

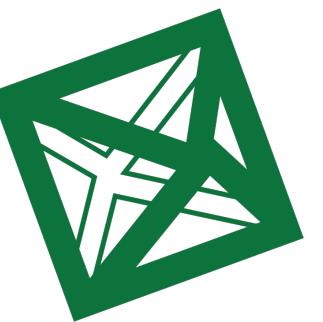


DESIGN OF CADe

DIGITAL SIGNAL AND IMAGE MANAGEMENT PROJECT

Bonardi Pietro - 859505
Ravazzi Lucia - 852646

Università degli studi
di Milano-Bicocca



INTRODUCTION

MAMMOGRAPHY

WHAT?

Examination of breast for diagnosis and screening.

HOW?

Low-energy X-ray create b/w images for CC and MLO views.

WHY?

Early detection of breast cancer.

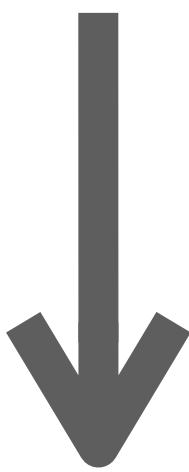
WHO?

Female after 40.



DATASET DESCRIPTION

For each patient
(total equal to 1950)



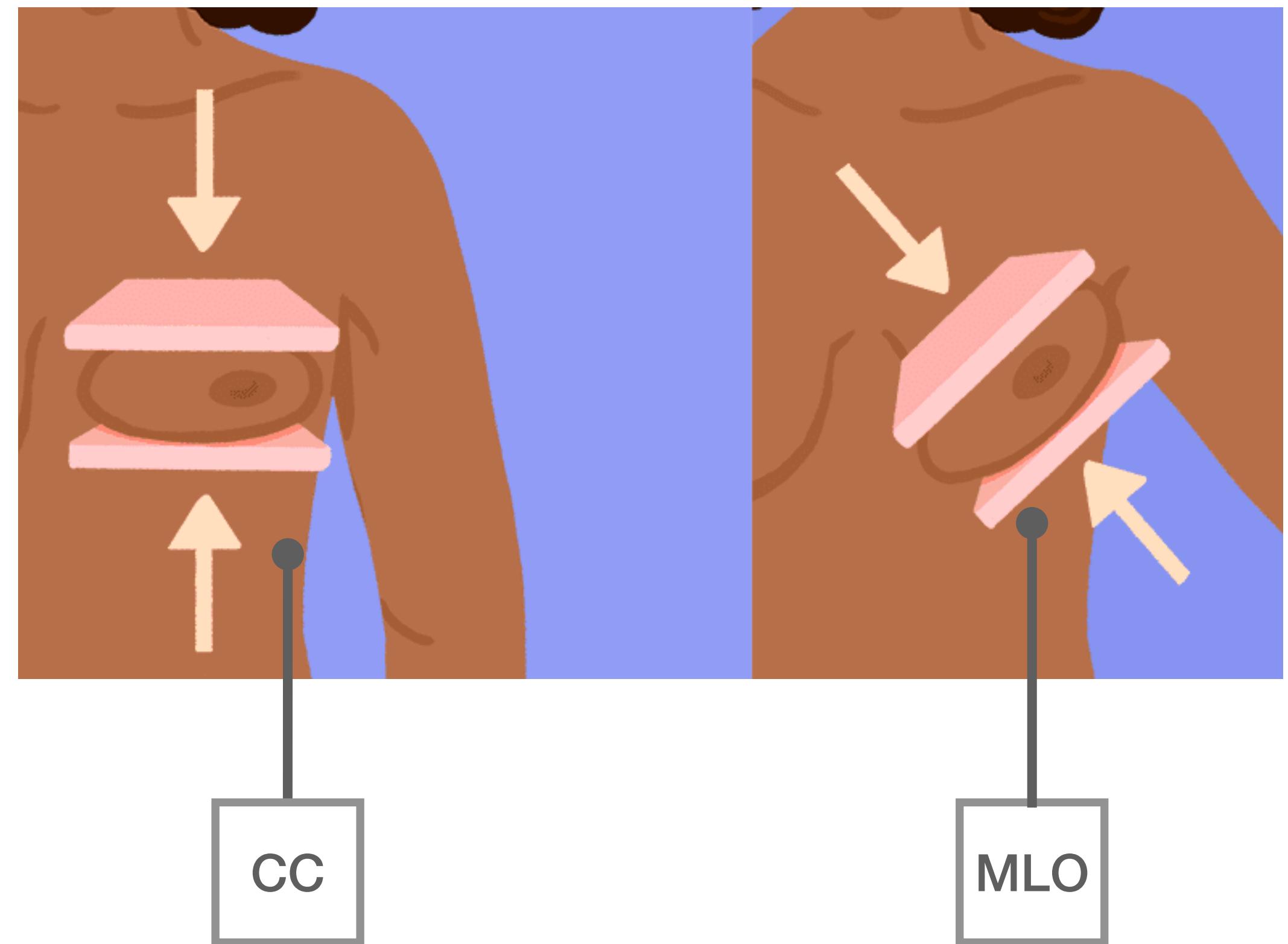
DATA

Age, Density of
tissue, etc.

+

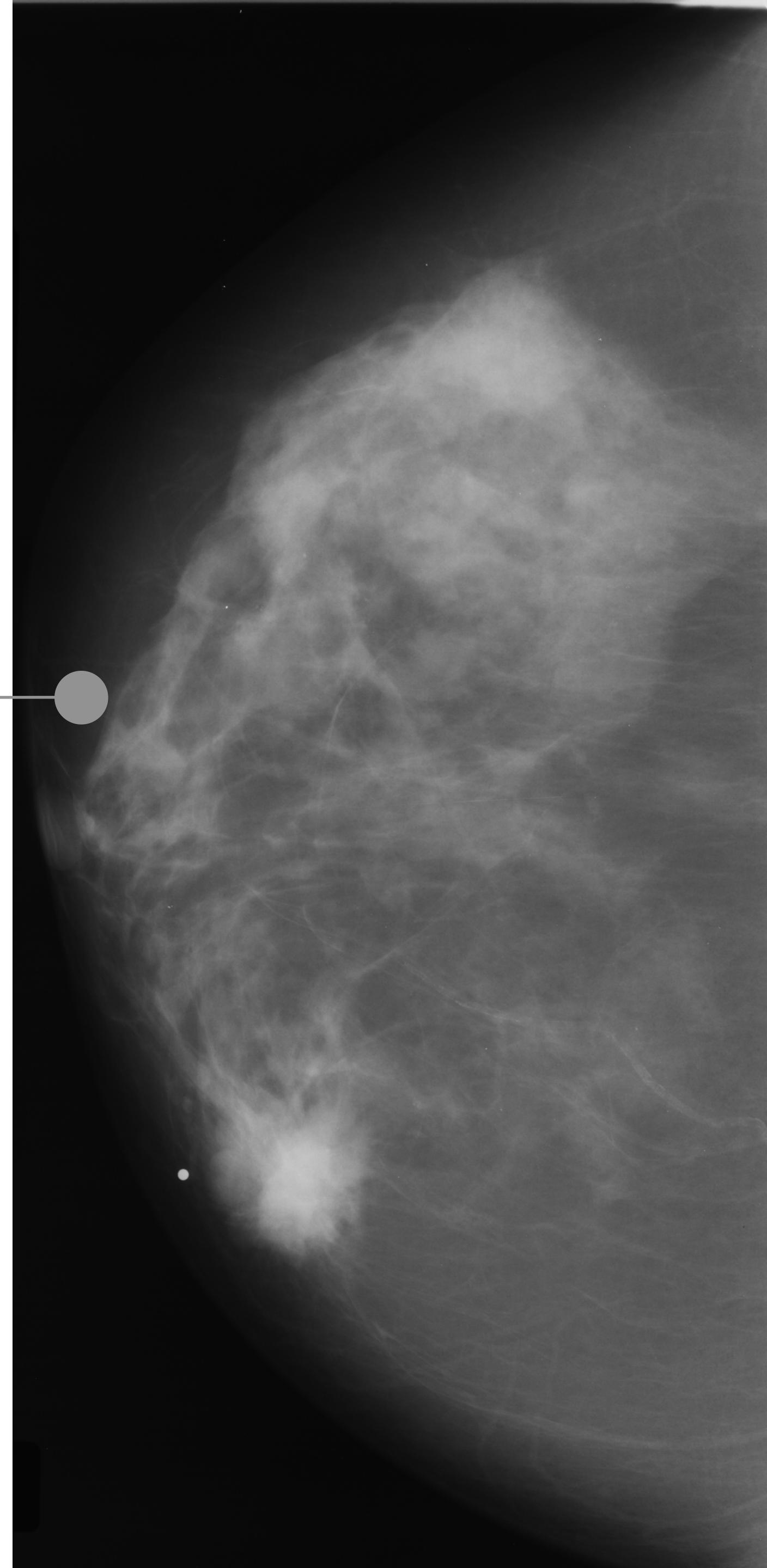
IMAGES

Left - [CC, MLO]
Right - [CC, MLO]

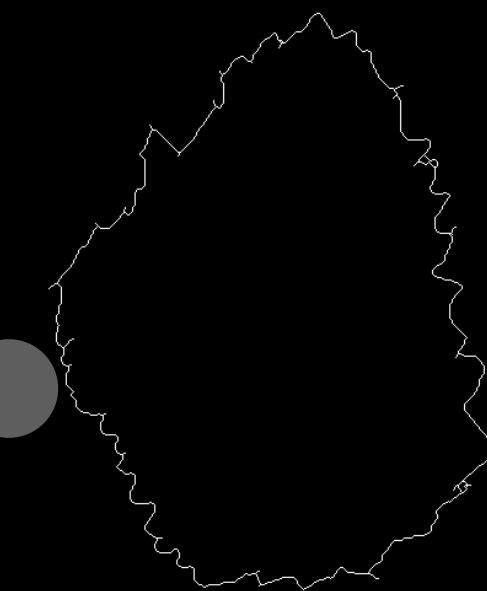


Malignant Cancer

right breast,
CC view



Mask



Tasks

- 1 | Classification
- 2 | Content Based
Image Retrieval

SEGMENTATION PROBLEM

HOW TO EXTRACT
THE DANGEROUS
ZONE?



If a **benign** or **malign** cancer occurs, each suspect breast have two **masks** (CC and MLO) provided by radiologist.

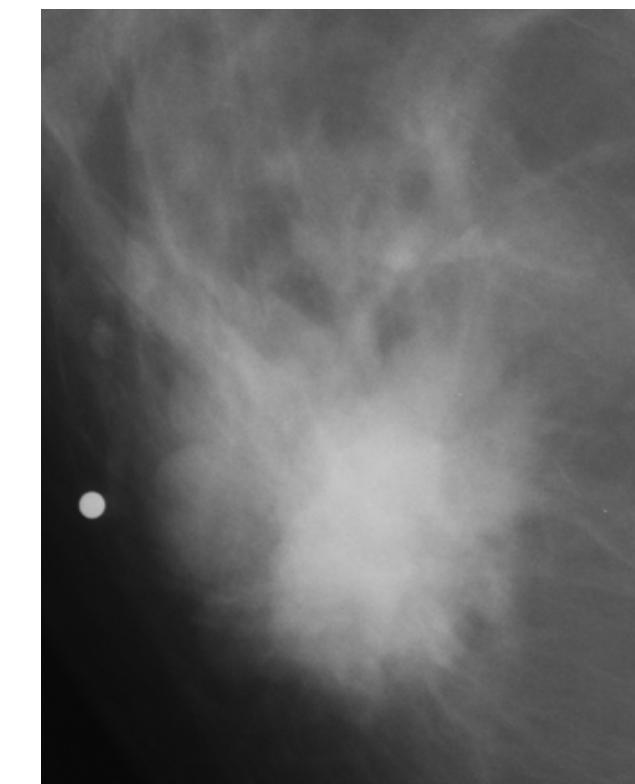
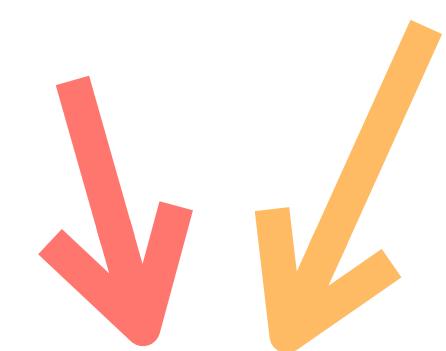
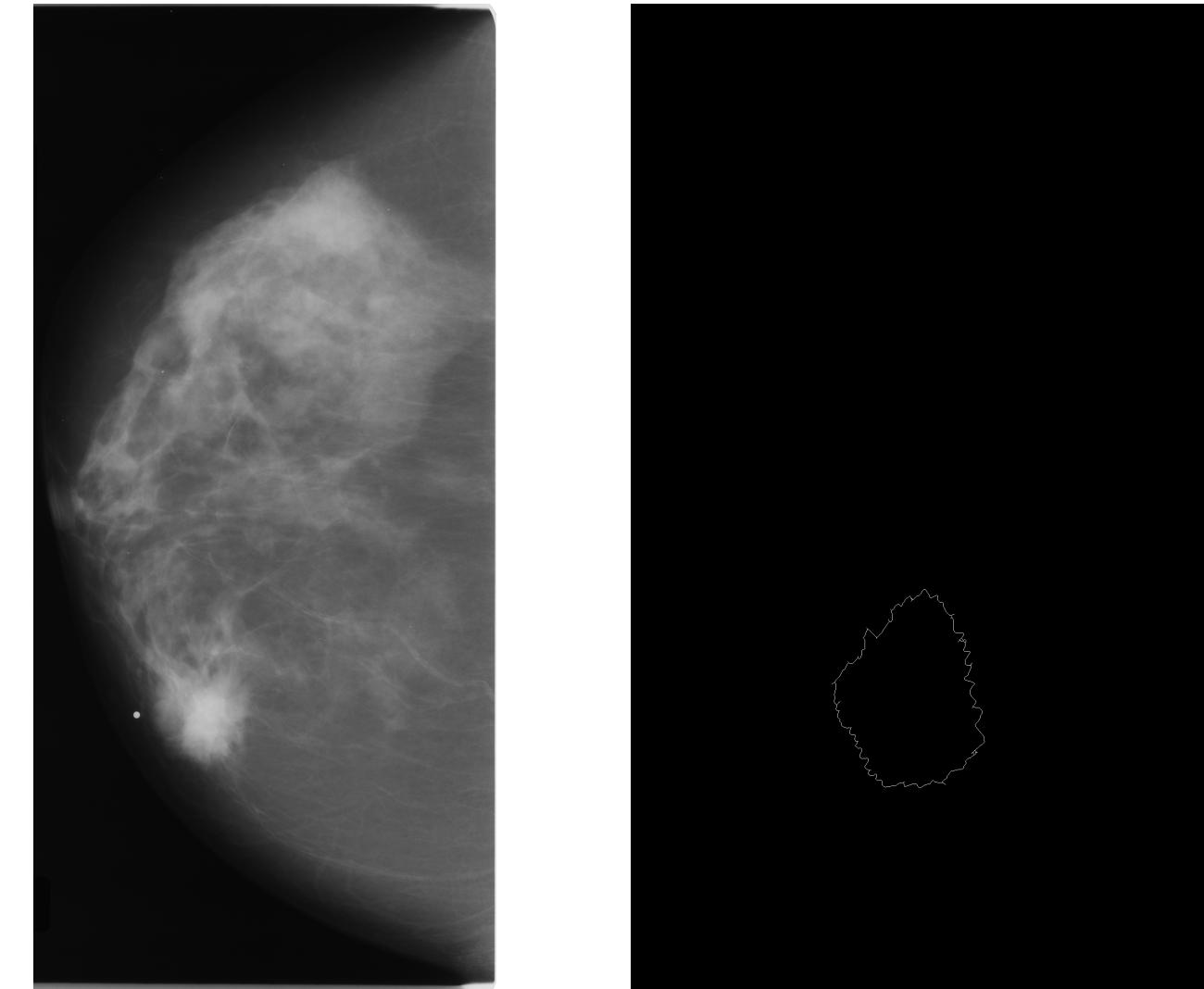
PREPROCESSING

1

Selection of all images provided by mask.

2

Crop images to select the suspicious zone.



CLASSIFICATION



TASKS

1

Classification

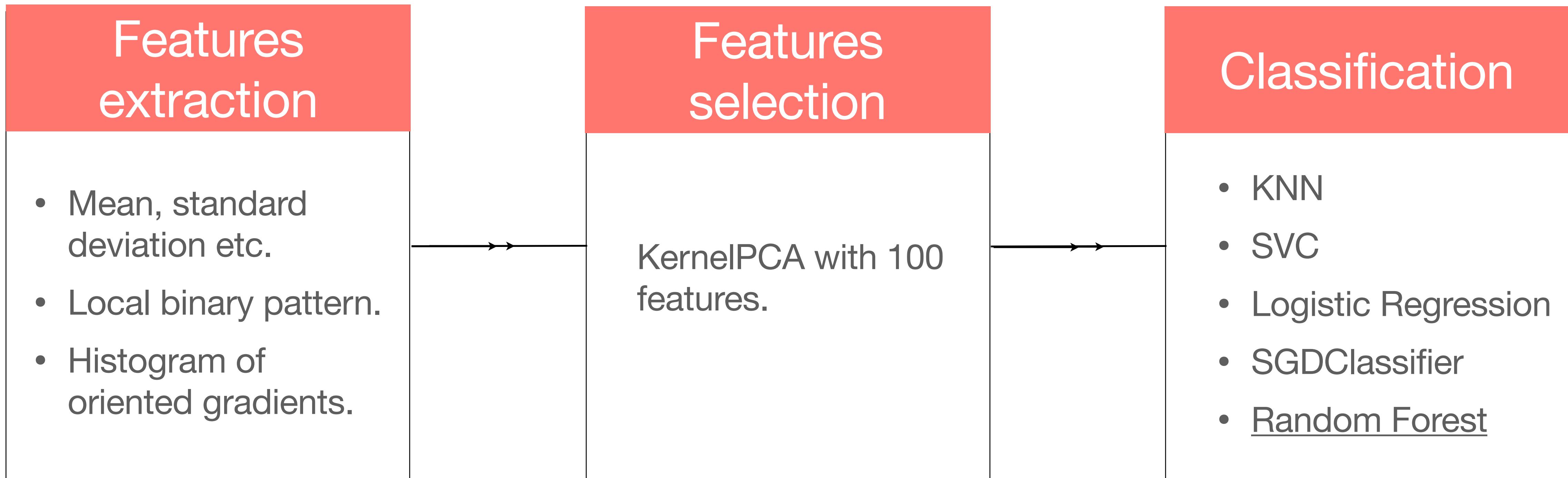
given an image with breast cancer,
is it malignant or benign?

2

CBIR

given an image with breast cancer,
what are the most similar previous cases?

HANDCRAFTED FEATURES



CNN AND PURE CLASSIFICATION

1

Features extraction
with VGG16 at
block4_pool layer

2

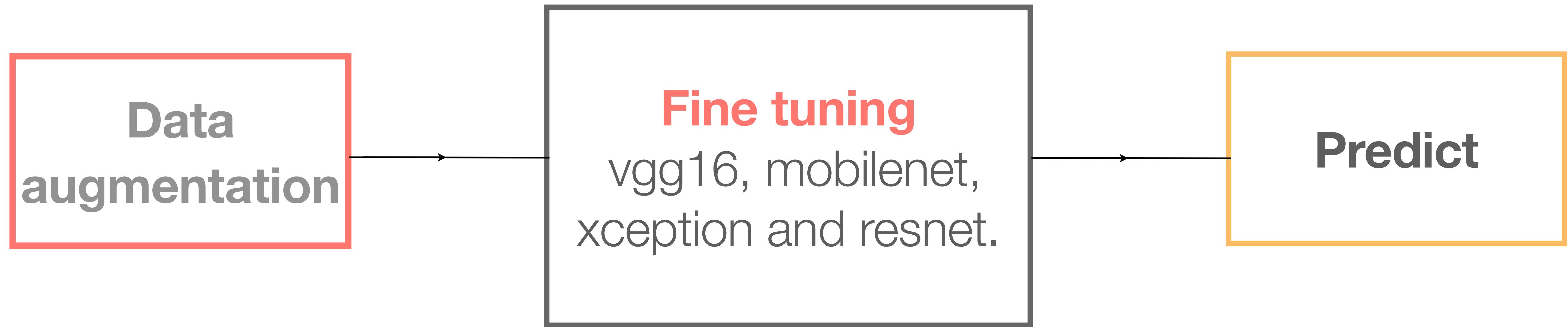
Train Naive Bayes,
Logistic Regression
and **SGDClassifier**

3

Test
models



FINE TUNING



Best results: VGG16: accuracy ~ 71%



CBIR

Tasks

1

Classification

given an image with breast cancer,
is it malignant or benign?

2

CBIR

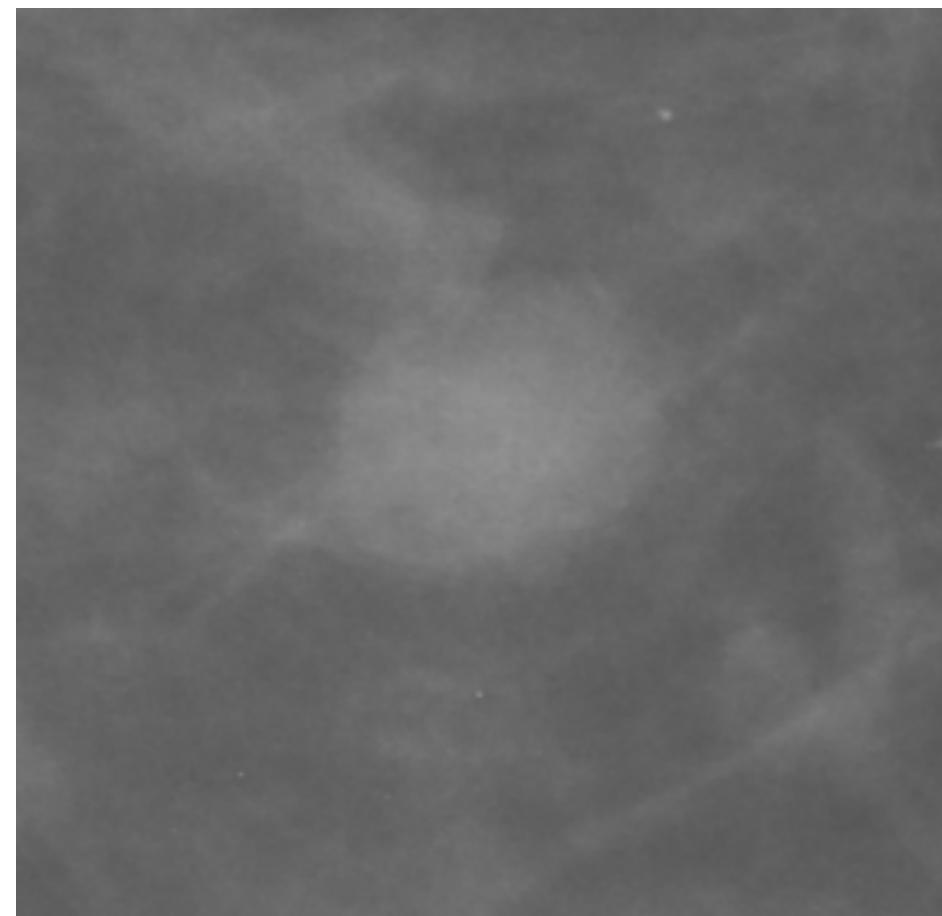
given an image with breast cancer,
what are the most similar previous cases?

CONTENT BASED IMAGE RETRIEVAL

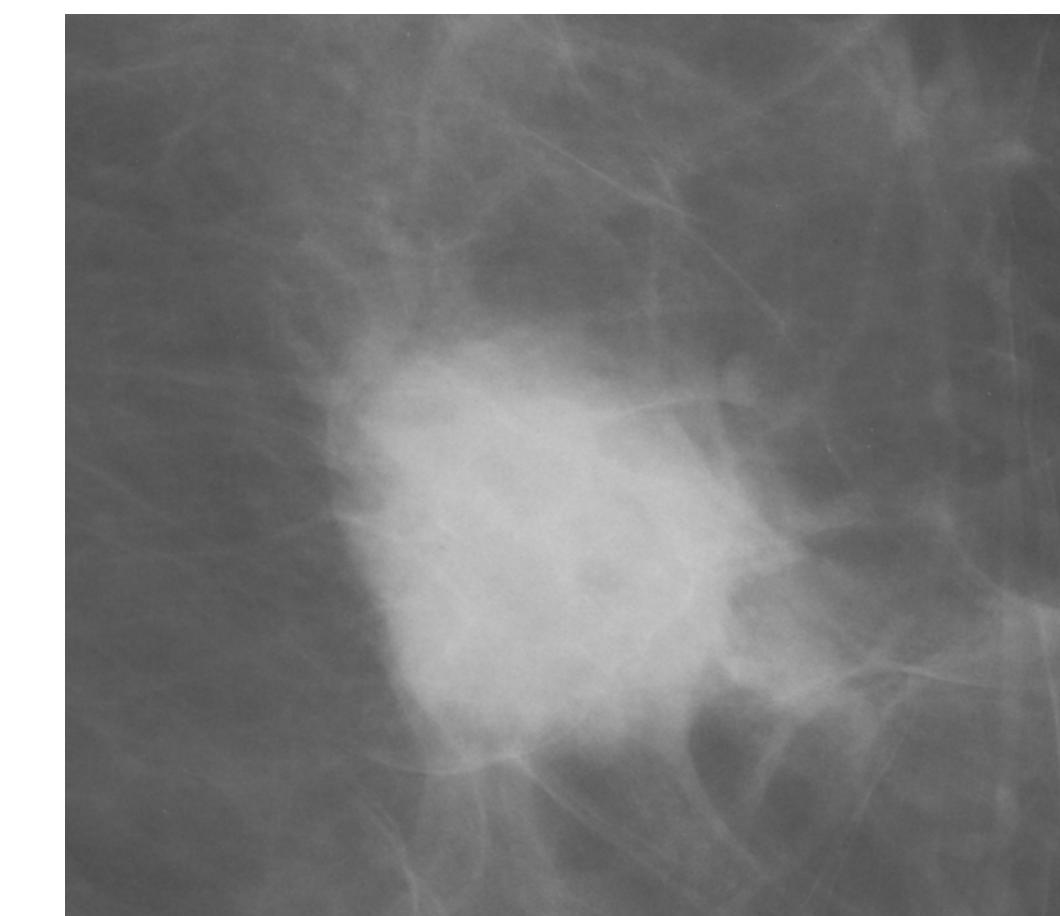
- 1** Take the best CNN from the classification task: VGG16
- 2** Cut the net and extract features
- 3** Retrieve the most similar images
- 4** Check their diagnosis and compare to the new one

Q
U
E
R
Y

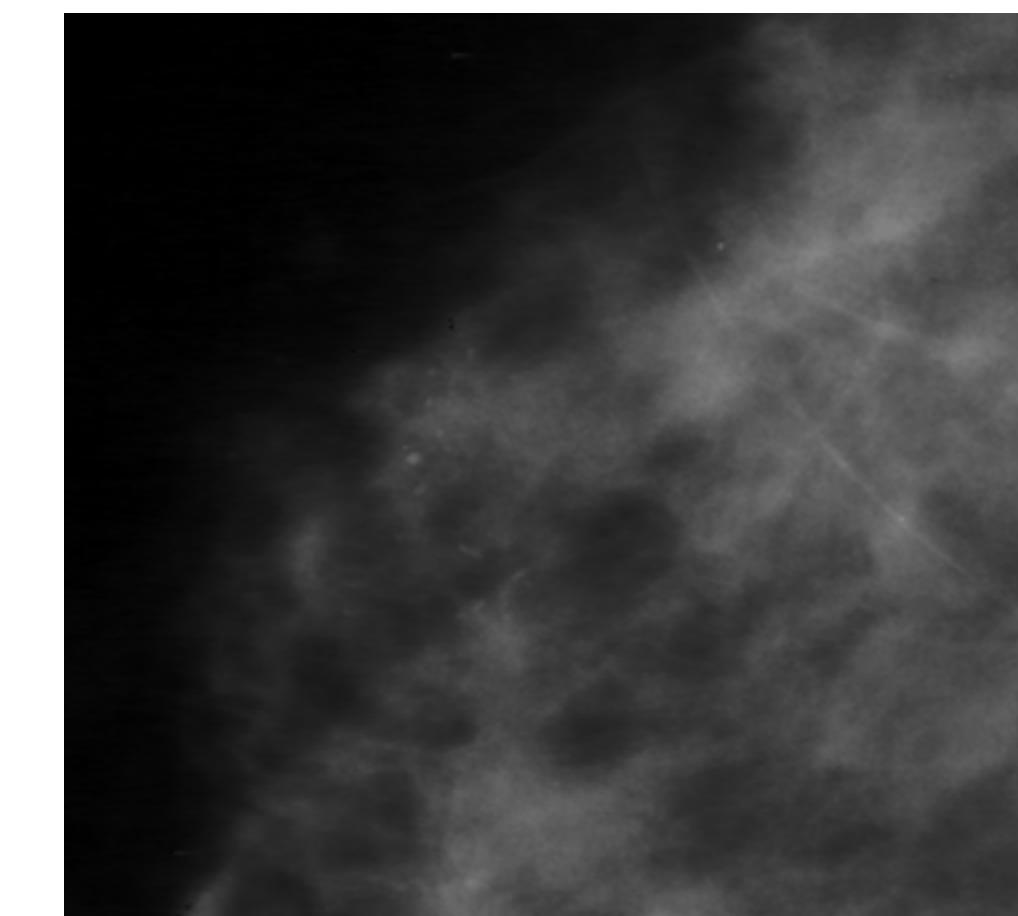
BENIGN



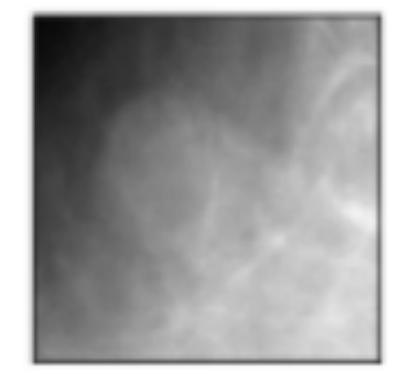
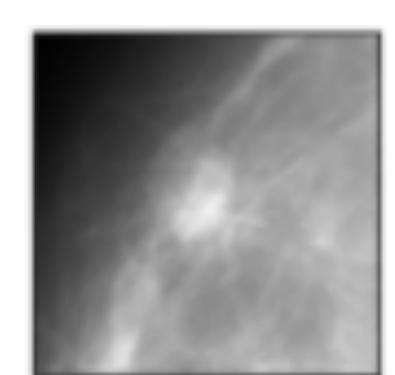
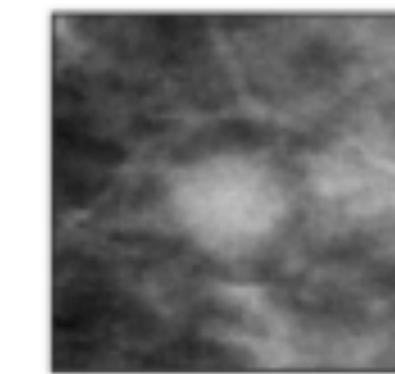
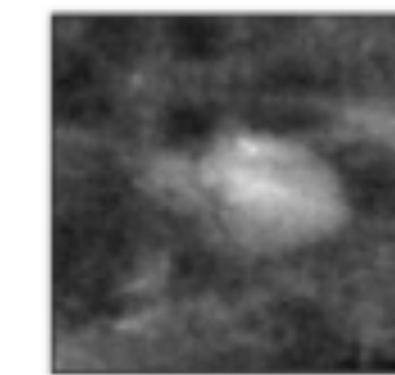
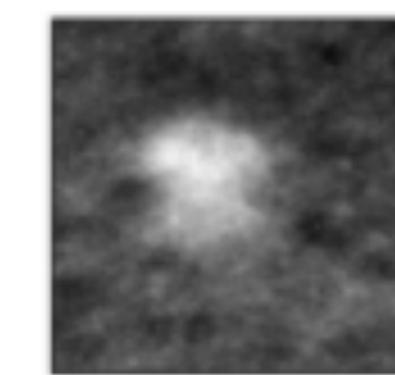
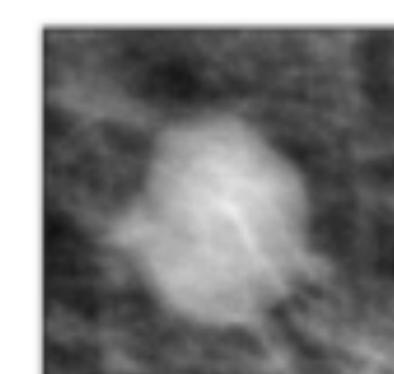
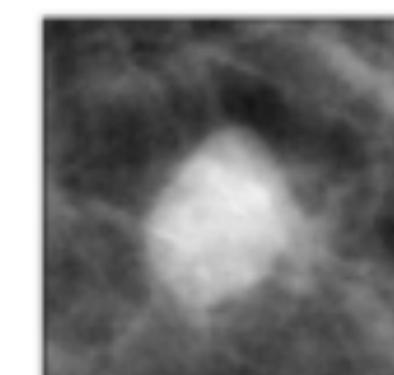
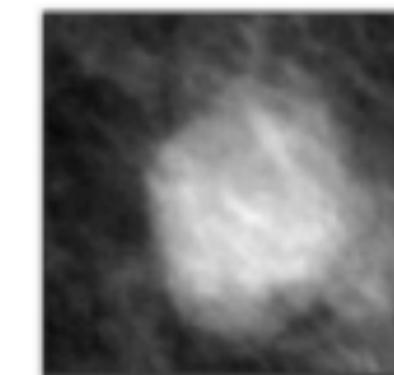
MALIGNANT



FLAT



R
E
T
R
I
E
V
E
D



(A LOT OF!) PROBLEMS

- **Informativeness** of focused images
- Dimension of dataset (2528 images)
- Computational resources
- Complexity of task

IMPROVEMENTS

- ⊕ Download the **full dataset** (~200 GB) with web scraping.
- ⊕ Extract **other regions** of breast.
- ⊕ **Fine tuning** on the medical X-ray dataset (chestXray).

GRAZIE PER
L'ATTENZIONE

