mean μ_c. In words, it is an iterative algorithm that minimizes the sum of the distance of each object to its cluster center. The procedure of K-means clustering starts with specifying number of clusters K, then randomly selected K centers, which are used as the beginning points for every cluster. The next step is to assign each object to the group that has the closest center. When all objects have been assigned, recalculate the positions of the K centroids. Generally, the procedure stops creating clusters until the centroids have stabilized.

The drawback of K-means clustering is that the initial center usually selected randomly with global cluster. On the other words, different initial center can result in different final clusters. To overcome this obstacle, these centers should be placed in a cunning way, which means placing these centers as much as possible far away from each other. This study utilized the local boundary to generate initial centers. For example, the number of classes NC was set as three which divided the pixel intensities into three parts, then randomly select initial centers in three parts, respectively. The pixels were combined into homogenous regions according to the centers. The proposed method performed the multi-thresholding method to divide the pixel intensity into three parts. As mention before, the dynamic algorithm and morphological operators were performed to extract the tumor region more completely.

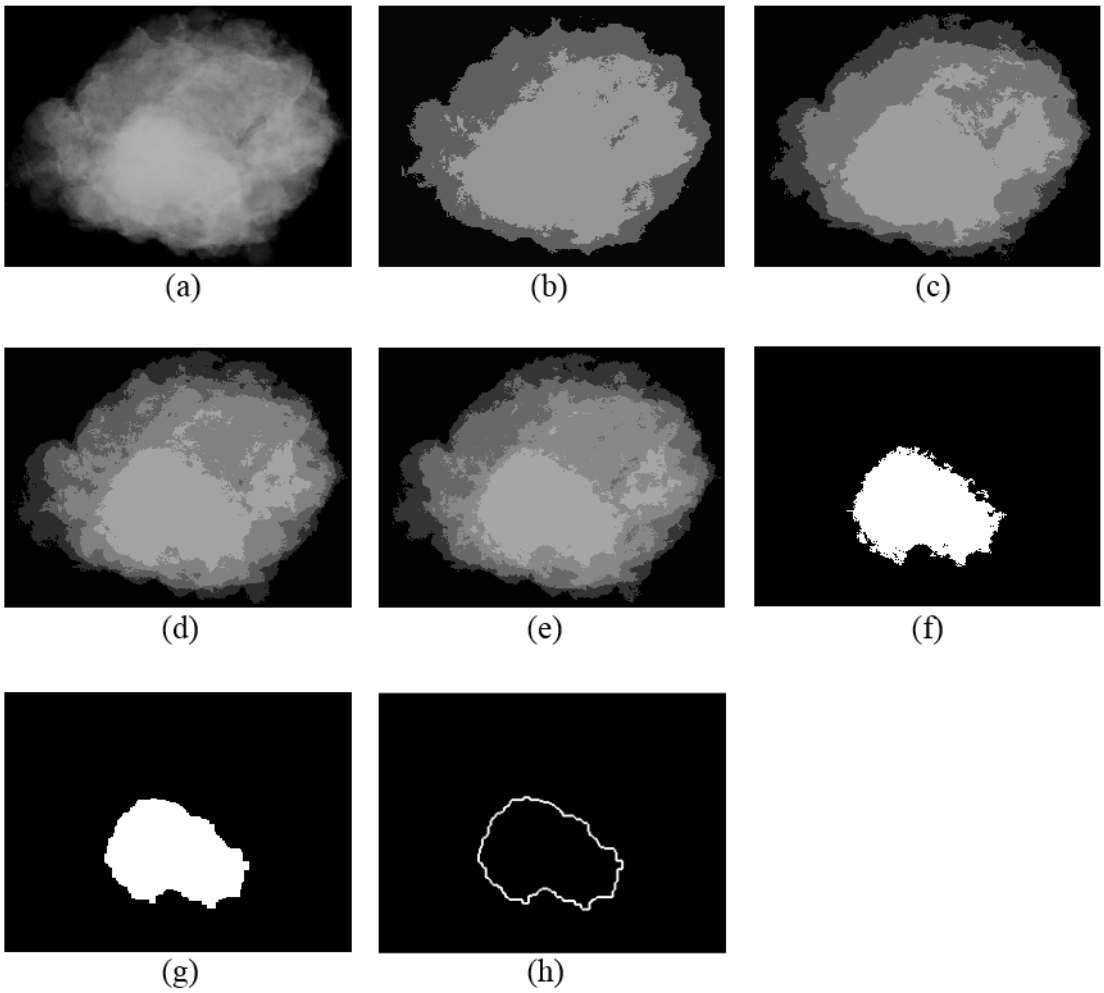


Figure 2.7. Example of tumor detection using k-means clustering method: (a) ROI result after specimen boundary detection, (b) ROI partitioned into three classes, (c) ROI partitioned into four classes, (d) ROI partitioned into five classes, (e) ROI partitioned into six classes (stop criteria are checked), (f) extraction of binary tumor image from (d), (g) extracted tumor area and (h) extracted tumor boundary

2.5.3. Region growing

Region growing [14] is one of the most simple and popular algorithms for region based segmentation. The process established from a seed point and then the region would grow by appending to each seed those neighboring pixels that similar to the seed. Let $\mathrm{I}(\mathrm{x},\mathrm{y})$ denotes an input image; $\mathrm{S}(\mathrm{x},\mathrm{y})$ denotes a seed array containing 1's at the locations of seed points and 0's elsewhere; and Q denotes a predicate to be apply at each location (x, y). Arrays I and S are assumed to be of the same size. A basic region-growing based on 8-connectivly may be stated as follows.

1. Selecting the initial seed point in $\mathrm{S}(\mathrm{x},\mathrm{y})$ and label as 1. All other pixels in S are labeled 0.
2. Form an image I_Q such that, at each point(x, y), $I_Q(x, y) = 1$ if the input image satisfies a given predicate *Q*, at those coordinates, and $I_Q(x, y) = 0$ otherwise.
3. Let g be an image shaped by adding to each seed point in s all the 1-vales points in I_Q that are 8-connected to that seed point.
4. Label each connected component in g with a different region label (e.g., 1, 2, 3, ...). This is the segmented image acquired by region growing.

However, most of the region growing methods require manually selecting the initial seed point. Considering the stability of the segmentation, the proposed method selected the initial seed point automatically. Tumor or dense tissue usually appear brighten than surrounding area on specimen mammography, this is because tumor and dense tissue are denser than fat which will stop more x-ray photons. Expecting the regions containing tumor, the K-means clustering algorithm is used to find the cluster with the highest intensity level, the center of mass in the cluster is calculated as a seed point.

The center of mass is a position defined relative to an object or system of objects. It is the average position of all the parts of the system, weighted according to their masses, that is

$$COM_x = \frac{1}{M}\sum_{i=1}^{M} m_i x_i \quad \text{and} \tag{8}$$

$$COM_y = \frac{1}{M}\sum_{i=1}^{M} m_i y_i, \tag{9}$$

where *COM* is the center of mass, *M* is the sum of the masses, m_i is the weight according to the masses, x_i and y_i are the position of the masses.

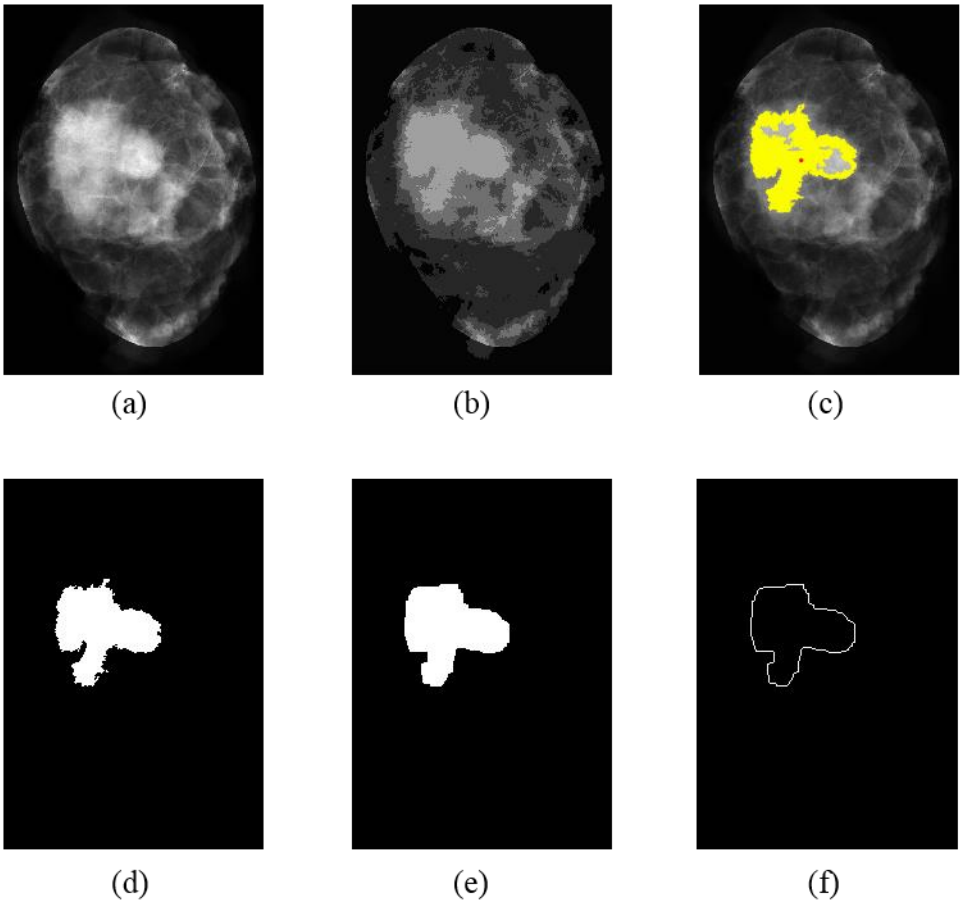


Figure 2.8. Example of tumor detection using region growing method: (a) ROI result after specimen boundary detection, (b) K-means algorithm was applied to find the cluster with the highest intensity level (light gray region), (c) center of mass were calculate as a seed point (red) and the region growing result (yellow) (d) extraction of binary tumor image from (d), (e) extracted tumor area and (f) extracted tumor boundary

2.5.4. U-net

Deep networks have seen huge success lately in the field of natural image segmentation, such as PSPNet [18], RefineNet [19] and Deeplab [20], while only minority deep models like U-net [15] achieved success in medical image segmentation. It is still a very challenging task due to two reasons. First, there are few available labelled data. Second, it is costly and time-consuming to collect images for medical segmentation task since the marking work must be done by professional radiologists or doctors. U-net has won two challenge at the ISBI 2015 and also has outstanding performance in biomedical image segmentation. There are many applications of U-Net in biomedical image segmentation, such as brain image segmentation [21, 22] and liver image segmentation [23]. The model is an improved version of the full convolutional neural network (FCN) [24], which means using convolution instead of the fully connected layer. This strategy allows input any size of images, and the output is also a picture.

The U-Net owes its name to its symmetric shape, and the architecture composes three parts: contracting path, bottleneck, and expanding path. The contracting path is composed of four blocks, each block is composed of convolution layer (with batch normalization) and max pooling. The number of feature maps doubles at each pooling, starting with 64 feature maps for the first block, 128 for the second, and so on. The contracting path aims to capture the context of the input image, and the contextual information will be transferred to the up-sampling path by means of skip connections. The bottleneck is between the contracting and expanding paths, it is built from simply two convolutional layers (with batch normalization), with dropout. The expanding path is also composed of four blocks, each of these blocks is composed of deconvolution layer (up-sampling) and concatenation followed by convolution layer. Concatenation with the corresponding cropped feature map from the contracting path is to enable precise localization combined with contextual information from the contracting path. The architecture of U-Net is shown in Fig. 2.9.

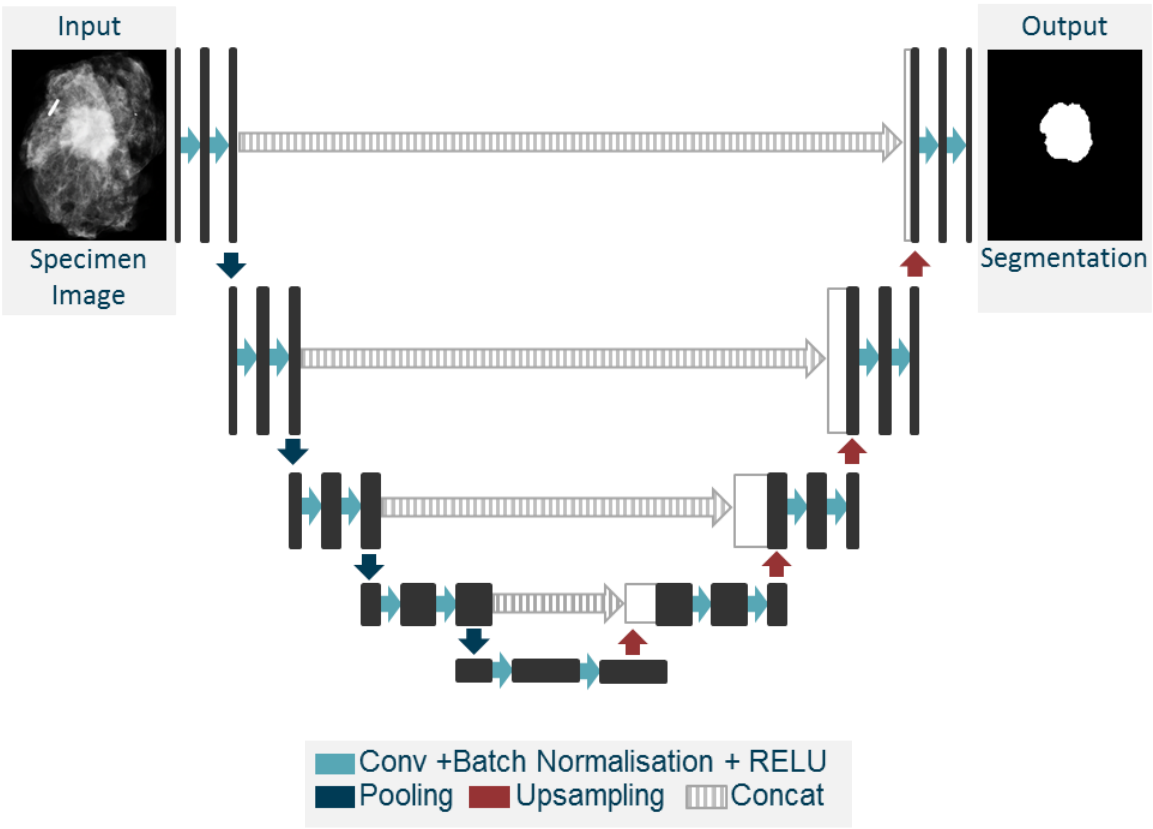


Figure 2.9. U-net architecture

2.5.5. SegNet

SegNet [16] is also a deep convolutional network proposed by Cambridge University to address image semantic segmentation for autonomous driving or intelligent robots. Due to the advantages of retaining high frequency details in the segmented images and also reducing the total number of trainable parameters in the decoders, SegNet has been used in medical image segmentation recently , such as gland segmentation in colon Cancer [25] and blood Cell Images Segmentation [26]. The model is designed based on FCN. SegNet is composed of a symmetry network: the encoder and the decoder. The architecture of SegNet is shown in Fig. 2.10.

The structure of Encoder is similar to VGG-16 [27], it composed of three kinds of network: convolution, batch normalization and pooling. Convolution layers are used to extract local features; Batch normalization layers are used to expedite learning; and Pooling layers are utilized to down sampling feature map. Decoder aims to map the low-resolution feature maps from the encoder to obtain the same resolution as the input image feature map for pixel-level classification. The highlight of SegNet is that the decoder utilized max-pooling indices from the corresponding encoder stage to up-sample, this gives reasonably good performance and is space efficient, and this is also why SegNet was selected as one of the methods in this study.

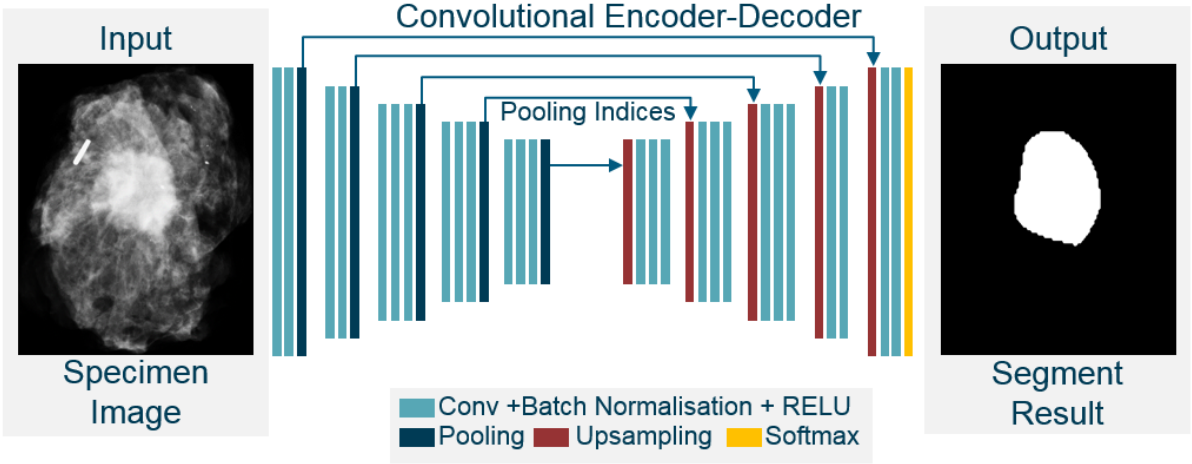
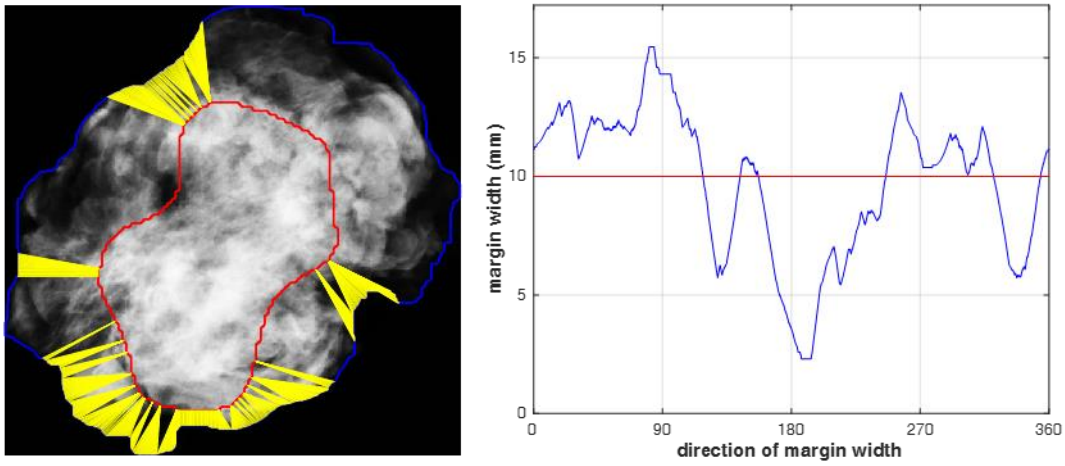


Figure 2.10. SegNet architecture

This study performed U-net and SegNet to contour the tumor in specimen mammography. Due to the small dataset, data augment technique was used to create new images [28]. In this work, 20 new images were generated from each case and resulting in 600 images. Combination of flipping, rotation, distortion and zoom transformations were performed randomly. Both U-net and SegNet utilized a pre-trained VGG16 model as the encoder part, thus could benefit from the features already created in the model and only focus on learning the specific decoding features. The proposed method used a mini-batch of 10 images, learning rate of 0.001 and the Adam optimizer. In order to maintain a fair measure of the performance of the convolutional networks, the leave-one-out cross validation was applied on U-net and SegNet segmentation. Finally, the morphological operator erosion was used to figure the obtained tumor boundaries.

2.6. Margin width evaluation

The distance between specimen boundary and tumor boundary was estimated as margin width. This study evaluated the margin width by the Euclidean distance [29]. In an image coordinate plane, the distance between two points is usually given by the Euclidean distance (2-norm distance). The distance from a point to a line is the shortest distance from a fixed point to any point on a fixed line in Euclidean geometry. In this study, the safety margin width recommended as 10mm. When the margin width is less than 10mm, the system would display the area in yellow. Figure 2.11 illustrates the segmentation and estimation results.



(a) (b)

Figure 2.11. (a) Segmentation result (extracted specimen boundary (blue), tumor boundary(red) and the region less than margin width (yellow)) and (b) the evaluated margin width

CHAPTER 3

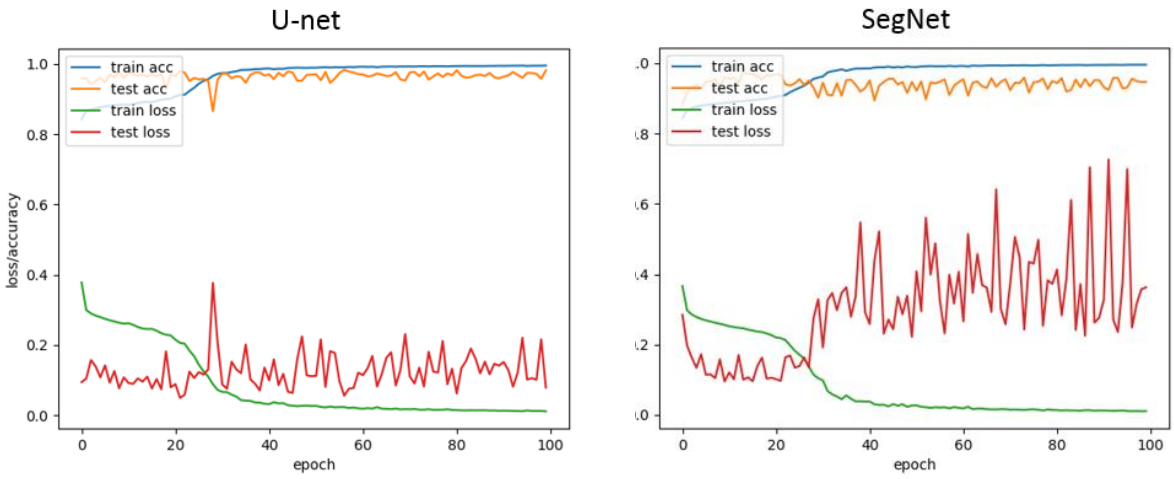
RESULTS

This study totally experimented 30 cases with manual sketched boundaries to evaluate the accuracy of the proposed method. In this work, measurement of pixel density was first performed to converted pixel resolution to millimeter and the specimen detection was applied to obtain the ROI. This study proposed the five contouring approaches to obtain the tumor boundaries. Figures 3.1-3.3 (b) demonstrate the final result applied the proposed contouring methods on various cases. The comparison of computational time consists of training time and testing time. Since traditional segmentation methods don't need training, only the execution time needs to be compared. As shown in Table 3.1, average execution time of each approaches are less than 7 seconds, which means the proposed system is suitable for intra-operative tumor margin evaluation. The traditional segmentation methods were implemented by Matlab (R2016a, MathWorks Inc., MA). The deep learning networks were trained on an Nvidia 1080Ti GPU. All methods were performed on a single CPU Intel i7 3.6 GHz personal computer with Microsoft Windows 7 operating system.

Table 3.1. A comparison of computational time

Method	**Average training time(sec)**	**Average execution/testing time (sec)**
Multi-thresholding	-	5.20
K-means	-	6.03
Region-growing	-	6.52
U-net	1148.47	4.82
SegNet	1059.60	4.77

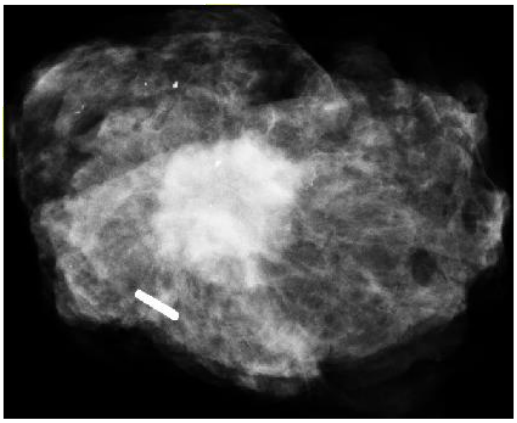
Learning curves show the performance of our models on training and validation sets. The basics of two types of learning curves are loss curve and accuracy curve. Loss curves, which compare the error of the loss function in the training and validation sets. Accuracy curves, which compare the performance of the model according to a specific metric (accuracy) on training and validation sets. The learning curves of U-net and SegNet are shown in Fig. 3.1. Although the plot of testing accuracy always close to training accuracy, the plot of training loss continues to decrease with experience and the validation loss decreases to a point and begins unstable. This situation could be identified as an overfitting model. There were many reasons give rise to overfitting, we inferred that is due to the size of dataset is small. Thus, this study expected the specimen mammography dataset would be expanded much more, then the algorithm has the potential to come up with a better model in the future.



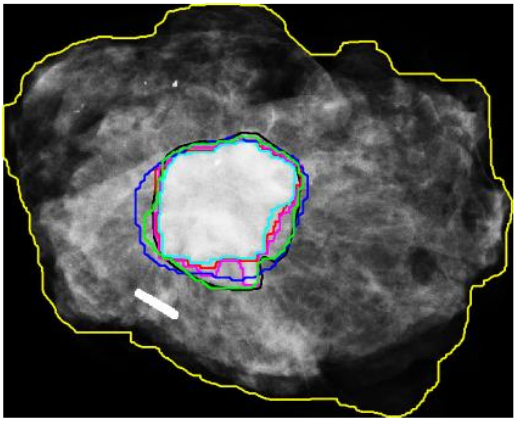
(a) (b)

Figure 3.1. Learning curve of (a) U-net and (b) SegNet

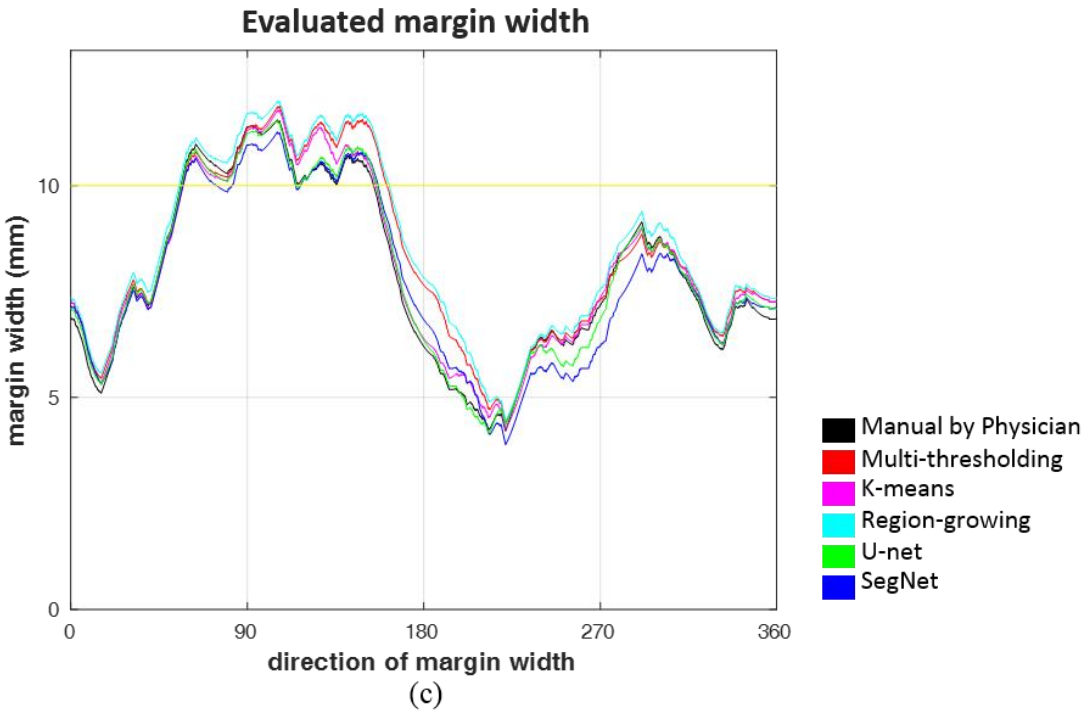
After obtaining the specimen boundary and tumor boundary, the margin width was estimated by Euclidean distance. Figures 3.2-3.4(b) show the evaluation results (in pixel) using the proposed five methods. The results were compared with manual delineations from experienced physician. The measurements of distance are the Euclidean distance. The length



(a)



(b)



(c)

unit is millimeter.

Figure 3.2. Case 01: (a) original image, (b) final segmentation result and (c) evaluated margin width between the five contouring methods and the manual sketching by physician (black: manual sketching, red: multi-thresholding, magenta: K-means, cyan: region-growing, green:

U-net, blue: Segnet)

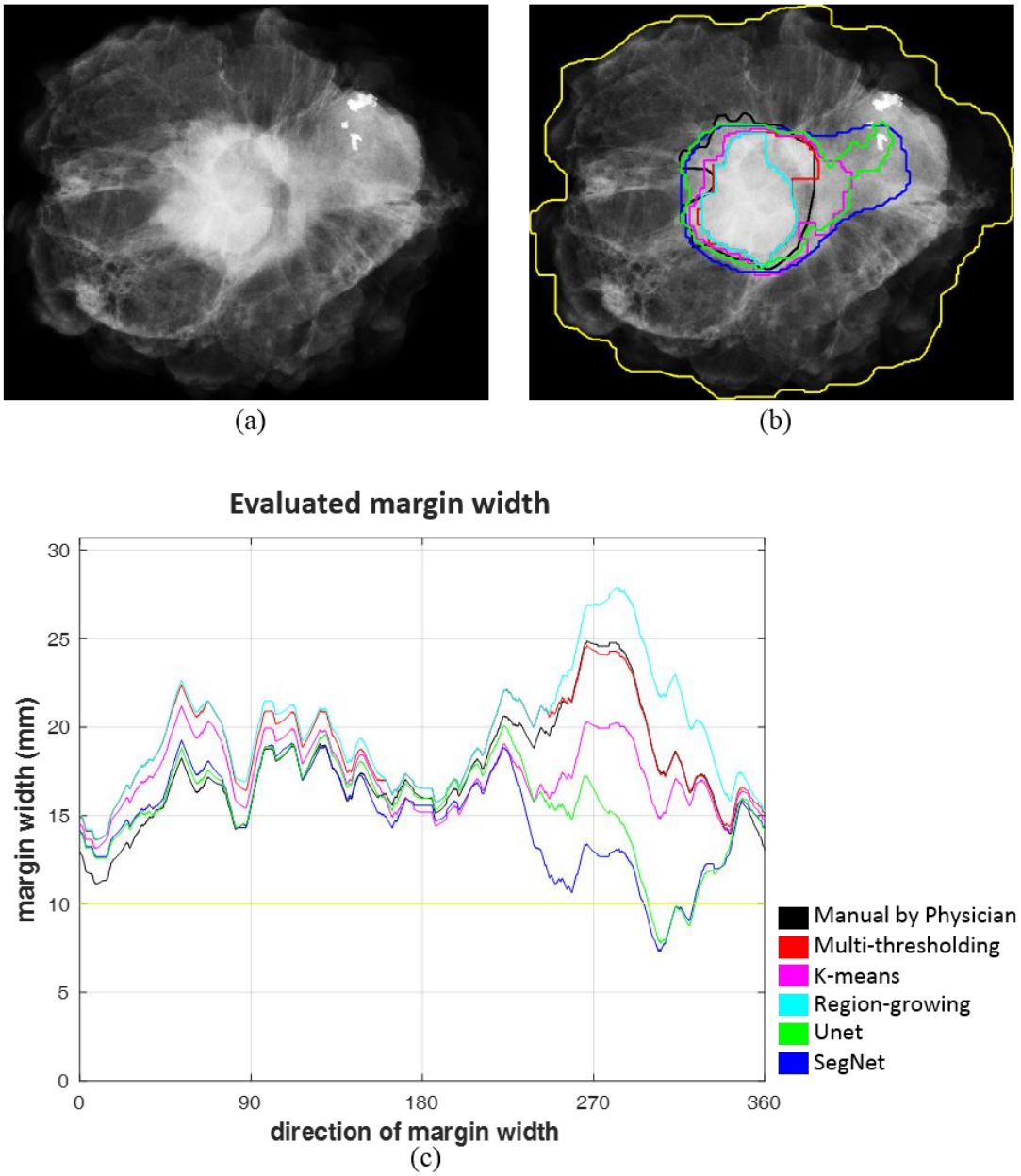


Figure 3.3. Case 02: (a) original image, (b) final segmentation result and (c) evaluated margin width between the five contouring methods and the manual sketching by physician (black: manual sketching, red: multi-thresholding, magenta: K-means, cyan: region-growing, green: U-net, blue: Segnet)

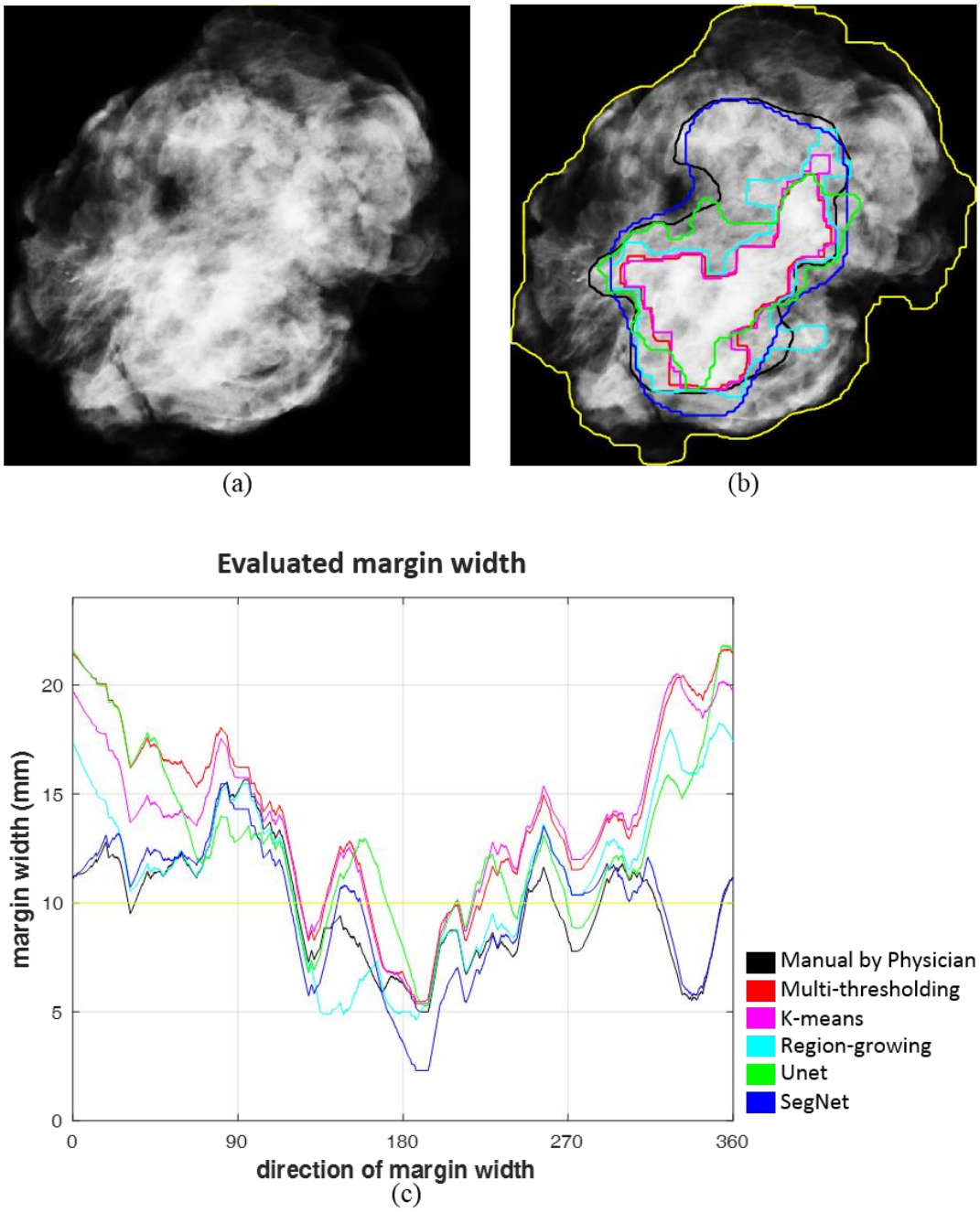


Figure 3.4. Case 05: (a) original image, (b) final segmentation result and (c) evaluated margin width between the five contouring methods and the manual sketching by physician (black: manual sketching, red: multi-thresholding, magenta: K-means, cyan: region-growing, green: U-net, blue: Segnet)

Four practical similarity measures [30], i.e. the similarity index (SI), overlap value (OV), overlap fraction (OF) and extra fraction (EF) between the manually determined boundaries and the automatically detected boundaries, were calculated for quantitative analysis of the contouring results. An abstract specimen mammogram is shown in Fig. 3.5 to illustrate the similarity measures. The REF indicates the results sketched by the experienced physician manually, and the SEG denotes the results describe by proposed method. Then the similarity index SI, OV, OF and EF are respectively defined as

$$\mathrm{SI} = \frac{2 \times (REF \cap SEG)}{REF + SEG} \times 100\%, \tag{10}$$

$$\mathrm{OF} = \frac{REF \cap SEG}{REF} \times 100\%, \tag{11}$$

$$\mathrm{OV} = \frac{REF \cap SEG}{REF \cup SEG} \times 100\%, \tag{12}$$

$$\mathrm{EF} = \frac{\overline{REF} \cap SEG}{REF} \times 100\%, \tag{13}$$

where SI, OF, OV approach to 1, and EF approach to 0, it indicates that the tumor area segmented by our method is similar to the physician manual sketch. Overlap area represents the area covered by SEG and REF, extra area represents the false positive area and missing area represents false negatives area. Tables 3.2-3.6 show the four similarity measures of all cases.

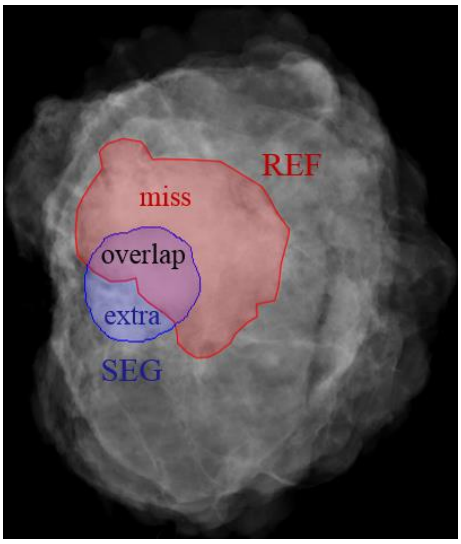


Figure 3.5. Abstract ROI. REF represent the results sketched by the experienced physician manually, and SEG represents the results describe by proposed method

Table 3.2. The similarity results of multi-threshold which compared with the physician manual sketching

Multi-threshold				
Case #	SI	OF	OV	EF
1	0.86	0.76	0.76	0.01
2	0.81	0.69	0.68	0.01
3	0.57	0.51	0.40	0.29
4	0	0	0	0.16
5	0.56	0.39	0.39	0
6	0.31	0.18	0.18	0
7	0.11	0.06	0.06	0
8	0.45	0.30	0.29	0.02
9	0.06	0.03	0.03	0
10	0.90	0.85	0.82	0.04
11	0.82	0.88	0.70	0.26
12	0.62	0.45	0.45	0
13	0.17	0.09	0.09	0
14	0.62	0.73	0.45	0.61
15	0.92	0.98	0.86	0.14
16	0.70	0.76	0.54	0.41
17	0.81	0.76	0.69	0.11
18	0.57	0.40	0.40	0
19	0.75	0.77	0.60	0.28
20	0.85	0.78	0.74	0.06
21	0	0	0	0.05
22	0.20	0.11	0.11	0
23	0.35	0.21	0.21	0
24	0.39	0.24	0.24	0
25	0.70	0.56	0.54	0.03
26	0.80	0.67	0.67	0
27	0.53	0.98	0.36	1.72
28	0.42	0.26	0.26	0
29	0.85	0.80	0.74	0.08
30	NaN	NaN	NaN	NaN
Average	**0.54**	**0.49**	**0.42**	**0.15**

Table 3.3. The similarity results of K-means which compared with the physician manual sketching

K-means				
Case #	SI	OF	OV	EF
1	0.86	0.76	0.76	0.01
2	0.81	0.69	0.68	0.01
3	0.57	0.51	0.40	0.29
4	0	0	0	0.16
5	0.56	0.39	0.39	0
6	0.31	0.18	0.18	0
7	0.11	0.06	0.06	0
8	0.45	0.30	0.29	0.02
9	0.06	0.03	0.03	0
10	0.90	0.85	0.82	0.04
11	0.82	0.88	0.70	0.26
12	0.62	0.45	0.45	0
13	0.17	0.09	0.09	0
14	0.62	0.73	0.45	0.61
15	0.92	0.98	0.86	0.14
16	0.70	0.76	0.54	0.41
17	0.81	0.76	0.69	0.11
18	0.57	0.40	0.40	0
19	0.75	0.77	0.60	0.28
20	0.85	0.78	0.74	0.06
21	0	0	0	0.05
22	0.20	0.11	0.11	0
23	0.35	0.21	0.21	0
24	0.39	0.24	0.24	0
25	0.70	0.56	0.54	0.03
26	0.80	0.67	0.67	0
27	0.53	0.98	0.36	1.72
28	0.42	0.26	0.26	0
29	0.85	0.80	0.74	0.08
30	NaN	NaN	NaN	NaN
Average	**0.66**	**0.66**	**0.53**	**0.36**

Table 3.4. The similarity results of Region-growing which compared with the physician manual sketching

Region-growing				
Case #	SI	OF	OV	EF
1	0.82	0.70	0.70	0
2	0.71	0.56	0.55	0
3	0.66	0.74	0.50	0.50
4	0.75	0.94	0.60	0.57
5	0.73	0.59	0.58	0.02
6	0.79	0.87	0.66	0.32
7	0.65	0.48	0.48	0
8	0.61	0.44	0.44	0
9	0.84	0.76	0.73	0.05
10	0.85	0.76	0.74	0.03
11	0.80	0.67	0.66	0.01
12	0.58	0.41	0.41	0
13	0.76	0.69	0.61	0.12
14	0.42	0.98	0.26	2.74
15	0.59	1	0.42	1.36
16	0.52	0.91	0.35	1.59
17	0.81	0.75	0.68	0.10
18	0.86	0.85	0.75	0.13
19	0.76	0.73	0.61	0.21
20	0.07	0.04	0.04	0.08
21	0.55	0.99	0.38	1.62
22	0.20	0.11	0.11	0
23	0.74	0.78	0.59	0.32
24	0.23	0.15	0.13	0.14
25	0.63	0.46	0.46	0.01
26	0.70	0.54	0.54	0
27	0.60	0.95	0.42	1.25
28	0.45	0.29	0.29	0
29	0.04	0.04	0.02	0.99
30	0.04	1	0.02	46.50
Average	**0.59**	**0.64**	**0.46**	**1.96**

Table 3.5. The similarity results of U-net which compared with the physician manual sketching

U-net				
Case #	SI	OF	OV	EF
1	0.96	0.95	0.93	0.02
2	0.83	0.95	0.71	0.35
3	0.73	0.99	0.57	0.74
4	0.83	0.96	0.71	0.34
5	0.68	0.53	0.51	0.03
6	0.81	0.97	0.68	0.42
7	0.68	0.52	0.52	0
8	0.37	0.23	0.23	0
9	0.89	0.92	0.79	0.16
10	0.81	0.7	0.68	0.04
11	0.49	0.87	0.32	1.68
12	0.88	0.79	0.79	0
13	0.81	0.74	0.68	0.08
14	0.51	0.96	0.34	1.82
15	0.44	1	0.28	2.53
16	0.31	0.89	0.18	3.83
17	0.86	0.93	0.76	0.22
18	0.83	0.99	0.71	0.39
19	0.65	0.96	0.48	0.99
20	0.46	0.30	0.30	0
21	0.68	0.91	0.52	0.75
22	0.28	0.16	0.16	0
23	0.87	0.98	0.77	0.28
24	0.27	0.15	0.15	0
25	0.84	0.74	0.72	0.04
26	0.84	0.73	0.72	0.01
27	0.52	0.98	0.35	1.77
28	0.11	0.06	0.06	0
29	0.33	0.98	0.20	3.85
30	0.01	0.02	0.01	1.68
Average	**0.62**	**0.73**	**0.49**	**0.73**

Table 3.6. The similarity results of SegNet which compared with the physician manual sketching

SegNet				
Case #	SI	OF	OV	EF
1	0.91	0.94	0.84	0.12
2	0.72	0.96	0.56	0.70
3	0.65	1	0.48	1.07
4	0.80	0.96	0.67	0.44
5	0.90	0.93	0.82	0.13
6	0.74	0.83	0.59	0.41
7	0.55	0.38	0.38	0
8	0.73	0.57	0.57	0
9	0.86	0.95	0.76	0.25
10	0.75	0.61	0.60	0.01
11	0.54	1	0.37	1.72
12	0.93	0.97	0.87	0.12
13	0.85	0.91	0.73	0.24
14	0.42	0.97	0.26	2.66
15	0.39	1	0.24	3.14
16	0.25	1	0.14	5.92
17	0.75	1	0.60	0.66
18	0.78	0.98	0.63	0.55
19	0.72	0.70	0.57	0.24
20	0.43	0.27	0.27	0
21	0.69	0.98	0.52	0.87
22	0.36	0.22	0.22	0
23	0.80	0.98	0.66	0.48
24	0.84	0.74	0.73	0.02
25	0.84	0.75	0.73	0.04
26	0.87	0.79	0.77	0.03
27	0.38	1	0.23	3.27
28	NaN	NaN	NaN	NaN
29	0.28	0.93	0.16	4.65
30	0.61	0.92	0.44	1.09
Average	**0.62**	**0.73**	**0.49**	**0.73**

However, it is difficult to obtain a perfect set of labels to serve as a ground truth, due to the complexity of the medical data and the large inter-rater variability [31]. Two expert physicians might disagree on a complex patient case due to differences in clinical training, prior experience and understanding of the disease. Annotation differences can also arise due to the limited amount of time available in annotating large number of cases. Therefore, pathology data was used as a ground truth to compare with proposed methods. Since three cases of the pathological data were incomplete, the other 27 pathological data are used to compare with the proposed method. The average difference between the proposed automatic segmentation and pathologist are listed in Table 3.7.

Table 3.7. The results of proposed five methods compared with pathology margin width

	Average Difference					
Pathology direction	**Manual sketch**	**Multi-threshold**	**K-means**	**Region-growing**	**U-net**	**SegNet**
3 o'clock	7.31± 6.25	11.76± 9.78	8.40± 5.34	7.91± 5.27	7.58± 6.32	7.29± 6.32
6 o'clock	5.44± 4.11	9.19± 9.45	6.22± 6.30	6.96± 7.15	6.02± 6.06	5.09± 4.76
9 o'clock	6.96± 5.85	7.72± 5.40	7.75± 5.88	8.24± 5.89	6.35± 5.16	6.18± 5.48
12 o'clock	6.78± 6.68	9.7± 7.49	9.18± 8.34	8.11± 7.18	7.89± 7.13	7.56± 6.81
Average	**6.62± 5.72**	**9.59± 8.03**	**7.89± 6.47**	**7.81± 6.37**	**6.96± 6.17**	**6.53± 5.84**

The experimental results revealed that SegNet has the smallest average difference. U-net and SegNet have similar resultant values to the manual sketch by doctor, which means U-net and SegNet have the ability to be comparable to doctor. The case with the desirable results using proposed all contouring methods is demonstrated in Fig.3.1 and Fig 3.6.

The case with the best segmentation results using SegNet and U-net is demonstrated in Fig. 3.7. Deep convolutional networks might identify many contrasting features that are highly complex and difficult to describe in words from medical images. However, in some cases, traditional approaches could obtain the better tumor region due to low contrast of the image. The case with the best segmentation results using traditional image segmentation is demonstrated in Figs. 3.8-3.9.

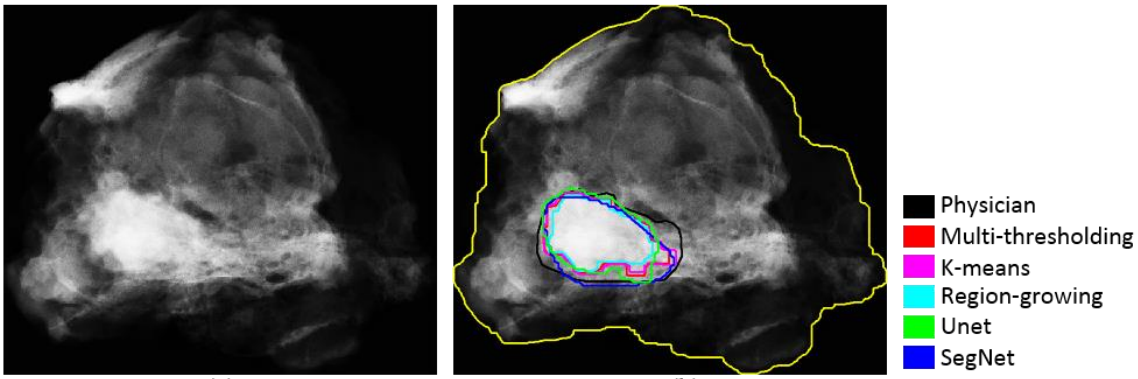


Figure 3.6. The case with the desirable results using the proposed all contouring methods

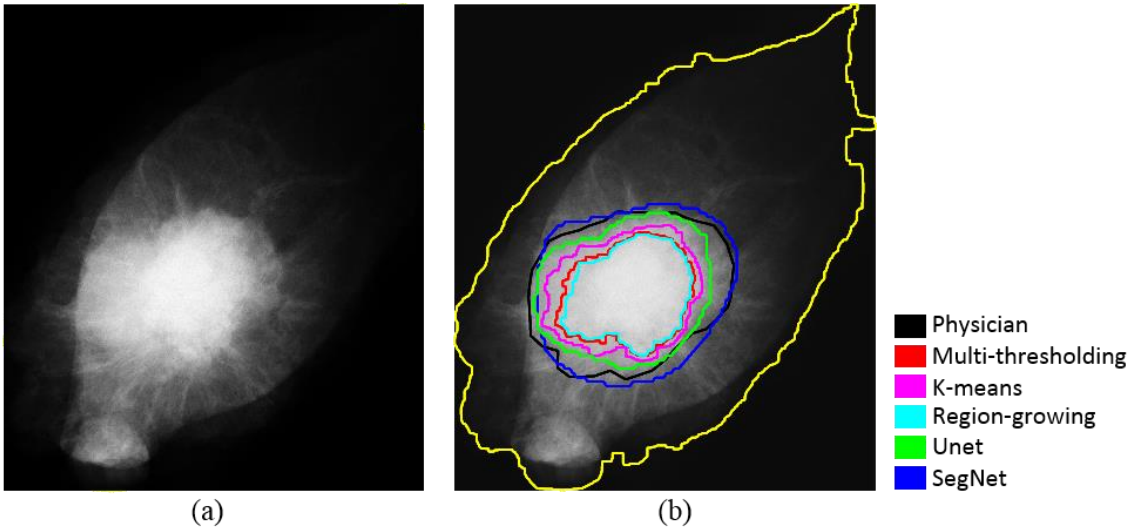


Figure 3.7. The case with the best segmentation results using SegNet and U-net: (a) original image and (b) the segmentation results

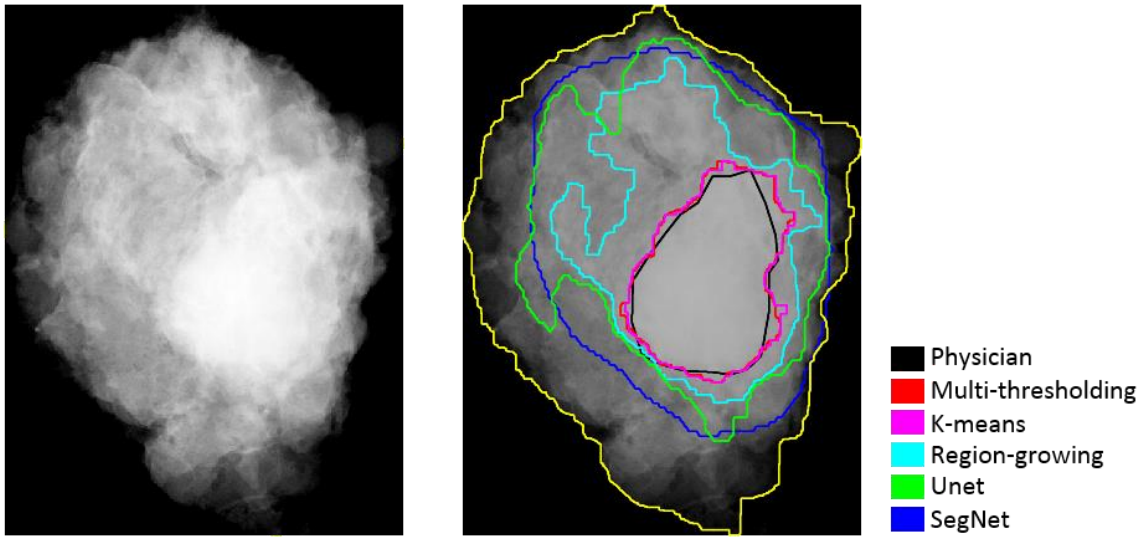


Figure 3.8. The case with the best segmentation results using the multi-thresholding and K-means methods: (a) original image and (b) the segmentation results

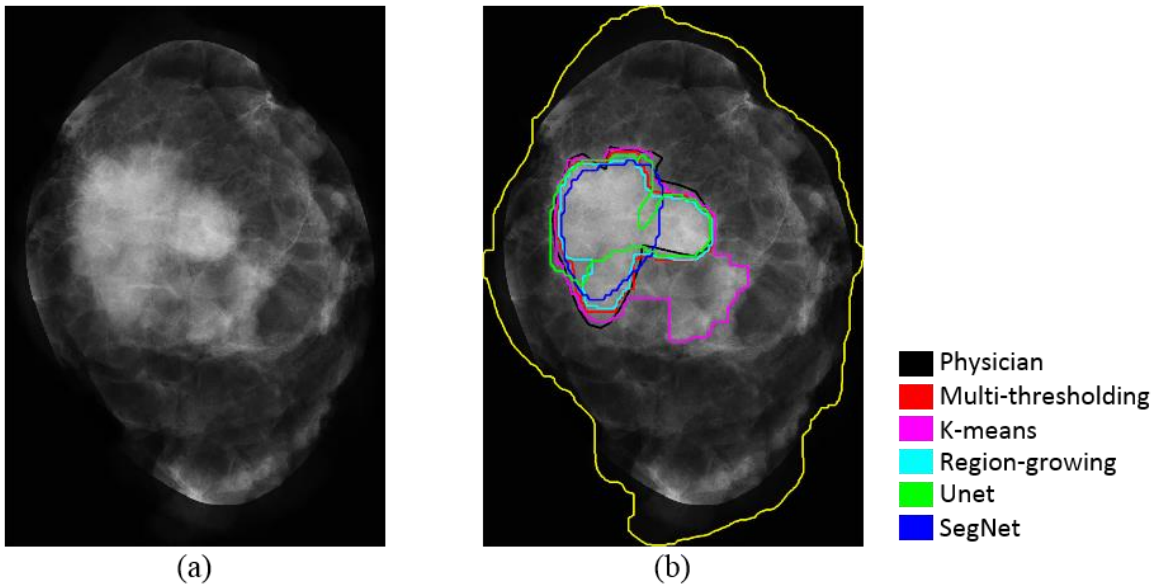
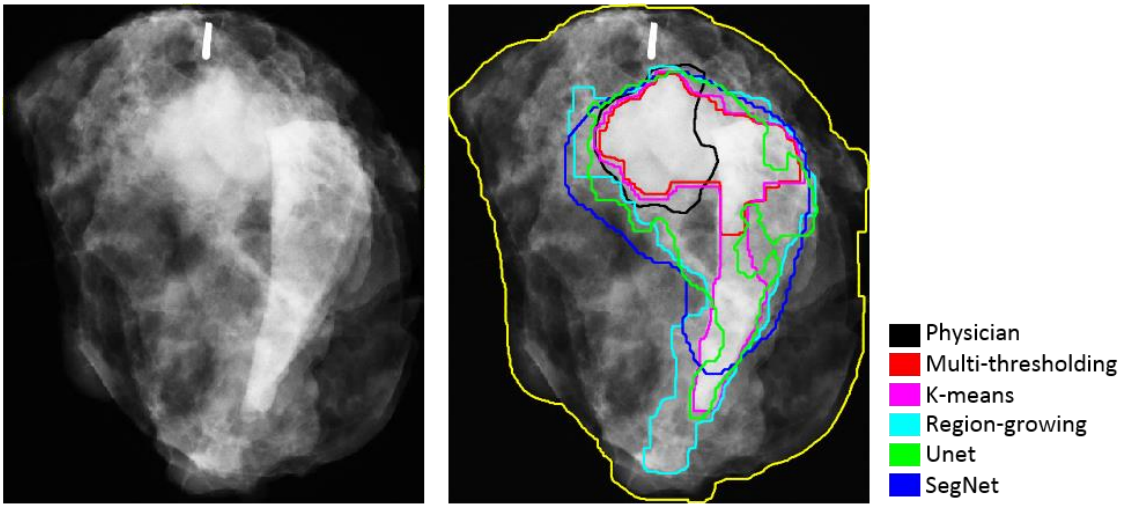


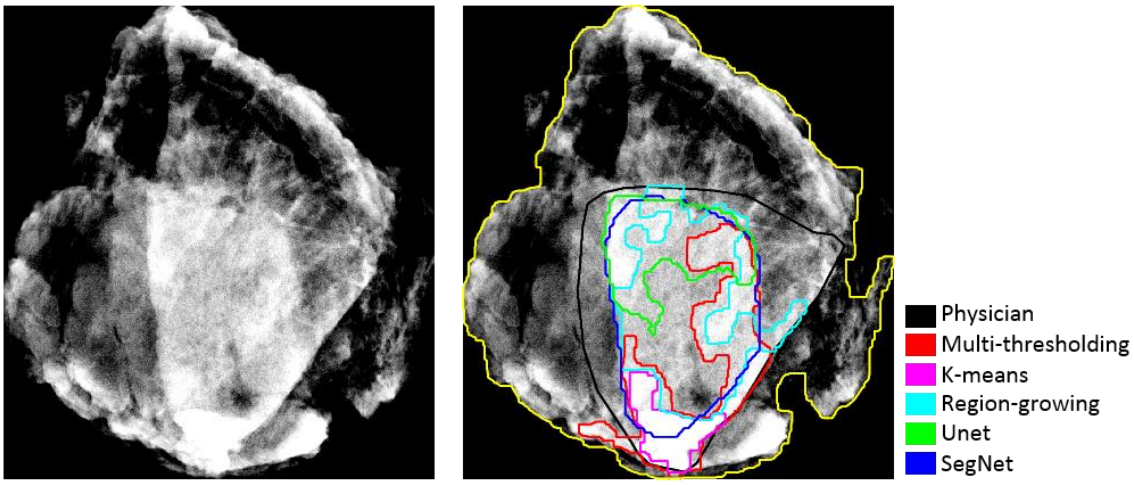
Figure 3.9. The case with the best segmentation results using the multi-thresholding and region growing methods: (a) original image and (b) the segmentation results

The case with the undesirable results using proposed all contouring methods is demonstrated in Fig. 3.10-3.11. The reason for the failure was mainly due to the poor quality of the specimen mammography. In a few special cases, due to the location of the tumor near the nipple, skin, or other high-brightness tissue, the system misidentified other tissues as tumors.



(a) (b)

Figure 3.10. The case with the undesirable results using the proposed all contouring methods: (a) original image and (b) the segmentation results



(a) (b)

Figure 3.11. The case with the undesirable results using the proposed all contouring methods: (a) original image and (b) the segmentation results

CHAPTER 4

CONCLUSIONS

Nowadays, BCT has been one of the common methods of treating breast cancer. During the surgery, the doctor usually measures the margin width to ensure that the tumor is removed clearly. A number of assisted diagnostic systems for measuring tumor margins have been proposed, however these methods require high cost equipment or contrast medium injection for patient. Therefore, this study proposes a fast, low-cost computer-aided methods for detecting tumor boundaries and estimating margin width.

Measurement of pixel density was first applied by estimating coin size. Adaptive thresholding was utilized to eliminate artifacts and obtain rough specimen region. Five contouring approaches were proposed to generate the tumor regions individually. Morphological operators were used to obtain desired specimen and tumor boundary. Evaluated the margin width by the Euclidean distance.

The experimental results revealed that the average difference of deep-learning techniques is more similar to doctor's manual sketching than the traditional approaches, which means the deep-learning techniques could sketch boundary more reasonably. That is the deep-learning techniques have the opportunity to automatically find new features without human intervention. However, in some case, traditional approaches could obtain the better tumor region due to low contrast of the image, which means the methods were complementary. With the aid of deep learning techniques, the proposed scheme would be a potential procedure in intra-operative measurement system.

Reference

[1] R. K. Benda, N. P. Mendenhall, D. S. Lind, J. C. Cendan, B. F. Shea, L. C. Richardson, *et al.*, "Breast-conserving therapy (BCT) for early-stage breast cancer," *J Surg Oncol,* vol. 85, pp. 14-27, Jan 2004.

[2] K. K. Hunt, B. D. Smith, and E. A. Mittendorf, "The Controversy Regarding Margin Width in Breast Cancer: Enough is Enough," *Annals of Surgical Oncology,* vol. 21, pp. 701-703, Mar 2014.

[3] F. T. Nguyen, A. M. Zysk, E. J. Chaney, J. G. Kotynek, U. J. Oliphant, F. J. Bellafiore, *et al.*, "Intraoperative evaluation of breast tumor margins with optical coherence tomography," *Cancer Res,* vol. 69, pp. 8790-6, Nov 15 2009.

[4] F. Schnabel, S. K. Boolbol, M. Gittleman, T. Karni, L. Tafra, S. Feldman, *et al.*, "A randomized prospective study of lumpectomy margin assessment with use of MarginProbe in patients with nonpalpable breast malignancies," *Ann Surg Oncol,* vol. 21, pp. 1589-95, May 2014.

[5] D. W. Shipp, E. A. Rakha, A. A. Koloydenko, R. D. Macmillan, I. O. Ellis, and I. Notingher, "Intra-operative spectroscopic assessment of surgical margins during breast conserving surgery," *Breast Cancer Res,* vol. 20, p. 69, Jul 9 2018.

[6] M. Koller, S. Q. Qiu, M. D. Linssen, L. Jansen, W. Kelder, J. de Vries, *et al.*, "Implementation and benchmarking of a novel analytical framework to clinically evaluate tumor-specific fluorescent tracers," *Nat Commun,* vol. 9, p. 3739, Sep 18 2018.

[7] M. Adhi and J. S. Duker, "Optical coherence tomography--current and future applications," *Curr Opin Ophthalmol,* vol. 24, pp. 213-21, May 2013.

[8] M. Terashima, H. Kaneda, and T. Suzuki, "The role of optical coherence tomography in coronary intervention," *Korean J Intern Med,* vol. 27, pp. 1-12, Mar 2012.

[9] T. H. Tsai, J. G. Fujimoto, and H. Mashimo, "Endoscopic Optical Coherence Tomography for Clinical Gastroenterology," *Diagnostics (Basel),* vol. 4, pp. 57-93, May 5 2014.

[10] J. T. McCormick, A. J. Keleher, V. B. Tikhomirov, R. J. Budway, and P. F. Caushaj, "Analysis of the use of specimen mammography in breast conservation therapy," *Am J Surg,* vol. 188, pp. 433-6, Oct 2004.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (4th Edition)*: Pearson, 2017.

[12] E. R. D. J. T. Astola, "An Introduction to Nonlinear Image Processing," *SPIE publication,* 1994.

[13] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-Means Clustering Algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics),* vol. 28, pp. 100-108, 1979.

[14] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 641-647, 1994.

[15] O. F. Ronneberger, Philipp; Brox, Thomas, "U-Net: Convolutional Networks for Biomedical Image Segmentation" *MICCAI*, 2015.

[16] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 2481-2495, 2017.

[17] A. Elmoufidi, K. El Fahssi, S. Jai-andaloussi, A. Sekkaki, Q. Gwenole, and M. Lamard, "Anomaly classification in digital mammography based on multiple-instance learning," *Iet Image Processing*, vol. 12, pp. 320-328, Mar 2018.

[18] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. (2016, December 01, 2016). Pyramid Scene Parsing Network. *arXiv e-prints*. Available: https://ui.adsabs.harvard.edu/abs/2016arXiv161201105Z

[19] G. Lin, A. Milan, C. Shen, and I. Reid. (2016, November 01, 2016). RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation. *arXiv e-prints*. Available: https://ui.adsabs.harvard.edu/abs/2016arXiv161106612L

[20] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. (2016, June 01, 2016). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *arXiv e-prints*. Available: https://ui.adsabs.harvard.edu/abs/2016arXiv160600915C

[21] Z. Meng, Z. Fan, Z. Zhao, and F. Su, "ENS-Unet: End-to-End Noise Suppression U-Net for Brain Tumor Segmentation," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 5886-5889.

[22] J. Chang, X. Zhang, J. Chang, M. Ye, D. Huang, P. Wang, *et al.*, "Brain Tumor Segmentation Based on 3D Unet with Multi-Class Focal Loss," in *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2018, pp. 1-5.

[23] X. Li, H. Chen, X. Qi, Q. Dou, C. Fu, and P. Heng, "H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation From CT Volumes," *IEEE Transactions on Medical Imaging*, vol. 37, pp. 2663-2674, 2018.

[24] J. Long, E. Shelhamer, and T. Darrell. (2014, November 01, 2014). Fully Convolutional Networks for Semantic Segmentation. *arXiv e-prints*. Available: https://ui.adsabs.harvard.edu/abs/2014arXiv1411.4038L

[25] J. Tang, J. Li, and X. Xu, "Segnet-based gland segmentation from colon cancer histology images," in *2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2018, pp. 1078-1082.

[26] T. Tran, O. Kwon, K. Kwon, S. Lee, and K. Kang, "Blood Cell Images Segmentation using Deep Learning Semantic Segmentation," in *2018 IEEE International Conference*

on Electronics and Communication Engineering (ICECE), 2018, pp. 13-16.

[27] K. Simonyan and A. Zisserman. (2014, September 01, 2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv e-prints*. Available: https://ui.adsabs.harvard.edu/abs/2014arXiv1409.1556S

[28] H. Salehinejad, J. Barfett, S. Valaee, and T. Dowdell, *Training Neural Networks with Very Little Data -- A Draft*, 2017.

[29] M. M. D. Deza, Elena, "Encyclopedia of Distances," 2009.

[30] P. Anbeek, K. L. Vincken, M. J. van Osch, R. H. Bisschops, and J. van der Grond, "Probabilistic segmentation of white matter lesions in MR imaging," *Neuroimage*, vol. 21, pp. 1037-44, Mar 2004.

[31] Y. Dgani, H. Greenspan, and J. Goldberger, "Training a neural network based on unreliable human annotation of medical images," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 39-42.