EFFECT OF TUMOR DELINEATION STRATEGIES ON ANN CLASSIFICATION ACCURACY IN LUNG CAD

by

Adem Cihan Arslan

B.S in Mathematics, Boğaziçi University, 2012

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2015

ACKNOWLEDGMENTS

I would like to thank my thesis supervisor Prof. Dr. Albert Güveniş for his support and motivation during my thesis study. Without his valuable academic support, experience, and knowledge, this study would not be completed. I would like to thank Dr. Dilek Şahin who segments X-ray images, for her support in ÇAPA Oncology Institute.

I would like to thank Boğaziçi University Mathematics Department that gives me opportunity to study. Many thanks to my friends in both B.U. Biomedical Engineering Institute and B.U. Mathematics. Many thanks to Merve Arız who is my fionce. Working and sharing life with you is unforgettable.

And to those who directly and indirectly help me, thank you.

ABSTRACT

EFFECT OF TUMOR DELINEATION STRATEGIES ON ANN CLASSIFICATION ACCURACY IN LUNG CAD

Lung Cancer is a serious illness and patient survival rate depends on early and accurate detection. CAD systems are commonly used for detection and characterization of nodules. The type of tumor segmentation algorithm or radiologist segmentation may affect the accuracy when characterizing lung nodules on chest x-ray images. In order to segment and classify nodules better, preprocessing step is needed. Histogram equalization, fuzzy minimization, bone subtraction, cropping can be some steps of preprocessing.

In this study, the main object is to evaluate the accuracy of the characterization of lung nodules on bone subtracted chest x-ray images by using different types of boundary segmentation algorithms and an artificial neural network based classification method. Another aim is to evaluate the contribution of CAD systems and accuracy of radiologist segmentation on raw chest X-ray. The standard digital image database with chest lung nodules (JSRT database) that was created by the Japanese Society of Radiological Technology in cooperation with the Japanese Radiological Society (JRS) is used. To subtract the bones a bone shadow elimination algorithm is used. Preprocessing and look up tables are used if nodule is not clearly seen. Active contour, spline active contour and radiologist based delineation methods are used. Artificial neural network classifications are used and their accuracy is evaluated. At the end, high specificity and sensitivity ratios are obtained and different segmentation techniques are compared. As a result, results are satisfying and interesting. Future work is possible to extend the study to other segmentation techniques and modalities.

Keywords: CAD, ANN, MATLAB, Lung Cancer, X-Ray.

ÖZET

AKCİĞER KANSERİ BİLGİSAYAR DESTEKLİ TEŞHİSİNDE TÜMÖR ÇİZİM STRATEJİLERİNİN YAPAY NÖRAL AĞ SINIFLANDIRMASI ÜZERİNDEKİ ETKİSİ

Akciğer kanseri ciddi bir hastalıktır ve hastaların hayatta kalma oranı erken ve doğru saptanmasına bağlıdır. Bilgisayar teşhis metotları tümörün saptanmasında ve sınıflandırılmasında yaygın olarak kullanılmaktadır. Tümör çizim algoritmalarının çeşidi veya radyoloğun çizimi göğüs x-ray filmlerinde bulunan akciğer tümörlerinin sınıflandırılmasına etki edebilir. Tümörleri daha iyi çizmek ve sınıflandırmak için ön işleme gereklidir. Histogram eşitlemesi, bulanık minimizasyonu, kemiklerin çıkarılması ve tümör etrafının kesilerek filmden çıkarılması ön sürecin bazı işlemleri arasında sayılabilir.

Bu çalışmada temel amaç, kemikleri çıkartılmış göğüs filmleri üzerinde akciğer tümörlerinin doğru bir şekilde sınıflandırılmasını sağlamak, değişik tümör çizim algoritmaları ve yapay nöral ağ sınıflandırılmaları oluşturmak ve bu ağların başarımını ölçmektir. Aynı zamanda da bilgisayar teşhis metodunun katkısını ve işlenmemiş filmler üzerinde yapılan radyolog çizimlerine göre oluşan teşhisinin doğruluğunu hesaplamaktır. Çalışmada JSRT(Japanese Society of Radiological Technology) tarafından oluşturulmuş tümör içeren standart dijital film verisi ve bu verinin kemikleri çıkartılmış halleri de çalışmada kullanılmıştır. Ön süreç ve pencere ayarları tümör tam olarak görülmediğinde kullanılmıştır. Yapay nöral ağlar sınıflandırma ve değerlendirmede kullanılmıştır. Sonunda, yüksek duyarlılık ve özgünlük oranları elde edilmiş ve değişik çizim tekniklerinin istatistiksel olarak farklı olabileceği bulunmuştur. Özet olarak, sonuçlar tatmin edici ve ilgi çekicidir. Çalışma gelecekte genişletilebilir ve başka çizim algoritmaları kullanılabilir.

Anahtar Sözcükler: Bilgisayarlı Teşhis, Yapay Nöral Ağ, MATLAB, Akciğer Kanseri, X-Ray.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii
ABSTRACT iv
ÖZET v
LIST OF FIGURES viii
LIST OF TABLES ix
LIST OF ABBREVIATIONS x
1. INTRODUCTION 1
  1.1 Background and Motivation 1
  1.2 Computer Aided Diagnosis 1
  1.3 Bone Subtraction 2
  1.4 Artifical Neural Network Systems 3
  1.5 Objectives 4
  1.6 Outline of The Thesis 4
2. PREVIOUS WORK 5
3. METHOD 7
  3.1 Data Set 7
  3.2 Bone Subtraction 9
  3.3 Pre-processing 9
  3.4 Types of Segmentation 11
    3.4.1 Active Contour Segmentation 11
    3.4.2 Spline Active Contour Segmentation 11
    3.4.3 Manuel Segmentation 12
  3.5 Feature Extraction 13
  3.6 Classification 13
  3.7 Evaluation 14
4. RESULTS AND DISCUSSION 15
  4.1 Artificial Neural Network Results 15
    4.1.1 AC Type 1 15
    4.1.2 SAC Type 1 15

4.1.3 Radiologist Segmentation 18

4.1.4 AC Type 2 18

4.1.5 SAC Type 2 18

4.2 Similarity Index Results 22

4.2.1 Area of Measure 22

4.2.2 Dice Similarity Index 22

4.2.3 Haussdorff Distance 23

4.2.4 Discussion 24

5. CONCLUSION AND FUTURE WORK 25

REFERENCES 26

LIST OF FIGURES

Figure 1.1	A basic multilayer neural network [1]	4
Figure 3.1	Some examples of raw JSRT data.	8
Figure 3.2	Some examples of bone suppressed JSRT data.	10
Figure 3.3	Bone suppressed (left), histogram equalized (middle), fuzzy minimization (right)	11
Figure 3.4	Some examples of nodule segmentations by AC	12
Figure 3.5	Some examples of nodule segmentations by SAC	12
Figure 3.6	Some examples of nodule segmentations by R	13
Figure 4.1	testing table of AC Type 1 (1: malignant 2: benign)	16
Figure 4.2	testing table of SAC Type 1 (1: malignant 2: benign)	17
Figure 4.3	confusion table of RS classification (1: malignant 2: benign)	19
Figure 4.4	confusion table of AC Type 2 classification (1: malignant 2: benign)	20
Figure 4.5	confusion table of SAC Type 2 classification(1:malignant 2:benign)	21

LIST OF TABLES

Table 4.1	AOM Means and Standart Deviations	22
Table 4.2	DSI Means and Standart Deviations	23
Table 4.3	Haussdorff Means and Standart Deviations	23

LIST OF ABBREVIATIONS

CAD	Computer Aided Diagnosis
ANN	Artificial Neural Network
ROI	Region of Interest
AC	Active Contour
SAC	Spline Active Contour
R	Radiologist
RS	Radiologist Segmentation
AOM	Area of Measure
DSI	Dice Similarity Index
HD	Haussdorff Distance
3D	Three Dimension
CT	Computerized Tomography
MATLAB	Matrix Laboratory

1. INTRODUCTION

1.1 Background and Motivation

In today's world, cancer has become one of the most common death reason in the whole world. Currently, cancer is the second leading reason of death in the United States. Moreover, cancer can surpass heart diseases as the leading cause of death in the next few years. In the United States, estimated deaths for all cancers are 589,430 in 2015. With respect to statistics, lung cancer is the most occurred and most deadly cancer type with the estimated number of death 158,040 in 2015. Cancer death rates decline for the past 2 decades. Cancer death rate decreased 22% from 1991(when most deaths occurred) up to now [2]. However, cancer, especially lung cancer, is still a serious illness and it must be investigated. Early detection and treatment of lung cancer are crucial for improvement of the patient survival rate. The patient survival rate can be improved if the carcinoma is found and removed at this early stage. Thus, early detection is the key factor to cure the lung cancer. Chest radiographs are the most widely used imaging technique for detection and classification of lung cancers. However, studies show that radiologists could fail to detect pulmonary nodules in chest x-ray images in up to 30% actually abnormal cases [3].

1.2 Computer Aided Diagnosis

Computer Aided Diagnosis (CAD) is one of the major research subjects in medical imaging. CAD is a decision support system that helps radiologists to make their final decisions. CAD is like a second opinion which takes into account equally the roles of radiologist and computers. With CAD, the performance of computers are complementary to radiologists, not comparable or better. Automated computer diagnosis is based on computer algorithms only [4]. To assist radiologists in the diagnosis of lung diseases, thus, computer-aided diagnosis (CAD) schemes can be developed for the detection of pulmonary nodules in digital chest X-Rays. The output from the CAD scheme would be used to alert radiologists to potential nodule locations, and the final diagnostic decision would then be made by the radiologists. A major obstacle in obtaining high performance for a computerized detection scheme is related to the problem of how to reduce the false-positive detection rate. Biopsy costs are a main proportion of total cost of diagnosing lung cancer today. Health sector needs to develop more precise risk stratification tools to better identify patients who require referrals for lung biopsy. This has the potential to reduce costs and improve patient outcomes. Biopsy rates may decline if the classification of the CAD diagnosis is better [5]. CAD scheme contains many sub works to improve the schemes. The main objective of this thesis was to implement the role of segmentation in the CAD schemes. To obtain the role of segmentation, active contour and spline active contour methods were used. The detection of the lung nodules can be found out by introducing Computer Aided Diagnosis (CAD) scheme. But segmentation is a real problem to understand the nature of the tumour. The aim is to develop a CAD scheme with improved sensitivity and specificity to classify the lung nodules as benign or malignant by two methods namely active contour segmentation, spline active contour segmentation and also we have manual radiologist segmentation.

1.3 Bone Subtraction

In an X-Ray image, bones and clavicles are obstacles to detect or to segment tumor areas. Eliminating bone shadows from chest radiographs can greatly improve the accuracy of lesion detection and segmentation. To free images from rib and clavicle shadows, they are first segmented using a dynamic programming approach. The segmented shadows are eliminated in difference space. The cleaned images are processed by a hybrid lesion detector based on gradient convergence, contrast and intensity statistics. False findings are eliminated by a Support Vector Machine. Bone subtraction method can eliminate approximately 80% of bone shadows (84% of the posterior part) with an average segmentation error of 1 mm. With shadow removal the number of false findings dropped from 2.94 to 1.23 at 63% of sensitivity for cancerous tumors.

The output of the improved system showed much less dependence on bone shadows. Bone shadow elimination can lead to great benefits for computer aided detection CAD systems [6].

1.4 Artifical Neural Network Systems

Artificial neural network systems is commonly used in many areas. It is a system that has common properties compared to neurons in brain. ANN systems have many neurons to process. ANN systems have neurons and each neuron has different weight and also they are all connected by communication links. Weights are related to solution of the network problem. There is no restriction about hidden layers. Moreover, there is no rule to determine the neurons in the hidden layers. ANNs has an iterative learning process and weights related to the input values are adjusted each time. ANN systems have a great generalization capacity by using the given training data. This is a great advantage for ANN systems to use various areas include medical diagnosis. ANN systems have two different steps namely training and testing. In our thesis, training and testing are done simultaneously by using the MATLAB pattern recognition toolbox. Feedforward backpropagation network type is used. Details can be checked by the reference [1].

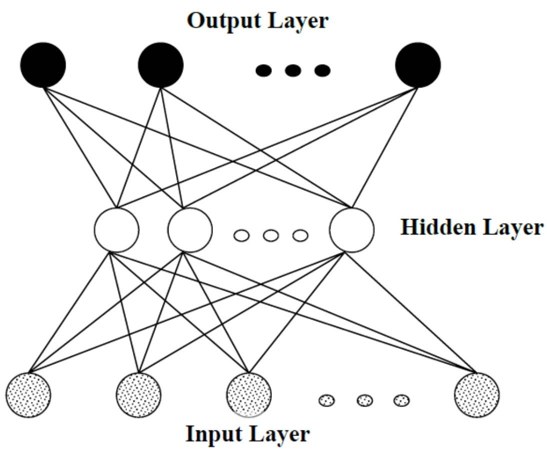

Figure 1.1 A basic multilayer neural network [1]

1.5 Objectives

The main objective of this thesis was to study the role of segmentation in classification accuracy. To obtain the role of segmentation, active contour, spline active contour methods and radiologist segmentation method was used. The detection of the lung nodules can be obtained out by introducing CAD scheme. But segmentation is a real problem to understand the nature of the tumor. The aim is to develop a CAD scheme with improved sensitivity and specificity to detect the lung nodules by two methods namely active contour and spline active contour. Then, this thesis focuses on the similarity indexes between these segmentation algorithms and the radiologist segmentation.

1.6 Outline of The Thesis

Chapter 1 introduces the subject and gives outline of this thesis. In Chapter 2, previous work are given. In Chapter 3, artificial neural network is discussed. Chapter 4 contains the whole methodology. In Chapter 5, segmentation and classification results are given. Future work and conclusion can be found in the Chapter 6.

2. PREVIOUS WORK

CAD is one of the major research subjects in medical imaging and diagnostic radiology. Computer-aided diagnosis has a part of clinical work of CAD that is a concept based on the equal roles of physician and computer, and thus is distinctly different from automated computer diagnosis [7]. Previous work on CAD can be summed recently. The feasibility and efficiency of CAD were proved by Kobayashi et al. and MacMahon et al. They proved that a CAD scheme could help radiologists in the detection of pulmonary nodules. De Boo et al. [8] provided a detailed review of CAD schemes. They showed that multiple studies could improve the detection performance of radiologists especially less experienced ones. Giger et al.[9] built a difference-image technique to reduce complexity of anatomic background structures. This helps to detect initial nodule candidates. Lo et al. [10] used a similar technique to the difference-image to create nodule-enhanced images. Then, images are processed by a feature extraction technique based on edge detection, gray-level thresholding and sphere profile matching. They also studied on a difference-image process, followed by feature extraction and classification processes. Yoshida et al [11] reported the wavelet transform to detect the subtle nodules missed in the difference-image technique to initial nodule candidate detection. Vittitoe et al. [12] employed fractal texture characterization to raise the detection accuracy for solitary pulmonary nodules. Recently, Carreira et al. [13] studied on detection of nodule candidates with normalized cross-correlation images and classification of the candidates in curvature space. Then, Penedo et al. [14] improved the performance by appliyng two-level Artificial Neural Network(ANN).Coppini et al. [15] provided a biologically inspired ANNs with fuzzy coding to develop a CAD scheme. Shiraishi et al. [16] incorporated a localized searching method based on anatomical classification and automated techniques. Schilham et al. [17] create a new initial nodule candidate detection method by using multiscale techniques. The performance of a CAD scheme improved by Campadelli et al. [18]. They introduce a new lung segmentation method. Hardie et al. [19] incorporates a CAD scheme based on a weighted-multiscale convergence index filter for initial nodule candidate detection. The scheme also based on an adaptive distance-based threshold algorithm for candidate segmentation. Nodule detection in chest x-rays is a great deal of work have to be done by researchers. However, false positive rates are still high. False positives could confuse radiologists to decide normal areas as suspicious. Moreover, radiologists can lose their confidence CAD tools. Matsumoto et al. [20] provided an observer performance study to investigate the effect of false positives. They showed that if a CAD scheme had a high false-positive rate, accuracy of radiologists in detecting pulmonary nodules was not improved with CAD, although the scheme had a sensitivity of 80%. Scientists have developed many false positive reduction methods. Yoshida and Keserci [21] applied an edge-guided wavelet snake model to the extraction a weighted overlap between the snake and the multiscale edges. Yoshida developed a CAD scheme based on local contralateral subtraction[22]. They aims to remove normal anatomic structures in chest x-rays based on the symmetry between the left and right lung regions for false positive reduction. Suzuki et al. proposed a multiple massive-training ANN to decrease the number of false positives in a CAD scheme [23]. El-Baz, Ayman, et al. overviews the current state-of-the-art techniques that have been developed to implement each of these CAD processing steps [4]. Htike, Zaw Zaw, et al. proposes a three-layered framework to perform automatic diagnosis of pulmonary nodules [24]. Meziane, Moulay, et al. demonstrates good detection of moderately subtle lesions with a relatively low false positive rate [25].

3. METHOD

3.1 Data Set

The standard digital image database with and without chest lung nodules (JSRT database) was created by the Japanese Society of Radiological Technology (JSRT) in cooperation with the Japanese Radiological Society (JRS) in 1998. Since then, the JSRT database has been used many researchers for various research aims such as image processing, computer-aided diagnosis (CAD), picture archiving and communication system (PACS), and for training and testing.Citations for this database was 35 in 2006. After 10 years they decided to release the JSRT database with free of charge in order to facilitate potential users around the world.

Descriptions of the JSRT database:

- Useful for ROC analysis (154 nodule and 93 non-nodule images)
- High resolution (2048 x 2048 matrix size, 0.175mm pixel size)
- Wide density range (12bit, 4096 gray scale)
- Universal image format (no header, big-endian raw data)
- Useful for diagnostic training and testing
- Patient age, gender diagnosis (malignant or benign), X and Y coordinates of nodule, simple diagram of nodule location, degree of subtlety in visual detection of nodules.

However, clavicles and ribs are huge obstacle to segmentation. Eliminating bone shadows from chest Radiographs can greatly improve the accuracy of automated lesion detection. Bone shadow elimination can lead to great benefits for computer aided detection CAD systems. Thus, bone suppressed images is used for semi-automated systems. However, radiologists used raw version of the data [26].

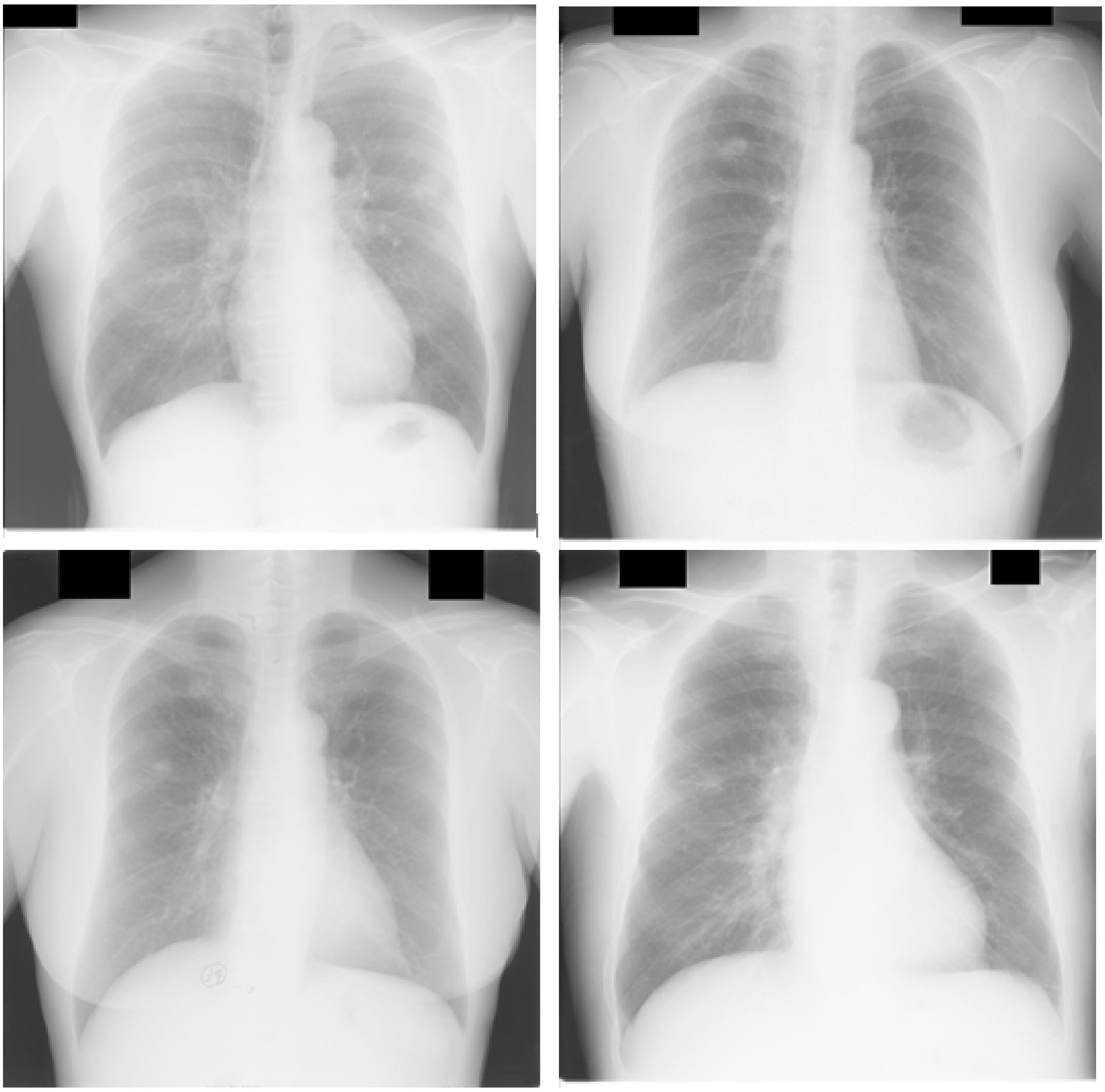

Figure 3.1 Some examples of raw JSRT data.

3.2 Bone Subtraction

Bone subtraction or eliminating ribs and clavicles gives good results in another study that is discussed in the Chapter 1.3. They used Active Shape Model to segment and eliminate bones and to obtain soft tissue only. More information can be found at the related reference. In this thesis, bone subtracted images is used for semi-automated segmentations. Tumour areas can be seen well in these images. Bone subtraction is the initial process of our CAD system.

3.3 Pre-processing

The mass segmentation method employed in this study starts with the initial detection of a mass shape by the user. CAD systems have much to study and this thesis focus on segmentations. Thus, all values that will be given in the next chapters have the assumption of correct nodule findings. In order to see nodules clearly and for the initiation, automated and semi-automated look up tables are used. Changing window-level values provides more contrast. Also, histogram equalization and fuzzy minimization is used in computer software if necessary. In some images cropping the nodule is a good way to use histogram equalization and fuzzy minimization. This can give better images to segment nodules. After that, initial polygonal ROI is given the computer. This ROI is only a simple version not a complete segmentation. If the nodule on the image is too complicated to give an initiation, radiologist segmentation is a guide to give some initial contour. The final segmentation is given by computer algorithms with using the software MIPAV.

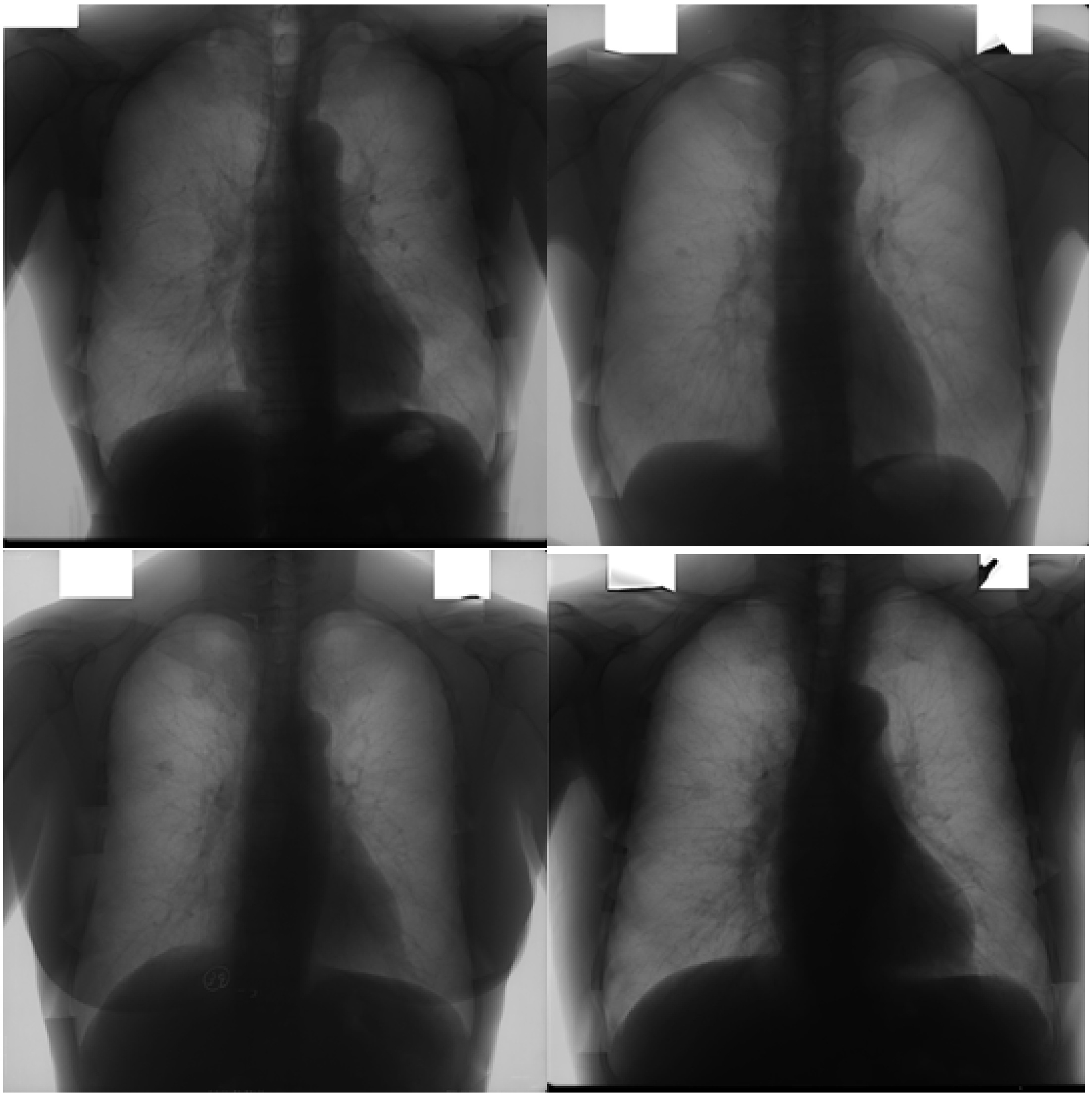

Figure 3.2 Some examples of bone suppressed JSRT data.

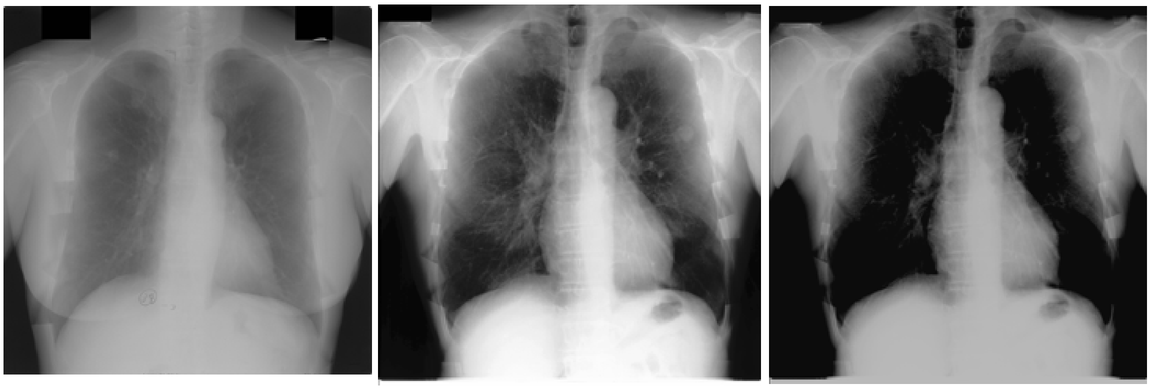

Figure 3.3 Bone suppressed (left), histogram equalized (middle), fuzzy minimization (right)

3.4 Types of Segmentation

3.4.1 Active Contour Segmentation

The initial polygonal mass segmentation process is described in pre-processing. Further refinement was necessary before classifications of the nodules. An AC is a deformable continuous curve, whose shape is controlled by internal forces and external forces. The internal forces impose a smoothness constraint on the contour, and the external forces push the contour toward salient image features, such as edges. To solve a segmentation problem, an initial boundary is iteratively deformed so that the energy due to internal and external forces is minimized along the contour. Further explanations is discussed in the reference [27]. AC gives the segmentation of the nodule. Rest of the non-nodule area is shaded to save for feature extraction. Examples can be seen on the figures.

3.4.2 Spline Active Contour Segmentation

The initial polygonal mass segmentation comes from preprocessing again. Further refinement procedure is applied in the same way to AC.Spline Active contour

Figure 3.4 Some examples of nodule segmentations by AC

Figure 3.5 Some examples of nodule segmentations by SAC

method (SAC) is based on A parametric active contour method based on highly reduce the computational cost. Further explanations is discussed here [28]. SAC gives also the segmentation of the nodule.

3.4.3 Manuel Segmentation

Manuel segmentation is done by a radiologist. Radiologist uses raw images without any processing. Radiologist segmentation (RS) results is also classified by ANN.

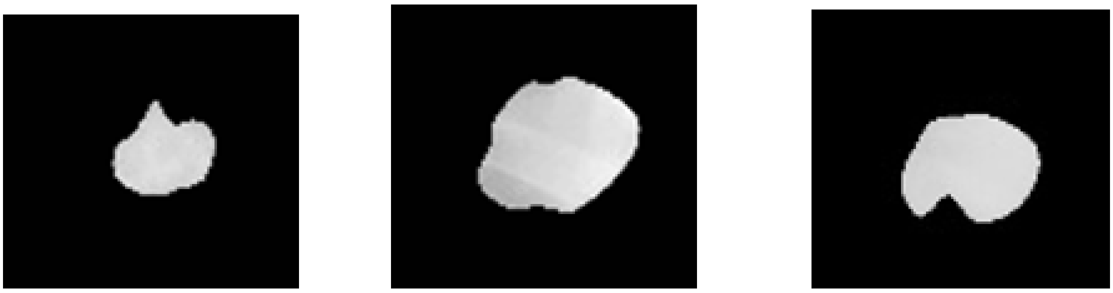

Figure 3.6 Some examples of nodule segmentations by R

3.5 Feature Extraction

Region properties can be ordered as Area, MajorAxisLength, MinorAxisLength, Eccentricity, Orientation, ConvexArea, FilledArea, EulerNumber, EquivDiameter, Solidity, Extent, and Perimeter. Results are saved in a structure for each image by using a loop. Also, patient age and gender are the last two feature of the nodule. Features are used from MATLAB. More information can be provided by the reference [29]. Finally, matrix of the feature table is obtained.

3.6 Classification

This step has done straightforwardly. It starts with nprstart toolbox in MATLAB and pattern recognition is applied. The table from feature extraction is given as input and nodule information from the JSRT data source is given as target. Neural network tool uses 70% of the data for training,15% for validation and 15% for testing. After training, the network is created and it is ready to show results. They will be given in the Results Chapter.

3.7 Evaluation

We can evaluate the CAD system basically in terms of sensitivity, accuracy, and specificity.

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.1}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.3}$$

TP: the number of true positives

FP: the number of false positives

FN: the number of false negatives

TN: the number of true negatives

For the comparison of the segmentations Area of Measure, Dice Similarity and Haussdorff distance is used.

4. RESULTS AND DISCUSSION

4.1 Artificial Neural Network Results

Artificial Neural Network results can be explained by confusion tables that contains sensitivity and specificity ratios.

4.1.1 AC Type 1

Non-specialist draws contours on the bone subtracted x-ray image. The image is cropped around the tumour (approximately twice the diameter of the tumour).The cropped image is subject to histogram equalization. Fuzzy minimization is used if necesssary. Active Contour Algorithm used to segment the tumour using the Non-specialist drawn contour as the initial region of interest.14 features are extracted. ANN is used for classification (14 input layer, 10 hidden layer, 70% of the data for training,15% for validation and 15% for testing).

4.1.2 SAC Type 1

Non-specialist draws contours on the bone subtracted x-ray image. The image is cropped around the tumour (approximately twice the diameter of the tumour).The cropped image is subject to histogram equalization. Fuzzy minimization is used if necessary. Spline Active Contour Algorithm used to segment the tumour using the Non-specialist drawn contour as the initial region of interest.14 features are extracted. ANN is used for classification (14 input layer, 10 hidden layer, 70% of the data for training,15% for validation and 15% for testing).

Figure 4.1 testing table of AC Type 1 (1: malignant 2: benign)

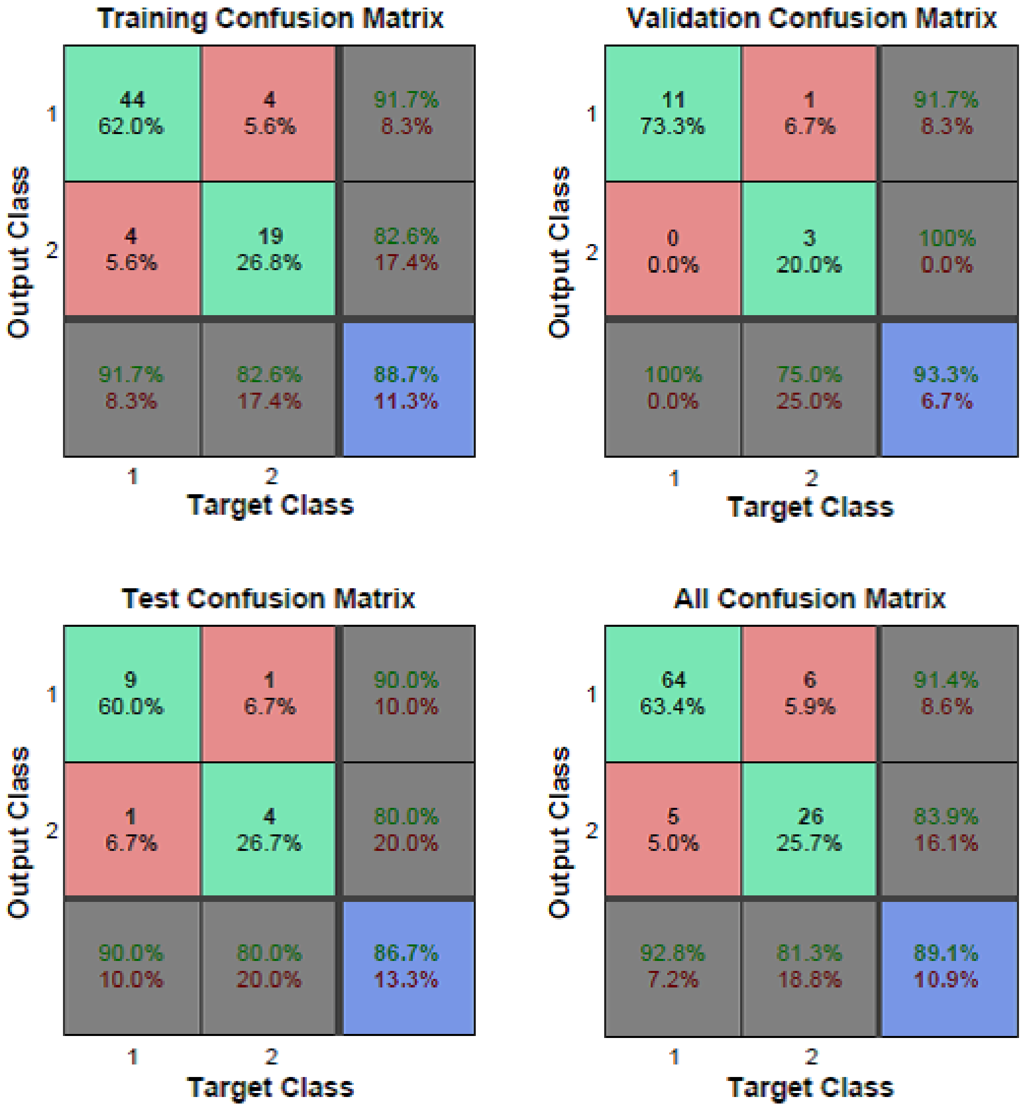

Figure 4.2 testing table of SAC Type 1 (1: malignant 2: benign)

4.1.3 Radiologist Segmentation

Radiologist draws contours on original chest x-ray image. Contour is pasted on the bone subtracted corresponding chest x-ray image.14 features are extracted. ANN is used for classification (14 input layer, 10 hidden layer, 70% of the data for training,15% for validation and 15% for testing).

4.1.4 AC Type 2

Radiologist draws contours on original chest x-ray image. Contour is pasted on the bone subtracted corresponding chest x-ray image. The image is cropped around the tumor (approximately twice the diameter of the tumour). The cropped image is subject to histogram equalization. Active Contour Algorithm used to segment the tumour using the radiologist drawn contour as the initial region of interest.14 features are extracted. ANN is used for classification (14 input layer, 10 hidden layer, 70% of the data for training,15% for validation and 15% for testing).

4.1.5 SAC Type 2

Radiologist draws contours on original chest x-ray image. Contour is pasted on the bone subtracted corresponding chest x-ray image. The image is cropped around the tumour (approximately twice the diameter of the tumour). The cropped image is subject to histogram equalization. Active Contour Algorithm used to segment the tumour using the radiologist drawn contour as the initial region of interest.14 features are extracted. ANN is used for classification (14 input layer, 10 hidden layer, 70% of the data for training,15% for validation and 15% for testing).

Figure 4.3 confusion table of RS classification (1: malignant 2: benign)

Figure 4.4 confusion table of AC Type 2 classification (1: malignant 2: benign)

Figure 4.5 confusion table of SAC Type 2 classification(1:malignant 2:benign)

4.2 Similarity Index Results

4.2.1 Area of Measure

Area of Measure (AOM) is the basic way of getting similarity ratios. It can be calculated as follows:

$$AOM = \frac{A \cap B}{A \cup B} \tag{4.1}$$

where A, B are the different binary segmentations.
Mean and standart deviation for all comparisons are in the table.

Table 4.1
AOM Means and Standart Deviations

	Mean	**Standart Deviation**
AC-SAC	76,94%	11,94%
AC-R	68,13%	12,96%
SAC-R	76,63%	11,28%

4.2.2 Dice Similarity Index

Dice similarty index is common tool to calculate similarities.Formula is given as follows:

$$DSI = \frac{2 \mid A \cap B \mid}{\mid A \mid + \mid B \mid} \tag{4.2}$$

where A,B are the different binary segmentations.
Mean and standart deviation for all comparisons are in the table.

Table 4.2
DSI Means and Standart Deviations

	Mean	**Standart Deviation**
AC-SAC	86,07%	9,14%
AC-R	78,87%	10,49%
SAC-R	84,61%	8,55%

4.2.3 Haussdorff Distance

Hausdorff distance is a metric and it satisfies the identity and symmetry equalities and the triangle inequality and it is nonnegative.More information can be found on the reference [27].

Table 4.3
HaussdorffMeans and Standart Deviations

	Mean	**Standart Deviation**
AC-SAC	2,49%	0,61%
AC-R	2,71%	0,65%
SAC-R	2,36%	0,56%

4.2.4 Discussion

The mass boundaries obtained by the AC, SAC and R is compared with one another in terms of the HD, the AOM, and DSI for the 100 masses that were segmented by both radiologists. Tables shows the mean and standard deviation of these measures computed using the three possible paired comparisons of the segmentations (AC-R, AC-SAC, and SAC-R). Results are checked by a paired t-test. All tests, 95% confidence interval is used. In all similarity indexes gives the same result. Only AC-R and SAC-R are not statistically different. Thus, we can conclude that AC and R segmentations are different. With the results of training, we can conclude that AC segmentation gives us a better segmentation compared to others.

5. CONCLUSION AND FUTURE WORK

We conclude that two main types of results. Firstly, we get good specificity and sensitivity ratios in the classification of AC segmentation. In evaluation part, we get 83.3% sensitivity and 75% specificity and 81.3% accuracy. In the whole data of AC we get 86.5% accuracy. These results implies that AC segmentation can help radiologist to classify nodules as malignant or benign. In SAC classification, results are not too satisfying compared to AC. However, SAC is still good R segmentation classification. Thus, radiologist should use CAD system to segment nodules. In this way, biopsy rates can be avoided more. Moreover, early detection can be possible too. Bone subtracted images, when used with local cropping, histogram equalization and active contour segmentation result in a high tumour characterization accuracy. Clinical validation is necessary as future work in order to assess the use of this method in lung nodule management and prevention of unnecessary biopsies.

This study gives good results, nevertheless a better result can be possible by contributing more study. We used only X-ray Images, but CT is a better way to detect nodules since CT provides 3D images. In this study, 100 data is used for training. More data can provide better results. To get more data, hospitals and universities should work cooperatively. Another thing for future work can be using more features to train ANN. Even anamnesis can be a part of neural network features.

Secondly, segmentation algorithms are compared by using three methods. Then pairs are used to compare and we get that AC and R are statistically different. Similarity indexes are meaningful when we have a grand truth. However, comparing algorithms can helps us with the training results. If we have more segmentation algorithms we can use indexes to determine the best one. It can help merchandising of hospitals in the future. To sum up, CAD is a huge area to study and researchers must provide a lot, since cancer is a current health problem. This study area can save too many cancerous or probable cancerous people.

REFERENCES

1. Bomma, P., *COMUPTER-AIDED DIAGNOSIS TOOL FOR THE DETECTION OF CANCEROUS NODULES IN X-RAY IMAGES*. PhD thesis, Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering in The Department of Electrical and Computer Engineering by Pallavi Bomma BE, Andhra University Visakhapatnam, India, 2005.

2. Siegel, Rebecca L., K. D. M., and A. Jemal., "Cancer statistic," *CA: a cancer journal for clinicians 65.1 (2015)*, Vol. 5-29, 2015.

3. Agrawal, S. K., "International journal of advance research in computer science and management studies," *International Journal*, Vol. 1, no. 7, 2013.

4. El-Baz, A., G. M. Beache, G. Gimel'farb, K. Suzuki, K. Okada, A. Elnakib, A. Soliman, and B. Abdollahi, "Computer-aided diagnosis systems for lung cancer: challenges and methodologies," *International journal of biomedical imaging*, Vol. 2013, 2013.

5. Lokhandwala, T., R. Dann, M. Johnson, and A. D'Souza, "Costs of the diagnostic workup for lung cancer: A medicare claims analysis," *International Journal of Radiation Oncology, Biology, Physics*, Vol. 5, no. 90, pp. S9–S10, 2014.

6. Juhász, S., Á. Horváth, L. Nikházy, and G. Horváth, "Segmentation of anatomical structures on chest radiographs," in *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*, pp. 359–362, Springer, 2010.

7. Doi, K., "Current status and future potential of computer-aided diagnosis in medical imaging," *The British Journal of Radiology*, 2014.

8. De Boo, D. W., M. Prokop, M. Uffmann, B. van Ginneken, and C. M. Schaefer-Prokop, "Computer-aided detection (cad) of lung nodules and small tumours on chest radiographs," *European journal of radiology*, Vol. 72, no. 2, pp. 218–225, 2009.

9. Giger, M. L., N. Ahn, K. Doi, H. MacMahon, and C. E. Metz, "Computerized detection of pulmonary nodules in digital chest images: Use of morphological filters in reducing false-positive detections," *Medical Physics*, Vol. 17, no. 5, pp. 861–865, 1990.

10. Shih-Chung, B. L., M. T. Freedman, J.-S. Lin, and S. K. Mun, "Automatic lung nodule detection using profile matching and back-propagation neural network techniques," *Journal of Digital Imaging*, Vol. 6, no. 1, pp. 48–54, 1993.

11. Yoshida, H., X.-W. Xu, T. Kobayashi, M. L. Giger, and K. Doi, "Computer-aided diagnosis scheme for detecting pulmonary nodules using wavelet transform," in *Medical Imaging 1995*, pp. 621–626, International Society for Optics and Photonics, 1995.

12. Vittitoe, N. F., J. A. Baker, and C. E. Floyd, "Fractal texture analysis in computer-aided diagnosis of solitary pulmonary nodules," *Academic radiology*, Vol. 4, no. 2, pp. 96–101, 1997.

13. Carreira, M. J., D. Cabello, M. G. Penedo, and A. Mosquera, "Computer-aided diagnoses: Automatic detection of lung nodules," *Medical physics*, Vol. 25, no. 10, 1998.

14. Penedo, M. G., M. J. Carreira, A. Mosquera, and D. Cabello, "Computer-aided diagnosis: a neural-network-based approach to lung nodule detection," *Medical Imaging, IEEE Transactions on*, Vol. 17, no. 6, pp. 872–880, 1998.

15. Coppini, G., S. Diciotti, M. Falchini, N. Villari, and G. Valli, "Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms," *Information Technology in Biomedicine, IEEE Transactions on*, Vol. 7, no. 4, pp. 344–357, 2003.

16. Shiraishi, J., Q. Li, K. Suzuki, R. Engelmann, and K. Doi, "Computer-aided diagnostic scheme for the detection of lung nodules on chest radiographs: localized search method based on anatomical classification," *Medical Physics*, Vol. 33, no. 7, pp. 2642–2653, 2006.

17. Schilham, A. M., B. Van Ginneken, and M. Loog, "A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database," *Medical Image Analysis*, Vol. 10, no. 2, pp. 247–258, 2006.

18. Campadelli, P., E. Casiraghi, and D. Artioli, "A fully automated method for lung nodule detection from postero-anterior chest radiographs," *Medical Imaging, IEEE Transactions on*, Vol. 25, no. 12, pp. 1588–1603, 2006.

19. Hardie, R. C., S. K. Rogers, T. Wilson, and A. Rogers, "Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs," *Medical Image Analysis*, Vol. 12, no. 3, pp. 240–258, 2008.

20. MATSUMOTO, T., H. YOSHIMURA, M. L. GIGER, K. DOI, H. MacMAHON, S. M. MONTNER, and T. NAKANISHI, "Potential usefulness of computerized nodule detection in screening programs for lung cancer.," *Investigative radiology*, Vol. 27, no. 6, pp. 471–475, 1992.

21. Keserci, B., and H. Yoshida, "Computerized detection of pulmonary nodules in chest radiographs based on morphological features and wavelet snake model," *Medical Image Analysis*, Vol. 6, no. 4, pp. 431–447, 2002.

22. Yoshida, H., "Multiscale edge-guided wavelet snake model for delineation of pulmonary nodules in chest radiographs," *Journal of Electronic Imaging*, Vol. 12, no. 1, pp. 69–80, 2003.

23. Chen, S., K. Suzuki, and H. MacMahon, "Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification," *Medical physics*, Vol. 38, no. 4, pp. 1844–1858, 2011.

24. Htike, Z. Z., W. Y. N. Naing, S. L. Win, and S. Khan, "Computer-aided diagnosis of pulmonary nodules from chest x-rays using rotation forest," in *Computer and Communication Engineering (ICCCE), 2014 International Conference on*, pp. 96–99, IEEE, 2014.

25. Meziane, M., P. Mazzone, E. Novak, M. L. Lieber, O. Lababede, M. Phillips, and N. A. Obuchowski, "A comparison of four versions of a computer-aided detection system for pulmonary nodules on chest radiographs," *Journal of thoracic imaging*, Vol. 27, no. 1, pp. 58–64, 2012.

26. Shiraishi, J., S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-i. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," *American Journal of Roentgenology*, Vol. 174, no. 1, pp. 71–74, 2000.

27. Sahiner, B., N. Petrick, H.-P. Chan, L. M. Hadjiiski, C. Paramagul, M. A. Helvie, and M. N. Gurcan, "Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization," *Medical Imaging, IEEE Transactions on*, Vol. 20, no. 12, pp. 1275–1284, 2001.

28. Precioso, F., M. Barlaud, T. Blu, and M. Unser, "Smoothing b-spline active contour for fast and robust image and video segmentation," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, Vol. 1, pp. I–137, IEEE, 2003.

29. *www.mathworks.comhelpvisionfeature-detection-extraction-and-matching.html.* June 2015.