Computer-aided diagnosis of subtle signs of breast cancer: Architectural distortion in prior mammograms

Rangaraj M. Rangayyan

Department of Electrical and Computer Engineering
University of Calgary, Calgary, Alberta, CANADA

UNIVERSITY OF CALGARY

MO SHÙILE TOGAM SÙAS

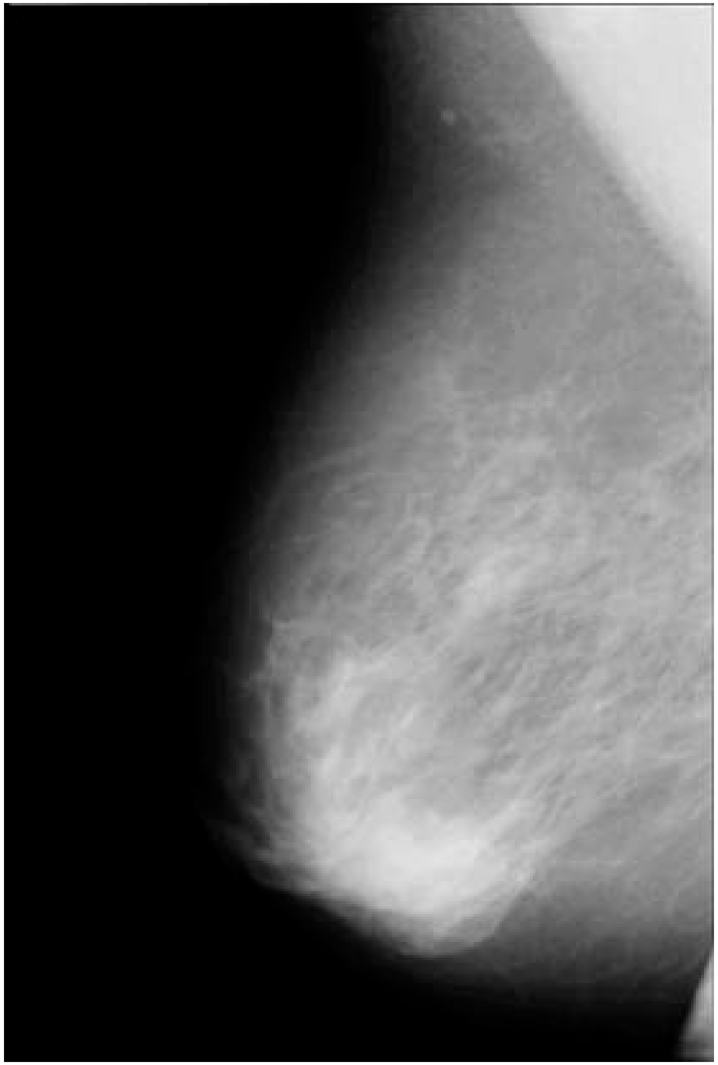
screentest
ALBERTA PROGRAM FOR THE EARLY DETECTION OF BREAST CANCER

SCHULICH School of Engineering

UNIVERSITY OF CALGARY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Mammography

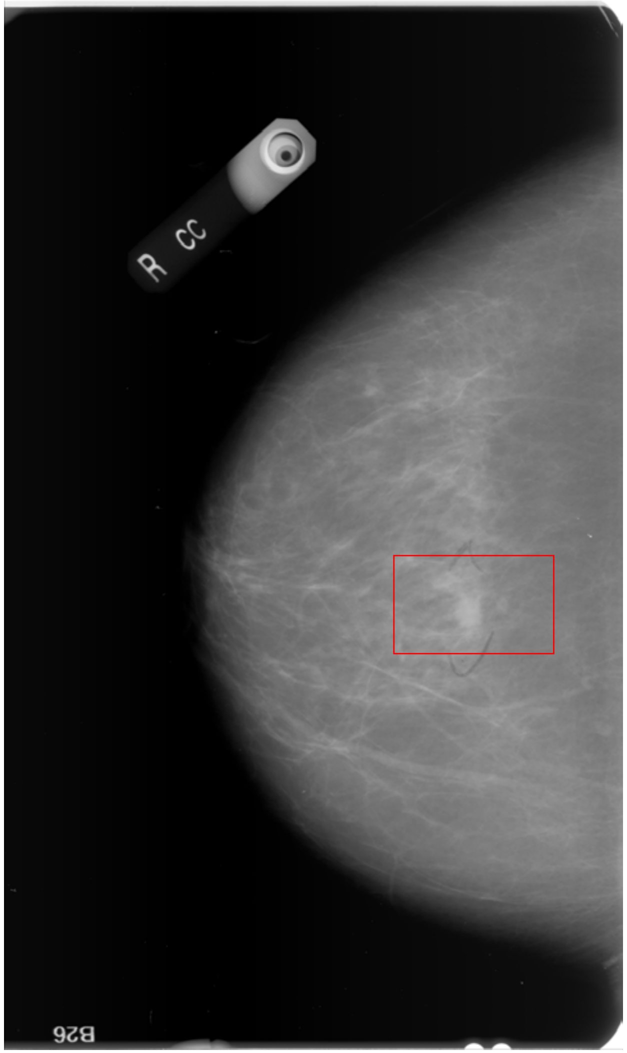


Signs of Breast Cancer:

- Masses
- Calcifications
- Bilateral asymmetry
- Architectural distortion (often missed)

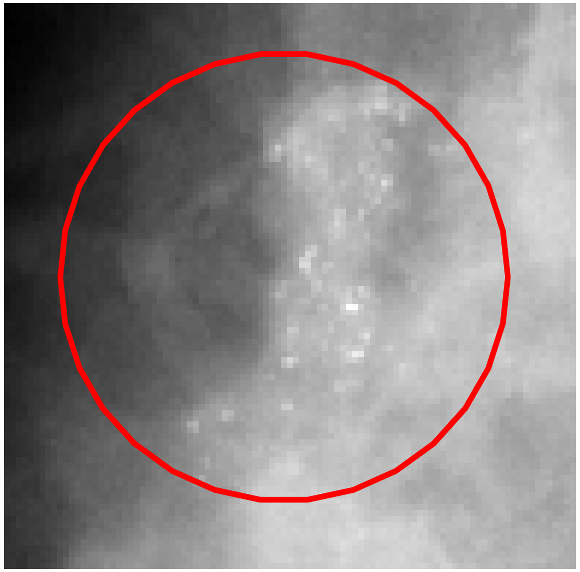
Masses

- Breast cancer causes a desmoplastic reaction in breast tissue
- A mass is observed as a bright, hyperdense object

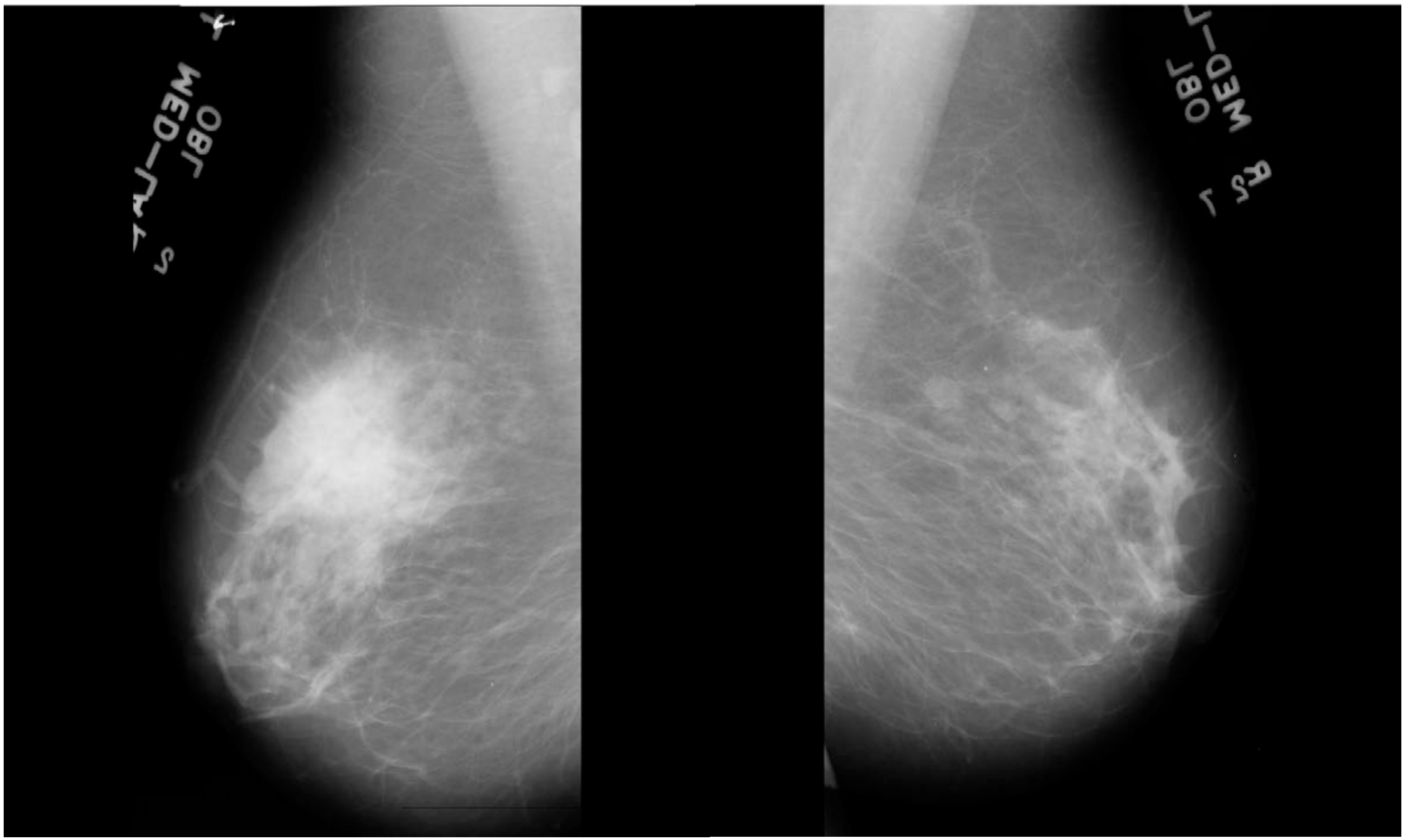


Calcification

Deposits of calcium in breast tissue



Bilateral asymmetry



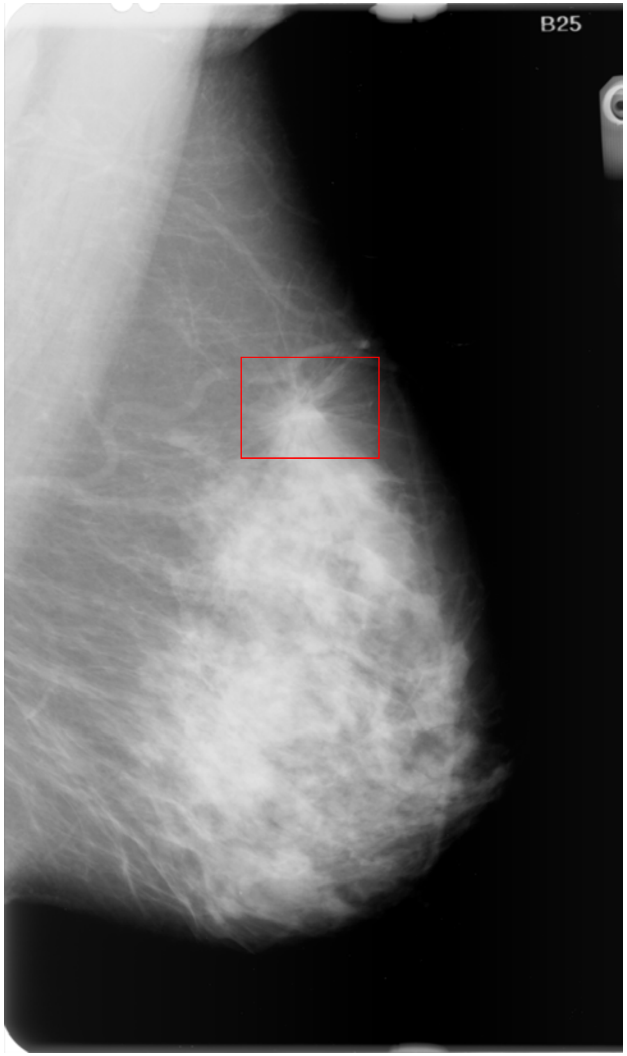
Differences in the overall density distribution in the two breasts

Computer-aided diagnosis

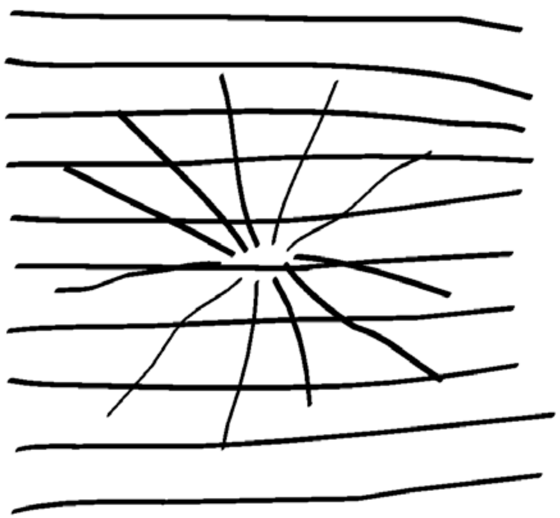
- Increased number of cancers detected
- Increased early-stage malignancies detected
- Increased recall rate
- Missed cases of architectural distortion

Architectural distortion

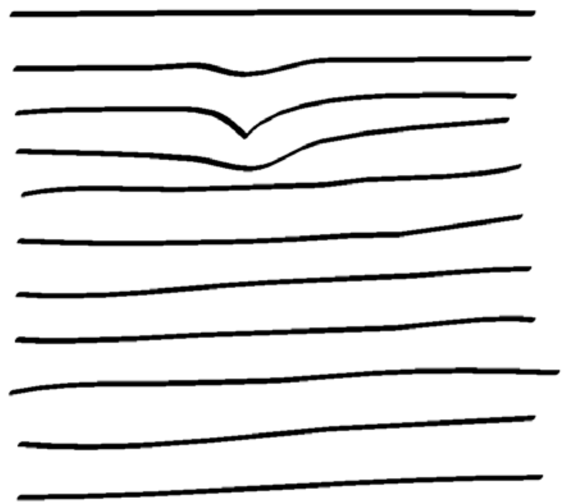
- Third most common mammographic sign of nonpalpable breast cancer
- The normal architecture of the breast is distorted
- No definite mass visible
- Spiculations radiating from a point
- Focal retraction or distortion at the edge of the parenchyma



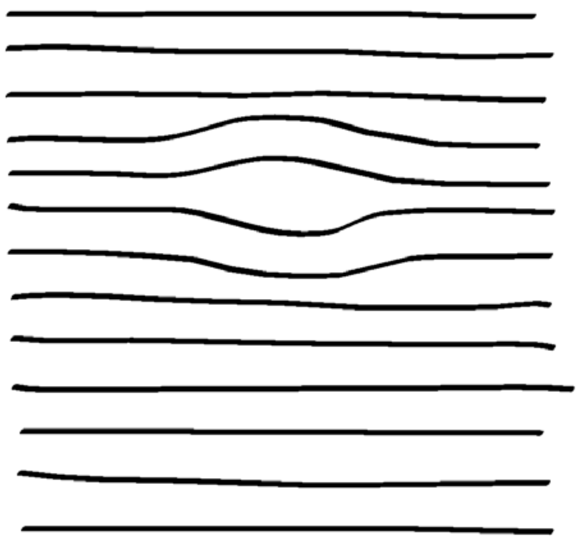
Architectural distortion



spiculated

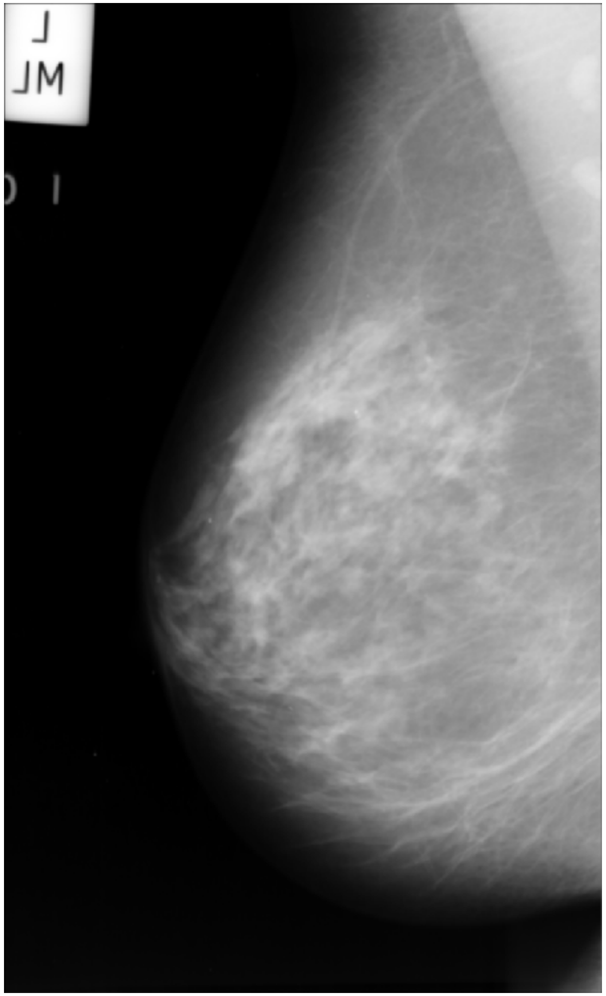


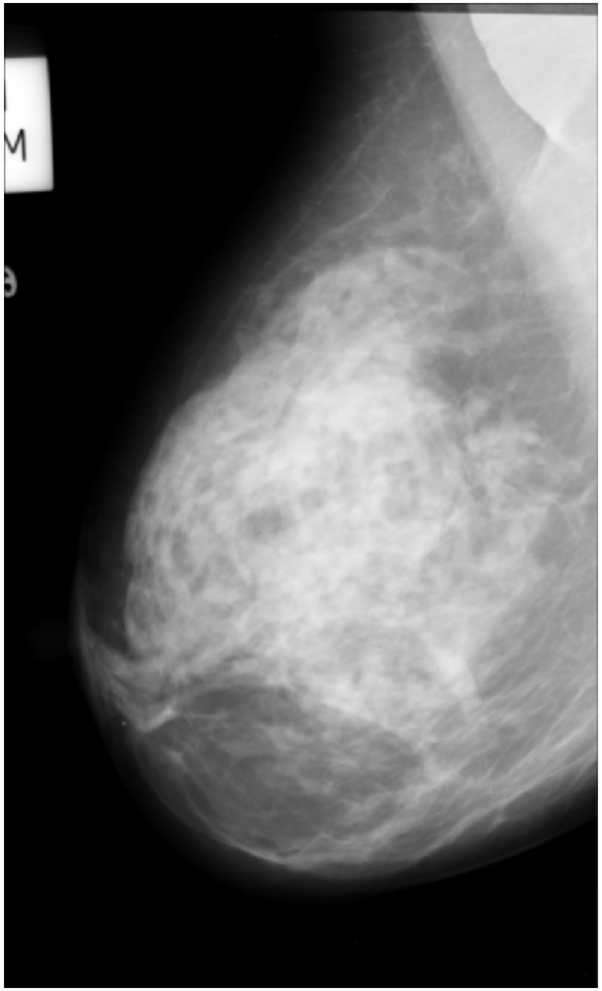
focal retraction



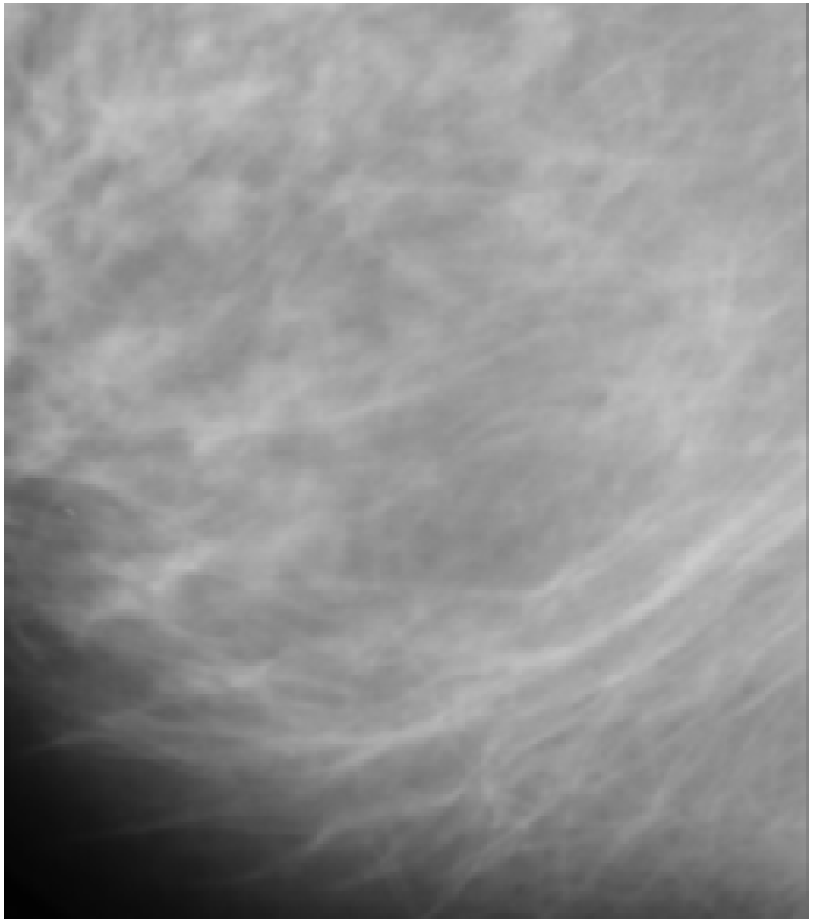
incipient mass

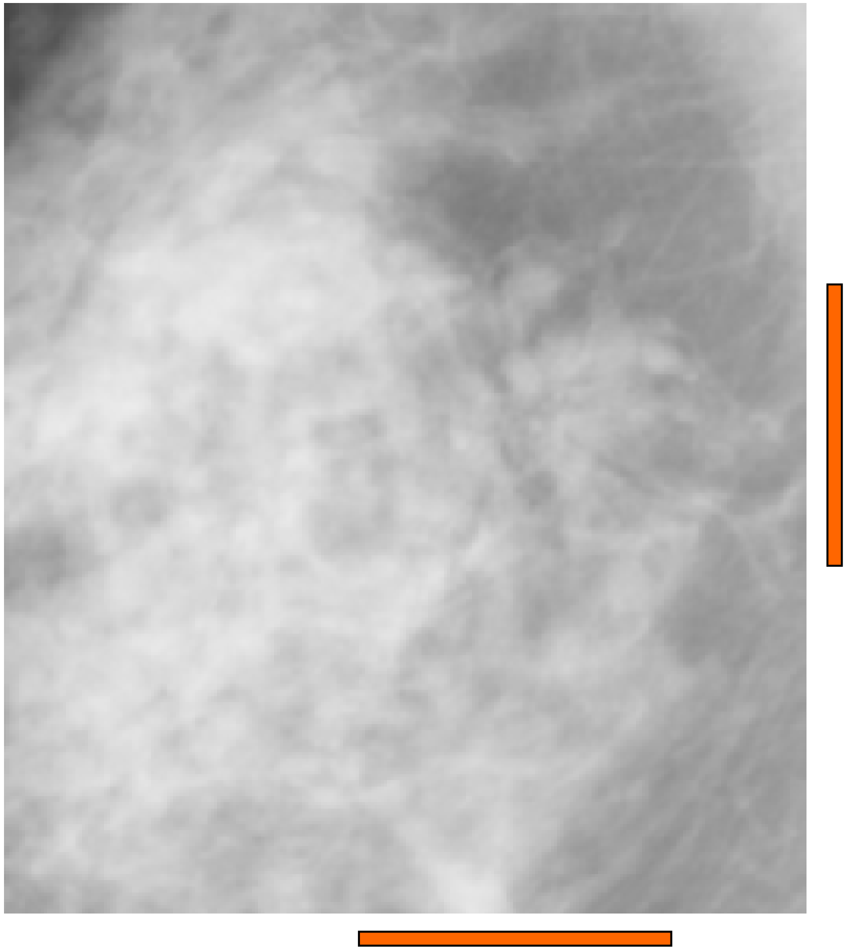
Normal vs architectural distortion





Normal vs architectural distortion





Initial algorithm for detection of architectural distortion

1. Extract the orientation field
2. Filter and downsample the orientation field
3. Analyze orientation field using phase portraits
4. Postprocess the phase portrait maps
5. Detect sites of architectural distortion

Gabor filter

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]\cos(2\pi f x)$$

Design parameters

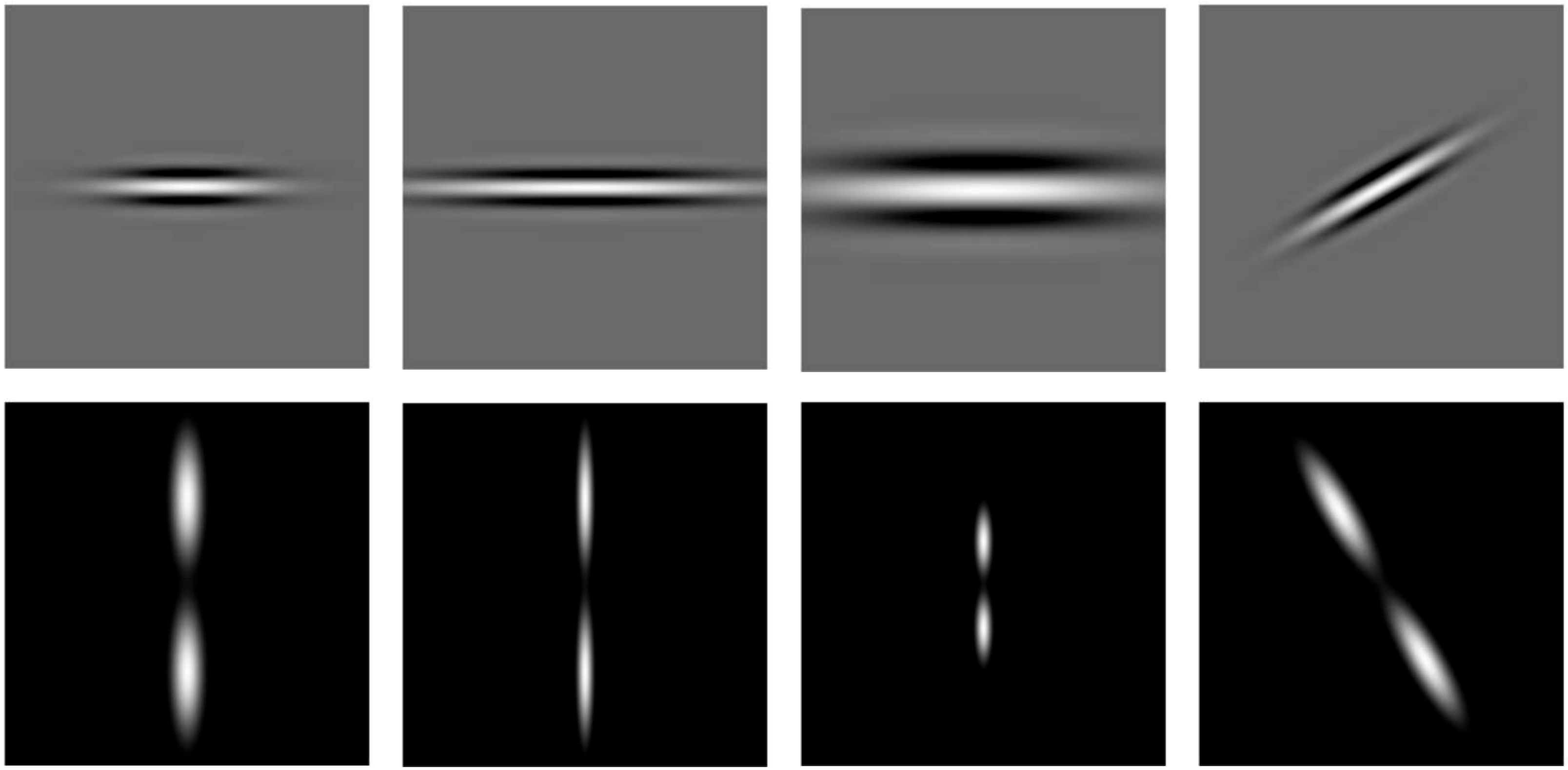
- line thickness τ
- elongation l
- orientation θ

Gabor parameters

$$f = \frac{1}{\tau}; \qquad \sigma_x = \frac{\tau}{2\sqrt{2\ln 2}}$$

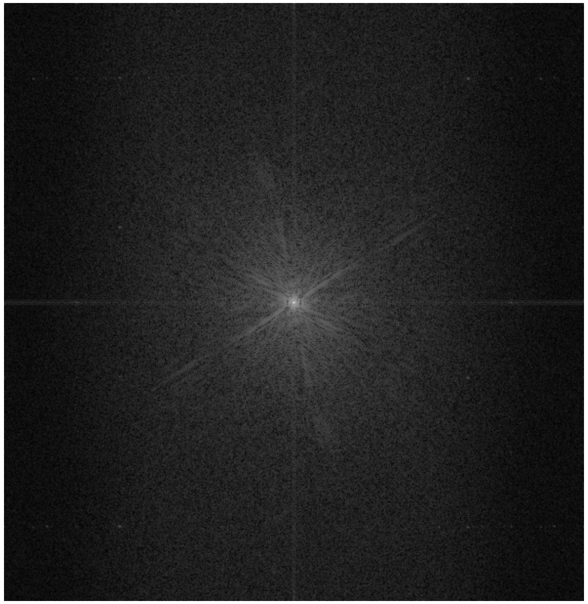
$$\sigma_y = l\sigma_x; \qquad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Design of Gabor filters

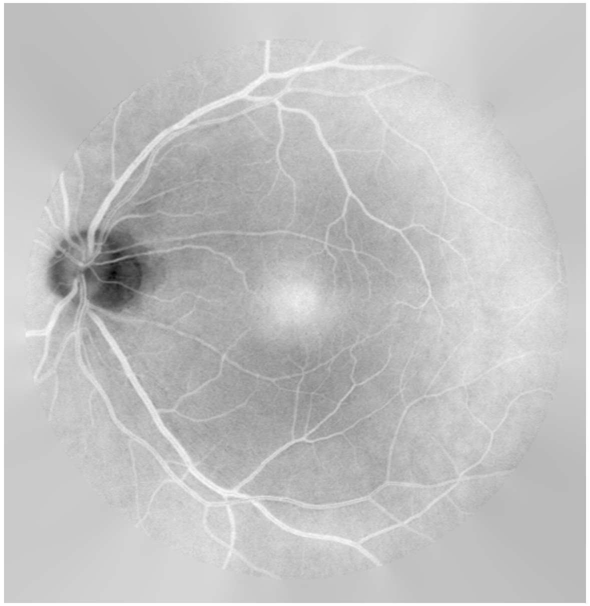


$l = l_0$	$l > l_0$	$l = l_0$	$l = l_0$
$\tau = \tau_0$	$\tau = \tau_0$	$\tau > \tau_0$	$\tau = \tau_0$
$\theta = \theta_0$	$\theta = \theta_0$	$\theta = \theta_0$	$\theta > \theta_0$

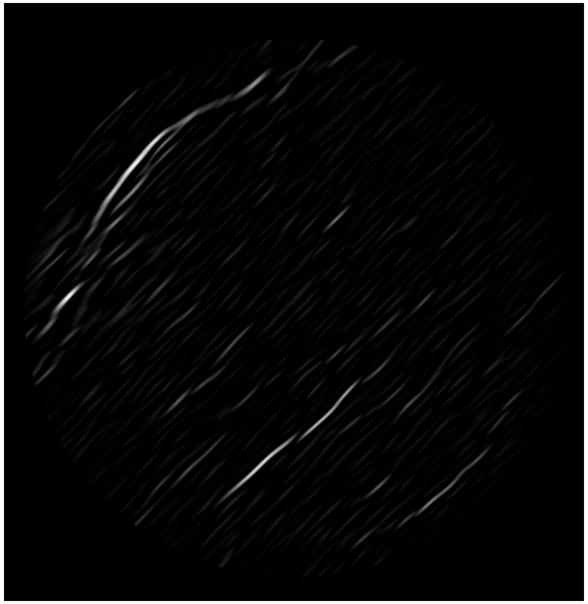
Example of Gabor filtering



Log-magnitude Fourier spectrum



Inverted Y channel of retinal fundus image

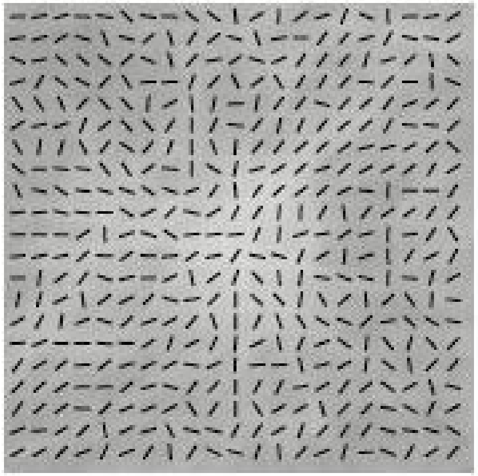


Magnitude response of a single Gabor filter:
$\tau = 8,\ l = 2.9,\ \theta = 45^{o}$

Extracting the orientation field

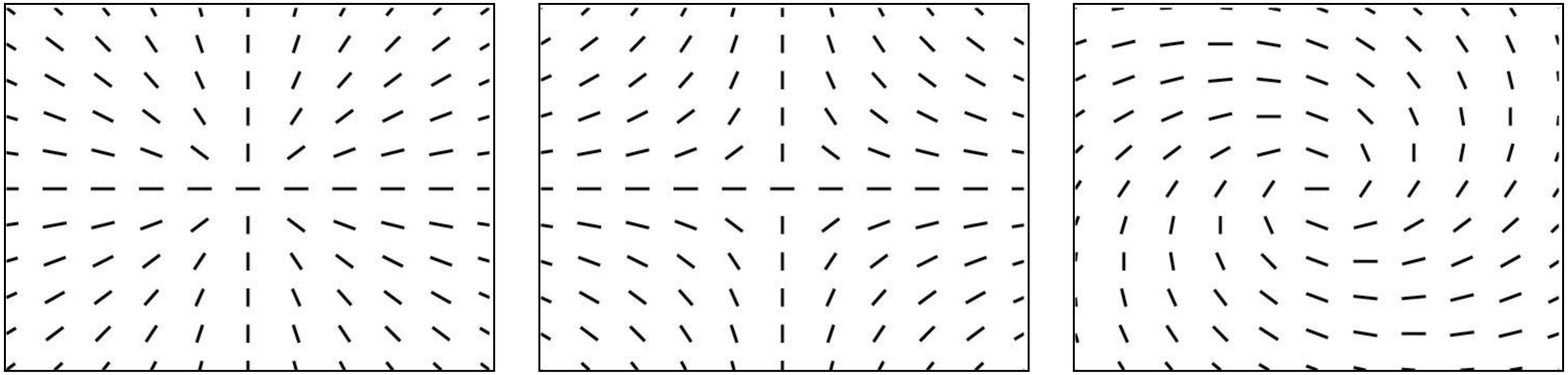
Compute the texture orientation (angle) at each pixel

Gabor filtering
(line detection)



Phase portraits

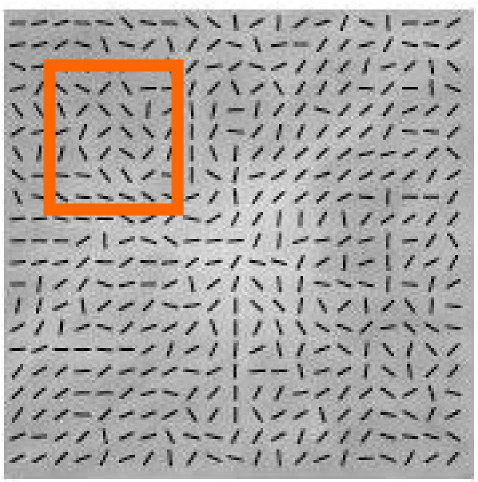
$$\vec{\mathbf{v}}(x, y) = \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{b}$$



node saddle spiral

Texture analysis using phase portraits

Fit phase portrait model to the analysis window



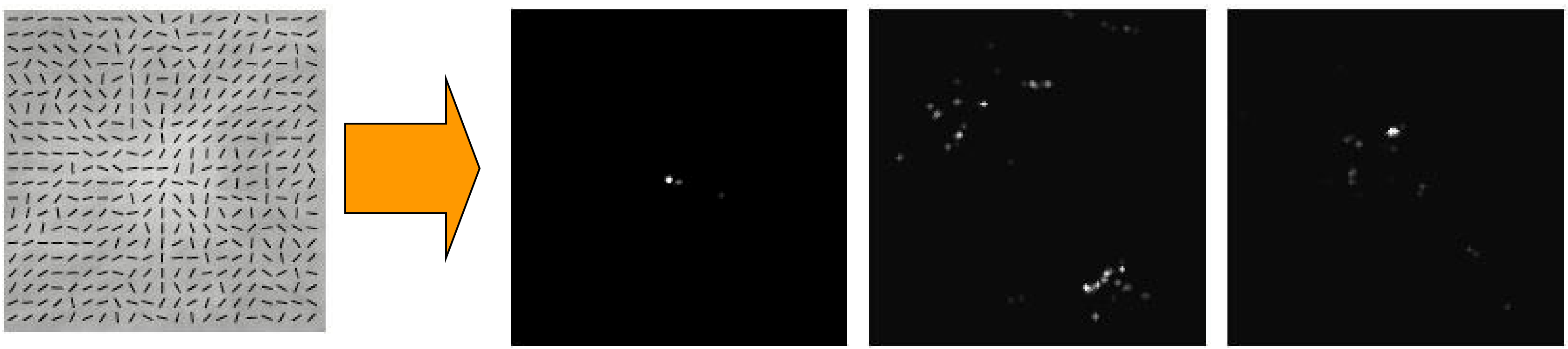
Nonlinear least squares optimization

$$\mathbf{A} = \begin{bmatrix} 1.1 & 0.3 \\ -0.2 & 1.7 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} -4.8 \\ -7.9 \end{bmatrix}$$

Texture analysis using phase portraits

Cast a vote at the fixed point = $\mathbf{A}^{-1}\,\mathbf{b}$ in the corresponding phase portrait map

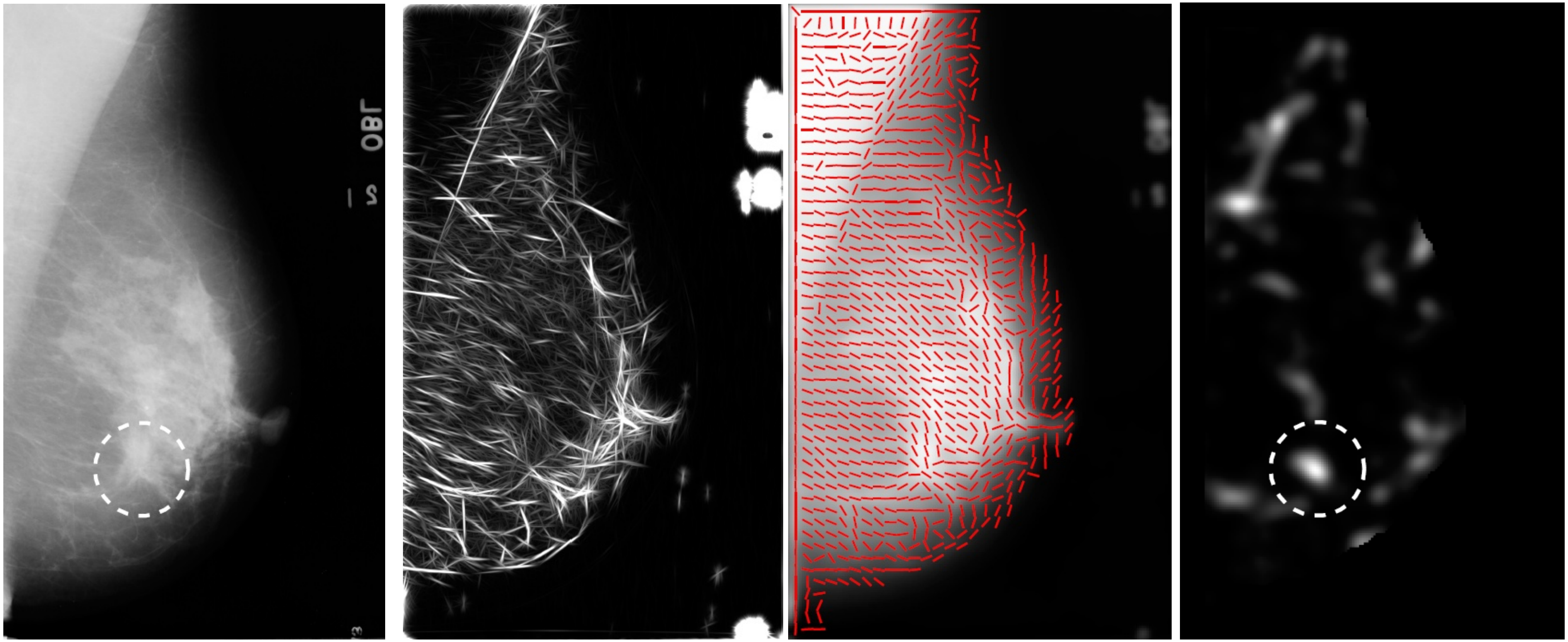


Orientation field — Node — Saddle — Spiral

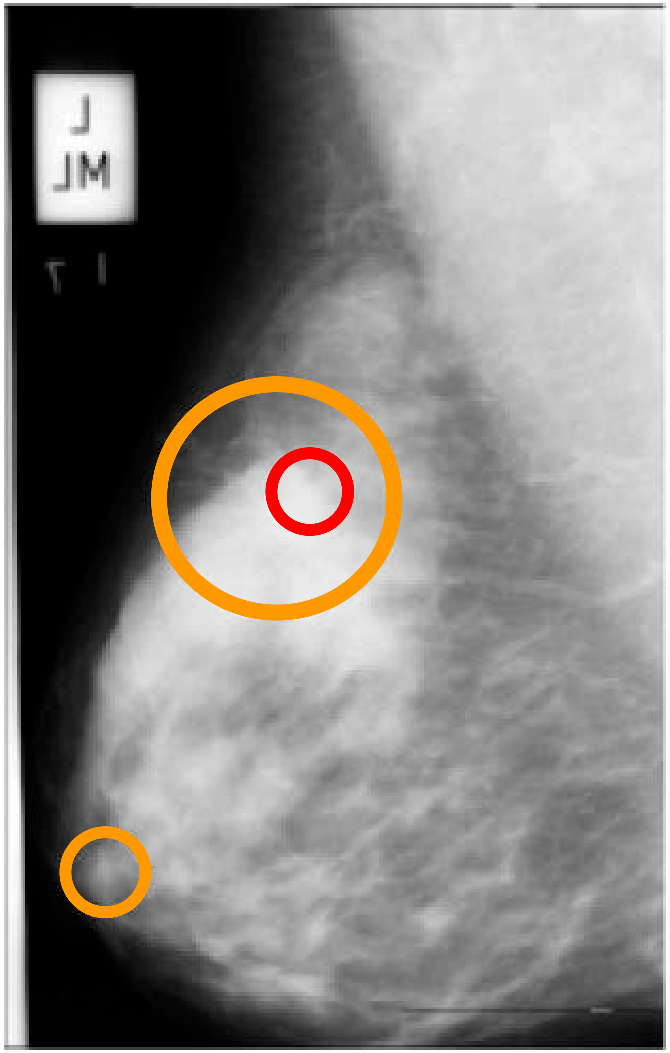
real eigenvalues of same sign

UNIVERSITY OF CALGARY

Detection of architectural distortion



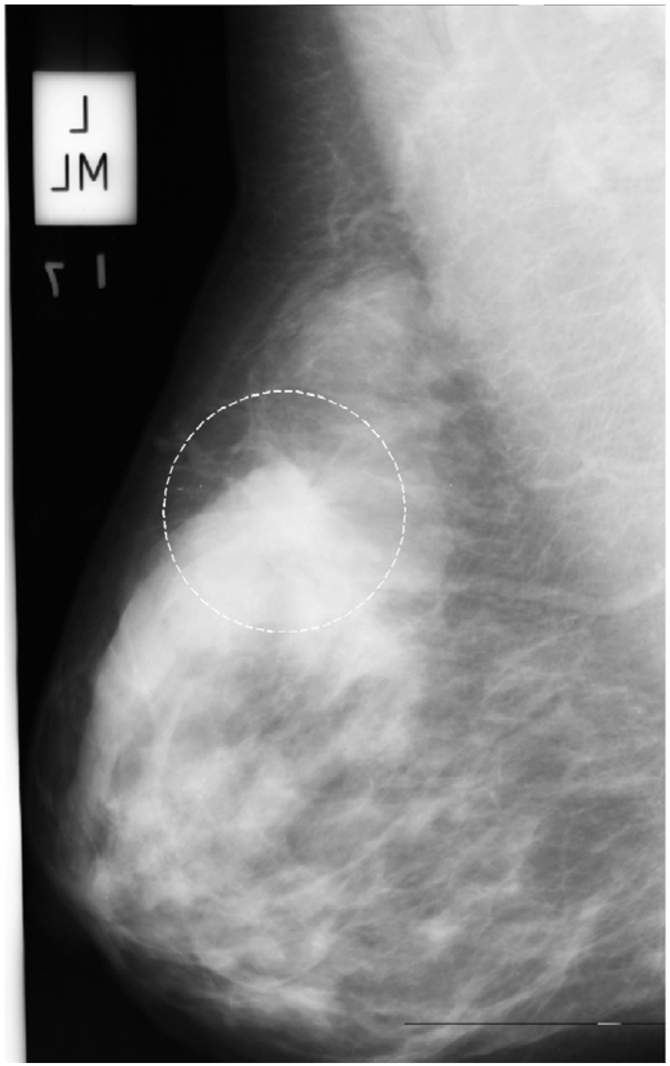
Initial results of detection

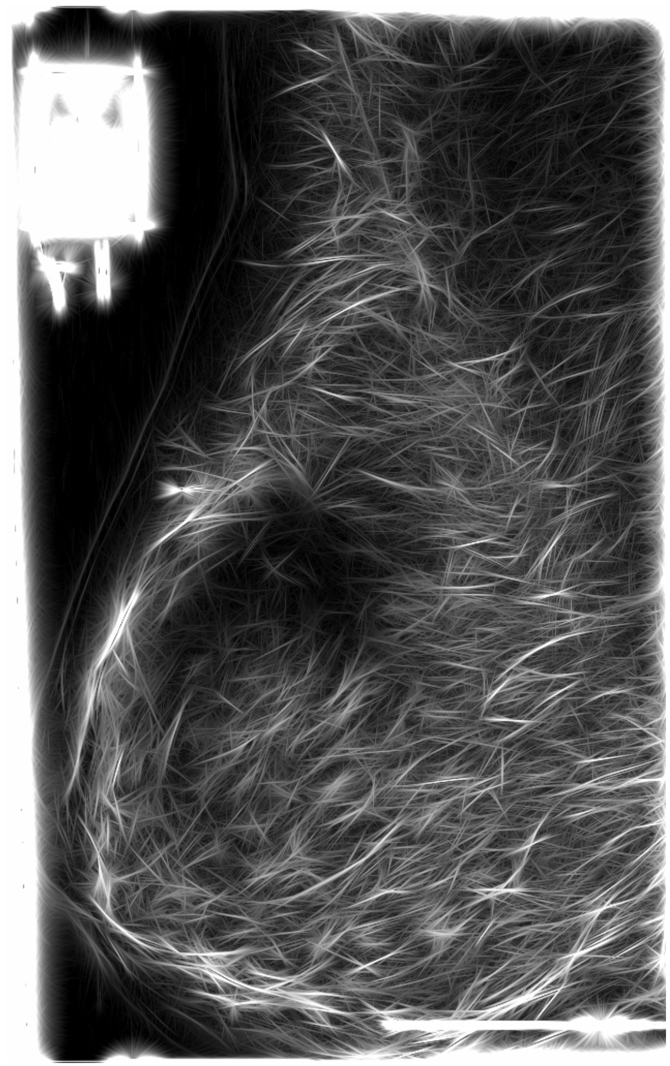


- Test dataset: 19 mammograms with architectural distortion (MIAS database)
- Sensitivity: 84%
- 18 false positives per image!

UNIVERSITY OF CALGARY

Reduction of false positives





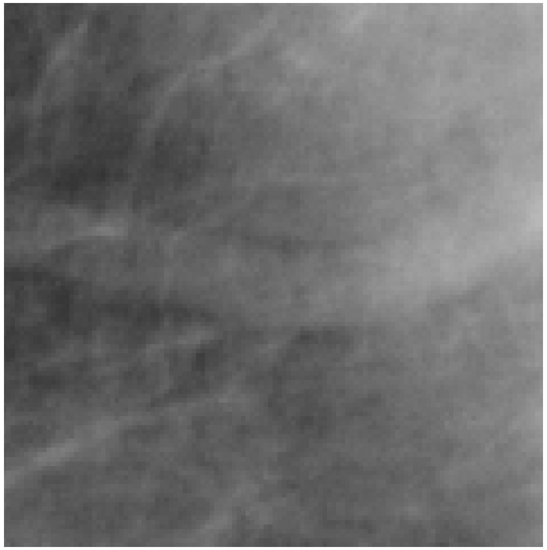
UNIVERSITY OF CALGARY

Rejection of confounding structures

- Confounding structures include
  - Edges of vessels
  - Intersections of vessels
  - Edge of the pectoral muscle
  - Edge of the fibroglandular disk

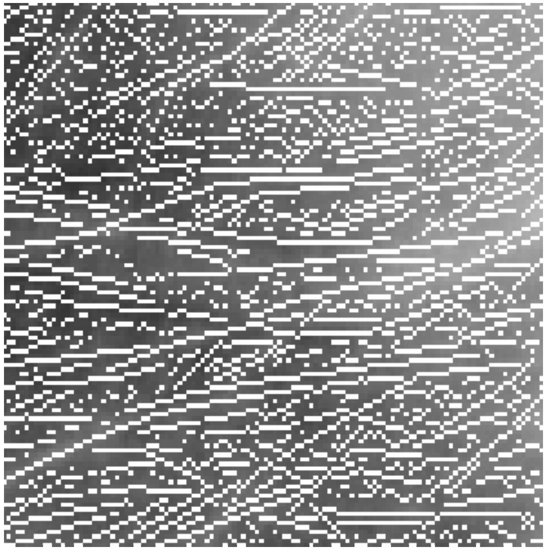
“Curvilinear Structures”

Nonmaximal suppression



ROI with a vessel

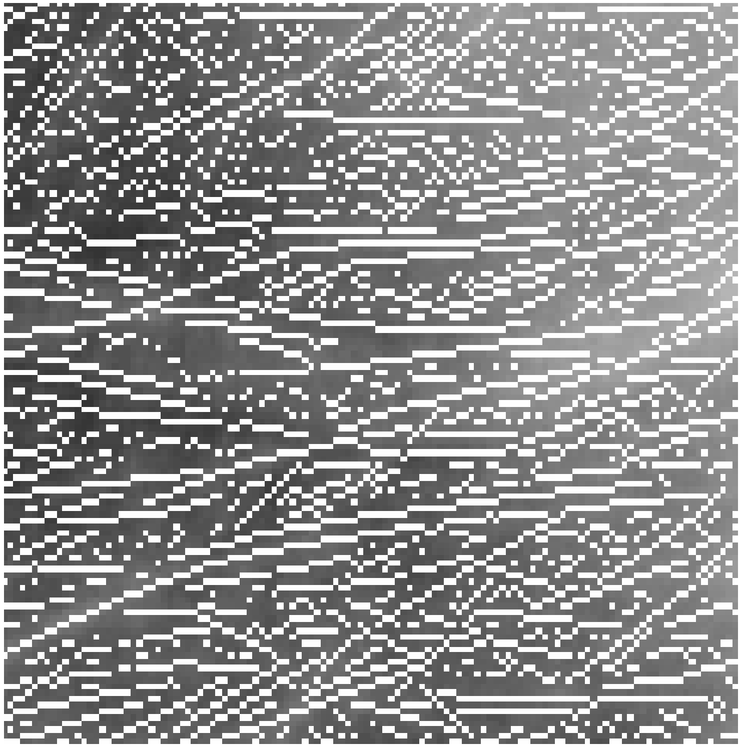
Gabor magnitude output



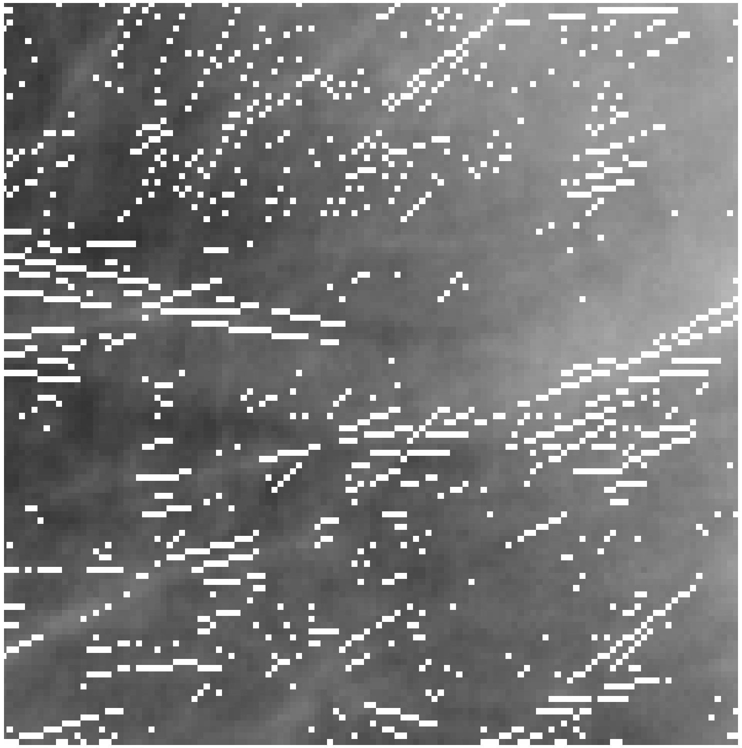
Output of nonmaximal suppression (NMS)

Rejection of confounding CLS

Output of NMS

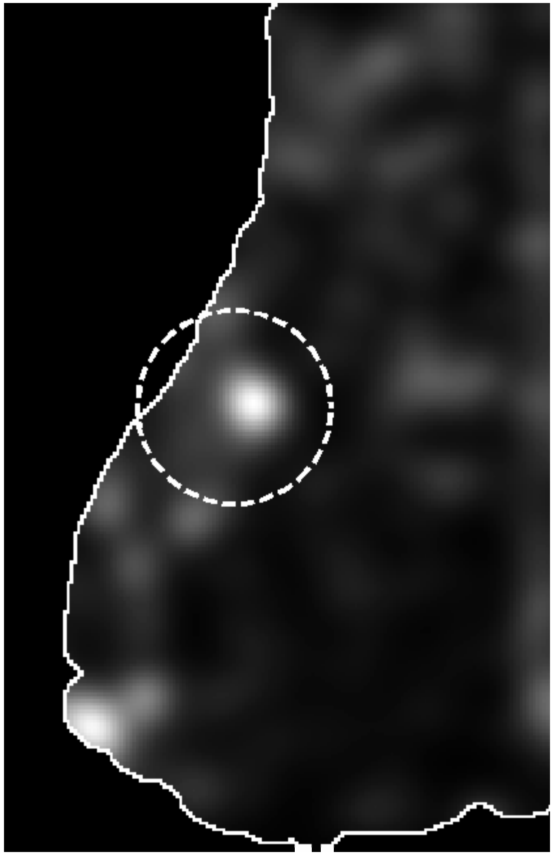


CLS Retained

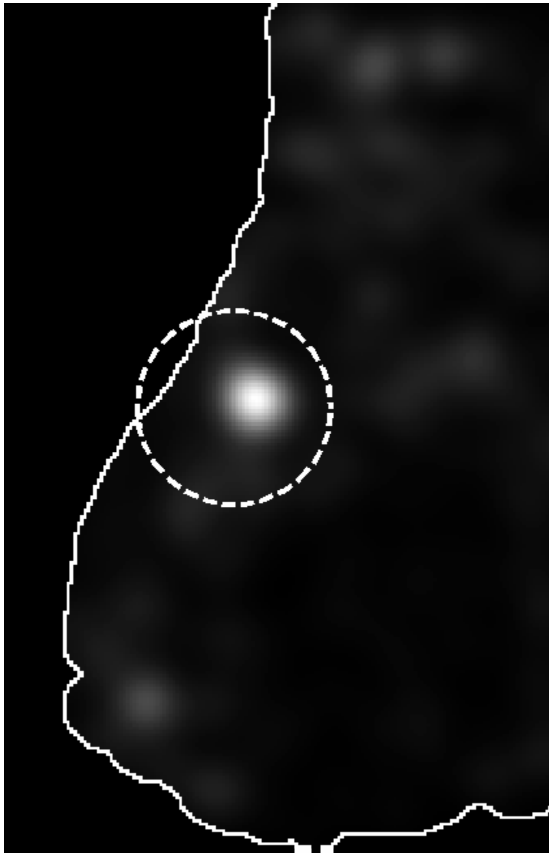


Angle from the orientation field and direction perpendicular to the gradient vector differ by < 30º

Improved detection of sites of architectural distortion

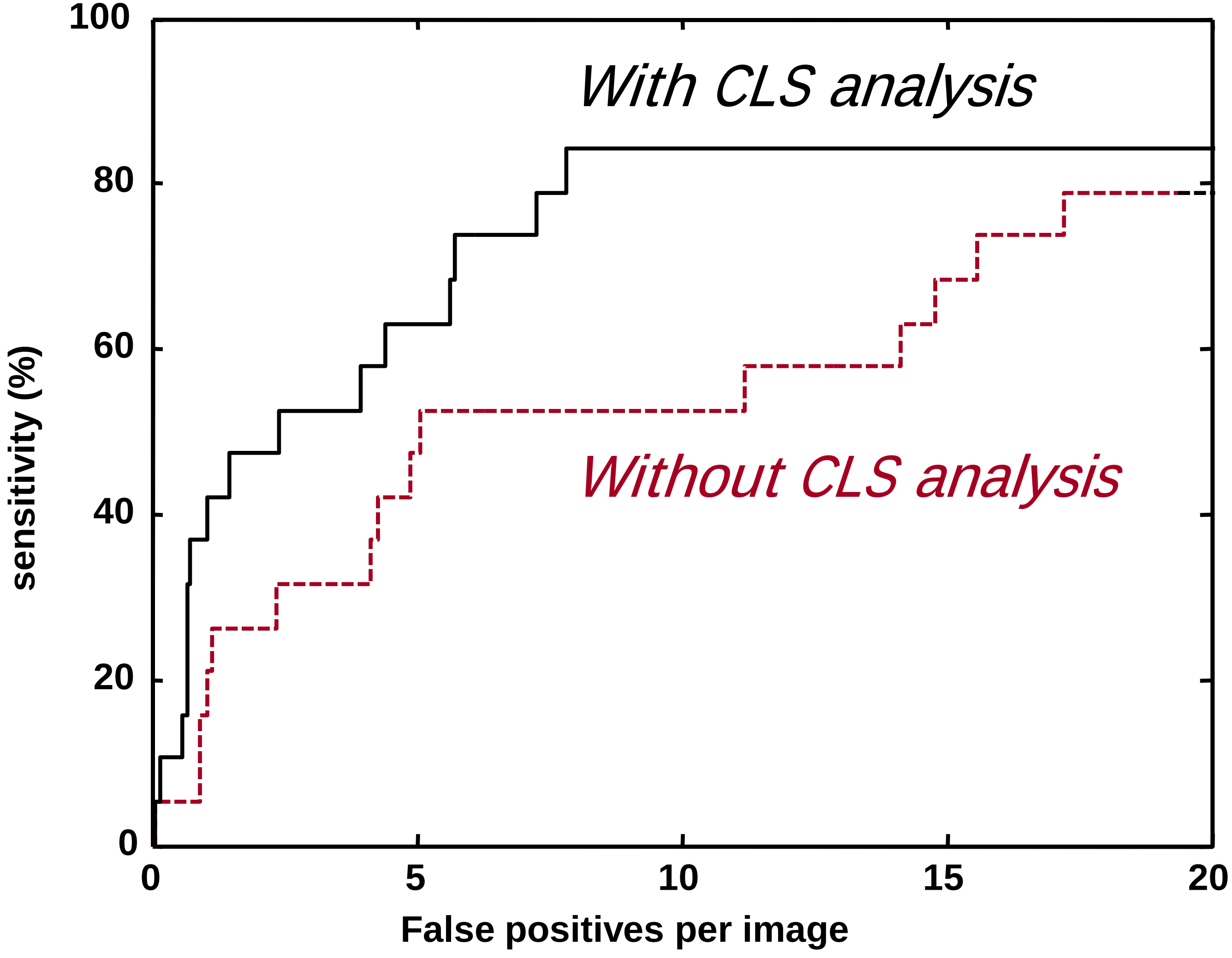


Node map
(without CLS analysis)

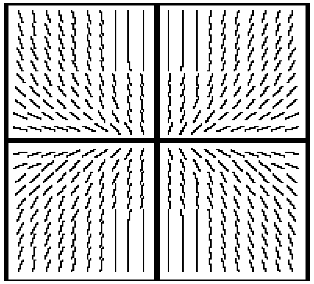
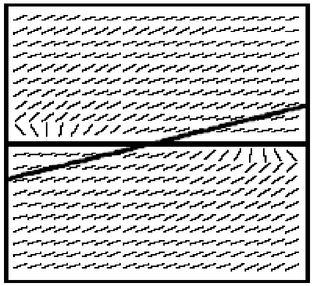
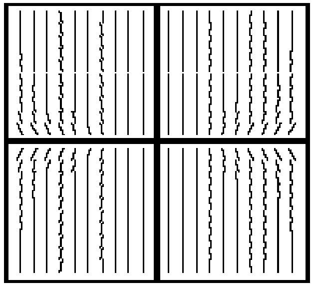


Node map
(with CLS analysis)

Free-response ROC analysis



Effect of condition number of matrix *A* on the orientation field

Example	Matrix **A**	Eigenvalues	Angle between principal axes	Condition number	Orientation field
A	$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$	$\lambda_1 = 1$ $\lambda_2 = 3$	90°	3	
B	$\begin{bmatrix} 1 & 7.46 \\ 0 & 3 \end{bmatrix}$	$\lambda_1 = 1$ $\lambda_2 = 3$	15°	21.85	
C	$\begin{bmatrix} 1 & 0 \\ 0 & 20 \end{bmatrix}$	$\lambda_1 = 1$ $\lambda_2 = 20$	90°	20	

Condition Number: The ratio of the largest to smallest singular value of a matrix

Results

- 19 cases of architectural distortion
- 41 normal control mammograms (MIAS)

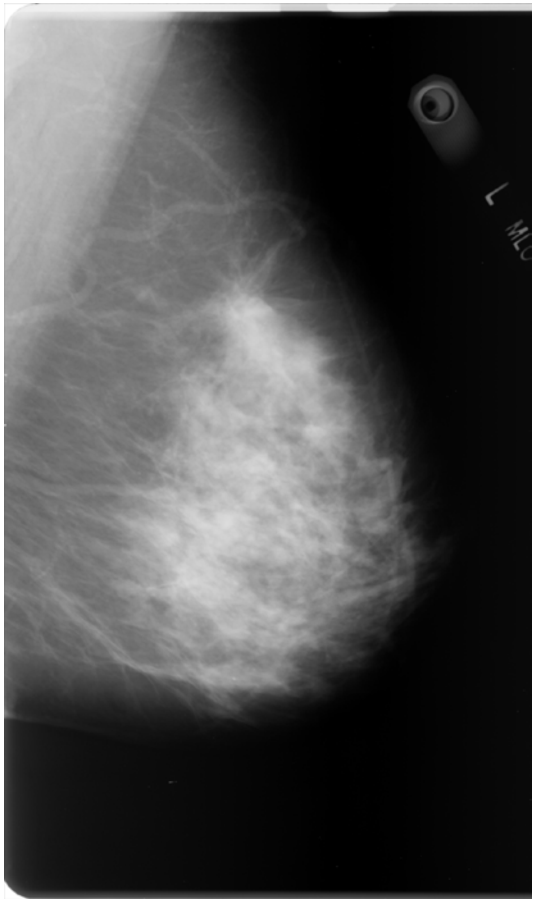
- Symmetric matrix A: node and saddle only
- Condition number of A > 3: reject result

- *Sensitivity: 84% at 4.5 false positives/image*

- *Sensitivity: 95% at 9.9 false positives/image*

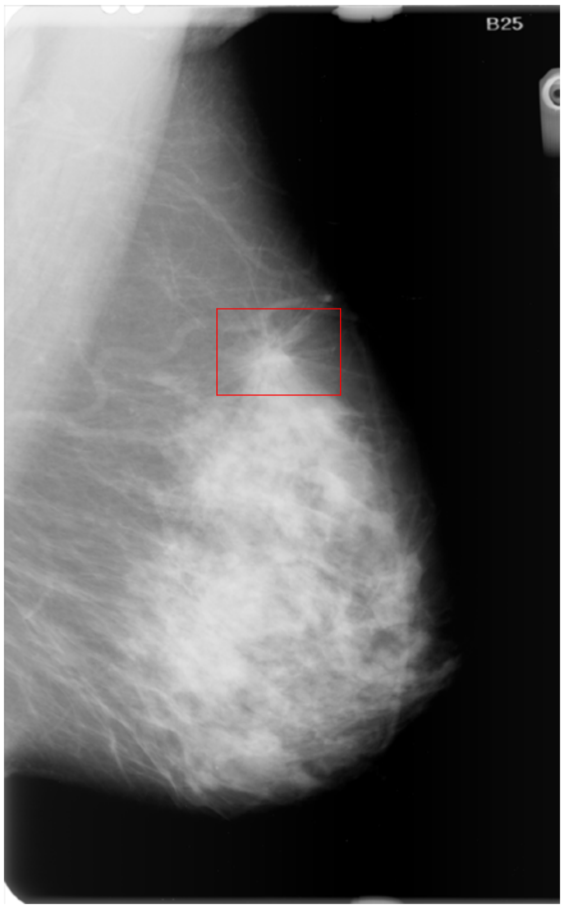
Prior mammograms

Detection mammogram 1997

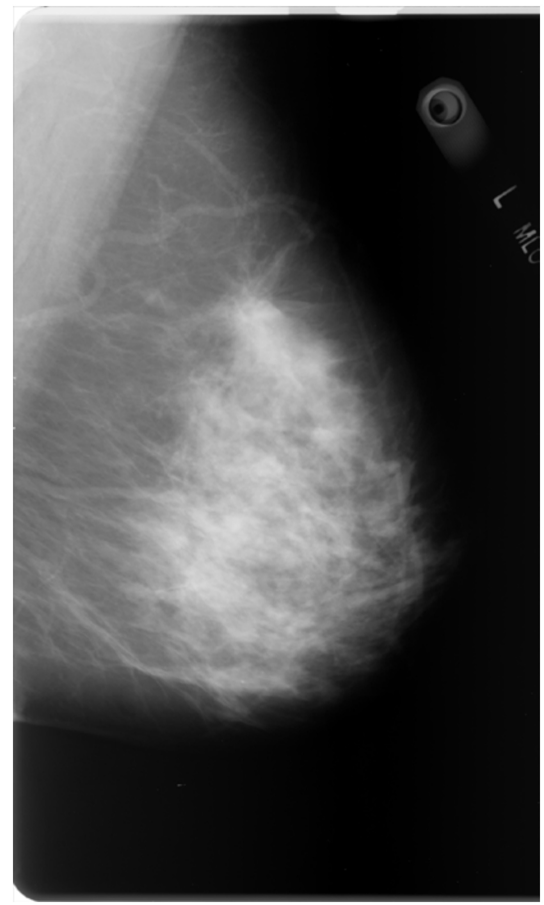


Prior mammogram 1996

Prior mammograms

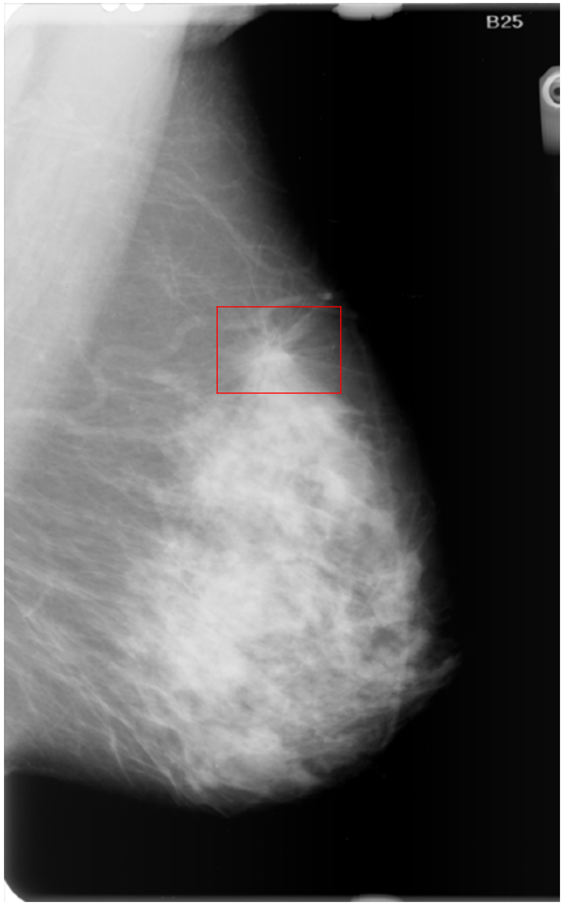


Detection mammogram 1997

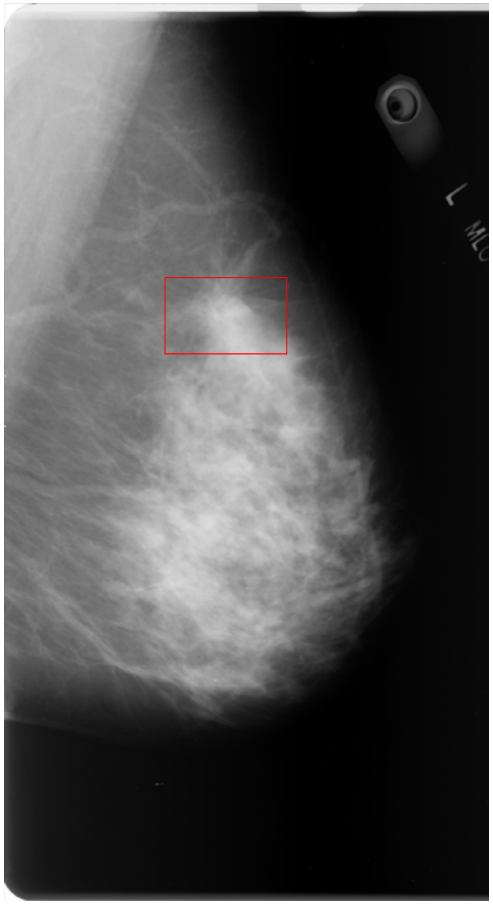


Prior mammogram 1996

Prior mammograms



Detection mammogram 1997



Prior mammogram 1996

Interval cancer

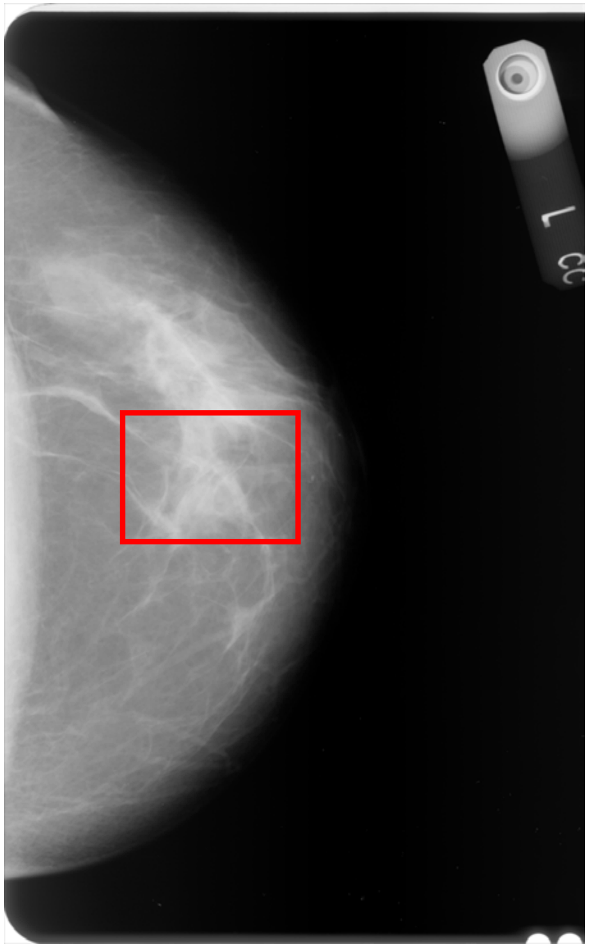
- Breast cancer detected outside the screening program in the interval between scheduled screening sessions
- "Diagnostic mammograms" not available

UNIVERSITY OF CALGARY

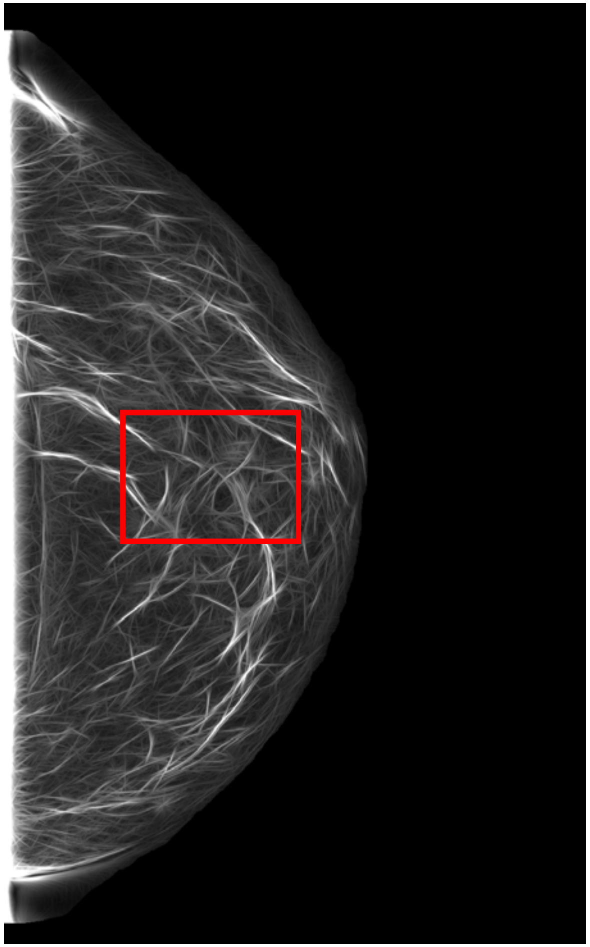
Dataset

- 106 prior mammographic images of 56 individuals diagnosed with breast cancer (interval-cancer cases)
- Time interval between prior and detection (33 cases)
  average: 15 months, standard deviation: 7 months
  minimum: 1 month, maximum: 24 months
- 52 mammographic images of 13 normal individuals
- Normal control cases selected represent the penultimate screening visits at the time of preparation of the database

Interval cancer: site of architectural distortion

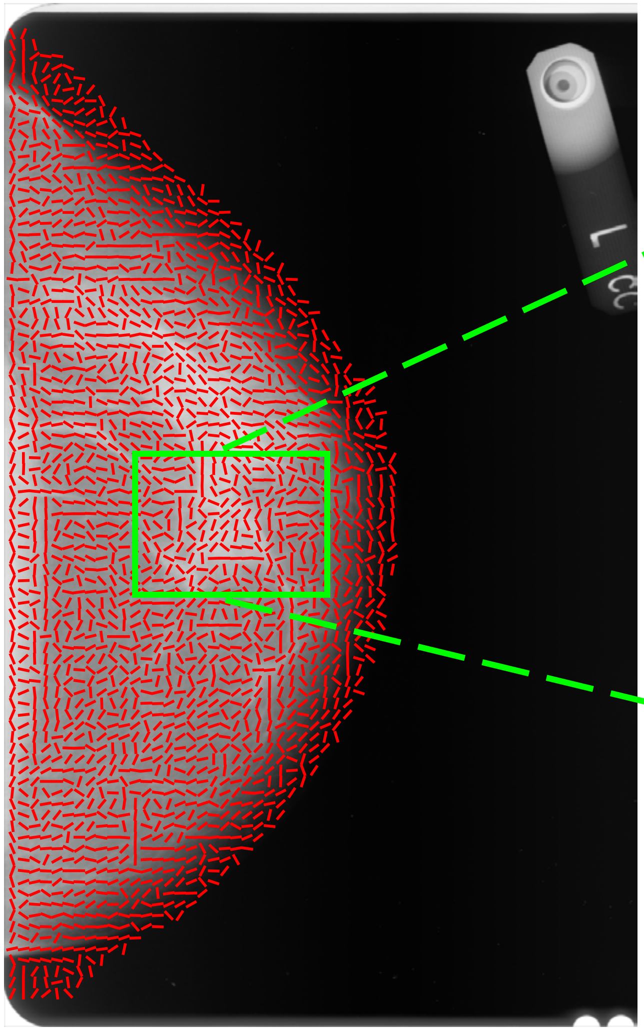


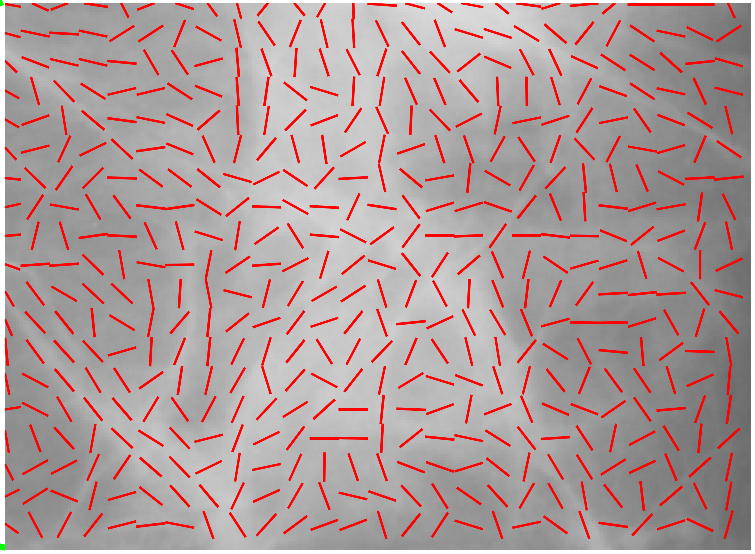
Mammogram



Gabor Magnitude

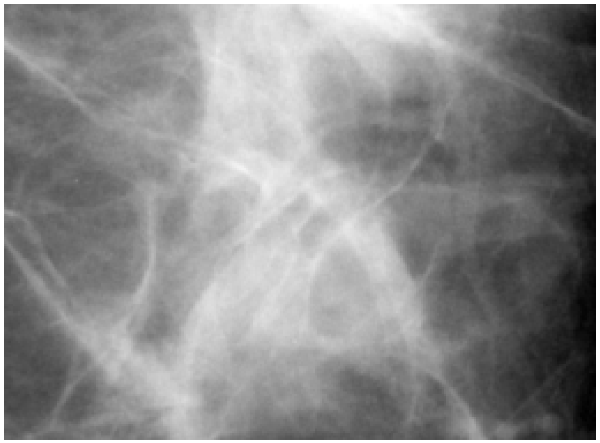
Interval cancer: site of architectural distortion



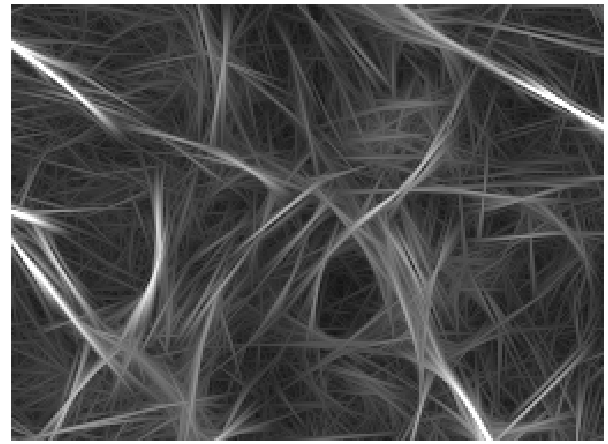


Orientation field

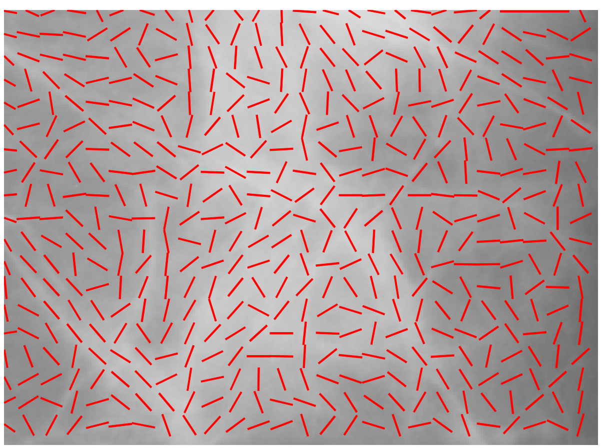
Site of architectural distortion



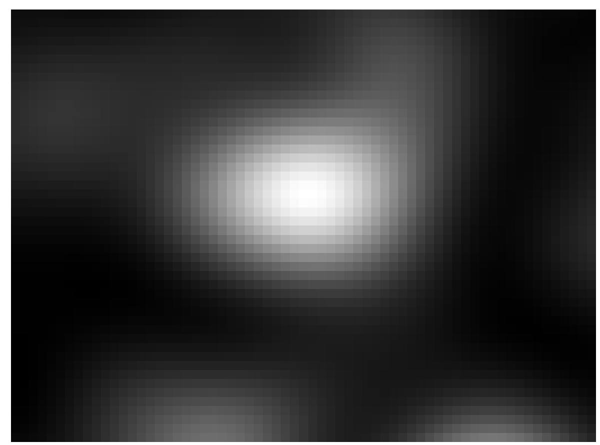
Mammogram



Gabor magnitude

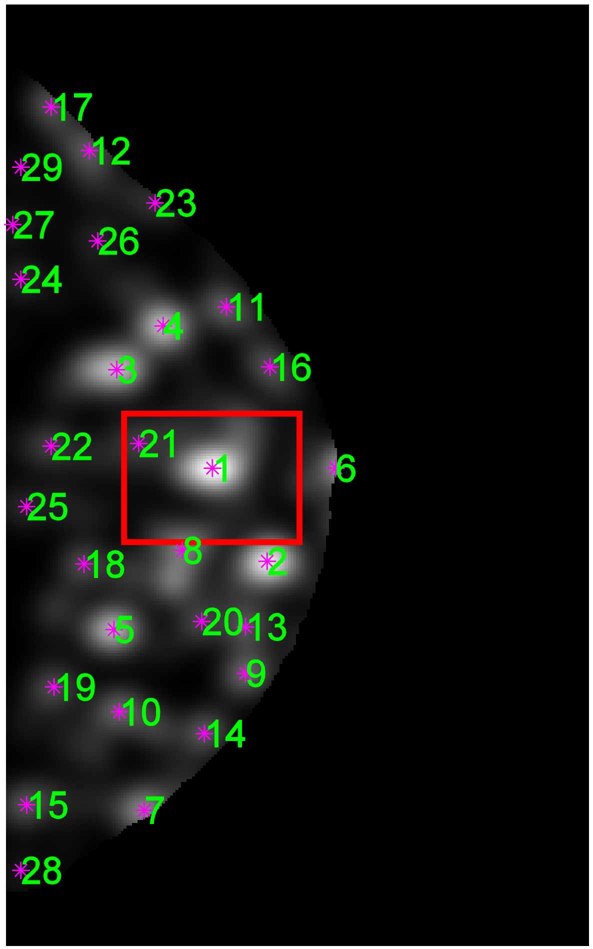


Orientation field

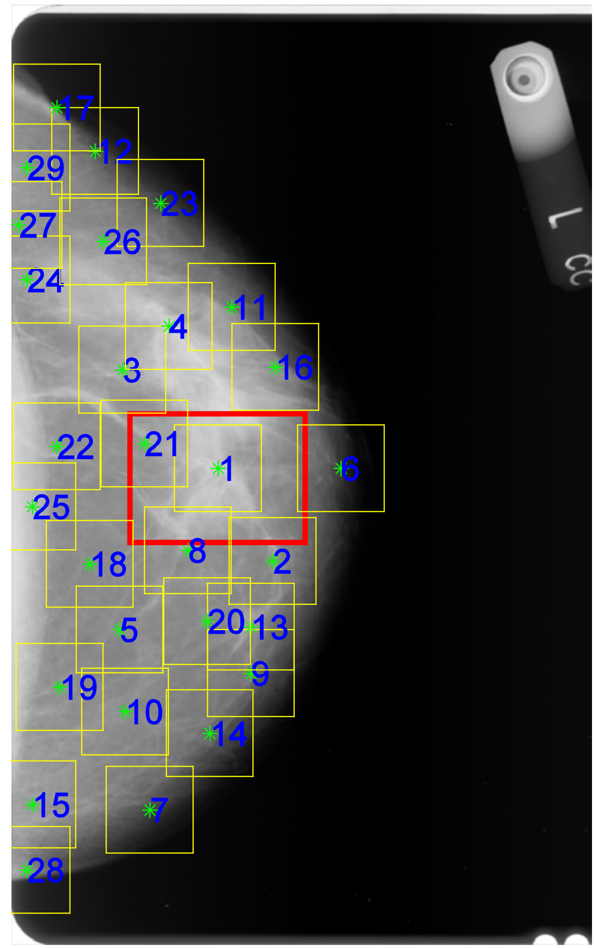


Node map

Interval cancer: potential sites of architectural distortion



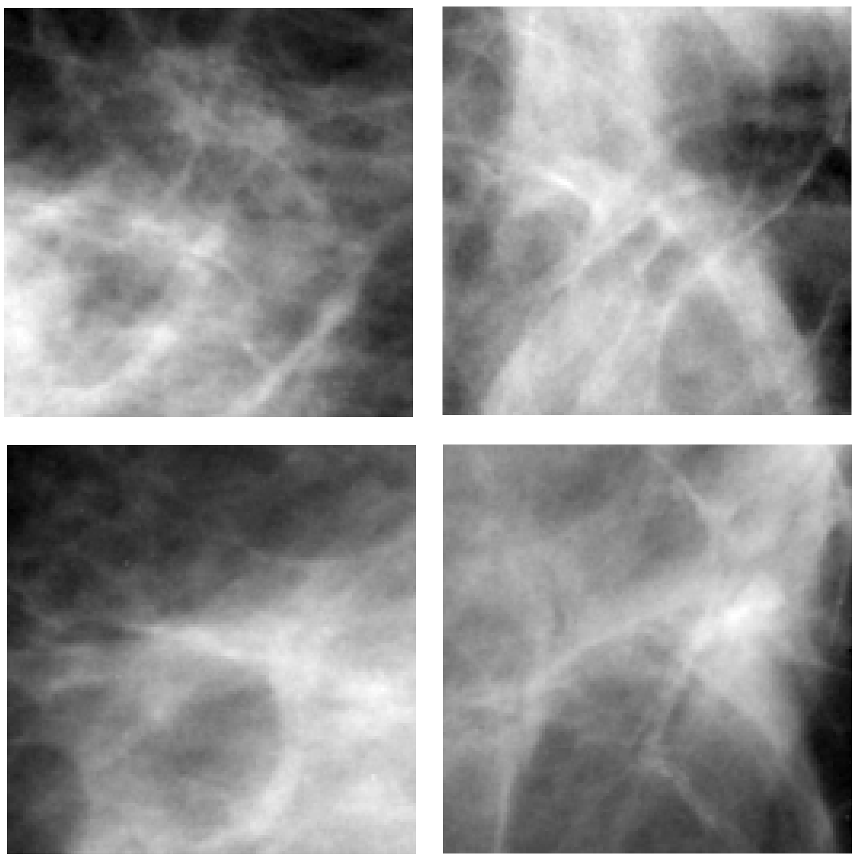
Node map



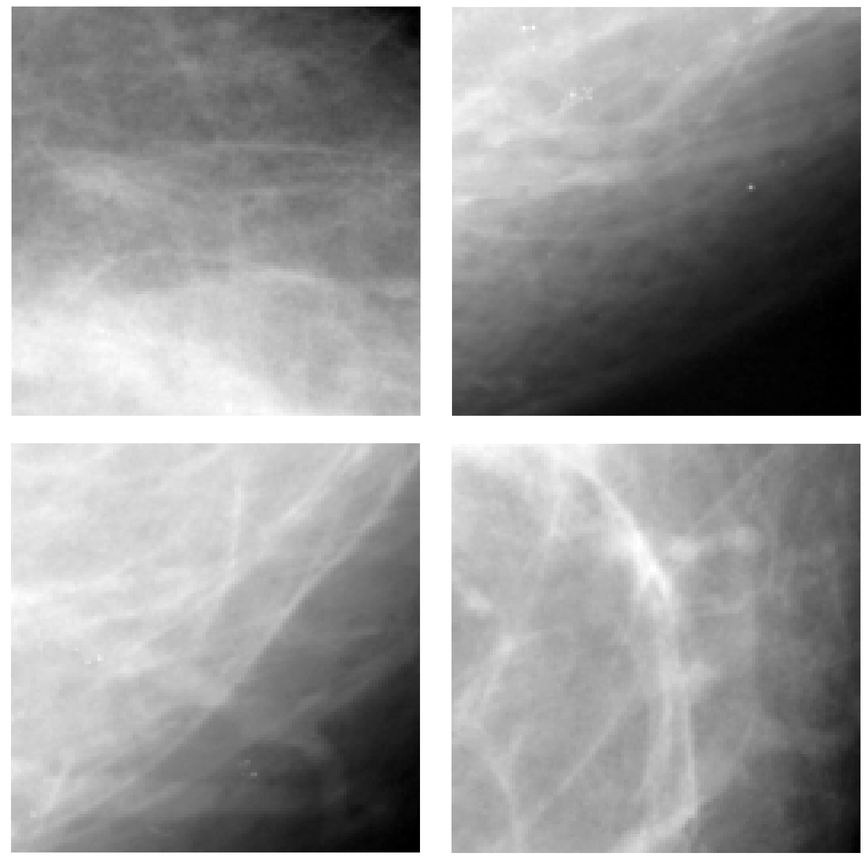
Automatically detected ROIs

Examples of detected ROIs

True-positive



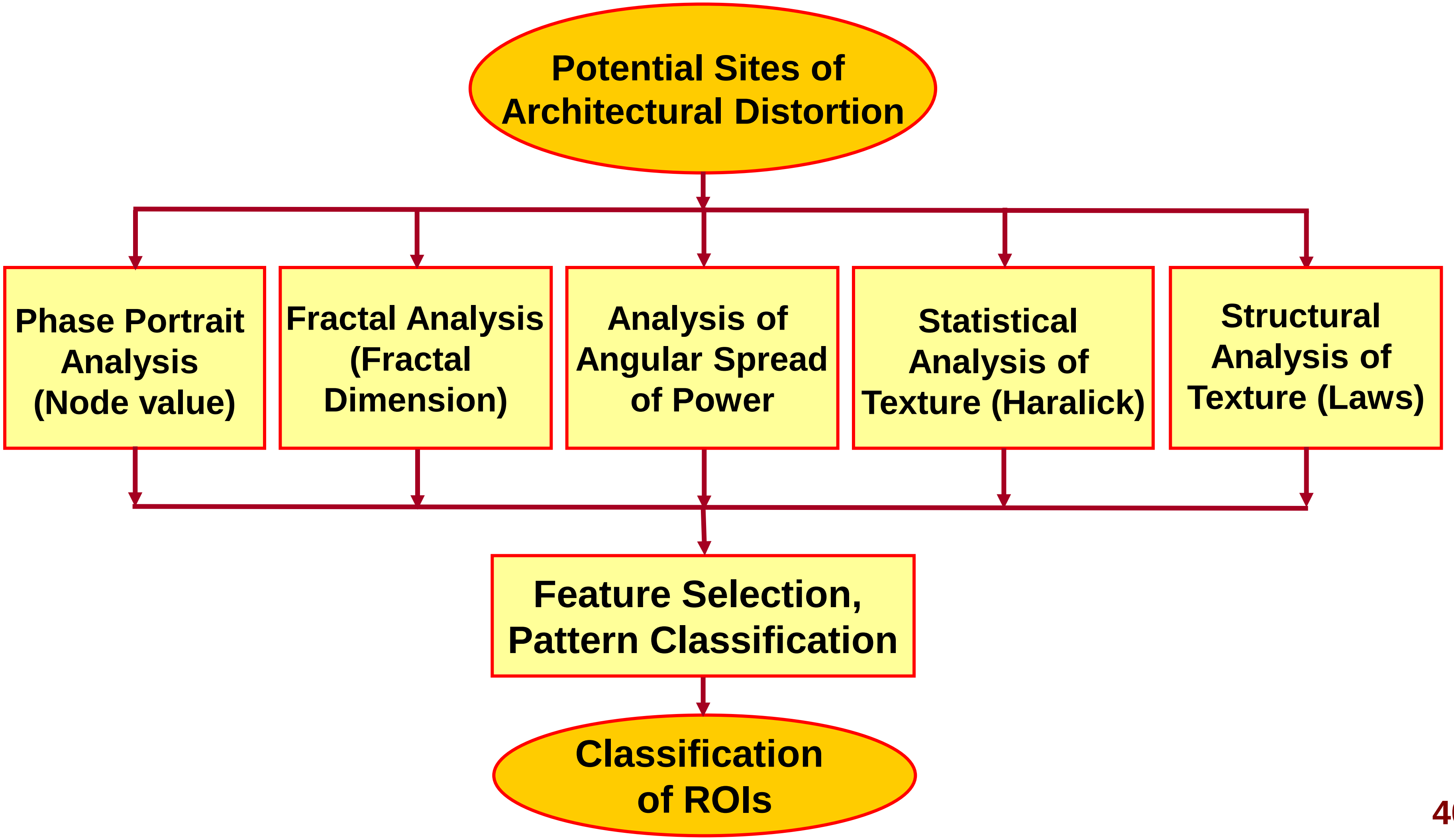
False-positive



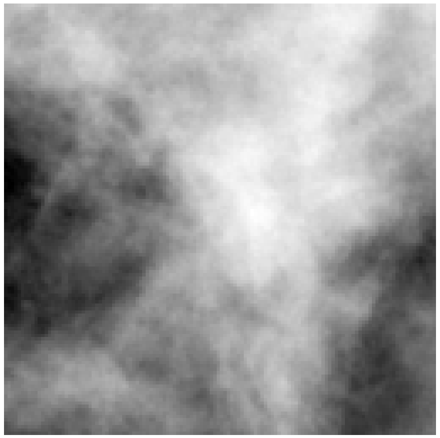
Automatically detected ROIs

Data Set	No. of Images	No. of ROIs 128 x 128 pixels at 200 µm/pixel	No. of True-Positive ROIs	No. of False-Positive ROIs
Prior mammograms of 56 interval-cancer cases	106	2821	301	2520
Penultimate mammograms of 13 normal cases	52	1403	0	1403
Total	158	4224	301	3923

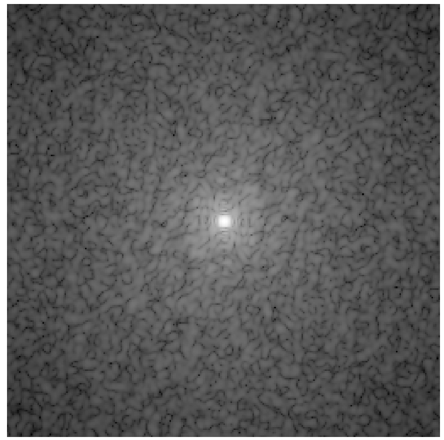
Feature extraction from ROIs



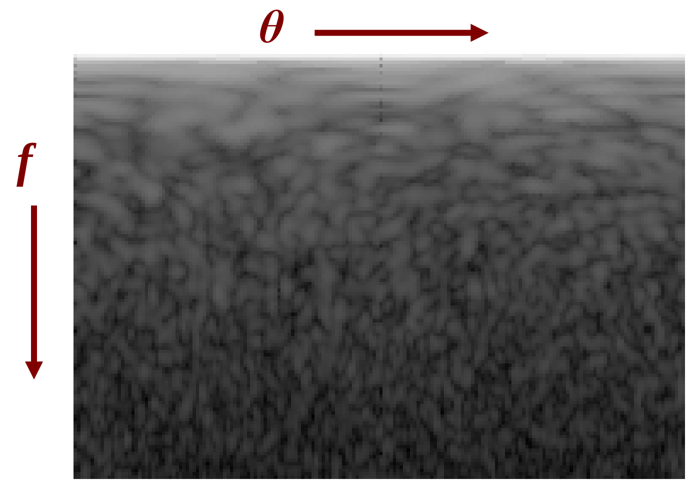
Fractal and spectral analysis



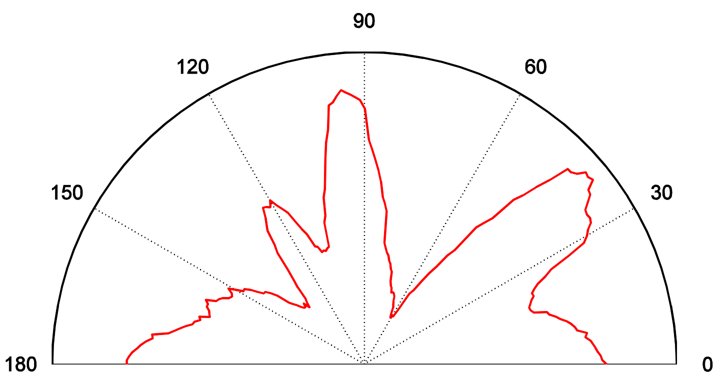
TP ROI, $s(x, y)$



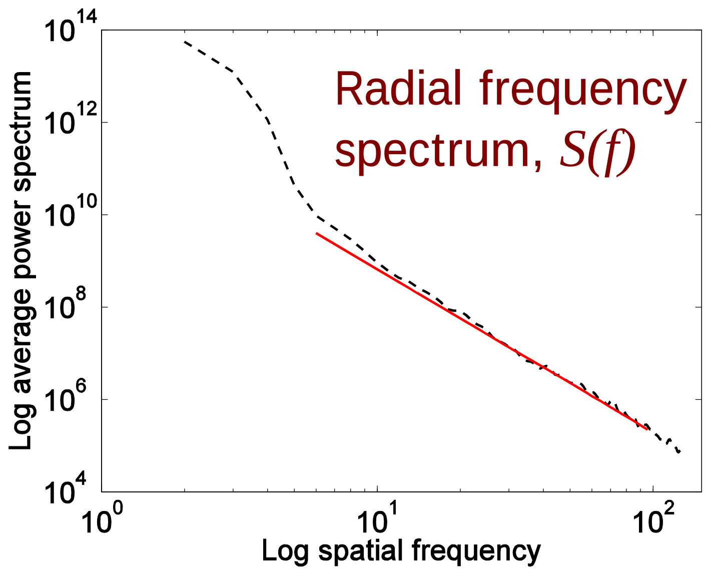
Fourier power spectrum, $S(u, v)$



Power spectrum in polar coordinates, $S(f, \theta)$



Angular spread of power, $S(\theta)$



Radial frequency spectrum, $S(f)$

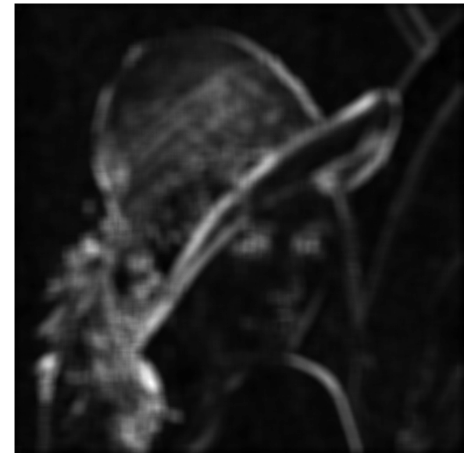
Laws' texture energy measures

- Operators of length five pixels may be generated by convolving the basic L3, E3, and S3 operators:
  - $L5 = L3 * L3$ = [1 4 6 4 1] (local average)
  - $E5 = L3 * E3$ = [-1 -2 0 2 1] (edges)
  - $S5 = -E3 * E3$ = [-1 0 2 0 -1] (spots)
  - $R5 = -S3 * S3$ = [1 -4 6 -4 1] (ripples)
  - $W5 = -E3 * S3$ = [-1 2 0 -2 1] (waves)
- 2D 5×5 convolution operators:
  - $L5L5 = L5^T L5$
  - $W5W5 = W5^T W5$
  - $R5R5 = R5^T R5$ etc.

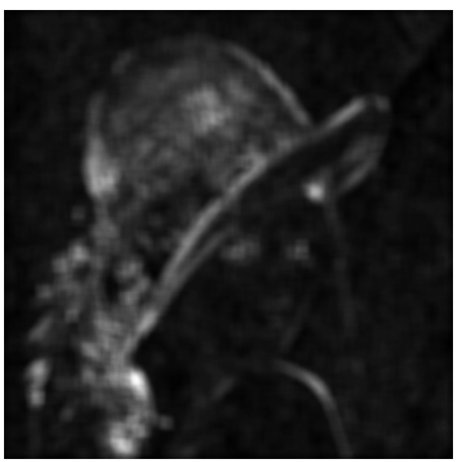
Laws' texture energy

Sum of the absolute values in the filtered images in a 15×15 window

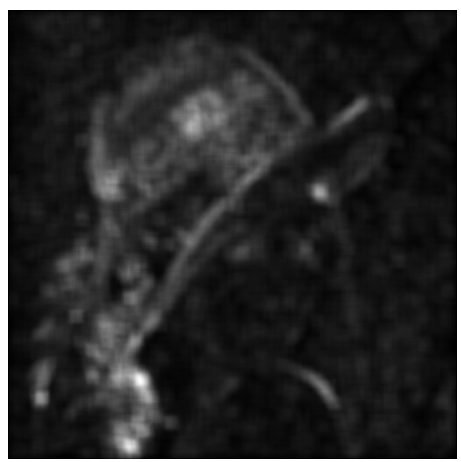
L5L5



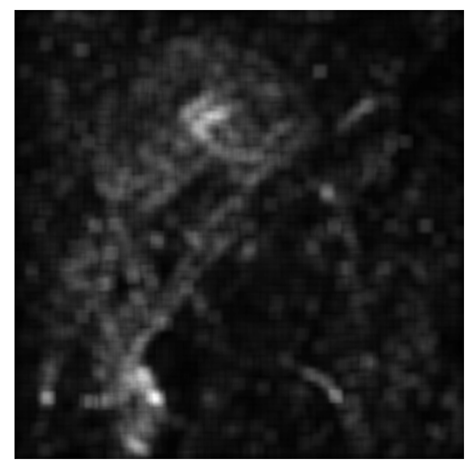
E5E5



S5S5

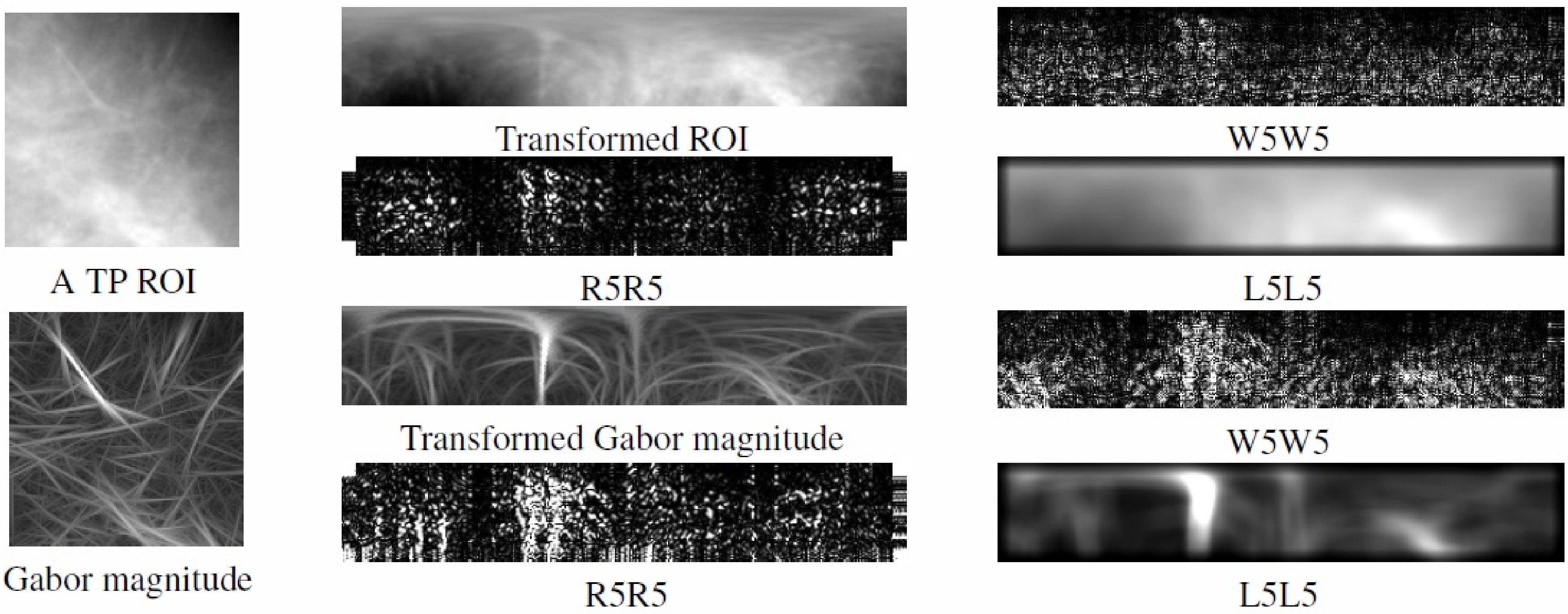


W5W5

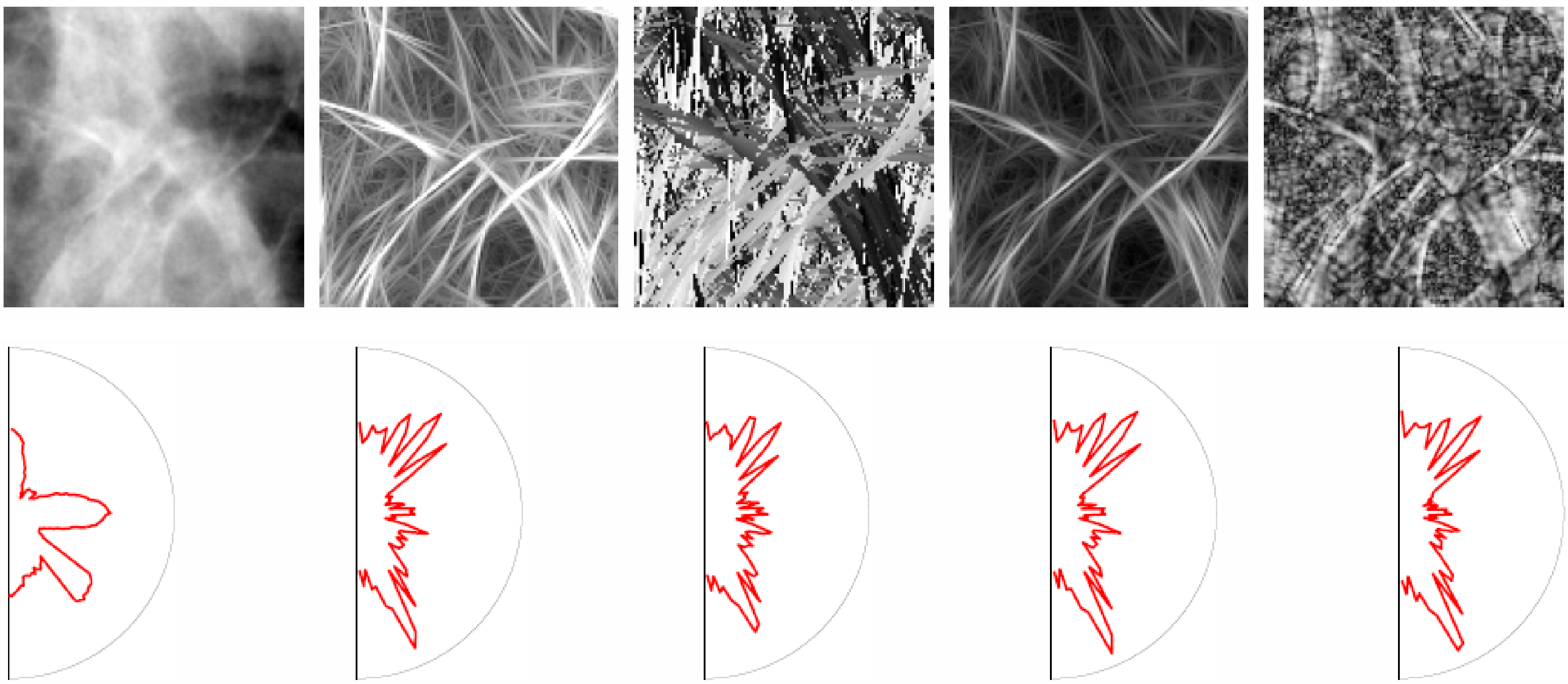


R5R5

Geometrical transformation for Laws' feature extraction



Analysis of angular spread: True-positive ROI



Frequency domain

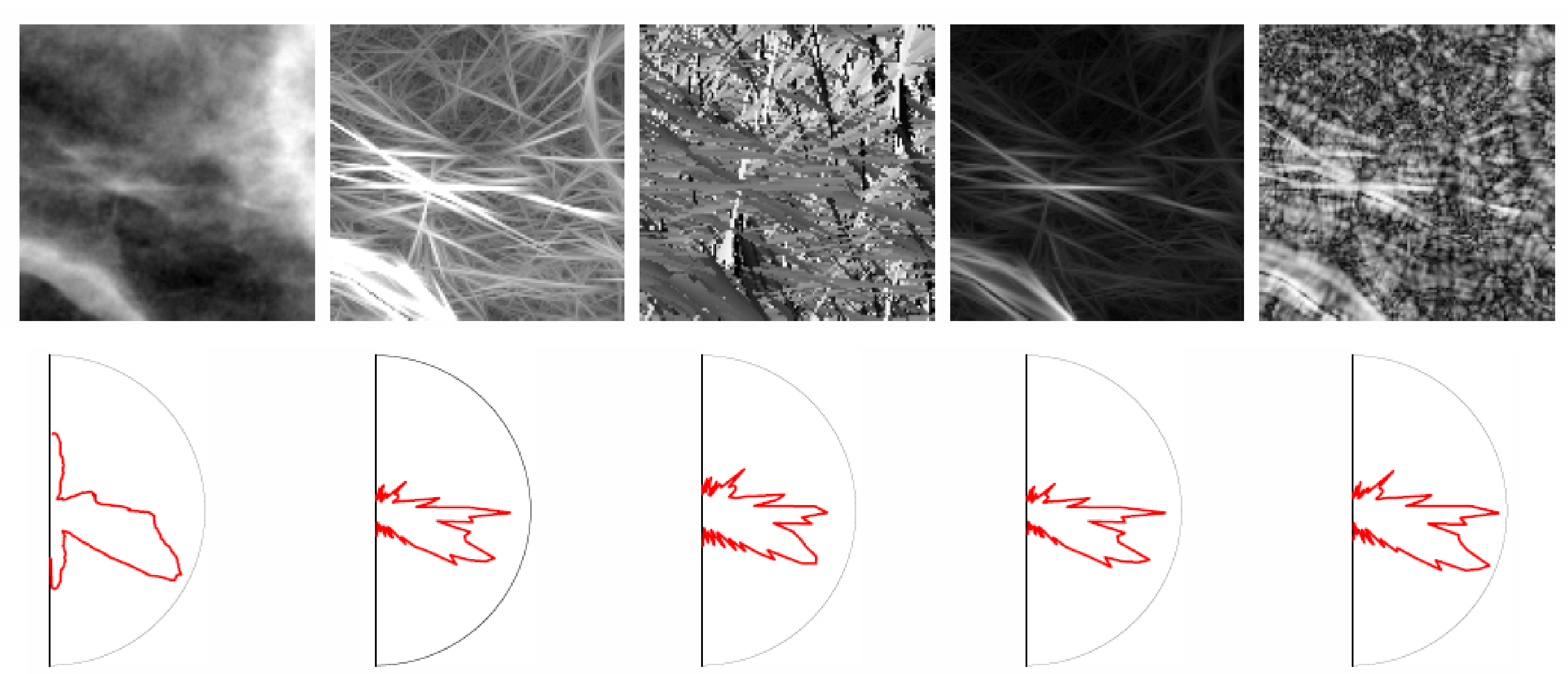
Gabor magnitude

Gabor orientation

Coherence

Orientation strength

Analysis of angular spread: False-positive ROI



Frequency domain | *Gabor magnitude* | *Gabor orientation* | *Coherence* | *Orientation strength*

UNIVERSITY OF CALGARY

Results with selected features

Classifiers	AUC using the selected features with stepwise logistic regression
FLDA (Leave-one-ROI-out)	0.75
Bayesian (Leave-one-ROI-out)	0.76
SLFF-NN (Single-layer feed forward: tangent-sigmoid)	0.78
SLFF-NN* (Single-layer feed forward: tangent-sigmoid)	**0.78** ± 0.02

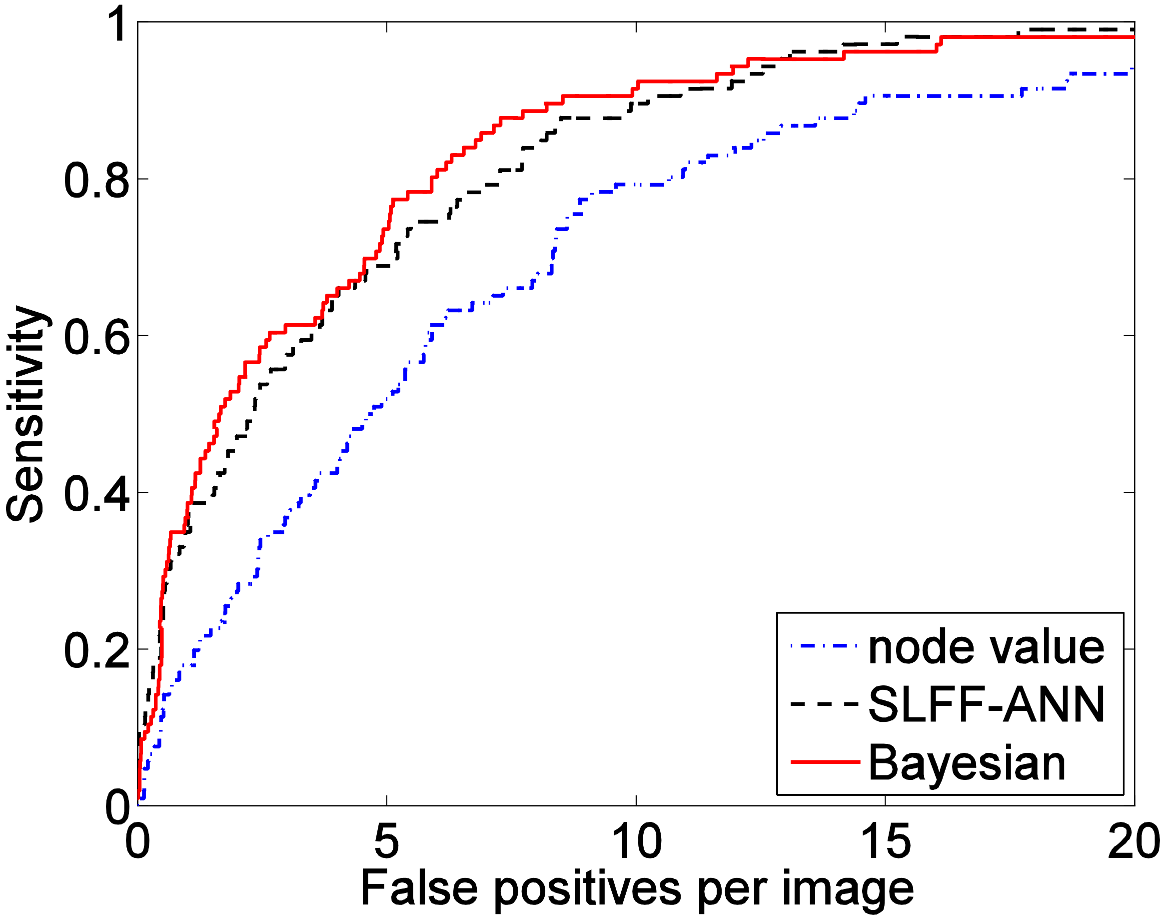
* Two-fold random subsampling, repeated 100 times

Free-response ROC

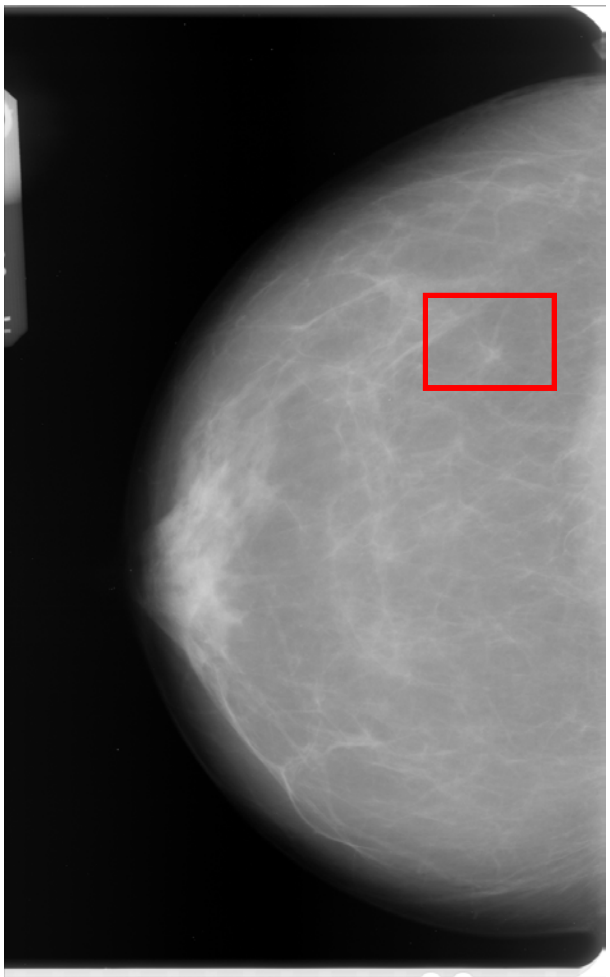
Sensitivity

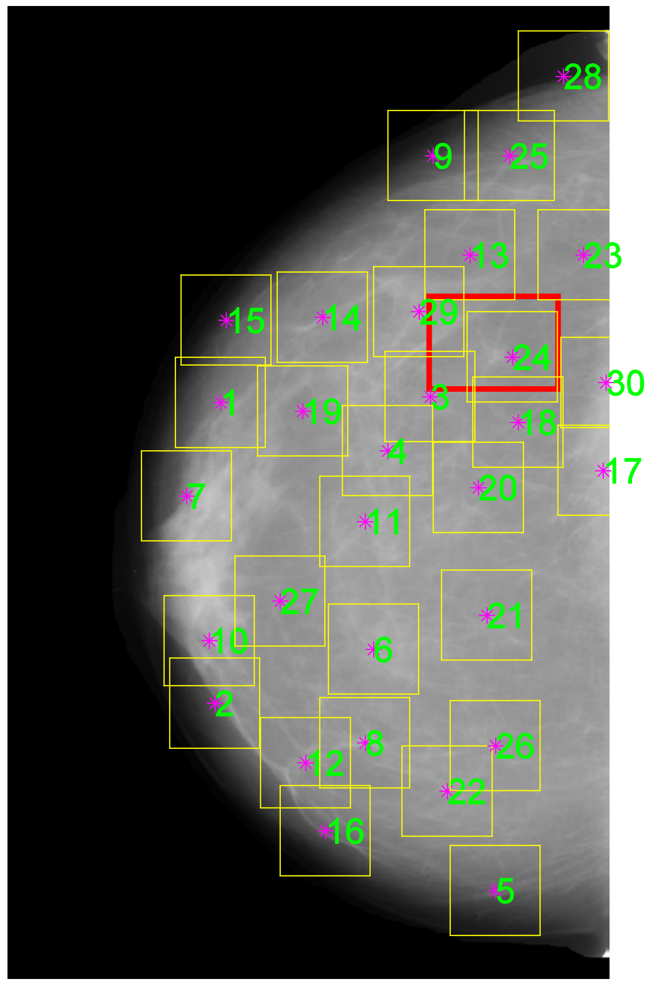
80% at 5.8 FP/image
90% at 8.1 FP/image

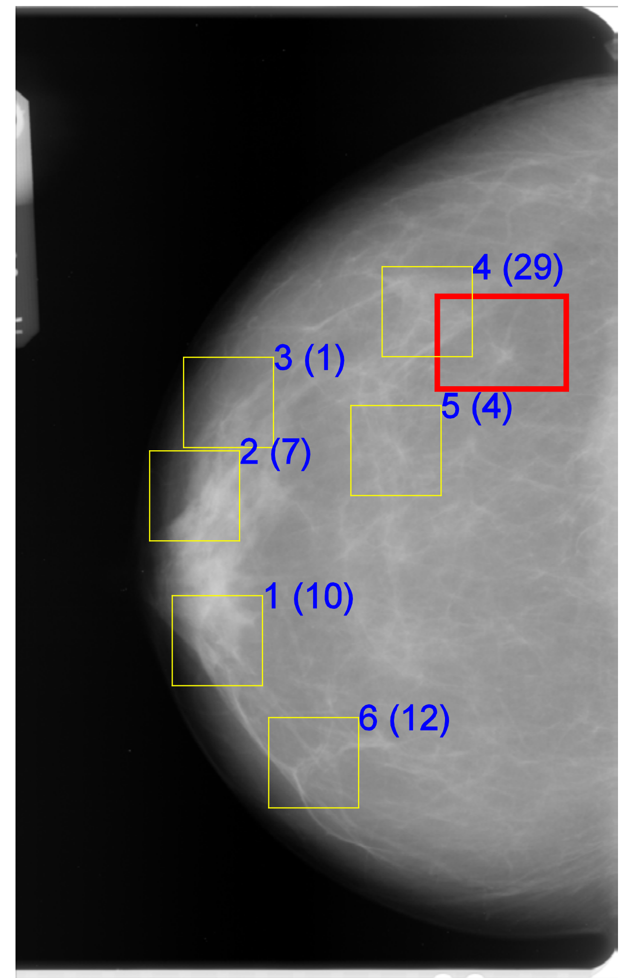
using features selected with stepwise logistic regression, the Bayesian classifier, and the leave-one-image out method



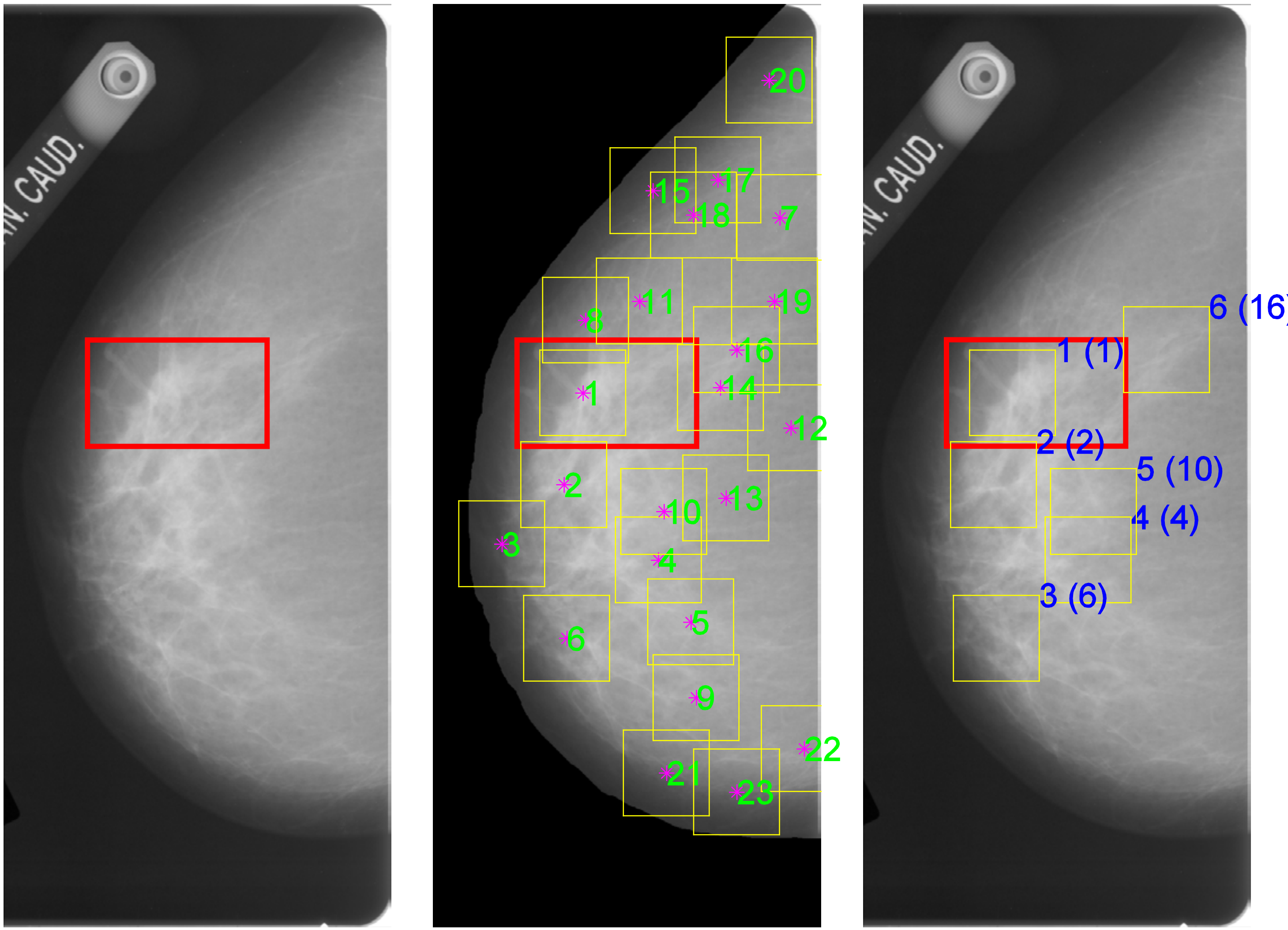
Bayesian ranking of ROIs: unsuccessful case







Bayesian ranking of ROIs: successful detection



Geometrical analysis of spicules and Gabor angle response

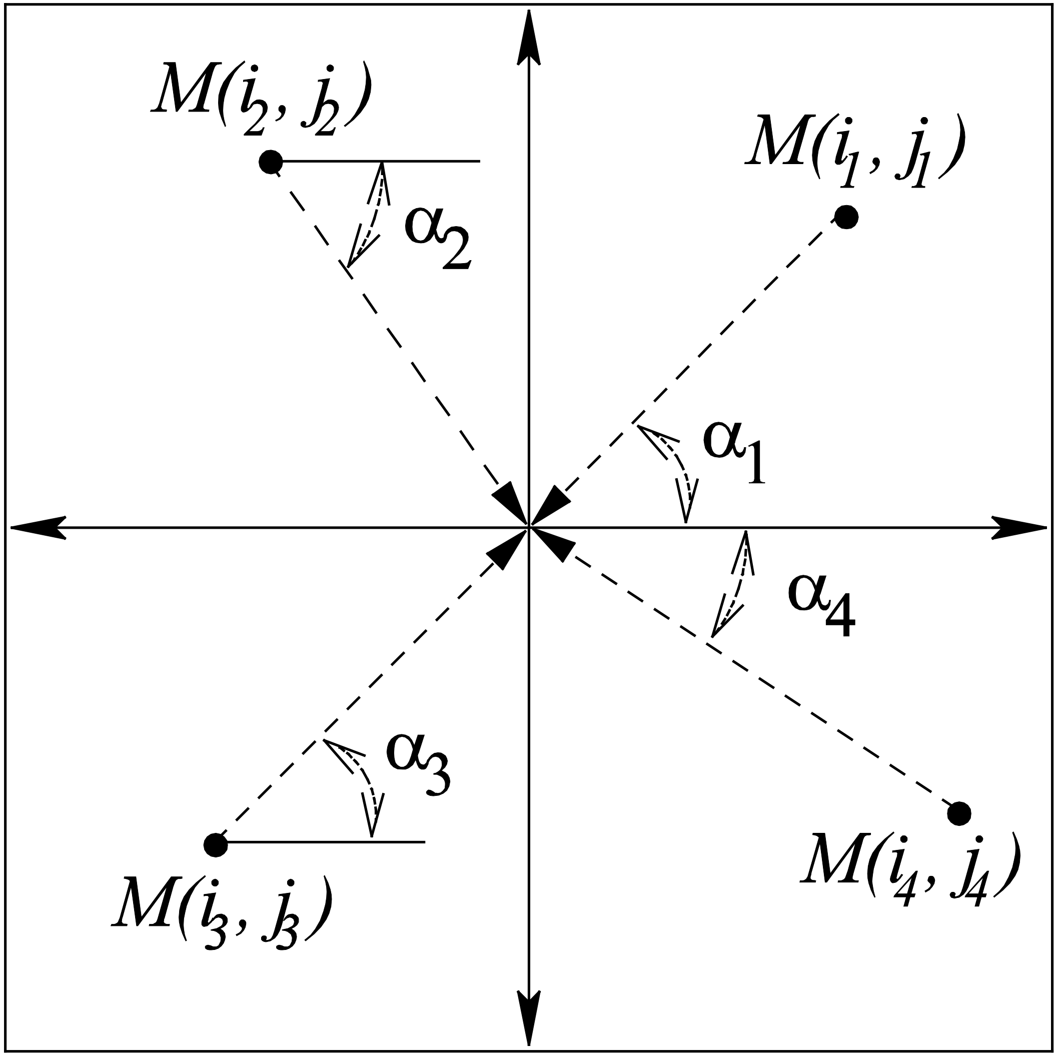
Index of convergence of spicules

$$\mathrm{ICS} = \sum_{i=1}^{P} \sum_{j=1}^{Q} M(i,j) \left| \cos[\theta(i,j) - \alpha(i,j)] \right|$$

$P \times Q$: size of the ROI
$\theta(i, j)$: Gabor angle response within the range [-89°, 90°]
$M(i, j)$: Gabor magnitude response
$\alpha(i, j)$: angle of a pixel with respect to the horizontal toward the center of ROI, in the range [-89°, 90°]

Index of convergence of spicules

ICS quantifies the degree of alignment of each pixel toward the center of the ROI weighted by the Gabor magnitude response



FROC analysis

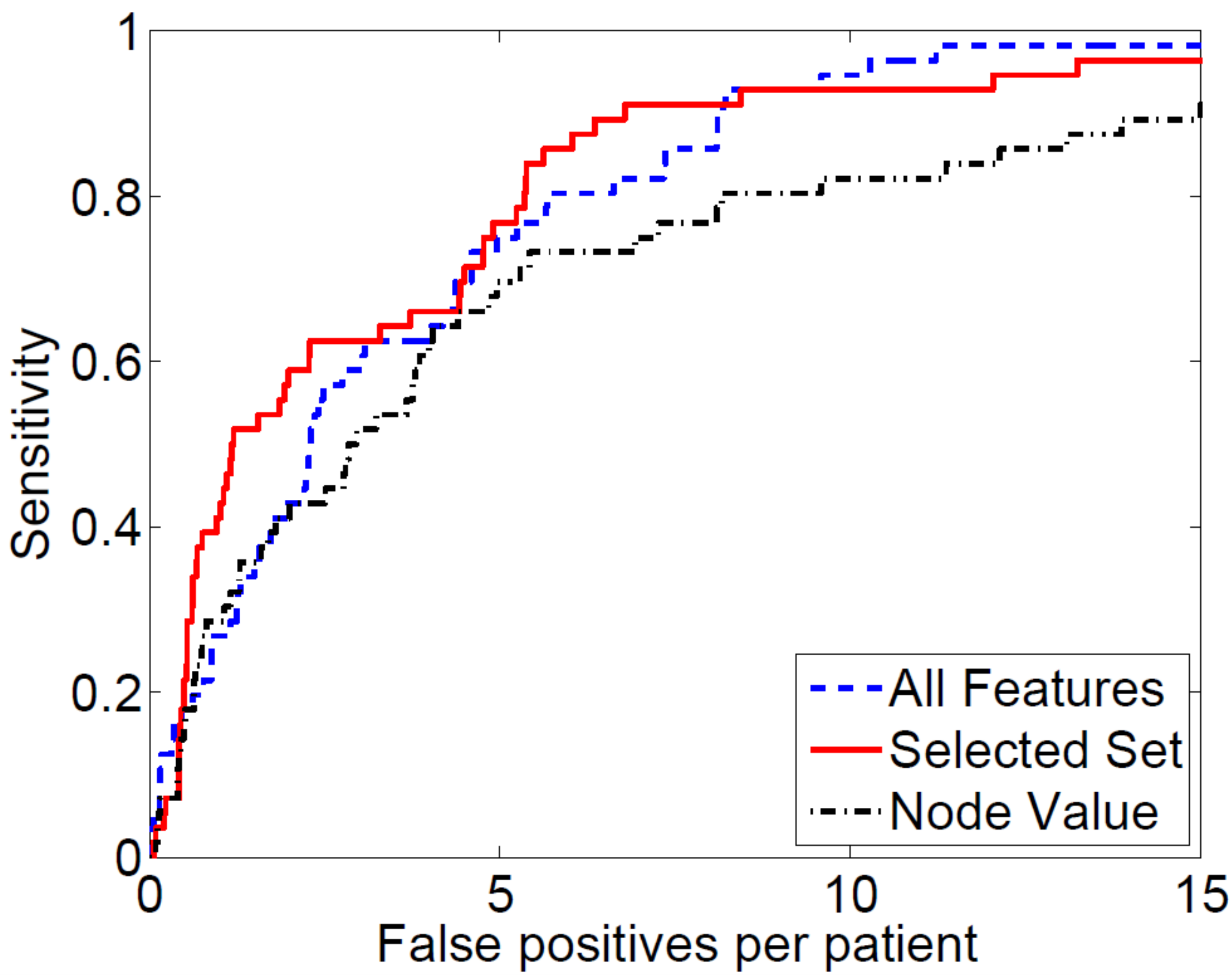
Sensitivity

80%

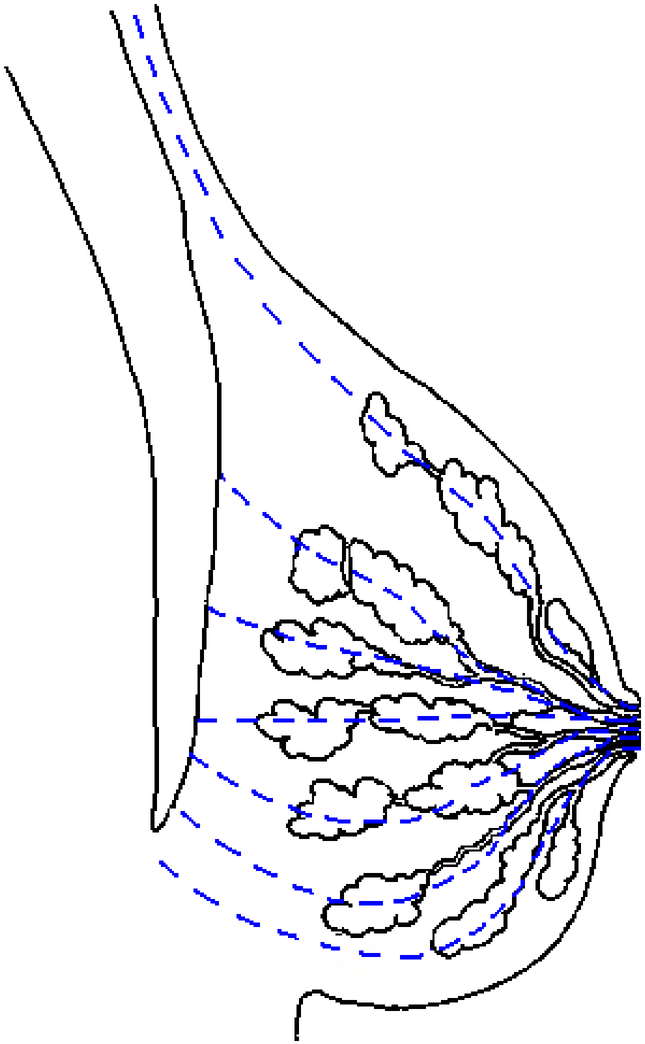
5.3 FP/patient

90%

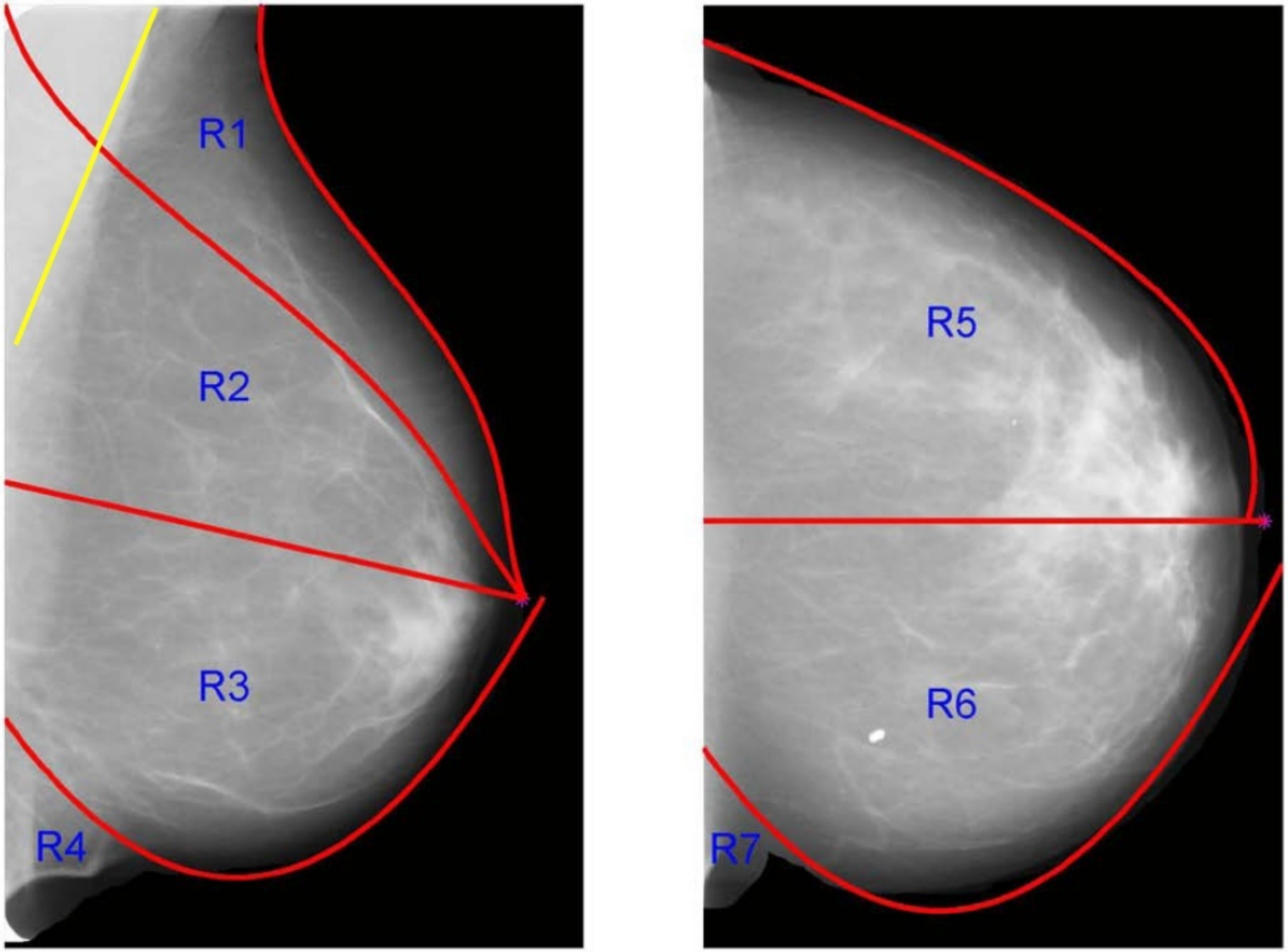
6.3 FP/patient



Expected loci of breast tissue

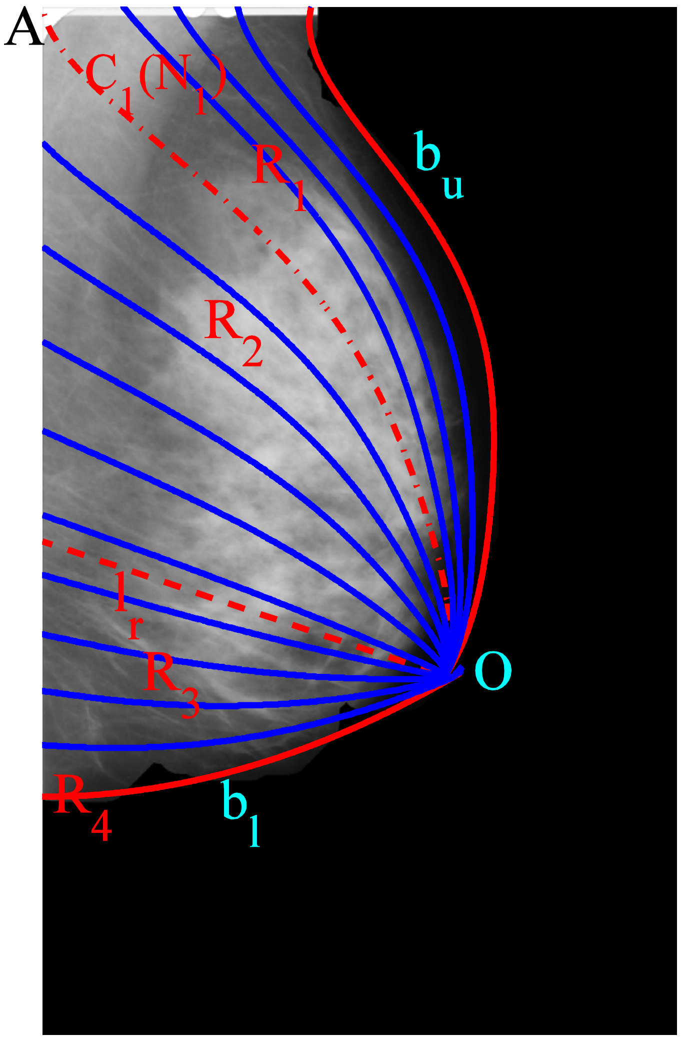


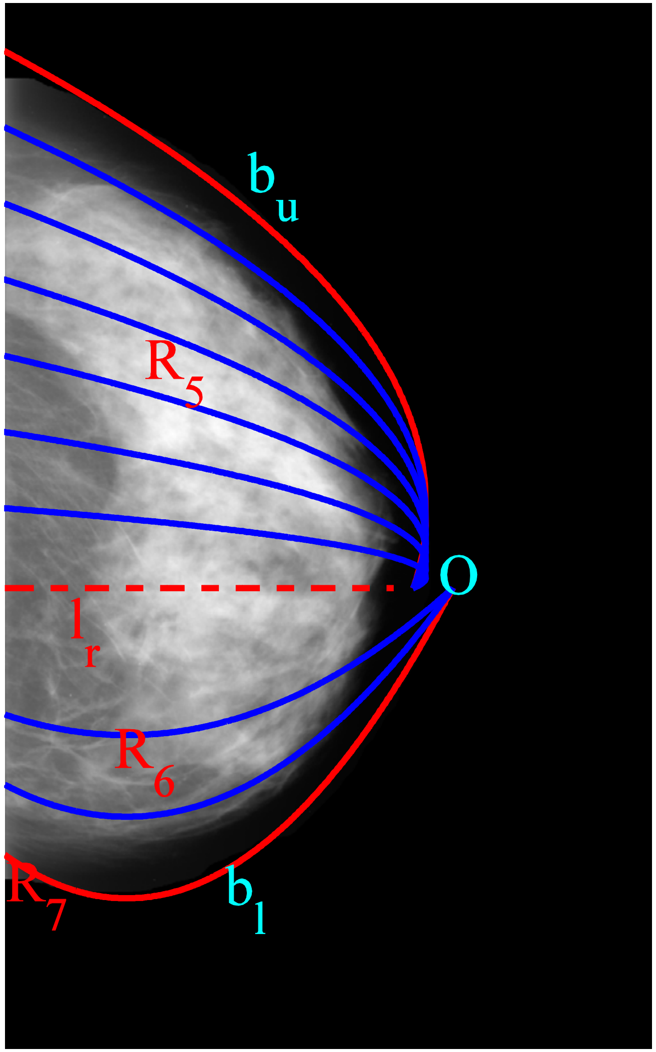
Landmarking of mammograms:
breast boundary, pectoral muscle, nipple

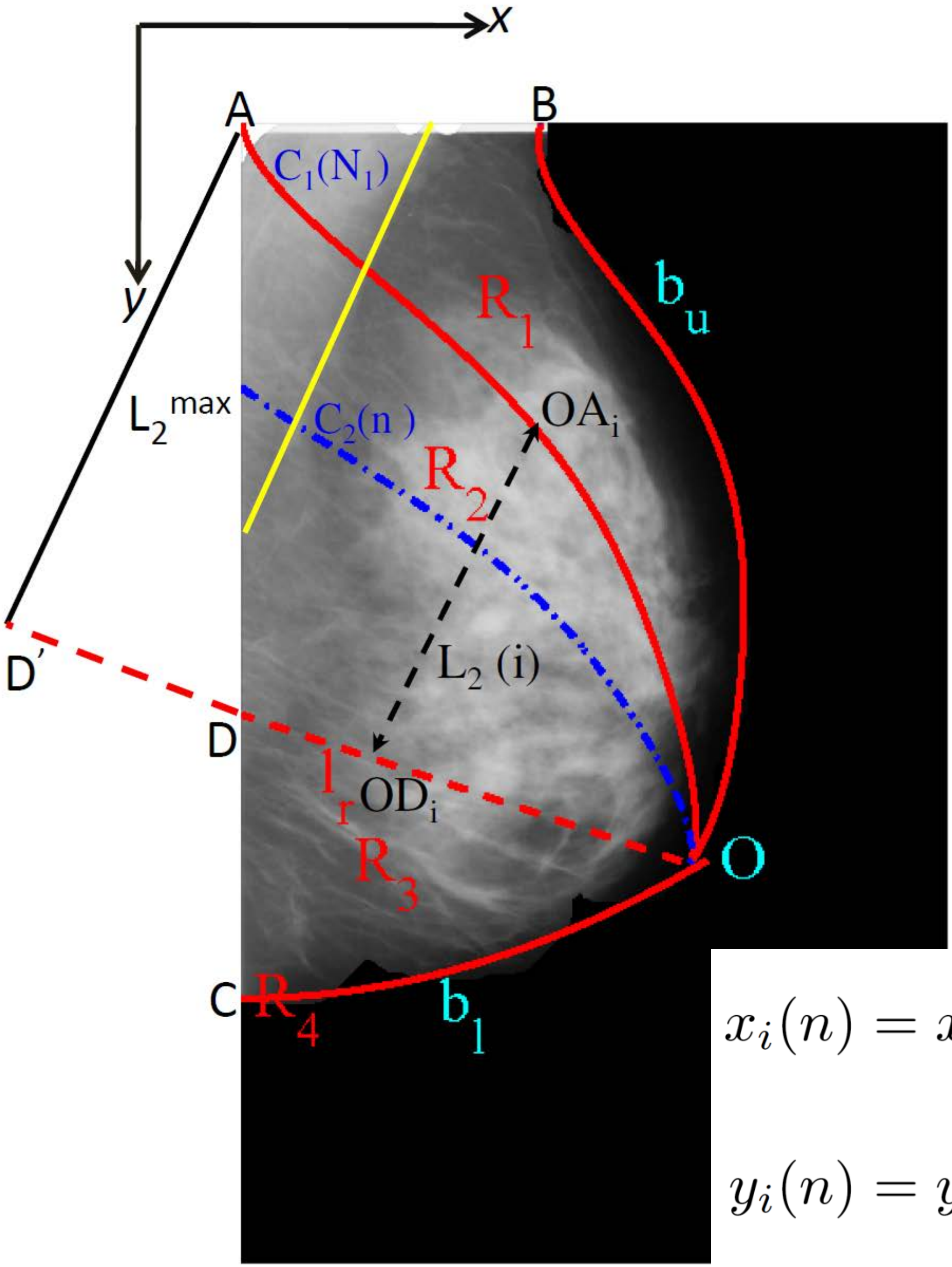


Second- and fifth-order polynomials fitted to parts of breast boundary

Derivation of expected loci of breast tissue: interpolation







Number of points in curve = M

L_i = ⊥ length between two curves at the i-th point

$L_{max} = max(L_i)$

Number of curves = N = L_{max}+1

Distance at i-th point = L_i / L_{max} = L_i /(N-1)

i-th point of n-th curve:

$$x_i(n) = x_i(1) - [x_i(1) - x_i(N_2)]\left(\frac{n-1}{N_2-1}\right)$$

$$y_i(n) = y_i(1) - [y_i(1) - y_i(N_2)]\left(\frac{n-1}{N_2-1}\right)$$

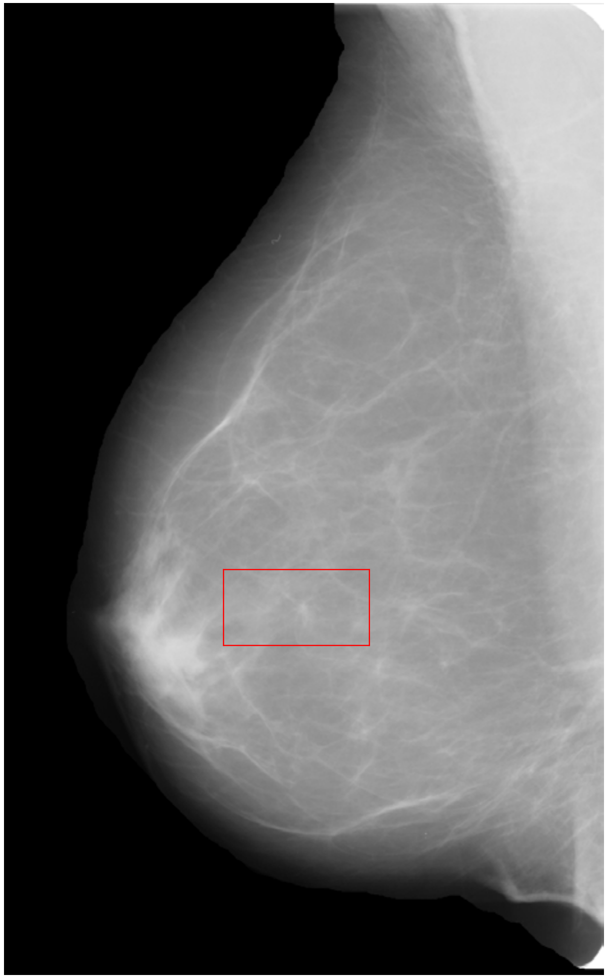
Divergence with respect to the expected loci of breast tissue

$$\gamma(i,j) = \frac{\sum_{m=1}^{L}\sum_{n=1}^{L}|M(m,n)\cos[\theta(m,n)-\phi(i,j)]|}{\sum_{m=1}^{L}\sum_{n=1}^{L}M(m,n)}$$

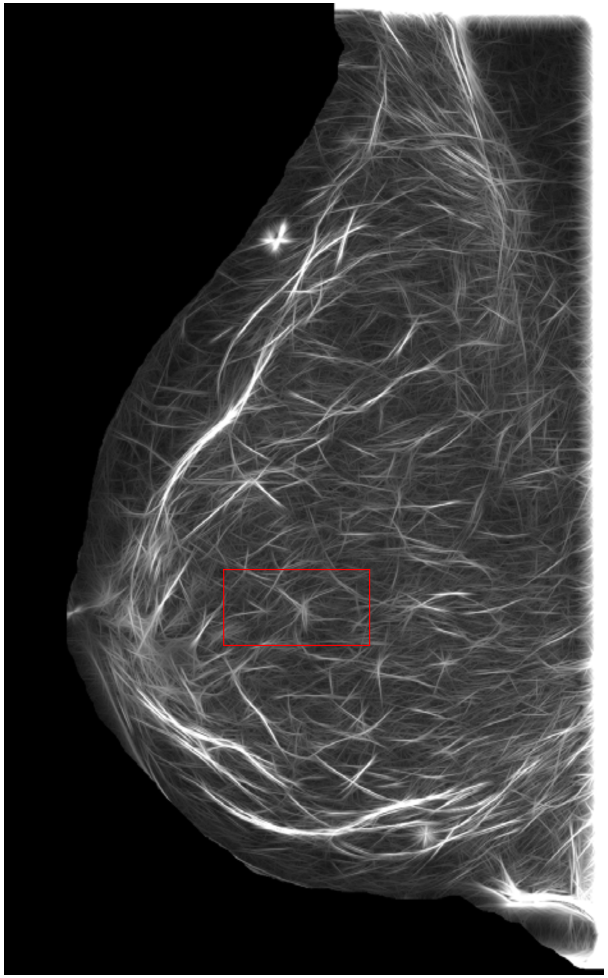
M: Gabor magnitude response
θ: Gabor angle response
ϕ: expected orientation of breast tissue
L: 25 pixels at 200 µm/pixel
180 Gabor filters used over [-90, 90] degrees

$$D(i,j) = 1 - \gamma(i,j)$$

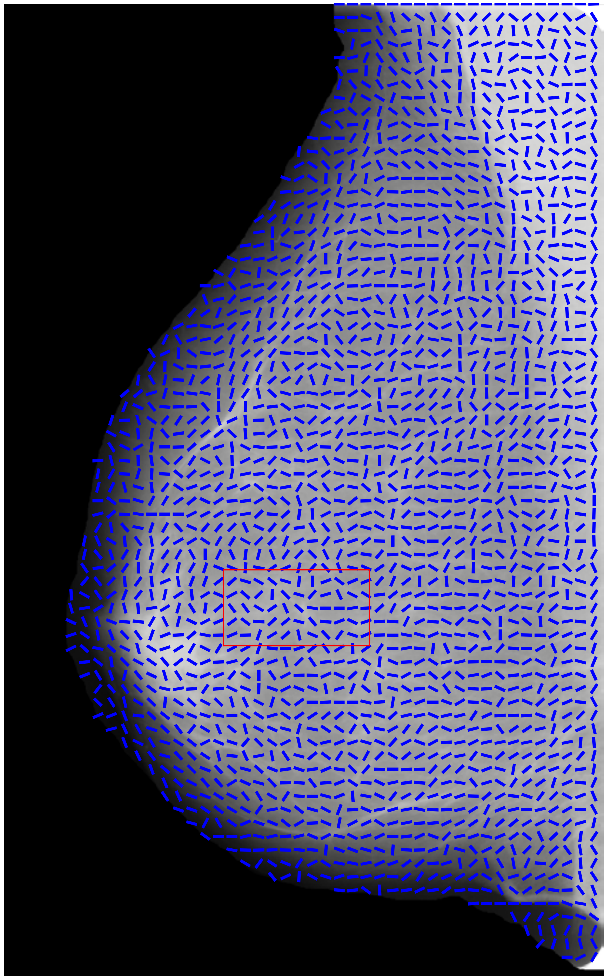
Orientation field of breast tissue obtained using Gabor filters



Original image

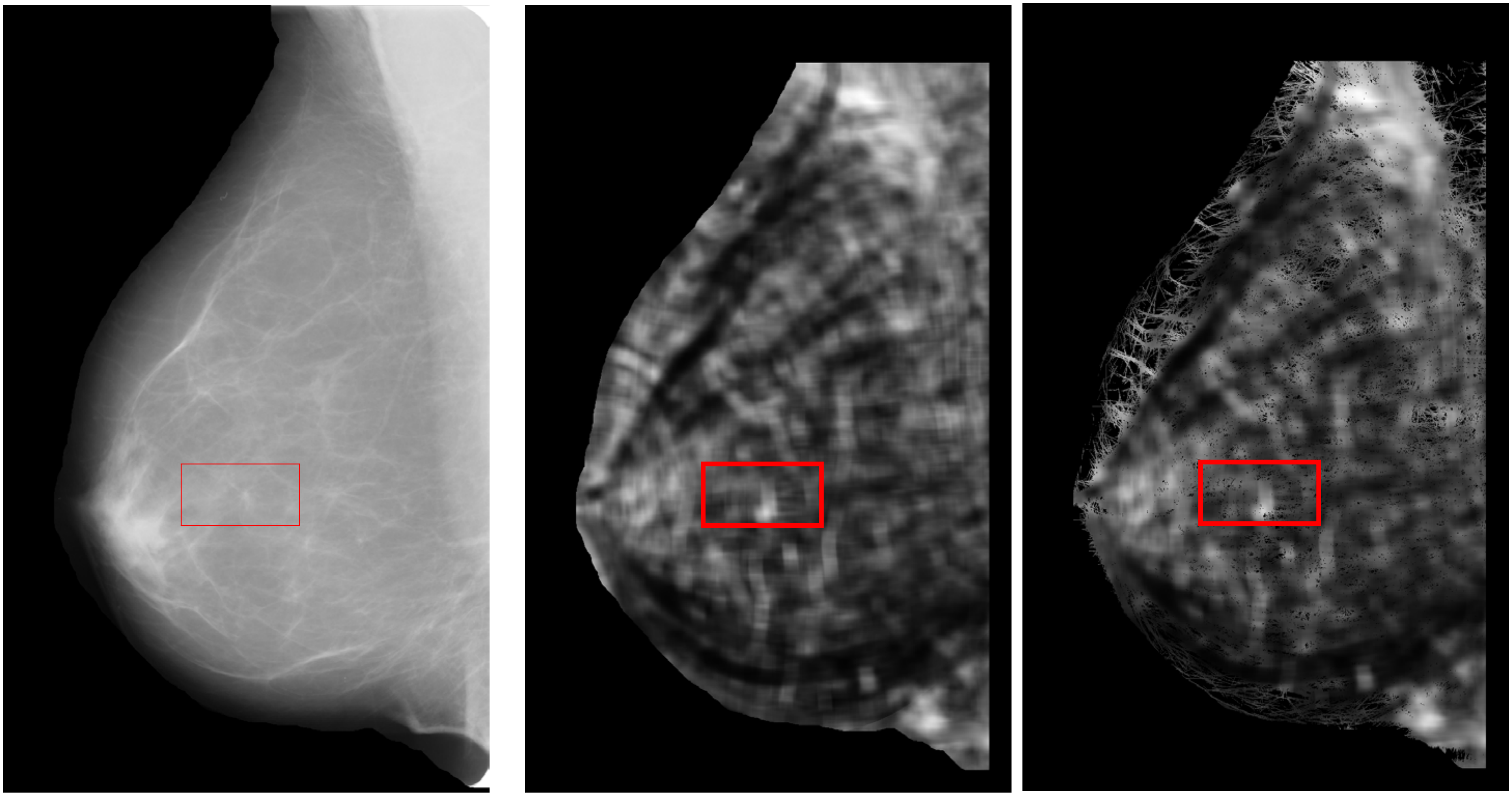


Gabor magnitude



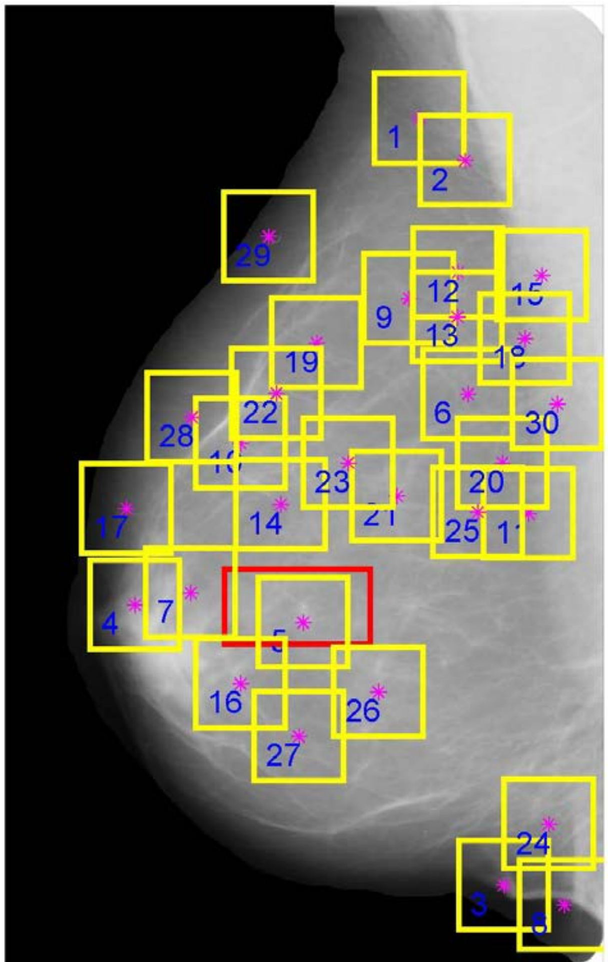
Gabor angle

Divergence with respect to the expected loci of breast tissue



Original image *Divergence map* *Thresholded map*

Automatically detected regions of interest



ROC: AUC = 0.61

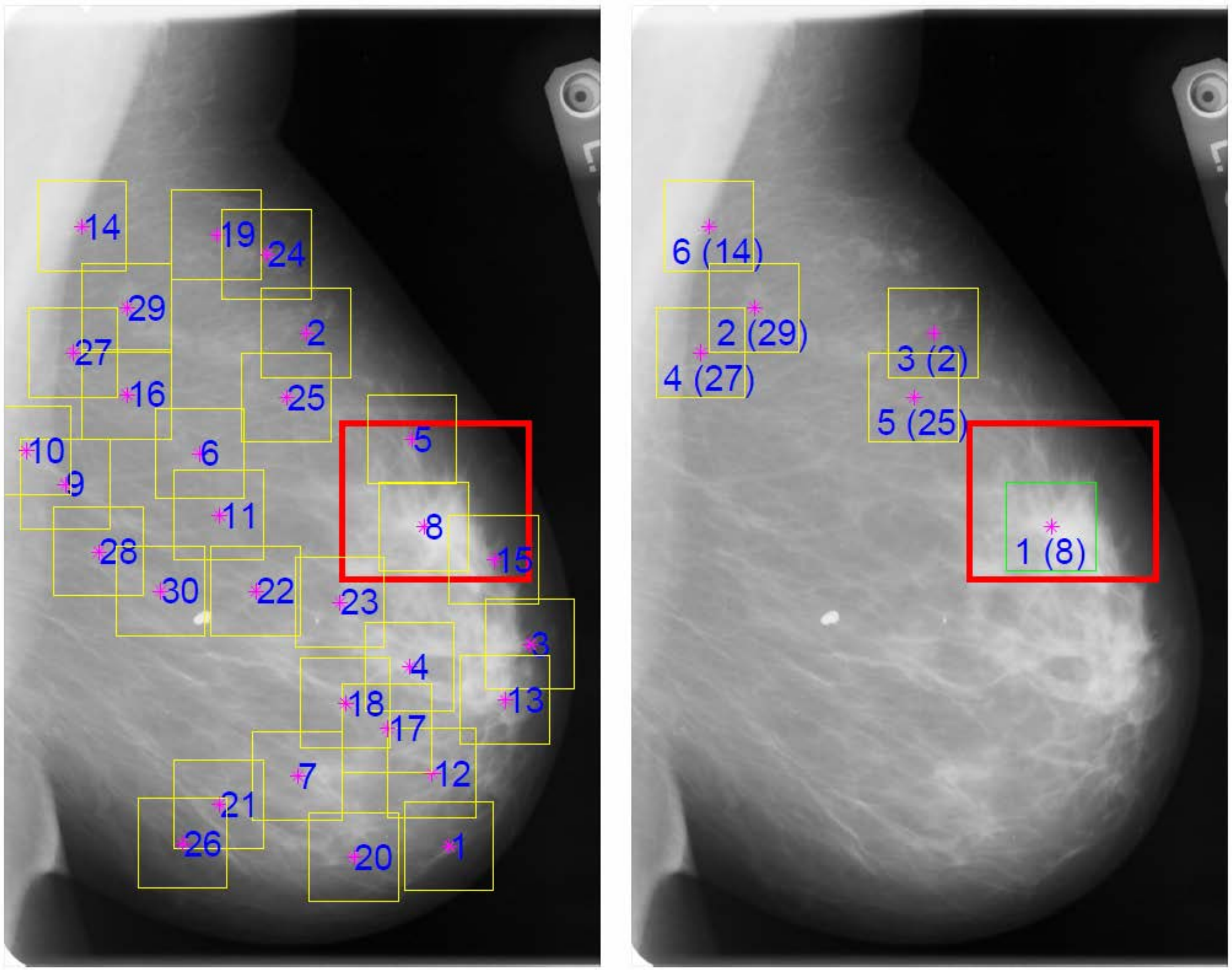
FROC:
Sensitivity = 80%
at 9.1 FP/patient

Combination of 86 features

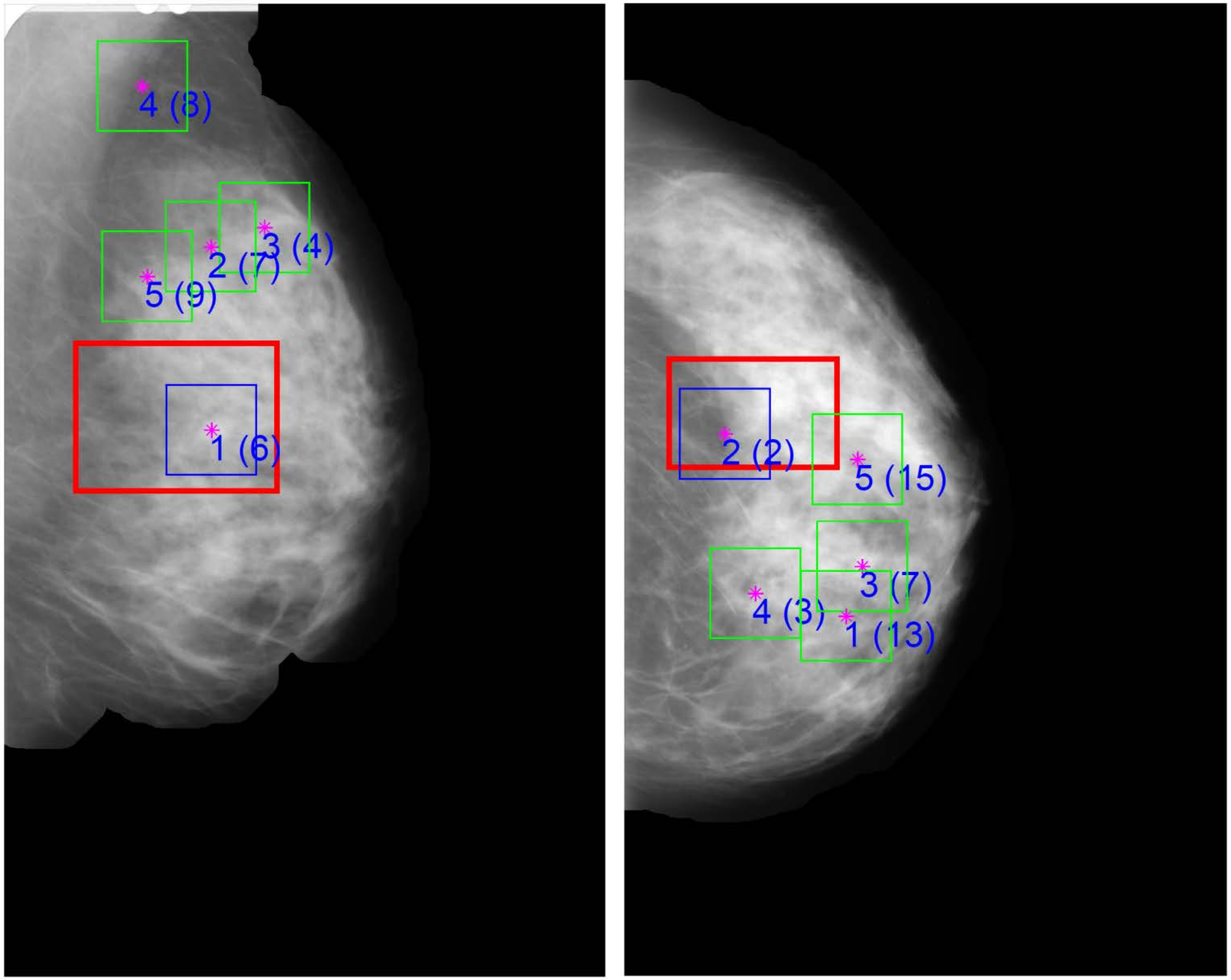
- Geometrical features of spicules: 12
- Haralick’s and Laws’ texture features, fractal dimension: 25
- Angular spread, entropy: 15
- Haralick’s measures with angle cooccurrence matrices: 28
- Statistical measures of angular dispersion and correlation: 6
- Feature selection with stepwise logistic regression
- Bayesian classifier with leave-one-patient-out validation:
  80% sensitivity at 3.7 FP/patient

UNIVERSITY OF CALGARY

Reduction of false positives



Reduction of false positives



UNIVERSITY OF CALGARY

Conclusion

"Our methods can detect early signs of breast cancer 15 months ahead of the time of clinical diagnosis with a sensitivity of 80% with fewer than 4 false positives per patient"

- Further work required:
  - Detection of sites of architectural distortion at higher sensitivity and lower false-positive rates
  - Application to direct digital mammograms and breast tomosynthesis images

Thank You!

- Natural Sciences and Engineering Research Council (NSERC) of Canada
- Alberta Heritage Foundation for Medical Research
- Alberta and Canadian Breast Cancer Foundation
- Screen Test: Alberta Program for the Early Detection of Breast Cancer
- Indian Institute of Technology Kharagpur
- Shastri Indo-Canadian Institute
- University of Calgary International Grants Committee
- Department of Information Technology, Government of India
- My collaborators and students:

  Dr. J.E.L. Desautels, N. Mudigonda, H. Alto, F.J. Ayres, S. Banik, S. Prajna, J. Chakraborty, Dr. S. Mukhopadhyay

http://people.ucalgary.ca/~ranga/