



Automatic Detection of Lung Cancer

Team Members

Ahmed Hassaan

Ahmed Sallam

Abdelrahman Hagag

Amina Alaa

Taghreed Fayez

Assis. Prof. Ahmed Elnakib

Table of Contents

Chapter 1: Introduction	1
1.1 Overview of Lung cancer	2
1.1.1 Lung	2
1.1.2 Lung Cancer	2
1.1.3 There are three main types of lung cancer	
1.1.4 Lung cancer causes	
1.1.5 Symptoms and signs	5
1.1.6 Lung Nodules	5
1.2 Overview of current technologies for lung cancer detection	7
1.2.1 Biopsy	7
1.2.2 Chest X-ray	7
1.2.3 Computed Tomography (CT)	
1.2.4 Positron emission tomography (PET) imaging/CT	
Chapter 2: System Design	
2.1 Design Objectives	9
2.2 CAD System	9
2.3 Visualize	
2.4 Framework	
Chapter 3: Data Collection	
Chapter 4: Lung Segmentation	13
Chapter 4. Lung Segmentation	13
4.1 Lung Segmentation	13
4.2 Methods of Segmentation	13
4.2.1 Problems of threshold-based segmentation	14
4.3 Method we used	14
4.3.1 Iterative Threshold	14
4.3.2 Algorithms	15
2D	15
3D	15
4.4 Performance Evaluation	
4.4.1 Sørensen–Dice coefficient	
4.4.2 Formula	16
4.4.3 Applications	17

4.5 Segmentation Results	
4.6 Discussion	
Chapter 5: Nodule Detection	19
5.1 Detection framework	19
5.2 Candidate nodule selection	19
5.3 Feature extraction	
5.4 Nodule detection and classification	
Chapter 6: Visualization and GUI	
Chapter 7: Conclusion and Future Work	

1 Introduction

Lung cancer is the second most common cancer in both men and women. About 14% of all new cancers are lung cancers. According to the American Cancer Society's, the estimates for lung cancer in the United States for 2016 are (see Fig. 1):

- About 224,390 new cases of lung cancer (117,920 in men and 106,470 in women)
- About 158,080 deaths from lung cancer (85,920 in men and 72,160 in women)

Lung cancer mainly occurs in older people. About 2 out of 3 people diagnosed with lung cancer are 65 or older, while less than 2% are younger than 45. The average age at the time of diagnosis is about 70.

Early detection of lung cancer can enhance the survival rate. Currently, the gold standard for lung cancer detection is using the Biopsy, which is an invasive procedure. Research studies aims to stage the lung cancer using noninvasive images of the lung. In this project we aim to develop an automated tools

This chapter is organized as follows, first we will give an overview of the lung and the cancerous cells inside the lung. Then, we will present the current technologies for lung cancer detection.

1.1 Overview of Lung cancer

1.1.1 Lung





The chest contains two lungs, one lung on the right side of the chest, the other on the left side of the chest. Each lung is made up of sections called lobes. The lung is soft and protected by the ribcage. As Fig. [1]

The purposes of the lungs are to bring oxygen (O2), into the body and to remove carbon dioxide (CO2). Oxygen is a gas that provides us energy while carbon dioxide is a waste product or "exhaust" of the body.

1.1.2 Lung Cancer

Lung cancer is the uncontrolled growth of abnormal cells in one or both lungs. These abnormal cells do not carry out the functions of normal lung cells and do not develop into healthy lung tissue. As they grow, the abnormal cells can form tumors and interfere with the functioning of the lung.

At a later stage of disease, some cells may travel away from the original tumor and start growing in other parts of the body. This process is call metastasis and the new distant sites are referred to as metastases.

Metastatic spread

Bronchial carcinoma commonly spreads to mediastina, cervical and even axillary or intra-abdominal nodes. In addition, the liver, adrenal glands, bones, brain and skin are frequent sites for metastases:

Liver	Common symptoms are anorexia, nausea and weight loss. Right upper quadrant pain radiating across the abdomen is associated with				
	liver capsular pain.				
Bone	Bony pain and pathological fractures as a result of tumor spread occur. If the spine is involved, there is a risk of spinal cord compression.				
Adrenal glands	Metastases to the adrenals do not usually result in adrenal insufficiency and are usually asymptomatic.				
Brain	Metastases present as space-occupying lesions with subsequent mass effect and signs of raised intracranial pressure. Less common presentations include carcinomatous meningitis with cranial nerve defects, headache and confusion.				

1.1.3 There are three main types of lung cancer:

- Non-Small Cell Lung Cancer (NSCLC)
- Small Cell Lung Cancer (SCLC)
- Lung Carcinoid Tumor

Table 1.1 Lung cancer cell types and clinical features

Cell type	Incidence in UK (%)	Features
Squamous cell carcinoma	35	Remains the most common cell type in Europe Arises from epithelial cells, associated with production of keratin Occasionally cavitates with central necrosis Causes obstructing lesions of bronchus with post-obstructive infection Local spread common, metastasizes relatively late
Adenocarcinoma	27–30	 Likely to become the most common cell type in the UK in the near future (most common cell type in the USA) Increasing incidence over last 10 years possibly linked to low tar cigarettes Originate from mucus-secreting glandular cells Most common cell type in non-smokers Often causes peripheral lesions on chest X-ray/CT Subtypes include bronchoalveolar cell carcinoma (associated with copious mucus secretion, multifocal disease) Metastases common: pleura, lymph nodes, brain, bones, adrenal glands
Large cell carcinoma	10–15	Often poorly-differentiated Metastasize relatively early
Small cell carcinoma	20	Arise from neuroendocrine cells (APUD cells) Often secrete polypeptide hormones Often arise centrally and metastasize early

1.1.4 Lung cancer causes

Tobacco	Tobacco use is the major preventable cause. Just as tobacco use and
	cancer rates are falling in some developed countries, both smoking and
	lung cancer are rising in Eastern Europe and in many developing
	countries. The great majority of tumors in the lung are primary
	bronchial carcinomas and, in contrast to many other tumors, the
	prognosis remains poor, with fewer than 30% of patients surviving at 1
	year and 6–8% at 5 years.
	The vast majority (85%) of cases of lung cancer are due to long-term
	tobacco smoking. About 10–15% of cases occur in people who have
	never smoked. These cases are often caused by a combination of
	genetic factors and exposure to radon gas, asbestos, second-hand
	smoke, or other forms of air pollution.
Exposure to	Radon is one of the biggest causes of lung cancer after smoking.
radon gas	Radon gas is a naturally occurring radioactive gas which comes from
<u> </u>	tiny amounts of uranium present in all rocks and soils.
Exposure to	A number of substances that occur in the workplace may cause lung
certain	cancer. In particular, these include asbestos, silica, and diesel exhaust.
chemicals	And smokers are at even higher risk.
Air pollution	We know that air pollution can cause lung cancer. The risk depends on
	the levels of air pollution you are regularly exposed to. At UK levels,
	the extra risk is likely to be small - and much smaller than being a
	smoker.
Previous lung	Having had a disease that caused scarring in the lungs may be a risk
disease	factor for a type of lung cancer called adenocarcinoma of the lung.
	Tuberculosis (TB) can make scar tissue form in the lungs. People who
	have had TB have double the risk of lung cancer. This risk continues
	for more than 20 years.
A family	Researchers are looking into the impact of family history on lung
history of lung	cancer. They have found that a family history does increase your risk.
cancer	Between 2 and 6 people out of 100 who do not have a first degree
	relative with lung cancer are diagnosed with lung cancer.
Lowered	HIV and AIDS lower immunity and so do drugs that people take after
immunity	organ transplants. An overview of research studies shows that people
	with HIV or AIDS have a risk of lung cancer that is 3 times higher
	than people who do not have HIV or AIDS. People who take drugs to
	suppress their immunity after an organ transplant have double the
	usual risk of lung cancer.

1.1.5 Symptoms and signs

The most common symptoms of lung cancer are:

- A cough that does not go away or gets worse.
- Coughing up blood or rust-colored sputum (spit or phlegm).
- Chest pain that is often worse with deep breathing, coughing, or laughing.
- Hoarseness.
- Weight loss and loss of appetite.
- Shortness of breath.
- Feeling tired or weak.
- Infections such as bronchitis and pneumonia that don't go away or keep coming back.
- New onset of wheezing.

If lung cancer spreads to distant organs (Metastatic spread), it may cause:

- Bone pain (like pain in the back or hips)
- Nervous system changes (such as headache, weakness or numbness of an arm or leg, dizziness, balance problems, or seizures), from cancer spread to the brain or spinal cord
- Yellowing of the skin and eyes (jaundice), from cancer spread to the liver
- Lumps near the surface of the body, due to cancer spreading to the skin or to lymph nodes (collections of immune system cells), such as those in the neck or above the collarbone.

1.1.6 Lung Nodules

Nodules are relatively low-contrast white circular objects within the lung fields. They are usually smaller than 3–4 cm in diameter (no larger than 6 cm) and are always surrounded by normal, functioning lung tissue. Their intensity in CT scans is from -300 to 0 HU.

A solitary pulmonary nodule:

- > Parenchymal, non-pleural nodule is a small, round or egg-shaped lesion in the lungs.
- > Juxtapleural pulmonary nodule is a small, worm-shaped lesion connected to pleura.

Difference between Benign and Malignant nodules:

Size	Benign lesions are smaller overall than malignant ones. As the		
	lesion increases in size, its probability of being malignant		
	increases.		
	Nodules more than 3 cm in diameter, 93% to 97% are malignant.		
Growth Rate	A very rapid growth rate or stability over a 2-year time period are		
	accepted indicators of benignity.		
Calcification	Benign nodules often contain dense, central, laminated, popcorn,		
	and punctate patterns of calcification.		
	Malignant pulmonary nodules are typically small and contain		
	central niduses with speckled patterns		
Edge	Benign nodules have well-defined borders while malignant nodules		
characteristics	are irregular or elongated.		

1.2 Overview of current technologies for lung cancer detection

Lung cancer may be seen on chest radiographs and computed tomography (CT) scans. The diagnosis is confirmed by biopsy which is usually performed by bronchoscopy or CT-guidance.

1.2.1 Biopsy

Biopsy may be done during a bronchoscopy. With the patient under sedation, the doctor inserts a small tube through the mouth or nose and into the lungs. The tube, which has a light, small camera and a surgical instrument on the end, allows the doctor to see inside the lung and remove a small tissue sample.

Disadvantages (complications) of lung biopsy:

- Lung biopsy may enhance your probability of developing a collapsed lung (pneumothorax) during the biopsy. Your physician may need to put a tube in your chest to sustain the inflation of your lung while the biopsy site heals
- Severe bleeding (hemorrhage) may be witnessed
- An infection like pneumonia may occur and such infections can ideally be gone if treated with antibiotics
- Spasms of the bronchial tubes can erupt, which can lead to breathing troubles just after the biopsy
- Irregular heart rhythms, which is also acknowledged as arrhythmias, may surface
- People with serious lung disorder have a very small probability of dying from the biopsy. If you get general anesthesia, there is rare chance of death from complications allied with general anesthesia.

1.2.2 Chest X-ray

Plain chest radiographs show obvious evidence of lung cancer or nonspecific appearances. In some cases, the initial radiograph is normal, either because the lesion is small, or the disease is confined to central structures.

X-ray is less effective in detection lung cancer compared to CT because it doesn't provide clear show of lung due to the appearance of the Thoracic Cage. As Fig. [2]



Fig. [2] Chest X-ray image

1.2.3 Computed Tomography (CT)

CT is considered to be the most accurate imaging modality available for early detection and diagnosis of lung cancer. It allows detecting lung nodules as small as 1mm in diameter.

CT provides excellent images of the lungs and mediastinal structures. Narrow slice, high-resolution CT scans show the lung parenchyma well, while thicker slice staging CT scans are used for diagnosis of malignant disease.

1.2.4 Positron emission tomography (PET) imaging/CT

This is the investigation of choice for characterizing extent of mediastinal nodal involvement and highlighting distant metastases either not visualized or indeterminate on CT. Most commonly, PET images are combined with CT for best correlation. Negative predictive value is high but a positive node on PET-CT should prompt sampling for confirmation of presence of malignant cells as the positive predictive value is relatively low

2 System Design

In this chapter we will talk about our project objectives, then we will give a note about CAD system and how we will visualize our system. Finally, we will show our framework.

2.1 Design Objectives

The objective of our project is to develop an automated computer aided detection (CAD) system that can detect lung nodules using computed tomography (CT) lung images, and can provide the following features:

- 1) Detection of lung nodules in early diagnosis for lung cancer.
 - Early detection for lung cancer increase the likelihood of treatment.
- 2) Accurate 3D segmentation of the lung fields.
 - A step which is required in most CAD systems for the detection and diagnosis of lung diseases.
- 3) Visualization of 3D lung fields and 3D lung nodules.
 - An additional feature for the radiologist to check the 3D organs of the patients.

2.2 CAD System

Computer-aided diagnosis (CAD) systems for lung cancer from a host of different image modalities. The success of a particular CAD system can be measured in terms of accuracy of diagnosis, speed, and automation level.

Image processing and visualization techniques for volumetric CT data sets may improve the radiologist's ability to detect small lung nodules.

CAD System steps:

- 1. Lung segmentation.
- 2. Nodule detection.
- 3. Nodule segmentation.
- 4. Nodule diagnosis.

CAD systems for detection (CADe): categorizes the nodule candidates identified in the previous step into nodules or non-nodules

CAD system for diagnosis (CADx): classifies detected nodules into benign or malignant nodules.

2.3 Visualize

We implement our algorithm in the MATLAB R2015b. It is divided into functions and



3 Data Collection

Recall that the purpose of this CAD system is to automatically detect nodules from CT images, so we had to collect lung CT data, also we needed to have ground truth for nodules in each CT sample, in order to train and test our system.

Data affected our procedure to some extent, we decided to depend on growth rate of candidate nodules to diagnose it into benign and malignant, that method which requires data distributed on several times, for example we need to have the patient's scan now, and his scan after 6 months, in order to measure the change of volume in each candidate, but we did not manage to find the suitable data "pre & post data" for the same patient, so we changed our procedure into feature extraction as illustrated above in chapter 3.

Our test data consist of 10 CT volume scans (10 patients, 2830 slices), we have 8 volume scans (8 patients, 2080 slices) for which ground truth nodule information was available. In total, there were 25 known nodules, size of slices is 512*512, all data was acquired on Siemens CT machine. Resolution of images was 0.76*0.76*1.25.

Our data mainly acquired from online platforms, such as: ELCAP Public Lung Image Database [], and Cancer imaging archive [].

We tried to get some local data, from Mansoura University hospitals, but we did not find the data with its ground truth in order to be able to test it, we contacted 57357 Children Cancer hospital, but the hospital policy did not allow them to export such data about their patients.







Fig. [3] Data samples

4 Lung Segmentation

The segmentation of lungs from chest images is a crucial step in any CAD system that can lead to the early diagnosis of lung cancer.

In this chapter we will talk about the first step (lung segmentation) of our system and its importance, then we will have a glimpse about segmentation method, after that we will present our proposed method, performance evaluation and our result.

4.1 Lung Segmentation

In CT images, Healthy lung tissues form dark regions compared to other parts of the chest such as the heart and the liver which appears lighter.

In segmentation step we separate the lung from the background.

Substance	HU
Air	-1000
Nodules	-150
Fat	-180
Water	0
Muscle	40
Bone	1000

The intensity in CT scans is from -300 to 0 HU.

4.2 Methods of Segmentation:

Techniques of lung segmentation can be classified into four categories:

- 1) Signal threshold
- 2) Deformable boundaries
- 3) Shape models
- 4) Edges models

Hu et al.	Computed iteratively such a threshold to get an initial lung region.
	Then the initial segmentation was refined by opening and closing
	morphological operations.
Yim et al.	Extracted the lung fields by the region growing followed by
	connected-component analysis.

Armato et al.	Used gray-level thresholding to segment the thorax from the
	background first and then the lungs from the thorax.
Wei et al.	Selected a threshold to segment the lung regions using histogram
	analysis. The segmented lungs were then refined using connect-
	component labeling (CCL) and circular morphology closing.
Ye et al.	Used 3D adaptive fuzzy thresholding to segment the lung region
	from CT data. The segmentation was followed by smoothing the
	segmented lung contour, represented as chain code, by 1D
	Gaussian smoothing. They further applied a methodology to detect
	the lung nodules in the segmented lung fields.

4.2.1 Problems of threshold-based segmentation:

- Accuracy affected by many factors, including image acquisition protocol and scanner type.
- Densities (in Hounsfield units) of some pulmonary structures such as arteries, veins, bronchi, and bronchioles, are very close to densities of chest tissue.
- Needs further intensive post-processing steps to overcome the inhomogeneity of densities in the lung region.

4.3 Method we used:

We used single threshold method in our system.

4.3.1 Iterative Threshold:

The simplest method of lung segmentation, it changes the grayscale image to a binary image, by determine the threshold value then apply the following equation to the image pixels

$$f(x,y) \begin{cases} 1 & if(x,y) \ge Threshold \\ 0 & if(x,y) < Threshold \end{cases}$$

In this method threshold value for segmentation is calculated using iterative method.

4.3.2 Algorithms:

2D

Algorithm 1 2D Lung Segmentation

Input: Sample 2D CT Image to be segmented **Initialization:**

• Select a random initial threshold value (T_i) (optional: *thresh* may be set as the average gray value of image)

Processing Steps:

Consider the two image regions R1 and R2 that are determined by T_i . (i.e. pixels above *thresh* and pixels below *thresh*):

- 1) Compute the average gray values (avg1, avg2) for R1 and R2
- 2) Compute a new threshold using:

$$T_i = \frac{avg_1 + avg_2}{2}$$

3) **Termination criterion:** Repeat (1-2) until avg1 and avg2 do not change between successive iterations, and output the optimal threshold T_f

Output: Output the 2D segmented lung using T_f

Note that you will have to save copies of avg1 and avg2 (old1, old2) between iterations to see if they have changed.

3D

Consider all the volume



Fig. [4] Flow Chart of iterative Algorithm

4.4 Performance Evaluation:

We used Sørensen–Dice coefficient to calculate the accuracy of our lung segmentation algorithm.

4.4.1 Sørensen–Dice coefficient

The **Sørensen–Dice index**, is a statistic used for comparing the similarity of two samples. It was independently developed by the botanists Thorvald Sørensen and Lee Raymond Dice, who published in 1948 and 1945 respectively.

4.4.2 Formula

Sørensen's original formula was intended to be applied to presence/absence data, and is

$$QS = \frac{2 |X \cap Y|}{|X| + |Y|}$$

Where |X| and |Y| are the numbers of species in the two samples. QS is the quotient of similarity and ranges between 0 and 1. It can be viewed as a similarity measure over sets.

Similarly to the Jaccard index, the set operations can be expressed in terms of vector operations over binary vectors *A* and *B*:

$$s_v = rac{2|A \cdot B|}{\left|A
ight|^2 + \left|B
ight|^2}$$

Which gives the same outcome over binary vectors and also gives a more general similarity metric over vectors in general terms.

For sets *X* and *Y* of keywords used in information retrieval, the coefficient may be defined as twice the shared information (intersection) over the sum of cardinalities:

When taken as a string similarity measure, the coefficient may be calculated for two strings, *x* and *y* using bigrams as follows:

$$s = rac{2n_t}{n_x + n_y}$$

Where n_t is the number of character bigrams found in both strings, n_x is the number of bigrams in string x and n_y is the number of bigrams in string y.

4.4.3 Applications

The Sørensen–Dice coefficient is commonly used in Image segmentation, in particular for comparing algorithm output against reference masks in medical applications.

4.5 Segmentation Results



Table.1. Images pre and post applying iterative threshold



Table.2. Result of Dice coefficient after applying on our segmented images

Slice	dice (2d segmentation)	3d segmentation
60	97.7787	97.778
70	98.3375	98.0559
80	92.4392	92.5941
90	93.5027	93.7528
100	97.7681	97.7771
110	97.565	97.5768
120	85.12	85.1587
130	88.9309	88.8657
140	79.9389	92.0865
150	85.6751	85.9825
Mean	91.70561	92.96281
std dev.	6.511614408	4.961351001

4.6 Discussion

Our accuracy of system performance in 2D lung segmentation is $91.7\pm6.5\%$, in terms of detection of nodules.

In 3D segmentation, the accuracy is $93\pm4.9\%$, the system approximately 2 minutes to segment a volume of 250 slices.

5 Nodule Detection

At the beginning of this chapter we will present the detection framework, then we will explain our detection algorithm and the set of feature used to extract nodules, after that we will identify neural network (NN) the proposed classification method, finally discuss the detection results.

CAD systems for detection of lung nodules in thoracic CT (CADe) generally consist of two major stages:

- 1) Initial candidate nodules are selected (identification of nodule candidates)
- 2) The false positive nodules (FPNs) are partially eliminated while preserving the true positive nodules (TPNs) (classification of the identified nodule candidates into nodules or non-nodules (normal anatomic structures)).

5.1 Detection framework

This the final stage of the project, where all candidates are judged to be whether nodule or non-nodule. So far the previous chapters described the methods used to read, handle and segment the lung in DICOM volumes. In this chapter we are going to discuss how to extract the connected components from the lung, classify them and finally detect which of these components are nodules.



5.2 Candidate nodule selection

Candidate nodules are all components that are suspected to be nodules. The previous step resulted in segmenting the lung (Fig.5.2) from the volume. Beginning with the segmented lung we create a mask (Fig.5.3) using opening and closing operations.









By multiplying the mask and the segmented lung image, we get the inside components -candidate nodules- of the lung (Fig.5.4)



Fig.5.4

Of course these candidates has a quite large number, so they have to be reduced. To do so, we used connected components analysis to learn more about these candidates' properties such as area (no. Of pixels), centroids, bounding boxes and so on.

Depending on our training data the smallest nodule area was about 38 pixels, and the largest was about 1580 pixels. So we eliminated all pixels that contain number of pixels more than 1600 pixels, and less than 30 pixels.

5.3 Feature extraction

There many criteria on which the nodule specification depends. In our project we concentrated on the Ellipsoid fitting. We focused on calculating the ellipsoid parameters a, b and c neglecting the effect of the rotation angle of the ellipsoid. We assumed that each parameter has value that nearly equals the half of its corresponding length of the '*BoundingBox*' obtained from '*regionprops*' function where:

a = (bounding box length)/2.

b = (bounding box width)/2.

c = (bounding box height)/2.

After calculating ellipsoid parameters, we are going to calculate the features that will be used in nodule detection.

Effective radius	$d1 = (abc)^{\frac{1}{3}}$
Dis-circularity	d2 = max (a, b, c) - d1.
Elongation	$d3 = \max(a, b, c) / \min(a, b, c).$
Number of pixels	$d4 =$ number of pixels \in E.
Mean intensity	$d5 = \left(P(x, y, z) \in E f(x) \right) / d4.$
Intensity sum	$d6 = P(x, y, z) \in E * f(x).$
Variance of intensity	d7 = var { $f(x, y, z) (x, y, z) \in E$ }
Threshold	d8 = the intensity threshold T.

We calculated the values of these feature for both nodules and non-nodules in our training data. In table.1 you will find samples of these values of selected sample of nodules..

Table 5.1 Typical feature values extracted from a sample of seven nodules

Nodules	d1	d2	d3	d4	d5	d6	d7	d8
1	7.5199	2.9801	2.3333	1581	55.9221	32767	227.8576	395.2344
2	6.5693	2.4307	2	927	50.3358	32767	344.7356	389.1318
3	2.3811	0.6188	2	58	34.6903	2012	197.3387	282.5238
4	5.0016	1.4983	1.8571	545	45.2861	24681	359.5723	390.1904
5	2.7589	0.7410	1.75	85	40.7648	3465	316.8475	332.4029
6	3.1748	0.8251	2	121	32.0376	3877	249.7301	319.3260
7	4.1212	0.8787	1.4285	255	89.7848	22895	575.4879	609.5091
		•	•		•	•	•	•

Table 5	•	1
---------	---	---

And for non-nodules..

Non-Nodules	n1	n2	n3	n4	n5	n6	n7	n8
1	2.9624	3.5375	6.5	52	47.1904	2973	66.2534	219
2	3.2514	2.2485	2.2	38	44.2682	41922	46.966	224
3	23.4964	0.5035	1.0434	1548	53.6204	2419946	117.9027	57
4	3.402	1.09795	1.8	42	50.3923	202527	94.7427	213
5	2.28942	1.7105	4	46	43.567	4226	62.998	288
6	8.47628	6.0237	2.4166	345	48.338	684467	92.4125	254

Table 5.2 Typical feature values extracted from a sample of six non-nodules..



5.4 Nodule detection and classification

In this section we will use neural networks to classify and detect the nodules depending on the values of the features calculated in the previous section. The neural network use these features as input. In our project, we used 60% of nodules as a training data, 20% for testing and 20% to validate.

We tried variable number of hidden layers, in order to find t the optimum accuracy (Fig.5.5).



Fig.5.5

The best results were achieved at 10 hidden layers. The network structure is 10*2, which means we have 10 neurons in the hidden layer (Fig.5.6) and two neurons in the output layer.



Fig.5.6

The results (Fig.5.7) showed accuracy of 85.7% in detecting nodules which can be increased by adding more training data in the future work.



Fig.5.7

6 Visualization and GUI

Our GUI interface Fig. [6.1]



Fig. [6.1]

Our GUI have three buttons as Fig. [6.2]

- Load Volume load the patient CT images
- Segment process of the first step –lung segmentaion-
- Classify detect the nodules

Pan	el	
	Load Volume	
	Segment	
	Classify	

After loading the volume folder, the CT image appear as Fig. [6.3].





After press the segment button, the Segmented image appear for each CT image (input), there is a slider to move through the slices easily as Fig. [6.4]

You can also know which slice is observed by the slice number





3D Nodule Visualization

7 Conclusion and Future Work

7.1 Conclusion:

To find reliable method for nodule detection is an important problem in medicine. By starting a screening program based on this method, survival rate should be improved because of early detection of lung cancer.

We have developed an automatic method for the nodule detection. We are using two steps for ROI detection and lowering number of FPs among the detected nodule candidates. Also some volumetric information about nodules can be obtained from our method. The method was validated on data with 25 real nodules.

Our system performance in 2D lung segmentation, accuracy is $91.7\pm6.5\%$, where in 3D segmentation it is $93\pm4.9\%$, in terms of detection of nodules, more training data is required to improve reliability of the system.

In 3D segmentation, the system approximately 2 minutes to segment a volume of 250 slices.

By adding some features and modifications in future work, we claim that our system can be used in the Egyptian lung cancer hospitals and clinics to facilitate the detection of lung nodules automatically.

7.2 Future Work:

Acceleration. System can be significantly accelerated by new fitting algorithm, and by using C programming language "low level language" in some parts of code.

New descriptors. Classification can be improved by new shape descriptors of nodule.

Training. System reliability can be enhanced by training more data.

Diagnosis. We can add parts to diagnose the selected nodules into benign and malignant

References

- Kumar and Clark's Clinical Medicine, Respiratory Diseases.
- Computer-Aided Diagnosis Systems for Lung Cancer: Challenges and Methodologies Ayman El-Baz, in international journal of biomedical imaging volume 2013.
- Detection of Pulmonary Nodules from CT Scans, Martin Dolej's'I, January 19, 2007
- 3D Shape Analysis for Early Diagnosis of Malignant Lung Nodules.
- Techniques for Detection of Solitary Pulmonary Nodules in human lung and their classifications a survey, Bhavanishankar .K and Dr. M.V.Sudhamani, International Journal on Cybernetics & Informatics (IJCI) Vol. 4, No. 1, February 2015
- Segmentation of Lung Structures in CT, PhD thesis Pechin Chien Pau Lo, Submitted: 25/03/10