Lung nodule detection in low-dose and high-resolution CT scans

Alessandra Retico
INFN Pisa

P.Delogu[12], M.E. Fantacci[12], I. Gori[3],
A. Preite Martinez[1], A. Retico[2], A.Tata[2]
[1] *University of Pisa*
[2] *INFN Pisa*
[3] *Bracco Imaging S.p.A. Milano*

Frontier Science 2005 – Milano, 12-17 September

Lung cancer

- Lung cancer is the main cause of cancer death in the male population and the second one in the female population

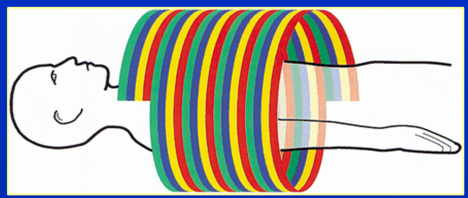
- The 5-year survival rate is 60-70% for early-stage cancers (stage IA). Most cancers are detected in advanced stages.
  - The overall 5-year survival rate is only 10-15%
  - No improvement has occurred in the last 20 years

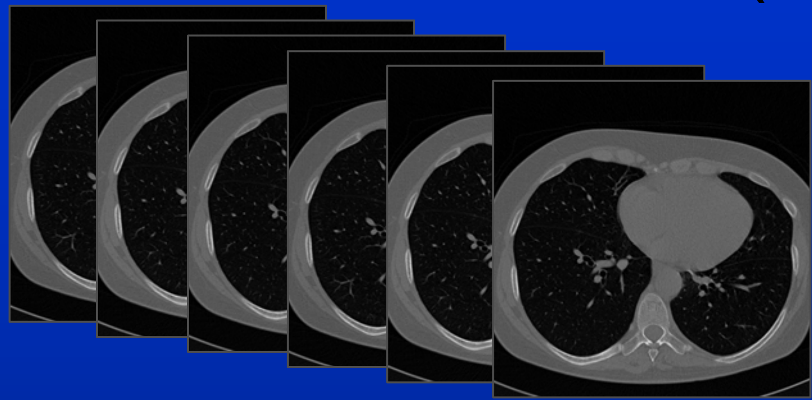
- It has been demonstrated that screening programs with chest X-ray do not lead to a reduction of the mortality rate

- Computed Tomography (CT) has proved to be more sensitive in detecting small, early-stage cancers

- The efficacy of CT-based screenings in reducing mortality rate for lung cancer has not been proved yet.
  - Trial screening programs are being carried out in the US, Canada, Japan and Europe

Screening & computer-aided detection (CAD)





High resolution CT:

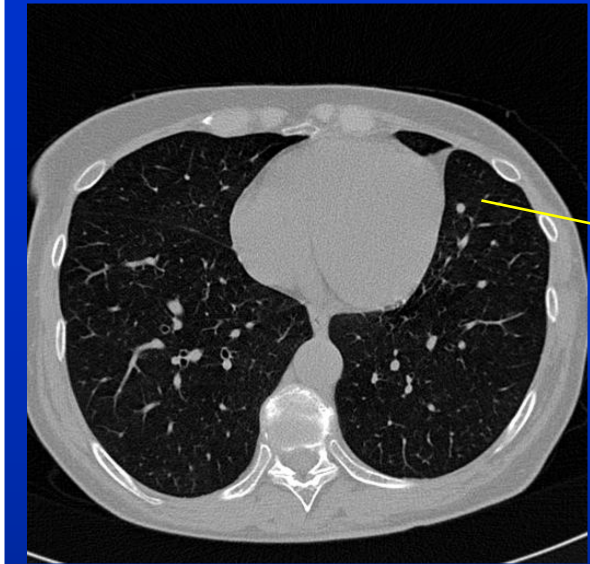
slice thickness =~1mm → ~300 slices/scan

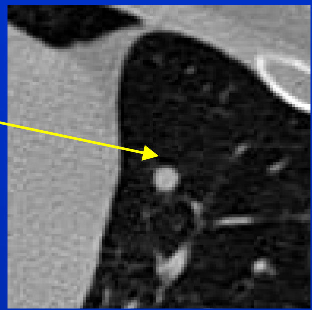
LD helical multi-slice CT	0.6 mSv
LD helical single-slice CT	1.2 mSv
Standard dose helical CT	5.0 mSv
Rx torace 2 views	0.1 mSv

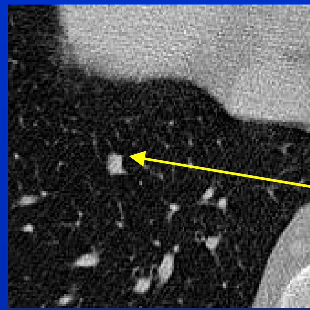
- Non-calcified nodules with diameter ≥ 5mm have to be detected
- A large number of nodules (~20-35%) are missed by radiologists in screening programs [Roberts et al, CARS 2005, pp 1137-42]
- A CAD system could be useful as second or third reader.
  It should be characterized by:
  - high sensitivity
  - low number of false-positive findings per scan

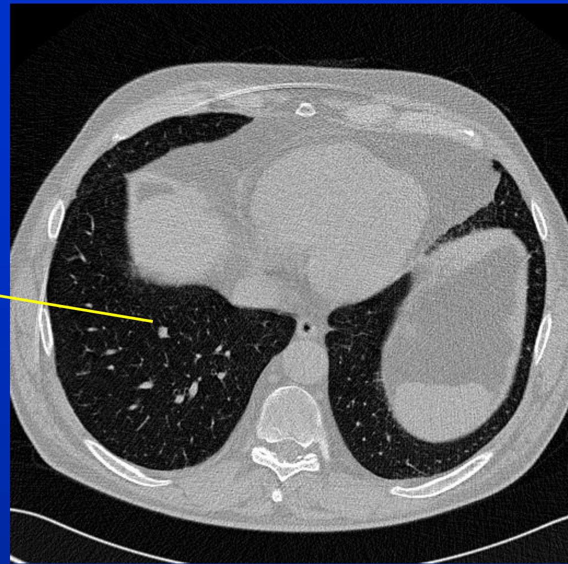
Searching for lung nodules …

Nodules





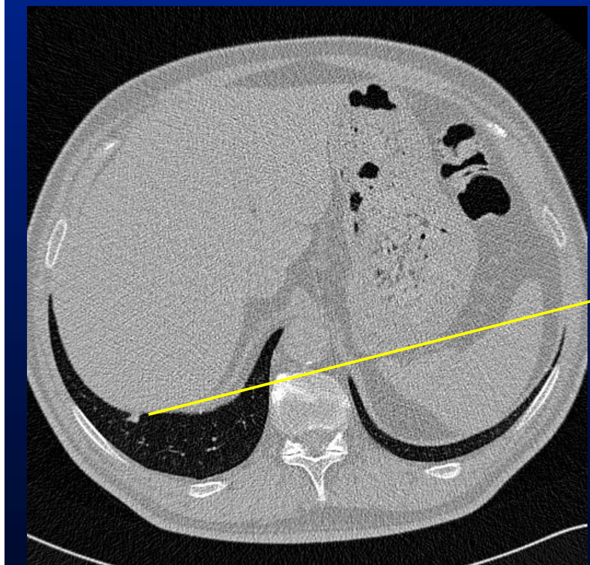


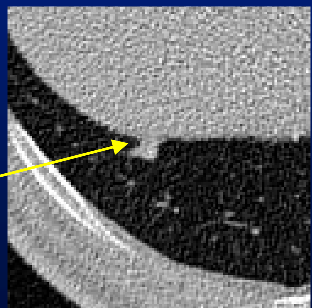


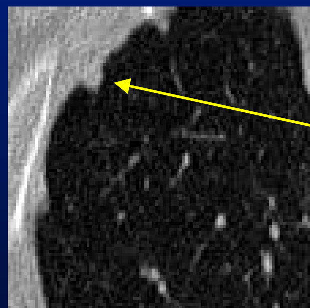
Almost spherical shape

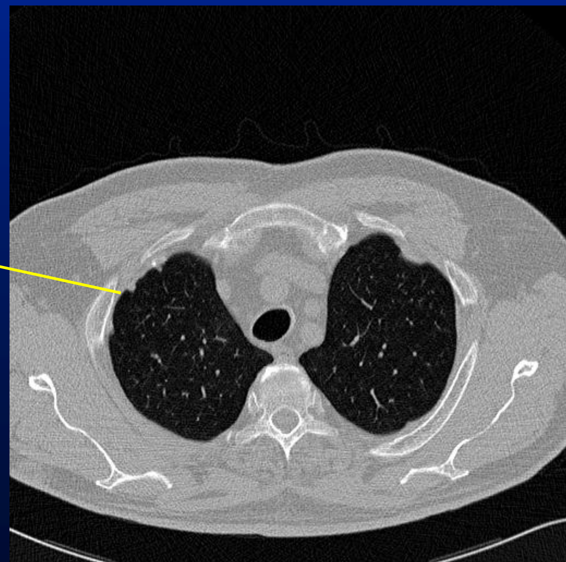
lobulated, popcorn shaped

A pulmonary nodule is an object with a diameter ≥ 5mm



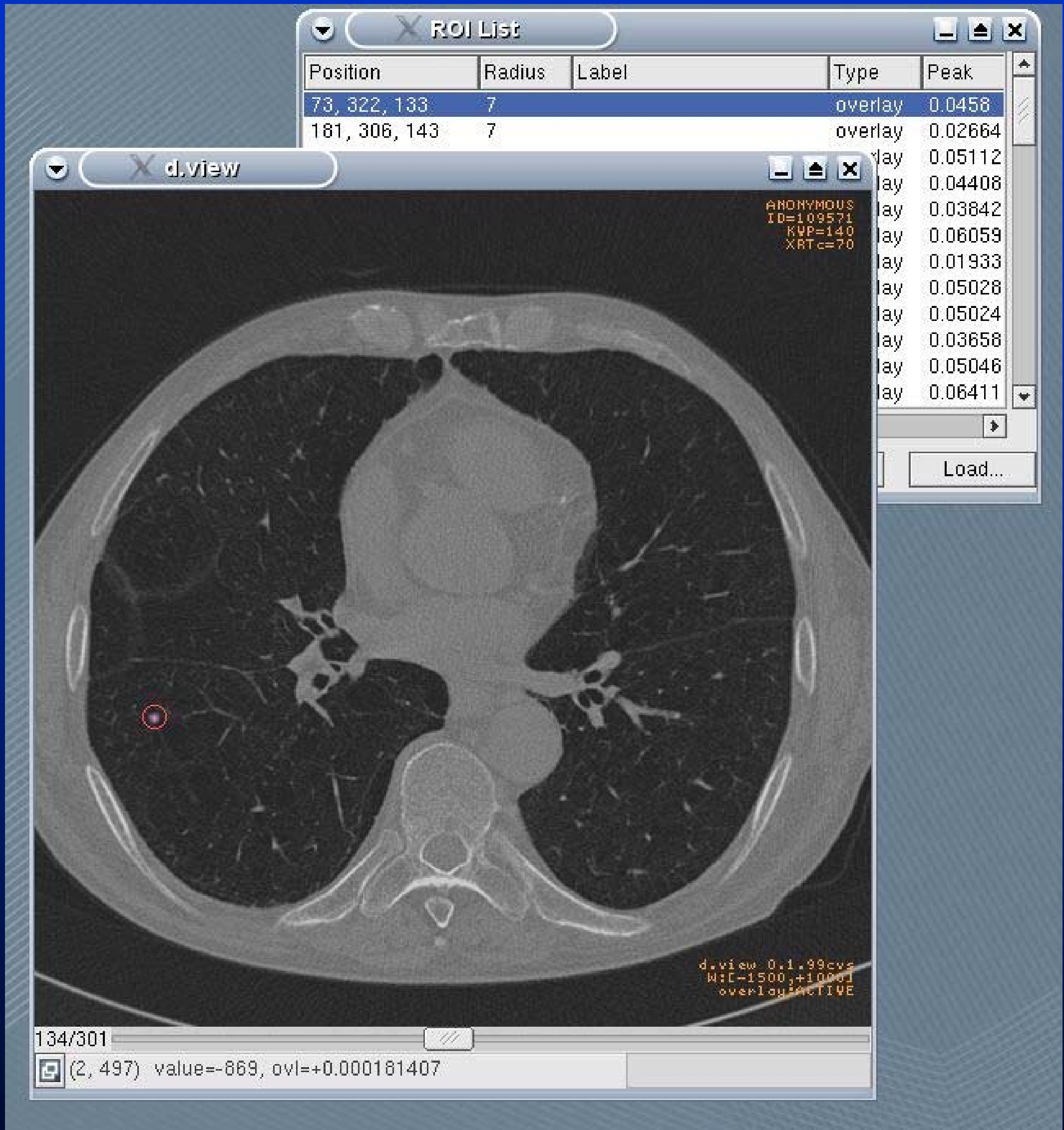






More complicated shapes

Operative framework

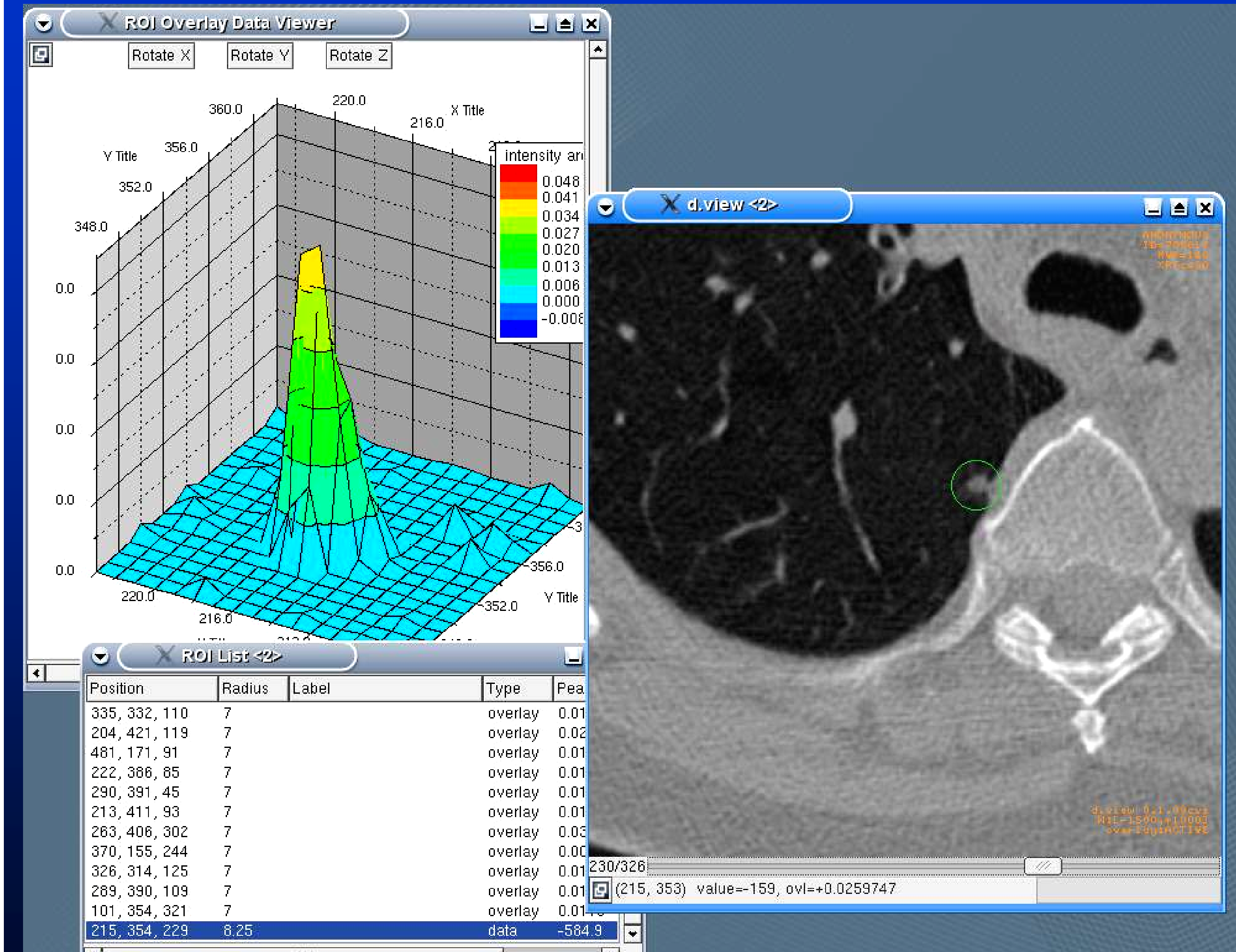


DICOM (Digital Imaging Communication in Medicine)

Simple DICOM Tools:

- based on DCMTK open-source software
- works on UNIX OS
- Software already implemented:
  - anonymizer
  - sorter
  - 3D matrix builder
  - viewer

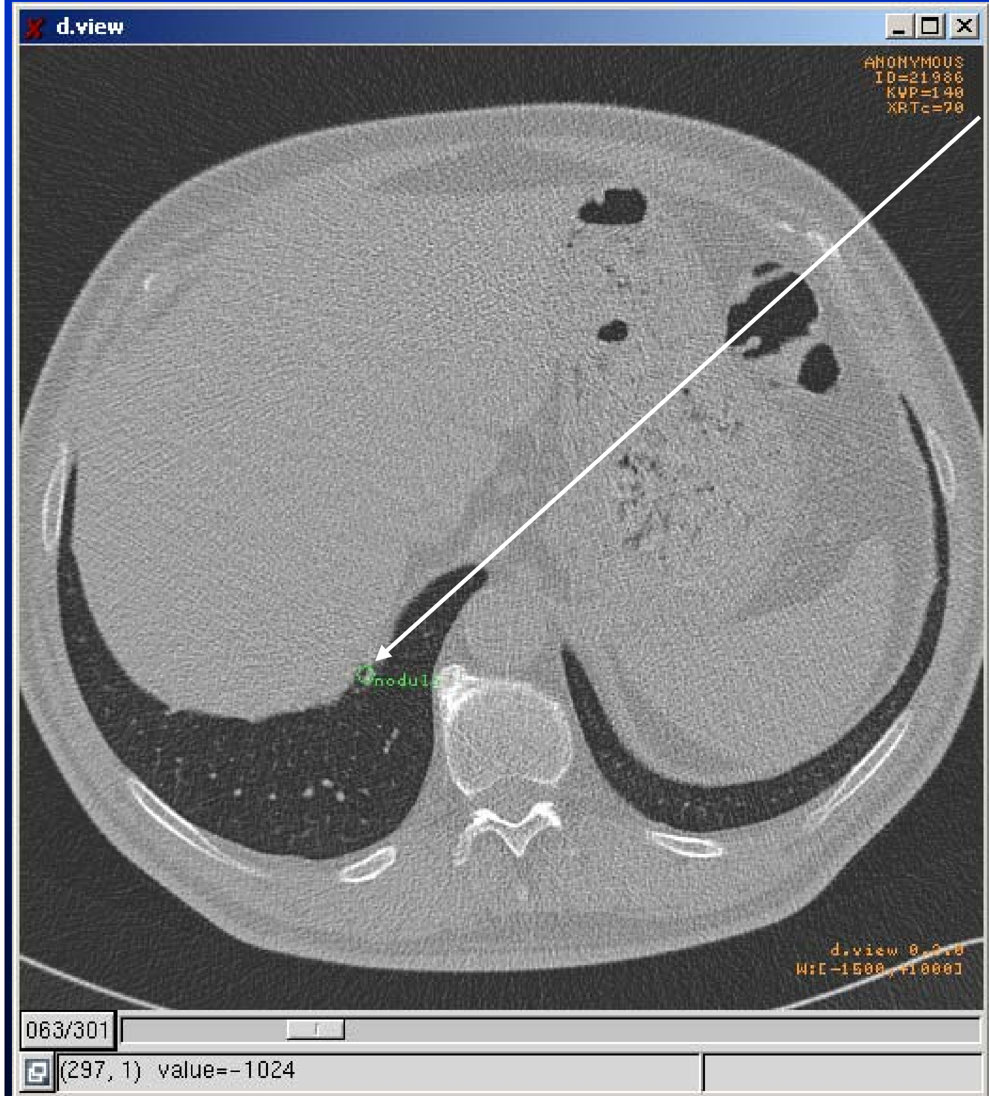
Operative framework



The viewer allows:

- to browse the scans
- to choose the visualization window (range of Hounsfield units)
- to zoom in/out
- to analyze the statistics in the regions of interest
- to measure geometrical distances
- …

Radiologist’s annotation



ROI List

Position	Radius	Label	Type	Peak	Average
74, 289, 217	4.92	nodulo	data	360	-431
97, 349, 67	4.03	nodulo	data	363	-402
78, 352, 61	5.66	nodulo	data	443	-355.2
184, 334, 62	5.39	nodulo	data	589	-396.2
140, 279, 270	7.43	micronodulo	data	752	-351
111, 258, 254	6.8	micronodulo	data	269	-344.5
119, 209, 232	6.26	micronodulo	data	474	-385.8
141, 188, 226	4.72	micronodulo	data	315	-278.5
118, 188, 116	5.5	micronodulo	data	292	-448.6
97, 377, 54	7.43	micronodulo	data	867	-371.7

Close | Add | Remove | Histograms | Save... | Load...

A radiologist draws circular regions of interest (ROIs)

Each ROI is identified by:

- the center coordinates (x,y,z)
- the radius
- the radiological class

CT database (at present)

The database is acquired in the framework of the first Italian Randomized Controlled Trial for lung cancer screening (Firenze, Pisa, Pistoia):

- 3000 high-risk individuals are enrolled
  (smokers ≥ 20 pack-year; 55→69 year-old men and women)
- ~400 scans/year are expected to be acquired in Pisa Hospitals

Total number of available scans	Total number of fully annotated scans	Number of fully annotated scans with nodules
103	57	16 (28%)

The CT scans are annotated by radiologists of the U.O. Radiodiagnostica 2 of the Azienda Ospedaliera Universitaria Pisana

Database features

CT model	KV	mA	Slice thickness	Slices per scan
Multi-slice	120÷140	20÷80	1.25 mm	about 300

- Low dose CT:
  - more noisy with respect to standard-dose CT
- High resolution CT:
  - less partial volume artifacts
  - more accurate 3D reconstruction

  - high number of slices per scan
  - high number of false positives

CAD schema

3-stage algorithm:

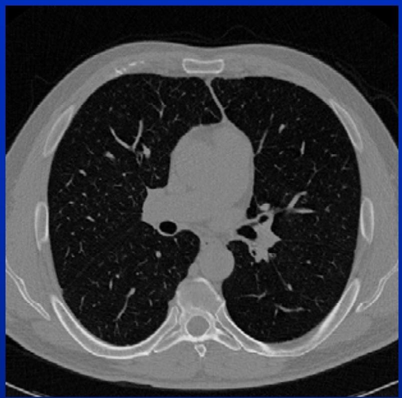
1. Segmentation: identification of the lung parenchyma

2. Identification of nodule candidates

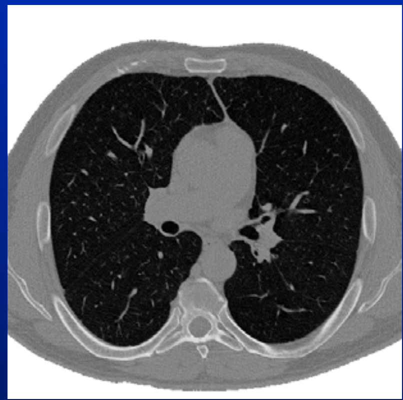
3. Reduction of false positive findings

1) Segmentation of the lung parenchyma

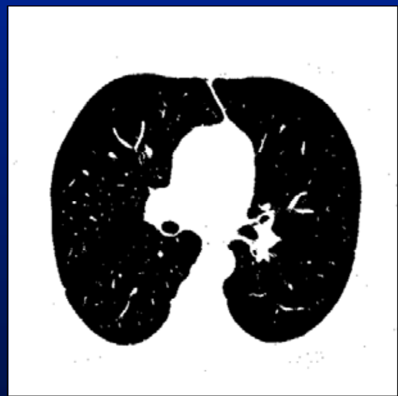
It is based on a combination of several image-processing techniques

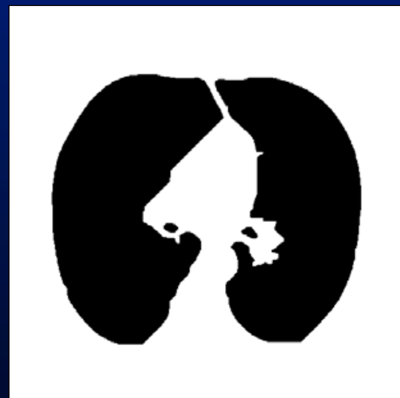


External background identification and removal
(grey-level thresholding outside the chest)



Lung identification (local/global
thresholding - binarization)





Improvement of lung border definition
morphological operators:
erosion (kernel 5) + dilation (kernel 5) + smoothing

2) Nodule candidate identification

3D isometric matrix

Modeling the lung structures:

- Nodules → spherical shapes
- Blood vessels and airway walls → elongated shapes
- Fissures → planar shape

$$d(x,y,z) = \exp\left\{ -\frac{x^2+y^2+z^2}{2\sigma^2} \right\}$$

$$l(x,y,z) = \exp\left\{ -\frac{x^2+y^2}{2\sigma^2} \right\}$$

$$p(x,y,z) = \exp\left\{ -\frac{x^2}{2\sigma^2} \right\}$$

[Q. Li, S. Sone and K. Doi, Med. Phys. 30 (8) 2003]

σ_1 .... σ_n — Convolution with 3D gaussians

$\lambda_{11}, \lambda_{12}, \lambda_{13}$ .... $\lambda_{n1}, \lambda_{n2}, \lambda_{n3}$ — Eigenvalues of the Hessian matrix

ζ_1 .... ζ_n — Likelihood function

$\zeta_i = |\lambda_{i3}|^2 / |\lambda_{i1}|$

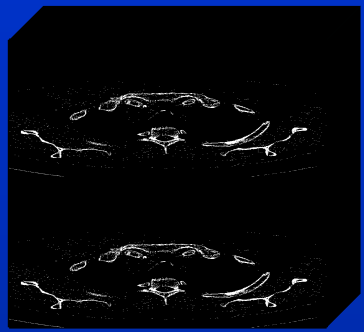
If $\lambda_{i1} < 0$, $\lambda_{i2} < 0$, $\lambda_{i3} < 0$

$\zeta_n = 0$

otherwise

$\zeta_{max} = \max [\zeta_1 \ldots\ldots \zeta_n]$ for each voxel

Nodule candidate identification



3D matrix of the ζ_{max}

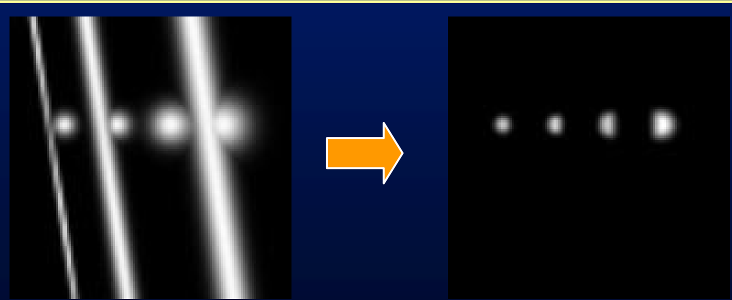
Peak detector

Finds the local maxima

ROIs

The peak list is sorted by value

Example of a synthetic image



Results on lung internal nodules

Nodule ID	Scan ID	Peak number
1,2	002	4,51
3	003	31
4	005	1
5-7	012	3,5,11
8	014	12
9	016	7
10	018	1
11	019	15
12	022	29
13-15	026	2,10,12
16-20	027	1,11,23,40,47
21,22	030	9,42
23	046	54
24	056	13

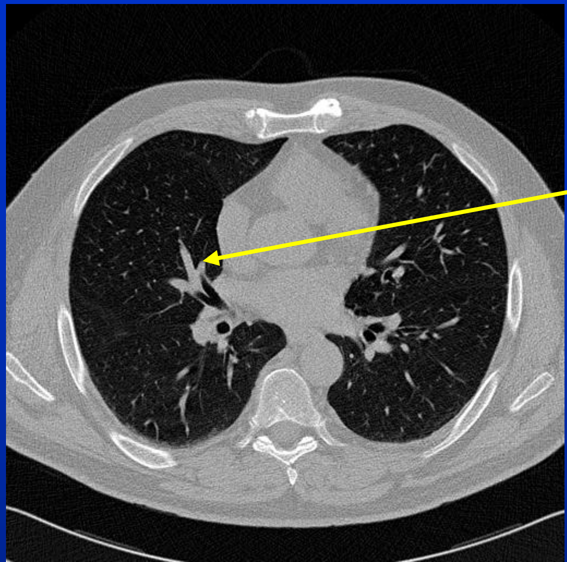
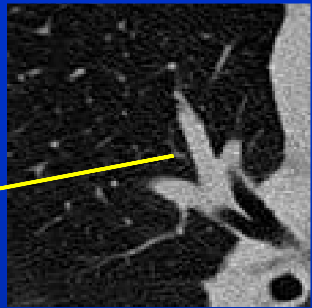
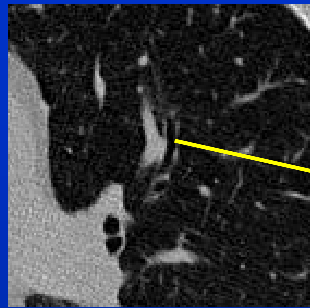
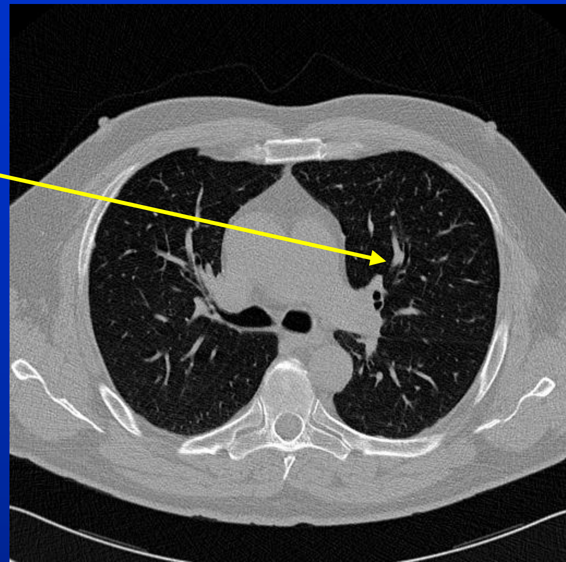
$2\sigma_1 = 4$mm; $2\sigma_n = 9$mm; n=5

If we take the first 54 findings for each scan, we obtain a sensitivity of 100%

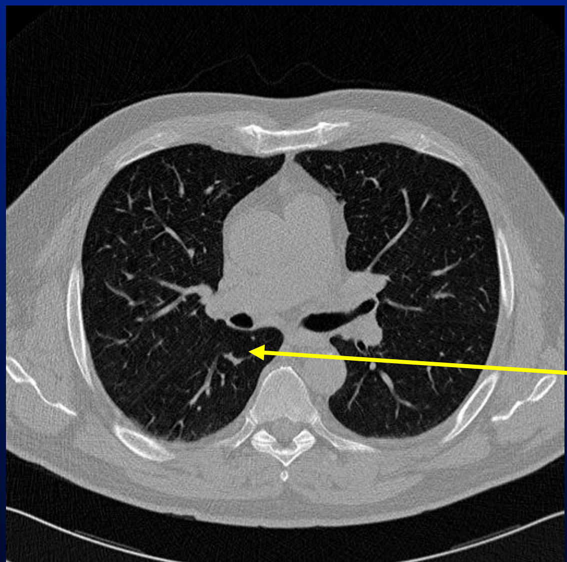
In other words:

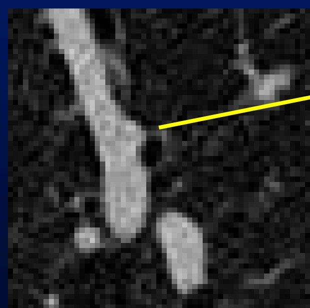
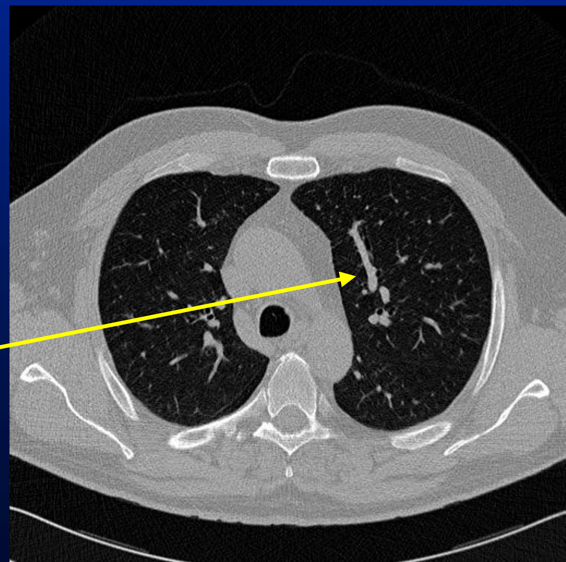
a 100% sensitivity is obtained at a maximum number of 54 false positive findings per scan (~0.2 FP per slice)

False positive findings

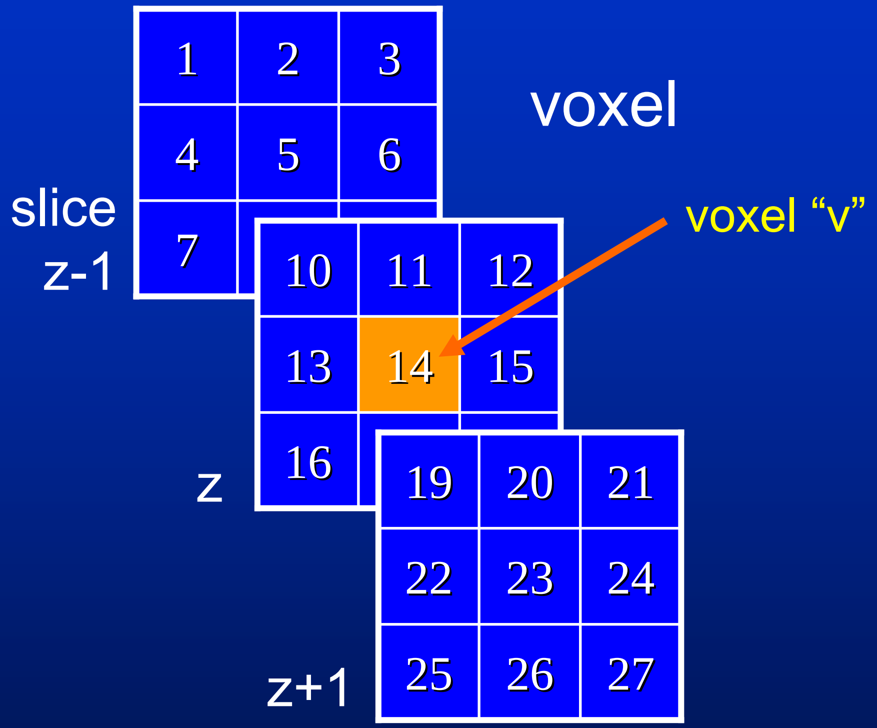





Most false positives are blood vessel crossings



3) FP reduction: voxel-based approach (VBA)



Voxel "v"

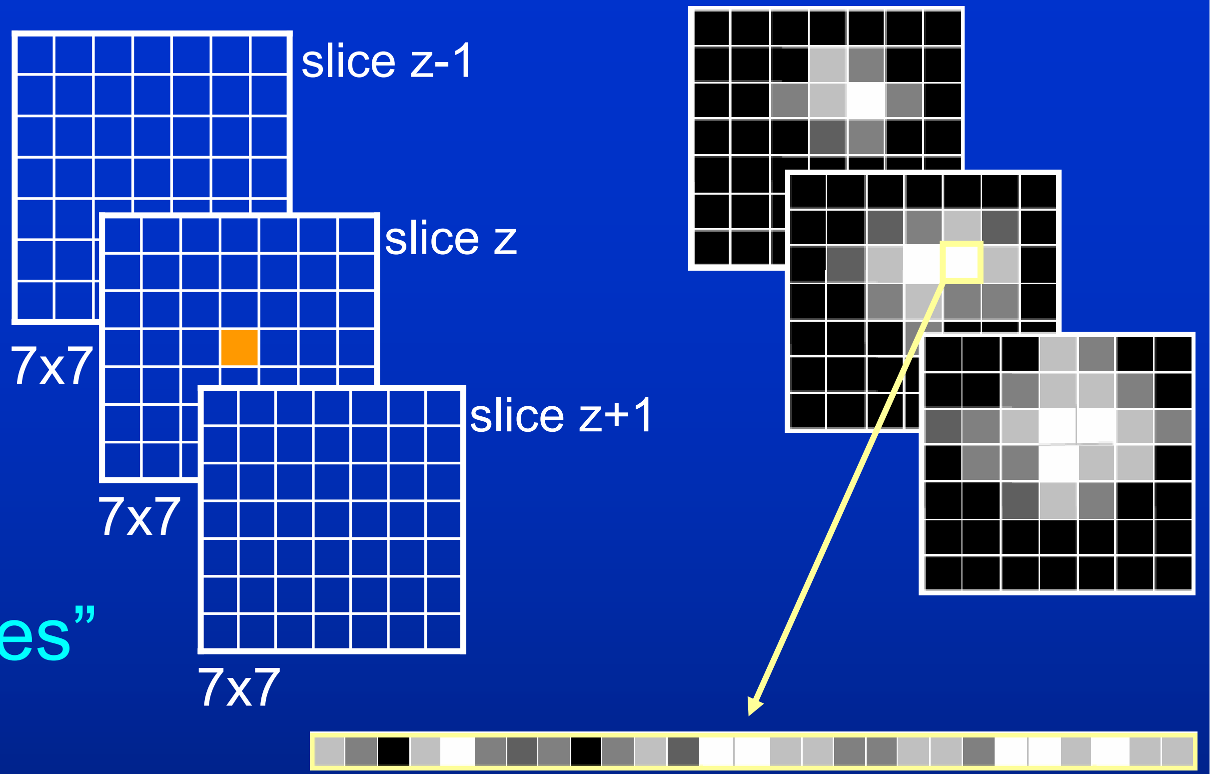
Vector of "features" (i.e. sequence of grey-level intensity values of the 3D neighborhood of "v")

Neural classification

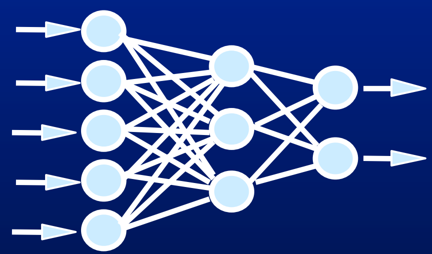
VBA

Neighborhood size:
- 7 pixels ~ 4 mm
- 3 slices = 3.75 mm

7x7x3 = 147 “features”



Feed-forward back-propagation neural network



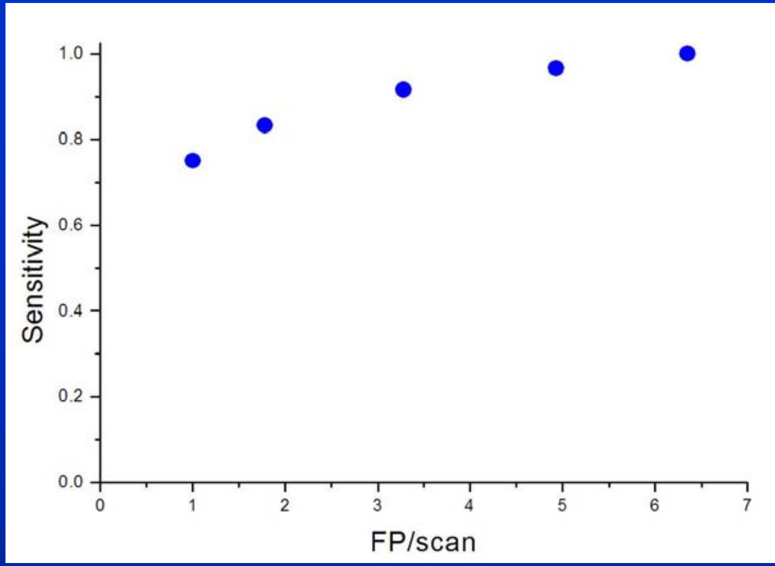
	Sensitivity (%)	Specificity (%)
1	85.12	84.46
2	86.95	85.28
3	82.5	84.9
4	83.8	81.4
5	80.3	82.8

5x2 cross-validation method

	on test set (TS)		on train set (TR)		on TR+TS	
	sens	spec	sens	spec	sens	spec
1	85.12	84.46	80.76	88.38	82.41	86.45
2	86.95	85.28	88.61	84.11	87.96	84.89

Results

- Sensitivity =100% @ 6.4 FP/scan
  Sensitivity =75% @ 1 FP/scan
  (57 patients)



Comparison with commercial systems:

- The ImageChecker® CT Lung system (R2 technology)
  [Roberts, CARS'05, pp 1137-42]
  - is the first clinically validated CAD system for chest CT
  - Sensitivity=73% 3 FP/scan (250 patients)
- Prototype of LungCAD CT (Siemens) [Wolf, CARS'05, pp 1143-5]
  - Sensitivity=77.1% 2.7 FP/scan (185 patients)

Conclusions

- Computer-aided detection of lung nodules:
  - The dot-enhancement pre-processing algorithm has a good sensitivity in the identification of nodule candidates
  - The VBA for FP reduction is a “trivial” but effective approach

<u>These results are preliminary and need to be validated on a more populated database</u>

Thanks for your attention!

This work has been carried out in the framework of the INFN (Istituto Nazionale di Fisica Nucleare)-funded MAGIC-5 (Medical Applications in a GRID Infrastructure Connection) project.

We would like to acknowledge:

Dr. L. Battolla, Dr. F. Falaschi and Dr. C. Spinelli of the “Unità Operativa Radiodiagnostica 2” of the “Azienda Ospedaliera Universitaria Pisana” and Prof. D. Caramella and Dr. T. Tarantino of the Diagnostic and Interventional Radiology Division of the “Dipartimento di Oncologia, Trapianti e Nuove Tecnologie in Medicina” of the Pisa University who have made available the full annotated CT scans.

E-mail: retico@df.unipi.it