Semi-automatic feature-based machine learning classification models for breast lesions

A. Hasan, S. Hasan

Objectives

- *To discriminate* between **benign** and **malignant** breast lesions
  - Clinicians have found that:
    - Malignant lesions appear larger in strain images compared to B-Mode images.
    - Benign lesions appear to be of similar size.
- *Developed* a semi-automatic algorithm to compute the **size difference** in strain and B-mode images.
  - Classification depends on physician expertise
  - Many images to manually analyze
  - even experts can sometimes misread read conventional ultrasounds images.

Garra et. al. (1997), *Radiology*.
Barr et. al. (2012) *J Ultrasound Med*.

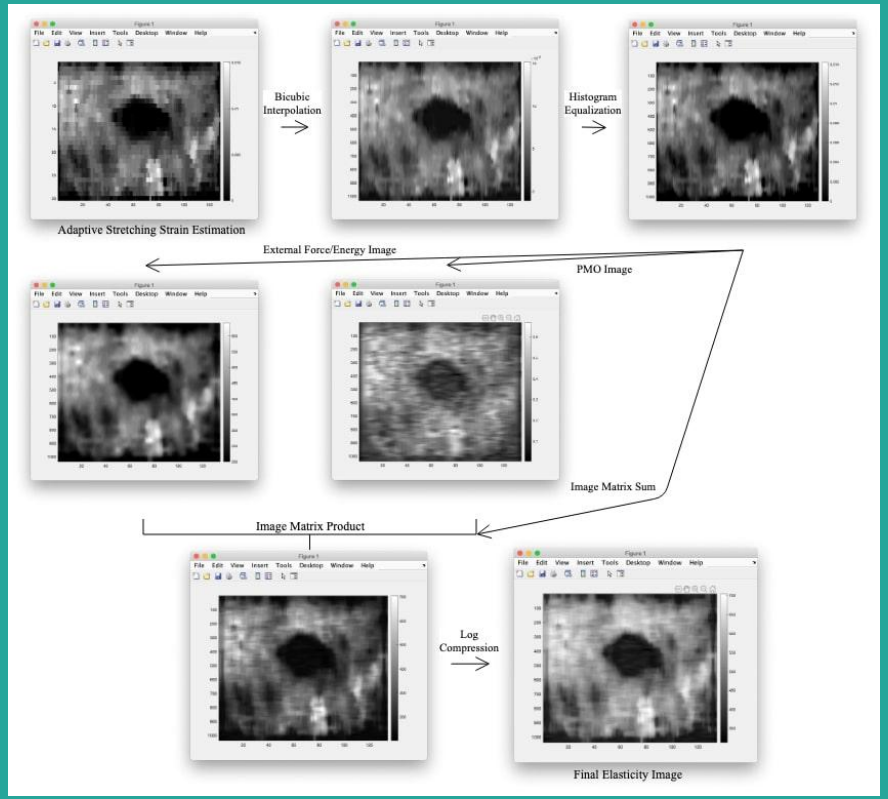
Methods — Images

- RF echosignal → **novel two-step adaptive-stretching strain estimation algorithm** → elastograms
  - Strain map used in the first step was used as an input to refine the strain map in the second step.
- RF echosignal → Hilbert Transform → B-Mode Image
- Elastogram/B-Mode Processing

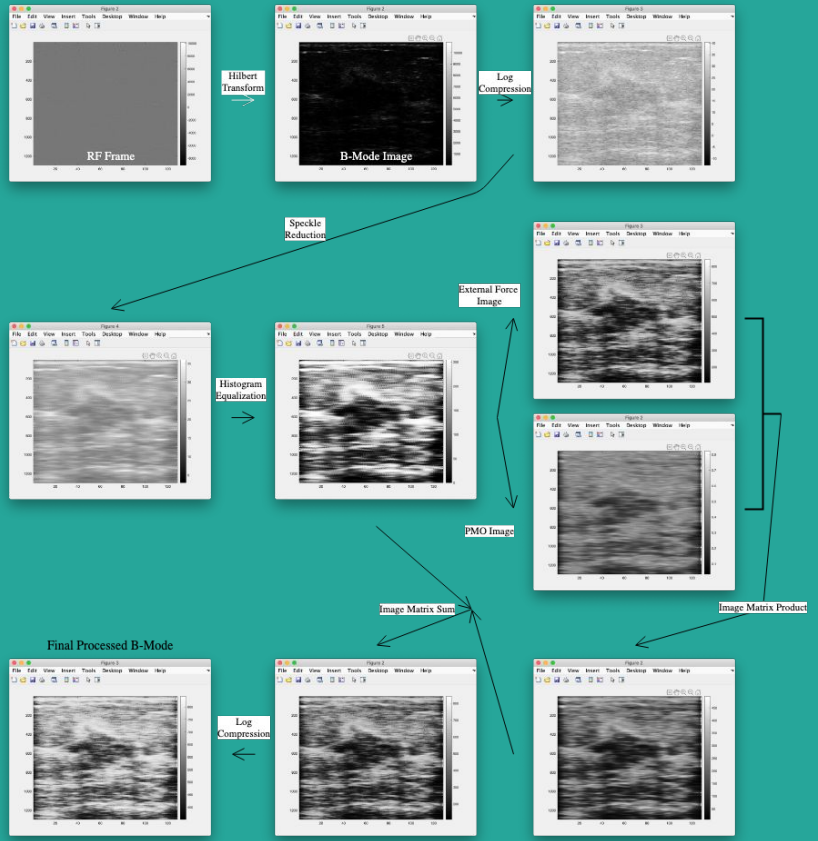
Alam et. al. (1998) *IEEE-TUFFC*.

Processing Algorithms

Strain Image

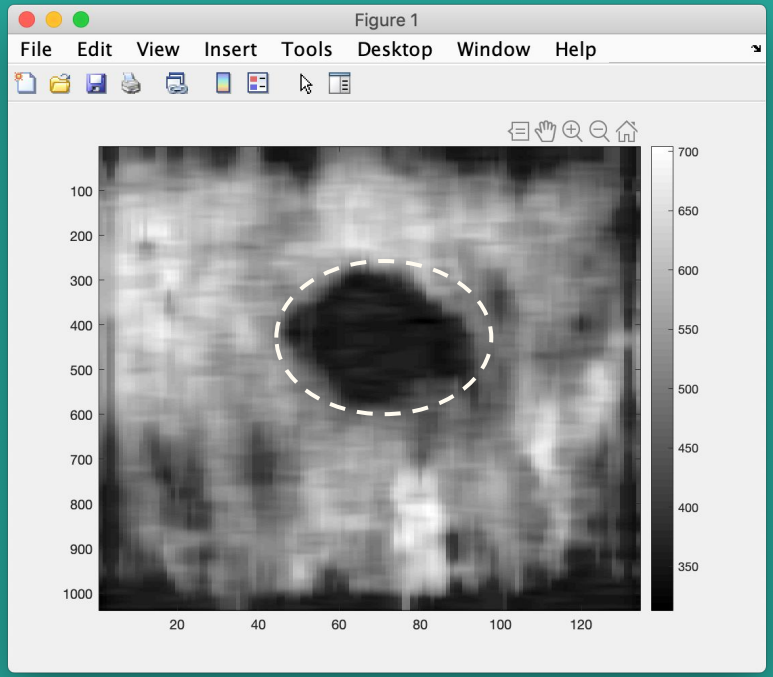


B-Mode Image

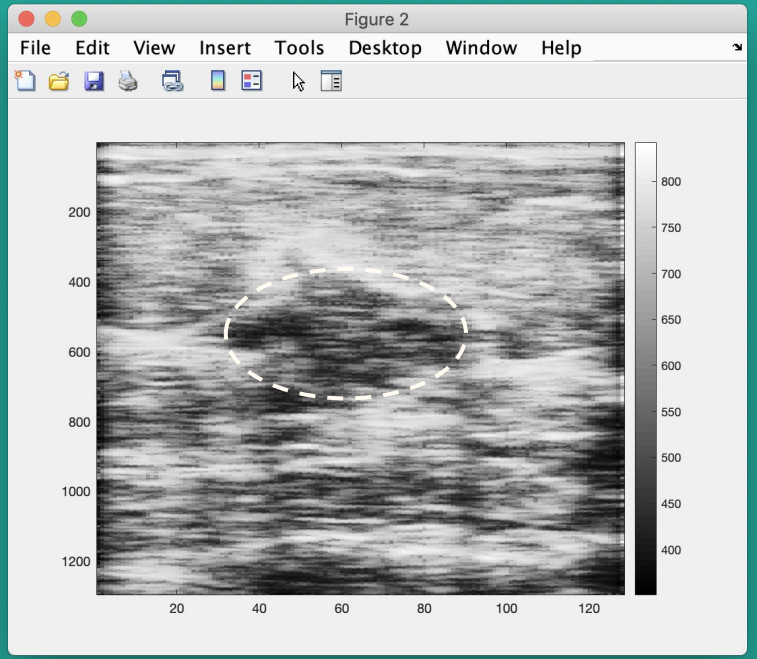


Final Processed Images

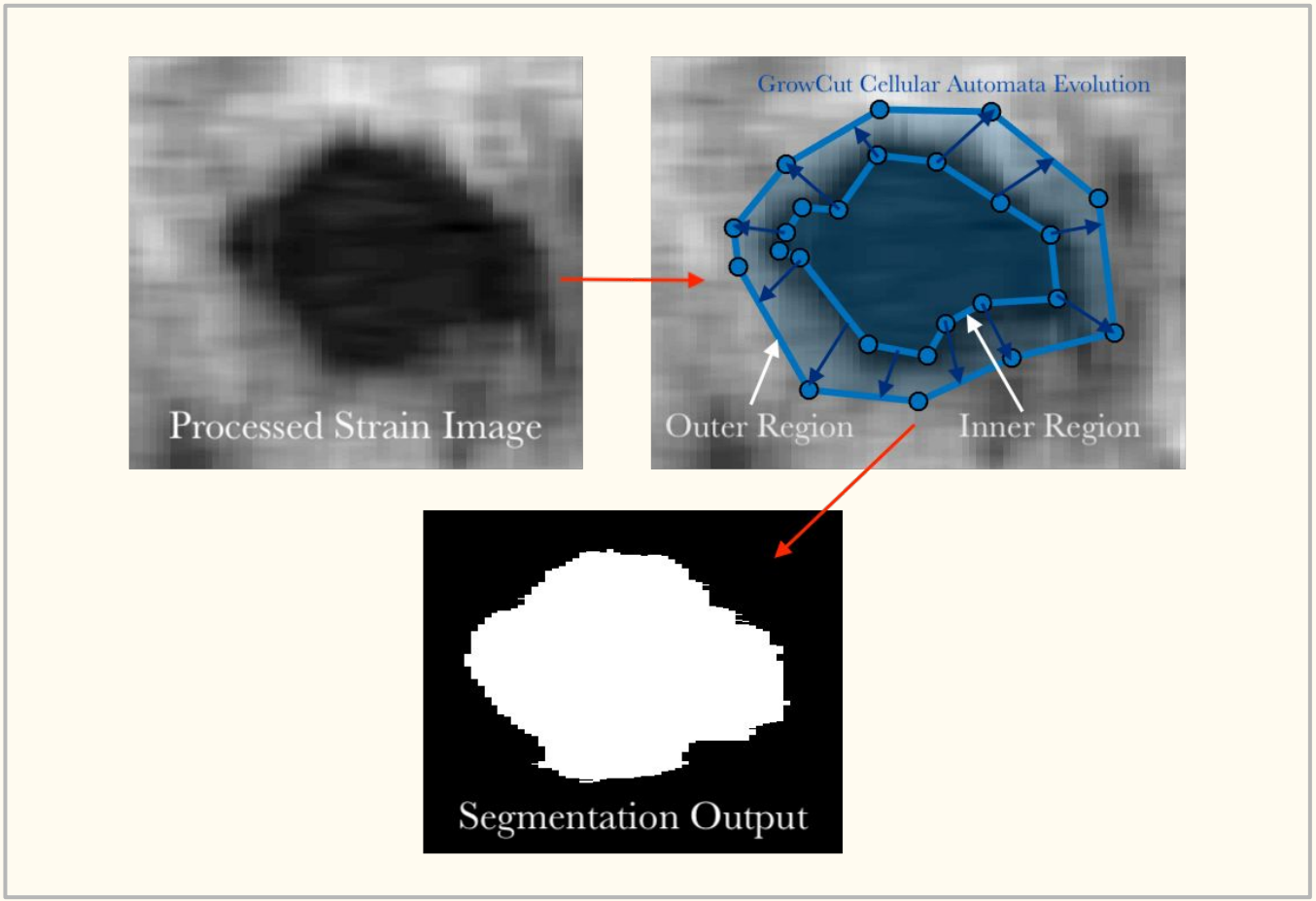
Strain Image



B-Mode Image



Label Selection



Methods — Segmentation, Features, Classification

- GrowCut/Active Contour Segmentation
- Calculated 16 features (area ratio & intensity contrast w/different parameters)
  - Area calculated in mm^2.
- Assessed classification performance by determining p-values for double-tailed t-tests between benign and malignant lesions
- Used several machine learning methods (e.g., SVM) to assess classification performance of these features.
  - Validated using ground truth (biopsy).

Methods — Segmentation, Features, Classification

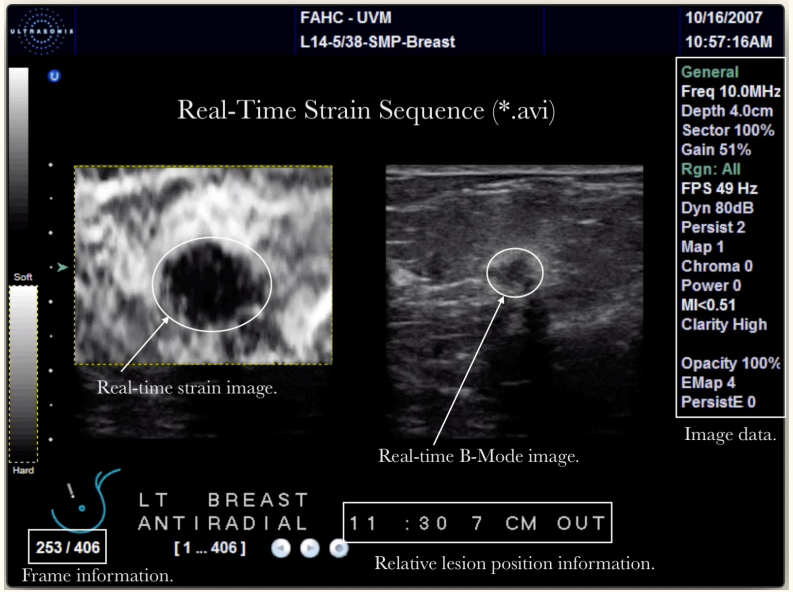
Feature	Processing	Segmentation	P-value %
Area Ratio	*Processed*	*GrowCut*	0.0008
	Processed	Active Contour	0.0022
	Unprocessed	GrowCut	0.0684
	Unprocessed	Active Contour	0.3900
Strain Contrast	Processed	GrowCut	0.0781
	Processed	Active Contour	0.2053
	Unprocessed	*GrowCut*	0.0188
	Unprocessed	Active Contour	0.0290
B-Mode Contrast	*Processed*	*GrowCut*	63.04
	Processed	Active Contour	94.38
	Unprocessed	GrowCut	70.67
	Unprocessed	Active Contour	82.72
Contrast Ratio	Processed	GrowCut	79.42
	Processed	Active Contour	74.25
	Unprocessed	*GrowCut*	6.588
	Unprocessed	Active Contour	14.71

** "Active Contour" refers to the segmentation by means of Active Contour for which empirical segmentation is processed by GrowCut.*

** Note that the table lists p-value percentages rather than p-values for the double-tailed t-tests between fibroadenoma and adenocarcinoma, thus, a p-value of 0.05 is equivalent to a p-value % of 5% on this table.*

Methods — Data

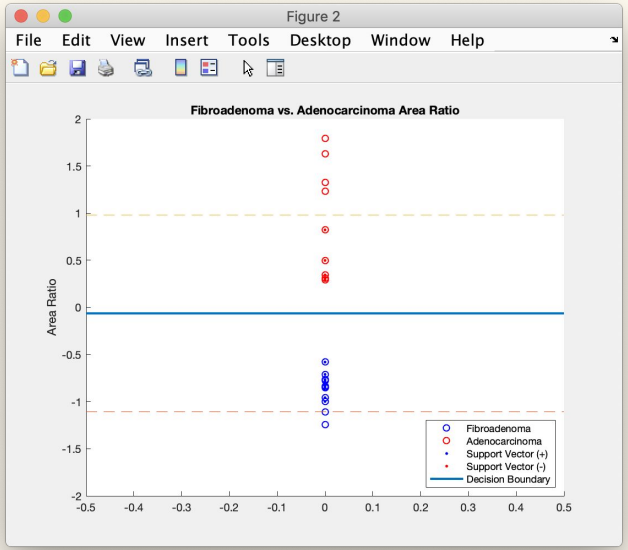
- *in vivo* patient data using Sonix-500RP w/ L14–5/38 probe @ University of Vermont Medical Center (UVM).
- Dataset of 48 patients of which 22 were used.
  - Fibroadenoma: 11
  - Adenocarcinoma: 11



Results

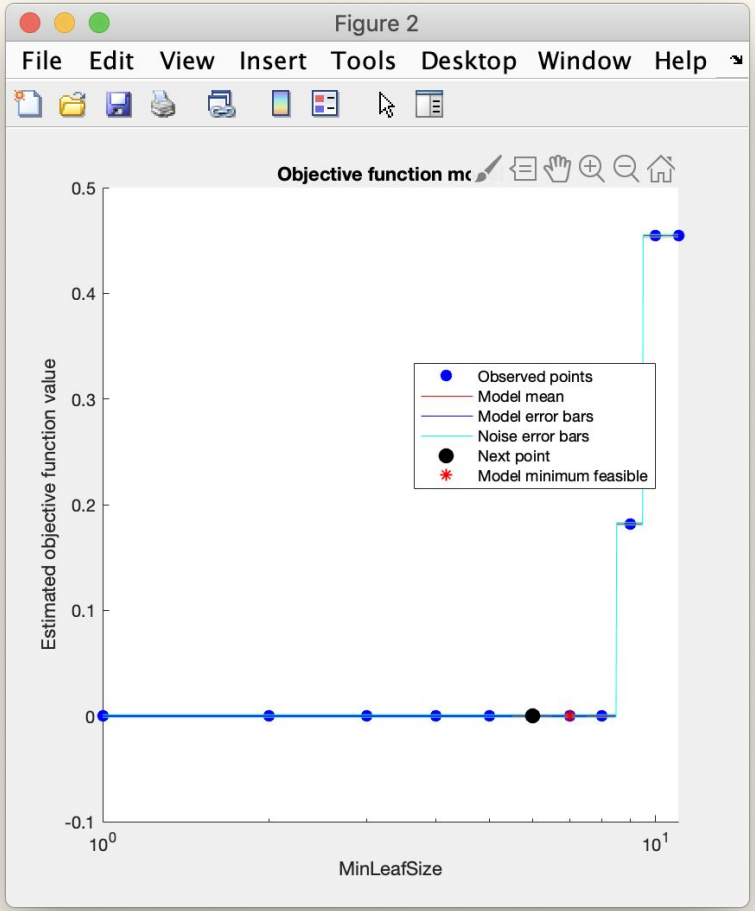
- Our results show that **area ratio is the most effective feature** ($p < 0.02$).
- Strain contrast, while less effective, can also help to discriminate between benign and malignant lesions.
- Support-vector machine (SVM) classifier yields accuracies of 100%, 81.82%, and 54.55% on the basis of area ratio, elasticity image contrast, and B-Mode image contrast, respectively.
- Validation assessed using Leave-one-out

Feature	Threshold	Accuracy (%)
Area Ratio	2.768960	100%
Strain Contrast	1.17×10^{8}	81.82%
Contrast Ratio	2.30×10^{18}	81.82%
B-Mode Contrast	1.2×10^{-5}	54.55%



Results

- Collectively implementing all 16 features as classification criteria in a Bayesian hyperparameter-optimization binary decision tree yielded a validation accuracy of 100% (between adenocarcinoma and fibroadenoma), the area ratio being the most significant criterion.



Conclusions

- Our semi-automated size-discrepancy method produces excellent discrimination between benign and malignant lesions, which will assist clinicians in discriminating breast lesions.

Future Work

- Applying the algorithm to independent datasets to more comprehensively evaluate the classifiers.
  - More difficult cases to more accurately reflect clinical experience e.g. where surrounding breast tissue is hard, single/multiple machine + multiple operators
- **Fully automating** the method using RegionGrowing and other automatic segmentation algorithms.
  - How many label selection points/how much granular noise is the algorithm able to handle