People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Ibn Khadoun university – Tiaret

Mathematics and Computer Science faculty

Computer Science department

Theme

Automatic Breast Cancer CAD

For obtaining Master's degree

Specialty: Software Engineering

By: Lebani Ali Zakaria and Hatem Tayeb

Publicly supported on 26/ 06/2022 in Tiaret in front of the jury composed of :

Mr. CHENINE Abdelkader	MAA	Tiaret University	President
Mr. MERATI Medjeded	MCA	Tiaret University	Supervisor
Ms. BENATHMANE Lalia	MAA	Tiaret University	Examiner

Academic year 2021-2022

Acknowledgements

Dedication

To the owner of a fragrant biography and enlightened thought, for he had the first credit in attaining higher education (my beloved father), May God have mercy on him and grant him vast paradise.

To the one who set me on the path of life, made me calm, and took care of me until I became old (my mother) .

To my brother Walid and my sisters Naima and sara ,my close friend yacine , who have had a great impact in many obstacles and difficulties and who gave me all the support to keep working hard .

To all my dear professors, who did not hesitate to extend a helping hand to me,also to my classmates and my partner Ali Zakaria for helping me to succeed Throughout the course.

<div align="right">

Hatem Tayeb.

</div>

Dedication

I would like to dedicate this modest work to:

My very dear parents who supported me a lot, encouraged me and who made me who I am today;

My dear sisters: Sara, Amel and Hadjer ;

My dear brothers: Mohamed and Anas;

My best friends and my colleagues; mustapha, Lahcen, Hakim, Azhar, Kamelia, Bouchra and sara,....;

M.Sara; for the moral support throughout my course; To all the 2022 promotion; May Allah keep them...

<div align="right">

Lebani Ali Zakaria

</div>

Abstract

Breast cancer is still one of the most frequent malignancies in the world, having claimed the lives of millions of people. Clinical specialists identify suspicious lumps to help diagnose breast problems. Due to the large number of mammograms and the time and effort required to study each view of a mammogram, this activity poses a daily challenge for radiologists. The failure of mammography to detect abnormalities is increasing due to poor image quality, eye tiredness, or radiologists' negligence.

Breast cancer identification at an early stage is crucial for lowering the mortality rate among women. That's why, in order to construct an autonomous system that aids radiologists in diagnosing breast cancer, we employed YOLO (a real-time object identification algorithm) and VGG16 to detect and pinpoint malignant lesions on mammograms.

Key words: Breast cancer, Mammograms, Masses, Object detection, YOLO, VGG16, classification, Malignant, Diagnose.

Résumé

Le cancer du sein est toujours l'une des tumeurs malignes les plus fréquentes dans le monde, ayant coûté la vie à des millions de personnes. Les spécialistes cliniques identifient les masses suspectes pour aider à diagnostiquer les problèmes mammaires. En raison du grand nombre de mammographies et du temps et des efforts nécessaires pour étudier chaque vue d'une mammographie, cette activité représente un défi quotidien pour les radiologues. L'échec de la mammographie à détecter des anomalies est en augmentation en raison de la mauvaise qualité des images, de la fatigue oculaire ou de la négligence des radiologues.

L'identification du cancer du sein à un stade précoce est cruciale pour réduire le taux de mortalité chez les femmes.

C'est pourquoi, afin de construire un système autonome qui aide les radiologues à diagnostiquer le cancer du sein, nous avons utilisé YOLO (un algorithme d'identification d'objets en temps réel) et VGG16 pour détecter et localiser les lésions malignes sur les mammographies.

Mot clé : Cancer du sein, Mammographies, Masses, Détection d'objets, YOLO, VGG16, Classification, Malin, Diagnostiquer.

الملخص

لا يزال سرطان الثدي من أكثر الأمراض الخبيثة انتشارًا في العالم ، حيث أودى بحياة الملايين من الناس. يحدد المتخصصون السريريون الكتل المشبوهة للمساعدة في تشخيص مشاكل الثدي. نظرًا للعدد الكبير من صور الثدي بالأشعة السينية والوقت والجهد اللازمين لدراسة كل صورة من صور الماموجرام ، فإن هذا النشاط يمثل تحديًا يوميًا لأخصائيي الأشعة. يتزايد فشل التصوير الشعاعي للثدي في اكتشاف العيوب بسبب سوء جودة الصورة أو إجهاد العين أو إهمال أطباء الأشعة.

يعد التعرف على سرطان الثدي في مرحلة مبكرة أمرًا حاسمًا لخفض معدل الوفيات بين النساء.

لهذا السبب ، من أجل بناء نظام مستقل يساعد أخصائيي الأشعة في تشخيص سرطان الثدي ، استخدمنا YOLO (خوارزمية تحديد الكائن في الوقت الفعلي) و VGG16 لاكتشاف وتحديد الآفات الخبيثة في تصوير الثدي بالأشعة السينية.

الكلمات المفتاحية: سرطان الثدي ، تصوير الثدي بالأشعة السينية ، الكتل ، كشف الأجسام ، التصنيف ، الخبيث ، التشخيص. VGG16 ، YOLO ،

Contents

List of Figures

3

List of Tables

Abbreviations list

ANN   Artificial neural network.

BRCA   BReast CAncer.

CAD   Computer Aided Diagnosis.

CCD   Charge Coupled Device.

CNN   Convolutional Neural Network.

CONV   Couche de convolution d'un CNN.

CPU   Processing Unit.

CUDA   Compute Unified Device Architecture.

DDSM   Digital Database for Screening Mammography.

GPU   Graphics Processing Unit.

IoU   Intersection over Union.

mAP   mean Average Precision.

OPENCV Open Computer Vision.

IARC   International Agency for Research on Cancer.

ILSVRC   ImageNet Large Scale Visual Recognition Challenge.

PAN   Path Aggregation Network.

ReLU   Rectified Linear Unit.

RNN   Recurrent Neural Network.

SPP   Spatial Pyramid Pooling.

VGG   Visual Geometry Group.

YOLO   You Only Look Once.

GENERAL INTRODUCTION

Introduction général

Background

Cancer is a condition in which cells multiply uncontrollably and invade other tissues. It is also one among the most common and well-studied causes of death in affluent countries. Cancerous (malignant) or non-cancerous (benign) tumor cells exist. Cancer can strike people of any age. As people get older, they are more prone to get most types of cancer. In 2018 China reported 2.9 million deaths [36].

Breast cancer is defined as a malignant tumor that develops from cells in the breast. Tumors can infect nearby tissues or spread to other parts of the body [37]. Only around one-tenth of all breast cancers (10%) are connected to known faulty genes passed down from parents (inherited). Breast cancer symptoms include lumps in the breast, dimpling, and nipple retraction. The survival percentage of breast cancer is largely determined by early detection procedures. Mammography, magnetic resonance imaging, and biopsy are some of the most essential tools for the identification of breast cancer [1]. Mammography is a sort of specialized medical imaging that scans the breast with a low-dose x-ray equipment and is considered the most accurate.

Aside from finding tumors in the breast, medical professionals must appropriately diagnose breast cancer and provide the necessary treatment alternatives. In the realm of medical imaging, computer-aided diagnosis (CAD) is a computer-based system that assists clinicians in making quick choices [2]. Medical imaging involves the evaluation and analysis of abnormalities in pictures by medical practitioners and doctors in a short amount of time.

CAD is a system that combines principles like artificial intelligence (AI), computer vision, and medical image processing.

Machine learning is a technique that can be used to identify, detect and classify abnormality in an image of breast. Standard machine learning techniques require the extraction of features such as shape and size of tumors from images, to perform a detection and classification task. However, the feature extraction task requires experience and expertise.

Deep learning algorithms have been developed to perform detection and classification tasks for pattern recognition, computer vision and image recognition [3]. In particular, a deep learning algorithm known as a Convolutional Neural Network (CNN) is popularly known for exceptional performance in image recognition tasks. In this essay, two CNN models will be utilized for the detection and prediction of tumors located in the breast.

Aims and Objective

The main goal of this memory is the implementation of CAD systems for detection of abnormal signs at an earliest that a human professional fails to find. In mammography, identification of small lumps in dense tissue, finding architectural distortion and detecting location of mass and prediction of mass type as benign or malignant by its shape, size, etc.

Memory organization

This thesis is a joint of the medical and computer fields and a grouping of several research axes including medical imaging, artificial vision. It is organized into two main parts:

- **The First part**

 This part is devoted on the one hand to the medical context of the problem treated and on the other hand to machine learning and its ramifications. Includes two chapters.

 - Medical Context.
 - Deep learning for image processing.

- **The second part**

 The second part of this thesis focuses on our contributions, through the implementation of a model and a detection assistance system, for the prevention of breast cancer by screening mammography. We organize it into a chapter.

 - Realization,Test, Results.

First part

MEDICAL CONTEXT

Chapter I

MEDICAL CONTEXT

I.1   Introduction

In the digital age, information technology has been integrated into medicine and is currently involved in all phases of the work of the medical profession, such as the management of medical records, surgical operations, image collection and processing, etc. Indeed, daily, the medical community creates views, exchanges and archives large quantities of images, extracts measurements and inserts them into reports.

The medical image is the only communication tool par excellence. It has become, not only essential for the diagnosis of a large number of pathologies, but also for the control of an action or a therapy, by a very precise monitoring of the evolution of a disease. In its raw state, however, an image is hardly meaningful, you have to know how to communicate with it and thus interpret it. In recent years, many teams have contributed by introducing diagnostic aid systems to provide the doctor with high-level symbolic information on the content of the image, especially when it comes to tumoral conditions such as cancer Breast.

In this first chapter, we will present the medical approach to breast imaging to situate the context in which our research work will be used. Not being exhaustive, this presentation is given only for the sole purpose of facilitating the understanding of the problem and thus acquiring a certain knowledge, necessary to carry out our study to deduce the approaches of analysis, detection and classification.

I.2   The breast, organ of the human body

"The medical Larousse defines the breast as being a name given to each mammary gland. Two in number in humans, the breasts are glandular[1]

organs whose function is milk secretion in women. Hemispherical in shape, they have a more or less turgid nipple[2] located in the center of a pigmented area called the areola.[4]

According to the sex and the phase of the genital life, the mammary gland is of very variable morphology. It develops in the subcutaneous cell tissue of the anterolateral wall of the chest, between the third and fifth ribs, above the pectoralis major muscle[3].

The skin encompassing this gland is smooth, supple, fine, mobile and slides easily over the gland. Beneath this cutaneous level there is a more or less developed adipose panniculus which gives it its shape and volume.

The mammary gland, immersed in this fatty tissue, is made up of about twenty lobules which are activated during the lactation period and produce a milky secretion, discharged by the milk ducts at the level of the nipple. [figure 1.1] further summarizes these different breast components.

---

[1]Gland: Organ whose function is to produce a secretion.
[2]Nipple or end of the udder: milk secretion organ in mammals.

[3]Important muscle contributing to the movements of the arm.
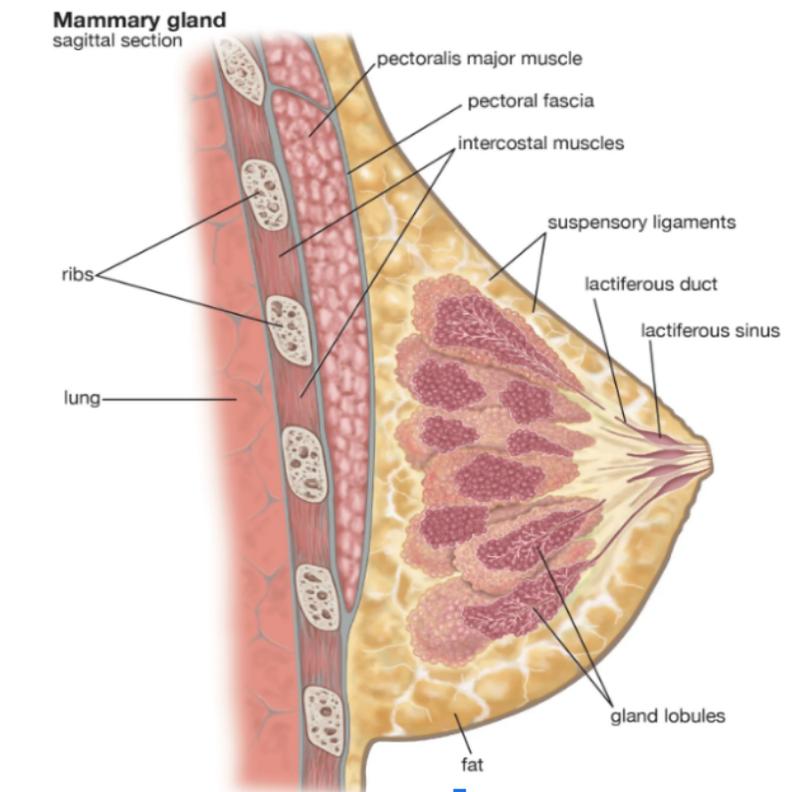
Figure I.1: Internal structure of the breast (mammary gland.)

This organ must be monitored regularly with the sole aim of pre-venting any anomaly, first by the woman herself, approximately every six months (self-examination of the breasts) in order to detect benign lesions, then by the gynecologist.

"Particularly frequent, these anomalies relate to the volume of the gland. One of the malformations present in women is Breast Ptosis or descent of the breast. Other abnormalities include mammary hypertrophy (breasts that are too small) as well as abnormalities in the nipple region." [Domart and Bouneuf, 99].

What becomes more serious for this organ is the infectious state and several inflammatory states of the breasts can be observed especially during breastfeeding. The breast is also an organ like the others which can present a risk of cancer, which continues to represent the first cause of death in women.
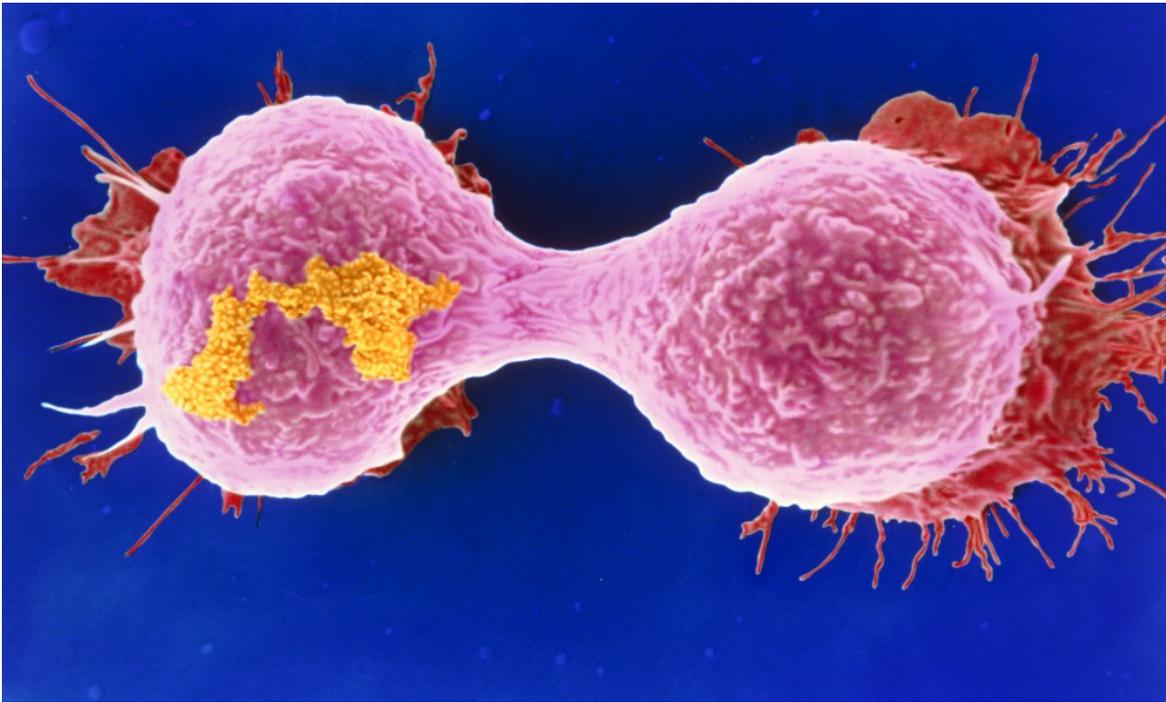
Figure I.2: Image of Cancer cell division.[Lund University Research Magazine]

I.3   Breast cancer

Our body is made up of a multitude of cells. Each individual contains one billion per gram of weight. They are all different from each other but nevertheless have one thing in common; they come from the same egg. These cells grow continuously and normally. However, there may be abnormal cells that can grow out of control. Others manage to spend a long time before beginning their anarchic development. Nevertheless, they can remain dormant throughout our lives without ever manifesting themselves.

This anarchic development can lead to a cancerous cell or a tumor. From this moment, it multiplies in an uncontrolled manner and ends up forming a cancerous tumor whose cells will destroy neighboring cells and gradually invade neighboring tissues with the help of metastases [figure 1.2].

Some people can have tumors without it being cancer. Often it is only

a benign tumor which is a localized and slow growth. However, the corresponding tissue remains normal, i.e. the cells retain their characteristics: this is only an increase in the total number of cells.

"The term breast cancer refers only to malignant, potentially aggressive tumors of the breast that frequently develop either in the glandular lobules (lobular carcinomas) or in the milk ducts (ductal carcinomas)."
[5]

I.3.1   Some numbers

"Breast cancer, almost exclusively female, is a major public health problem throughout the world. According to the press release, dated September 12, 2018, from the IARC, the incidence rates of breast cancer far exceed those of other cancers, both in developed and developing countries. Development [Figure 1.3], followed by colorectal cancer in developed countries and cervical cancer in developing countries." .
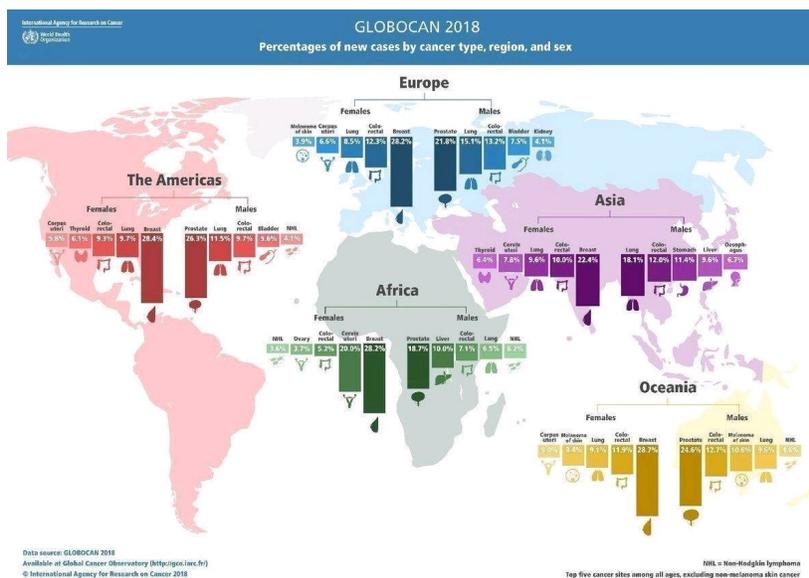
Figure I.3: Percentages of new cases by cancer type, region and sex [CLOBOCAN 2018].

According to GLOBOCAN [4], breast cancer accounts for 24.2%, or

---

[4]GLOBOCAN is an online database providing estimates of incidence and mortality in 185 countries for 36 types

about

one in four, of newly diagnosed cancer cases among women worldwide in 2018, and it is the most common in 154 of the 185 countries covered by this database. Given,it is also the leading cause of cancer death in women (15.0%), followed by lung cancer (13.8%) and colorectal cancer (9.5%), which are also the third and second most common types of cancer among them, respectively [Figure 1.4].
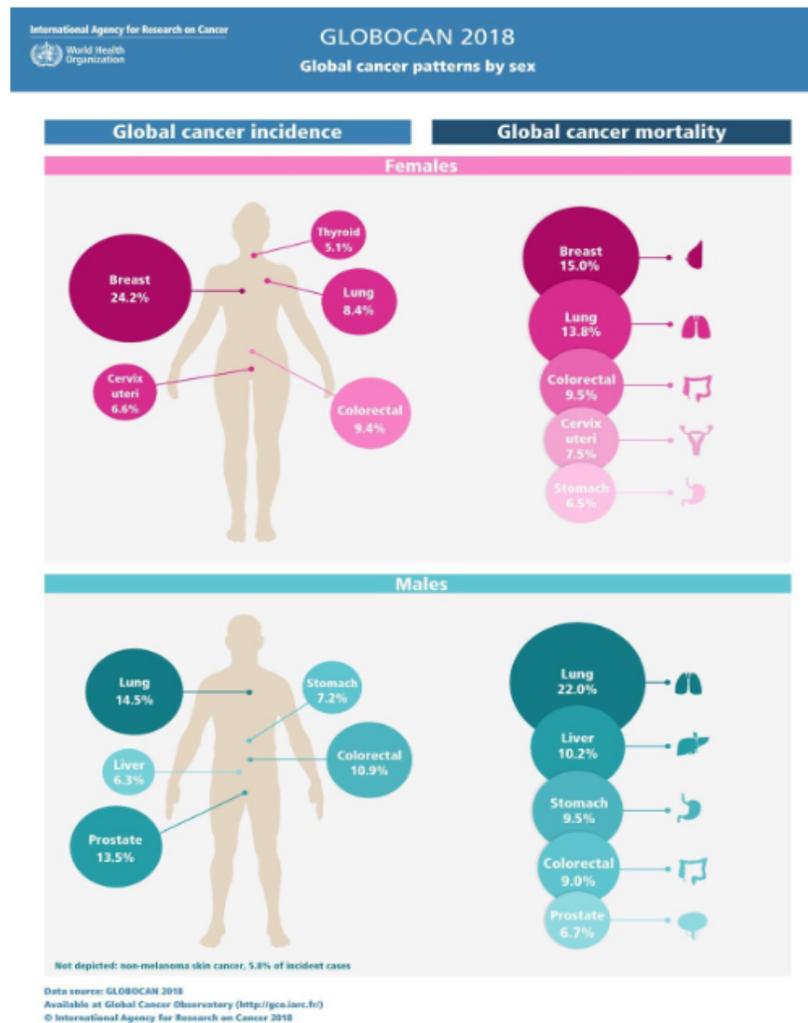
Figure I.4: Incidence and mortality rates by cancer type [GLOBOCAN 2018].

of cancer and for all cancer sites combined. The data is part of the Global Cancer Observatory of the International Agency for Research on Cancer (IARC) and is available online at http://gco.iarc.fr/today/home

" Algeria, where breast cancer is a cancer of young women, is no exception to this scourge. 10910 new cases were reported in 2014 with a rate of 59% of all tumors in women."[6]

Even if it occupies an intermediate position between the industrialized countries and the Arab countries, its risk factors must be highlighted and primary prevention must be established. To this end, the Ministry of Health of Algeria has put in place the anti-cancer plan (2015-2019) for breast cancer, calling for more efforts by different actors to promote the prevention and early detection of breast cancer sickness.

I.3.2   Primary prevention of breast cancer

"Several investigations have been undertaken to try to discover the risk factors likely to play a role in the genesis of breast cancer. Among these factors are those of age, heredity, history of certain types of benign diseases and exposures of the chest to high doses of irradiation. Obesity in postmenopausal women, early onset of menstruation and late menopause are other risk factors, the effect of which is less marked."[7]

"There is no primary prevention, despite the efforts made to highlight these risks factors.Many clinical studies of prevention with tamoxifen[5]have been published, however this drug is recommended for women at high risk to prevent breast cancer but it is not without serious potential consequences."[8]

"Other studies are needed for genetic tests to detect BRCA[6] 1 or 2 genes to be implemented." [9]

I.3.3   Breast cancer screening

Currently, primary prevention of breast cancer is not possible. Given the mortality rate of this type of cancer, screening remains the only practical way to reduce it in the long term. According to the World Health Organization, "screening consists of detecting at an early sub-

---

[5]Drug recommended in the treatment of breast cancer in the early or advanced phase in pre- and post-menopausal women.

[6]Acronym for BReast CAncer which is a tumor suppressor gene.

clinical (asymptomatic) stage a disease to enable early diagnosis and therapy with a view to improving the prognosis of the disease and the health status of individuals and the population."

"Detection of breast cancer therefore amounts to carrying out control examinations which make it possible to detect abnormalities without even the presence of detectable signs. As such, mass screening campaigns are organized with the aim of detecting malignant lesions before they are known by symptoms." [10]

Another important issue in these campaigns is the detection of tumors prior to their metastatic[7] dissemination i.e. at a stage where the size of the tumor is still very small. Indeed, the earlier the disease is identified, the more effective the treatment is. Early detection also makes it possible to reduce the use of radical surgery or radiotherapy and thus to avoid undesirable side effects, which improves the quality of life of patients.

I.4   Breast Cancer Imaging

Several medical imaging tools are used to examine the mammary gland such as ultrasound, Magnetic Resonance Imaging (MRI), mammography and others. Each of them has limitations and advantages, but mammography currently remains the only effective method of screening for breast cancer in asymptomatic women and remains the one with the greatest sensitivity for detecting tumors at an early stage.

I.4.1   Digital mammography

Mammography, also called mastography, is an X-ray imaging technique that provides an image of the internal structure of the mammary gland. Mammography is performed in two circumstances; in the context of screening or diagnosis, hence the appearance of the types of screen-

---

[7]Metastases: Cancer cells that spread over a distance.

ing mammograms[8], diagnostic mammograms[9]and preventive mammograms[10].

In the all-digital era, mammography is no exception to the rule, because with the generalization of image processing techniques carried out by computers and the appearance of computer-assisted diagnostic tools, algorithms for mammography analysis appeared. There was therefore the need to have digital support to be able to apply these methods.

The interest of such digitization is multiple. Indeed, with a digital mammography [figure 1.5] there is a reduction in the dose of radiation absorbed, due to the low noise and the high sensitivity of the CCD (Charge Coupled Device) system[11]. In other words, the total exposure time is determined by the computer when the signal-to-noise ratio has reached a predefined level and not by the optical density of conventional X-ray films.

---

[8]Screening mammography seeks to detect cancer even before any sign or symptom suggests its presence. It is performed in women between the ages of 50 and 69 by doctor's recommendation.

[9]Diagnostic mammography is required by the doctor who suspects breast cancer. This suspicion may arise from the discovery of a breast lump following a self-examination.

[10]Preventive mammography can be requested by the prescribing doctor for any woman outside the age of 50 to 60 without any particular symptoms, in order to ensure that there is no breast cancer call point. .

[11]A charge-coupled device (CCD) is a light-sensitive integrated circuit that captures images by converting photons to electrons. A CCD sensor breaks the image elements into pixels. Each pixel is converted into an electrical charge whose intensity is related to the intensity of light captured by that pixel.

Figure I.5: Digital mammography.

A second interest concerns near real-time manipulation, making the examination more rapid. Thus, it is easier to make a good acquisition of images with better contrast and therefore reduce the number of shots. This is difficult with emulsion films, where it is not possible to process images a posteriori to improve the quality of the image or to help detect a tumor or a lesion. That said, with the arrival of a new all-digital mammography technique, cancer detection should be possible at much earlier stages than with conventional mammography devices.

I.4.2 Different Mammography Scanning Processes

Cites two ways to obtain digital mammograms: either by digitization of traditional mammographic film, or by acquisition of mammograms which are directly in digital form.[11]

In the first case, the already formed image is then digitized using a camera, camera or scanner. However, a scanned radiograph will never contain more information than the original print, or even less due to loss phenomena on the scanner (sampling, noise, etc.) and above all retain all the defects of the original image (noise, dynamics, development artifacts,

etc.).

The second technique has been particularly topical for some time. Digitization is carried out at the time of image acquisition, and there is no longer any prior printing of the mammogram: the mammography device has sensors which more or less directly transform the X photons into digital information.

I.4.3 Normal appearance of a mammogram

"There is no normal breast, but an infinite number of variants of the norm". Indeed, the appearance of the normal breast on a mammogram can be quite variable from one woman to another. The breasts of younger women normally tend to be dense since they usually have more fibroglandular tissue. However, a minority of young women may have relatively fatty breasts and that dense normal tissue may predominate on the mammograms of some older women, resulting in a wide variation in normal patterns.[12]

"By analyzing the images in figure 1.6 we notice that it is extremely difficult to define normality on mammograms. The appearance of the mammary gland is indeed extremely variable depending on the patient, her age, the period during which the mammography is performed in relation to the menstrual cycle and according to the periods of pregnancy and breastfeeding."[13]

Indeed, the radiological image of the breast tends to change during life. The pubertal breast has a very dense and homogeneous appearance and should never be X-rayed. The adult female breast has a distribution of connective and adipose tissue that varies from woman to woman.

I.4.4 Reading mammography images

Due to the low contrast of many cancerous lesions and the glaring similarity of some opacities with breast tissue, the interpretation of the mammogram image becomes a very difficult task that requires great skill and concentration. Different inspection methods are used by radiologists
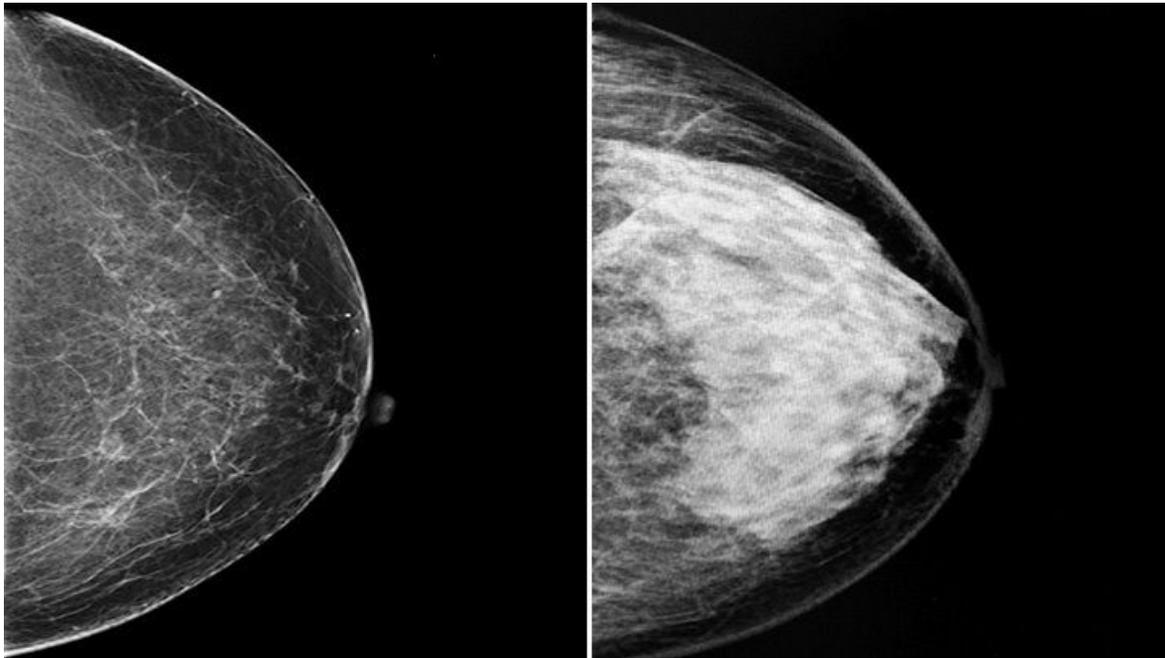
Figure I.6: Different types of breast composition. Breast with fibro glandular fatty tissue (left) and extremely dense breasts (right).

to interpret the mammography image and are as follows:

- The first way is based on the independent analysis of each mammogram available and this according to the abnormal structure to be localized, namely calcium structures (microcalcifications and/or foci of microcalcifications), structures having a nodular appearance (opacities and microcalcifications) and those presenting linear characteristics (stellar images, architectural disorganizations).

- The second way concerns the comparison of several mammograms available and is more dedicated to the detection of masses. We distinguish between the comparison of mammograms with the same incidence of the right and left breasts of the same patient (comparison of pairs of bilateral mammograms) and the comparison of mammograms with the same incidence of the same breast acquired on two separate dates (comparison of temporal mammograms).

"It should be noted that, like any radiological examination, the inter-pretation of the mammogram strongly depends on the rate of absorption of the X-rays by all the components of the breast. However, there are several constraints against this. We cite those related to the material, the clinical condition of the patient, technical artifacts (nose, ear, hands, ribs, dust, breast implants, etc.), breast compression, etc." [14]

- Cases of false positives, which have negative consequences, are generated when benign areas are identified as cancerous areas;

- More serious, false-negative cases endanger the patient's life and occur when an abnormality exists and is not detected by the radiologist;

Other reading error classifications have been mentioned in [Richard, 00] by some researchers and relate to:

- Research or exploration errors (these errors relate decision-making errors (localized but poorly classified anomaly).to examples in which the radiologist does not explore an area containing a significant abnormality);

- errors in identification (they occur when the sectors concerned are being sought, but the anomalies are not identified).

I.5   Computer Aided Diagnosis

By studying the mammary gland and its different pathologies in the previous sections, we note, on the one hand, the difficulty of the detection stage and, on the other hand, the complexity of the diagnostic decision-making stage (malignant aspect or benign). In addition, the prevention of breast cancer requires the processing of a large volume of mammographic images as well as the effort and intervention of different radiologists, in order to help each other in this decision-making. For these reasons, several studies have focused on automating mammogram reading and decision-making.

The automatic processing of images, at a high level, which is the responsibility of Computer Aided Diagnosis (CAD) systems, is today one of the most emerging areas of research in medical imaging. The role of the CAD system is to provide the doctor with symbolic information about the content of the image. It may happen that the radiologist does not see a lesion due to lack of concentration or fatigue due to reading several images (more than 150 images), or the shape of the lesion may be unusual on the screen.

The interest of an automated diagnosis in the systematic screening of breast cancer in asymptomatic women has been demonstrated. The main diagnostic aid software in mammography makes it possible to increase the sensitivity of the reading by increasing the detection of potentially significant lesions and the specificity by reducing the interpretations of false positives and thus the number of punctures of benign lesions.

I.5.1 Brief history of CAD in mammography

Research in the development of methodologies and instruments for diagnostic or therapeutic applications are among the main responsibilities of physicists in the medical field. Indeed, since the discovery of X-rays by Wilhelm Conrad Roentgen in 1985, physicists have continually contributed to the improvement of clinical decisions in medical diagnosis, through the study of imaging systems, the evaluation of the quality and the perception of images, as well as the development of methods for analyzing them.

"R2 Technology's ImageChecker M1000 is the first commercial CAD system that was approved in 1998 by the US Food and Drug Administration (FDA). It was developed at the University of Chicago, on digitized screening mammograms and subjected to a processing unit with an integrated neural network. The output of the system is expressed using annotations indicating suspicious locations of microcalcifications and masses, on a thermal paper carrier or monitor."[15]

The growth of work on automatic diagnostic tools in recent years has been considerable, ranging from the digitization of films and limited com-

putation time to the current state, where the approaches are developed in a very rigorous way on major clinical projects and very high quality digital image databases. In the context of CAD systems, most work has focused on detecting foci of microcalcifications or lesions in the form of masses or even architectural distortions.

As such, several international conferences and workshops specific to this issue are organized to evaluate this research. The series of conferences was launched in San José (Canada) in 1993 and every two years researchers from all over the world continue to organize it. The most recent is the IWBI 2018 (International Workshop on Breast Imaging 2018) which brought together a multidisciplinary group of researchers, clinicians and industry representatives, who are jointly committed to developing technologies that enable the early detection and management of subsequent breast cancer by the patient. IWBI was designed to discuss the latest technological advances and clinical experiences with new breast imaging technologies including digital mammography, tomosynthesis, CT scan, MRI, ultrasound, optical and molecular imaging.

I.5.2   Architecture of a CAD in mammography

The task assigned to a medical diagnostic support system on the breast image is a very difficult and complex function at the same time, given the varied appearance of the organ concerned. As a result, the architecture of the system greatly depends on the type of mammography to be processed. Figure 1.7 shows a basic diagram of a screening or diagnostic mammography interpretation incorporating a CAD system.
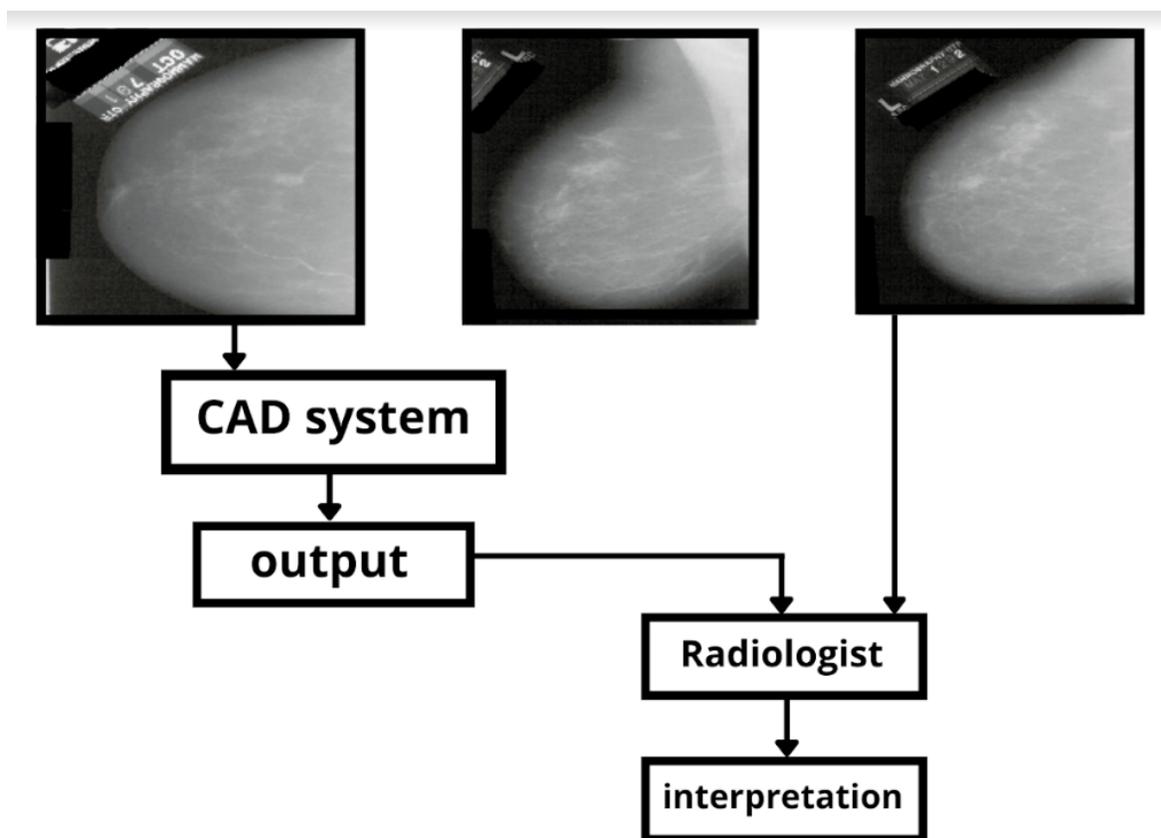
Figure I.7: Diagram of a basic scheme of an interpretation of screening/diagnostic mammography.

CAD systems can be classified into two classes: CADe and CADx. CADe systems have, fundamentally, a detection task, i.e. a localization task. Indeed, radiologists use the location of suspicious regions as an output of the CADe ignoring the characterization and diagnostic functions. On the other hand, CADx systems extend automatic analyses to produce results on the characterization of a region or a lesion, initially localized by the human or by the CADe.

The steps generally involved in the design of CADe and CADx are preprocessing, segmentation, description and classification. They correspond exactly to those of an artificial vision system where the data to be processed is the medical image. Also called computer vision or digital vision, this discipline of artificial intelligence aims to make useful decisions about real objects and physical scenes from captured images.
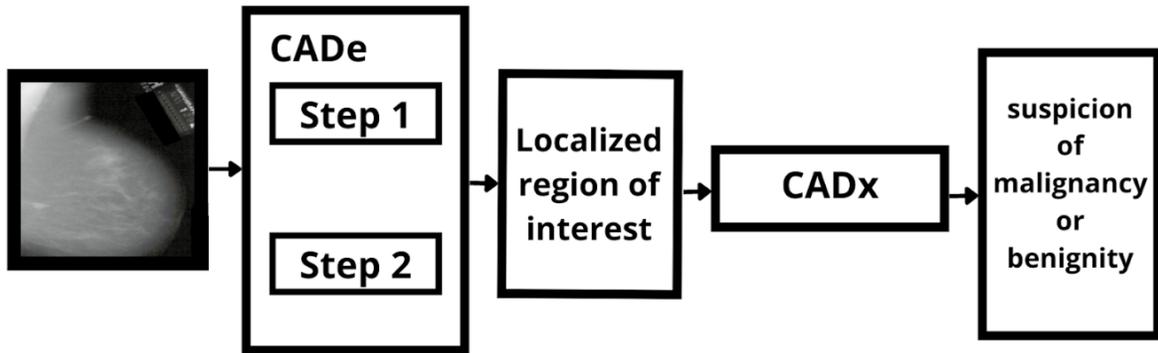
Figure I.8: Synoptic diagram of a CAD system.

I.6   Conclusion

To facilitate understanding of the problem, we have discussed in this chapter the notion of breast cancer as well as the medical approach of breast imaging allowing its diagnosis. Particular attention has been devoted to the study of the mammographic image by focusing on the one hand on the specifications of the mammary pathologies and on the other hand on their appearance in mammography.

Deep learning for image processing

Chapter II

Deep learning for image processing

II.1 Introduction to Machine Learning

There has been a significant expansion and development of approaches that help in the discovery and understanding of complicated and hidden patterns in data in the field of computers and information technology. Standard machine learning approaches have been created as a result of these improvements.

Machine Learning is the study of teaching computers to learn and act like people, and to enhance their learning over time in an autonomous manner, using data and information fed to them in the form of observations and real-world interaction [16].

Several machine learning techniques, ranging from simple classifiers to artificial neural networks, have been created throughout the years[17]. The need to categorize cancer patients into high and low risk categories has prompted numerous biomedical and bioinformatics research teams to investigate the use of machine learning (ML) technologies. These approaches are utilized in illness detection, medical imaging diagnosis, medication development, early cancer outcome prediction, and medical research[18].

Feature engineering is required for ordinary machine learning algorithms to reach state-of-the-art performance. Feature engineering is the process of creating or transforming features from data. The engineer's experience, devotion, and skill are required for feature transition.[19]

The performance of these standard machine learning algorithms depend on how accurately each feature is discerned and extracted. The feature engineering process which requires a deliberate effort on the part of an engineer may sometimes result in mistakes. Deep learning is an advanced technique in machine learning that was developed to reduce the need for feature engineering when processing data. Deep learning algorithms automatically learn to extract useful features directly from data and make useful predictions using the extracted features. 4

II.2   Deep Learning

Deep learning, a subset of the machine learning field, is designed to learn hierarchically from a problem and perform classification based on the learned features[20]. Deep learning algorithms utilize a layered structure of algorithms known as an Artificial Neural Network (ANN). The ANN's design was inspired by the structure of the biological neural network of the human brain[21]. A deep learning algorithm's ability to imitate how the human brain works by learning and making intelligent decisions on its own is what distinguishes it from standard machine learning algorithms.

A basic ANN structure shown in Figure 2.2 consists of interconnected neurons arranged in at least three layers. Each neuron in the network represents one aspect of the input data and each connection between neurons is associated with a weight that dictates the importance of an input value and a bias that controls the rate at which a neuron is fired. The layers of an ANN include the input layer, the hidden layers, and an output layer. The input layer receives raw input data and passes the input to the first hidden layer. The hidden layers perform mathematical computations on the input they receive. The output layer receives an input from the preceding hidden layer, performs computation on the input and returns an output value.

An ANN with two or more hidden layers forms a deep neural network. In a deep neural network, features are extracted hierarchically. The net-

work initially learns low-level features and further combines the learned features to form more abstract representations as it progresses into each layer. Features may include shape, orientation, color, texture, position, and pixel values. Deep neural networks have achieved state-of-the-art performance in various fields such as computer vision, natural language processing, and medical diagnosis.

Deep learning has been applied in the field of image classification[22], pattern recognition and other domains. Specifically, in the medical domain Deep learning techniques are revolutionizing the field of medical image analysis and hence in this study, a deep neural network will be applied for the location of tumors detected in the breast and prediction of normal and abnormal tumors, using images from mammograms. [23]
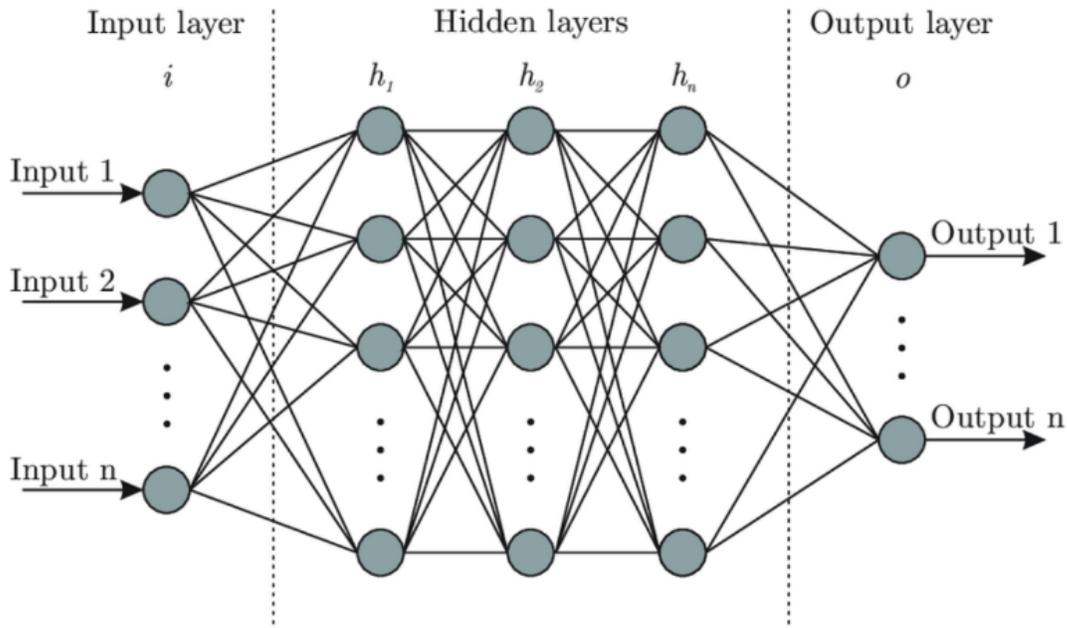
Figure II.1: The structure of an ANN with input, hidden and output layers.[24]

II.2.1 Deep Neural Network Models

There are various types of neural network models that have been developed over the years for use in classification and regression problems,

language processing, speech synthesis, and other relevant areas. Three major neural network models have achieved good performance on varying problems.

- **Convolutional Neural Network (CNN):**

    CNN is one of the most often used deep neural networks. The structure of the CNN allows it to perform well in image data categorization and computer vision. The convolutional layer, the pooling (sub-sampling) layer, and the fully connected (dense) layer[25] are the three layers that make up the CNN's structure. The convolution process conducted in the convolutional layer is the main feature of a CNN network's great performance, thus the term convolutional neural network. In this manuscript, CNN models are used to detect and forecast benign and malignant breast tumors using mammography images.

- **Recurrent Neural Network (RNN):**

    RNN is a deep neural network that detects patterns in sequential data. An input, hidden, and output layer make up a simple RNN structure. The inclusion of a hidden state, which retains sequential information[26], is a key differentiating feature of an RNN network. The network can maintain or store information from previous inputs, process that information, and forecast sequence in the future time step[27] thanks to the hidden state structure. In domains where the sequence of processed information is crucial, such as natural language processing, speech synthesis, and machine translation[28], RNNs have a high percentage of success.

II.3   Convolutional Neural Network

II.3.1   Definition

"A convolutional neural network or convolutional neural network (CNN) is a class of deep neural networks specialized in processing

data having a grid-like topology, like an image. A digital image is a binary representation of data It contains a series of pixels arranged in a grid pattern which in turn contains pixel values to indicate the brightness and color of each pixel."[29]

"Convolutional neural networks were invented with the aim of solving the problem of translation invariance of the objects constituting the images." [30]

II.3.2 Working

Each neuron in a CNN processes data only in its receptive field. By stacking multiple convolution layers one after another and applying a set of operations called convolution operations, the network is able to detect an increasingly complex hierarchy of features. It first detects simpler patterns (lines, curves, etc.) and more complex patterns (faces, objects, etc.) later going to deeper layers. By using a CNN, one can enable the view to computers. A query layer is usually added between successive convolutional layers to summarize important features. This makes it possible to reduce the number of parameters transmitted to the downstream layers and, at the same time, to introduce a translation invariant (capable of recognizing the learned models, independently of their geometric transformations) in the network.

II.3.3 Architecture of a CNN

There are four primary operations involved in the convolutional neural network. Those operations include convolution, non-linearity, pooling or sub-sampling and classification (fully connected layer).
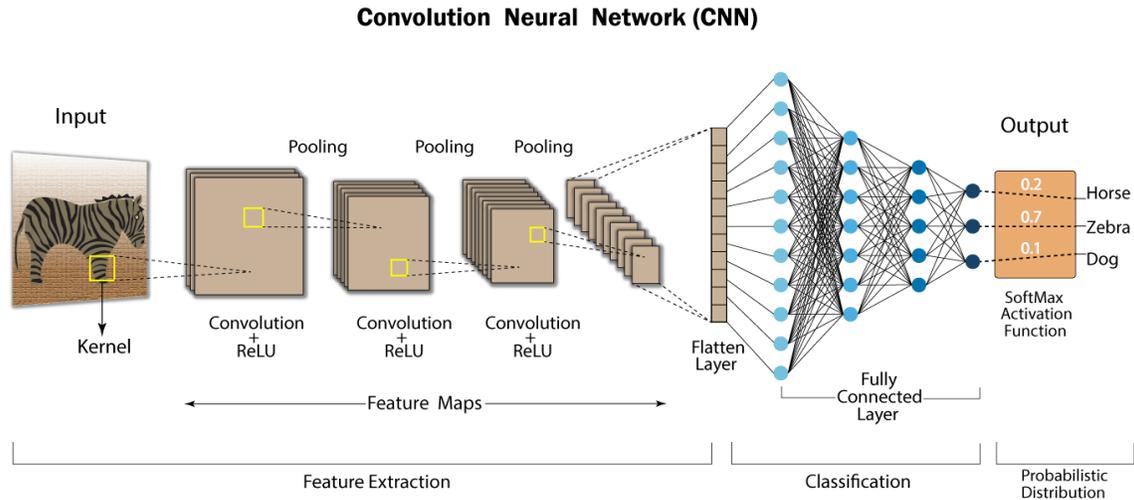
Figure II.2: The CNN architecture takes an input and passes it through a stack of convolutional layer, pooling layer and fully connected layer to produce the desired classes of output values.

II.3.3.1 Activation Functions

Activation functions transform the activation level of a unit into an output signal. Activation functions decide the output of a neuron based on the provided information [31].Three most commonly used activation functions include:

- **Sigmoid Function**

  The sigmoid function takes any range of real numbers and returns an output value which lies within the range of 0 and 1[31]. Shown in Figure 2.3, it is mathematically defined by the equation:

$$f\left(x\right) = \frac{1}{1 + e^{-x}}$$

- **Tanh Function**

  Tanh is a zero cantered function whose range lies between -1 and 1. It has a higher range of derivatives than the sigmoid function which results in a better learning rate[31]. Tanh is shown in Figure 2.3 and given by the equation:

$$f\left(x\right) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- **Rectified Linear Unit Function:**

    The Rectified Linear Unit (ReLU) function, shown in Figure 2.3 has been proved to be the most successful and widely used activation function in deep learning. ReLU imposes a non-saturating non-linearity which enhances better gradient flow with improved calculation efficiency[31]. The range of ReLU lies between 0 and infinity and is mathematically given by the equation:

$$f\left(x\right) = \max\left(0, x\right) = \begin{cases} x_i, & x_i \geq 0 \\ 0, & x_i < 0 \end{cases}$$

(i) Sigmoid Function   (ii) Tanh Function   (iii) ReLU Function
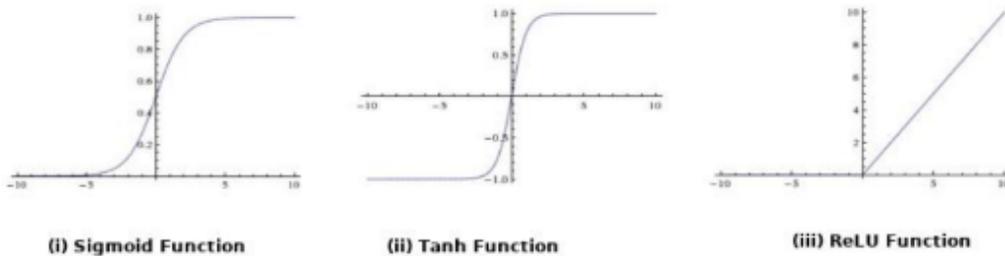
Figure II.3: The sigmoid, Tanh and ReLU activation functions[37]

II.3.3.2 Convolutional Layer

The convolutional layer is one of the fundamental layers of CNN. Convolutional layer performs feature extraction on input and aims to learn feature representations of the input. There are two main functions performed in this layer:

- **Convolution Operation**

  The input to the convolution layer is an array of pixel values ranging from 0 to 255. By sifting through the input data to filter out integral sections of the input, the convolution process aids in extracting the characteristics of the input. Kernels (filters) are an array of integers used as feature extractors during the convolution procedure. For each location, the filters convolve over the input with a specified amount of pixels (stride). The element-wise product of the input matrix and the filter is calculated and saved in a feature map, which is an empty array (activation map). Each convolution operation's exact placement is determined by the stride value. With several filters, this step is repeated indefinitely. A feature map will be created for each computation between the filters and the input array, which will be layered to generate the whole output volume from the convolutional layer. The convolution procedure between a three-dimensional input tensor and three filters is shown in Figure 2.4.

  Mathematically, the discrete convolution between two functions f and g is generally defined by:

  $$(f * g)(x) = \sum_t f(t)\, g(x + t)$$

  Given a 2-dimensional input image tensor, the convolution operation between an input array and the filter can be expressed as:
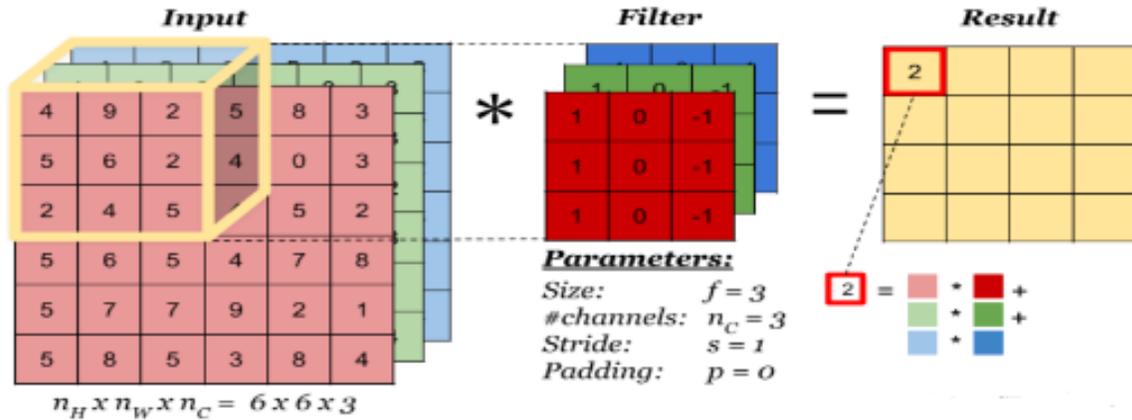
Figure II.4: convolution operation of a 3-Dimensional image tensor with a 3 x 3 kernel size. The kernel is convolved across the input tensor, the dot product yields a 4 x 4 feature map[8].

$$(f * g)(i, j) = \sum_{m} \sum_{n} f(m, n) g(i - m, j - n)$$

where the input image and the kernel are denoted by f and g respectively. i and j are the indexes of rows and columns of the result matrix. This generalizes to higher dimensions. The size of the feature map is regulated by the following three (3) parameters.

- Depth: The depth which is also known as channels is the number of filters used for the convolution operation. (Figure 2.4);

- Stride: The stride is the number of pixels by which the kernel is convolved over the input tensor (larger strides produce smaller feature maps). (Figure 2.5);

- Zero-padding: Zero-padding involves the insertion of a border of pixels with value zero around the edges of an input image to preserve original input size for avoidance of loss of information. (Figure 2.6).
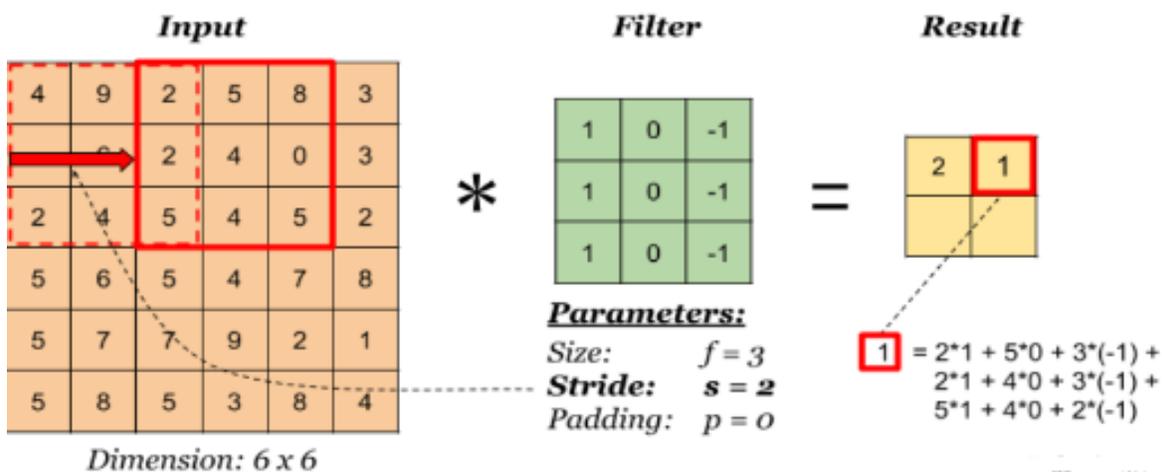
Figure II.5: Filter with stride (s) = 2.[39]

- **Non-Linearization:**

  Activation functions execute a nonlinear modification on the input they receive during non-linearization and deliver the altered output to the next layer as input. The normalization procedure boosts the network's non-linearity without changing the convolutional layer's receptive fields (an area of the input space that affects a specific unit of the network). In the buried layer, ReLU activation is frequently employed for non-linear transformations.

II.3.3.3 Subsampling (Pooling) Layer

Pooling reduces the dimensionality of each feature map albeit preserving the best information of each feature within a region. The pooling operation involves computing the maximum, average or sum over a small window of the feature map. The small window is applied across an image in different regions determined by a chosen value of stride, and the result is retained as shown in Figure 2.6 Pooling can be of different types:

- **Max Pooling:** The max pooling operation computes the maximum of elements in each region;

- Average Pooling: The average pooling operation computes the av-

40
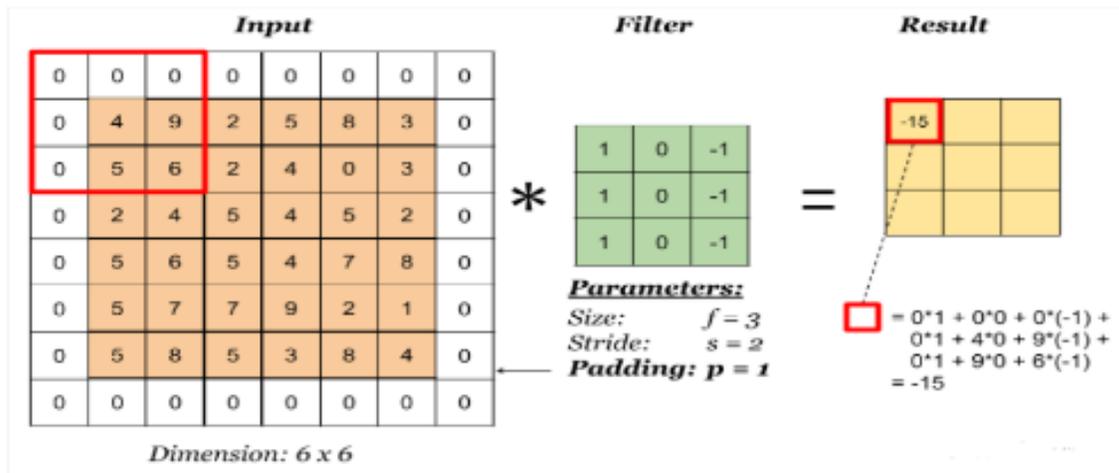
Figure II.6: Padding the dimension of the result has changed due to padding.

erage of elements in each region;

• Sum Pooling: The sum pooling operation computes the sum of all elements in each window.
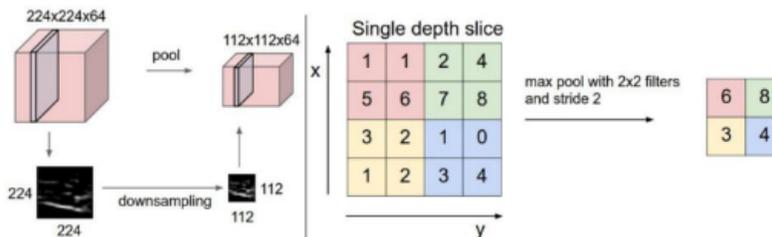
Figure II.7: Left side: A pixel size of $224 \times 224 \times 64$ down sampled to a pixel size of $112 \times 112 \times 64$, Right side: A 4×4 tensor max pooled with stride 2 using a 2×2 filter size to produce a 2×2 feature map [40].

II.3.3.4 Fully-Connected Layer

The fully-connected layer, also known as the dense layer, performs classification based on the features extracted by the other layers. The fully connected layer classifies images by applying the softmax activation function. The output has a probability of numbers ranging from zero to one for each classification label the model is trying to predict.

II.4 Learning Process

In a deep neural network, learning refers to the process of optimizing the network's weights to minimize the error in an algorithm. An error function is created to quantify the difference between the desired output value and the projected output value in order to monitor the progress of a network. The backpropagation method is used to update the weight and bias values in relation to the error function.

II.4.1 Backpropagation

The backpropagation algorithm aims to understand how a slight change in the values of the weight and bias affect the error function. Through backpropagation, the computation of error terms will proceed backward from the output layers to the input layers iteratively. The values of the weight and bias will be adjusted with each iteration until the error function is minimized. There are many ways in which an error function may be represented. The choice of an error function depends on the choice of user or the type of network to be used.

II.4.2 RegulariZation techniques

When given a big dataset to train on, a deep neural network may deliver excellent results. As a result, if the dataset isn't large enough, the network will overfit or underfit the training set. When a neural network is overfitted to the training data, it loses its ability to generalize and generate effective predictions on fresh data[32]. Similarly, underfitting

happens when the model is unable to train due to a lack of data. Some regularization approaches are mentioned in this part to avoid overfitting and underfitting in a network. These strategies have been presented to increase the generalization and performance of a model by modifying an algorithm.

- **L1/L2 Regularization:**

  L1 and L2 regularization techniques reduce overfitting in a network by penalizing the loss function as shown in the equations:

$$L1 : w^* = Loss\,(y, \hat{y}) + \lambda \sum_{i=1}^{k} |w_i|$$
$$L2 : w^* = Loss\,(y, \hat{y}) + \lambda \sum_{i=1}^{k} (w_i)^2$$

  The loss function is updated by the addition of a regularization term which regulates the values of the weight matrices[33]. The regulation parameter modulates how strongly weights are penalized and can be optimized for better results. L1 regularization adds an absolute magnitude of coefficient as a penalty term to the loss function to reduce weight values of less important features in the network. L2 regularization adds the squared magnitude of coefficient as a penalty term to the loss function. These regularization terms can both be used or employed separately to improve a network's performance.

- **Batch Normalization:**

  The normalization of an input data to have zero mean and unit variance is often implemented to speed up the learning process of a network. Batch normalization is a regularization technique proposed by Szegedy et al. (2016)[34] to further improve the generalization ability of a network by

  applying normalization to each hidden layer. The batch normalization technique is executed by applying normalization to every value from the preceding layer before it is passed on to the next layer[34].

Each value passed on to the batch normalization layer represents a batch of values to be normalized. The normalization process is described thus: Given values from a previous layer for each observation i in a dataset. The mean and variance are computed as:

$$\mu = \frac{1}{n} \sum_i z_i^l$$

and

$$\sigma^2 = \frac{1}{n} \sum_i \left(z_i^l - \mu \right)$$

Using these values, Z^l *normalized as* :

$$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 - \varepsilon}}$$

ε: small number to prevent zero division.

Practically, restriction of activation values to zero mean and unit variance could limit the network's optimal performance. Thus, parameters γ and β are applied *to scale normalized values by* α *and shift by*

$$\beta \text{ such that } \tilde{z}^{(i)} = \gamma z_{norm}^{(i)} + \beta$$

- **Drop-Out:**

  Dropout is another regularization technique proposed by Srivastava et al. (2014)[21], *used to improve a network's learning capability.*

  The Dropout process can be applied to both the input layer and the hidden layers. Srivastava et al. (2014)[35] describe the technique as a process

of random selection of some nodes which are dropped out of the network with a chosen probability 'p' during the training process. By dropping a unit out, all its incoming and outgoing connections are temporarily removed. The dropout process is performed iteratively, each with different nodes to be dropped out and formally dropped out nodes restored. By repeating this process, each node learns independently of other nodes and increases the optimal performance of the network.

- **Data Augmentation:**

  One of the major challenges in the training implementation in a deep neural network is the lack of sufficient datasets for training the network. Data augmentation can reduce overfitting and underfitting by artificially creating new data samples through transformation. The technique increases the size of the training data by the generation of relevant samples in various orientations and sizes. The transformation process includes image rotation, scaling, shifting, flipping, etc... In general, data augmentation increases a network's ability to generalize well.

II.5 Pretrained CNN Models

CNN models of varying structures have been developed to solve image classification and detection problems.

II.5.1 VGG 16 for image classification

II.5.1.1 Introduction

VGG stands for Visual Geometry Group; it's a well-known deep CNN structure with more than one layer. The "deep" refers back to the range of layers with VGG-16 or VGG-19 along with 16 and 19 convolutional layers. The VGG structure is the premise of ground-breaking item reputation models. Developed as a deep neural network, the VGGNet additionally surpasses baselines on many responsibilities and datasets past ImageNet. Moreover, it's miles now nevertheless one of the most famous picture reputation architectures.

II.5.1.2. VGG architecture

- **Input**

  The VGGNet takes in an image input size of 224×224. For the ImageNet competition, the creators of the model cropped out the center 224×224 patch in each image to keep the input size of the image consistent.

- **Convolutional Layers**

  VGG's convolutional layers leverage a minimum receptive field, i.e., 3×3, the smallest possible length that also captures up/down and left/right. Moreover, there also are 1×1 convolution filters appearing as a linear transformation of the center. This is accompanied through a ReLU unit, that's a big innovation from AlexNet that reduces schooling time. ReLU stands for rectified linear unit activation characteristic; it's a piecewise linear characteristic on the way to output the center if positive; otherwise, the output is zero. The convolution stride is constant at 1 pixel to preserve the spatial decision preserved after convolution (stride is the variety of pixel shifts over the enter matrix).

- **Hidden Layers**

  All the hidden layers withinside the VGG community use ReLU. VGG does now no longer commonly leverage Local Response Normalization (LRN) because it will increase reminiscence intake and education time. Moreover, it makes no enhancements to general accuracy.

- **Fully-Connected Layer**

  The VGGNet has 3 completely related layers. Out of the 3 layers, the primary has 4096 channels every, and the 0.33 has a thousand channels, 1 for every class.

II.5.1.3 VGG16 architecture

The wide variety sixteen withinside the call VGG refers back to the reality that it's miles sixteen layers deep neural community (VGGnet). This means that VGG16 is a quite tremendous community and has a total of around 138 million parameters. Even in accordance with fashionable standards, it's miles from a large community. However, VGGNet16 architecture's simplicity is what makes the community extra appealing. Just with the aid of searching at its architecture, it may be stated that its miles are pretty uniform. There are some convolution layers accompanied with the aid of using a pooling layer that reduces the peak and the width. If we examine the wide variety of filters that we are able to use, around sixty four filters are to be had that we are able to double to approximately 128 after which to 256 filters. In the closing layers, we are able to use 512 filters.

II.5.1.4. Configuration of VGG16

We can see that there are 2 variations of VGG-16 (C and D). There isn't a distinction among them besides for one which besides for a few convolution layers there is (3, 3) clear out length convolution is used as a substitute of (1, 1). These carry 134 million and 138 million parameters respectively.

II.5.1.5 Complexity and challenges

The variety of filters that we will use doubles on each step or through each stack of the convolution layer. This is a primary precept used to layout the structure of the VGG16 community. One of the vital downsides of the VGG16 community is that it's a large community, because of this that it takes greater time to teach its parameters. Because of its intensity and variety of absolutely linked layers, the VGG16 version is greater than 533MB. This makes enforcing a VGG community a time-eating task. The VGG16 version is utilized in numerous deep studying

photograph class problems, however smaller community architectures inclusive of GoogLeNet and SqueezeNet are frequently preferable. In any case, the VGGNet is a first-rate constructing block for studying functions because it is easy to implement.

II.5.1.6. Performance of VGG Models

VGG16 incredibly surpasses the preceding variations of models withinside the ILSVRC[12]-2012 and ILSVRC-2013 competitions. Moreover, the VGG16 end result is competing for the category project winner (GoogLeNet with 6.7% error) and notably outperforms the ILSVRC-2013 prevailing submission Clarifai[13]. It received 11.2% with outside schooling records and around 11.7% without it. In terms of the unmarried-internet performance, the VGGNet-sixteen version achieves the first-rate end result with approximately 7.0% check error, thereby surpassing an unmarried GoogLeNet via means of around 0.9

II.5.2 Yolo v5 for object detection

II.5.2.1 Introduction

YOLO is a shortened form for the term 'You Only Look Once'. This is an algorithm that detects and recognizes different objects in an image (in real-time).

II.5.2.2 YOLO history

- **YOLO v1**

  YOLO was the first object detection network to combine the problem of drawing bounding boxes and identifying class labels in one end-to-end differentiable network.

---

[12]ILSVRC : evaluates algorithms for object detection and image classification at large scale.
[13]Clarifai : an artificial intelligence (AI) based solution, which helps businesses manage processes related to data organization,

- **YOLO v2**

  YOLOv2 made some iterative enhancements on the pinnacle of YOLO along with BatchNorm, better resolution, and anchor boxes.

- **YOLO v3**

  YOLOv3 constructed upon preceding models by including an objectness score[14] to bounding box prediction, delivered connections to the spine community layers and made predictions at 3 separate ranges of granularity to enhance overall performance on smaller objects.

- **YOLO v4**

  On YOLOv4, it's far from a two-degree detector with numerous additives to it. Each factor might be damaged down similarly within-inside the later segment of the blog.

II.5.2.3 YOLO v5

YOLOv5 is a family of object detection architectures and models pre-trained on the COCO dataset, and represents Ultralytics open-source research into future vision AI methods, incorporating lessons learned and best practices evolved over thousands of hours of research and development.

YOLO v5 Architecture

Until now we've mentioned techniques used to enhance the version accuracy and extraordinary components of an item detector (backbone, neck, head). Let us now communicate approximately what's used within-inside the new YOLO.

---

[14]the objectness score is defined to measure how well the detector identifies the locations and classes of objects during navigation.

- **Backbone**

  A convolutional neural network that aggregates and forms image features at different granularities.

- **Neck**

  A series of layers to mix and combine image features to pass them forward to prediction.

- **Head**

  Consumes features from the neck and takes box and class prediction steps.

Models Available in YOLOv5

YOLOv5 contains 5 models in total. Starting from YOLOv5 nano (smallest and fastest) to YOLOv5 extra-large (the largest model):

- **YOLOv5n**

  It is a newly introduced nano model, which is the smallest in the family and meant for the edge, IoT devices, and with OpenCV DNN support as well. It is less than 2.5 MB in INT8 format and around 4 MB in FP32 format. It is ideal for mobile solutions.

- **YOLOv5s**

  It is the small model in the family with around 7.2 million parameters and is ideal for running inference on the CPU.

- **YOLOv5m**

  This is the medium-sized model with 21.2 million parameters. It is perhaps the best suited model for a lot of datasets and training as it provides a good balance between speed and accuracy.

- **YOLOv5l**

  It is the large model of the YOLOv5 family with 46.5 million cparameters. It is ideal for datasets where we need to detect smaller objects.

- **YOLOv5x**

  It is the largest among the five models and has the highest mAP among the 5 as well. Although it is slower compared to the others and has 86.7 million parameters.

Model Name	Params (Milion)	Accuracy (mAP 0.5)	CPU Time(ms)	GPU Time(ms)
YOLOv5n	1.9	45.7	45	6.3
YOLOv5s	7.2	56.8	98	6.4
YOLOv5m	21.2	64.1	22.4	8.2
YOLOv5l	46.5	67.3	430	10.1
YOLOv5x	86.7	68.9	766	12.1

Table II.1: Better overview of all YOLOv5 models including CPU and GPU inference speed and also the number of parameters with an image size of 640.

Second part

Realization,Test, Results

Chapter III

Realization,Test, Results

III.1   Introduction

We give our contribution in this chapter after a review of the state of the art in computer-assisted systems in breast imaging and an examination of the various detection methodologies. Our main goal is to create a prototype of a fully automated digital mammography detection assistance system that can be connected to a digital mammography device and used by radiologists as a second reader. The main challenge is to provide the doctor with a high-performance, intelligent, and scalable tool to increase sensitivity to image readings while highlighting potentially significant lesions.

III.2   Description of the proposed detection aid system

The proposed system to help detect suspicious areas, such as breast masses, which automatically identifies abnormal regions on screening mammograms. It is designed to provide a second opinion, to help rather than replace practitioners. Our CAD system, applied to a database of digital mammograms, consists of two parts:

- The first part locates the location of the affected mass using YOLOv5;

- The second part classifies the type of mass, benign or malignant using VGG16.
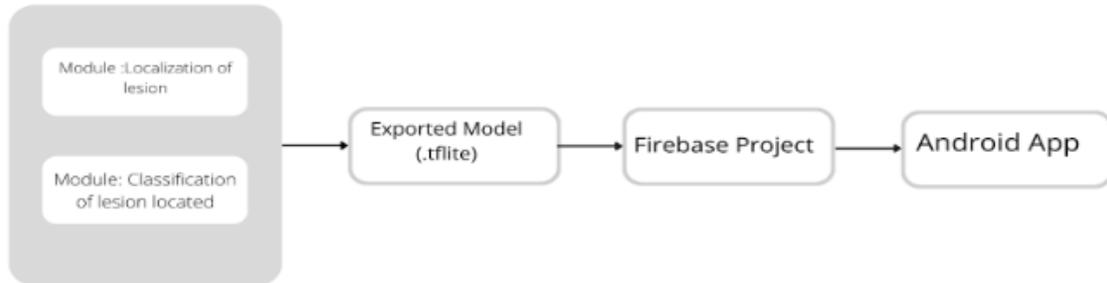
Figure III.1: An overview of the proposed CAD.

Where :
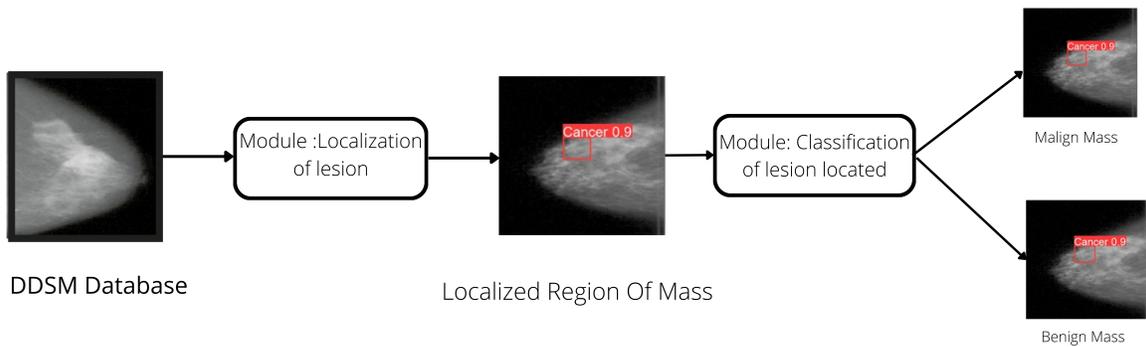
Figure III.2: Localization and classification of mass in DDSM database.

Module 1 : Localization of lesion
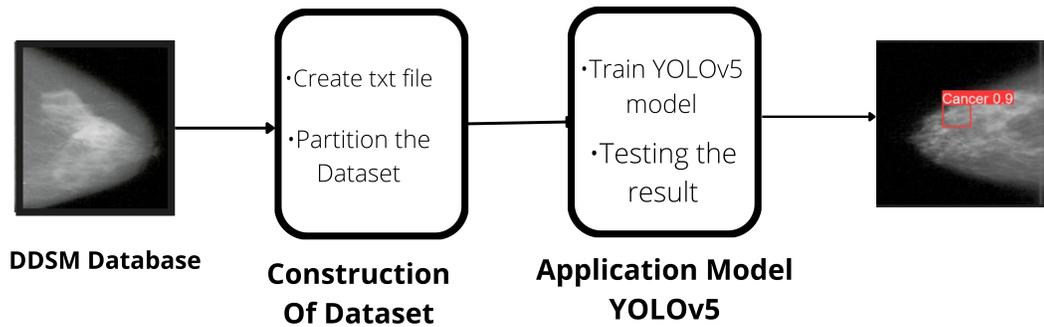
Figure III.3: Applying YOLOv5 for mass detection.

Module 2 : Classification of lesion located
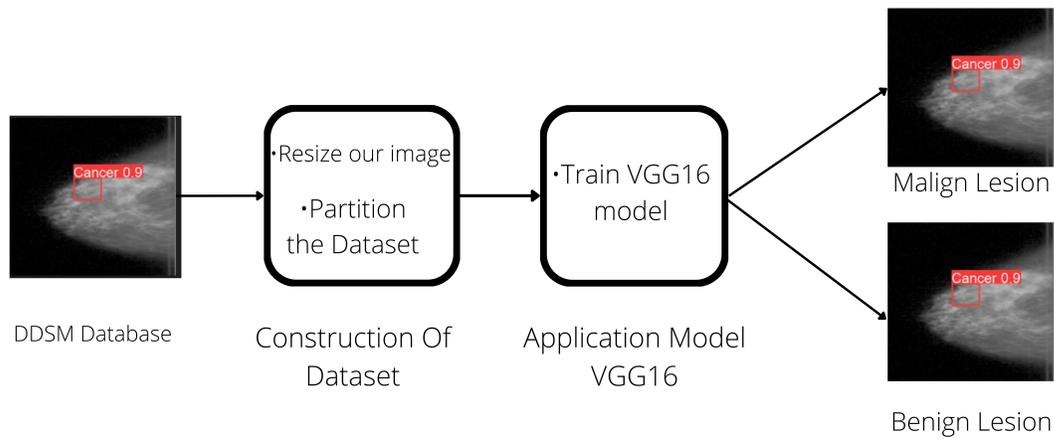
Figure III.4: Applying YOLOv5 for mass detection.

III.3   Presentation of the image database used

Various works dealing with the issue of breast cancer prevention, through automatic diagnostic assistance systems, make use of mammography image databases, some of which are public and others restricted to individual resear ch groups in the field.

The DDSM presented in [Heath et al, 1999, which is the largest and most widely used public database. It contains 2620 cases with two images for each breast, including a total of 10,480 images of all types.

We used 778 images from the DDSM dataset.

III.4   Used tools

Visual Studio

Visual Studio Code is an extensible code editor developed by Microsoft.

Latex

LaTeX is a document composition language and system.

LabelImg

A graphical image annotation tool for obtaining bounding box coordinates (in YOLO and PASCAL formats).

Google Colab

Colab is a free notebook environment running wholly in the cloud. Most importantly,Colab does not require a setup, plus the notebooks that you will create can be simultaneously edited by your team members. The greatest advantage is that Colab supports the most popular machine learning libraries which can be easily loaded in your notebook.

Figma

Figma is a vector graphics editor and prototyping tool. It is primarily web-based, with additional offline functionality enabled by desktop apps for macOS and Windows.

Android studio

Android Studio is a development environment for developing Android mobile applications. It is based on IntelliJ IDEA and uses the Gradle build engine.

Firebase

Firebase is a set of hosting services for any type of application. It offers to host in NoSQL and in real time databases, content, social authentication, and notifications, or even services, such as a real-time communication server.

Canva

Canva is a graphic design platform that allows users to create social media graphics, presentations, posters, documents and other visual content.

III.5   Implementation

III.5.1   Module 1: localization of lesion

First, we'll take processed mammography images and create.txt files for each one in this module. Second, we'll use the data we gathered to train theYOLOv5 model in order to determine the cancer's location. After receiving the results, we discuss them before putting them to the test.

III.5.1.1 Construction of dataset

We used LabelImg to label the photographs we intended to utilize after gathering them.
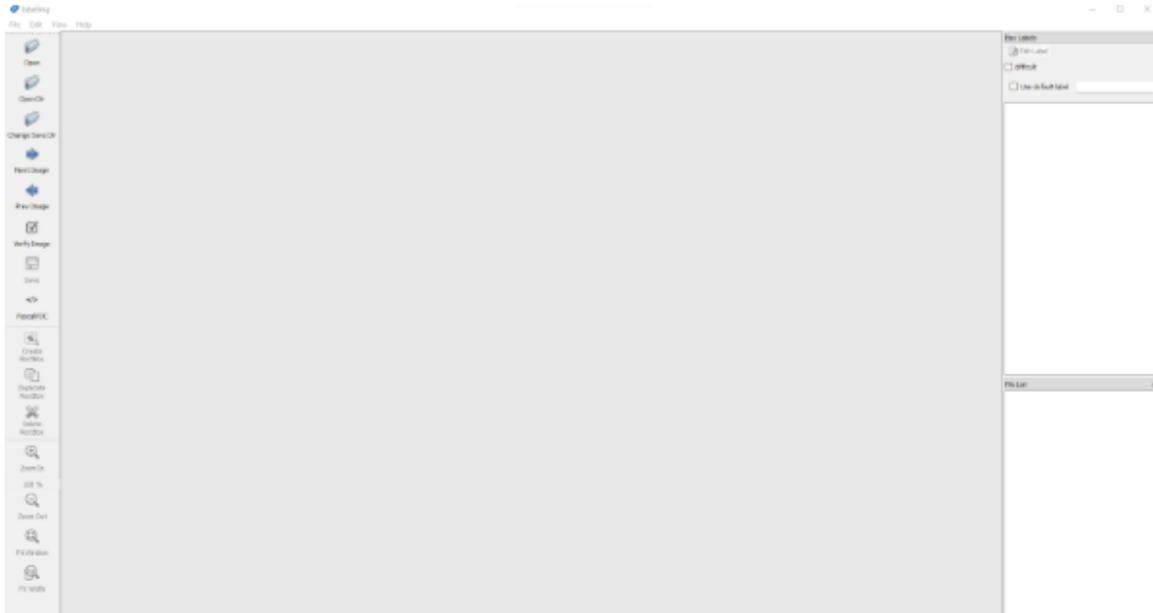
Figure III.5: LabelImg .

Because the default format is PASCAL, we modify it to YOLO, draw a bounding box around the ROI, name the label, and save it.

Figure III.6: Phase of annotation. .

When we save a txt file is automatically created containing the coordinates of the bound-ing box. We must keep our annotations and our images in the same directory.

.

Figure III.7: Annotations.txt. .

Where:

- 1 : The class id;

- 2 :x (center of the bounding box) ;

- 3 :y (center of the bounding box) ;

- 4 :w (width of the bounding box) ;

- 5 :h (height of the bounding box) .

III.5.1.2 partition of data

We divide the data into two parts: 90% for training and 10% for validation.
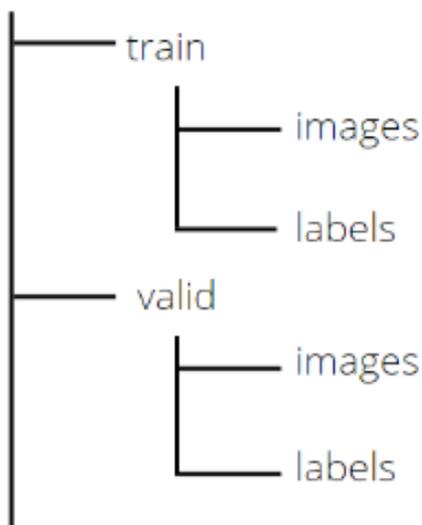

Figure III.8: Diagram of divided data. .

where : The images folder contains the mammography images and the labels folder contains the txt files.

III.5.1.3 Environment Setup

The important thing to notice is that you will require PyTorch version 1.5, Python version 3.7, and CUDA version 10.2. The rest of the dependencies can easily be installed using pip or the requirement.txt file.

III.5.1.4 Applying YOLOv5

- The first step is to enable GPU in colab;

- Then with cloning repository for YOLOv5 on our drive:
  !git clone https://github.com/ultralytics/yolov5.git;

- We need to add a YAML file to describe the settings for our dataset;

61

```
%%writefile data.yaml
path: "../" # Path relative to the `train.py` script.
train: images/train
val: images/val

# Classes
nc: 1
names: ["Cancer"]

Writing data.yaml
```

Figure III.9: YAML FILE..

Now everything is configured and we are ready to train our YOLOv5 model.

- use the following command to start training:

```
RES_DIR = set_res_dir()
if TRAIN:
    !python train.py --data ../data.yaml --weights yolov5l.pt \
    --img 640 --epochs 32 --batch-size 24
```

Figure III.10:   Script to start training.

Where:

- img: size of the input image;

- batch: batch size;

- epochs: number of epochs;

- data: YAML file;

- Weights: model selection YAML file. we have chosen "l" .

III.5.1.5 Result

After the training is completed, the model will be saved in our "Weight" directory and the resulting matrix diagram will be generated as follows.
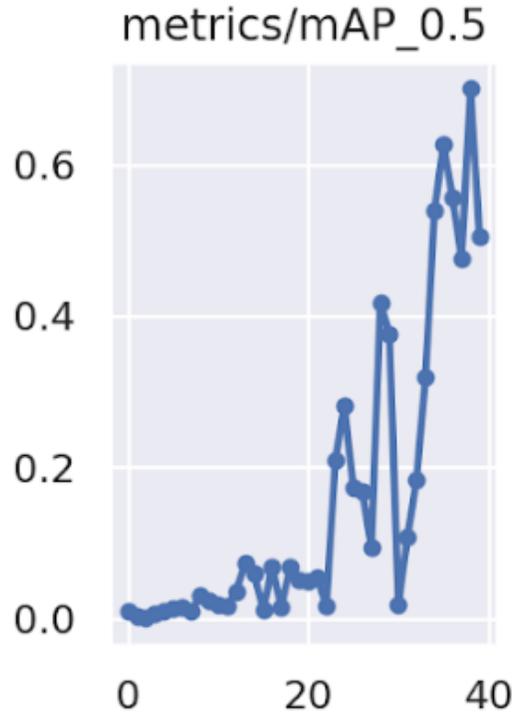
Figure III.11:   Result/metrics mAP.

As we can see the model is trained pretty well the mAP@0.5 is almost 0.7 in 40 epoch which is pretty good.

Mean Average Precision(mAP) is the current benchmark metric used by the computer vision research community to evaluate the robustness of object detection models.The mean of average precision(AP) values are calculated over recall values from 0 to 1.

Where:

$$AP = \int\limits_{0}^{1} P(r)dr \quad mAP = \frac{1}{N} * \sum_{i=1}^{N} AP_i$$

and

N: The number of classes. **P:** Precision. **r:**Recall.

mAP formula is based on the following sub metrics:

- **Confusion Matrix**

    To create a confusion matrix, we need four attributes:

True Positives (TP):   The model predicts a label and matches correctly as per ground truth.

True Negatives (TN):   The model does not predict the label and is not a part of the ground truth.

False Positives (FP):   The model predicted a label, but it is not a part of the ground truth (Type I Error).

False Negatives (FN):   The model does not predict a label, but it is part of the ground truth. (Type II Error).

in our model:
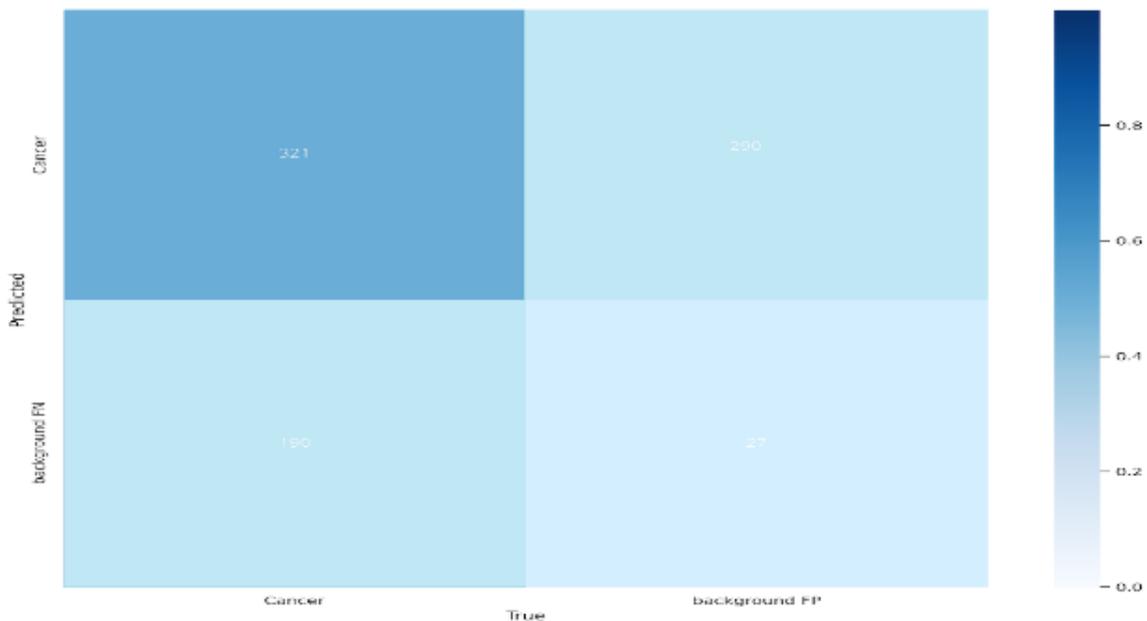
Figure III.12:   Confusion matrix/YOLOv5.

Where:

	Positive	Negative
Positive	TP=321	FN=190
Negative	FP=27	TN=290

Table III.1: Confusion matrix/YOLOv5.

• **Precision**

Precision measures how well you can find true positives(TP) out of all positive predictions. (TP+FP).

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

in our model:
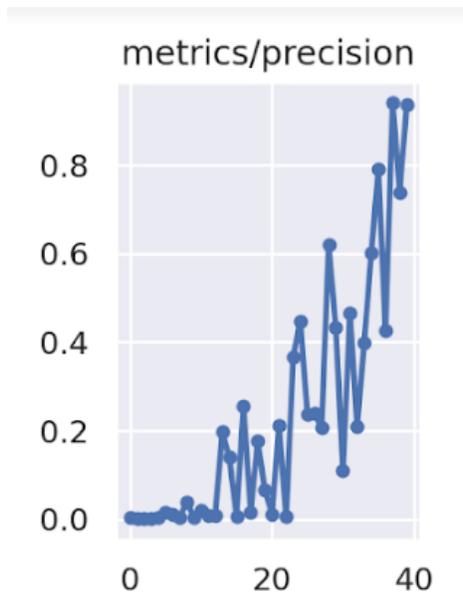
Figure III.13:   Result/metrics precision.

As can be seen, the model's precision is around 0.92 in 40 epochs, which is rather impressive.

• **Recall**

Recall measures how well you can find true positives(TP) out of all predictions(TP+FN).

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

in our model:
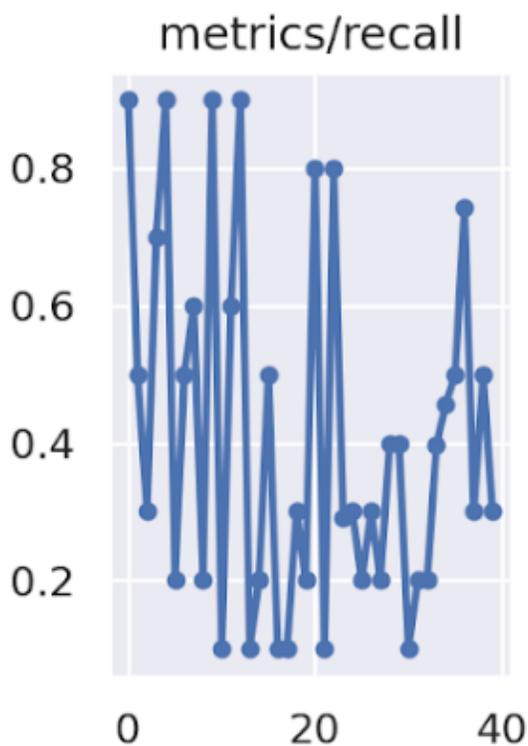
metrics/recall

Figure III.14:   Result/metrics recall.

As we can see the recall of the model is almost 0.62 in 40 epoch which is average.

Note: Before obtaining these results, we made more than 100 attempts.The changement of the results depend principally on the weight of the chosen pretrained model. This is a comparison between some of them:

Where:

Model	Epochs	Batch	Weight	mAP	Précision	Recall
model1	40	24	YOLOv5l	70%	92%	62%
model2	25	8	YOLOv5m	25%	52%	48%
model3	25	8	YOLOv5l	20%	39%	27%
model4	36	24	YOLOv5m	49%	63%	51%

Table III.2: Evaluation Metrics/YOLOv5.

III.5.1.6 Testing

To run the inference form, we use the following command.

python detect.py —source sample$_i$$mg/--weightsweights/best.pt--conf 0.4$

Where

- **source: input images directory or single image path;**

- **weights: trained model path;**

- **conf: confidence threshold.**

Figure III.15:   Prediction.

Here we used a set of mammography images that contain a lesion.

Figure III.16:   Prediction 2.

**Here we used a set of mammographic images that did not contain a lesion.**

Figure III.17:   Prediction 3.

Here we used a set of mammographic images from outside the DDSM dataset that we did not process and the rest were in normal size.

Note: The prediction wavered because we used 90% of the dataset for training and only 10% for validation (because the

dataset is small, if we had a larger dataset we would split the train and test set differently).

III.5.2   Module 2: Classification of lesion

III.5.2.1 Data preparation

 The dataset is divided into two categories: benign and malignant tumors. We used the three-way data split to divide the dataset for the training procedure. We separated the benign and malignant tumors in a ratio of 49:21:30 in the train, validation, and test sets to achieve a balanced distribution of both benign and malignant tumors. According to the stated ratio, the benign and malignant tumors in each dataset were well represented in each collection. We had 1960 samples in the training set, 840 samples in the validation set, and 1200 samples in the test set after dividing a total of 4000 samples.

III.5.2.2 Result

   The performance of VGG16 model was evaluated using the confusion matrix, accuracy, precision, and recall. The next figures show the graphical representation of the training and validation accuracies with respect to the number of epochs for the model.

predicted

	Positive	Négative
Positive	530	67
Négative	77	523

Table III.3: Evaluation Metrics.

Where

Figure III.18:   The Confusion Matrix/VGG16.

- Accuracies

Where:

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}}$$
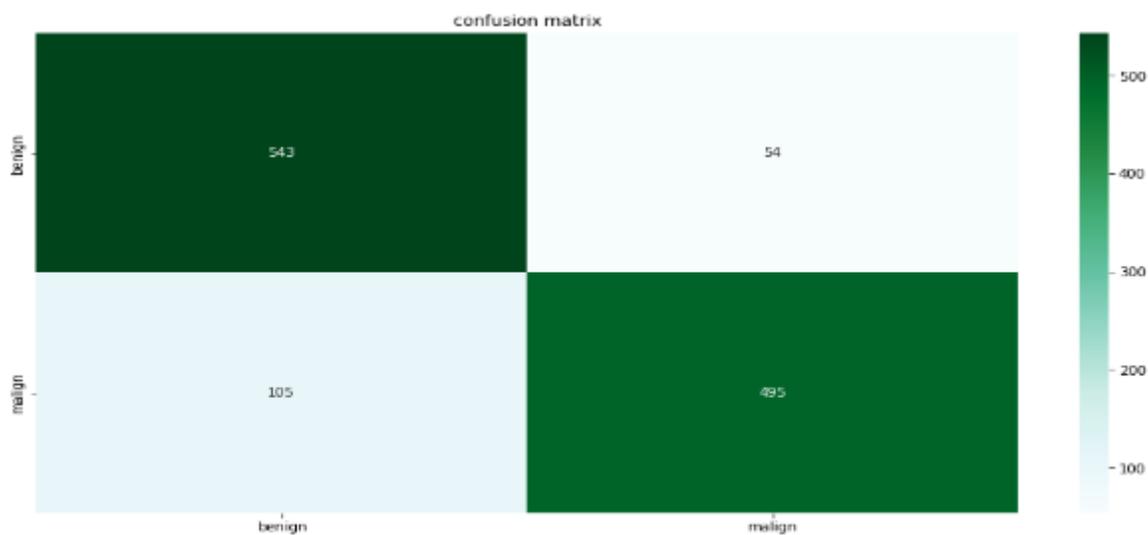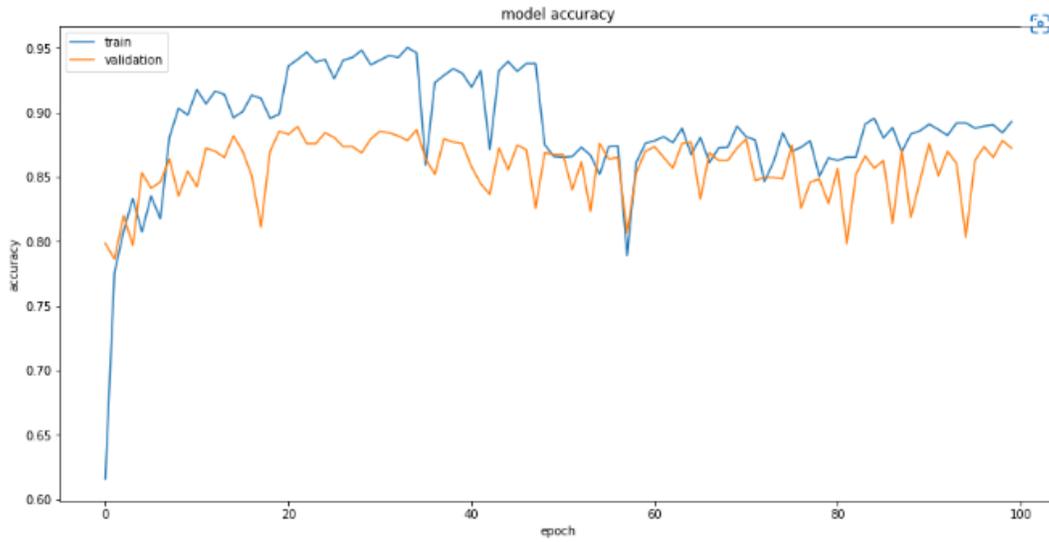
Figure III.19:   Training and Validation accuracies.

Model	Accuracy	Précision		Recall	
		Benign	Malignant	Benign	Malignant
VGG16	88%	87%	89%	89%	87%

Table III.4: Evaluation Metrics.

III.6   Prototype Spiro APP

III.6.1   MOBILE DEVELOPMENT

The process of developing software for use on a mobile device is known as mobile application development. The number of cellphones in use has increased by 2.5 billion in the last ten years. Consumers spend 380 billion on new electronics every year. Each offers applications that make life easier, such as counting, communicating, and ordering a cab or meals. A mobile app is software designed for smartphones and tablets running a variety of operating systems, including iOS, Android, KaiOS, Harmony OS, Tizen, and others. If it's connected to remote cloud services, it'll almost always need Internet connectivity.

III.6.2   APP IMPLEMENTATION

The Android mobile platform was chosen since it is popular among smartphone users. Android Studio is the official Integrated Development Environment (IDE) for developing Android apps. It includes a robust code editor as well as developer tools. We picked Kotlin as our programming language since it is a cross-platform programming language that can be utilized instead of Java for Android app development. In 2017, it was also launched as a supplementary "official" Java language. Kotlin is a Java-like language that runs on the Java Virtual Machine.

III.5.2.1 Spiro logo

Simple logo idea Colors symbolizing the hope of recovery from breast cancer.

Figure III.20:   Logo spiro.

III.5.2.2 . Spiro :UI,UX
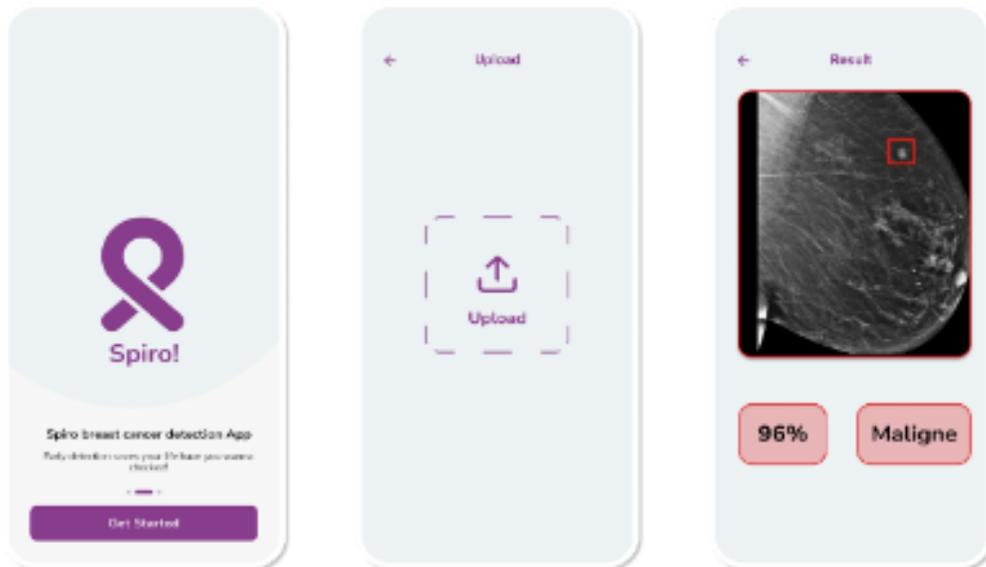
Figure III.21:   Spiro :UI,UX .

III.5.2.3 Models integration on a mobile application
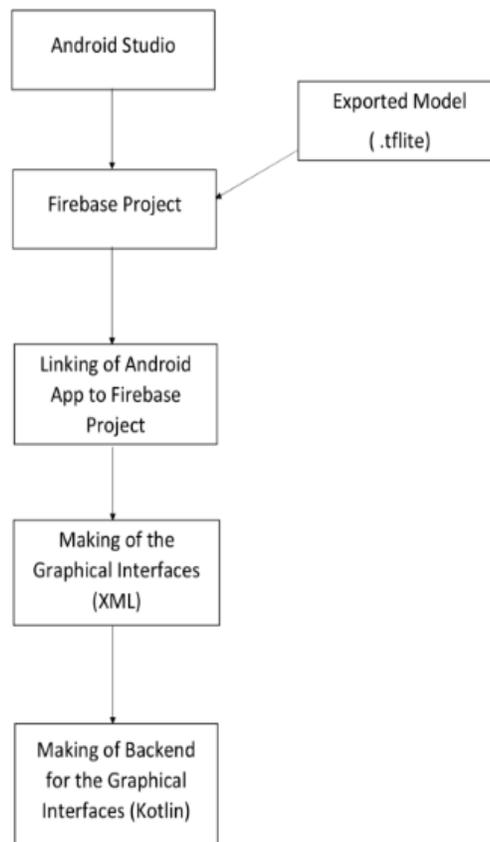
Figure III.22:   Spiro application Development workflow.
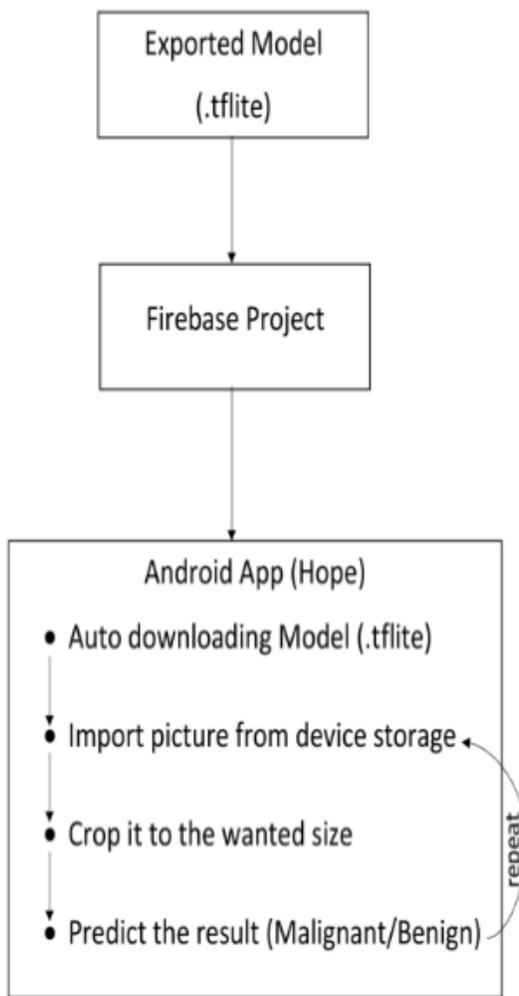
Figure III.23:   Spiro application's Life Cycle.

III.7   Conclusion

We demonstrated the tools we utilized in our research, how we implemented YOLOv4 and VGG16 on our custom dataset, the challenges we experienced with the original image size, and our final findings in this chapter. We didn't get high scores due to the lack of a powerful computer and the need to use Google Colab, which only allows access to the GPU for 12 hours straight, but we were able to get an accuracy of 70 percent in YOLOv4 and 88 percent in VGG16, which I felt was the result of Respectful given all that is described above.

General Conclusion and Prospects

General Conclusion and Prospects

General conclusion

We set out to create a "Computer-Aided Diagnostic" (CAD) Diagnostic Support System to aid in the early diagnosis of breast cancer in this project. We had a lot of problems using our resources because we didn't have a powerful machine or a strong GPU, but we did manage to create a diagnostic support system that identifies malignant lesions in mammograms. We didn't get very excellent results, but with the correct machine, GPU, and dataset, we could obtain pretty good accuracy. We'd like to have better results in the future, so we'll concentrate on improving accuracy and detecting other sorts of lesions (malignant, benign and normal).

Perspectives

Improvements can be made to our solution in order to improve it. Among the most relevant:

- Increase the volume of processed data;

- Increasing the amount of resources;

- Création d'un système de connexion d'empreintes digitales;

- Programming the app mobile prototype;

- Connecting the application to the mammogram device;

- Get a desktop app;

- Writing and publishing a scientific article;

- Giving more credibility to research by trying to contact the Ministry of Health and presenting it to it.

Bibliography

[1] . Sekaran, K., Ramalingam, S. P., and P.V.S.S.R., C. M. Breast Cancer Classification Using Deep Neural Networks, pages 227–241. 02 2018. ISBN 978-981-10-6679-5. doi: 10.1007/978-981-10-6680-1$_1$2.

[2] . Halalli, B. and Makandar, A. Computer Aided Diagnosis - Medical Image Analysis Techniques. In Breast Imaging. InTech, Jan 2018. doi: 10.5772/intechopen.69792. URL http://www.intechopen.com/books/breast-imaging/computer-aided-diagnosis-medical-image-analysis-techniques.

[3] . Wang, J., Ding, H., Bidgoli, F. A., Iribarren, B. Z. C., Molloi, S., and Baldi, P. Detecting cardiovascular disease from mammograms with deep learning. IEEE Transactions on Medical Imaging, 36:1172–1181,2017.

[4] . A. Domart, J. Bouneuf. "New Larousse Medical". Larousse bookstores, 1999.

[5] . Mohamed Ridha Hachana. "Contribution à l'étude de l'implication des virus et évaluation de leur valeur pronostique dans le cancer du sein de la femme dans la région du centre tunisien". Thèse de Doctorat, Institut Supérieur de Biotechnologie de Monastir, Université de Monastir, 2009.

[6] . Hamdi Cherif M, Bidoli E et al. "Cancer estimation of incidence and survival in Algeria 2014 Journal of Cancer Research Therapy". Vol. 3(9), October 2015, pp. 100–104.

[7] . J.L. Kelsey, L. Bernstein, "Epidemiology and prevention of breast cancer", Public Health, No. 17, pp. 47-67, 1996.

[8] . B. Fisher, J. Constantino, D.Wickerham, et al., "Tamoxifen for prevention of breast cancer: Report of the national sugical adjuvant breast and bowel project p-1 study", National Journal of Cancer, Vol. 90(18), pp. 1371-1388, 1998.

[9] . D. Easton, D. Bishop, D. Ford, et al. "Genetic linkage in familial breast and ovarian cancer: result from 214 families". American Journal of Human Genetic, No. 4, pp. 678-701, 1993.

[10] . Patricia Soler-Michel, Christine Lasset. "Chapter 3: Cancer Screening of the breast". Diseases of the breast, 2nd edition, Elsevier Masson, pp 29-40, April 2011.

[11] . M. Yaffe. "Development of full field digital mammography". Proceeding of the 4th International Workshop on Digital Mammography (Nijmegen, The Netherlands), pp. 3-10, 1998.

[12] . Domenico Lepori. "Cours coordonné de sénologie : Imagerie du sein". Service de radiodiagnostic et radiologie interventionnelle, Faculté de Médecine 4ème année, Université de Lausanne, Janvier 2002

[13] . C. Vachier, S. Muller and S. Bothorel. "Morphological analysis of mammographic images for the detection of breast opacities". In Proceeding of ADEBIO, April 1997.

[14] . J.P. Frederic Richard. "Elastic image fitting models and application to bilateral and temporal pairs of mammograms". Doctoral thesis, University René Descartes Paris V, December 2000.

[15] . R. Winn Hardin. "Computer-aided diagnosis becomes a reality in mammography". From OE Reports Number 186, SPIE Newsroom, 31 June 1999.

[16] . Ayodele, T. Introduction to Machine Learning. 02 2010. ISBN 978-953-307-034-6. doi: 10.5772/9394.

[17] . Jatana, V. Machine Learning For Beginners. 10 2018.

[18] . Kouro, K., P.Exarchos, T., P.Exarchos, K., V.Karamouzis, M., and I.Fotiadis, D. Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal,13:8–17, 2015.

[19] . Nargesian, F., Samulowitz, H., Khurana, U., Khalil, E., and Turaga, D. Learning feature engineering for classification. pages 2529–2535, 08 2017. doi:

[20] . Deng, L. Deep Learning: Methods and Applications. Foundations and Trends R in Signal Processing,7(3-4):197–387, 2014. ISSN 1932-8346. doi: 10.1561/2000000039. URL

[21] . Manickam, M., Mohanapriya, M., Kale, S., Uday, M., Kulkarni, P., Khandagale, Y., and Patil, S.Research study on applications of artificial neural networks and e-learning personalization. International Journal of Civil Engineering and Technology, 8:1422–1432, 08 2017.

[22] . Krishna, M., Neelima, M., Mane, H., and Matcha, V. Image classification using deep learning. International Journal of Engineering and Technology, 7:614, 03 2018. doi: 10.14419/ijet.v7i2.7.10892.

[23] . Ragab, D. A., Sharkas, M., Marshall, S., and Ren, J. Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ, 01 2019. doi: https://doi.org/10.7717/peerj. 6201.

[24] . Bre, F., Gimenez, J., and Fachinotti, V. Prediction of wind pressure coefficients on building surfaces using artificial neural networks. Energy and Buildings, 158, 11 2017. doi: 10.1016/j.enbuild.2017.11.045.

[25] . O'Shea, K. and Nash, R. An introduction to Convolutional Neural Networks. CoRR, abs/1511.08458, 2015. URL http://arxiv.org/abs/1511.08458.

[26] . Salehinejad, H., Baarbe, J., Sankar, S., Barfett, J., Colak, E., and Valaee, S. Recent advances in Recurrent Neural Networks. CoRR, abs/1801.01078, 2018. URL http://arxiv.org/abs/1801.01078.

[27] . Jones, T. Deep learning architectures: The rise of artificial intelligence. URL https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures/.

[28] . Gibson, A. and Patterson, J. Deep Learning: Major Architectures of Deep Networks. O'Reilly Media, Inc, 2017.

[29] . Adrian Rosebrock Deep Learning for Computer Vision with Python 2017.

[30] . Somayeh Mahmoudi IGARSS 2018 ,IEEE International Geoscience and Remote Sensing Symposium.

[31] . Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S. Activation functions: Comparison of trends in practice and research for deep learning. CoRR, abs/1811.03378, 2018. URL http://arxiv.org/abs/1811.03378.

[32] . Nusrat, I. and Jang, S.-B. A Comparison of Regularization Techniques in Deep Neural Networks. Symmetry, 10(11):648, Nov 2018. ISSN 2073-8994. doi: 10.3390/sym10110648. URL http://www.mdpi.com/2073-8994/10/11/648.

[33] . Krogh, A. and Hertz, J. A. A simple weight decay can improve generalization. In Moody, J. E., Hanson, S. J., and Lippmann, R. P., editors, Advances in Neural Information Processing Systems 4, pages 950–957. Morgan-Kaufmann, 1992. URL http://papers.nips.cc/paper/563-a-simple-weight-decay-can- improve-generalization.pdf.

[34] . Szegedy, C., Ioffe, S., and Vanhoucke, V. Inception-v4, Inception-resnet and the impact of residual connections on learning. CoRR, abs/1602.07261, 2016. URL http://arxiv.org/abs/1602.07261.

[35] . Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15:1929 – 1958,2014.

Webography

[36] . World Health Organization. Cancer. Wikipedia, the Free Encyclopedia, http://www.who.in/cancer/en/, Accessed February 2022.

[37] Rohith Gandhi. R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms. 2018. url: https://towardsdatascience.com/r-cnn-fast-r-cnn- faster-r-cnn-yolo-object-detection-algorithms-36d53571365e, Accessed February 2022.

[38] . Karn, U. A quick introduction to neural networks – the data science blog, Accessed December 2021.URL https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/.

[39] . Mesin, B. P. Student notes: Convolutional Neural Networks (CNN) Introduction, Accessed April 2022. Student Notes: Convolutional Neural Networks (CNN) Introduction.

[40] . Araujo, L. Pooling layer: Artificial intelligence, Accessed May 2022. URL https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/.