BI-RADS CLASSIFICATION OF BREAST CANCER: A NEW PRE-PROCESSING PIPELINE FOR DEEP MODELS TRAINING

CISUC

Inês Domingues[1;2], Pedro H. Abreu[1] and João Santos[2]

[1]CISUC, Department of Informatics Engineering, University of Coimbra, Portugal

[2]CI-IPOP, IPO-Porto Research Center, Portugal

IPOPORTO

Goal

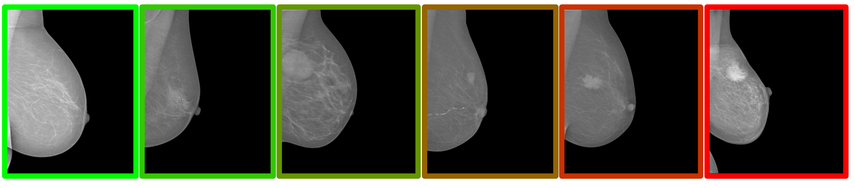
To classify mammograms from the InBreast database into the BI-RADS scale

Motivation

- Breast cancer remains the leading cause of death in women Worldwide
- Screening mammography has been shown to be an effective method for diagnosis of abnormalities in the breast and to increase the survival rate by detecting the cancer in its initial stage
- Manual analysis is time consuming, expensive, and subjective to variability and human error
- Deep learning based strategies have shown to provide near-human performance

Context

BI-RADS	Description
0	the exam is not conclusive
1	no findings
2	benign findings
3	probably benign findings
4	suspicious findings
5	high probability of malignancy
6	proved cancer



Acknowledgment

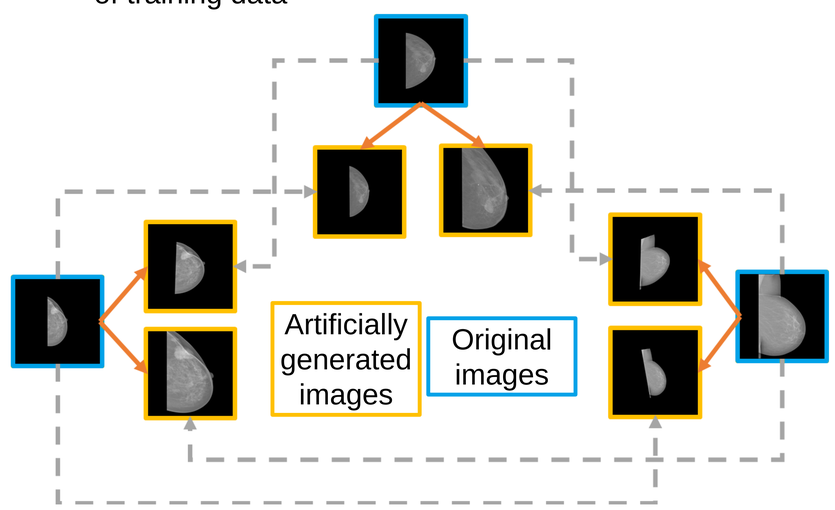
This article is a result of the project NORTE-01-0145-FEDER-000027, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).

Our proposals

We propose two preprocessing techniques:

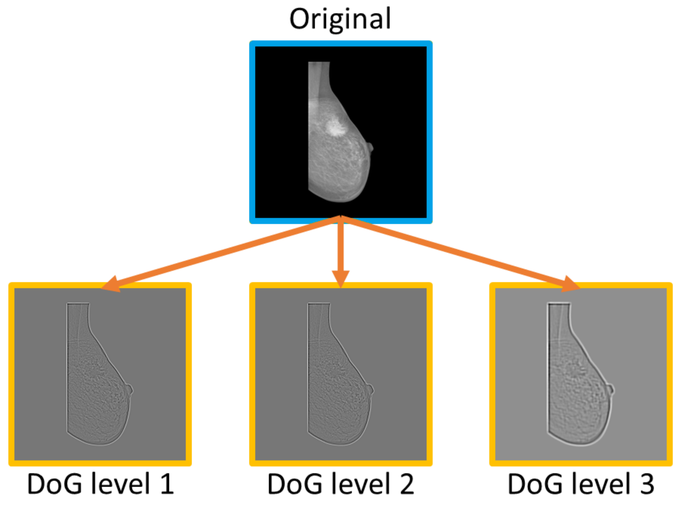
1. A new data augmentation technique that takes into consideration the anatomic soundness

   Why? Deep learning techniques perform better with a lot of training data

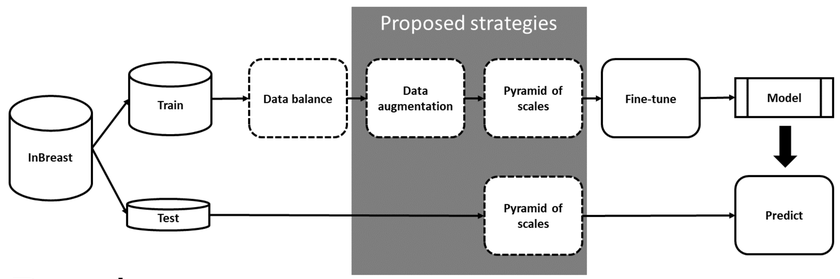


2. A pyramid of scales to be given as input to the model

   Why? Lesions visible in mammograms can have sizes ranging from 15 to 500mm2



Experimental setting



Results

- Column 1: Input
  - Original image *versus* pyramid of scales (DoG)
- Column 2: Data augmentation
  - No data augmentation *versus* image mirroring *versus* proposed method
- Column 3: Imbalance
  - Original proportion *versus* undersampling

input	augmentation	balance	MaxAcc	MAE
image	none	yes	77.0(11.1)	1.2(0.4)
image	none	no	70.4(5.8)	0.5(0.1)
image	mirrored	yes	69.1(7.0)	1.1(0.1)
image	mirrored	no	70.9(6.6)	0.5(0.0)
image	proposed	yes	81.9(4.9)	0.7(0.1)
image	proposed	no	83.2(4.3)	**0.4**(0.0)
DoG	none	yes	73.5(16.4)	1.0(0.2)
DoG	none	no	70.5(2.3)	0.5(0.0)
DoG	mirrored	yes	68.3(7.4)	1.1(0.2)
DoG	mirrored	no	71.2(3.8)	0.6(0.1)
DoG	proposed	yes	76.0(4.2)	0.7(0.1)
DoG	proposed	no	**83.4**(3.6)	**0.4**(0.0)

Conclusions

- Data augmentation technique improves the accuracy by more than 33% when compared with the results with no data augmentation and more than 32% when compared with augmenting the train dataset by mirroring the images
- The pyramid of scales has a more mild effect with improvements in accuracy of 3%