Ensemble Learning Breast Lesion Classification based on Magnetic Resonance Spectroscopic Imaging and Diffusion-Weighted Imaging

Marlene Lin[1], Ajin Joy[2], Andres Saucedo[2], and M. Albert Thomas[2]

University of California, Los Angeles (UCLA)

1 Department of Computational and System Biology

2 Department of Radiological Sciences, David Geffen School of Medicine

Introduction

Breast cancer screening and diagnostics

· One of the most prevalent cancers in females and one of the leading causes of cancer deaths worldwide.

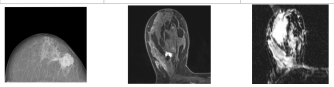
• Lifetime risk of 12.92% and an estimated 297, 790 new cases to be diagnosed in females in the U.S. in 2023.

• Early detection of malignancy before tumor metastasis outside the breast regions improve treatment outcomes.

• Small sizes and slow proliferation of early tumors — difficult to detect.

Existing breast cancer screening techniques

X-ray based	Magnetic Resonance Imaging (MRI) based	
Mammography	Dynamic-contrast Enhanced MRI (DCE-MRI)	**Diffusion-Weighted Imaging (DWI)**
· Gold Standard · Fast Scanning	· High sensitivity	· High sensitivity in conjunction with DCE-MRI ⟹ Multiparametric MRI (mpMRI)
· Increased false findings · Use ionizing radiation	· Wide range of **specificity** · Contrast potentially harmful	· Wide range of **specificity** · Not protocol for breasts



RESEARCH GOAL

Address the knowledge gap in using ensemble learning-based machine learning classifiers to classify breast tumors (malignant v.s. benign) using features from both multi-voxel, multi-dimensional MRSI and DWI data.

Methodology

Magnetic Resonance Spectroscopic Imaging (MRSI)

An efficient biochemical tool for quantifying metabolite and lipid concentrations in human tissues non-invasively.

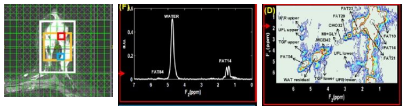
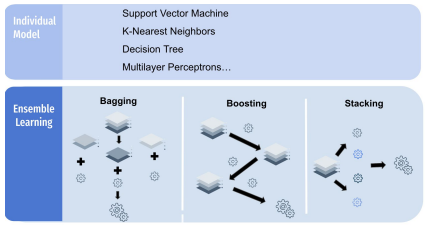


Figure 1. 5D Echo Planar-Correlated Spectroscopic Imaging (EP-COSI) scan of a benign lesion of a 32-year-old female. Left: Volume-of-interest placement; Middle: 1D spectrum (the red-outlined voxel); Right: 2D spectrum.

Datasets

· 5D EP-COSI and DWI data from 23 subjects
· malignant breast masses (mean age 53 [range:33–71] y.o.
· 17 benign breast masses (mean age 37 [range:19–60] y.o.
· 10~12 voxels are selected for each MRSI dataset based on the water-corrected choline map,
· 241 malignant and 195 benign voxels after outlier removal.

Models



Features

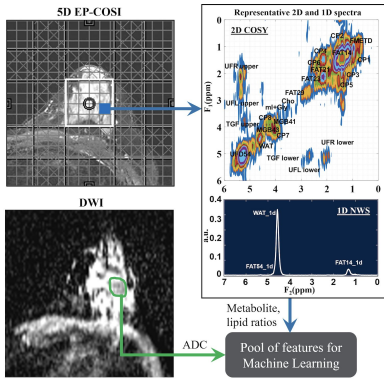


Figure 2. An example of 2D correlated spectroscopy (COSY) spectrum and Apparent Diffusion Coefficient (ADC) map. Localizer image for 5D EP-COSI acquisition in shown on the top left panel with ROI in white box. An extracted COSY spectrum and 1D non-water-suppressed (NWS) spectrum are on the right. Bottom-left panel shows the corresponding ADC map for the same subject with lesion marked in green.

Methyl Fat	FMETD
Methylene Fat	FAT
Water	WAT
Olefinic Fat	UFD
Methylene Glycceral Backbone	MGB
Cross-peak	CP
Unsaturated Fatty Acid Cross Peak	UFR(R: Right; L: Left)
Triglyceride Fat Cross Peak	TGF

Table 1. Metabolite and lipid acronyms

Results

Significant features

9 features, including ADC, ratios of FAT21(2.1-2.1), FAT23(2.3-2.3), CP2(1.32-0.9), CP6(2.06-1.32), CP8(4.2-3.9), and TGFUpper w.r.t. 1D FAT (1.4), ratio of FMETD w.r.t. 1D water, and the ratio of FAT21(2.1-2.1) with respect to 1D UFD were selected by statistical tests (p-value < 0.01) and recursive feature elimination (numbers in parentheses indicate peak locations).

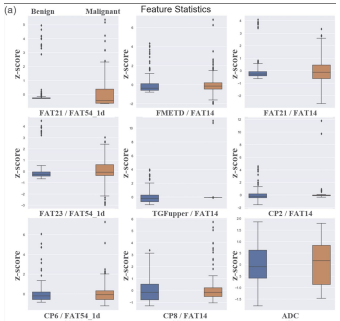


Figure 2. Box plots of the most important features (standardized).

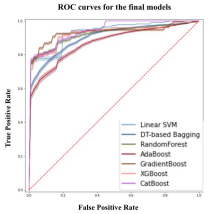


Figure 3. Receiver Operating Characteristic curves of different models

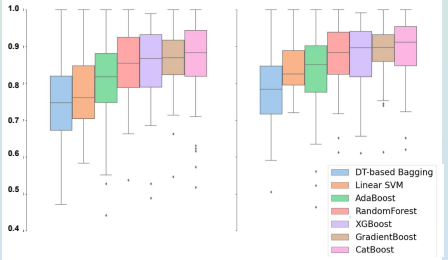


Figure 4. The box plots of accuracy and F1 scores of different models during the testing stage (rep. = 100).

Model	AUC (%)	Accuracy (%)	F1 score (%)
AdaBoost	88.54 ± 10.79	81.08 ± 10.81	83.46 ± 10.39
CatBoost	**95.17 ± 05.94**	86.93 ± 09.58	88.44 ± 09.87
DT-based Bagging	84.08 ± 11.88	74.61 ± 11.09	78.62 ± 09.71
GradientBoost	94.37 ± 06.35	**86.97 ± 07.90**	**89.39 ± 06.73**
Linear SVM	94.77 ± 06.44	77.43 ± 09.86	83.94 ± 06.37
RandomForest	92.24 ± 07.30	85.31 ± 09.15	87.59 ± 08.06
XGBoost	92.98 ± 08.18	85.65 ± 09.61	87.64 ± 08.58

Table 2. Model performance, best scores in bold.

Model Comparison

Dataset was split into 80% for training, 20% for testing and cross-validated using the shuffled grouped 5-fold approach.

One-way ANOVA + Tukey's post-hoc tests showed CatBoost, Gradient Boost, XGBoost, Random Forest to be the best-performing methods from repeated testings with randomized model initializations and data splits.

Discussion

· Unused features from the imaging data:
Incorporate more image-based radiomics features from DWI, or potentially DCE-MRI

· Limited datasets and data source:
Collect more datasets, ideally from different machines/sites, to ensure model generalizability

· Classification task:
Use models to classify other types of tumors and tumors of different grades aiming at higher specificity by establishing connections between MRSI spectra quantitations and cancer metabolism

Conclusion

· Highlighted the potential of using features from both five-dimensional 5D EP-COSI and DWI for accurate breast lesion classification.

· Gradient Boost-based ensemble learning and random forest classifiers achieved accuracies of 85 - 87% and F1 scores of 87 - 89%.

· ADC values from the DWI and lipid cross-peaks unique to 2D COSY spectra were identified as important features.

Dr. Melissa Joines, Dr. Stephanie Lee-Felker and Dr. Maggie DiNome were responsible for patient recruitment. This work was supported by a CDMRP grant from the US Army Breast Cancer Research 340 Program:# W81XWH-16-1-0524.