An Analysis of

Breast Cancer Diagnostic Modalities

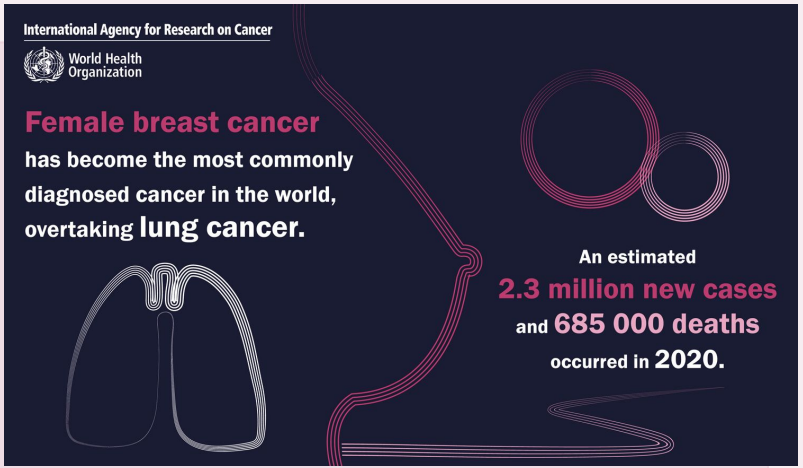
Team 11

Shijie Tang · Zhenhao Liu · Jiaqi Li
Zhijian Chen · Juliana Miller

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Department of Electrical and Computer Engineering

Motivation

- One of the **most commonly diagnosed** cancers among women worldwide and a major contributor to the global cancer morbidity
- Ranked **second in overall incidence** in women (WHO, 2022)
- **Second leading cause of cancer death** among women in the United States (CDC, 2025)



Source: https://www.iarc.who.int/featured-news/world-cancer-day-2021

Sources: https://www.who.int/news-room/fact-sheets/detail/breast-cancer, https://www.cdc.gov/breast-cancer/statistics/index.html

What We're Exploring

Which breast cancer diagnostic modality proves to be the most effective?
Which cell characteristics are highly correlated to malignancy?

- Identify which features most strongly distinguish malignant from benign tumors across different diagnostic modalities
  - Quantify how malignancy rates vary with tumor characteristics
  - Identify malignancy-metabolic marker trends

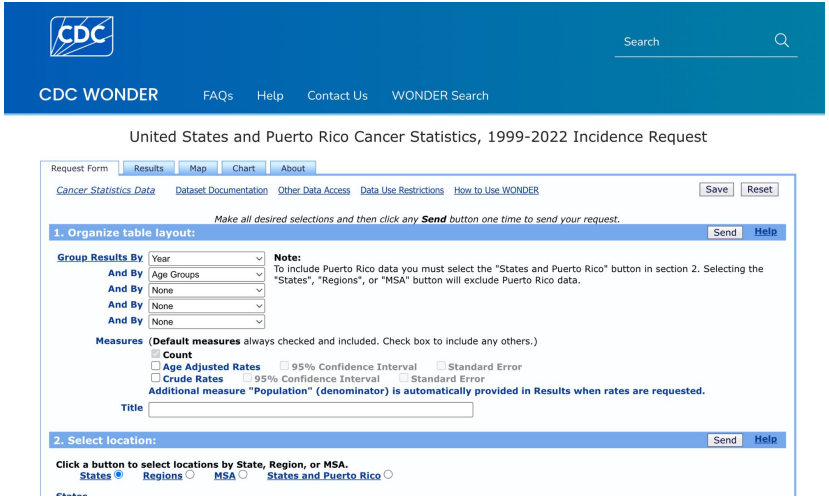
Data Overview

University of Wisconsin Hospitals, Kaggle
Cytological features from microscopic examination of samples obtained from surgery (e.g.Uniformity of Cell Size).

University of Wisconsin Hospitals, Kaggle
Morphological features computed from digitized fine needle aspirate (FNA) images (e.g. Radius, Texture).

UC Irvine Breast Cancer Coimbra Dataset, Kaggle
Blood-based metabolic and anthropometric biomarkers (e.g. Age, Glucose).

CDC Wonder Database, Cancer Statistics Incidence Report
Female breast cancer incidences from 1999-2022 grouped by year and age groups.



Methodology

Breast Cancer Wisconsin (Original) Data Set

	Clump Thickness	Uniformity of Cell Size	Uniformity of Cell Shape	Marginal Adhesion	Single Epithelial Cell Size	Bare Nuclei	Bland Chromatin	Normal Nucleoli	Mitoses	Diagnosis	Is_Malignant
0	5	1	1	1	2	1.0	3	1	1	2	0
1	5	4	4	5	7	10.0	3	2	1	2	0
2	3	1	1	1	2	2.0	3	1	1	2	0
3	6	8	8	1	3	4.0	3	7	1	2	0
4	4	1	1	3	2	1.0	3	1	1	2	0

Breast Cancer Wisconsin (Diagnostic) Data Set

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean
0	1	17.99	10.38	122.80	1001.0	0.11840	0.27760
1	1	20.57	17.77	132.90	1326.0	0.08474	0.07864
2	1	19.69	21.25	130.00	1203.0	0.10960	0.15990
3	1	11.42	20.38	77.58	386.1	0.14250	0.28390
4	1	20.29	14.34	135.10	1297.0	0.10030	0.13280

Database

.csv files

Data Profiling and Cleaning

Missing values, duplicates, empty cells, and useless columns

Preprocessing

Target Variable Standardization
0 → Healthy Patient
1 → Cancer Patient

Analysis Overview

Three different diagnosis modalities:

1. Microscopy-based morphology
2. Imaging-derived morphology
3. Metabolic biomarkers

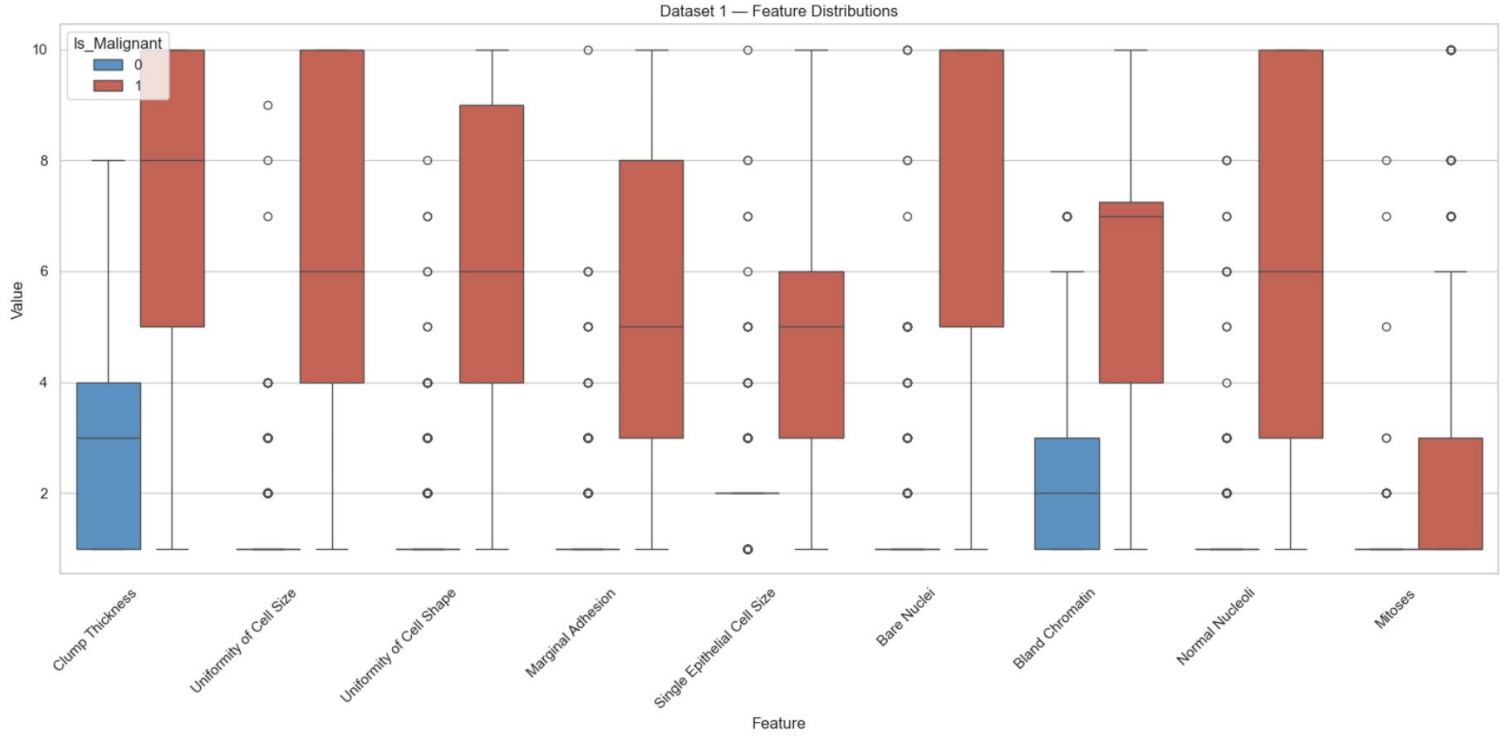
Invasive

Non-invasive

1. Microscopy-Based Morphology

A tissue sample obtained through a biopsy is examined under a microscope by a pathologist.

Morphology of the Cell

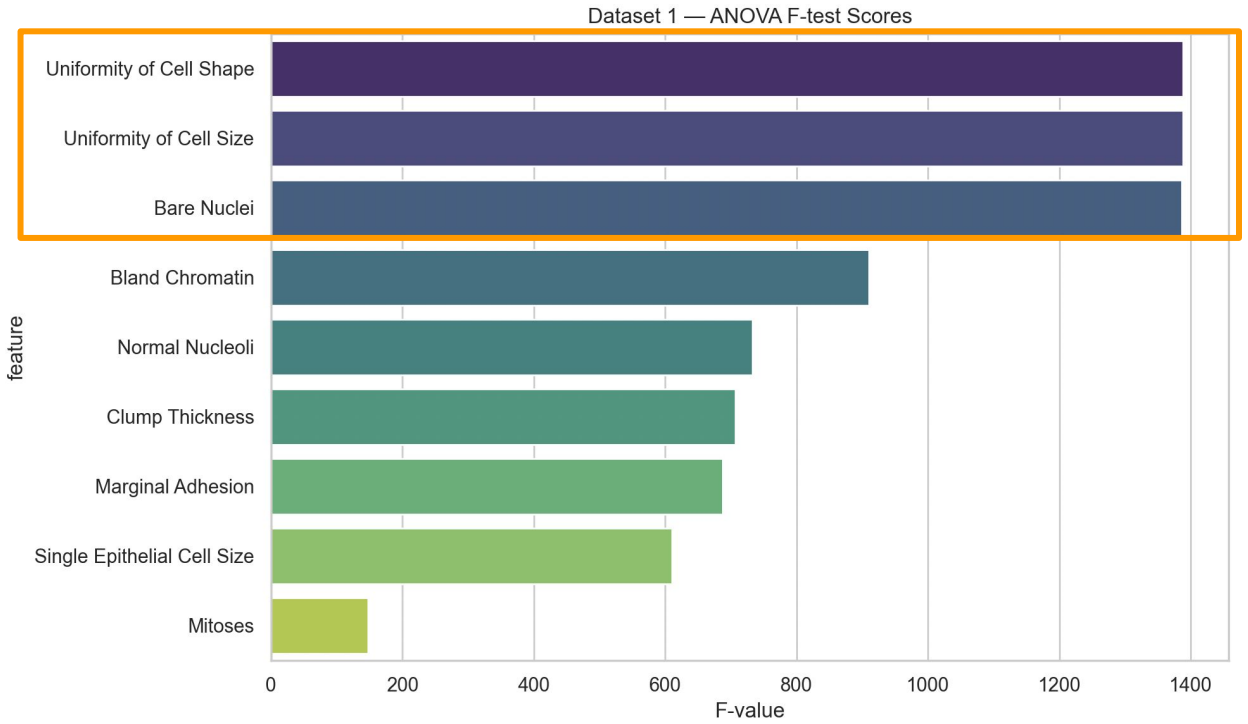


Morphology of the Cell

Normal tissue: uniform and organized

Malignant tissue: chaotic and disorganized

Cancerous cells **divide rapidly** without proper regulation resulting in **highly variable** cell sizes and shapes.



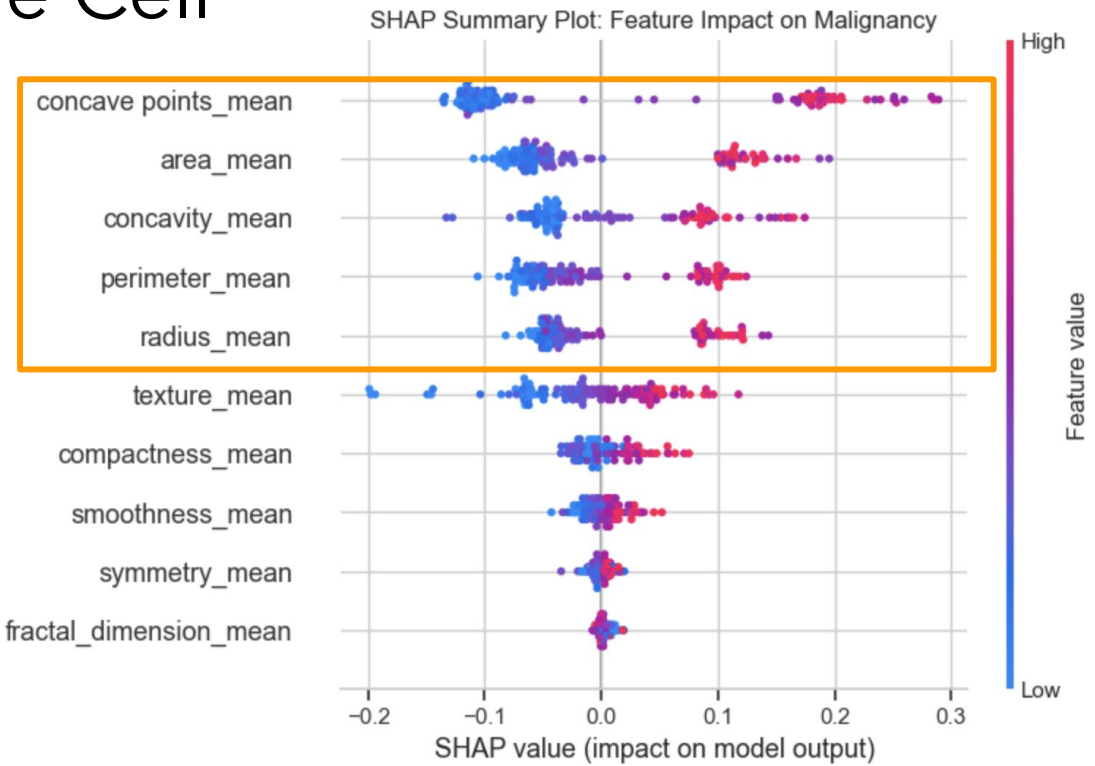
2. Imaging-Based Morphology

A sample of cells obtained using FNA followed by digital image analysis where algorithms extract quantitative features from cellular images.

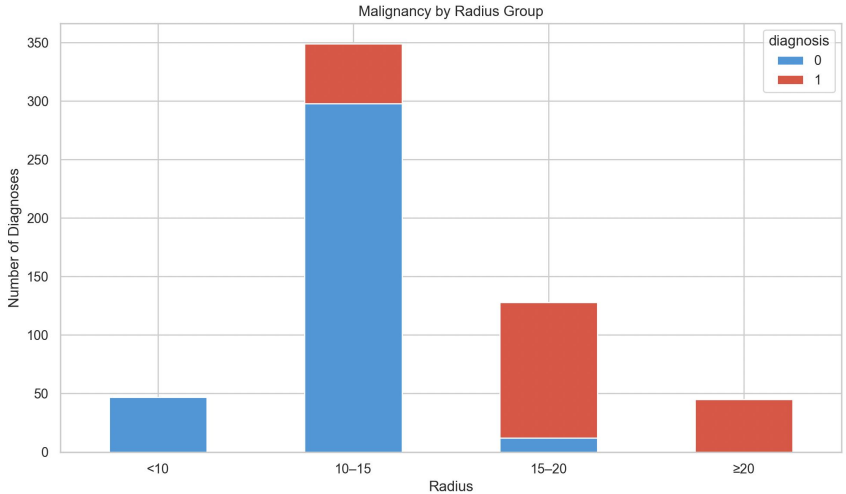
Morphology of the Cell

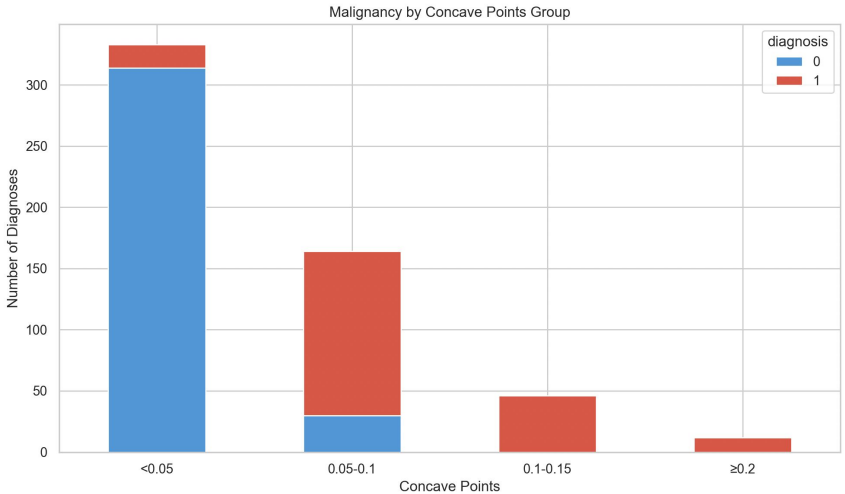
Concave Points: Number and severity of "dents" or "notches" in the tumor perimeter

- Reflects **how irregular and complex** the tumor boundary is
- Malignant tumors have **invasive growth pattern** and **loss of cohesion**



Morphology of the Cell





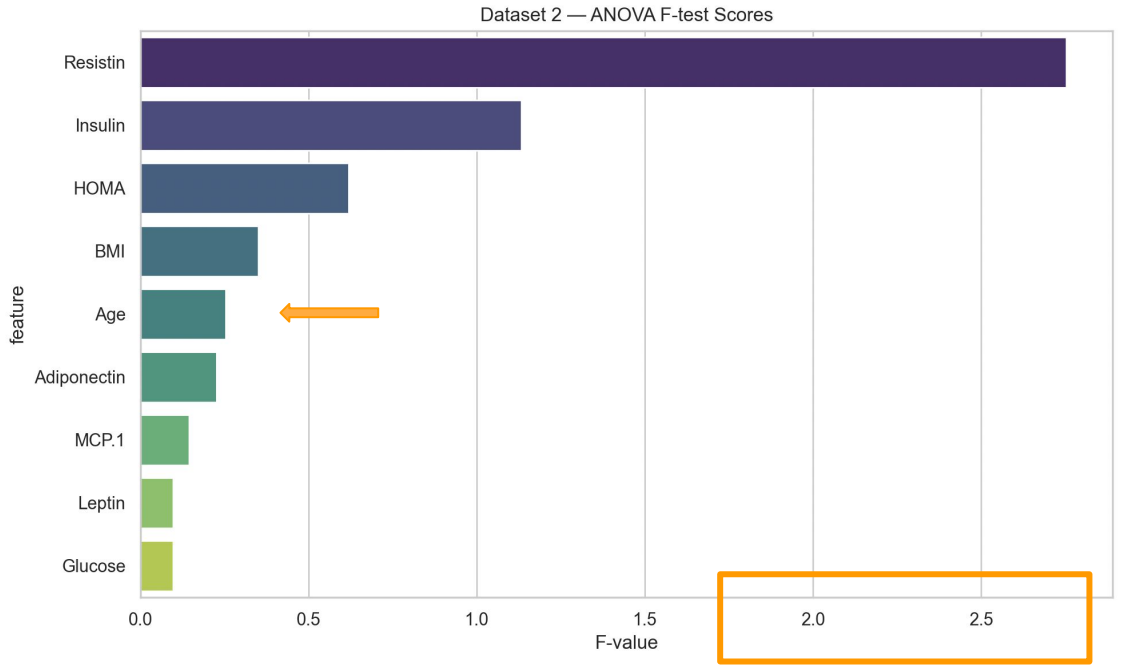
3. Metabolic Biomarkers

Non-invasive screening approach using blood tests to measure metabolic and inflammatory markers. Potential to identify at-risk individuals before clinical symptoms appear.

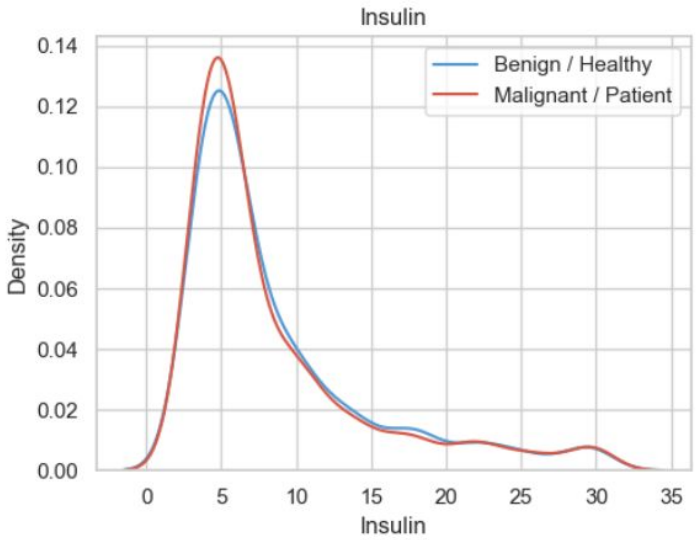
Metabolic Characteristics

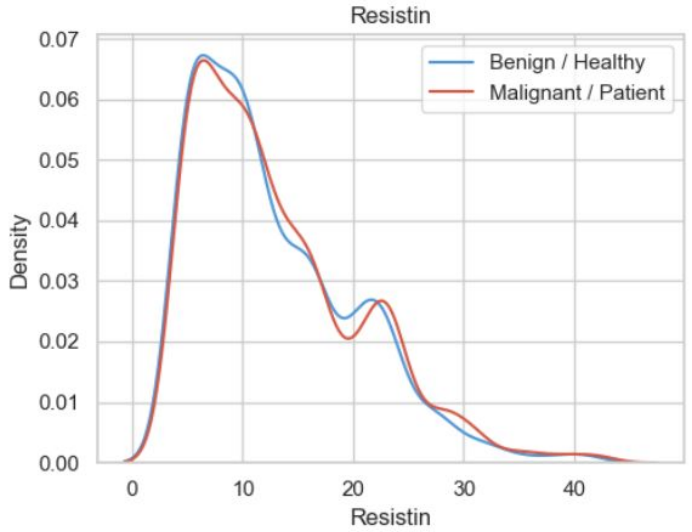
Low F-values from biomarkers, low efficacy for determining malignancy.

- Resistin is hormone released by fat tissue that prompts inflammation



Metabolic Characteristics



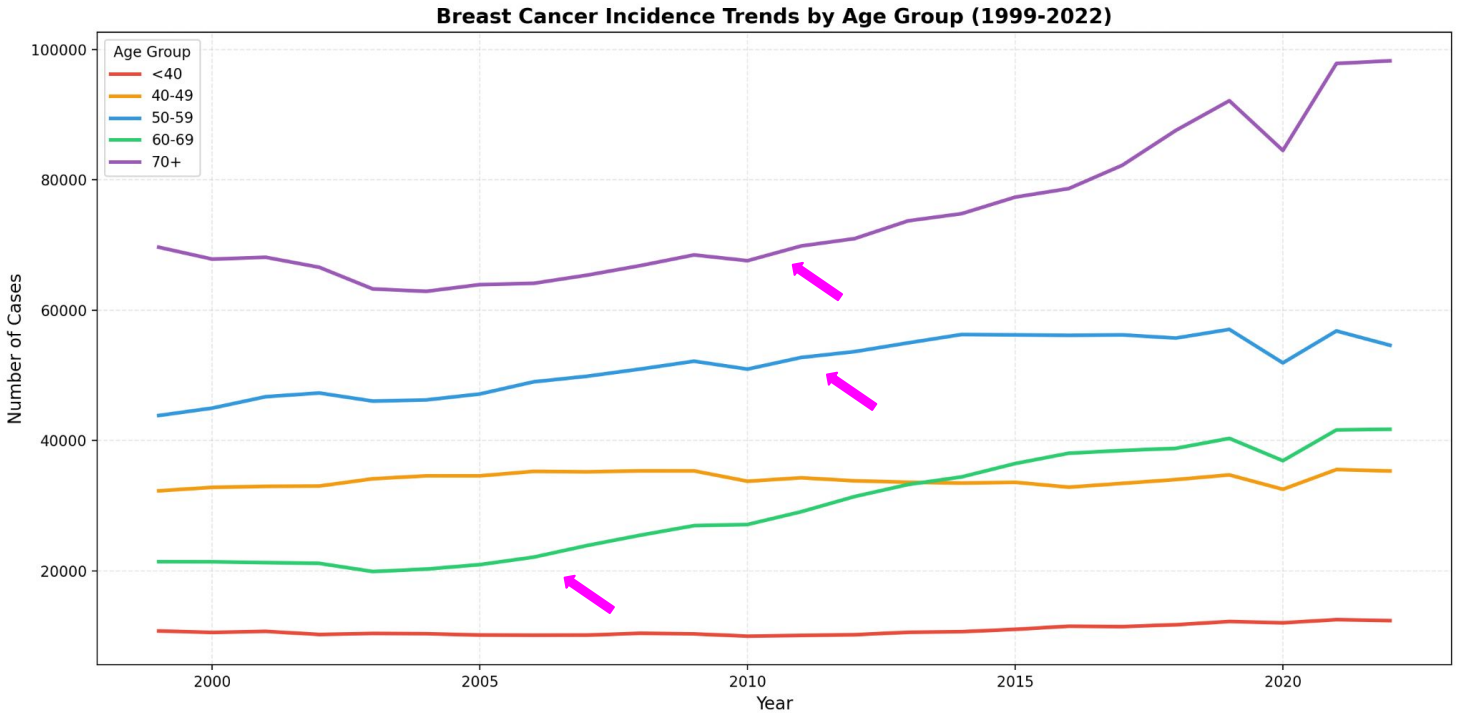


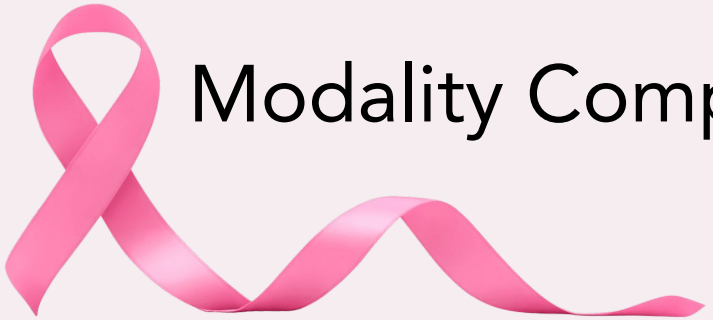
Large overlap of benign and malignant lines.

Metabolic Characteristics

Breast cancer incidences have **increased over time in patients age 50+**.

50+ age group is at potentially higher risk of diagnosis.

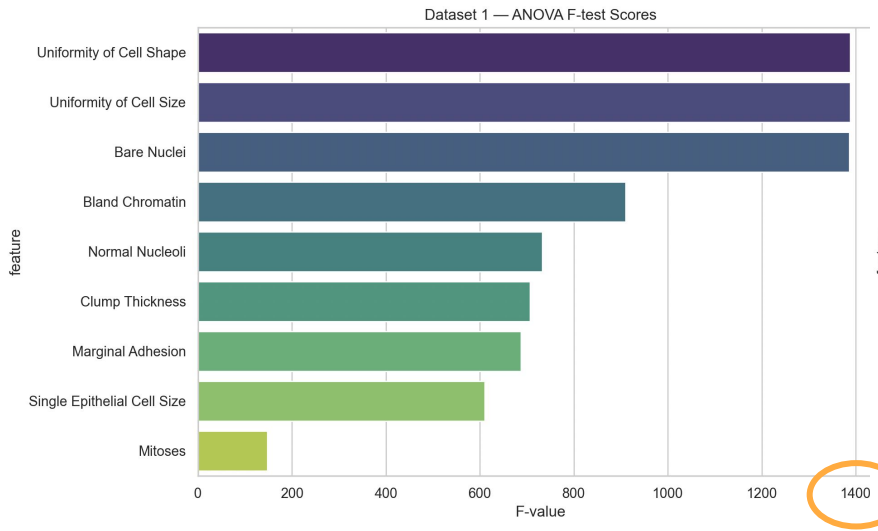




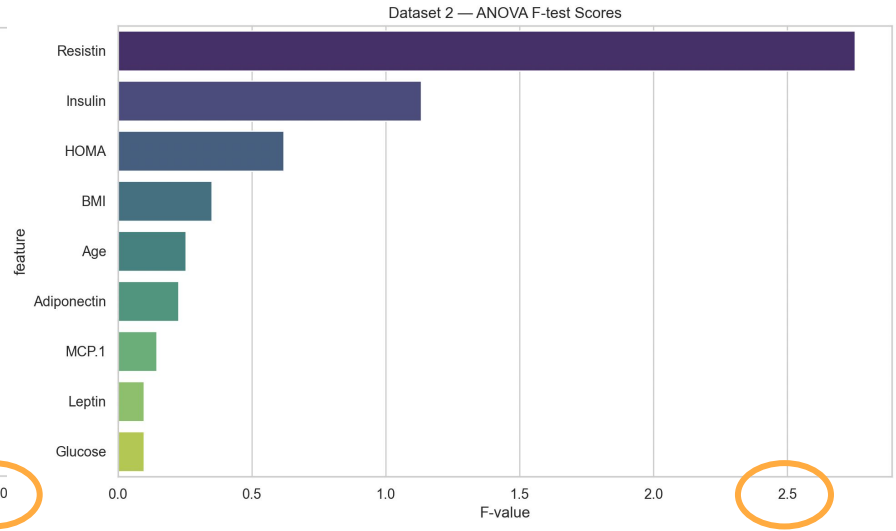
Modality Comparison

Comparison of F-Scores

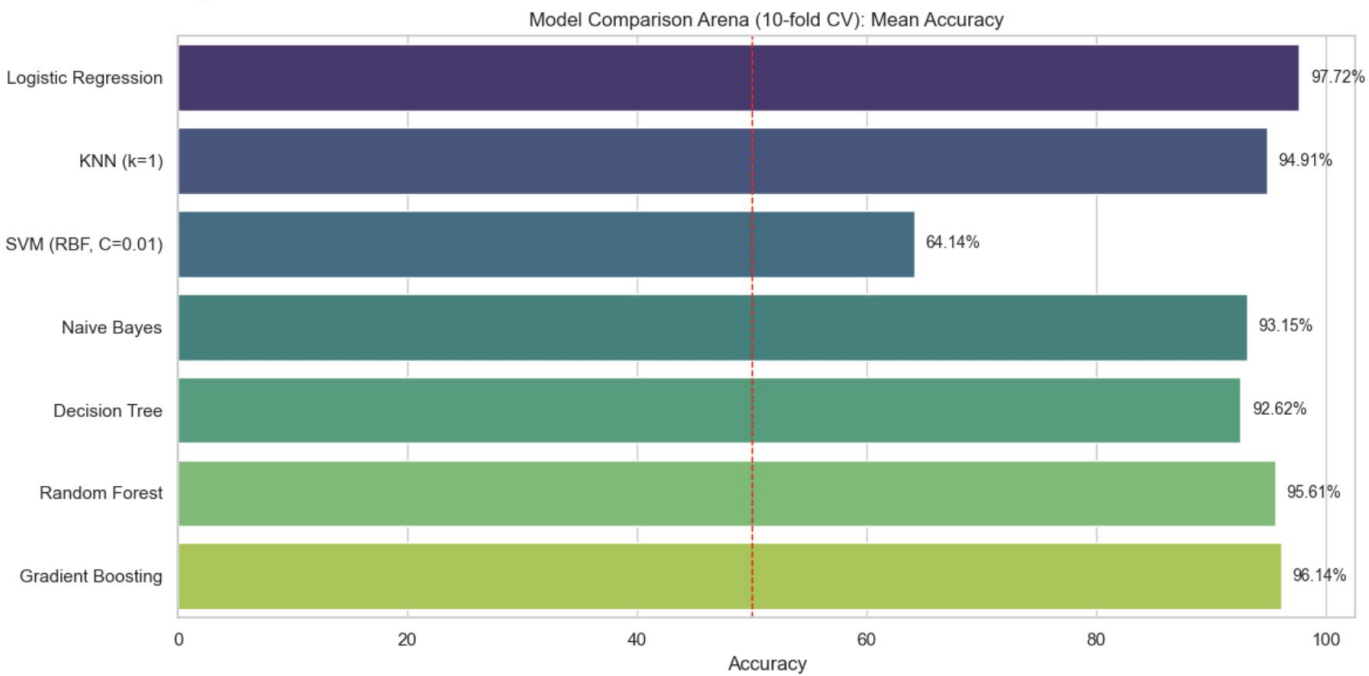
Microscopy-based Morphology



Metabolic Markers



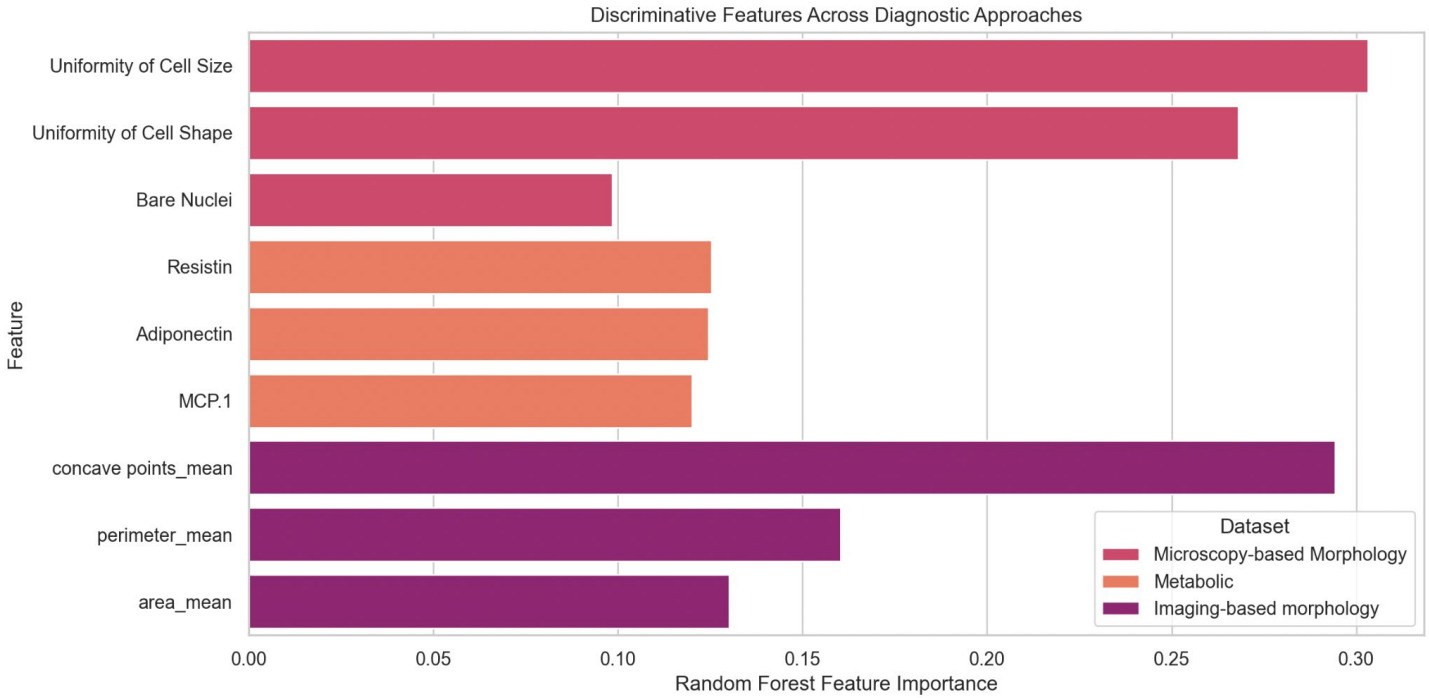
Model Comparison



Modality Comparison

Top 3 Features from each modality (dataset).

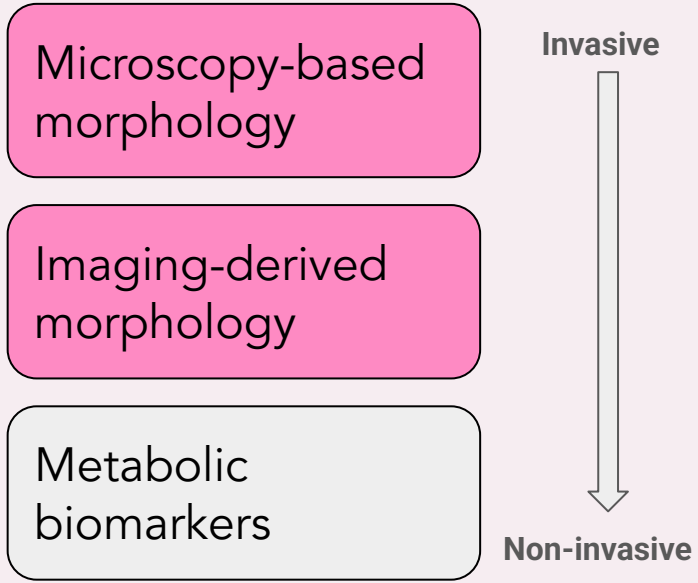
Both **morphology modalities outperform** the **metabolic biomarkers.**



Conclusion

1. Microscopy-based and imaging-based morphology **vastly outperform** metabolic markers for breast cancer diagnosis.

2. **Most discriminative features** are cell uniformity and tumor border irregularity.



Thank you for listening!