



UNIVERSITÀ DI PISA

# COMPUTATIONAL INTELLIGENCE AND DEEP LEARNING

Educational Version

2020

A. Schiavo - M. Gómez - M. Daole

January 4, 2021



# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>5</b>  |
| 1.1      | Notebooks . . . . .  | 5         |
| <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> |
| 3.1      | Masses and Calcifications . . . . .                              | 11        |
| 3.1.1    | Data Preparation . . . . .                                       | 11        |
| 3.1.2    | Building CNN Architecture . . . . .                              | 11        |
| 3.1.3    | Visualization of the Data . . . . .                              | 12        |
| 3.1.4    | Fighting Overfitting . . . . .                                   | 12        |
| 3.1.5    | Hyperparameter Tuning . . . . .                                  | 12        |
| 3.1.6    | Testing the best 3 models . . . . .                              | 12        |
| 3.1.7    | Choose the best model . . . . .                                  | 12        |
| 3.2      | Benign and Malignant . . . . .                                   | 12        |
| <b>4</b> | <b>Pre-Trained CNN</b>   | <b>17</b> |
| <b>5</b> | <b>Baseline Abnormality detection in mammography</b>             | <b>19</b> |
| <b>6</b> | <b>Ensemble of Neural Networks</b>                               | <b>21</b> |
| 6.1      | Examples of Predictions . . . . .                                | 21        |



# Chapter 1

## Introduction

Breast cancer is one of the most common types of cancer in women. Early detection and treatment can effectively improve cure rates and reduce mortality. Detecting breast cancer using mammographic images is a cost-effective technique, and radiologists can make a diagnosis by analyzing these images. However, the large number of mammographic images produced day by day has brought a huge workload on radiologists and also increased the rate of misdiagnosis. Therefore, developing a computer-aided diagnosis (CAD) system can significantly relieve the pressure on radiologists and improve the diagnosis accuracy.

Machine learning therefore quickly enters the picture, based on large, heterogeneous data sets, the automatic analysis for mammography images needs to be analyzed and make predictions from the regions of interest and classify these regions into normal or abnormal (benign and malignant).

### 1.1 Notebooks

#### 1. Knowing the Dataset

- [Knowing-DDSM\\_dataset.ipynb](#)

#### 2. CNN from Scratch

- [Scratch\\_CNN\\_benign\\_vs\\_malign.ipynb](#)
- [Scratch\\_CNN\\_masses\\_vs\\_calc.ipynb](#)

#### 3. PreTrained

- [PreTrained\\_CNN\\_benign\\_vs\\_malign.ipynb](#)
- [PreTrained\\_CNN\\_masses\\_vs\\_calc.ipynb](#)

#### 4. Baseline

- [BaselineCNN.ipynb](#)

#### 5. Esemble

- [Ensemble.ipynb](#)



## 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.

In the subsequent sections, data source, data preprocessing, data augmentation, model development and evaluation will be delineated. A simple example of the image provided from the original dataset:

|                              |   |
|------------------------------|---|
| <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.

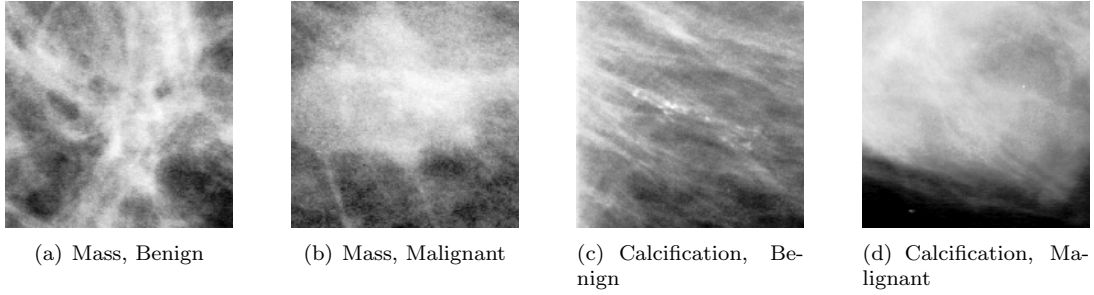


Figure 2.1: Medical Image Representation

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

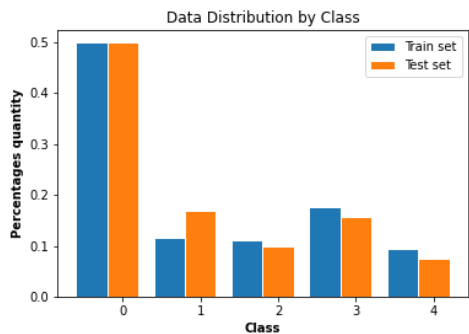
## 2.2 Data Records

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.

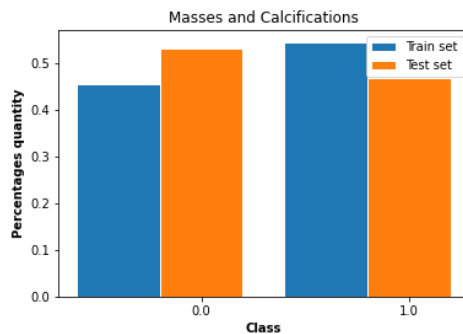
As we mention previously, *CBIS DDSM*: Curated Breast Imaging Subset of Digital Database for Screening Mammography.

A description of the dataset is provided in:

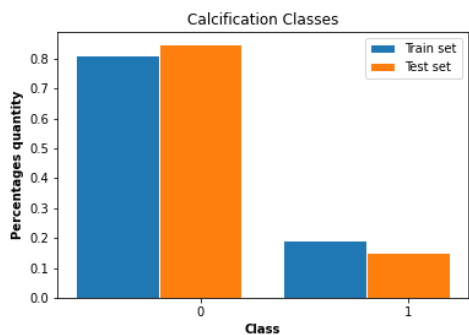




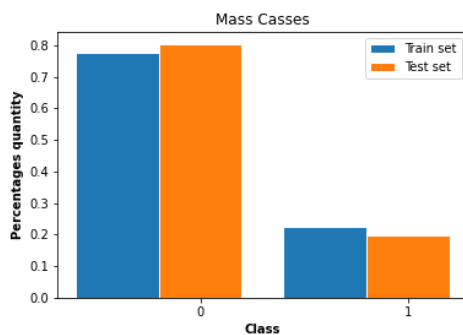
(a) Classes: Baseline patch (0), Mass, benign (1), Mass, malignant (2), Calcification, benign (3), Calcification, malignant (4)



(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

### 3.1 Masses and Calcifications

#### 3.1.1 Data Preparation

- Data Loading
- Data Reshape
- Data Normalization
- Splitting into Training and Validation data

This can be easily done with the Python data manipulation. Modern deep learning provides a very powerful framework for supervised learning, we introduce on this step the convolutional network for scaling to large images.

#### 3.1.2 Building CNN Architecture

- Definition CNN Model
- Data Optimizer

| Test Num | Batch size | Epochs | Optimizer | Loss   | Accuracy | Precision | Recall | AUC    |
|----------|------------|--------|-----------|--------|----------|-----------|--------|--------|
| 1        | 20         | 100    | RMSprop   | 0.7444 | 0.8313   | 0.8519    | 0.7718 | 0.8803 |
| 2        | 20         | 100    | Adam      | 0.8893 | 0.8031   | 0.7829    | 0.7987 | 0.8761 |
| 3        | 20         | 100    | Nadam     | 0.7338 | 0.8156   | 0.7885    | 0.8255 | 0.9022 |

Table 3.1: Comparation Optimizer

| batch size | num epochs | units per layer | optimizer | learning rate | mean val acc | mean val loss | num folds |
|------------|------------|-----------------|-----------|---------------|--------------|---------------|-----------|
| 20         | 100        | [32 64 128 128] | RMSprop   | 0.01          | 0.8615       | 0.3377        | 5         |
| 20         | 100        | [32 64 128 256] | RMSprop   | 0.01          | 0.8777       | 0.3445        | 5         |
| 20         | 100        | [32 64 128 128] | Adam      | 0.01          | 0.8746       | 0.3198        | 5         |
| 20         | 100        | [32 64 128 256] | Adam      | 0.01          | 0.8712       | 0.3218        | 5         |
| 20         | 100        | [32 64 128 128] | RMSprop   | 0.001         | 0.8865       | 0.3032        | 5         |
| 20         | 100        | [32 64 128 256] | RMSprop   | 0.001         | 0.8865       | 0.3096        | 5         |
| 20         | 100        | [32 64 128 128] | Adam      | 0.001         | 0.8808       | 0.3143        | 5         |
| 20         | 100        | [32 64 128 256] | Adam      | 0.001         | 0.8719       | 0.3136        | 5         |
| 20         | 100        | [32 64 128 128] | RMSprop   | 0.0001        | 0.8796       | 0.3183        | 5         |
| 20         | 100        | [32 64 128 256] | RMSprop   | 0.0001        | 0.8704       | 0.3300        | 5         |
| 20         | 100        | [32 64 128 128] | Adam      | 0.0001        | 0.8815       | 0.3031        | 5         |
| 20         | 100        | [32 64 128 256] | Adam      | 0.0001        | 0.8554       | 0.3688        | 5         |

Table 3.2: Comparation Optimizer

### 3.1.3 Visualization of the Data

### 3.1.4 Fighting Overfitting

- CNN Model No.1 Data Augmentation
- CNN Model No.2 K-fold cross validation

### 3.1.5 Hyperparameter Tuning

### 3.1.6 Testing the best 3 models

### 3.1.7 Choose the best model

## 3.2 Benign and Malignant

| batch size | num epochs | units per layer | optimizer | learning rate | mean val acc | mean val loss | num folds |
|------------|------------|-----------------|-----------|---------------|--------------|---------------|-----------|
| 32         | 100        | [32 64 128 128] | RMSprop   | 0.01          | 0.8609       | 0.3520        | 5         |
| 32         | 100        | [32 64 128 256] | RMSprop   | 0.01          | 0.8723       | 0.3259        | 5         |
| 32         | 100        | [32 64 128 128] | Adam      | 0.01          | 0.8711       | 0.3289        | 5         |
| 32         | 100        | [32 64 128 256] | Adam      | 0.01          | 0.8859       | 0.3226        | 5         |
| 32         | 100        | [32 64 128 128] | RMSprop   | 0.001         | 0.8582       | 0.3498        | 5         |
| 32         | 100        | [32 64 128 256] | RMSprop   | 0.001         | 0.8613       | 0.3446        | 5         |
| 32         | 100        | [32 64 128 128] | Adam      | 0.001         | 0.8723       | 0.3278        | 5         |
| 32         | 100        | [32 64 128 256] | Adam      | 0.001         | 0.8699       | 0.3309        | 5         |
| 32         | 100        | [32 64 128 128] | RMSprop   | 0.0001        | 0.8727       | 0.3319        | 5         |
| 32         | 100        | [32 64 128 256] | RMSprop   | 0.0001        | 0.8664       | 0.3441        | 5         |
| 32         | 100        | [32 64 128 128] | Adam      | 0.0001        | 0.8688       | 0.3460        | 5         |
| 32         | 100        | [32 64 128 256] | Adam      | 0.0001        | 0.8730       | 0.3359        | 5         |

Table 3.3: Comparison Optimizer

| batch size | num epochs | units per layer | optimizer | learning rate | mean val acc | mean val loss | num folds |
|------------|------------|-----------------|-----------|---------------|--------------|---------------|-----------|
| 64         | 100        | [32 64 128 128] | RMSprop   | 0.01          | 0.8742       | 0.3405        | 5         |
| 64         | 100        | [32 64 128 256] | RMSprop   | 0.01          | 0.8125       | 0.4174        | 5         |
| 64         | 100        | [32 64 128 128] | Adam      | 0.01          | 0.8621       | 0.3646        | 5         |
| 64         | 100        | [32 64 128 256] | Adam      | 0.01          | 0.8387       | 0.3795        | 5         |
| 64         | 100        | [32 64 128 128] | RMSprop   | 0.001         | 0.8527       | 0.3954        | 5         |
| 64         | 100        | [32 64 128 256] | RMSprop   | 0.001         | 0.8684       | 0.3482        | 5         |
| 64         | 100        | [32 64 128 128] | Adam      | 0.001         | 0.8414       | 0.3946        | 5         |
| 64         | 100        | [32 64 128 256] | Adam      | 0.001         | 0.7875       | 0.4502        | 5         |
| 64         | 100        | [32 64 128 128] | RMSprop   | 0.0001        | 0.8613       | 0.3523        | 5         |
| 64         | 100        | [32 64 128 256] | RMSprop   | 0.0001        | 0.8562       | 0.3706        | 5         |
| 64         | 100        | [32 64 128 128] | Adam      | 0.0001        | 0.8109       | 0.4174        | 5         |
| 64         | 100        | [32 64 128 256] | Adam      | 0.0001        | 0.8363       | 0.4063        | 5         |

Table 3.4: Comparison Optimizer

| batch size | num epochs | units per layer | optimizer | learning rate | mean val acc | mean val loss | num folds |
|------------|------------|-----------------|-----------|---------------|--------------|---------------|-----------|
| 128        | 100        | [32 64 128 128] | RMSprop   | 0.01          | 0.8441       | 0.4371        | 5         |
| 128        | 100        | [32 64 128 256] | RMSprop   | 0.01          | 0.8398       | 0.4141        | 5         |
| 128        | 100        | [32 64 128 128] | Adam      | 0.01          | 0.8477       | 0.4169        | 5         |
| 128        | 100        | [32 64 128 256] | Adam      | 0.01          | 0.8297       | 0.4262        | 5         |
| 128        | 100        | [32 64 128 128] | RMSprop   | 0.001         | 0.7937       | 0.5057        | 5         |
| 128        | 100        | [32 64 128 256] | RMSprop   | 0.001         | 0.8520       | 0.4182        | 5         |
| 128        | 100        | [32 64 128 128] | Adam      | 0.001         | 0.7594       | 0.4973        | 5         |
| 128        | 100        | [32 64 128 256] | Adam      | 0.001         | 0.8152       | 0.4529        | 5         |
| 128        | 100        | [32 64 128 128] | RMSprop   | 0.0001        | 0.8301       | 0.4478        | 5         |
| 128        | 100        | [32 64 128 256] | RMSprop   | 0.0001        | 0.8195       | 0.4629        | 5         |
| 128        | 100        | [32 64 128 128] | Adam      | 0.0001        | 0.8316       | 0.4299        | 5         |
| 128        | 100        | [32 64 128 256] | Adam      | 0.0001        | 0.8316       | 0.4338        | 5         |

Table 3.5: Comparison Optimizer

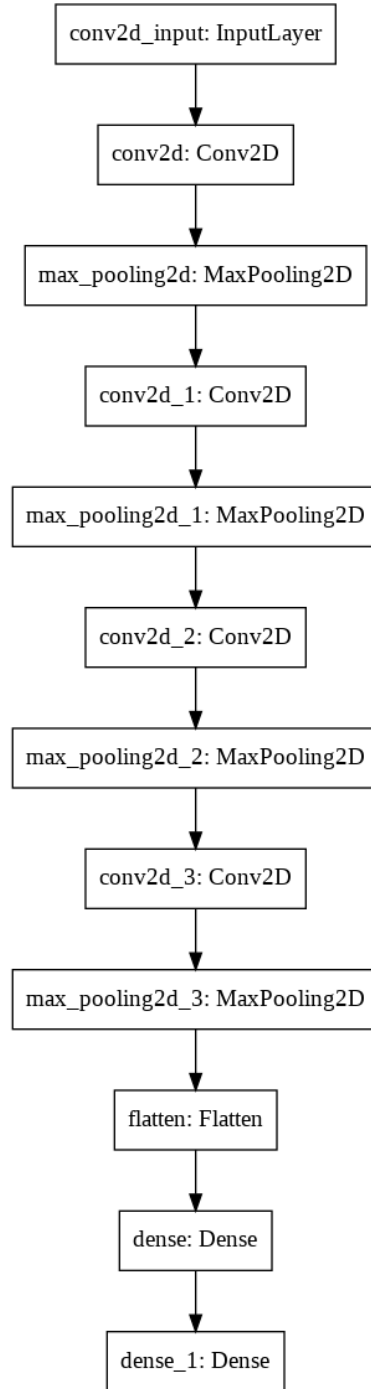


Figure 3.1: The schematic diagram of two biological neurons.





## Chapter 4

# Pre-Trained CNN



## Chapter 5

# Baseline Abnormality detection in mammography



## Chapter 6

# Ensemble of Neural Networks

### 6.1 Examples of Predictions