



UNIVERSITÀ DI PISA

# COMPUTATIONAL INTELLIGENCE AND DEEP LEARNING

Educational Version

2020

A. Schiavo - M. Gómez - M. Daole

December 31, 2020



# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Convolutional Neural Network for Medical Imaging Analysis</b>	<b>7</b>
2.1	Original Dataset . . . . .	7
2.2	Data Records . . . . .	8
<b>3</b>	<b>CNN from Scratch with Python</b>	<b>11</b>
<b>4</b>	<b>Pre-Trained CNN</b>	<b>13</b>
<b>5</b>	<b>Baseline Abnormality detection in mammography</b>	<b>15</b>
<b>6</b>	<b>Ensemble of Neural Networks</b>	<b>17</b>
6.1	Examples of Predictions . . . . .	17



# Chapter 1

## Introduction

Deep learning rose to its prominent position in computer vision when neural networks started outperforming other methods on several high-profile image analysis. Healthcare providers generate and capture enormous amounts of data containing extremely valuable signals and information, at a pace far surpassing what *traditional* methods of analysis can process. Machine learning therefore quickly enters the picture, as it is one of the best ways to integrate, analyze and make predictions based on large, heterogeneous data sets.



## Chapter 2

# Convolutional Neural Network for Medical Imaging Analysis

On this investigation, the main objective is to perform abnormality classification in mammography using Convolutional Neural Networks for Medical Imaging Analysis. This laboratory research will be development with a standart evaluation data set in the area of decision support systems in mammography, the *Digital Curated Breast Imaging Subset of Database for Screening Mammography* (CBIS DDSM)

### 2.1 Original Dataset

The dataset we will focus on is an updated and standardized version of the Digital Database for Screening Mammography (DDSM). The DDSM is a database of 2,620 scanned film mammography studies. It contains normal, benign, and malignant cases with verified pathology information. Few well-curated public datasets have been provided for the mammography community. These include the DDSM, the Mammographic Imaging Analysis Society (MIAS) database, and the Image Retrieval in Medical Applications (IRMA) project. Although these public data sets are useful, they are limited in terms of data set size and accessibility.

Lee, Rebecca Sawyer, et al. 'A curated mammography data set for use in computer-aided detection and diagnosis research' Scientific data 4 (2017): 170177.

The images have been decompressed and converted to DICOM format. Updated ROI segmentation and bounding boxes, and pathologic diagnosis for training data are also included. The data set contains 753 calcification cases and 891 mass cases, providing a data-set size capable of analyzing decision support systems in mammography.

Considering the benefits of using deep learning in image classification problem (e.g., automatic feature extraction from raw data), develop a deep Convolutional Neural Network (CNN) that will be trained to read mammography images and classify them into the following five instances:

- Normal

<b>Design Types</b>	-Design Types and Parallel group design. -Feature extraction objective. -Image processing objective
<b>Measurement Type</b>	Mammography
<b>Technology Type</b>	Digital curation
<b>Factor Type</b>	Diagnosis
<b>Sample Characteristic</b>	Homo sapiens

Table 2.1: DDSM Data Set description.

- Benign Calcification
- Benign Mass
- Malignant Calcification
- Malignant Mass

A simple example of the image provided from the original dataset:

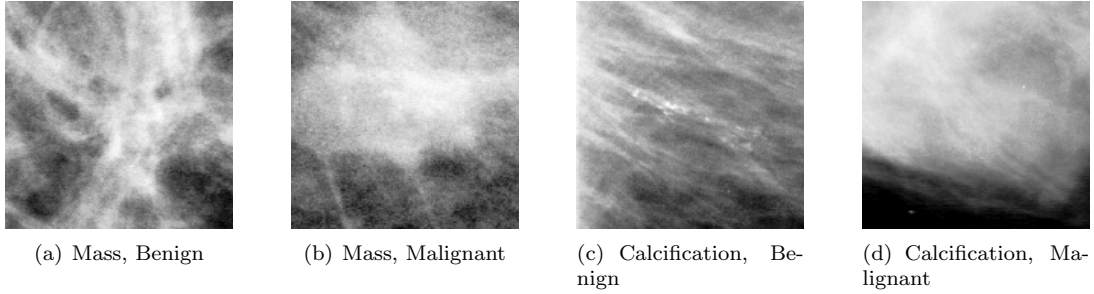


Figure 2.1: Medical Image Representation

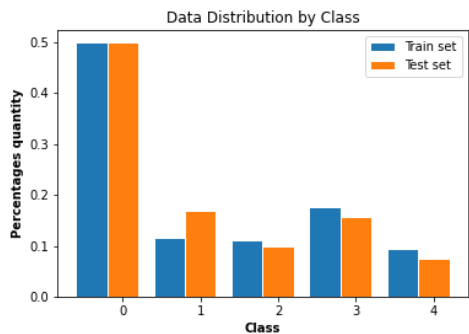
## 2.2 Data Records

The images are distributed at the full mammography and abnormality level as DICOM files. Full mammography images include both MLO and CC views of the mammograms. Abnormalities are represented as binary mask images of the same size as their associated mammograms.

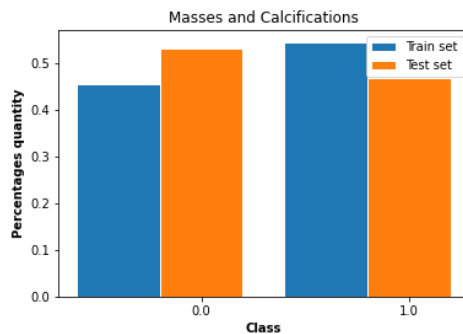
As we mention previously, *CBIS DDSM*: Curated Breast Imaging Subset of Digital Database for Screening Mammography.

A description of the dataset is provided in:

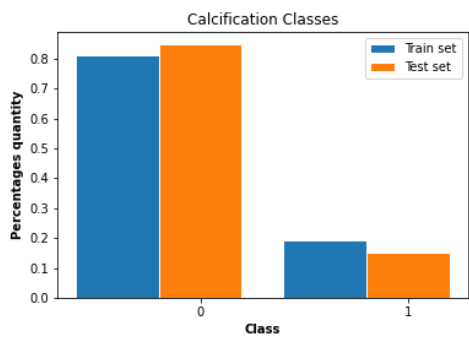




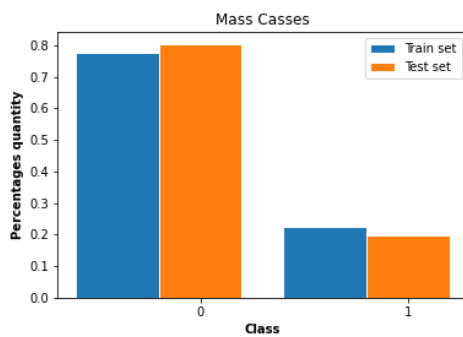
(a) Classes: Baseline patch(0), Mass, benign(1), Mass, malignant(2), Calcification, benign(3), Calcification, malignant(4)



(b) Pathology: Benign(0), Malignant(1)



(c) Pathology: Benign(0), Malignant(1)



(d) Pathology: Benign(0), Malignant(1)

Figure 2.2: Data Representation



## Chapter 3

# CNN from Scratch with Python



## Chapter 4

# Pre-Trained CNN



## Chapter 5

# Baseline Abnormality detection in mammography





## Chapter 6

# Ensemble of Neural Networks

### 6.1 Examples of Predictions