# UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY

**Escuela de Ciencias Biológicas e Ingeniería**


## TITULO: Detection of breast cancer using the INCEPTION V3 neural network

Trabajo de integración curricular presentado como requisito para la obtención del título de Ingeniera Biomédica


**Autor:**

Elizabeth del Cisne Macas Aguilar

**Tutor:**

M.Sc. Jonathan Cruz Varela

**Co-tutor:**

Ph.D. Andrés Tirado Espín



Urcuquí, Abril 2024

# AUTORÍA

Yo, **Elizabeth del Cisne Macas Aguilar**, con cédula de identidad 0705631307, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el/la autora (a) del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, abril del 2024

Elizabeth del Cisne Macas Aguilar
CI:0705631307

# AUTORIZACIÓN DE PUBLICACIÓN

Yo, **Elizabeth del Cisne Macas Aguilar,** con cédula de identidad 0705631307, cedo a la Universidad de Investigación de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior

Urcuquí, Abril del 2024.

Elizabeth del Cisne Macas Aguilar
CI:0705631307

# Dedication

*This is dedicated to my parents, Fidel and Patricia, and my brothers, Jhon, Elvis, Mayerly, and Matias, the most important people in my life, my support and inspiration.*

# Acknowledgment

Agradezco profundamente a mis amados padres Fidel y Patricia por ser el pilar fundamental de cada uno de mis logros. Su amor incondicional, sacrificio y apoyo han sido la brújula que ha guiado mi camino. Son mi mayor inspiración para superar cualquier obstáculo y alcanzar mis sueños. A mis queridos hermanos, gracias por ser un ejemplo de tenacidad y por impulsarme siempre a dar lo mejor de mí. Sin la unión y fortaleza de mi familia, ninguno de mis logros hubiera sido posible.

Además, quiero agradecer a Dios y la vida por haber puesto a personas maravillosas en mi vida universitaria en mi querida Yachay Tech. En especial, a mi gran amiga Marilyn por ser un apoyo invaluable estando lejos de mi familia. A Emilia, mi mejor amiga, que pese a la distancia siempre me brindó su respaldo, apoyo y cariño. Y a cada una de las personas que compartieron conmigo diferentes momentos de mi vida universitaria; su amistad y tiempo compartido han sido un regalo invaluable. Cada uno aportó una enseñanza significativa en mí, con altibajos, pero también con grandes alegrías y satisfacciones. Todos han hecho de mi paso por la universidad una experiencia maravillosa e inolvidable, que atesoro profundamente en mi memoria y corazón. Cada conversación, cada instante compartido, cada palabra de aliento en los momentos difíciles han enriquecido mi formación y me han ayudado a crecer tanto profesional como personalmente. Llevo conmigo todos esos recuerdos que han marcado mi vida.

Expreso mi más sincero agradecimiento a mi tutor Jonathan por haber sido mi guía y haber compartido generosamente conmigo sus valiosos conocimientos e impulsarme a alcazar este logro. A mi co-tutor Andrés, por su apoyo y motivación para superarme continuamente y luchar por mis sueños. Asimismo, a todos y cada uno de los profesores que tuve a lo largo de mi carrera, cada uno contribuyó con un valioso granito de arena en mi formación, llevo conmigo todos sus aprendizajes.

# Resumen

El cáncer de mama es la principal causa de mortalidad en las mujeres. Su detección temprana junto con un tratamiento eficaz podría ayudar a reducir su letalidad. El medio de detección más común es la mamografía, sin embargo, su interpretación por parte de los especialistas puede ser desafiante, sobre todo en las fases tempranas de la enfermedad debido a limitaciones de este método. Por ello, los radiólogos y oncólogos buscan apoyarse en medios de detección usando inteligencia artificial. En esta investigación, se propone una arquitectura de red neuronal convolucional inception V3, la cual permitirá realizar la clasificación de las mamografías en dos categorías: con cáncer y sin cáncer. El modelo fue entrenado con una base de datos de mamografías llamada VinDr-Mammo obtenida del portal physionet. Los resultados demuestran un buen desempeño en la tarea de detección de cáncer de mama, ya que el modelo alcanza una precisión de 95.3%, lo cual demuestra su potencial.

**Palabras Clave:**

*Cáncer de mama, Redes neuronales convolucionales, Inteligencia artificial, Aprendizaje profundo, Clasificación de imágenes médicas.*

# Abstract

Breast cancer is the main cause of mortality in women. Its early detection, along with effective treatment, could help reduce its lethality. The most common means of detection is mammograms, however, their interpretation by specialists can be challenging, especially in the early phases of the disease due to limitations of this method. For this reason, radiologists and oncologists seek to rely on detection means using artificial intelligence. In this thesis, an inception V3 convolutional neural network architecture is proposed, which will allow the classification of mammograms into two categories: With cancer and without cancer. The model was trained with a mammography database called VinDr-Mammo obtained from the physionet portal. The results demonstrate good performance in the breast cancer detection task, since the model reaches an accuracy of 95.3%, which demonstrates its potential.

**Keywords:**

*Breast cancer, Convolutional neural networks, Artificial intelligence, Deep learning, medical image classification.*

# Contents

# List of Abbreviations

WHO: World Health Organization

PIAA: Physician's Insurance Association of America

CAD: Computer Aided Design

CNN: Convolutional neural networks

MC: Microcalcifications

AD: Architectural distortions

R-MLO: Right Mediolateral Oblique View

L-MLO: Left Mediolateral Oblique View

R-CC: Right Cranial Caudal Compression View

L-CC: Left Cranial Caudal Compression View

BI-RADS: Breast Imaging Reporting and Data System

AI: Artificial Intelligence

ML: Machine learning

SVM: Support Vector Machines

RNNs: Recurrent Neural Networks

SVR: Support Vector Regression

PCA: Principal Component Analysis

SVD: Singular value decomposition

NLP: Natural language processing

CV: Computer vision

SGD: Stochastic Gradient descent

RMSProp: Root Mean Square Propagation

Adagrad: Adaptive Gradient

DDSM: Digital Database for Screening Mammography

CBIS-DDSM: Curated Breast Imaging Subset of DDSM

MRI: Magnetic resonance imaging

AUC: Area under the curve

ROC: Receiver operating characteristic curve

GRU:  Gated Recurrent Units

LSTM: Long Short-Term Memory

ABUS: Automated breast ultrasound

DICOM: Digital Imaging and Communications in Medicine

GPU: Graphic Processing Unit

CUDA: Compute Unified Device Architecture

ANN: Artificial Neural network

# List of figures

# List of tables

# Chapter 1

# 1. Introduction

## 1.1 Problem Statement

Breast cancer is the most common type of cancer and is the leading cause of death in women (WHO). Early detection along with proper treatment could reduce the alarming statistics of this disease.

The most common means of detection is mammography however this method is not the most effective method is due to the complexity in the analysis by the specialists already in many cases can result in cases can be confused and therefore result in a misdiagnosis erroneous.

Therefore, the present study proposes the development of a computer algorithm through medical image processing, which allows a more complete and comprehensive analysis of mammograms and thus detects the signs of the disease and thus an early and accurate diagnosis of the same.

## 1.2 Justification

Breast cancer is considered one of the most deadly diseases worldwide. According to the World Health Organization, about one in 12 women will develop breast cancer in their lifetime. In 2020, about 685 000 women died as a result of this disease [1].

Timely and early detection is one of the most relevant aspects that could help reduce the mortality rate of this disease by subjecting the patient to appropriate treatment. One of the techniques most commonly used in the detection of this type of cancer, in addition to self-examination, are mammograms.

Mammograms are images of breast tissue obtained through a process of ionizing radiation. The information that can be obtained by analyzing these medical images makes it possible to detect mucinous and lobular cancers in particular more easily [2].

The existence of false positives in the detection of this type of cancer is quite significant and in many cases is due to the fact that the analysis performed is incorrect. This may be due to misinterpretation and difficult differentiation of the characteristics of the anomalies that are present in the breast tissue when there is a tumor. The sensitivity of mammography is 79% and is influenced by a number of variables, most of which are the size of the tumor, its degree of visibility, the density of the breast tissue, the characteristics of the patient and the experience of the radiologist [3]. According to the Physician's Insurance Association of America (PIAA), delay in breast cancer diagnosis and diagnostic errors are common causes of medical malpractice litigation [4].

The final diagnosis is made by professionals in this area: a radiologist and an oncologist, they are in charge of determining the existence or not of a tumor, and whether the diagnosis is 100% reliable will depend a lot on the degree of experience among other factors. In general, mammograms have major limitations, as there is data showing that physicians will recommend about 15 out of every 1000 women screened with mammography for a needle biopsy, and 10 to 13 of those biopsies will show that cancer is not present (false positives) [5]. Because of these errors in the detection of cancer by human analysis of mammograms alone, radiologists and oncologists have always sought to improve the performance of these mammograms and thus generate an accurate diagnosis by which they have relied on computer-aided diagnosis systems but these traditional CAD systems fail to significantly improve detection performance, mainly due to their low specificity [6,7], meaning how well they detect and classify abnormalities in mammograms.

Trister et al. said, "report that novel algorithms based on convolutional neural networks (CNNs) improve the performance of screening mammography and increase the efficiency of breast imaging radiologists" [8]. Thus, this paper proposes the development of a CNN-based breast cancer detection model that detects abnormalities present in breast tissue and classifies the patient's mammograms as benign or malignant.

## 1.3 Objectives

### 1.3.1 General Objective

Develop and evaluate a neural network model based on the Inception v3 architecture to improve accuracy and efficiency in breast cancer detection through mammography image analysis.

### 1.3.2 Specific Objetives

- Obtain a dataset of mammogram images, including both positive breast cancer cases and negative controls.
- Configure, implement and adapt the Inception v3 neural network architecture for mammogram processing and analysis.
- Train the model with the prepared data set, adjusting parameters to improve its performance and generalization capability.
- Test the model using an independent test data set to evaluate its accuracy, sensitivity and specificity in detecting breast cancer.

## 1.4 Tesis Overview

In the development of this work, the first chapter addresses and justifies the importance of early and accurate detection of breast cancer, emphasizing that it is the most common type of cancer and the leading cause of death among women according to the WHO. In addition, it is pointed out that current screening methods, such as mammography, have limitations due to the complexity of image interpretation, which can lead to misdiagnosis. The development of a neural network based on medical image processing is proposed to analyze mammograms in a more complete and exhaustive way and to classify them into two categories with cancer and without cancer. The limitations of the current methods are also pointed out and the objectives that were set for the development of this work are introduced.

Chapter 2 establishes a solid foundation on the current knowledge in the field of breast cancer and related convolutional neural networks. A comprehensive and critical review of the existing literature is conducted, identifying both advances and limitations in recent developments in our area of research. This review allows us to identify challenges and opportunities for future research and practical applications.

In the materials and methods chapter, the methodology applied in this research is demonstrated, highlighting the importance of each step from the preparation of the data set to the evaluation of the model. It details each method and justifies the choices and thus provides a solid basis for the validity of the findings which contributes significantly to the field of biomedicine and early detection of breast cancer.

In chapter 4, the experiments and analyses performed are presented, with reference to the results and conclusions. It includes the testing of the deep learning classification model with different optimizers, as well as the results evaluation phase. In this section, we describe in detail how the model was evaluated, the metrics used, the additional parameters employed, among other relevant aspects. In addition, the generalizability of the CNN to new data is discussed and a comparison and dialogue of authors of similar works is presented in order to highlight the strengths of our model.

The conclusions chapter summarizes the most relevant results of the research, highlighting the most significant findings, their implications, limitations and future work. This section provides an overview of the main contributions and findings of the study, as well as the opportunities and challenges for future research in the area.

The final chapters include the bibliographical references, where all the works cited or mentioned throughout the research can be found. In addition, annexes containing fundamental sections of the code developed are attached, which provides greater transparency of the study.

# Chapter 2

# 2. State of Art

Obtaining an accurate diagnosis for the early detection of breast cancer is fundamental to be able to intervene with an adequate treatment. Currently in the field of neural networks it has been possible to demonstrate a great efficiency in terms of models that perform the processing and classification of medical images for the detection of breast cancer.

## 2.1 Breast Cáncer

According to the World Health Organization, breast cancer is the most common type of cancer, with more than 2.2 million cases in 2020 [1]. On the other hand, in Ecuador, according to statistics presented by the Global Cancer Observatory, 3,563 women were diagnosed with breast cancer in 2020 [9]. The main factors for the development of breast cancer may be DNA damage and genetic mutations that may be influenced by estrogen exposure. In some cases, there will be an inheritance of DNA defects or pro-cancer genes such as BRCA1 and BRCA2 [10]. Frequently detected cancer cells are found in the milk-forming centers, lobules, and ductal canals [11].

Figure 1 shows the most common areas of the breast where cancerous tissue usually develops.

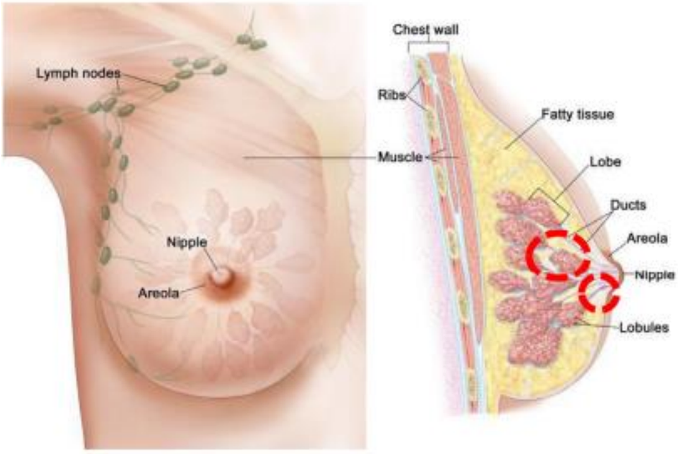



*Fig 1. Most common areas of the breast where cancer develops. Source of own elaboration from [12].*

## 2.2 Mammography

Mammography is one of the most common methods to perform early detection of breast cancer, since after obtaining the images, they are analyzed by radiologists who aim to detect critical features such as microcalcifications (MC), architectural distortions (AD) and asymmetries as biomarkers of cancer or cancer risk [13].

There are mainly two different approaches to cancer detection: disease prevention by finding and eliminating premalignant precursors of cancer; and early cancer detection where the goal is to treat invasive cancer at an early curable stage [14]. According to several studies it has been shown that the sensitivity of mammography could reach 90% and a specificity that could reach 95%, it is important to mention that these states may vary depending on various factors such as population, age, country, etc. [15].


*Fig 2. Image of the breast obtained by mammography. Source of own elaboration from [58].*

### 2.2.1 Mammography views

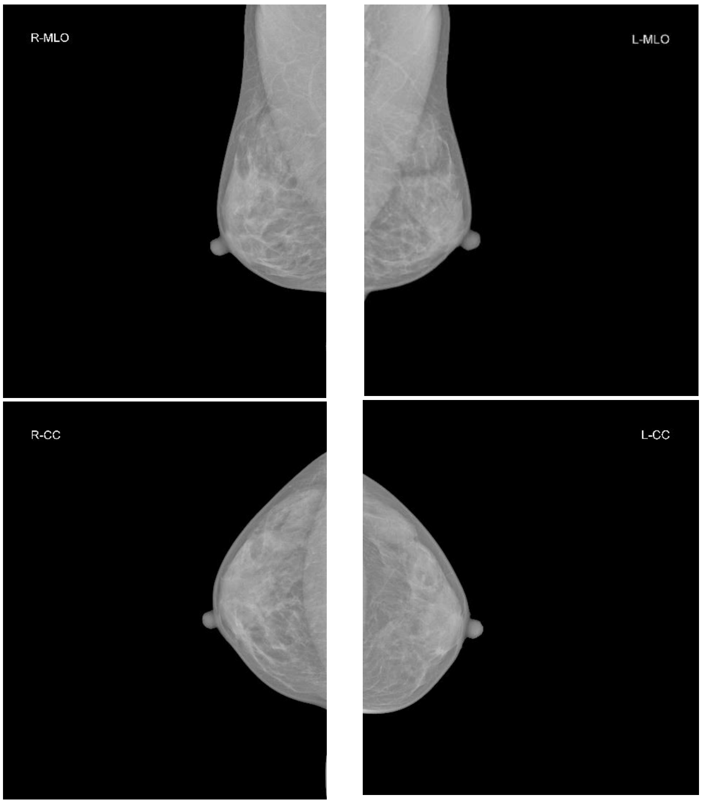



*Fig 3: Four views obtained during mammography (R-MLO, L-MLO, R-CC and L-CC). Source of own elaboration from [58].*

In the process of interpreting mammograms, it is critical to keep in mind that four distinct images are obtained, each corresponding to a specific view of the breast tissue. This diversity of views is crucial for a thorough and accurate evaluation, as illustrated in Figure 3. Each image acquired on mammography is labeled according to the specific view from which it was captured:

**R-MLO (Right Mediolateral Oblique View):** This image shows an oblique view of the right side of the breast, providing an angular perspective that is vital for identifying abnormalities that may not be visible in direct views.

**L-MLO (Left Mediolateral Oblique View):** Similar to the R-MLO, but for the left side of the chest. This angled view helps supplement the information obtained from the right view, ensuring complete coverage of the breast tissue.

**R-CC (Right Cranial Caudal Compression View):** This view provides a vertical compression perspective from the lower (caudal) to the upper (cranial) right breast. It is essential for identifying abnormalities that may be hidden in the oblique view.

**L-CC (Left Cranial Caudal Compression View):** Provides the same vertical perspective as the R-CC, but on the left chest. This view is equally important for a detailed and complete analysis.

Each of these views and their combination when performing the analysis on mammography is crucial for accurate detection and diagnosis. Each view provides a unique perspective of the breast tissue, allowing for a more complete evaluation and reducing the likelihood of overlooking certain abnormalities. In the process of training neural networks for mammogram classification, considering the four views and their specific labels is critical to developing an effective and efficient model.

### 2.2.2 Mammographic Density

Mammography is used to analyze the density of tissue in the breasts. This study shows a higher absorption of X-rays by stromal and epithelial tissues compared to adipose tissue, which translates into brighter areas in the mammographic examination; in contrast, fatty areas appear darker. This characteristic causes mammographic images to differ from one woman to another, based on the tissue structure of each breast. The fraction of the breast made up of epithelial and connective tissue is often identified as the percentage of breast tissue or the mammographic density ratio [16].

According to the review of all published studies investigating the association between qualitative or quantitative measures of radiological mammographic pattern and the risk of prevalent or incident breast cancer in adult women, it was evident that the percentage of mammographic density was strongly associated with an increased risk of breast cancer [17].

In addition, it is essential to keep in mind that there are several methods for measuring mammographic density, some of which are shown in Table 1.

*Table 1. Methods for measuring mammographic density. Source of own elaboration from [18].*

| **Method** | **Category** |
|---|---|
| Wolfe: | N1, predominantly fat<br>P1, ductal prominence <25%.<br>P2, ductal prominence of >25%.<br>DY, extensive dysplasia |
| BI-RADS: | A predominantly fatty<br>B scattered fibroglandular densities<br>C heterogeneously dense<br>D extremely dense |
| Visual estimation of the proportion of the breast occupied by breast tissue. radiológicamente denso: (Boyd) | 1- 0% de density<br>2- 0% a <10%<br>3- 10% a < 25%<br>4- 25% a <50%<br>5- 50% a <75%<br>6- 75% a > o = 75% |

The BI-RADS (Breast Imaging Reporting and Data System) represents a standardized approach to the classification of breast imaging findings and is considered the common language in the diagnosis of breast disease. This method standardizes the language and process of mammographic reporting, in addition to categorizing abnormalities according to their level of suspicion and determining the actions to be taken in each situation [19].

On the other hand, in order to understand certain points of the present work, the BI-RADS method is the one we should have a little more in-depth knowledge of, mainly with regard to the management of breast lesions according to this system, which we can observe in Table 2.

*Table2. Management of breast lesions according to the BI-RADS system. [19].*

| | |
|---|---|
| BI-RADS 0 | Inconclusive due to incomplete reading |
| BI-RADS 1 | Normal breast |
| BI-RADS 2 | Benign (probability of cancer similar to the general population). |
| BI-RADS 3 | Probably benign findings (<2% risk of malignancy) |
| BI-RADS 4 | Probably malignant (positive predictive value for cancer between 29-34% up to 70%) |
| BI-RADS 5 | Highly suggestive of malignancy (PPV for cancer greater than 70%) |
| BI-RADS 6 | Histologically confirmed malignancy, but before definitive treatment is initiated. |

## 2.3 Artificial Intelligence

According to the Encyclopedia Britannica, artificial intelligence is defined as the ability of a digital computer or computer-controlled robot system to perform tasks commonly associated with intelligent beings. The concept is often used to refer to the creation of systems that possess human-like cognitive processes, such as the ability to think, interpret, generalize or learn from previous experiences. Some programs have achieved levels of competence comparable to those of human experts and professionals in specific tasks, so artificial intelligence, in this narrow sense, has been implemented in a wide range of applications, including medical diagnosis, computer search systems, speech or handwriting recognition, and chatbots [20].

In the following figure we can see a representation of the components or subsections of the IA

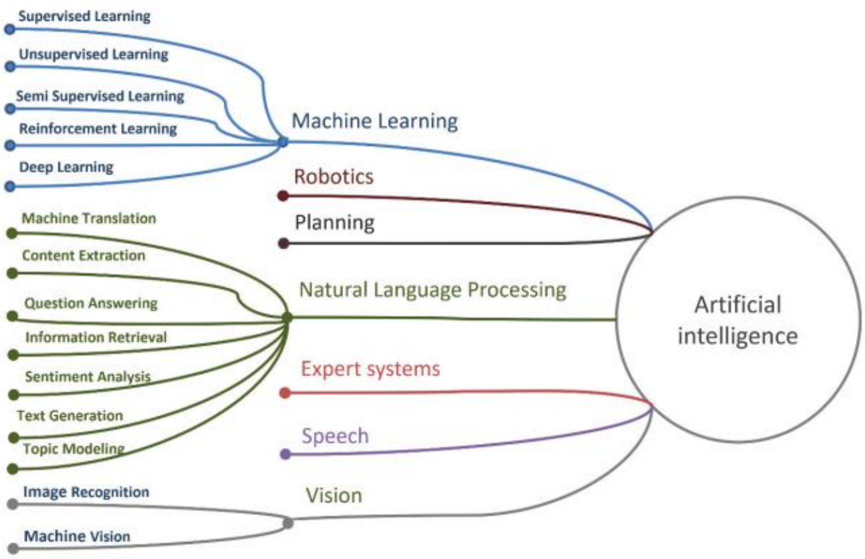



*Fig 4: Artificial Intelligence sub-sections. Source of own elaboration from [21].*

## 2.3.1 Artificial Intelligence in Medicine and Biomedicine

The rapid pace of change in medicine makes it a challenge for today's physicians to keep up with the times. Thus, artificial intelligence (AI) in medicine is having a major impact. By means of advanced algorithms and sophisticated software to replicate human reasoning and analyze medical data, AI promises a revolution in healthcare [22].

The development of AI systems designed for healthcare and medicine is one of the great achievements of science and technology today. Such systems learn to perform specific tasks, such as diagnosing diseases or recommending treatments, by processing large amounts of data from clinical practice and research, largely contained in large databases of biomedical information [23].

The application of artificial intelligence to a large amount of medical data has revolutionized every field of biomedicine, among which is clinical diagnosis, one of the main applications of AI in biomedicine, consists of the analysis of high-resolution medical images in 2 or more dimensions [23]. A large number of deep learning studies have demonstrated excellent and high predictive ability similar to that of specialist physicians in various diagnostic imaging tasks, for example, in the diagnosis of diabetic retinopathy and prostate or breast cancers [24].

## 2.4 Machine Learning

Machine learning is a field of artificial intelligence in which machines learn from past experiences to improve their performance. This process begins with the collection of data such as numbers, images or text. This data is then prepared and organized to train the machine learning model. The more data used to train the model, the better its performance. Other different data is also needed to test how well the model performs. The result is a model that can be applied to new data in the future [25].

On the other hand, we have that the ML are divided into five classes, which are presented below [26] Fig5.

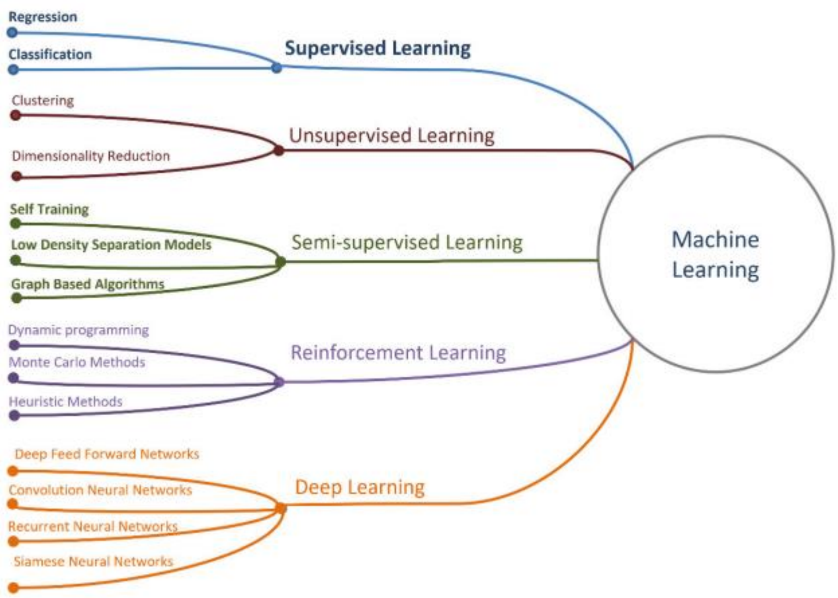


*Fig 5: Basic types of ML models. Source of own elaboration from [21].*

### 2.4.1 Supervised machine learning

Supervised machine learning models require training data that is already categorized or labeled. By training the model with this labeled data, it can learn to recognize specific patterns and features that correspond to each label [27].

In supervised machine learning there are two types:

**Classification**

Classification is a supervised machine learning task that assigns predefined labels or categories to new input data. The goal is to train a model to accurately predict the correct class for new instances it has not seen [28].

Within this learning task, various algorithms are used to solve classification problems:

1. Decision trees

   Predictive models that are generated by decision tree algorithms adopt a tree structure. This means that each of the internal nodes in this structure symbolizes a specific characteristic of the evaluated data. The decision rules associated with each attribute are represented through the various branches emerging from each of these nodes. At the ends of these branches are the leaf nodes, which are the ones that carry the final prediction or classification assigned to the data instance. Among the most prominent algorithms for the creation of decision trees are ID3, C4.5 and CART. These methods are recognized for their effectiveness in structuring and analyzing data for predictive decision making [28].

2. Support Vector Machines (SVM)

Support Vector Machines (SVM) operate by identifying a hyperplane that efficiently distinguishes the different categories within the feature space. This process consists of optimizing the space or margin between the hyperplane and the nearest points of each category. In doing so, a clear and extensive demarcation between categories is generated, which facilitates the model's task of classifying new data points with greater accuracy. SVMs prove to be highly effective in multi-dimensional contexts, since in these environments it is more feasible to find a suitable linear hyperplane compared to lower dimensional spaces. Concisely, SVMs focus on locating the ideal hyperplane that maximizes the spacing between categories, leading to the creation of highly accurate classification models, even in high-dimensional environments [28].

3. Random Forests

   Random forest models build a set of individual decision trees and aggregate their predictions to obtain the final output. Each tree is trained on a random subset of the data and considers only a random subset of features at each split. By averaging a large number of diverse trees, random forests avoid overfitting and are very robust to noise in the data. In short, random forests aggregate the outputs of multiple weak trees to create a strong and stable ensemble classifier [28].

4. Neural Networks

   Modern deep neural networks, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have revolutionized data classification in recent years. Unlike traditional algorithms, these deep learning architectures are capable of learning abstract representations and complex patterns present in large data sets on their own. Thanks to this ability to model nonlinear relationships, deep neural networks consistently outperform other approaches in various classification tasks, such as object detection in images, natural language processing, speech recognition, etc. Classification algorithms in general have applications in a wide variety of fields, from filtering spam and analyzing opinions in text, to diagnosing diseases, detecting fraudulent activities, and many more problems that require assigning items to defined categories [28].

**Regression**

Regression is a supervised learning technique used to predict continuous numerical values, in contrast to classification which predicts discrete categories. The objective in regression problems is to train a model capable of capturing the relationship between a set of independent variables or predictors and a target or dependent variable that we want to predict [28].

Commonly used algorithms for regression problems:

1. Linear Regression

A linear function $(y(\{xn\})$ is assumed to identify the relationship between the observations $\{xn\}$ and the target values $\{tn\}$, which are components of the training data sets [29]. To express the uncertainty about the value of $t$ for each value of $x$, a predictive distribution $p(t|x)$ must be modeled.

2. Decision trees

The decision tree model, as the name suggests, uses an inverted tree data structure to categorize information. The root node contains the most general value, and as you go down the tree, the classes become less generic and more specific until you reach the leaf nodes, which are the final level. The value of these leaf nodes will be the value predicted by the model. In the case of regression, the tree will return a continuous variable as the output [30].

3. Random trees

In random forest regression, we ensemble the predictions of several decision tree regressions [31].

4. Support Vector Regression (SVR)

The fundamental purpose of Support Vector Machines (SVM) is to generate a separating hyperplane that correctly divides two classes with the maximum possible margin in a high-dimensional feature space, using linear or nonlinear mapping functions. The extension of SVMs to Support Vector Regression (SVR) is achieved by introducing a loss function insensitive to an epsilon parameter (ε), which allows for a certain degree of error in data separation [32].

### 2.4.2 Unsupervised machine learning

In unsupervised machine learning, one works with data that is not previously categorized or labeled. Instead of relying on already defined examples, this type of learning identifies patterns and structures in the unlabeled data by itself. Thus, it can reveal connections, groupings, or trends that were not initially evident or not specifically sought after [27].

On the other hand, there are two main tasks for which unsupervised machine learning is used:

**Clustering**

Clustering consists of organizing unlabeled items into groups based on their characteristics and similarities [33].

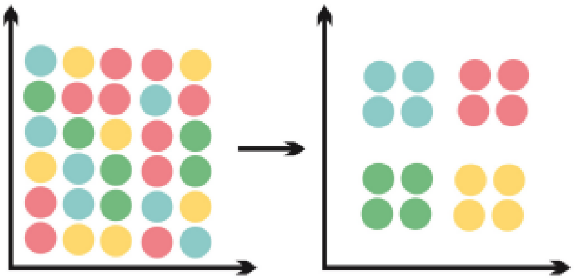


*Fig 6. Example of Grouping. Source of own elaboration from [33].*

Clustering algorithms are classified into several types according to their performance:

1. **Exclusive grouping**

This type of grouping states that a data element or object can only exist within one group. [34]

2. **Hierarchical clustering**

This clustering builds a hierarchy of clusters in the form of a tree. There are two approaches: agglomerative and divisive hierarchical clustering.

Agglomerative hierarchical clustering merges progressively smaller clusters into larger groups as you move up the hierarchy. Divisive clustering recursively separates each cluster into more specific subgroups as you move down the hierarchy. Both methods produce a tree structure showing the relationships between clusters at different levels of granularity [33].

3. **Overlapping clustering**

Overlapping clustering algorithms allow each piece of data to belong to more than one group or cluster, which is useful when the boundaries between groups are fuzzy or when the data has multiple facets [35].

## Dimensionality reduction

Dimensionality reduction is a technique used as a preprocessing step when working with data sets that have a very large number of variables or dimensions. Because having so many features can negatively affect the performance of machine learning algorithms, causing problems such as overfitting. In addition, it makes it difficult to visualize and interpret the data. Dimensionality reduction seeks to reduce the number of variables or features, preserving as much relevant information as possible [36].

In order to meet the dimensionality reduction objective, different methods can be used:

1. **Principal Component Analysis**

   PCA is a dimensionality reduction technique that transforms an original set of correlated variables into a new set of uncorrelated variables known as principal components. It basically applies a linear transformation that reorients the data in the directions of maximum variability, allowing dimensionality reduction by focusing on the first most representative principal [36].

2. **Singular value decomposition**

   The singular value decomposition (SVD) technique represents an alternative methodology for dimension reduction that decomposes a matrix, denoted A, into three matrices of reduced rank. This operation is symbolized by the equation A = USVT, where both U and V are orthogonal matrices and S is a diagonal matrix whose elements are known as the singular values of A. Similar to Principal Component Analysis (PCA), SVD is frequently employed to attenuate noise and perform data compression, including image file management [36].

### 3. Autoencoders

Autoencoders use neural networks to reduce the dimension of the data and subsequently generate a reconstructed version of the original input data. As illustrated in the accompanying figure, the intermediate layer or hidden layer essentially functions as a choke point that minimizes the input data prior to its reconstruction at the output layer. The process that takes the data from the input layer to the hidden layer is called "encoding", and the process that reconstructs the data from the hidden layer to the output layer is called "decoding" [36].

## 2.5 Deep learning

Deep learning is a subfield of machine learning that uses multi-layered artificial neural networks to learn representations of data. It is called "deep" because these neural networks have several hidden layers that allow them to extract complex features and relationships within large data sets. Deep learning methods can be supervised, semi-supervised or unsupervised, depending on whether the training data is labeled or unlabeled. In essence, deep learning leverages multilayer neural networks to discover subtle and abstract patterns in large volumes of data [37].

The four basic architectures of neural networks are shown in Fig. 6.

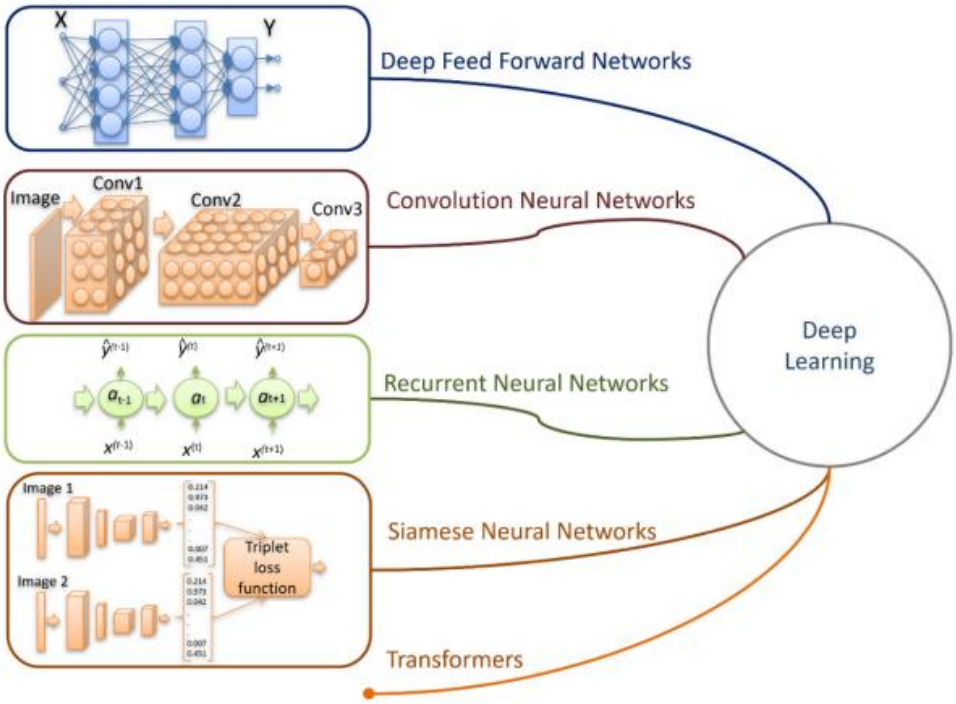



*Fig 7: Deep neural networks. Source of own elaboration from [21]*.

### 2.5.1 Deep Feed Forward Networks

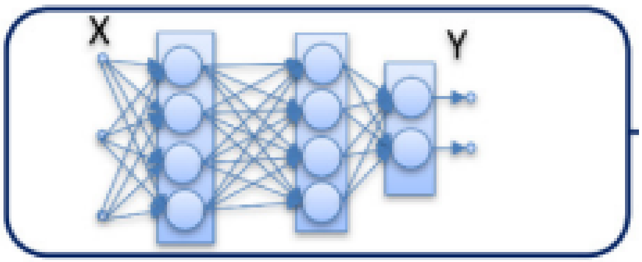



*Fig 8: Diagram indicating the basic architecture of a Deep feed neural networks. Source of own elaboration from [21].*

Deep Feed Forward Networks models known as multilayer perceptrons are currently the most widely used in many applications. These models are inspired by biology and consist of simple processing units connected in layers, mimicking neurons. Each unit in one layer connects to all units in the previous layer. These connections have different weights that encode the knowledge of the network [38].

Data enters at the input layer and flows through the network, layer by layer, until it reaches the output. After the input layer there is a hidden layer with several neurons in parallel. Each neuron does a weighted sum of the inputs using the connection weights. This sum goes through a nonlinear activation function. Normally there is no feedback between layers during operation. The layered structure and nonlinear activation functions give these networks the ability to learn complex pattern representations [38].

### 2.5.2 Convolutional Neural Networks

A convolutional neural network may contain numerous layers, with each layer charged with learning to identify different features present in an image. These filters are applied to each training image at various resolutions, and the results obtained after convolution are used as input for the next layer. The filters can initially detect basic features such as brightness and edges, but as they progress, they can become more complex and uniquely define specific objects [39].

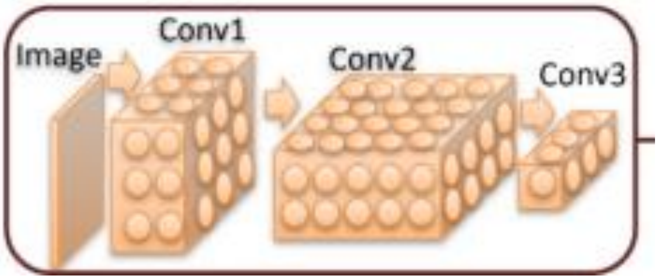



*Fig 9. Diagram indicating the basic architecture of a convolutional neural network. Source of own elaboration from [21].*

In the field of medicine, the application of a CNN allows the analysis of medical images in order to visually detect the presence of anomalies in a given organ. CNN-based models have a wide variety of applications among which the main ones are: image segmentation in medical image processing, image feature extraction, search for regions of interest, object detection, natural language processing, etc.

In breast cancer detection one of the applications that brings the most benefit is that CNNs have trainable parameters in the various layers that are applied to extract important features at various levels of abstraction [40]. In a research, a learning approach was presented that uses a deep learning framework capable of automatically learning features from mammography images to detect cancer [41].

## 2.5.3 Recurrent Neural Networks

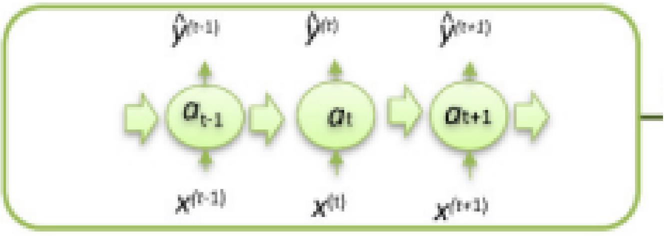


*Fig 10. Diagram indicating the basic architecture of a Recurrent Neural Networks. Source of own elaboration from [21].*

Recurrent neural networks (RNN) are a type of neural network specialized in modeling temporal sequences. Their distinguishing feature is the ability to propagate information over time. RNNs have additional connections between units that capture temporal dependencies. These recurrent connections train the network to generate predictions where each output depends on the current input and information from previous steps. In this way, they can analyze sequential data where context and order are important, as elements are correlated in time. In conclusion, recurrent neural networks use temporal memory through their recurrent connections, which makes them suitable for processing sequences of interrelated inputs such as time series [42].

### 2.5.4 Siamese neural network

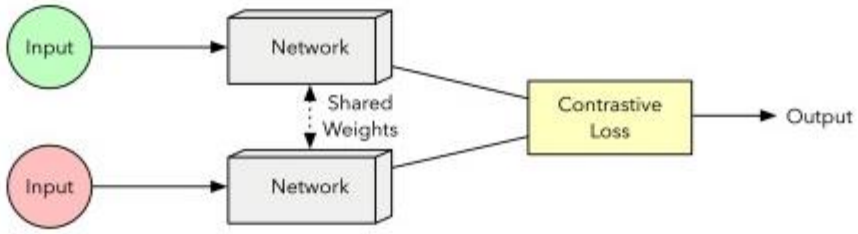



*Fig 11. Diagram indicating the basic architecture of a. Siamese neural network. Source of own elaboration from [21].*

Siamese networks, also known as twin networks, consist of two identical neural networks that share the same weights. Their goal is to calculate the similarity between pairs of input data. Basically, these networks seek to determine whether two inputs are different or not. They have two branches that process each input separately, but by sharing the same parameters, they learn a joint representation of the input space. Then a comparison layer connects both branches and determines the similarity between the two generated feature vectors. In this way, Siamese networks can tell whether two inputs are similar or not, which is useful in tasks such as identity verification, duplicate detection, image search, etc [43].

An interesting aspect of Siamese networks is that they can be constructed with different types of neural networks as components. Instead of standard multilayer perceptrons, convolutional or recurrent networks can be used in each Siamese branch. This flexibility increases the applicability of Siamese networks for various tasks. For example, using convolutional networks improves handwriting recognition and face detection. And with RNNs, search queries can be compared with documents taking into account the sequential context [43].

### 2.5.5 Transformes

Transformer is a leading deep learning model that has been widely adopted in various fields, such as natural language processing (NLP), computer vision (CV) and speech processing [44].

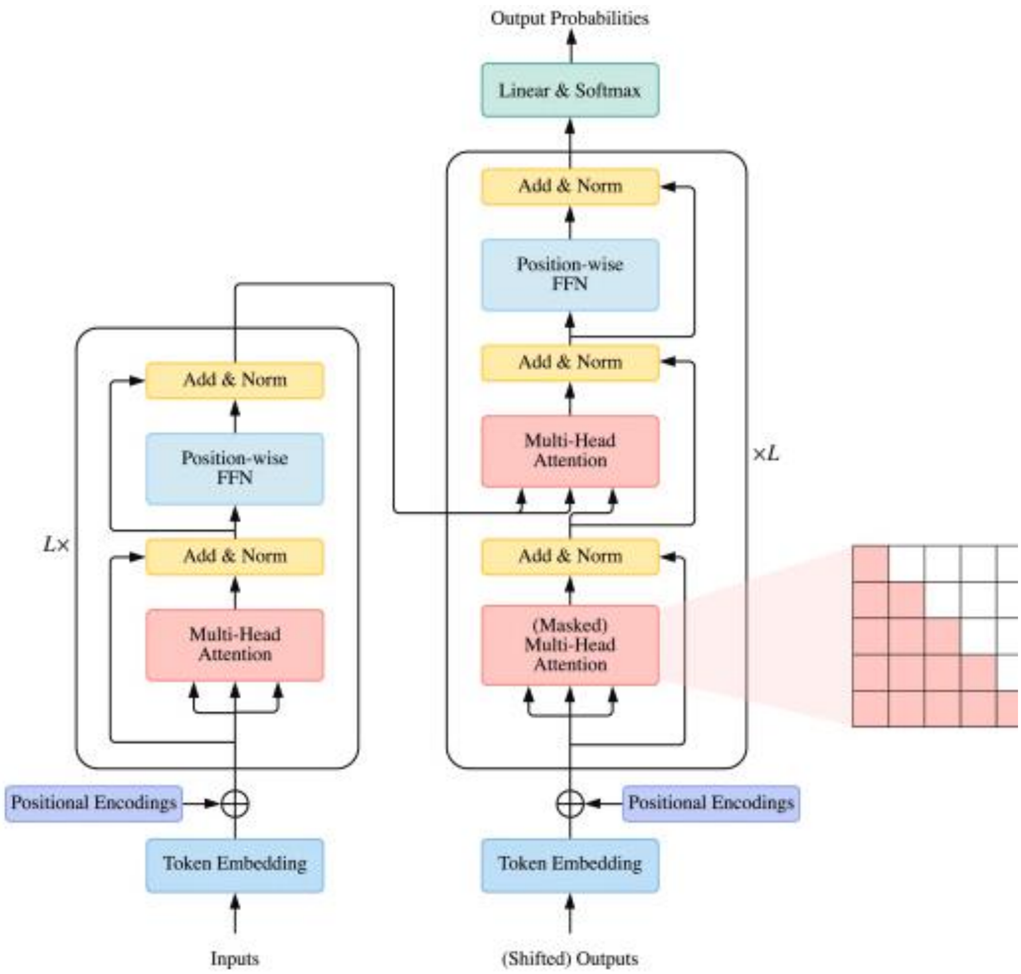



*Figure 12. Overview of the Vanilla Transformer architecture. Source of own elaboration from [45].*

## 2.6 Optimizers

There are algorithms that optimize the training of neural networks, allowing them to converge faster and achieve better performance, among these optimization algorithms are: Stochastic Gradient descent with momentum (SGD).

Stochastic gradient descent with momentum adds an additional term to the weight update rule. This term represents a velocity that accumulates from past gradients. By including a momentum, the gradient gains inertia during training, as if it were accumulating velocity. This velocity is calculated by summing the previous gradients weighted by a friction factor α. Momentum helps prevent the gradient from stalling at local minima or saddle points. This is because the step size depends not only on the current gradient but also on the velocity accumulated in previous iterations. In other words, momentum smooths the gradient trajectory and allows it to more easily escape from local optima to converge to better solutions. It gives stability and speed to the training [46].

**Adagrad**

AdaGrad maintains a cumulative sum of the squares of the historical gradients. When updating, it divides the gradient by the square root of that cumulative sum. This adaptively scales the learning rate for each parameter, further accelerates changes in directions with small gradients and slows down changes in directions with large gradients. Thus, AdaGrad balances the step size in all dimensions. In summary, this optimizer dynamically adjusts the learning rate parameter by parameter, but its decreasing aggressiveness can lead to premature convergence [46].

**RMSProp**

RMSProp is a modification of AdaGrad to prevent the learning rate from decreasing too quickly, what it does is to keep a moving average of the squares of the gradients instead of a cumulative sum and this is updated with a discount factor α which allows it to forget old gradients. In conclusion RMSProp adapts the learning rate parameter by parameter continuously during training [46].

**ADAM**

The Adam optimizer combines ideas from SGD with momentum and RMSProp in a single algorithm, which uses biased estimators of the first and second moments of the gradient to adapt the learning rate. The first moment provides the momentum effect as in SGD, while the second moment provides the gradient scaling effect as in RMSProp. It also includes bias corrections to avoid very large steps at the start of training. In summary, Adam combines momentum, gradient scaling and bias correction to obtain a robust and widely used optimization algorithm for deep learning [46].

## 2.7 Architecture of the Inception V3 model

The present project uses the Inception-v3 architecture, a convolutional neural network architecture from the Inception family that makes several improvements that are used for image classification. These include the use of label smoothing, 7 x 7 factored convolutions, and the use of an auxiliary classifier to propagate label information further down the network (along with the use of batch normalization for the lateral head layers). It has a total of 42 layers and a lower error rate than its predecessors [47].

In order to optimize the inception v3 model, a series of improvements were made to its predecessor models, among the most relevant ones it is important to mention:

**Factorization in smaller convolutions**

One of the pioneering models in the Inception family, known as Inception V1, was characterized by the use of large convolutions, which were subjected to factorization into convolutions of smaller dimensions. This approach was implemented in response to the considerable computational burden of the larger convolutions. In order to optimize computational efficiency, the strategy of replacing the 5x5 convolution layer with two

3x3 convolution layers was adopted. This choice allowed maintaining a significant level of representational capacity while reducing the computational requirements, resulting in a major advance in the Inception architecture [47].

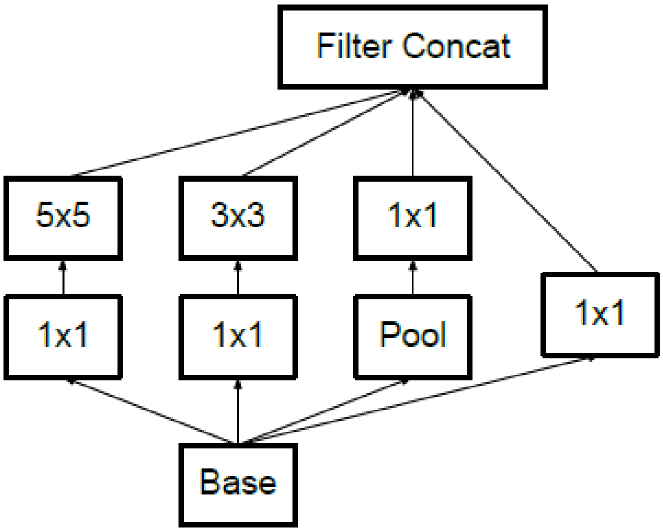



*Fig13. Inception V1 network model. Source of own elaboration from [47].*

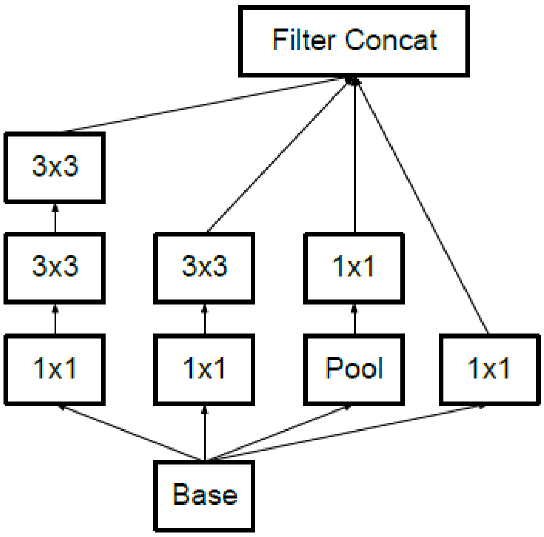



*Fig 14. Factored network model. Source of own elaboration from [47].*

### Spatial factorization in asymmetric convolutions

In order to improve the efficiency of the model, especially in the handling of asymmetric convolutions, a spatial factorization technique was implemented. This strategy focuses on optimizing convolutions of the form "nx1," where 'n' represents a considerable dimension in one direction and 1 in the other. The basic idea behind this technique was to replace the 3x3 convolutions by a combination of 1x3 convolutions followed by a 3x1 convolution [47]. The result of this improvement can be seen in the figure below.

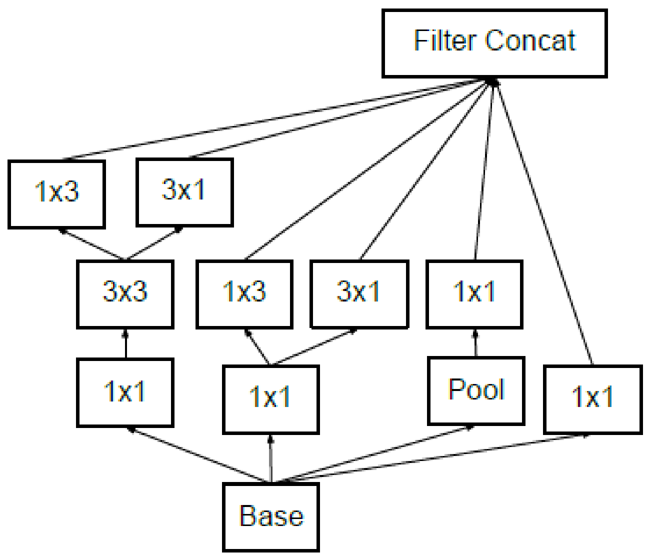



*Fig 15. Structure of asymmetric convolutions. Source of own elaboration from [47].*

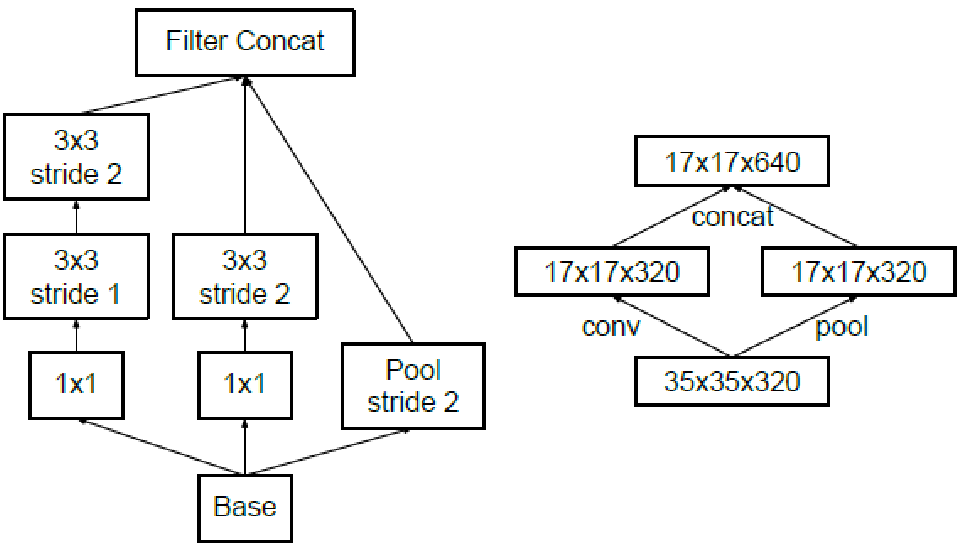
### Utility of auxiliary classifiers

The underlying purpose of incorporating an auxiliary classifier is to enhance the convergence of deep neural networks. This additional component plays a crucial role in addressing the evanescent gradient problem, a common obstacle in deep network optimization. In the initial stages of the training process, the inclusion of auxiliary classifiers brought no discernible improvements. However, as the training term approached, an appreciable increase in the accuracy of the model with these auxiliary classifiers was observed compared to the performance of the network lacking them [47].

**Efficient grid size reduction**

Historically, the maximum clustering and average clustering technique was used to reduce the grid size in the feature maps. However, in the initial iteration of the V3 model, a different strategy is chosen to achieve efficient grid size reduction. In this approach, the activation dimension of the grid filters is extended. For example, if we start from an ad×d grid with k filters after reduction, the result translates into an ad/2 × d/2 grid with a total of 2k filters. This process is executed by implementing two convolution blocks that operate in parallel and subsequently concatenate their results. This methodology represents a significant change in the dimension reduction strategy and offers a more efficient approach to adapt the grid size to the needs of the problem [47]

*Fig 16. Represents the new changes that were made in the model regarding the contact filter. Source*




*of own elaboration from [47].*

#### Network reduction

Diagram of the final model of inception V3. Finally, we can see the inception v3 model after all the improvements were made for better optimization.

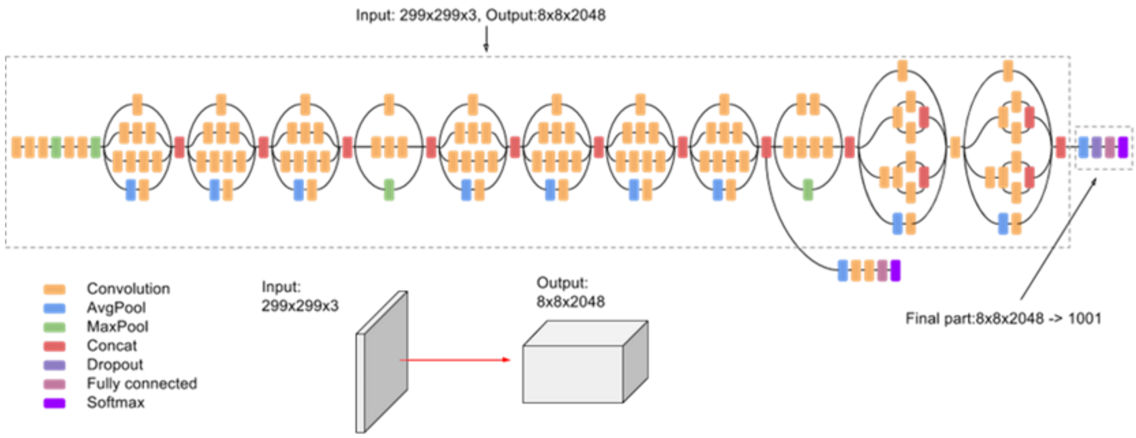



*Fig 17. Diagram of the structure of the final inception V3 model. Source of own elaboration from [48].*

The structure of the Inception V3 architecture is mentioned since it will be used later for the development of the neural network. As we can see, it has several features that make it ideal. Among them, it has faster inference that is achieved through the use of factorizations and batch normalization, which reduces the number of operations and parameters required without compromising performance. In addition, Inception V3 architecture enables highly discriminative feature extraction by effectively combining information at different scales. It also possesses flexibility and adaptability to different input sizes, which makes it suitable for medical images that may vary in dimensions and resolution. On the other hand, Table 3 presents a comparison with other architectures that also have high performance in learning techniques for mammography images, and among them, we observe that the Inception V3 architecture is very effective for this type of analysis, which is why it is used in this study.

## 2.8 Repositories

Over time, deep learning models involving CNNs have had a great impact on the classification of breast cancers due to their efficiency and efficacy, which is why the following are some of the research studies conducted in this field.

Authors in [49] this study proposes a CAD system using DCNN and SVM to classify breast tumors in mammography images on two datasets: DDSM and CBIS-DDSM. Using manual and threshold-based segmentation, different levels of accuracy were achieved. The highest accuracy was 87.2% with an AUC of 0.94 on the CBIS-DDSM set using SVM with mean Gaussian kernel. This approach stands out for its ability to improve the classification of benign and malignant tumors by using advanced machine learning techniques. On the other hand in [50] their researchers employed convolutional neural networks (CNN) to classify lesions as malignant or benign tumors in magnetic resonance imaging (MRI). A multilayer CNN architecture with online data augmentation was designed and the model was trained and tested. The results showed a high accuracy of 98.33%, sensitivity of 100%; and specificity of 96.88%. These findings suggest that CNNs may be a promising tool for accurate characterization and detection of breast tumors in magnetic resonance imaging.

The authors' research in [51] proposed the use of convolutional neural networks (CNNs) for automated detection of breast masses in mammography images in Ethiopia. The goal was to reduce overhead and improve the accuracy of manual diagnosis. The CNN architecture was designed to extract features and classify masses as benign or malignant. The model achieved a high detection accuracy of 91.86%, sensitivity of 94.67% and

AUC-ROC of 92.2% on the test data set. This suggests that CNNs have the potential to significantly improve early detection of breast cancer in Ethiopia

The authors in [52] describe a model based on deep recurrent neural networks for breast cancer prediction, specifically a stacked GRU-LSTM-BRNN model. This model uses medical records to predict the likelihood of breast cancer, outperforming other base classifiers such as simple RNN, stacked LSTM, and stacked GRU in terms of accuracy and other evaluation metrics. The Wisconsin breast cancer diagnostic dataset (Diagnostic) from UCI, consisting of 569 samples, was used. The proposed model achieved an accuracy of 97.34%, with an F1-score of 0.97, an MSE of 0.03 and a Cohen-Kappa score of 0.94, proving to be highly effective in the early detection of breast cancer.

The authors in [53] in this paper describe a study on breast cancer detection using thermography and thermal images processed by convolutional neural networks (CNN). A new algorithm is proposed to extract meaningful features from thermal images captured by a thermal camera, which are classified as normal or suspicious using CNNs optimized by Bayes algorithm. The study achieves an accuracy of 98.95% on a dataset of 140 individuals, demonstrating the effectiveness of thermography and CNNs in the early detection of breast cancer.

Authors in [54] conducted a study comparing shallow and deep convolutional neural networks (CNNs) using pretrained models tuned to classify whole mammogram images as benign or malignant. Two datasets, CBIS-DDSM and INbreast, were employed, achieving a significant accuracy of 82.1% and 83.3% respectively, especially with the tuned Xception model, which outperformed other classifiers. This approach proves to be effective for breast cancer detection, offering a promising path towards computer-aided diagnosis and personalized treatment in medicine.

The authors in [55] propose a Convolutional Neural Network (CNN) model to classify breast cancer lesions in automated breast ultrasound (ABUS) images, using multiview and transfer learning strategies with a modified Inception-v3 architecture. It was evaluated on 316 breast lesions, achieving an AUC of 0.9468, with sensitivity of 0.886 and specificity of 0.876. This approach outperforms traditional machine learning methods and proves to be a promising tool to assist in breast cancer classification on ABUS images.

Below is a table summarizing various research publications on the use of deep learning and machine learning techniques for mammography image analysis.

*Table3. Summarizes research publications on the use of deep learning and machine learning techniques for mammography image analysis. Source of own elaboration from [56].*

| Reference | Dataset | ML Method | Results |
|---|---|---|---|
| **[36]** | Mini-MIAS INBreast | SVM Classifier | Accurancy of 99% AUC value is 0.933 |
| **[52]** | DDSM | SVM Classifier | Sn value is 82.4% |
| **[37]** | DDSM | SVM Classifier | Accuracy is 98.9% |
| **[38]** | MIAS INBreast | SVM Classifier | Accuracy is 99% ± 0.50 AUC value 0.99 ± 0.005 |
| **[39]** | MIAS | SVM Classifier | Accuracy 93.17% |
| **[40]** | MIAS: 109 cases | SVM Classifier | Accuracy value from 68% to 100% |
| **[41]** | MIAS | RBFNN classifier | RBF (normal/abnormal) Accuracy is 93.9% Sn value is 97.2% RBF (benign/malignant) Accuracy is 94.3% Sn value is 100% |
| **[53]** | Private-1896 cases | GLCM SFFS (sequential floating forward selection) the bilateral CC and MLO view images | Sn-value is 68.8% Sp value is 95.0% The AUC value is 0.85 ± 0.046 |
| **[45]** | MIAS: 57 images 37 benign and 20 malignant | CNN classifier | Accuracy is 90.9% AUC value is 96.9% |
| **[46]** | MIAS -BancoWeb: 100 image | CNN and hybrid of K-means a | Accuracy 96% |
| **[42]** | DDSM | Fuzzy C-Means (FCM) | Accuracy is 87% Sn value is 90 to 47% Sp value is 84 to 84% [49] |

| | | | |
|---|---|---|---|
| | | | 252 images from Mini-MIAS -DDSM KNN Abnormality detecting: |
| **[49]** | 252 images from Mini-MIAS -DDSM | KNN | Abnormality detecting: Accuracy is 91.2% AUC value is 0.98 Malignancy detecting: Accuracy is 81.4% AUC value is 0.84 |
| **[50]** | 300 images from DDSM | Fuzzy Gaussian Mixture Model (FGMM) | Accuracy is 93% Sn value is 90% Sp value is 96% |
| **[44]** | IRMA-MIAS | k-NN | Accuracy is 92.8% ± 0.009<br>Sn value is 92.85% ± 0.01<br>AUC value is 0.971 |
| **[54]** | DDSM | CNN and transfer learning | Sensitivity of the mass 89.9% |
| **[55]** | DDSM, MIAS | LS SVM, KNN, Random Forest, and Naive Bayes | Accuracy 92% |
| **[56]** | (Mini-MIAS) DDSM | CNN | The accuracy of 0.936, 0.890, 0.871 on the DDSM, 0.944, 0.915, 0.892 on the Mini-MIAS for normal, benign, and malignant regions |
| **[48]** | MIAS | CNN a pre-trained architecture such as Inception V3, ResNet50, Visual Geometry Group networks (VGG)-19, VGG-16, and Inception-V2 ResNet | Overall Accuracy, Sn, Sp, precision, F-score, and AUC of 98.96%, 97.8%, 99.1%, 97.4%, 97.7%, and 0.995, respectively, for the 80–20 method and 98.87%, 97.3%, 98.2%, 98.84%, 98.04%, and 0.993 for the 10-fold cross-validation method, the TL of the VGG16 model is adequate for diagnosis |
| **[57]** | DDS | CNN | Accuracy 71.4% |

# CHAPTER 3

## 3. Materials and methods

In this chapter, we will focus our attention on presenting the development of the Breast Cancer Detection System using neural networks. In the following, the operation of this system is described in detail. First, it starts with the use of a database that acts as the main input of the system. Next, an artificial neural network is implemented to process the input data. Finally, the system produces a result that is related to the classification and recognition of the data obtained.

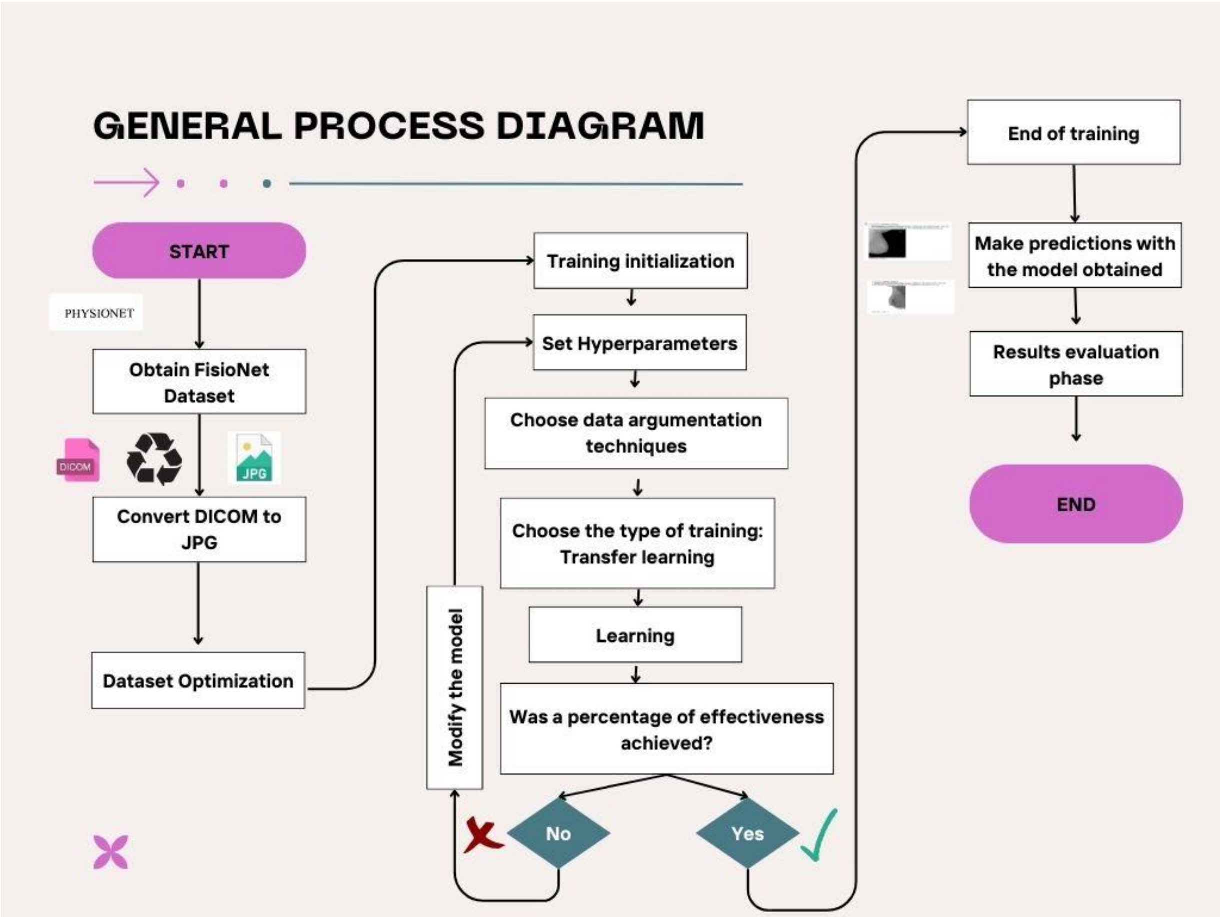



*Fig 18. Diagram detailing the CNN development process in general terms.*

## 3.1 Dataset

For the present project, we opted for the application of supervised automatic learning. In this approach, the algorithm is trained by means of a database which contains examples of desired inputs and outputs, and in this way, the main objective of the algorithm can learn to map the correct inputs and outputs and thus finally be able to make fairly accurate predictions in unseen data.

The database for this study was obtained from the FisioNet portal, which consists of a large-scale reference dataset of full-field digital mammography, called VinDr-Mammo consisting of 5000 four-view exams with breast-level assessment, giving a total of 20000 images. The result of reading this data set includes both general breast assessment and information about abnormal regions of the breast. This image output follows the Breast Imagin Reporting and Data System (BI-RADS) schema and lexicon [57].

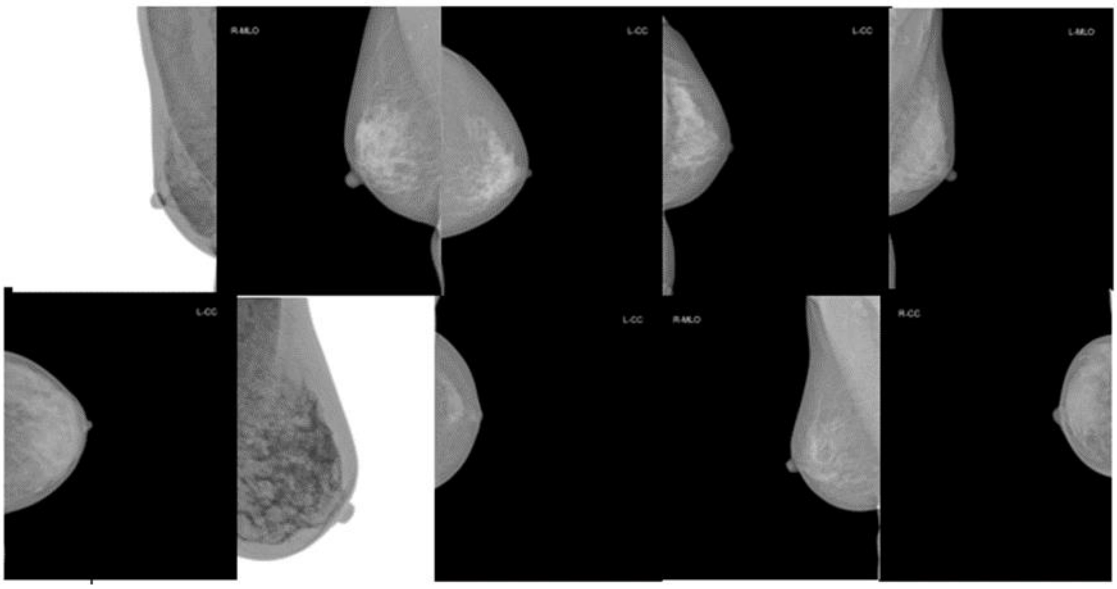


*Fig 18. Sample images of the complete Dataset. Source of own elaboration from [58].*

## 3.2 Dataset optimization

### 3.2.1 Dicom to jpg format conversion

The database obtained from the physionet portal presents the set of mammograms in DICOM format, since this is the most widely used format for the transmission, storage and management of medical images and related data. However, since we need a format that computers are able to read without any additional software, the decision was made to convert the images to jpg. In order to obtain the conversion, a Python code was created and after that the complete data set could be obtained in the desired format.

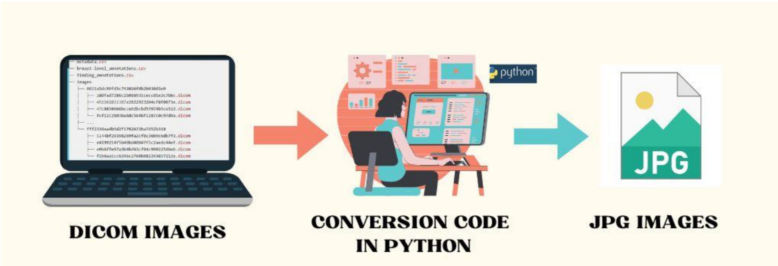



*Fig 19. Schematic of the dataset format conversion.*

### 3.2.2 Restructuring and simplification of the dataset

The database included a wide range of mammograms, of cases classified as negative and positive for breast cancer, as well as others that required additional analysis to confirm or rule out the presence of the disease.  For the ANN training and simulation process, a restructuring of the original database was performed. In order to carry out the selection process of the images of interest, a Python code was created, the complete Dataset was loaded and in parallel an Excel document was used where the codes corresponding to the different classes of images that needed to be classified were found, once the relevant images were identified through the code, they were extracted and stored in a new folder. Thus, after the segmentation and final categorization of the images, the new Dataset was organized as follows:

*Table 4: New Dataset after Restructuring.*

| | Classes | |
|---|---|---|
| **Dataset** | **Positive cases** | **Negative cases** |
| **Training** | 1000 | 1000 |
| **Validation** | 500 | 500 |
| **Testing** | 150 | 150 |

## 3.3 Hyperparameters

As we know the hyperparameters are parameters that the researcher or user establishes prior to the training of the neural network and in this case, they were the following:

**Number of epochs**

This parameter indicates the number of times the training and validation sets are processed by the convolutional neural network.

**Batch size**

Training convolutional neural networks using massive image data sets requires resource optimization. Processing all images simultaneously would overload memory and computational capacity, even on the most advanced GPUs. Therefore, strategically dividing the images into more manageable batches is ideal. Each batch contains a number of images that are fed together to the neural network for training. The optimal batch size depends on the balance between GPU capacity, amount of data and desired training time, so it is a trial and error process to find the ideal size, but identifying it makes it possible to train the neural network on massive data sets in an efficient way.

**Loss Function**

The Loss Function measures the discrepancy between predictions and actual labels.

**Optimizer**

The optimizer is considered as a kind of "engine" that drives the learning of the neural network. Its function is to be able to minimize the cost function to find the optimal weights and biases.

## 3.4 Framework and hardware acceleration

**ML framework**

The ML framework, or machine learning framework, to assist in the development, implementation and management of the machine learning model was PyTorch.

**Hardware acceleration**

For model training and validation, a high-performance computing environment equipped with an NVIDIA Tesla T4 GPU was used due to the Tesla T4's renowned processing capabilities, especially suited for intensive machine learning and medical image analysis applications.

The GPU has a substantial memory of 15360 MiB (approximately 15 GB), which enables the handling of the extensive mammography dataset. The GPU driver version installed is 535.104.05 and worked with CUDA version 12.2, ensuring compatibility and optimal utilization of the GPU's processing capabilities. This GPU technical description and additional details are shown in Fig19.

This hardware environment provides power and efficiency for processing and analyzing mammography images, as well as the stability and reliability required for the execution of deep learning algorithms in this case.

```
Mon Jan 15 13:47:46 2024
+---------------------------------------------------------------------------------------+
| NVIDIA-SMI 535.104.05             Driver Version: 535.104.05   CUDA Version: 12.2     |
|-----------------------------------------+----------------------+----------------------+
| GPU  Name                 Persistence-M | Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |         Memory-Usage | GPU-Util  Compute M. |
|                                         |                      |               MIG M. |
|=========================================+======================+======================|
|   0  Tesla T4                       Off | 00000000:00:04.0 Off |                    0 |
| N/A   35C    P8               9W /  70W |      0MiB / 15360MiB |      0%      Default |
|                                         |                      |                  N/A |
+-----------------------------------------+----------------------+----------------------+

+---------------------------------------------------------------------------------------+
| Processes:                                                                            |
|  GPU   GI   CI        PID   Type   Process name                            GPU Memory |
|        ID   ID                                                             Usage      |
|=======================================================================================|
|  No running processes found                                                           |
+---------------------------------------------------------------------------------------+
```

*Fig 20. Technical description of the GPU environment*

## 3.5 Data Argumentation

Data augmentation, an invaluable technique in the field of machine learning, is often used when there is an insufficient or small training data set when training the CNN. Artificial data augmentation allows increasing the size and diversity of the training set by means of certain common techniques such as:

**Random rotations:** This technique helps the neural network to be orientation invariant.

**Horizontal and vertical flips:** This technique helps the network learn robust features that are independent of image symmetry.

**Cropping with different width and height:** This method can be especially useful for teaching the network to identify important features without relying on the full context of the image.

**Changes in saturation and illumination:** This setting is of great relevance as it helps the network to be more robust to changes in the visual quality of the images.

**Noise:** Adding noise can help the network to be more resistant to variations and artifacts that might be present in the actual data.

**Blurring:** Blurring can help the network focus on important features while ignoring fine details that may not be relevant.

If these transformations are applied at the same time, new information is being introduced to the training data set. And in the end, you would obtain an extended Dataset that reduces overfitting and above all improves the generalization capability of the trained models.

## 3.6 Transfer learning

Transfer Learning is a technique that takes advantage of deep learning models previously trained for similar tasks, i.e. instead of creating a model from scratch, it trains current models with previously trained models. This saves time and resources, since the knowledge acquired in the previous model can be reused. Basically, what is done is to update and retrain the neural network for the new task. Transfer Learning is useful in areas such as image classification, object detection and speech recognition. It allows the creation of efficient models in a faster and slightly less complex way [59].

Here we can mention the classification models available in PyTorch Hub among these is Inception V3. These models have been previously trained to perform classification on a large amount of data, such as ImageNet.

Two types of transfer learning can be used, which are:

**Fine tuning:** In this type of tuning what is done is to update the parameters (weights and biases) of all or some of the layers of the model during training. This means that the pre-trained model is adapted more deeply to the new task, allowing the model to adjust all its learned features to the new data.

**Feature extract:** In this setting, only the parameters of the last added or modified layers of the model are updated, while the rest of the pre-trained model remains unchanged. This approach is commonly used in transfer learning, where the learned features of an already trained model are leveraged and customized for a new task with less data.

## 3.7 Confusion matrix

Once a supervised ML model is trained on a dataset, it is tested by using data that is retained from the training process. The confusion matrix is a tool that allows analyzing the classification accuracy and detecting error patterns. This matrix compares the classes predicted by the model with the actual classes in the test data, providing a comprehensive view of classification performance and failures. A well-analyzed confusion matrix identifies ways to improve the model and its generalizability to new data [60].

The comparatives are true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). In Figure 14 we can observe a confusión matrix.

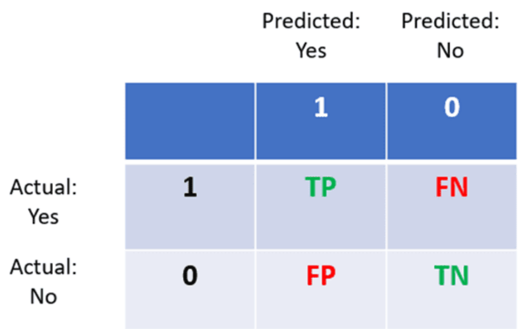



*Fig 21. Confusion matrix. Source of own elaboration from [61].*

## 3.8 Acquisition of Metrics

Metrics are functions, or measures, that are used to accurately assess the performance of a model. The confusion matrix acts as a foundation, providing a detailed breakdown of the results. This allows the calculation of specialized and detailed metrics tailored to the specific requirements of the analysis.

1. **Accuracy:** Accuracy is a metric that helps you know how exact or close the result is to the true value [62].

$$accuracy = \frac{all\ correct\ predictions}{all\ samples} \qquad accuracy = \frac{TP+TN}{TP+FN+FP+TN}$$

2. **Precision:** It measures how accurate the neural network's predictions are through the percentage of correct predictions [62].

$$precision = \frac{TP}{TP + FP}$$

3. **Recall:** It is a measure that allows us to know the proportion of positive cases that were correctly classified [63].

$$recall = \frac{TP}{TP + FN}$$

4. **F1-Score:** It is an average of precision and Recall, f1_score assigns the value 1 to the best value and 0 to the worst, it is represented as follows [62].

$$F1_{score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

5. **Specificity:** These are the negative cases that have been correctly classified [64].

$$Specificity = \frac{TN}{TN + FN}$$

# CHAPTER 4

## 4. Results and discussion

In this chapter we focus on presenting the tests, experiments and results obtained during the training, simulation and testing phase of the neural network model.

### 4.1 Experiment I: Testing the deep learning classification model with different optimizers with different optimizers

In order to optimize the performance and evaluate the performance of different optimizers on the mammogram classification task for breast cancer detection. of the model, tests were performed with various optimizers, specifically Adagrad, SGD, ADAM and RMSprop. The model was evaluated by training it for 25 complete epochs with each optimizer, keeping all other hyperparameters fixed. The objective of this study is to evaluate the performance of different optimizers on the task of mammogram classification for breast cancer detection.

**Loss during Training:**

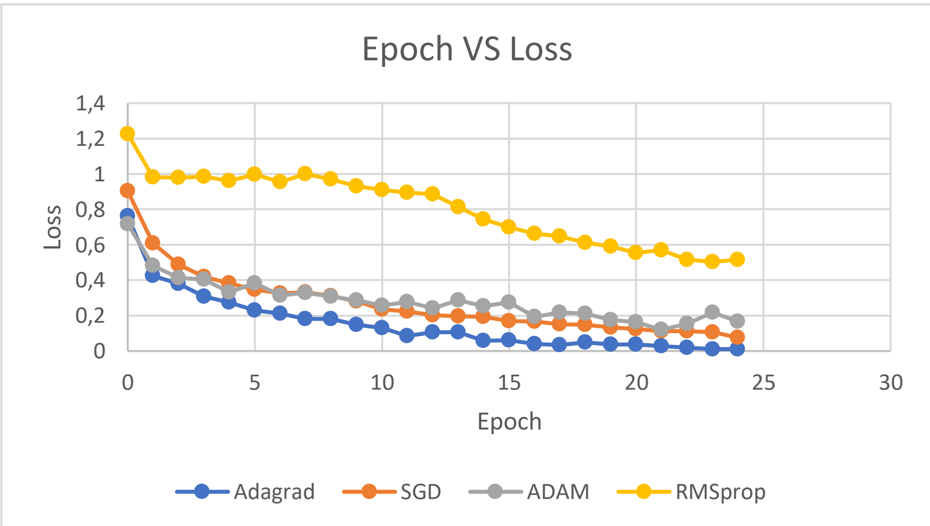



*Fig 22. Loss during Training: using 4 different optimizers.*

As seen in Fig21. During the 25 epochs, all four optimizers showed a decrease in the loss function. Adagrad started with a loss of 0.763 and ended at 0.0109, demonstrating a rapid initial improvement that plateaued toward the end of training. SGD showed a more moderate but consistent decline, starting at 0.9031 and reducing to 0.0757. ADAM and RMSprop started with losses of 0.7172 and 1.2263, respectively, and both showed significant improvements, with ADAM ending at 0.1661 and RMSprop at 0.5157, although the latter experienced some fluctuations throughout the process.

**Accuracy during training:**

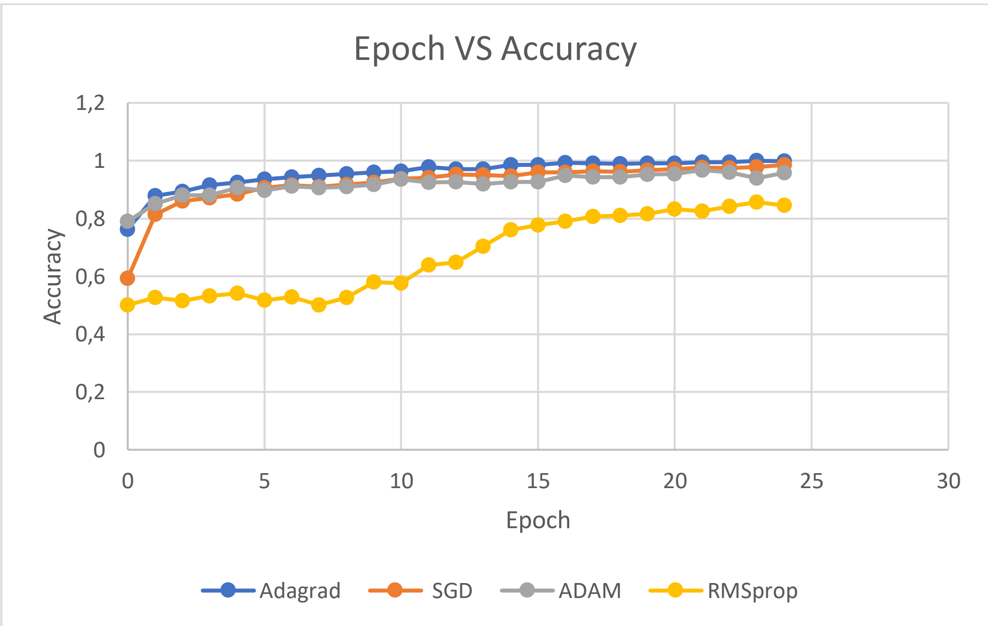



*Fig 23: Precision during Training, using 4 different optimizers.*

Adagrad accuracy increased significantly from 0.763 to 0.9985, suggesting effective learning capability. SGD improved from 0.5925 to 0.9850, ADAM from 0.79 to 0.957, and RMSprop from 0.501 to 0.846, indicating that all optimizers improved in correctly classifying images over time. However, the consistency of SGD and the high final accuracy of Adagrad were noteworthy.

**Model Validation Results**

**Loss during Validation:**

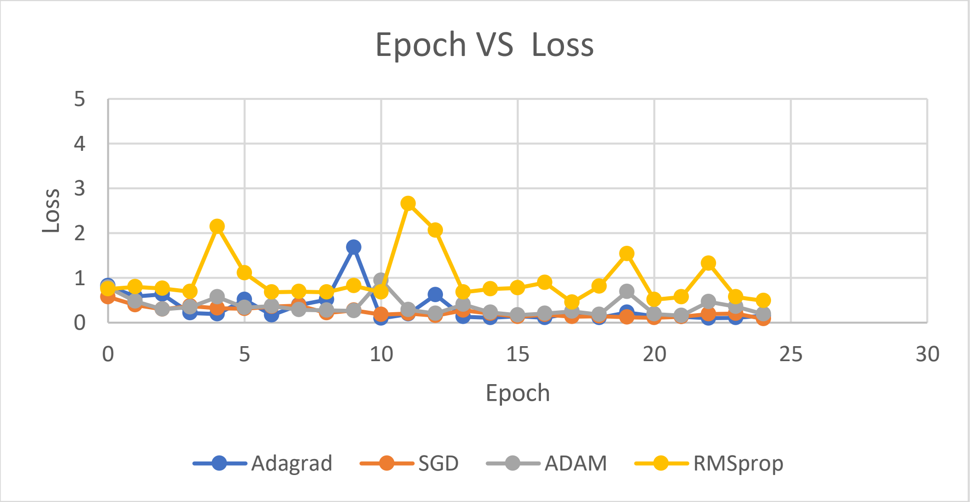



*Fig24: Loss during validation, using 4 different optimizers*

The validation loss reflects how the model generalizes to unseen data. Adagrad showed an initial loss of 0.8278 and, despite some fluctuations, improved to 0.1528. SGD started at 0.5757 and ended at 0.0829, showing an overall improving trend with less volatility than Adagrad. ADAM and RMSprop started with similar losses of 0.7907 and 0.7548, respectively, but ADAM showed a more consistent decrease in loss across epochs, while RMSprop experienced significant spikes, which may be an overfitting signal or sensitivity to the specific characteristics of the validation data set.

**Accuracy during Validation**

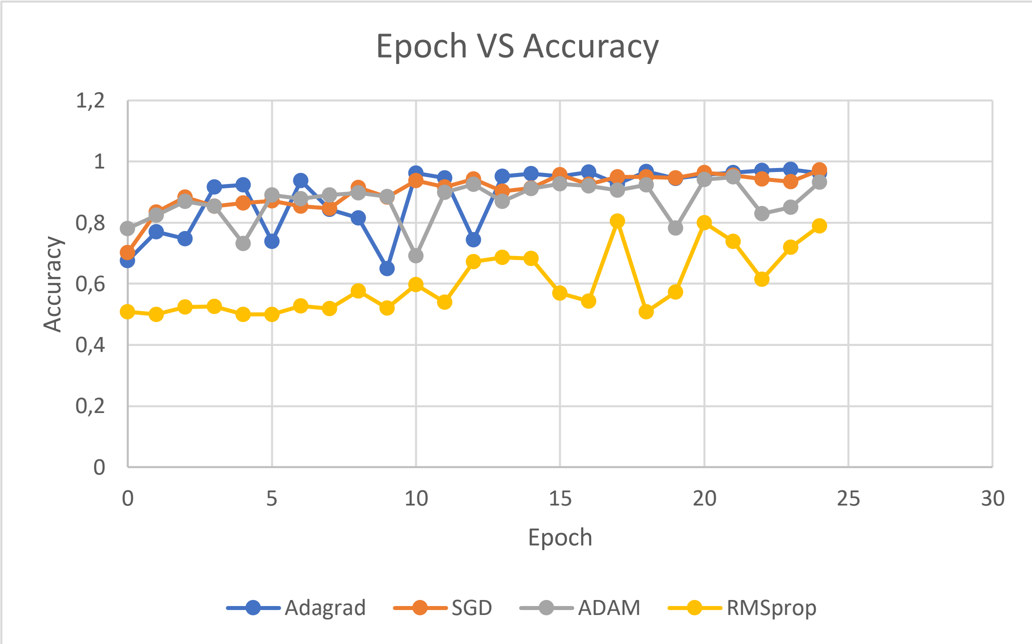



*Fig 25: Accuracy during validation, using 4 different optimizers.*

Adagrad improved from an accuracy of 0.676 to 0.962, while SGD showed a consistent improvement from 0.7020 to 0.9720. ADAM started with 0.78 and ended with 0.932, showing some fluctuations. RMSprop, although starting with the lowest accuracy of 0.509, improved to 0.79, but its trajectory was the least consistent, suggesting lower reliability in generalization.

**Analysis**

Overall, Adagrad and SGD were the optimizers that showed the most promising results, due to their high final accuracies and consistent reduction in loss in both training and validation. On the other hand, ADAM showed a good ability to minimize loss and a robust final accuracy, but exhibits higher variability than Adagrad and SGD. While RMSprop, although able to achieve a reduction in loss and an improvement in accuracy, showed the highest volatility, which raises some concern about its reliability in different data sets or under varying validation conditions.

Thus, the stability of SGD makes it a fairly reliable candidate for longer training and robust generalization, which is why it was chosen for neural network training. Adagrad seems to be suitable for situations where fast convergence and high accuracy are required. ADAM, with its overall balanced performance, may be preferred in situations where a compromise between convergence speed and stability is sought. Finally, RMSprop may be more suitable for initial explorations and experimentation with different hyperparameter settings due to its fast initial response, although its variability requires more cautious analysis.

With respect to data expansion for training the machine learning model, different Torchvision transformations were used. In addition, other techniques were performed as described below.

**For Training:**

- Random rotation of the images randomly up to 5 degrees. This helps the model to be more robust to variations in the orientation or angle of the images.
- Resize all images to 299x299 pixels. This change should be applied since the input size of the Inception model is 299x299.
- Horizontal flipping of the images with a probability of 50%. This also helps to make the model invariant to certain transformations.
- The images were normalized using the means and standard deviations given for each channel (RGB).

**For Validation:**

- Resizing of all images to 299x299 pixels.
- Normalization of the images, but no rotations or flips were applied, as the validation images should reflect real data without alterations.

### 4.1.2 Training

The training was performed with the inception V3 model and the components were the following:

**Loss function:** Cross Entropy

**Optimizer:** SGD (Stochastic Gradient Descent)

**Number of Epochs:** 25

**Batch Size:** 256

**Classes:** 2

**Full training time:** 100 mn 21s

## 4.2 Experiment II: Results evaluation phase

In this phase what will be done is to evaluate the trained model to verify the efficiency, but above all its ability to generalize, remember that its task is to make predictions regarding the detection of breast cancer through mammograms. For this purpose, a completely new test dataset was used, which includes 300 mammograms that the neural network has not seen before.

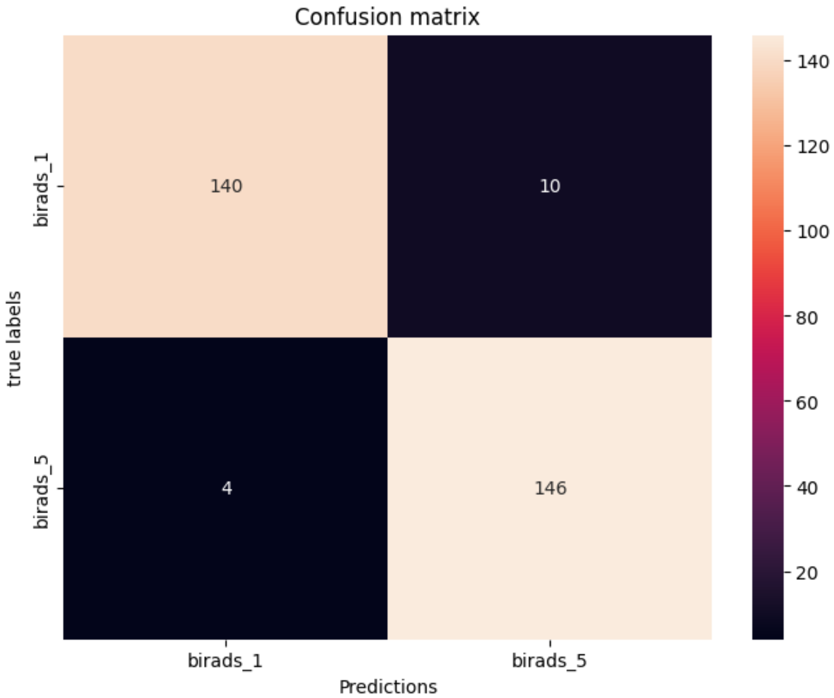



*Fig 26. Neural network confusion matrix using Inception V3 architecture.*

In Fig 25, we observe the confusion matrix that visualizes the performance of the classification model, basically it compares the true labels (real observations) with the predictions made by the neural network. In this case, the classes are:

**Birads_1:** corresponding to Normal category.

**Birads_5:** which corresponds to Carcinogenic category.

The analysis of the confusion matrix obtained from the model indicates that a fairly effective classification of mammograms into 'Normal' and 'Cancerous' categories was obtained. Out of a total of 300 images evaluated, the model correctly identified 140 as 'Normal' and 146 as 'Cancerous'. However, there were 10 cases where 'Carcinogenic' images were misclassified as 'Normal', and 4 cases of 'Normal' images classified as 'Carcinogenic'. These results imply a high accuracy rate of the model, with 95.33% (Fig

20) of classification accuracy, indicating the potential and reliability for breast cancer detection through mammography techniques.

```
Classes:  {0: 'birads_1', 1: 'birads_5'}
Number of images:  300
<class 'list'>
<class 'list'>
Len labels: 300
Len predictions: 300
Test accuracy : 0.9533333309491475
Model: inception-Cross_SGD_25epochs_breast_FT_transfors
<Figure size 640x480 with 0 Axes>
```

*Fig 27. Accuracy obtained by the neural network model with the test dataset.*

Below we can see in the table additional metrics such as precision, recall, F1-score and specificity that were calculated with results from the confusion matrix. These values obtained show that the model has a very high performance in detecting both classes. As we can see, the values are high in all metrics, which suggests that the model is reliable and effective in distinguishing between normal and carcinogenic mammograms. Furthermore, the balance between the metrics (as seen in the close F1-Score values) indicates that there is no significant bias toward detecting one class over the other, which is critical in medical applications where both false negatives and false positives can have serious consequences.

In the context of a medical diagnosis, these results are encouraging because they show that the model is able to identify with high precision and sensitivity cases that require additional attention, while minimizing cases that could lead to unnecessary concerns or additional procedures not necessary.

*Table 5. Additional metrics for model evaluation.*

| | Precision | Recall | F1-score | Specificity |
|---|---|---|---|---|
| Birads_1 | 0.972 | 0.933 | 0.952 | 0.973 |
| Birads_5 | 0.935 | 0.973 | 0.953 | 0.933 |

It is important to mention that something fundamental in the development of a deep learning model is the correct labeling of the data since this will allow the model to be trained correctly and produce real results, that is, accurate predictions.

## 4.3 Repositories Results

In this section, we will perform a critical review of four relevant works with an approach similar to ours, that is, breast cancer detection using neural networks. This review will allow us to establish a dialogue with the authors of these works and compare the results obtained in our own study.

The authors in [ 65] use a Deep Neural Network (DNN) for automated breast cancer classification, using recursive feature elimination (RFE) for feature selection automated breast cancer classification, employing recursive feature elimination (RFE) for feature selection. While in [ 66] they implement the MobileNetV2 model for binary classification of mammography images, exploring both transfer learning and learning from scratch. On the other hand, in [67 ] the authors propose a breast cancer detection framework based on RetinaNet and two-stage transfer learning to improve mammogram mass detection performance. Finally, the authors in [68] employ convolutional neural networks (CNNs) with transfer learning techniques, using pre-trained models such as VGG16, VGG19, ResNet-50, and Inception-V3.

*Table 6. Comparison of main features between the 4 repositories and the model developed in this work.*

| Characteristics | Repository1 | Repository2 | Repository3 | Repository4 | Developed model |
|---|---|---|---|---|---|
| Accuracy | 98.62% | 99.4% | Improvement in TPR and reduction of FPPI | Up to 97% with ResNet-50 | 95.3% |
| Dataset | WBCD (UCI) | DDSM | COCO, CBIS-DDSM, INbreast | BreaKHis (Kaggle) | VinDr-Mammo |
| Training Type | There is no mention of using pre-training techniques; It appears to be a model trained from scratch. | Transfer learning is explored using MobileNetV2, indicating the use of pre-trained models. | Two-stage transfer learning, which involves the use of pre-trained models as part of the approach. | Transfer learning techniques are used with pre-trained models such as VGG16, VGG19, ResNet-50, and Inception-V3. | Transfer learning with Inception v3 pre-trained model |
| Observations | High classification accuracy using the Wisconsin Breast Cancer dataset, demonstrating the effectiveness of DNN on tabular data sets. | This approach combines transfer and ground-up learning to achieve high accuracy on mammography images, indicating a strong ability to adapt to different data structures. | Using a two-stage approach to transfer learning allows for more refined and accurate detection of masses on mammograms. | The use of pre-trained models and transfer learning techniques to achieve high accuracies with histopathological images indicates excellent generalization capacity. | A high level of precision was achieved in the results, which is a solid indicator of the notable generalization capacity of the proposed model. |

## 4.4 Ethical Considerations

In the first instance we must be clear that data security and privacy are crucial in medical research, given that sensitive information from medical records (mammograms) of cancer patients is being manipulated. To protect individuals, the publicly accessible Dataset obtained from the physionet portal had anonymization and data encryption protocols, which ensured that personal information could not be linked to individuals, thus mitigating the risks of privacy violations and misuse of data. Another of the points to be considered within ethics was transparency and ease of understanding, i.e. explainable artificial intelligence techniques were used to break down the model's decisions so that it is easy to understand how it works. This not only facilitates understanding and confidence in the tools developed, but also by further developing this model and taking it to the application, doctors could make decisions based on the recommendations of the model, thus improving the collaboration between humans and machines in the diagnosis of breast cancer.

# 5. Conclusions

The present work shows the development of a convolutional neural network for breast cancer detection by analyzing mammograms. The model classifies the images into two categories: with cancer and without cancer.

The research began with an exhaustive literature review on breast cancer and the difficulty that specialists have in visually detecting certain subtle patterns indicative of the disease in its early stages. Hence the great relevance of having an artificial intelligence model to support and improve the effectiveness of medical diagnosis.

Subsequently, an extensive dataset of 20,000 images was obtained from the Physionet portal. This contained positive cases, discarded negative cases and others that required further studies to confirm or rule out the disease. A thorough cleaning and reorganization of the dataset was performed, selecting only positive and negative cases.

The Deep learning model implemented was based on Google's well-known Inception V3 architecture. Multiple optimizers were evaluated to measure their performance on the binary classification task, finally choosing SGD for achieving 97% accuracy in the tests. In addition, several Torchvision data transformations were employed in both training and validation to significantly improve accuracy and overall performance. Other techniques applied were: noise cleaning, manual relabeling, dataset balancing, and data augmentation using transformations. This allowed training a more robust and variation resistant model.

In this way, a highly optimized convolutional neural network model with Inception V3 architecture was obtained, which achieves a detection accuracy of 97%, using SGD as optimizer and certain finely tuned hyperparameters, as detailed in the results section.

In conclusion, this work presents a deep learning model capable of efficiently supporting the early diagnosis of breast cancer from mammograms, with a high level of accuracy. The results obtained are promising and have the potential to significantly improve the timely detection of this type of cancer.

## 5.2 Limitations

One of the significant limitations encountered during neural network development was the restricted access to large mammography datasets. The availability of a considerable volume of mammographic images is crucial for the effective training of neural network-based models, as these require a substantial amount of data to optimally learn relevant patterns and features. Above all, data imbalance is an unequal distribution of benign and malignant cases in the available datasets.

## 5.3 Future Work

For future work one can have the approach that as the model has demonstrated high accuracy, it would be important to validate its performance with clinical studies in collaboration with medical centers and healthcare professionals. This would help to evaluate the applicability of the model in a real clinical setting and its integration into existing workflows for breast cancer diagnosis. Similarly expanding the dataset to include mammograms from diverse breast cancer populations and subtypes may improve the model's ability to generalize and detect a wider range of disease manifestations and finally it would be of great importance to continue the search for optimal hyperparameters to further improve the accuracy and robustness of the model.

# 6. References


1. World Health Organization. (12 of July 2023). Breast cancer. https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer
2. National Cancer Institute (2019). Breast cancer. https://www.cancer.gov/espanol/tipos/seno/pro/deteccion-seno-pdq#_89_toc
3. Tabár L, Faberberg G, Day NE, Holmberg L. What is the optimum interval between mammographic screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. Br J Cancer. 1987 May;55(5):547-51. doi: 10.1038/bjc.1987.112.
4. Elmore JG, Jackson SL, Abraham L, Miglioretti DL.et al. (2009) Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy. Radiology. doi: 10.1148/radiol.2533082308.
5. W.A. Berg, R.E.H. PhD, D.B. Kopans, P. Robert, A. Smith, Frequently asked questions about mammography and the uspstf recommendations: a guide for practitioners, Soc. Breast Imag. (2020) 1–17.
6. J.J. Fenton, et al., Influence of computer-aided detection on performance of screening mammography, N. Engl. J. Med. 356 (14) (Apr. 2007) 1399–1409, https://doi.org/10.1056/nejmoa066099.
7. C.D. Lehman, R.D. Wellman, D.S.M. Buist, K. Kerlikowske, A.N.A. Tosteson, D. L. Miglioretti, Diagnostic accuracy of digital screening mammography with and without computer-aided detection, JAMA Internal Med. 175 (11) (Nov. 2015) 1828, https://doi.org/10.1001/jamainternmed.2015.5231.

8. A.D. Trister, D.S.M. Buist, C.I. Lee, Will machine learning tip the balance in breast cancer screening? JAMA Oncol. 3 (11) (Nov. 2017) 1463, https://doi.org/10.1001/jamaoncol.2017.0473.
9. Society for the Fight against Cancer in Ecuador (2022). International Day for the Fight against Breast Cancer. https://www.solca.med.ec/dia-internacional-de-la-lucha-contra-el-cancer-de-mama-19-de-octubre/#:~:text=Seg%C3%BAn%20estad%C3%ADsticas%20actuales%20del%20Observatorio,la%20cifra%20est%C3%A1%20en%20411.
10. Alkabban FM, Ferguson T. Breast cancer (2020). In: StatPearls [Internet]. Treasure Island (FL): Published by StatPearls.https://www.ncbi.nlm.nih.gov/books/NBK482286/.
11. Borchartt TB, Conci A, Lima RCF, et al. Breast thermography from an image processing viewpoint: a survey. Signal Process 2013;93:2785–803. https://doi.org/10.1016/J.SIGPRO.2012.08.012.
12. National Breast Cancer Foundation, Breast Anatomy and How Cancer Starts | About Breast Cancer, (2019). https://nbcf.org.au/about-national-breast-cancer-foundation/about-breast-cancer/what-you-need-to-know/breast-anatomy-cancer-starts/.
13. Leila Abdelrahman, Manal Al Ghamdi, et al.(2021).Mottaleb,Convolutional neural networks for breast cancer detection in mammography: A survey,Computers in Biology and Medicine,Volume 131,2021. https://doi.org/10.1016/j.compbiomed.2021.104248.
14. M Bretthauer , M Kalager, Principles, efficacy and caveats in cancer screening, British Journal of Surgery , volume 100, number 1, January 2013, pages 55-65, https://doi.org/10.1002/bjs.8995

15. Siu AL. U.S. Preventive Services Task Force, Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2016;164:279–96. https://doi.org/10.7326/M15-2886

*16.* Neyra,P. (2013, January). Breast density and risk of breast cancer. Revista Médica Clinica Las Condes,24(1),122. https://www.elsevier.es/es-revista-revista-medica-clinica-las-condes-202-articulo-densidad-mamaria-riesgo-cancer-mamario-S0716864013701378

*17.* McCormack,V. Dos Santos, I. (2006, June). Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiology, Biomarkers & Prevention,15(06), 1159-1169. https://doi.org/10.1158/1055-9965.EPI-06-0034

*18.* Boyd, N, Martin, L, Yaffe, Martin, et al. (2009, Diciembre). Mammographic density. *Breast Cancer Research,11*(S3). https://doi.org/10.1186/bcr2423

19. Aibar, L, Santalla, A, Criador, M, et al. (2011, July - August). Radiological classification and management of breast lesions.Elsevier, 38(4). 10.1016/j.gine.2010.10.016

20. Copeland B. Inteligencia artificial. Enciclopedia Británica. 2022. Recuperado de: https://www.britannica.com/technology/artificial-intelligence

21. Mukhamediev, R.I.;Symagulov, A.; Kuchin, Y.; Yakunin, K.; Yelis, M. From Classical Machine Learning to Deep Neural Networks: A Simplified Scientometric Review. Appl. Sci. 2021, 11, 5541. https://doi.org/10.3390/app11125541

*22. Lanzagorta-Ortega, Dioselina, Carrillo-Pérez, Diego L., & Carrillo-Esper, Raúl (2022). Artificial intelligence in medicine: present and future. Gaceta médica de México, 158(Suppl. 1),* 17-21. Epub January 20, 2023.https://doi.org/10.24875/gmm.m22000688.

23. Castelo, G, Urda,B, Nuñez, I. (2022). Artificial intelligence in biomedicine. AmbioSciences,20,7- https://revpubli.unileon.es/index.php/ambioc/article/view/7484/5935
24. Witowski, J., Heacock, L., Reig, B., Kang, S. K., Lewin, A. et al. 2022. Improving breast cancer diag-nostics with deep learning for MRI. Science Translational Medicine 14(664):eabo4802.
25. Lüthy, Isabel A.. (2022). Artificial intelligence and machine learning in cancer diagnosis and treatment. Medicina (Buenos Aires), 82(5), 798-800. Retrieved on January 13, 2024, from http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0025-76802022000900798&lng=es&tlng=es.
26. Nassif, A.B.; Shahin, I.; Attili, I.; Azzeh, M.; Shaalan, K. Speech recognition using deep neural networks: Una revisión sistemática. IEEE Access 2019, 7, 19143-19165. [CrossRef]
27. Brown, S. (April 21, 2021). Machine learning, explained. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained#:~:text=Machine%20learning%20is%20a%20subfield%20of%20artificial%20intelligence%2C%20which%20is ,to%20how%20humans%20solve%20pro%20blems
28. Shahzadi, N. (2023). Supervised Machine Learning: Classification and Regression. Medium.https://medium.com/@nimrashahzadisa064/supervised-machine-learning-classification-and-regression-c145129225f8
29. Bishop, C. M., Pattern Recognition and Machine Learning, Springer, USA, (2006).
30. Mostafa Ibrahim (February 28, 2023). A Gentle Introduction to Supervised Learning. Weights and Biases. https://wandb.ai/mostafaibrahim17/ml-

articles/reports/A-Gentle-Introduction-to-Supervised-Learning--VmlldzozMTcxOTk0

31. Publication, I. J. R. A. S. E. T. (2020). Regression Techniques in Machine Learning &Applications: A Review. International Journal For Research In Applied Science & Engineering Technology. https://doi.org/10.22214/ijraset.2020.32019
32. Onder, E, et al. (2022). Application of Supervised Machine Learning Regression Algorithms to Prediction of Dielectric Properties of PPy/Kufeki Stone Composites for Energy Implementations. *Journal of Science. 35*(1): 235-254. 10.35378/gujs.810948
33. Naeem, Samreen & Ali, Aqib & Anam, Sania & Ahmed, Munawar. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. IJCDS Journal. 13. 911-921. 10.12785/ijcds/130172.
34. M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep clustering for unsupervised learning of visual features," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 132– 149. https://doi.org/10.48550/arXiv.1807.05520)
35. H. Ma, "Design of chinese linguistics teaching system based on k-means clustering algorithm," in The International Conference on Cyber Security Intelligence and Analytics. Springer, 2022, pp. 424–431
36. What is unsupervised learning. s.f. https://www.ibm.com/topics/unsupervised-learning
37. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521 , 436-444 (2015). https://doi.org/10.1038/nature14539
38. Senthilkumar, M. (2010). Use of artificial neural networks (ANNs) in colour measurement. Colour Measurement, 125–146. doi:10.1533/9780857090195.1.125

39. MathWorks. Convolutional Neural Networks. What is a convolutional neural network?. https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html
40. X. Yang, T. Zhang, C. Xu, S. Yan, MS Hossain y A. Ghoneim, "Deep Relative Attributes", en IEEE Transactions on Multimedia , vol. 18, núm. 9, págs. 1832-1842, septiembre de 2016, doi: 10.1109/TMM.2016.2582379
41. Arevalo J, González FA, Ramos-Pollán R, Oliveira JL, Lopez MAG (2016) Learning representations for mammography mass lesion classification with convolutional neural networks. Comput Methods Programs Biomed 127:248-257. https://doi.org/10.1016/j.cmpb.2015.12.014.
42. Singh, E, Kuzhagaliyeva, N, Sarathy, S, etal. (2022). Artificial Intelligence and Data Driven Optimization of Internal Combustion Engines. 13-237. https://doi.org/10.1016/B978-0-323-88457-0.00005-9)
43. De Rosa, G, Papa, J. (2022). Optimum-Path Forest. Learning to weigh similarity measures with Siamese networks: a case study on optimum-pat forest. En Falcao, A (Ed), Optimum-Path Forest(pp.155-173). https://doi.org/10.1016/B978-0-12-822688-9.00015-3
44. Vaswani, A , Shazeer, N, Parmar, N, etal.(2017). Attention is all you need. NeurIPS proceedings ( 2017 ) , págs . 5998-6008 https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
45. Tian,Y, Yuxin,W, et al.(2022). A survey of Transformers. AI Open3, 111-132. https://doi.org/10.1016/j.aiopen.2022.10.001
46. Artemoppermann. s.f. https://artemoppermann.com/optimization-in-deep-learning-adagrad-rmsprop-adam/

47. NT, A. (2021). Inception V3 model architecture. Obtained from https://iq.opengenus.org/inception-v3-model-architecture/
48. Google cloude . Advanced guide to Inception v3. Retrieved from: https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=es-419
49. D. Ragab, M. Sharkas, S. Marshall and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," PeerJ, vol. 7, p. 19, 2019. (19)
50. A. Yurttakal, H. Erbay, T. ˙Ikizceli and S. Karacavus, "Detection of breast cancer via deep convolution neural," Multimedia Tools and Applications, vol. 79, pp. 15555-15573, 2019. 10.1007/s11042-019-7479-6
51. S. Hadush, Y. Girmay, A. Sinamo y G. Hagos, «BREAST CANCER DETECTION USING CONVOLUTIONAL NEURAL NETWORKS,» ICLR, p. 8, 2020.
52. Dutta, S, Kumar, J, Hoon, T, et al. (2020). Breast cancer prediction using stacked GRU-LSTM-BRNN. Scindo.25(2).https://doi.org/10.2478/acss-2020-0018
53. Ekici, S., & Jawzal, H. (2020). Breast cancer diagnosis using thermography and convolutional neural networks. Medical Hypotheses, 137, 109542. doi:10.1016/j.mehy.2019.109542
54. Das HS, Das A, Neog A, Mallik S, Bora K and Zhao Z (2023), Breast cancer detection: shallow convolutional neural network against deep convolutional neural networks based approach. Front. Genet. 13:1097207.doi: 10.3389/fgene.2022.1097207
55. Wang, Y., Choi, E. J., Choi, Y., Zhang, H., Jin, G. Y., & Ko, S.-B. (2020). Breast Cancer Classification in Automated Breast Ultrasound Using Multiview Convolutional Neural Network with Transfer Learning. Ultrasound in Medicine & Biology. doi:10.1016/j.ultrasmedbio.2020.01.001

56. Jalloul, R.; Chethan, H.K.; Alkhatib, R. A Review of Machine Learning Techniques for the Classification and Detection of Breast Cancer from Medical Images. Diagnostics 2023, 13, 2460. https://doi.org/10.3390/diagnostics13142460
57. Sickles, EA et al. ACR BI-RADS® Mammography (American College of Radiology, 2013), fifth ed.
58. Dataset: Pham, HH, Nguyen Trung, H. and Nguyen, HQ (2022). VinDr-Mammo: a large-scale reference dataset for computer-aided detection and diagnosis in full-field digital mammography (version 1.0.0). FisioNet . https://doi.org/10.13026/br2v-7517 .
59. Introducción a Transfer Learning. (s.f). MathWorks. https://la.mathworks.com/discovery/transfer-learning.html
60. Shivam, P. (2021). Results from machine learning models. https://github.com/MicrosoftDocs/Dynamics-365-Operations.eses/blob/live/articles/finance/finance-insights/confusion-matrix.md.
61. Selvaraj,N. (November 9, 2022). Confusion Matrix, Precision, and Recall Explained. *KD nuggets.* https://www.kdnuggets.com/2022/11/confusion-matrix-precision-recall-explained.html
62. Hernández, Adriana & Rosales, Gustavo & Santiago, Héctor & Lee, Maribel. (2022). Performance metrics for evaluating machine learning in oil image classification using convolutional neural networks. Latin Science Multidisciplinary Scientific Magazine. 6. 4624-4637. 10.37811/cl_rcm.v6i5.3420
63. Borja-Robalino, Ricardo & Monleon-Getino, Antonio & Benedé, Jose. (2020). Standardization of Performance Metrics for Machine and Deep Learning Classifiers

64. Barrios, J. (July 26, 2019). The confusion matrix and its metrics. Health big data. https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/

65. Sekaran, Karthik & Ramalingam, Srinivasa Perumal & P.V.S.S.R., Chandra Mouli. (2018). Breast Cancer Classification Using Deep Neural Networks. 10.1007/978-981-10-6680-1_12.

66. Bokade A, Shah A (2023) Breast Cancer Diagnosis in Mammography Images Using Deep Convolutional Neural Network-Based Transfer and Scratch Learning Approach. Indian Journal of Science and Technology 16(18): 1385-1394. https://doi.org/ 10.17485/IJST/v16i18.39

67. Habeeb, Z. Q., Vuksanovic, B., & Al-Zaydi, I. Q. (2023). Breast cancer detection using image processing and machine learning. Journal of Image and Graphics(United Kingdom), 11(1), 1-8. https://doi.org/10.18178/joig.11.1.1-8

68. Victor Guevara-Ponce, Ofelia Roque-Paredes, Carlos Zerga-Morales, Andrea Flores-Huerta, Mario Aymerich-Lau and Orlando Iparraguirre-Villanueva, "Detection of Breast Cancer using Convolutional Neural Networks with Learning Transfer Mechanisms" International Journal of Advanced Computer Science and Applications(IJACSA), 14(6), 2023. http://dx.doi.org/10.14569/IJACSA.2023.0140661

# 7. Anexos

## 1. Prediction tests

Prognostic prediction test in cancer-negative prognoses


Predictive test in cancer-positive prognoses

## 2. Most important parts of the code

- Preparing data

```
[ ]  !cp /content/drive/MyDrive/Eli-tesis/breast_cancerp2.zip.
     !mkdir /content/dataset
     !unzip jpg_dataset_by_class_balanced.zip -d ./dataset
     !mv ./dataset/jpg_dataset_by_class_balanced/test ./dataset
```

```
data_dir = "/content/drive/MyDrive/Eli-tesis/breast_cancerp2.zip"
# Models to choose from [resnet, alexnet, vgg, squeezenet,
densenet, inception]
model_name = "inception"

# Number of classes in the dataset
num_classes = 2

# Batch size for training (change depending on how much memory you
have)
batch_size = 256

# Number of epochs to train for
num_epochs = 25

# Flag for feature extracting. When False, we finetune the whole
model,
#   when True we only update the reshaped layer params
#feature_extract = True
feature_extract = False
```

- Data argumentation

```
# Data augmentation and normalization for training
# Just normalization for validation


data_transforms = {
    'train': transforms.Compose([
        transforms.Resize((299,299)),
        transforms.RandomRotation(5),
        transforms.RandomHorizontalFlip(0.5),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
    ]),
    'val': transforms.Compose([
        transforms.Resize((299,299)),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
    ]),
}

data_dir = '/content/drive/MyDrive/Eli-tesis/breast_cancerp2'
image_datasets = {x: datasets.ImageFolder(os.path.join(data_dir, x),
                                          data_transforms[x])
                  for x in ['train', 'val']}
dataloders = {x: torch.utils.data.DataLoader(image_datasets[x], batch_size=64,
```

```
                                             shuffle=True, num_workers=4)
              for x in ['train', 'val']}
dataset_sizes = {x: len(image_datasets[x]) for x in ['train', 'val']}
class_names = image_datasets['train'].classes

#use_gpu = torch.cuda.is_available()
# Detect if we have a GPU available
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
```

- Optimizers

```
# Observe that all parameters are being optimized
#optimizer_ft = optim.SGD(params_to_update, lr=0.001, momentum=0.9)

#optimizer_ft = optim.SGD(model_ft.parameters(), lr=0.001, momentum=0.0)
optimizer_ft = optim.SGD(model_ft.parameters(), lr=0.001, momentum=0.9)
#optimizer_ft = optim.Adagrad(model_ft.parameters(), lr=0.001)
#optimizer_ft = optim.Adam(model_ft.parameters(), lr=0.001)
#optimizer_ft = optim.RMSprop(model_ft.parameters(), lr=0.001)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=int((1/3)*num_epochs), gamma=0.1)

# Setup the loss fxn
criterion = nn.CrossEntropyLoss()
```

- Making predictions/ load the data

```
import torch
from torchvision import transforms, datasets
from torch.utils.data import DataLoader
from torch import nn

import matplotlib.pyplot as plt
%matplotlib inline
import numpy as np
from sklearn.metrics import classification_report

###
def imshow(inp, title=None):
    """Imshow for Tensor."""
    inp = inp.numpy().transpose((1, 2, 0))
    mean = np.array([0.485, 0.456, 0.406])
    std = np.array([0.229, 0.224, 0.225])
    inp = std * inp + mean
    inp = np.clip(inp, 0, 1)
    plt.imshow(inp)
    if title is not None:
        plt.title(title)
```

```
# Applying Transforms to the Data

#Original
image_transforms = {
    'test': transforms.Compose([
        #transforms.RandomSizedCrop(224),
        transforms.Resize(size=(299, 299)),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225]),
    ])
```

```
test_directory="//content/drive/MyDrive/Eli-tesis/breastcancertest"

# Batch size
bs = 256

# Number of classes
num_classes = 2

# Load Data from folders
data = {
    'test': datasets.ImageFolder(root=test_directory, transform=image_transforms['test']),
}

class_names = data['test'].classes
transform=image_transforms['test']


# Get a mapping of the indices to the class names, in order to see the output classes of the test images.
idx_to_class = {v: k for k, v in data['test'].class_to_idx.items()}
print('Classes: ',idx_to_class)
```

```
# Size of Data, to be used for calculating Average Loss and Accuracy
test_data_size = len(data['test'])

# Create iterators for the Data loaded using DataLoader module
test_data_loader = DataLoader(data['test'], batch_size=bs, shuffle=False)

# Print the test set data sizes
print('Number of images: ',test_data_size)

def computeTestSetAccuracy(model, loss_criterion, data_loader, data_size):
    '''
    Function to compute the accuracy on the test set
    Parameters
        :param model: Model to test
        :param loss_criterion: Loss Criterion to minimize
    '''

    test_acc = 0.0
    test_loss = 0.0

    model.eval()
    images_so_far = 0
    fig = plt.figure()
```