Mansoura University

Automatic Detection of Lung Cancer

Assis. Prof. Ahmed Elnakib

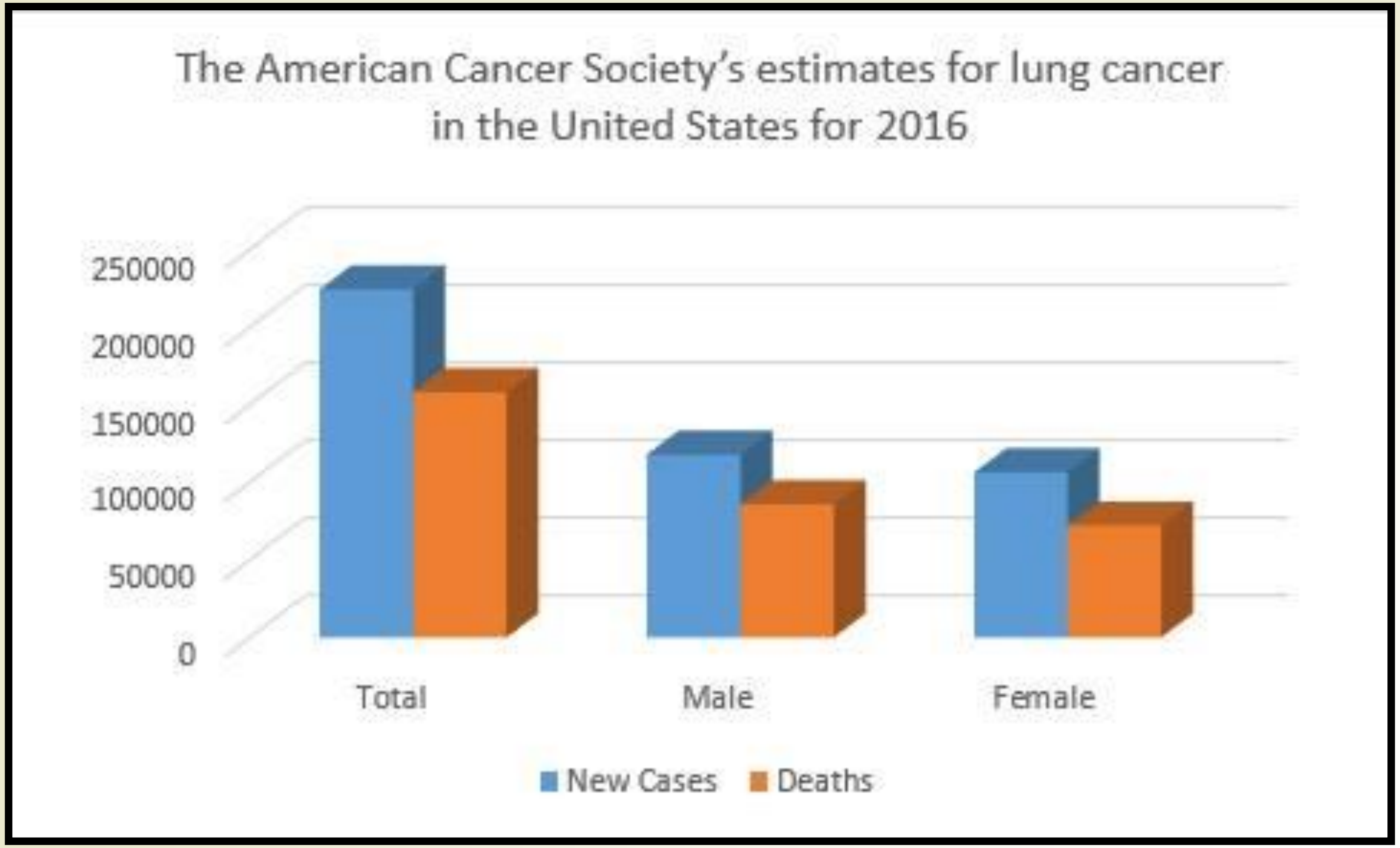
- Ahmed Hassaan
- Ahmed Sallam
- Abdelrahman Hagag
- Amina Alaa
- Taghreed Fayez

Motivation

- Lung cancer is the first most common cancer in men, and second most common cancer in women.
- Lung cancer 12.8% of cancer cases in the world and the proportion of 17.8% of deaths due to cancer around the world.
- About 14% of all new cancers are lung cancers.
- In our country, Egypt receives annually 150 new cases of cancer, for every 100 thousand inhabitants.

Motivation

	Total	Men	Female
New cases	224,390	117,920	106,470
Deaths	158,080	85,920	72,160



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:

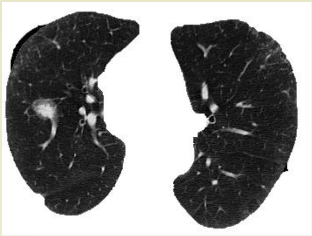
1. Detection of lung nodules in early diagnosis for lung cancer
    - Early detection for lung cancer increase the likelihood of treatment.
2. Accurate 3D segmentation of the lung fields
    - A step which is required in most CAD systems for the detection and diagnosis of lung diseases.
3. Visualization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

Related works..

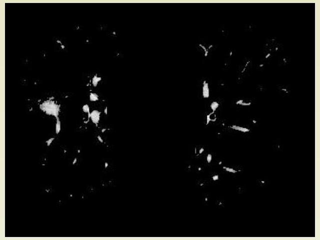
Author	Title	Place	Year
A. El-Baz, et. al	Computer-Aided diagnosis systems of lung cancer: challenges and methodologies	In international journal of biomedical imaging	2013
Ayman El-Baz	3D Shape Analysis for Early Diagnosis of Malignant Lung Nodules.	University of Louisville, Louisville, KY, USA	2007
Martin Dolejˇs´ I. Thesis Advisor: Dr. Ing. Jan Kybic.	Detection of Pulmonary Nodules from CT Scans.	Research Reports of CMP, Czech Technical University in Prague, No. 5, 2007	January 19, 2007

The Proposed Analysis Framework:

Input:
CT images.



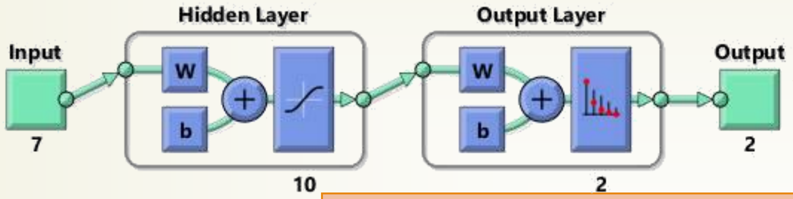
Step 1:
3D Lung segmentation.



Step 2:
Candidate Lung nodules segmentation.

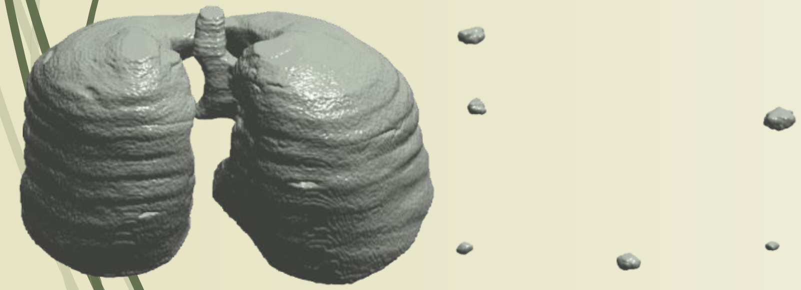
$$\begin{matrix} d_1 & d_2 & d_3 \\ d_4 & d_5 & d_6 \\ & d_7 & \end{matrix}$$

Step 3:
Features Extraction.



Step 4:
Neural Network Classification.

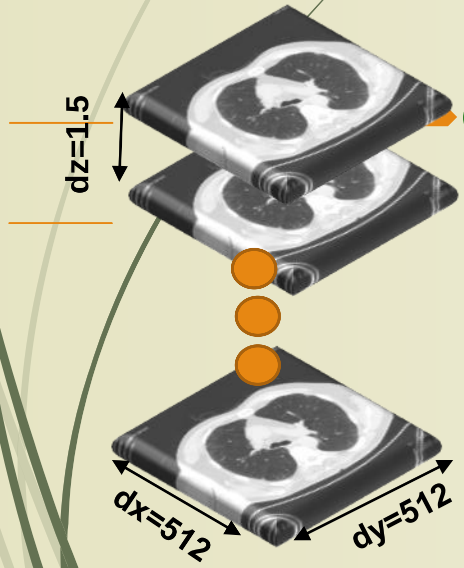
Output:
Nodule Detection
(Nodule or non Nodule)



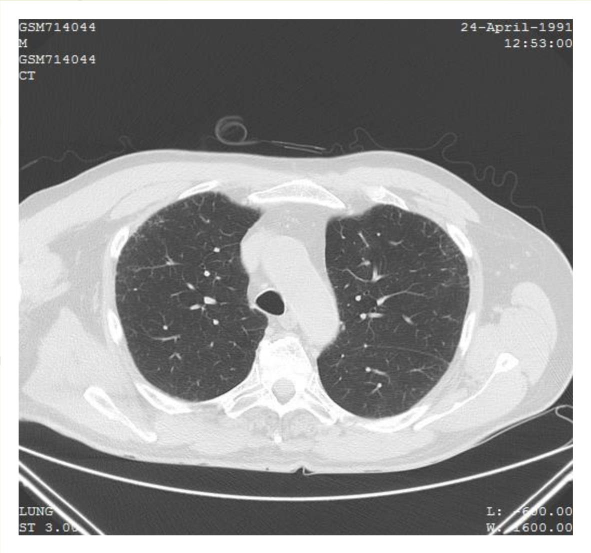
3D segmentation

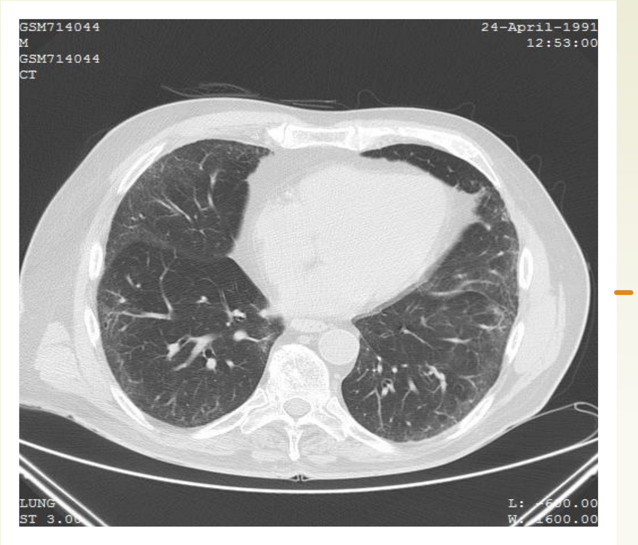
Data collection

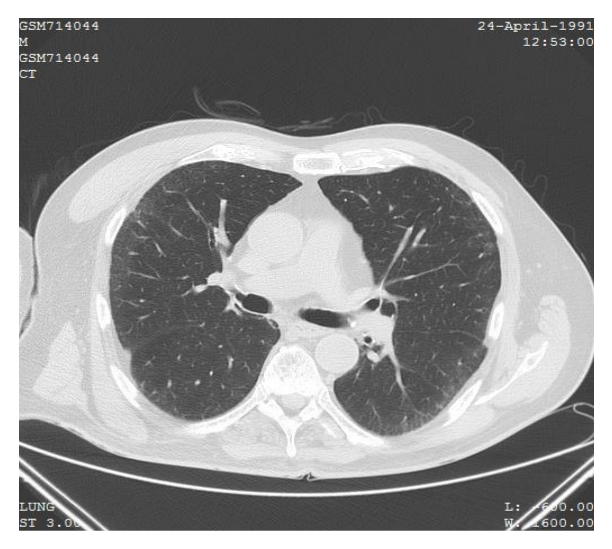
- Our data mainly acquired from online platforms, such as: ELCAP Public Lung Image Database , and Cancer imaging archive
- Using computed tomography scanning technique(CT imaging).
  - we needed to have ground truth for nodules in each CT sample, in order to train and test our system.
- Our test data:
  - 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.



Data samples







Step 1:

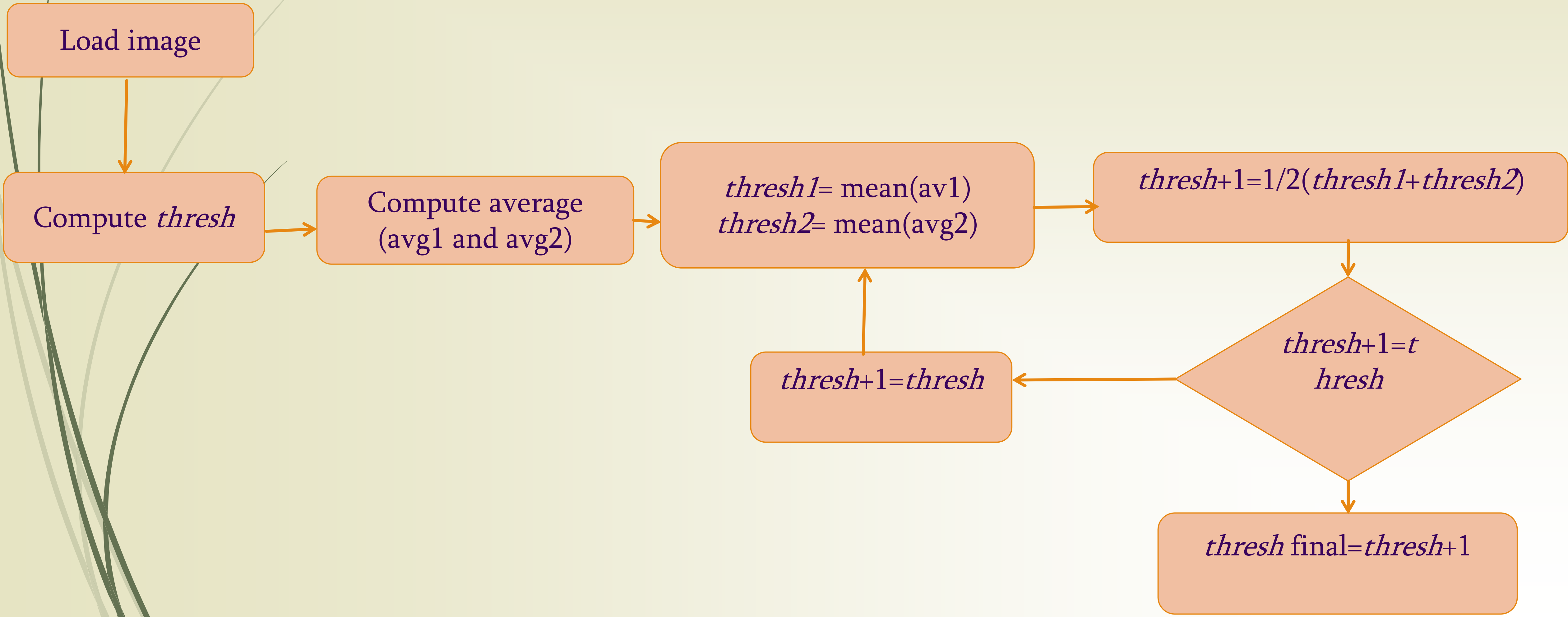
Lung segmentation.

- In segmentation step we separate the lung from the background.
- The lung region is segmented using iterative threshold to get an initial lung region. “Otsu’s method”.
- Iterative threshold changes the grayscale image to a binary image, by determine the threshold value, then apply the following equation to the image pixels

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

- We segment the lung in 2D & 3D.

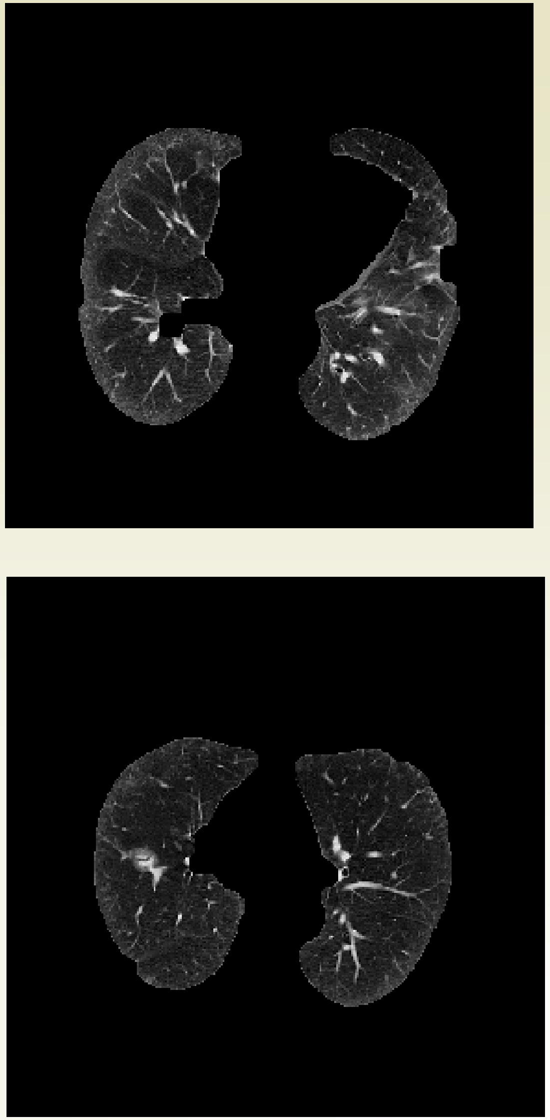
Our Algorithm for 2D Lung Segmentation..



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

Input image

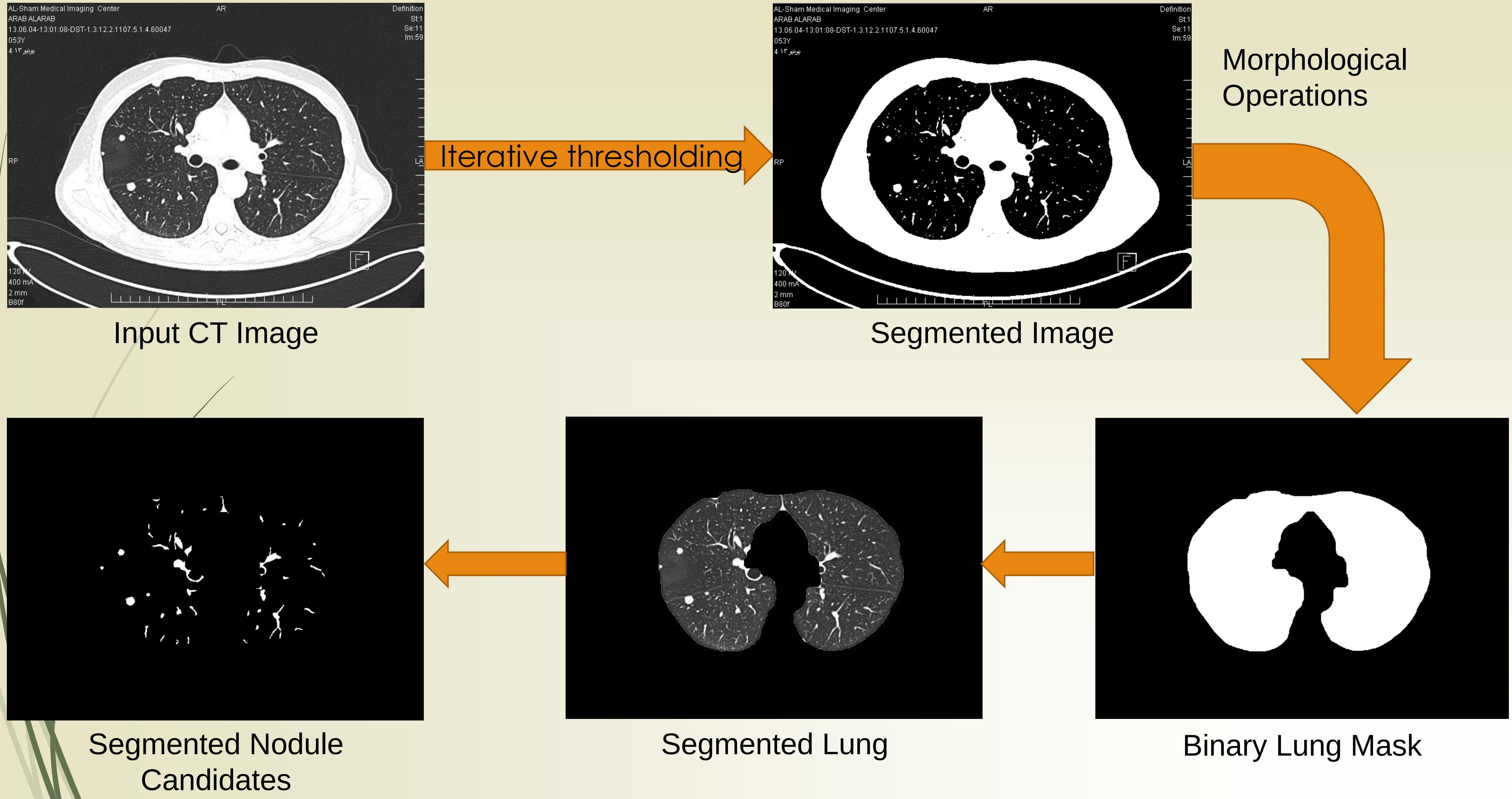


After segmentation

Step 2:

Candidate lung nodules segmentation.

- In 2D we use region growing to segment the nodules candidates.
- In 3D we use Connected component technique.


Input CT Image
Iterative thresholding
Segmented Image
Morphological Operations
Binary Lung Mask
Segmented Lung
Segmented Nodule Candidates

Step 3:

Features Extraction.

- Using feature extraction, we determine 8 features for each nodule candidate.

Effective radius	**d1 = (a*b*c)^(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 ∈ E
Mean intensity	d5 = P (x, y, z) ∈ E f(x)/d4
Intensity sum	d6 = P (x, y, z)∈E f(x)
Variance of intensity	d7 = var {f(x, y, z) : (x, y, z) ∈ E}
Threshold	d8 = the intensity threshold T

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

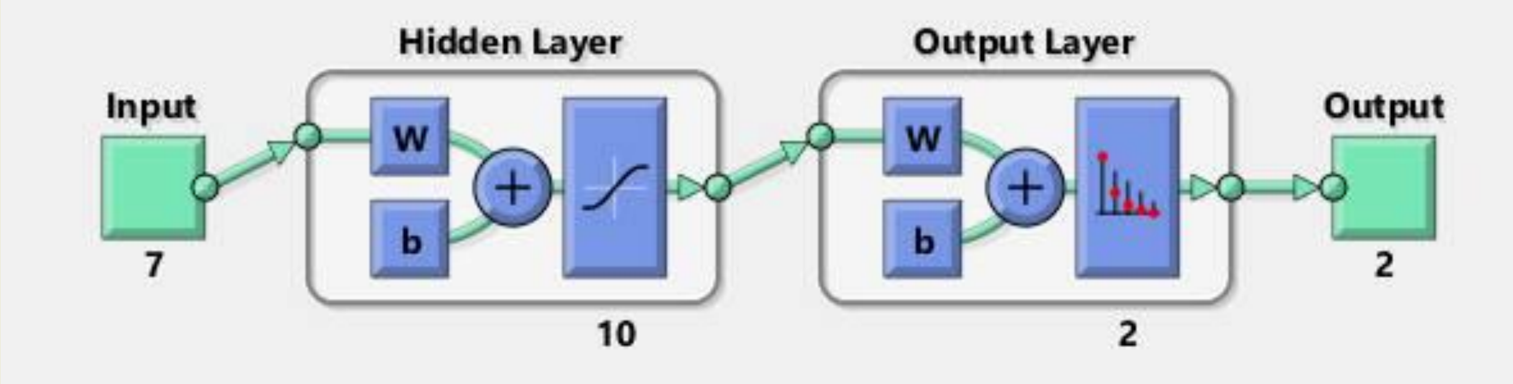
Typical feature values extracted from a sample of six non-nodules..

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

Step 4:

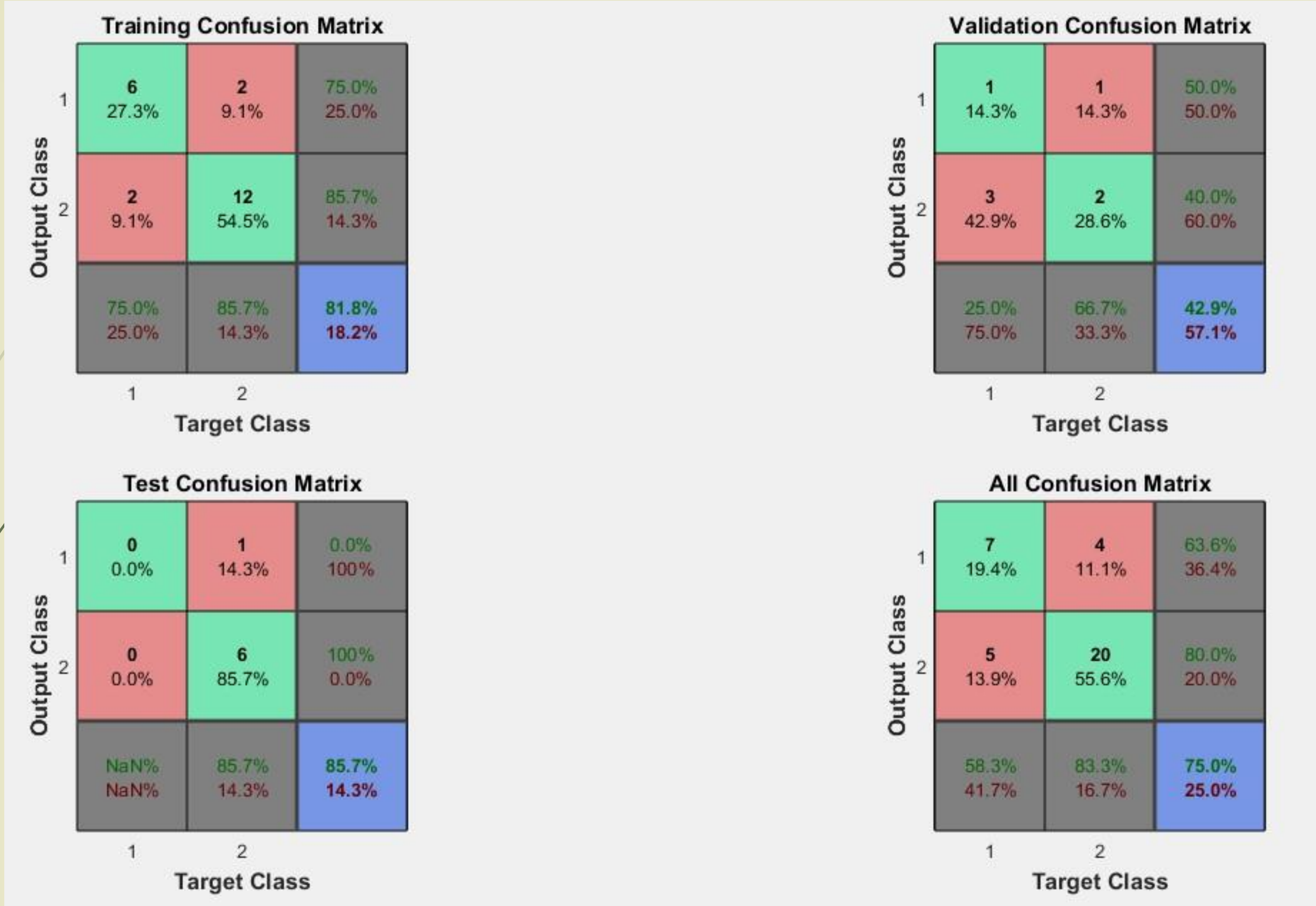
Neural Network Classification.

- Using supervised neural network to classify each nodule candidate is nodule or non nodule.
- We use 22 nodule candidates as training data, and 14 as testing data.
- We use 10*2 network.
- Experimental Results: When the network is 10*2, using 10 input neurons and 2 output, the accuracy is 85.7%



Structure	Test Data Accuracy
10*2	85.7%
8*2	71.4%
6*2	71.4%
4*2	85.7%
2*2	71.4%

Neural Network Results


Training Confusion Matrix
6
27.3%
2
9.1%
75.0%
25.0%
2
9.1%
12
54.5%
85.7%
14.3%
75.0%
25.0%
85.7%
14.3%
81.8%
18.2%
Output Class
Target Class
Validation Confusion Matrix
1
14.3%
1
14.3%
50.0%
50.0%
3
42.9%
2
28.6%
40.0%
60.0%
25.0%
75.0%
66.7%
33.3%
42.9%
57.1%
Output Class
Target Class
Test Confusion Matrix
0
0.0%
1
14.3%
0.0%
100%
0
0.0%
6
85.7%
100%
0.0%
NaN%
NaN%
85.7%
14.3%
85.7%
14.3%
Output Class
Target Class
All Confusion Matrix
7
19.4%
4
11.1%
63.6%
36.4%
5
13.9%
20
55.6%
80.0%
20.0%
58.3%
41.7%
83.3%
16.7%
75.0%
25.0%
Output Class
Target Class

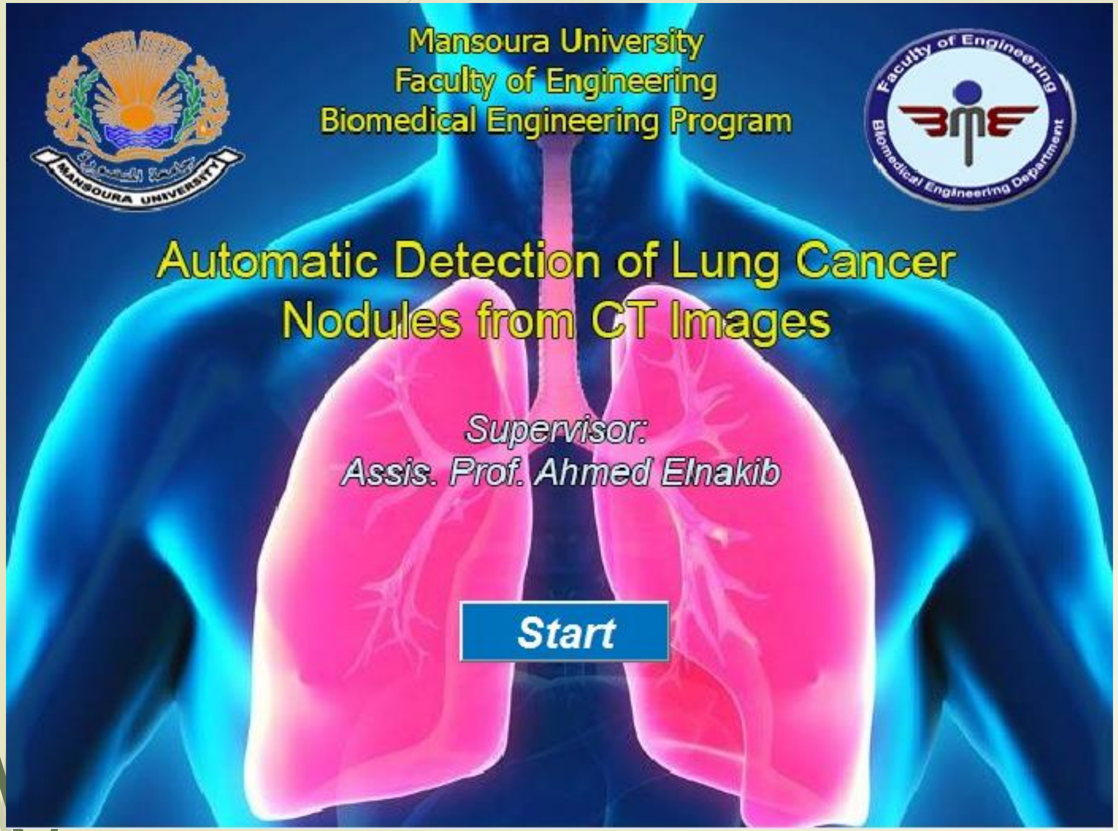
Visualization and GUI

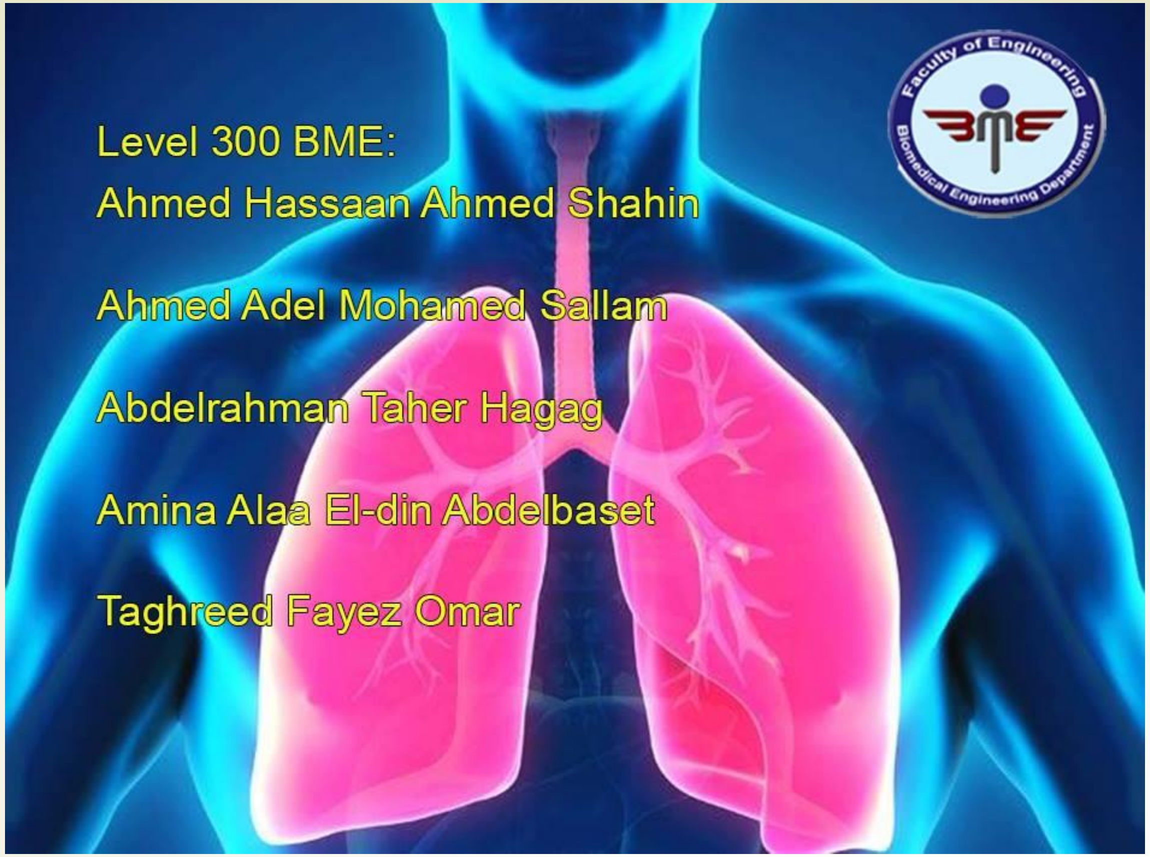
- We develop our own GUI for the project, 2D and 3D version.
- It consist of many patterns:

Allow the radiologist to load the lung volume.	**Import**
Separates the lung from background, and visualize it.	**segment**
Detect the candidate nodules.	**Detection**
Classify if the detected nodules is nodule or non nodule.	**Nodule classification**

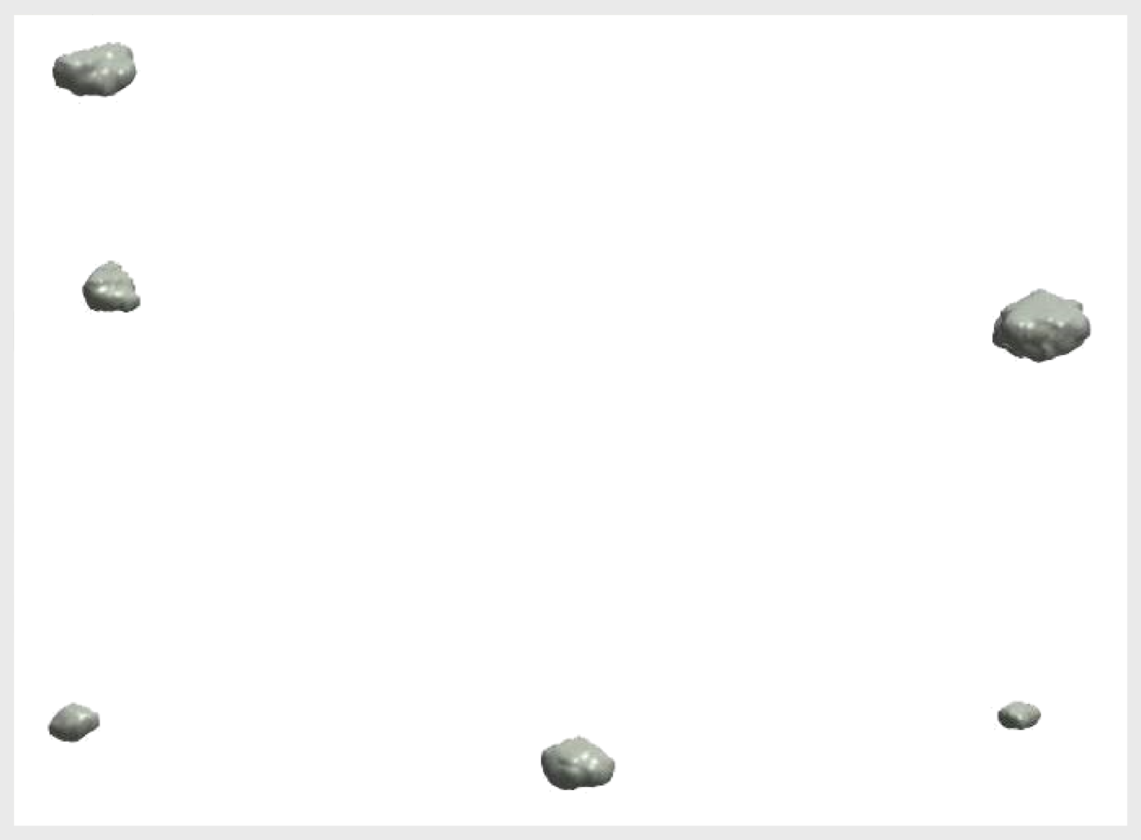
- It also contain a moving slider to move between the slices, and to be able to see the slices that contain the nodule.

Visualization and GUI

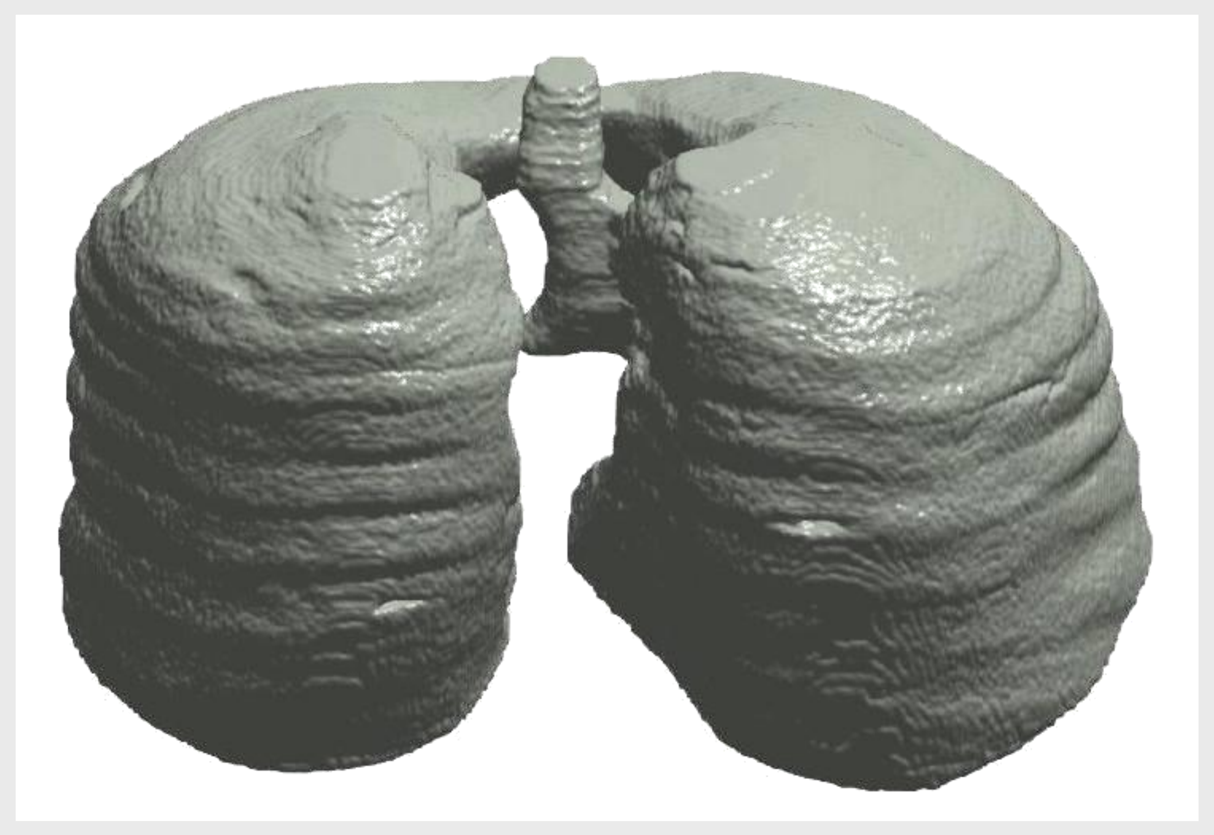





Panel
Load Volume
Segment
Classify
The Slice Number is :
29



3D nodule visualization



3D visualization for the lung

Conclusions and Future Work

- We have developed an automatic method for the nodule detection
- The method was validated on data with 25 real nodules, our accuracy of nodule detection is 85.7% for our samples.
- Our accuracy of system performance in 2D lung segmentation is 91.7±6.5%, in terms of detection of nodules.
- In 3D segmentation, the accuracy is 93±4.9%, the system approximately 2 minutes to segment a volume of 250 slices.
- Some volumetric information about nodules can be obtained from our method, which can be used in diagnosis.
- We plan to add more features and test our framework on more data in order to test the feasibility of the system to be used in the Egyptian lung cancer hospitals and clinics to facilitate the detection of lung nodules