



REPUBLIC OF TÜRKİYE
ALTINBAŞ UNIVERSITY
Institute of Graduate Studies
Information Technology

**PREDICTION BREAST CANCER IN FINE NEEDLE
ASPIRATION IMAGES USING MACHINE
LEARNING**

Layth Adnan Majeed ALABDLI

Master's Thesis

Supervisor

Prof. Dr. Osman UCAN

Istanbul, 2022

PREDICTION BREAST CANCER IN FINE NEEDLE ASPIRATION IMAGES USING MACHINE LEARNING

Layth Adnan Majeed ALABDLI

Information Technology

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titled PREDICTION BREAST CANCER IN FINE NEEDLE ASPIRATION IMAGES USING MACHINE LEARNING prepared by LAYTH ADNAN MAJEED ALABDLI and submitted on 20/12/2022 has been **accepted unanimously** for the degree of Master of Science in Information Technology.

Prof. Dr. Osman Nuri UCAN

the Supervisor

Thesis Defense Committee Members:

Prof. Dr. Osman Nuri UCAN

Department of Electrical-
Electronics Engineering ,

Altınbaş University

Asst.Prof.Dr.Oğuz ATA

Department of Software
Engineering ,

Altınbaş University

Assoc.Prof. Dr. Adil Deniz DURU

Department of Sports and
Health Sciences,

Marmara University

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis.

Submission date of the thesis to Institute of Graduate Studies: ____/____/____

I hereby declare that all information and data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Layth Adnan Majeed ALABDLI

ACKNOWLEDGEMENTS

To Prof.Dr. Osman UCAN, one of the most significant instructors in my life; it was from him that I learned to draw my first letter, acquire the talent of pronouncing Arabic, create my first complete sentence, and apply the first of many rules. To my colleagues, whose unwavering support I can always count on, and who provided me all the tools I need to complete my thesis without a second thought.



ABSTRACT

PREDICTION BREAST CANCER IN FINE NEEDLE ASPIRATION IMAGES USING MACHINE LEARNING

ALABDLI, Layth Adnan Majeed

M.Sc. Information Technology, Altınbaş University,

Supervisor: Prof. Dr. Osman UCAN

Date: 12/2022

Pages: 85

Breast cancer, with an estimated 1.5 million new cases per year, is a critical worldwide health problem owing to the disease's high fatality rate. It is also the most common cancer among women, accounting for 16% of all cancer cases in this demographic globally. Mammography is the most successful way for detecting breast cancer in its early stages, allowing for early intervention. Breast cancer may be detected in its early stages using ultrasound, magnetic resonance imaging, and computational tomography, albeit these methods are not as good in investigating and diagnosing these abnormalities as other ways. This work aims to develop a methodology that is capable of identifying and classifying cancerous tumors using digital mammography images. In this way, the diagnostic work of these images, carried out by radiologist specialists, will be supported by the system to be developed through the methodology presented in this work. Therefore, with the proposed methodology, it was expected an increase in the agility and effectiveness of the diagnosis and, thus, to increase the chances of cure of the patients, since a rapid and effective diagnosis is of fundamental importance for the treatment of the disease.

Keywords: CAD, CNN, ML, DL, AI.

TABLE OF CONTENTS

	<u>Pages</u>
ABSTRACT.....	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
ABBREVIATIONS.....	xi
1. INTRODUCTION.....	1
1.1 BACKGROUND.....	1
1.2 PROBLEM STATEMENT	6
1.3 OBJECTIVES	9
1.4 CONTRIBUTION.....	9
1.5 THESIS STRUCTURE.....	10
2. LITERATURE REVIEW.....	11
3. MATERIALS AND METHODS	24
3.1 BREAST CANCER	24
3.2 CUMPUTER AIDED DESIGN (CAD).....	26
3.2.1 MRI	29
3.2.2 Fine Needle Aspiration (FNA).....	32
3.2.3 Cat Scan for Breast Cancer	34
3.3 IMAGE SEGMENTATION	36
3.3.1 K-means Algorithm.....	38
3.3.2 Fuzzy C-Means FCM	42
3.4 DEEP LEARNING	47
3.4.1 CNN Algorithm.....	51

4. PROPOSED METHOD.....	56
4.1 PREPROCESSING THE IMAGES	56
4.2 SEGMENTATION	56
4.3 DEEP LEARNING DETECTION	63
4.3.1 Creating the Dataset	65
4.3.2 Data Validation.....	66
4.3.3 Image Classification	66
5. CONCLUSION.....	69
REFERENCES.....	71

LIST OF TABLES

	<u>Pages</u>
Table 2.1: Summary of the literature review	17
Table 4.1: Comparing SVM algorithm performance to other algorithms	67
Table 4.2: Time consumption for the execution by the SVM and NB algorithms.	67
Table 4.3: Confusion matrix of true positive labels	68

LIST OF FIGURES

	<u>Pages</u>
Figure 1.1: Classification Image Processing.....	3
Figure 1.2: Breast Cancer Rate	7
Figure 3.1: Screening for Breast Cancer	26
Figure 3.2: overall view of the device relative to an MRI	30
Figure 3.3: Survival rates for breast cancer	35
Figure 3.4: K-means algorithm the clustering	42
Figure 3.5: Illustration of The Network Architecture of Vgg-19 Model.....	53
Figure 4.1: Choosing region of interest ROI in the pre-processing phase.....	56
Figure 4.2: Lesion detection under the microscope	62
Figure 4.3: Segmentation of the lesions by the proposed algorithm.....	63
Figure 4.4: Breast FNA images dataset from the university of Wisconsin	65

ABBREVIATIONS

AI	:	Artificial Intelligence
CAD	:	Compute Aided Design
DDSM	:	Digital Database for Screening Mammography
DL	:	Deep Learning
CNN	:	Convolutional Neural Networks
EMV	:	Effective Magnetization Vector

1. INTRODUCTION

1.1 BACKGROUND

Breast cancer, with an estimated 1.5 million new cases per year, is a critical worldwide health problem owing to the disease's high fatality rate. It is also the most common cancer among women, accounting for 16% of all cancer cases in this demographic globally. Mammography is the most successful way for detecting breast cancer in its early stages, allowing for early intervention. Breast cancer may be detected in its early stages using ultrasound, magnetic resonance imaging, and computational tomography, albeit these methods are not as good in investigating and diagnosing these abnormalities as other ways. The presence of a few distinct types of lesions distinguishes breast cancer symptoms.

CT scans may detect malignant or noncancerous masses, which are big tumors with extremely low contrast images when compared to normal tissue. The formation of microscopic hydroxyapatite or calcium oxalate crystals in breast tissue, known as micro-calcifications, is another symptom that the breast is unwell or on the edge of disease. You will not see them unless there is a lot of noise in the picture, the contrast is low, or the spatial resolution of your equipment is inadequate. Furthermore, malignant infiltration in the galactophores ducts may result in problems. When it comes to mammograms, it's not only about the size and contrast of the breast; it's also about the radiologist's competence and the circumstances in which he or she analyses the picture.

To improve the sensitivity of mammographic tests, a rising number of computer-aided design (CAD) solutions are being created. These technologies may indicate mammograms as abnormal, assisting human professionals in making evaluations. The advantages of these systems have become increasingly evident as interest in artificial intelligence in medical imaging has grown. There has been an explosion of research into machine learning (ML) and a subset of this known as "deep learning" in the recent decade (DL).

This has been made feasible by advancements in processing power, storage capacity, and data set availability. CADs were originally composed of four phases: picture pre-processing, lesion segmentation, characteristic extraction, and classification, all of which occurred in four steps.

Prior to the introduction of machine learning (ML), classification techniques based on ML were the norm. The research, which used the MIAS database, is remarkable, with the top-performing model being a support vector machine (SVM) with the best accuracy of 84.57 percent, the highest sensitivity of 0.8261, and the highest specificity of 0.8889.

An LBP-based Bayesian classifier with a 75% efficiency rate gives the best results when used to extract feature information from mammography pictures in the MIAS database. Convolutional Neural Networks (CNN) are a kind of deep learning (DL) that is used in computer vision (CNN). Hidden layers from convolution and pooling layers have a unique connection design that makes them insensitive to displacement, scaling, and rotation of the input pictures, which explains their success. This implies that CNNs can tolerate minor differences in the pictures they receive. The fact that characteristics are extracted from inside the network is an important component of this method. In reality, this strategy has helped CAD systems in a variety of ways. The DL has fundamentally revolutionized computer-aided image analysis, setting it apart from more typical CAD approaches. According to the research, a CNN was used to classify around 29,000 mammograms from five different institutions.

The final results demonstrated a sensitivity of 76.1 percent, a specificity of 88.6 percent, and an AUC of 0.906 percent. We presented a CNN with 96 percent specificity and 98 percent specificity for breast cancer diagnosis for the Digital Database for Screening Mammography (DDSM). We were able to manufacture three false negatives and one false positive using our proposed technique. A CAD system was built using the Faster R-CNN object categorization platform. They finished in second place overall in the Digital Mammography DREAM Challenge competition based on their network design, with an AUC of 0.85. The DDSM database's AUC was 0.92, which is great. Using a CNN-based CAD system, it was possible to get an AUC of 0.98 in the DDSM database, 0.97 in the IN-breast database, 0.96 in the BCDR database, and 0.99 in the MIAS database. CNNs attain a high level of discriminating strength in classification because they automatically learn the most significant characteristics from the pictures they are taught to identify.

A CNN's purpose is to discover the kernel coefficient that minimizes classification error to the minimum amount possible. A huge number of annotated photographs are necessary for training techniques to work, which is rare in the medical field. Furthermore, building a CNN from the

ground up takes a long time and is computationally expensive. Using a short dataset to train a CNN, on the other hand, may result in the algorithm being overfitted, which is undesirable. These flaws may be remedied by using a method known as "transfer learning." When there is a limited amount of data available, this method has been demonstrated to be beneficial in the medical field. The ImageNet dataset, which contains over 1.2 million images classified into 1000 categories and can be obtained here, is a common technique for training CNN. The weights of the convolutional kernels may then be fine-tuned depending on the freshly obtained pictures.

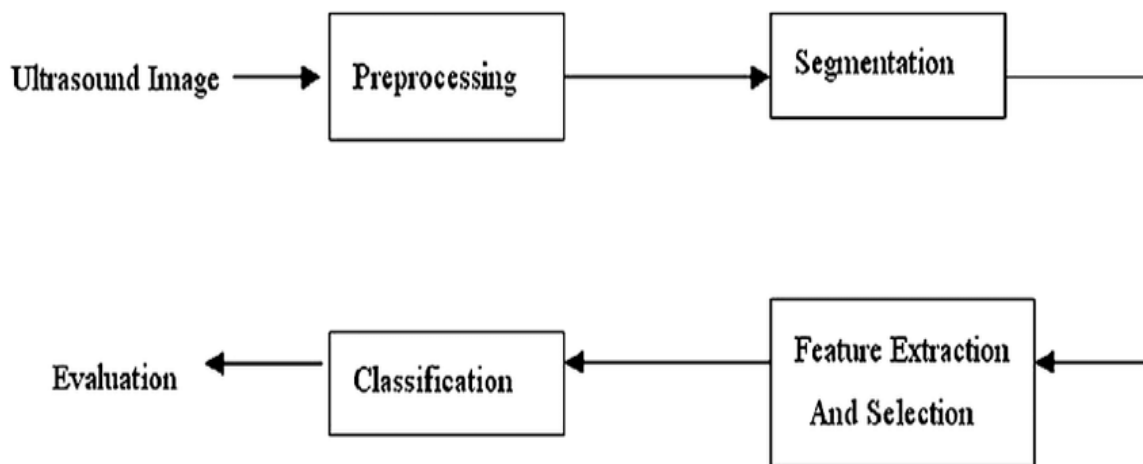


Figure 1.1: Classification Image Processing

Breast cancer represents a huge risk to the health of individuals all over the world because of the high death rate that is associated with the condition. There are around 1.5 million new cases of the disease that are diagnosed each year, according to estimates. In addition to this, it is the type of cancer that occurs most frequently in females and is responsible for 16% of all cases of cancer that are diagnosed in this demographic group all over the world. Mammography is now the most effective tool that is available for diagnosing breast cancer in its earliest stages. This paves the way for therapy to be administered in a more timely manner. However, these approaches are not as efficient as other ways when it comes to investigating and identifying these anomalies. Some of the technologies that can be used to detect breast cancer in its early stages include computational tomography, magnetic resonance imaging, and ultrasound. Other methods include the following: Lesions that are specific to breast cancer can help differentiate the symptoms of this disease from those of other diseases, but only if there are a certain number of them.

CT scans have the potential to detect cancerous and noncancerous masses, which are big tumors that have extremely low contrast pictures in comparison to normal tissue. Masses can be seen in both cancerous and noncancerous forms. There are two types of tumors: malignant and noncancerous masses. The growth of micro-calcifications, also known as microscopic crystals of hydroxyapatite or calcium oxalate, is another indicator that the breast is unhealthy or on the edge of disease. These crystals can also be seen under a microscope. These microscopic calcifications are frequently discovered in breast tissue. You won't be able to notice them if the image does not have a lot of background noise, the contrast is not strong enough, or your camera does not have a high enough spatial resolution. In addition to this, there is a possibility that problems will occur as a result of malignant infiltration in the ducts of the galactophores. When it comes to mammograms, it is not only about the size and contrast of the breast; it is also about the radiologist's competence and the circumstances in which he or she analyzes the picture. Other factors include the size of the breast and the contrast between the breast and the surrounding tissue. The size of the breast as well as its contrast are both crucial variables.

In order to improve the diagnostic accuracy of mammograms, an increasing variety of computer-aided design (CAD) tools are being developed. These tools are helping to make mammograms more sensitive. If these technologies were applied, abnormalities on mammograms might become more apparent, which would be helpful to medical professionals in their evaluations. The value of these systems has been more apparent as there has been a rise in interest in the use of artificial intelligence in medical imaging. In the last ten years, there has been a precipitous rise in the quantity of research into machine learning (ML), particularly a subsection of this topic known as "deep learning." This rise in research has been driven by a growing interest in developing artificial intelligence (DL).

The accessibility of data sets, advancements in computational power, and expansions in storage capacity have all played a part in making the prospect of this situation a reality. In the beginning, CADs were made up of four stages, which were the picture pre-processing stage, the lesion segmentation stage, the characteristic extraction stage, and the classification stage. Separate actions were required for each of these phases to be completed. The employment of classification methods that were predicated on machine learning was a prevalent practice before the emergence

of machine learning (ML). The research is of the highest quality; it was carried out using the MIAS database. The support vector machine (SVM) turned out to be the model that performed the best overall. It had an accuracy of 84.57 percent, the highest sensitivity of 0.8261, and the highest specificity of 0.8889 respectively.

An LBP-based Bayesian classifier with an efficiency rate of 75% gives the best possible results when used to the process of extracting feature information from mammography photographs stored within the MIAS database. This particular method produces the best possible outcomes. Through the implementation of convolutional neural networks (CNN), the discipline of computer vision makes use of deep learning (DL) (CNN). Because of their one-of-a-kind connection architecture, which makes them insensitive to the translation, scaling, and rotation of the input pictures, hidden layers that are derived from convolution and pooling layers are highly effective. This may be attributed to the fact that these layers are effective. This is the most important factor in determining how accurately they can provide outcomes. This demonstrates that CNNs are able to accept relatively little shifts in the images that they receive. The technique of extracting features from within the network itself is a crucial component of this strategy. In point of fact, the application of this technology has proven to be advantageous for CAD systems in a variety of various ways since its introduction. Computer-assisted image analysis has undergone a significant transformation as a result of the DL, setting it apart from other, more conventional CAD methods in the process. According to the results of the research, a CNN was employed in order to classify around 29,000 mammograms that were obtained from five distinct medical facilities.

According to the findings that were compiled in the end, the level of sensitivity was found to be 76.1 percent, the level of specificity was found to be 88.6 percent, and the area under the curve was found to be 0.906 percent. For the purpose of the Digital Database for Screening Mammography, we demonstrated a CNN that has a specificity of 96 percent for the diagnosis of breast cancer and a specificity of 98 percent respectively (DDSM). With the help of the procedure that we suggested, we were able to generate three false negatives in addition to one false positive. The creation of a computer-aided design (CAD) system made use of a platform known as Faster R-CNN, which is an object categorization method. They finished in second place overall in the Digital Mammography DREAM Challenge competition thanks to their AUC

score of 0.85, which allowed them achieve this position. This occurred as a result of the highly efficient network design that they had implemented. The area under the receiver operating curve (AUC) of the DDSM database came in at 0.92, which is an outstanding figure. When using a CNN-based CAD system, it was able to get an AUC of 0.98 in the DDSM database, 0.97 in the IN-breast database, 0.96 in the BCDR database, and 0.99 in the MIAS database. CNNs are able to achieve a high degree of discriminating strength in categorization due to the fact that they intuitively learn the most important characteristics from the photos they are trained to recognize. The ability to acquire such a high level of discriminating strength is made possible by CNNs.

The objective of a convolutional neural network, often known as a CNN, is to locate the kernel coefficient that produces the most accurate classification with the least amount of error possible. The training approaches need an exceptionally large quantity of labeled photographs in order to be effective, since captioned photos are often not required in the medical field. In addition, the process of creating a CNN from the ground up is one that is both time-consuming and expensive from a computer point of view. When training a CNN, on the other hand, with a limited dataset, the probability of the algorithm becoming overfitted is raised, which is something that should be avoided whenever possible. These weaknesses can be remedied by utilizing a method known as "transfer learning," which entails gaining knowledge from one context and applying it to another. It has been demonstrated that this tactic is effective in the realm of medicine, even when there is only a limited amount of data to work with. The ImageNet dataset is a method that is often used for training CNN. This dataset consists of around 1.2 million photographs that have been sorted into 1000 different categories, and it can be accessed from this location. After that, the weights of the convolutional kernels can be fine-tuned based on the most current photographs that were obtained from the source.

1.2 PROBLEM STATEMENT

Early identification and treatment are critical to reducing breast cancer deaths. Early detection and treatment of breast cancer are critical to optimal treatment. Breast cancer screening is one of the most effective techniques to detect the illness early.

Mammograms are specialized radiograms used by radiologists to detect and diagnose breast tissue irregularities. Prior to the widespread use of mammography for screening, most breast

cancers were only palpable (feeling an object in our body). Breast mammography uses less radiation than other imaging procedures. If the breast is abnormal owing to cancer, fatty cells, or cysts, digital mammography may detect a lump. Micro-calcifications, a kind of calcification, may be seen in digitized mammo-images.

Diagnostic mammography may identify up to 75% of breast cancers before the patient is aware. As a consequence, mammography is vital in lowering the incidence and mortality of breast cancer. In mammography, CAD helps radiologists find cancers that could otherwise go unnoticed.

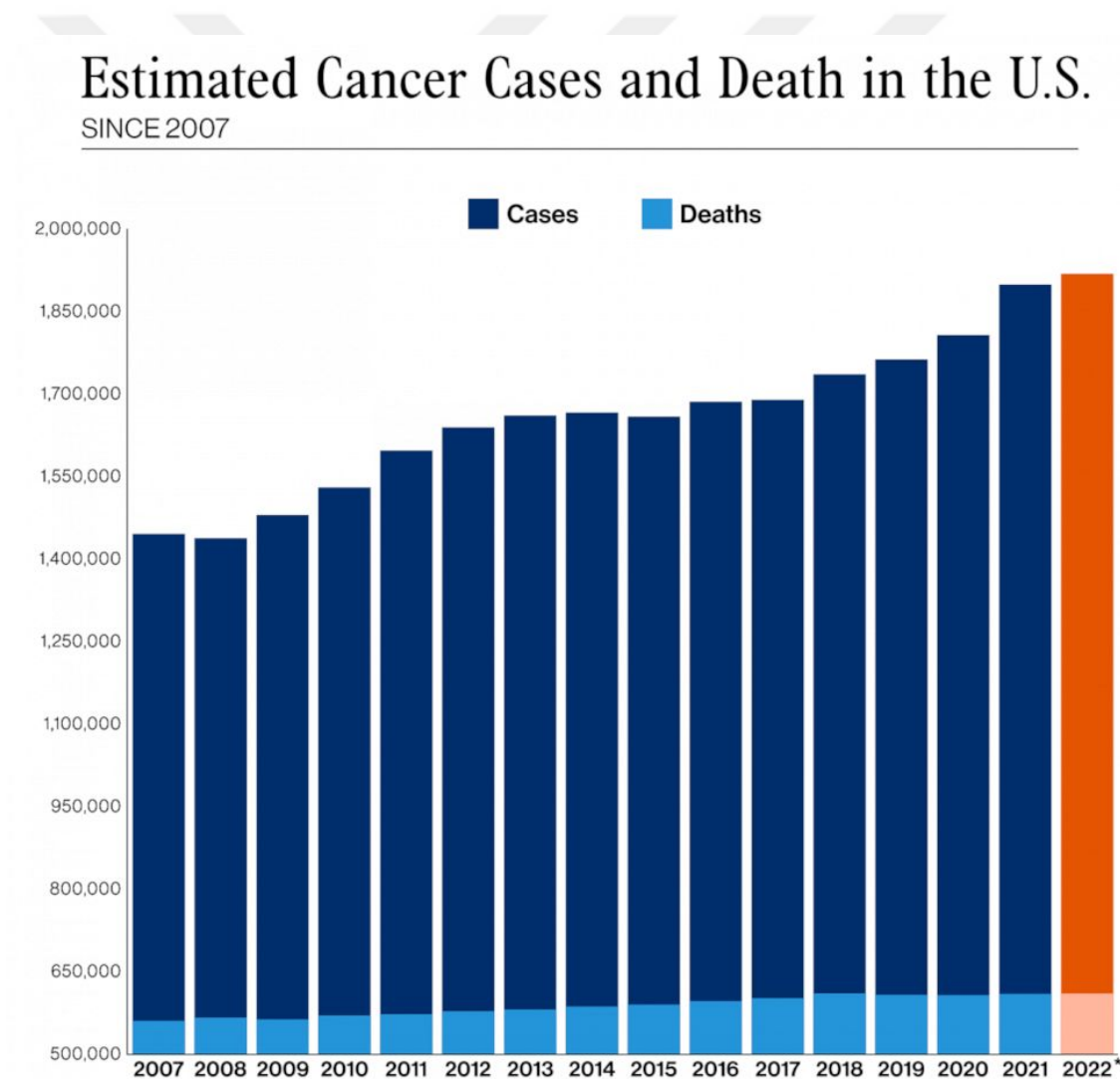


Figure 1.2: Breast Cancer Rate

In the early 1990s, radiologists used CAD software to improve the diagnostic and prediction accuracy of mammography screening approaches. CAD analyses the digitalized mammo images and highlights any flaws. In 2000 and 2001, CAD was used to analyze 5% of all screening mammograms in the United States. Following an increase in CAD funding by the CMS in 2002, CAD was used to evaluate almost 74% of all screen mammography performed on Medicare beneficiaries in 2008. Many women worldwide will be diagnosed with breast cancer at some point in their life.

The sooner this cancer is discovered and treated, the more lives may be spared. Additionally, mammography has long been acknowledged as the most common and straightforward means of identifying cancer in its earliest stages. Radiologists may accurately predict the findings of mammograms with a rate of more than 90% in certain circumstances. Despite this, radiologists may miss up to 10% of instances. Women account for around 15% of all breast cancer patients (Erickson, 2005). Re-examining the mammography pictures and doing a second check may assist to lessen the likelihood of having a false positive result in the future. To ensure accuracy, the same mammography must be evaluated by two radiologists at separate times.

When comparing correct detection accuracy to single-checking, research indicates that double-checking boosts accurate detection accuracy by 15%. While this is a time-consuming and expensive process, it is important. CAD, or computer-assisted diagnosis, has the potential to result in considerable cost reductions (CAD). In other words, computers may one day be used to diagnose medical concerns. This diagnosis is made using a range of approaches, including database management, image processing, machine learning, and data analysis.

How may these strategies be used to minimize the frequency of false positives in breast cancer screening using mammography? In this context, computer vision and image processing have been presented as viable answers. Image processing and machine learning techniques may be applied to improve the detection accuracy of mammography pictures. After attempting to identify masses in mammograms using a variety of image processing techniques, picture segmentation, edge detection, and region of interest (ROI) are used to identify masses in mammograms. After the images have been analyzed, feature retrieval is conducted to aid comparison and detection. Despite this, these systems' accuracy continues to be a cause of concern.

1.3 OBJECTIVES

This work aims to develop a methodology that is capable of identifying and classifying cancerous tumors using digital mammography images. In this way, the diagnostic work of these images, carried out by radiologist specialists, will be supported by the system to be developed through the methodology presented in this work. Therefore, with the proposed methodology, it was expected an increase in the agility and effectiveness of the diagnosis and, thus, to increase the chances of cure of the patients, since a rapid and effective diagnosis is of fundamental importance for the treatment of the disease.

Thus, one of the specific objectives of this work are:

- a. Develop an effective methodology for the tumor segmentation stage, which is an important stage, as its result can interfere in the following stages, since tumors with poorly segmented areas can lose their shape characteristics, making classification stages difficult.
- b. Evaluate the methodology in the tasks of identifying and classifying tumors in the regions extracted from the segmentation step.
- c. There are different methods to be analyzed, such as the possibility of performing a classification into three classes (normal, benign tumor or malignant tumor) or performing the classification in two stages: between normal and cancerous and then between benign and malignant.
- d. It will also be evaluated how the features of each image will be extracted, and how it will be possible to use a CNN network, which has the ability to unite the features extraction and classification steps.

1.4 CONTRIBUTION

In this project, we devised and built a system that can identify breast cancer masses in pictures of fine needle aspiration. After a clustering and feature extraction approach like CNN, we employ a classification process to determine cancer. Survival rates for women with breast cancer are significantly improved when the disease is caught early. Basically, there are two kinds of people:

- i. Treatment is made easier when people with symptoms have access to early diagnosis, symptoms knowledge, and a family medical history. Even in underdeveloped nations with little financial means, early diagnosis like this is critical.
- ii. Detecting breast cancer in women who are asymptomatic using a series of tests called "routine controls."

1.5 THESIS STRUCTURE

The thesis is structured as follows: Section 2 is where we review some of the previous work and implementation on smart grids. Section 3 is where we give a background about all the components. Section 4 is where we explain our model in details and implementation of that model in details, Section 5 is where we simulate our model and record the results obtained. Section 6 is where we conclude our work and put a future scope into perspective.

2. LITERATURE REVIEW

In thick breast tissue, macro calcification is difficult to detect on a mammography because it tends to show up as white pixels. Dense breast tissue has a greater rate of false positives. Masses and micro calcifications are two of the most common signs of malignancy. Mass detection is more difficult than micro calcification detection. Because of the wide range of size and form variations and the low picture contrast they often display.

Classification & pattern recognition systems are increasingly relying on grouping frameworks, particularly in cancer detection. For medical diagnostics, machine learning plays a crucial role in evaluation and decision-making. In certain cases, an intelligent classification system may assist the clinician uncover symptoms that may not be achievable using standard methods⁸. Feature alignment for segmentation and classification is essential for any visual analytic applications. In the field of pattern recognition, texture and statistical aspects are most useful for locating this alignment. In terms of cost and convenience, mammography is the best option. Multiple approaches are used to examine the mammogram picture, including locating boundaries, smoothing borders, and identifying structures and forms in the matrix. Finding the tissue size distribution in an image without specifically segmenting each object⁵ is now possible.

A variety of classification problems are used to the digital mammography picture in order to diagnose breast cancer. Breast cancer is diagnosed using a conventional approach that extracts a variety of characteristics. To avoid affecting the categorization issue, these characteristics are determined from the breast's sensitive portion. Its Maximum Likelihood Estimation is used to determine the tumor's size (MLE).

The saved database is used for all feature extraction algorithms. This research focuses on different image processing techniques that may be used to identify malignant cells. To get a deeper knowledge of tumors, further research is needed to extract more characteristics and discover patterns. It is possible to distinguish between benign or malignant mammographic masses using the texture analysis approach. "A diversity of modalities" may be critical when it comes to breast imaging. Mammography is the only technique that has been proved throughout time in the area of breast imaging.

Jaya Sherma, R. P. Tewari, and J. K. Rai published "Identification of Preprocessing Technique for Improving Mammogram Images" in 2014. Choosing the optimal preprocessing approach for a batch of mammography pictures based on the images' greatest peak signal-to-noise ratio. Ketan Sharma and Bobbin Preet's 2016 paper, "Classification of Mammogram Images by Using CNN Classifier," is titled "Classification of Mammogram Images." CNN is the acronym for the computer-aided design system that the authors of this study suggest. CNN was compared to the Logistic Regression technique, which was also evaluated in this study.

In 2017, Nathan Jacobs, Jinze Liu, and Erik Y. Han co-authored an article titled According to the authors of the study "Full Mammogram Image Classification Using Convolutional Neural Networks," the project's objective is to build and enhance machine learning models for mammography whole image classification. Additionally, it contains a report summarizing early results. When used in conjunction with a CNN, data augmentation and transfer learning techniques were shown to be the most effective methods for boosting classification performance across the seven distinct architectures examined in this research.

The article "Preprocessing Filters for Mammogram Images" by Kshema and Jayesh is accessible online. M. George (George M., 2017). Preprocessing is the single most critical step in improving the quality of mammography images. Additional mammography image processing and analysis are not feasible until the images have been reoriented. Enhance and smooth the picture by adding filters to it to increase the image's clarity, remove noise, and maintain the image's original boundaries. The Gaussian filter, wiener filter, or mean filter may be used to enhance the reconstruction of a speckled mammography picture. The mean filter, median filter, and adaptive median filter all outperform the others in this category when dealing with salt and pepper noise. The completion and publication of a study titled "Mass Classification in Mammograms Using Neural Networks" was announced in 2017.

It is critical to identify the most advantageous properties that impact a classifier's performance, since there are several factors to consider. Additionally, it is critical to remember that the design of the neural network has a direct impact on the classification process's success or failure. Any input feature fed into the classifier results in a 10% performance increase over the architecture with three hidden nodes for the design with 100 hidden nodes over time. Mammography may miss dense breast tissue because macro calcification appears as white pixels on the image. Dense

breast tissue produces more false positives than thin breast tissue. Masses and micro calcifications are two common cancer indications. Bulk calcification is harder to detect than micro calcification. These images are generally shown in low contrast because to the broad variety of size and shape changes.

Classification and pattern recognition algorithms, notably in cancer diagnosis, are becoming increasingly reliant on grouping frameworks. Machine learning is vital in medical diagnosis since it helps evaluate and decide. Intelligent algorithms may help clinicians discover symptoms that are difficult to detect using traditional approaches. In image processing and analysis, feature alignment is critical for classification and segmentation. In pattern recognition, textural and statistical characteristics help locate an alignment. Screening mammography is the most cost-effective and convenient option. The mammography image is analyzed by establishing boundaries, smoothing borders, and recognizing patterns and shapes in the matrix.

Macro calcification, which usually looks like white pixels on a mammography, could make it harder to notice in dense breast tissue while a mammogram is being performed. Dense breast tissue has a greater propensity to result in a higher rate of false positive breast cancer diagnoses. The development of tumors and the appearance of calcifications on a microscopic scale are two of the most common symptoms that point to the presence of cancer. When compared to the identification of micro calcification, the detection of bulk is a far more difficult task. Because of the wide range of sizes and shapes that can be exhibited by them, in addition to the low visual contrast that they typically portray.

Grouping frameworks are becoming an increasingly significant component of classification and pattern recognition systems, most notably in the context of cancer detection. [C]lassification and pattern recognition systems rely heavily on grouping frameworks. The evaluation and choice-making processes that are involved in medical diagnostics are greatly aided by the application of machine learning. When applied in particular contexts, the utilization of an intelligent classification system may be capable of assisting the physician in the recognition of symptoms that may not be discoverable by employing the standard methods of investigation. 8. The prioritization of feature alignment for the purposes of segmentation and classification is a fundamental requirement for every application that uses visual analytics. When it comes to the identification of patterns, the aspects of texture and statistics that are statistical are the ones that

are the most beneficial in locating this alignment. Mammography is the procedure that is considered to have the greatest number of advantages in terms of cost and convenience of use. In the process of interpreting the mammography picture, a few different approaches are considered and applied. Methods such as defining boundaries, smoothing border edges, and detecting structures and shapes inside the matrix are included in these approaches. It is now possible to identify the size distribution of tissue in an image without first having to manually segment each individual object⁵, which was previously required in order to do so.

In order to arrive at an accurate diagnosis of breast cancer, a variety of different classification issues are applied to the picture obtained from the digital mammogram. Obtaining a variety of distinct characteristics from the patient is a standard step in the process that is used to diagnose breast cancer. These characteristics are determined from the sensitive area of the breast since there is a possibility that this will change the classification difficulty. The estimation of the tumor's size is referred to as its Maximum Likelihood, and it is used in the calculation (MLE).

Every one of the strategies for extracting features makes use of the database that was previously saved. The objective of this study is to evaluate a variety of image processing techniques that have the potential to be applied in the localization of malignant cells. In order to have a more in-depth understanding of malignancies, additional study is required to discover trends and extract additional characteristics. It is possible to successfully identify between benign and malignant mammographic masses by using the approach of texture analysis. When it comes to breast imaging, having "a diversity of modalities" could very well be a component that plays a very significant role. Mammography is the only imaging technique for the breast that has been shown to be successful throughout the course of time. This makes it the gold standard in the field of breast imaging.

The article "Identification of Preprocessing Technique for Improving Mammogram Images" was written by Jaya Sherma, R. P. Tewari, and J. K. Rai, and it was published in 2014. looking at the mammogram images to see which ones have the highest peak signal-to-noise ratio as a means of deciding which preprocessing technique would be best suited for the collection of mammography pictures. "Classification of Mammogram Images" is the title of the research article that was published in 2016 by Ketan Sharma and Bobbin Preet under the title "Classification of Mammogram Images by Using CNN Classifier." The authors of the article are Bobbin Preet and

Ketan Sharma. The abbreviation CNN refers to the computer-aided design (CAD) system that was used in this research and is recommended by the researchers who conducted it. Alongside the Logistic Regression approach, which was also researched over the course of this study, CNN was analyzed as part of the research process.

An article that was written collaboratively by Nathan Jacobs, Jinze Liu, and Erik Y. Han and published in 2017 under the title According to the authors of the article titled "Full Mammogram Image Classification Using Convolutional Neural Networks," the purpose of the project is to create and improve machine learning models for mammography full image classification. This objective was stated in the study. In addition to that, it consists of a report that provides a summary of the preliminary findings. Data augmentation and transfer learning techniques were shown to be the most effective methods for boosting classification performance across all seven different architectures that were examined in this research. These techniques were shown to be the most effective methods for boosting classification performance. It was found that utilizing these strategies in conjunction with a CNN resulted in the greatest amount of success.

The article "Preprocessing Filters for Mammogram Images" that was written by Kshema and Jayesh is available in an online version, which readers are able to view. M. George (George M., 2017). The preprocessing stage of mammography is the single most critical one, hence improving the image quality is best accomplished by focusing on this aspect of the procedure. After the images have been rotated, any additional mammography image processing or analysis will no longer be available. Until then, nothing will be able to be done. By applying various filters to the image, it is possible to enhance it and make it more balanced. These filters have the ability to make the image sharper, remove noise from the image, and preserve the original borders of the image. The reconstruction of a speckled mammography picture can be improved with the assistance of the mean filter, the Wiener filter, or the Gaussian filter. All of these filters are named after statistical distributions. The mean filter, the median filter, and the adaptive median filter all perform better than the other available choices in this area when it comes to dealing with salt and pepper noise. In 2017, it was stated that a study with the working title "Mass Classification in Mammograms Using Neural Networks" had been completed and would soon be published. The study had been conducted to classify masses in mammograms using neural networks.

It is necessary to establish which of a classifier's qualities have the most beneficial influence on its performance because there are many various aspects to take into consideration. This is because there are many different things to take into account. In addition, it is crucial to keep in mind that the topology of the neural network has a direct influence on the success or failure of the classification process. This is an important point to keep in mind since it is essential to keep in mind. It is impossible to place sufficient emphasis on this particular subject. Any feature that was input into the classifier resulted in a 10% improvement in performance over the architecture that had three hidden nodes for the design that had 100 hidden nodes over time. This improvement was shown over the course of the experiment. Due to the fact that macro calcification shows up as white pixels on the mammogram, there is a possibility that dense breast tissue will be missed by the mammography procedure. It has been discovered that dense breast tissue produces a greater number of false positive results when compared to thin breast tissue. Both the existence of masses and tiny calcifications are frequently seen as indicators of the presence of cancer. When compared to micro calcification, determining the presence of bulk calcification is a significantly more challenging endeavor. These photographs virtually never have a high level of contrast because of the wide variety of scaling and contouring options available to the user.

Grouping frameworks are becoming an increasingly significant component of classification and pattern recognition algorithms, particularly those that are utilized in the process of diagnosing cancer. [C]ase in point: Because it assists with analyzing and making decisions, machine learning plays an extremely essential role in the field of medical diagnosis. The application of intelligent algorithms may be of assistance to clinical practitioners for the detection of symptoms that are difficult to detect using more conventional methods. When it comes to the processing and analysis of images, feature alignment is a phase that plays a critical role in both the segmentation and the classification processes. Textural and statistical features can provide a helping hand in the process of pattern recognition when it comes to the position of an alignment. Screening with mammography is the procedure that satisfies both requirements of being the most practical and the most economical option. As part of the analysis process, you will establish boundaries within the mammography image, smooth down its edges, and look for patterns and forms within the matrix.

It can now estimate tissue size distribution inside an image without having to separate discrete items. Various classification difficulties are used to digital mammography images to detect breast cancer. Breast cancer is diagnosed using a conventional approach that extracts certain features. For the sake of categorization, these characteristics are taken from the breast's sensitive zone. The tumor's size is estimated using Maximum Likelihood (MLE). All feature extraction processes start with the database picture.

This study examines many image processing techniques for identifying cancerous cells. More study is needed to understand malignancy and trends. Using texture analysis, it is feasible to identify benign from malignant mammographic masses.

Table 2.1: Summary of the literature review

Author	Year	Findings
Jaya Sherma, R. P. Tewari, and J. K. Rai	2014	Choosing the best preprocessing method for a batch of mammography images based on their peak SNR
Sharma and Bobbin Preet's	2016	"Classification of Mammogram Images by Using CNN Classifier" The study's authors propose the abbreviation CNN for computer-aided design system. This research compares CNN to Logistic Regression.
Nathan Jacobs, Jinze Liu, and Erik Y. Han	2017	Co-wrote a piece named The goal of the research is to construct and improve machine learning models for mammography entire picture categorization, according to the study's authors.

Table 2.1: Summary of the literature review "table continued"

J. C. Tobias Charistian Cahoon, Melanie A.Sutton	2020	These frameworks are increasingly used in classification and pattern recognition algorithms, including cancer detection. Medical diagnosis requires machine learning to examine and decide.
Kui YI, Yue-Hua DING	2019	The mammography picture is evaluated by identifying patterns and structures in the matrix. No longer must it separate distinct objects to estimate tissue size distribution.
Hua, K.-L., et al	2015	Image processing strategies for detecting malignant cells are examined. More research is required to understand malignancy. Texture analysis can distinguish benign from malignant mammographic masses.
Zhang, Y., et al. DeepSplice	2016	These features are extracted from the breast's sensitive zone for classification. Maximum Likelihood (MLE) estimates the tumor size.

Macro calcification, which usually looks like white pixels on a mammography, could make it harder to notice in dense breast tissue while a mammogram is being performed. Dense breast tissue has a greater propensity to result in a higher rate of false positive breast cancer diagnoses. The development of tumors and the appearance of calcifications on a microscopic scale are two of the most common symptoms that point to the presence of cancer. When compared to the identification of micro calcification, the detection of bulk is a far more difficult task. Because of the wide range of sizes and shapes that can be exhibited by them, in addition to the low visual contrast that they typically portray.

Grouping frameworks are becoming an increasingly significant component of classification and pattern recognition systems, most notably in the context of cancer detection. [C]lassification and pattern recognition systems rely heavily on grouping frameworks. The evaluation and choice-making processes that are involved in medical diagnostics are greatly aided by the application of machine learning. An intelligent categorization system may be able to assist the physician in detecting symptoms that may not be accessible through standard methods⁸, in certain conditions. This may be the case when the system is used. [Further citation is required] [Further citation is required] Any application that makes use of visual analytics has an unquestionable need to give feature alignment the utmost priority in order to fulfill the requirements of segmentation and classification. When it comes to the identification of patterns, the aspects of texture and statistics that are statistical are the ones that are the most beneficial in locating this alignment. Mammography is the most financially and practically sound option, and it is also the one that is recommended. In the process of interpreting the mammography picture, a few different approaches are considered and applied. Methods such as defining boundaries, smoothing border edges, and detecting structures and shapes inside the matrix are included in these approaches. It is now possible to identify the size distribution of tissue in an image without first having to manually segment each individual object⁵, which was previously required in order to do so.

In order to arrive at an accurate diagnosis of breast cancer, a variety of different classification issues are applied to the picture obtained from the digital mammogram. Obtaining a variety of distinct characteristics from the patient is a standard step in the process that is used to diagnose breast cancer. These characteristics are determined from the sensitive area of the breast since there is a possibility that this will change the classification difficulty. The estimation of the tumor's size is referred to as its Maximum Likelihood, and it is used in the calculation (MLE).

Every one of the strategies for extracting features makes use of the database that was previously saved. The objective of this study is to evaluate a variety of image processing techniques that have the potential to be applied in the localization of malignant cells. In order to have a more in-depth understanding of malignancies, additional study is required to discover trends and extract additional characteristics. It is possible to successfully identify between benign and malignant mammographic masses by using the approach of texture analysis. When it comes to breast

imaging, having "a diversity of modalities" could very well be a component that plays a very significant role. Mammography is the only imaging technique for the breast that has been shown to be successful throughout the course of time. This makes it the gold standard in the field of breast imaging.

The article "Identification of Preprocessing Technique for Improving Mammogram Images" was written by Jaya Sherma, R. P. Tewari, and J. K. Rai, and it was published in 2014. Looking at the mammogram images to see which ones have the highest peak signal-to-noise ratio as a means of deciding which preprocessing technique would be best suited for the collection of mammography pictures. "Classification of Mammogram Images" is the title of the research article that was published in 2016 by Ketan Sharma and Bobbin Preet under the title "Classification of Mammogram Images by Using CNN Classifier." The authors of the article are Bobbin Preet and Ketan Sharma. The abbreviation CNN refers to the computer-aided design (CAD) system that was used in this research and is recommended by the researchers who conducted it. Alongside the Logistic Regression approach, which was also researched over the course of this study, CNN was analyzed as part of the research process.

An article that was written collaboratively by Nathan Jacobs, Jinze Liu, and Erik Y. Han and published in 2017 under the title "Full Mammogram Image Classification Using Convolutional Neural Networks," the purpose of the project is to create and improve machine learning models for mammography full image classification. This objective was stated in the study. In addition to that, it consists of a report that provides a summary of the preliminary findings. Data augmentation and transfer learning techniques were shown to be the most effective methods for boosting classification performance across all seven different architectures that were examined in this research. These techniques were shown to be the most effective methods for boosting classification performance. It was found that utilizing these strategies in conjunction with a CNN resulted in the greatest amount of success.

The article "Preprocessing Filters for Mammogram Images" that was written by Kshema and Jayesh is available in an online version, which readers are able to view. M. George (George M., 2017). The preprocessing stage of mammography is the single most critical one, hence improving the image quality is best accomplished by focusing on this aspect of the procedure. It is not possible to continue the mammography image processing and analysis until the images have been reoriented. By applying various filters to the image, it is possible to enhance it and make it more balanced. These filters have the ability to make the image sharper, remove noise from the image, and preserve the original borders of the image. The reconstruction of a speckled mammography picture can be improved with the assistance of the mean filter, the Wiener filter, or the Gaussian filter. All of these filters are named after statistical distributions. When it comes to dealing with salt and pepper noise, the mean filter, the median filter, and the adaptive median filter all perform superiorly than the other available choices in this area because of their ability to utilize more sophisticated statistical techniques. In 2017, it was stated that a study with the working title "Mass Classification in Mammograms Using Neural Networks" had been completed and would soon be published. The study had been conducted to classify masses in mammograms using neural networks.

It is necessary to establish which of a classifier's qualities have the most beneficial influence on its performance because there are many various aspects to take into consideration. This is because there are many different things to take into account. In addition, it is crucial to keep in mind that the topology of the neural network has a direct influence on the success or failure of the classification process. This is an important point to keep in mind since it is essential to keep in mind. This is an essential fact that should not be forgotten, so make sure to keep it in mind at all times. Any feature that was input into the classifier resulted in a 10% improvement in performance over the architecture that had three hidden nodes for the design that had 100 hidden nodes over time. This improvement was shown over the course of the experiment. Due to the fact that macro calcification shows up as white pixels on the mammogram, there is a possibility that dense breast tissue will be missed by the mammography procedure. It has been discovered that dense breast tissue produces a greater number of false positive results when compared to thin breast tissue. Both the existence of masses and tiny calcifications are frequently seen as

indicators of the presence of cancer. When compared to micro calcification, determining the presence of bulk calcification is a significantly more challenging endeavor. These photographs virtually never have a high level of contrast because of the wide variety of scaling and contouring options available to the user.

Grouping frameworks are becoming an increasingly significant component of classification and pattern recognition algorithms, particularly those that are utilized in the process of diagnosing cancer. [C]ase in point: Because it assists with analyzing and making decisions, machine learning plays an extremely essential role in the field of medical diagnosis. Clinical practitioners may profit from the use of intelligent algorithms for the discovery of symptoms that are difficult to identify using more conventional diagnostic approaches. These symptoms can be difficult to detect because they are not easily identifiable. When it comes to the processing and analysis of images, feature alignment is a phase that plays a critical role in both the segmentation and the classification processes. The process of finding an alignment is made easier in pattern recognition thanks to the contributions of textural and statistical features. Screening with mammography is the procedure that satisfies both requirements of being the most practical and the most economical option. The process of analyzing a mammogram image consists of several processes, including determining the boundaries of the matrix, smoothing out the edges of the matrix, and locating patterns and structures in the matrix.

It is now capable of estimating the size distribution of tissues contained within an image without first requiring the individual objects to be segmented into their component parts. Images obtained from digital mammography are put through a number of different classification tests in order to arrive at a conclusion on the presence or absence of breast cancer. Traditionally, a method that detects certain features has been used to diagnose breast cancer. This method takes into account a variety of factors. For the purpose of classification, these traits are obtained from a region of the breast known as the sensitive zone. The size of the tumor is estimated using maximum likelihood, which is a statistical method (MLE). All of the procedures for feature extraction will first start using the photo that is stored in the database.

In this study, a wide variety of various methods of image processing are studied to determine whether or not they have the ability to identify cancerous cells. Additional research is required in order to get a complete understanding of malignancy and trends. Texture analysis makes it possible to discern between benign and malignant mammographic masses. This distinction can be made with mammograms.



3. MATERIALS AND METHODS

3.1 BREAST CANCER

In recent years, there has been a surge in understanding of the significance of breast cancer as a threat to the general health of the population. In the vast majority of geographic places across the globe, cancer is the disease that strikes females more frequently than it does males. There were 2.1 million new cases of cancer that were identified across the entire world, and it is projected that 627 thousand people passed away as a direct result of the disease. The most recent data that is currently available reveals that there are. In 2019, it is anticipated that an extra 59,700 cases of breast cancer would be identified in Brazilian women for the first time. With the exception of cancers of the skin that are not melanoma, this proportion accounts for 29.5% of all cancers that are diagnosed in Brazilian women. Melanoma-related cancers of the skin are excluded. In the United States in the year 2016, breast cancer was the underlying cause of death for a total of 16,069 female patients who succumbed to their illness and died. Despite the fact that estimates of survival after five years show an upward trend in industrialized countries, there is still a significant disparity between countries on a worldwide scale. According to Concord-3, the estimations for the five-year survival rate for the period from 2005 to 2009 were anywhere from 75.5 to 78 percent, while the projections for the period from 2010 to 2014 ranged anywhere from 73.9 to 76.5 percent. A number of factors connected to the awareness of the disease as well as the challenges that women face in gaining access to diagnostic technologies as well as adequate and early treatment contribute to the arrival of patients at more advanced stages of breast cancer, which results in a worse prognosis. Breast cancer patients tend to be diagnosed at later stages of the disease as a result of these factors. Patients who present themselves at later stages of the illness have a poorer prognosis, and this is one of the reasons why.

The process of diagnosing breast cancer, as well as both local and systemic treatments for the disease, are all currently going through considerable advancement at a fast pace. The natural history of the disease as well as the molecular properties of the tumors have been studied in greater depth, which has led to a clearer understanding of how this is even possible. In addition to that, the molecular characteristics of the cancers have been taken into consideration. The design of breast cancer preventive measures, with an emphasis on early diagnosis, is an absolute

necessity in this kind of circumstance. When an invasive tumor is found early on in the diagnostic process and treatment is started at that time, the patient's prognosis is much improved and they have a much better chance of making a full recovery. As a direct result of this, a variety of various approaches have been investigated and tested in an effort to improve the precision of cancer diagnosis in its earliest stages. On the other hand, breast cancer is regarded as a disease of dynamic behavior due to its organic individualization and severe tumor heterogeneity, both of which are connected with the existence of known and undiscovered risk factors. This is because breast cancer is a disease that exhibits organic individualization and severe tumor heterogeneity. This is the primary reason why breast cancer is classified as a disease of dynamic behavior. There is a wide variety of histological and molecular subtypes that can be found in both in situ and invasive breast carcinomas. It is much more frequent to have an in situ breast carcinoma than an invasive breast carcinoma. Infiltrating lobular carcinoma is the second most prevalent invasive histologic type after nonspecific infiltrating ductal carcinoma, which accounts for 70 to 80 percent of all breast cancers. Infiltrating lobular carcinoma is the most common type of breast cancer. It is estimated that anywhere between 5 and 15 percent of all breast tumors are caused by it, which places it as the second most prevalent type of breast cancer overall. Breast tumors can also be classified as belonging to other histologic types, which accounts for the remaining 5–80 percent of cases. 2012).

The seventh edition of the TNM – Classification of Malignant Tumors uses clinical anatomical staging to define the extent of a disease based on the size of the tumor, the presence or absence of compromised homolateral axillary lymph nodes, and the presence of disease in locations other than the breast. This staging system was first introduced in the fifth edition of the TNM – Classification of Malignant Tumors. This staging approach was initially presented for the very first time in the fifth edition of the TNM, which is known formally as the Classification of Malignant Tumors. After the staging has been determined in accordance with the TNM, it is feasible to classify them into stages ranging from 0 to IV. In situ breast cancer is referred to as being in Stage 0 of the categorization system, Stage I is the earliest stage, and Stage IV is the most advanced stage of invasive carcinomas. The table that follows includes Stages 0 through IV for your convenience. Following the completion of the TNM staging, the stages can be arranged into ranges beginning with 0 and continuing all the way up to IV.

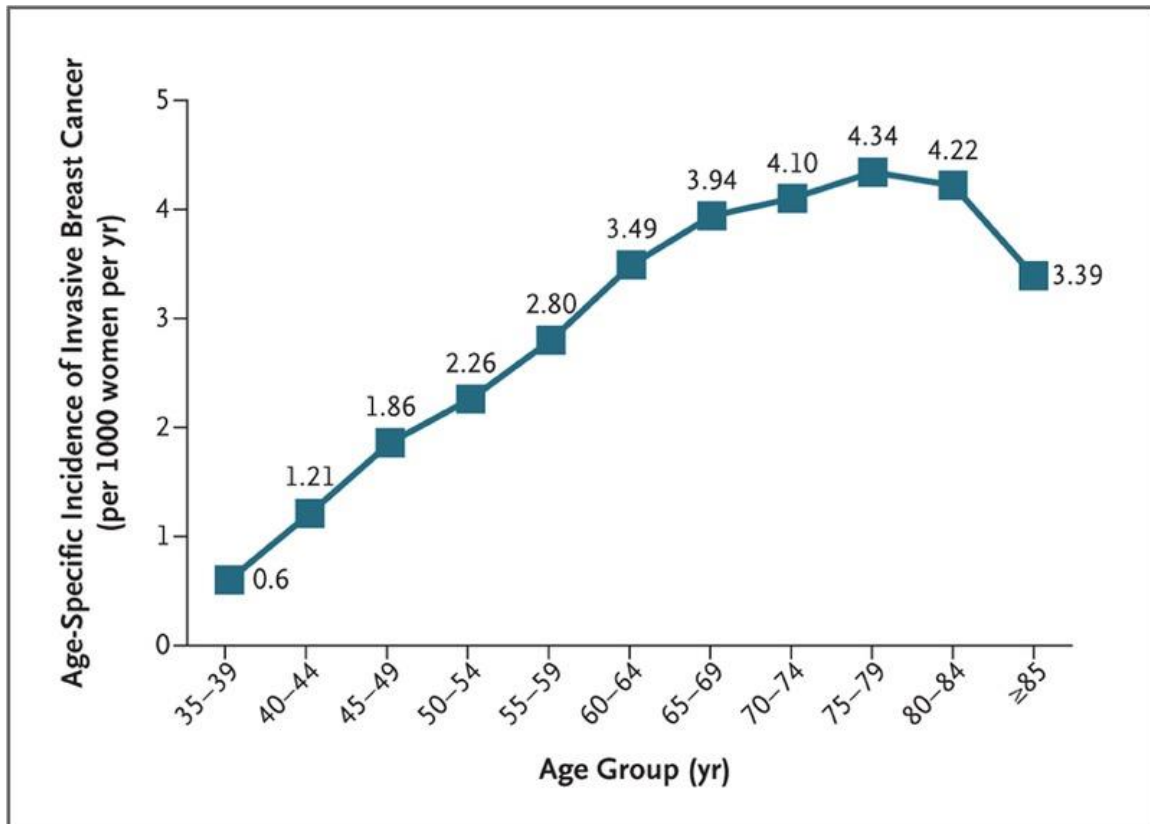


Figure 3.1: Screening for Breast Cancer

The evaluation of the molecular pattern of the tumor is also of extreme importance, because it predicts the response to specific therapies and the prognosis of the disease. It can be performed by immunohistochemistry, which quantifies estrogen (ER) and progesterone (RP) receptors and human epidermal growth factor receptor type 2 (HER2), or by microarray, which classifies tumors into molecular subtypes, according to their gene expression, in Luminal (A and B), HER and Basal-Like, with different clinical and prognostic implications.

3.2 COMPUTER AIDED DESIGN (CAD)

Drawings are often used by designers to depict the product they are developing and to convey the design to the others. In addition to drawing, they may also utilize symbolic & mathematical models, or even three-dimensional physical models – yet drawing is undoubtedly the most versatile and convenient method of representation. Drawings are mostly helpful for depicting the geometric shape and look of a design, but they may also be used for other purposes. As a result, computer graphics, or the creation of visual pictures using a computer, is critical in computer-

aided design (CAD). Technical drawings are a common tool in the design process. They describe the design and show how it relates to manufacturing. During the design and detailing stages, the level of competence and experience of the designers is critical. It takes a long time to make changes to prior designs since new drawings need to be created. Computer-aided drawings is a process that uses a computer to make engineering drawings and, like manual drawing, solely represents a design graphically.

A computer-aided design approach, on the other hand, combines the computer's capabilities with those of its user to create a problem-solving partnership. When the word "computer-aided design" (CAD) is used, it usually refers to the graphical system that allows three-dimensional modeling of parts and assemblies. The phrase design, on the other hand, encompasses the functions of modeling and analysis as well. Nowadays, the abbreviation CADD (computer-aided drug-finding and design) is increasingly often used. A CADD program is one that provides all drug-finding facilities as well as some or all of the design capabilities.

2D computer drawings depict objects using a single-view or motive format that each disclose two of the object's three dimensions in two dimensions. Only two values are stored in the database for each coordinate of the item. If the database just includes X and Y coordinates, it may also be a visual depiction. The coordinate format is three-dimensional (3D) computer-generated drawings. Geometric models of the product or component can be created using three-dimensional computer-aided drawing (CAD) software for visual and spatial analysis. In engineering, there are so many different types of drawings that it would be impossible to include them all here. The flexibility of contemporary CAD systems will be shown with a few example from the engineering fields.

- a. Planned layouts and parts of assembly are included in the Aerospace category.
- b. Chemical: Schematics and layout drawings for process plants.
- c. Structural details drawings, site plans, building drawings, and road-use plans are all examples of civil and construction drawings.
- d. VLSI (Very Large System Integration) design, circuit board design.
- e. Industrial and Manufacturing: Schematics for manufacturing facilities, components, and final assemblies.
- f. Subassemblies and assemblies of mechanical parts with in near future, computer-aided design (CAD) will play a significant role in engineering.

However, understanding the fundamentals of engineering sketching is a must. To put it another way, if we don't know these rules, we won't be able to draw anything, even with the assistance of a computer.

Computers, on the other hand, have simplified the process of drawing. The following are some of the many benefits of working with CADD systems: a drawing of consistent quality there is no correlation between the draughtsman's own skill level and the quality of the lines, measurements, symbols, and notes.

- i. Draughtsman and other processors may get relevant information by creating a database that they can then access.
- ii. It is possible to create a library of often used symbols such as bolts, screws and bearings that may be saved in the computer's memory and quickly retrieved when required. The library of symbols can be placed wherever on the screen and repainted to any size or angle of inclination.
- iii. the usage of layers drawing may be done on any one or more layers, which can be regarded a stack on transparent sheets and each individual sheet can be picked for drawing construct lines, grids, measurements, notes and hatching but can be combined to form one full drawing when necessary.
- iv. Speed and productivity are markedly increased as a consequence of saving on repeating repetitious labor on identical features of drawings and the fatigue and tedium that results is substituted by automated redrawing.
- v. An accurate material as well as cost estimate can be made due to greater computer mathematical accuracy, which leads to a reduction in the amount of mistakes.
- vi. Computers' ability to produce multi-colored drawings has made it simpler for people to grasp drawings, and the variety of colors available has made it even easier for people to comprehend them.
- vii. The ability to make edits with the use of computers, the advanced editing operations of repairing errors, removing and adding new features, copy, moving & rotating features, scaling etc., can only be accomplished.

3.2.1 MRI

The atom is made up of a central nucleus and electrons orbiting it. The atomic nucleus is made up of smaller particles that are neutrons and protons. Electrons have a negative electrical charge, protons have a positive charge and neutrons have no electrical charge. In the atomic structure one can observe movements like that of electrons around the nucleus.

These are the movements of electrons rotating on their own axis and of nuclei also rotating around their axes. Active nuclei in MR are those that tend to align their axis of rotation to an applied external magnetic field, due to the laws of electromagnetic induction. These nuclei therefore possess an effective charge and, when rotating within a magnetic field, acquire a magnetic moment, or angular momentum, or spin rotation. Alignment, or the summation of magnetic moments within a magnetic field, can be expressed as a summation vector.

Breast cancer is now a relevant public health problem. It is the most common malignancy in women in most parts of the world. According to the latest global statistics from, 2.1 million new cases of cancer and 627 thousand deaths from the disease were estimated. In Brazil, estimates of breast cancer incidence for the year 2019 are 59,700 new cases, which represents 29.5% of cancers in women, with the exception of non-melanoma skin cancer. In 2016, there were 16,069 deaths of women from breast cancer in the country. Although the five-year survival estimates show an increasing trend in developed countries, there is still a large global disparity. According to Concord-3 the five-year survival estimates were 76.9% (75.5 - 78) for the period from 2005 to 2009 and 75.2% (73, 9 - 76.5) for the period from 2010 to 2014. Factors related to knowledge of the disease and women's difficulties in accessing diagnostic methods and adequate and timely treatment result in the arrival of patients in more advanced stages of breast cancer, worsening the prognosis.

Currently, diagnosis, local treatment and systemic treatment for breast cancer are being rapidly improved, due to a better understanding of the natural history of the disease and the molecular characteristics of the tumors. In this scenario, planning breast cancer control strategies through early detection is essential. The earlier an invasive tumor is detected and treatment is started, the more likely it is to be cured. For this reason, several actions have been implemented to diagnose cancer in the early stages. However, due to organic individualization and extreme tumor heterogeneity associated with the presence of known and unknown risk factors, breast cancer is

considered a disease of dynamic behavior, in constant transformation. There is a wide variety of histological and molecular types of in situ and invasive breast carcinomas. The most common invasive histologic type is unspecified infiltrating ductal carcinoma, which represents 70 to 80% of all breast tumors, followed by infiltrating lobular carcinoma, with about 5 to 15%, and the other histologic types. 2012).

Clinical anatomical staging, according to the seventh edition of the TNM – Classification of Malignant Tumors defines the extent of the disease according to the size of the tumor, the presence or absence of compromised homolateral axillary lymph nodes and the presence of disease outside the breast. . Once the staging according to the TNM has been assigned, it is possible to group them into stages ranging from 0 to IV, with 0 being the classification referring to in situ breast carcinoma, I the most initial stage and IV the most advanced of invasive carcinomas

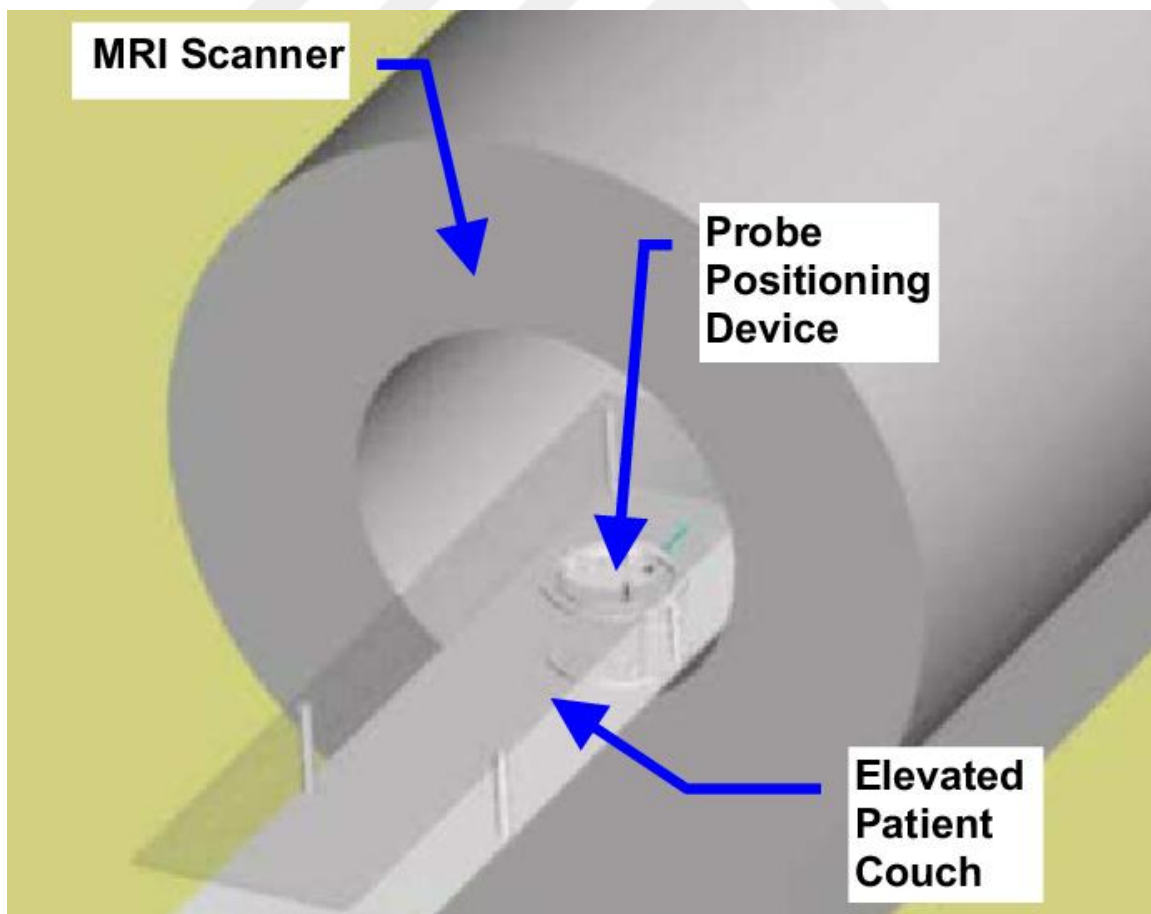


Figure 3.2: overall view of the device relative to an MRI

The vast majority of nuclei align in the same direction (parallel) and the other small part in the opposite direction (antiparallel) to the axis of the magnetic field. In the case of MR images, their generation is through the process of nuclear alignment, due to the excitation of protons by radiofrequency, spatial encoding and image formation.

The magnet aligns the nuclei with low energy (spins parallel to the magnetic field) and high energy (spins antiparallel to the magnetic field). By means of a radiofrequency source, the longitudinal vector is excited to the transverse plane, where the resulting signal is reached by a radiofrequency receiving antenna. Thus, the MRI signal comes from the nuclei of atoms in a certain region of the patient's body, due to the action of a homogeneous and uniform magnetic field. Carbon. The hydrogen atom is the most important for producing the highest signal strength.

Representation of the hydrogen atom. The nucleus is composed only of the proton the isotope of hydrogen is an active nucleus in MR, being formed only by one proton. Its atomic number is equal to the mass number. The hydrogen proton is used because it is more abundant in the human body, and because its individual proton provides a relatively large magnetic moment.

According to the laws of electromagnetism, a magnetic field is created when an electrically charged particle moves. When we talk about the hydrogen nucleus, it has a proton with a positive electrical charge and in rotation, that is, with movement. In this way, the hydrogen core has an induced magnetic field around it acting like a small magnet. The magnet of each hydrogen nucleus contains a north pole and a south pole of equivalent strength. Represents the hydrogen proton can be seen as a small sphere (1) that has a spin motion around its own axis (2). As it is a positively charged particle (3), it will generate its own magnetic field around it (4) behaving like a small magnetic dipole (4) or like a magnet (5) with a magnetic moment.

Is placed in an external magnetic field with fixed power. The resultant of this field is represented by a single vector, called the effective magnetization vector (EMV). The VME represents the energy difference between low and high energy hydrogen protons. In order for a change in the direction of the VME of a given patient tissue to occur, from a low energy state (parallel) to a high energy state (antiparallel), it is enough to add to the protons in question, energy in the form of radio waves represents hydrogen protons under the action of the applied external magnetic field. Protons are distributed in two energy levels, with a small larger number of protons lining up in parallel. As more energy is fed into the system, the greater the amount of proteomic

magnetic fields that change in the opposite direction, from bottom to top, and hence the intensity of the VME. Therefore, the VME is greater the greater the magnetic field in which the patient is inserted. Consequently, with high power fields, better signals are obtained.

3.2.2 Fine Needle Aspiration (FNA)

The increase in the incidence of breast carcinoma in recent decades, with all its implications in terms of morbidity and mortality for women, as well as the costs resulting from its treatment and complications, has demanded efforts to improve the propaedeutic arsenal for diagnosis. Early in this condition. The palpable breast lump is the main clinical situation to be investigated in order to differentiate between benign and malignant pathological processes.

For this, in addition to the clinical examination and mammography, fine needle aspiration (FNA) has been used alone or together, therefore, it is necessary to evaluate its performance. FNA was proposed as a diagnostic test. The evaluations started with studies of the validation type of diagnostic tests, mainly in European countries. Subsequently, North American studies using this test were resumed, based on the excellent experience obtained in Scandinavian countries⁶. In a recent publication prepared by the National Cancer Institute of the United States of America, recommendations for the use of FNA in breast lesions were gathered. The importance of this method has been reconfirmed, provided that it is performed by a professional with adequate training. A rate of slides with insufficient material of less than 20%, a rate of false positives of less than 1%, in addition to a regular evaluation of the rate of false-negative results are part of fundamental quality criteria for using the method

The FNA, which is being increasingly used in clinical practice, provides material for examining the cell morphology of a given lesion. If analyzed by an experienced cytopathology's, a cytological diagnosis of a series of benign and malignant lesions is possible, based on changes in cellularity, cell cohesion, characteristics of the nucleus and chromatin distribution¹. As long as it is performed with an appropriate technique, it has a high rate of agreement with the histological diagnosis, and can be repeated, if necessary, in addition to being an outpatient procedure, with minimal material expenditure. The use of this method has expanded in the approach of patients with palpable breast nodules, the main form of presentation of breast neoplasms in our country.

Among the methods used for the differential diagnosis of palpable breast lesions, FNA seems to be the single test with the best performance.

A prospective study was carried out to analyze the performance of this method in the diagnosis of palpable breast lumps, with evaluation of several aspects related to its diagnosis. The anatomy pathological examination of the nodules was performed after excisional biopsy in most cases, with only four cases in which it was based on core needle puncture material. The results of this examination were divided into benign (including non-neoplastic and neoplastic lesions) or malignant (compatible with malignant epithelial, mixed epithelial/mesenchymal tumors or of diverse origin). To assess the average diameter of the nodules, the mammographic image was measured in centimeters. This measurement was chosen, although this test was not part of the current study, as it is not possible to accurately measure all lesions in the histological examination (some carcinomas underwent chemotherapy treatment prior to surgery). To calculate the statistical significance of the differences found in the analysis of age and mean diameter of the nodules in both groups (malignant or benign disease), Student's t test was used. The statistical package Epi. Info version 6.02 was used for the different analysis procedures. The study project was previously approved by the Institution's Research Committee and Ethics Committee.

According to the laws of electromagnetism, a magnetic field is created when an electrically charged particle moves. When we talk about the hydrogen nucleus, it has a proton with a positive electrical charge and in rotation, that is, with movement. In this way, the hydrogen core has an induced magnetic field around it acting like a small magnet. The magnet of each hydrogen nucleus contains a north pole and a south pole of equivalent strength. Represents the hydrogen proton can be seen as a small sphere (1) that has a spin motion around its own axis (2). As it is a positively charged particle (3), it will generate its own magnetic field around it (4) behaving like a small magnetic dipole (4) or like a magnet (5) with a magnetic moment.

Is placed in an external magnetic field with fixed power. The resultant of this field is represented by a single vector, called the effective magnetization vector (EMV). The VME represents the energy difference between low and high energy hydrogen protons. In order for a change in the direction of the VME of a given patient tissue to occur, from a low energy state (parallel) to a high energy state (antiparallel), it is enough to add to the protons in question, energy in the form

of radio waves represents hydrogen protons under the action of the applied external magnetic field. Protons are distributed in two energy levels, with a small larger number of protons lining up in parallel. As more energy is fed into the system, the greater the amount of proteomic magnetic fields that change in the opposite direction, from bottom to top, and hence the intensity of the VME. Therefore, the VME is greater the greater the magnetic field in which the patient is inserted. Consequently, with high power fields, better signals are obtained

3.2.3 Cat Scan for Breast Cancer

Breast cancer is a condition that affects a large percentage of people throughout the globe and may be fatal. Breast cancer affects around 10% of women at some point in their life. Breast cancer will be found in the breasts of around 25,700 Canadian women this year. This means that for every 100,000 women, there are 130.1 instances of breast cancer. More than a quarter of all Canadian women will be diagnosed with breast cancer over their lifetime. Breast cancer is more common as women become older. Canadian women so over age of 50 are projected to account for 83 percent of all new instances of breast cancer.

Eighty-seven percent of women who were diagnosed with prostate cancer five years ago are still alive, according to statistics. Women of all ages may benefit from early detection, diagnosis, and treatment for breast cancer, which can extend their lives following a diagnosis. As a result, it is widely agreed that early identification of breast cancer is critical to the patient's survival. Breast cancer detection requires a high degree of high sensitivity and specificity since a false positive result might necessitate the surgical excision of the breast. Breast cancer may develop in males as well, as the tissues involved are the same in both sexes.

One percent of males are at risk for breast cancer, although the survival statistics for both women and men are the same. Detection of breast cancer is the topic of this research, which examines several strategies for analyzing breast pictures. Breast radiologists now have access to a wide variety of imaging modalities that may aid in the early detection of breast cancer. Imaging methods have recently seen significant advancements that have the potential to revolutionize medicine in the future. In addition to the clinical breast exam, there are a variety of imaging modalities to choose from, including mammography, breast ultrasonography, computed

tomography, positron emission tomography, magnetic resonance imaging (MRI), electrical impedance tomography, and breast thermography.

When it comes to breast cancer diagnosis, breast thermography is a novel and innovative screening tool that relies on the warmth a tumor may emit to assist identify the disease early. A range of other breast problems may also be detected using these imaging methods, which assess anomalies suggesting the presence of malignancy.

In certain cases, these imaging modalities are utilized for screening purposes, while in others they are employed for diagnostic purposes. Effective screening methods must have several key characteristics including accuracy, high sensitivity, and adequate specificity as well as being simple to use and using a comfort aid throughout the test. They must also be inexpensive. These imaging approaches are discussed in depth in the next parts of this article, which is the primary goal of this work. Additionally, it will explain the pros and disadvantages of each imaging method.

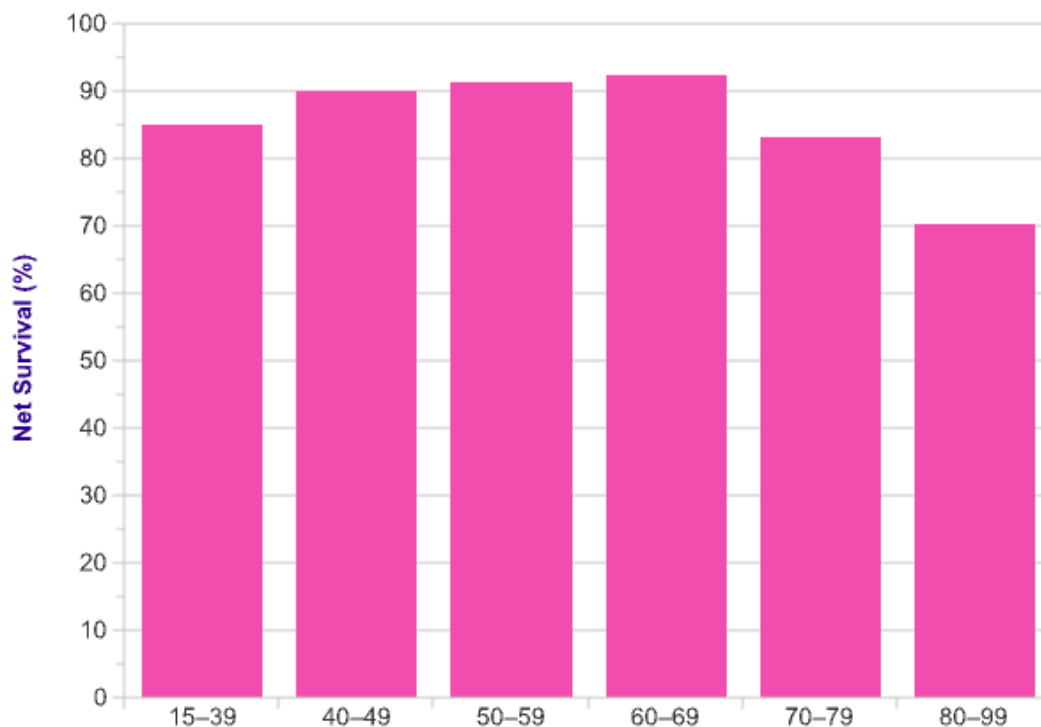


Figure 3.3: Survival rates for breast cancer

One percent of males are at risk for breast cancer, although the survival statistics for both women and men are the same. Detection of breast cancer is the topic of this research, which examines several strategies for analyzing breast pictures. Breast radiologists now have access to a wide variety of imaging modalities that may aid in the early detection of breast cancer. Imaging methods have recently seen significant advancements that have the potential to revolutionize medicine in the future. In addition to the clinical breast exam, there are a variety of imaging modalities to choose from, including mammography, breast ultrasonography, computed tomography, positron emission tomography, magnetic resonance imaging (MRI), electrical impedance tomography, and breast thermography.

When it comes to breast cancer diagnosis, breast thermography is a novel and innovative screening tool that relies on the warmth a tumor may emit to assist identify the disease early. A range of other breast problems may also be detected using these imaging methods, which assess anomalies suggesting the presence of malignancy.

In certain cases, these imaging modalities are utilized for screening purposes, while in others they are employed for diagnostic purposes. Effective screening methods must have several key characteristics including accuracy, high sensitivity, and adequate specificity as well as being simple to use and using a comfort aid throughout the test. They must also be inexpensive. These imaging approaches are discussed in depth in the next parts of this article, which is the primary goal of this work. Additionally, it will explain the pros and disadvantages of each imaging method

3.3 IMAGE SEGMENTATION

Images are being used in a variety of spheres of human endeavor right now for the purposes of storing information, displaying information, and providing information about the world that we live in. Finding a way to quickly extract useful information from raw image data presents a problem for researchers, engineers, and businesspeople. Converting images into information is the primary purpose of image processing. This is the primary objective of image processing. In general, processing functions that include recognition and interpretation are associated with image analysis applications. The goal of these applications is the automatic, or even partially automatic, extraction of information from an image.

In other words, processing functions like these are associated with image analysis applications. The reduction of the information contained within a digital image into more or less homogenous sections, in order to provide a more appropriate analysis of the image, is one of the most important aspects in the process of extracting information from a digital image. This is the segmentation problem, which has led to the creation of a number of potential solutions. In the context of this central theme, the purpose of this work is to present a discussion about some factors that are considered relevant for the conceptualization and rationale of segmentation, as well as to carry out a brief review of the algorithms that are most commonly used in the segmentation of digital images.

The quality of the processing that is carried out on the data source is intimately connected to the ability to obtain better results from the analysis of information that has been retrieved from an image. According to, the methods that are used for extracting information from images are classified as falling under the category of processing that is known as image analysis. In order for image analysis to be carried out, the first step that must be taken is the performance of a segmentation. The field to which it will be applied has a direct bearing on how picture segmentation is conceptualized and understood.

Within the realm of computer vision, the process of breaking down a digital image into its component parts, or segments, is known as segmentation. These segments are often referred to as regions. Image segmentation is the component of image analysis that deals with the delineation of geographic objects or regions in an image. It is a subfield of digital image processing that is used in the field of remote sensing image processing. Two fundamental qualities are shared by the elements that compose a picture: first, those elements demonstrate some degree of internal uniformity with regard to a quality of the image, and second, those elements contrast with the elements that surround them.

These features, on the other hand, are affected by noise, and as a result, they cannot be represented in a deterministic way. Image segmentation has two primary goals at its core: the first is to break the image down into its component pieces in order to conduct additional analysis, and the second is to effect a change in the image's representation. It is reasonable to anticipate that the end result of a segmentation process will be a finite number of homogenous regions, which will individualize the many regions that are contained inside a picture. Therefore, the

image would be composed of a collection of regions, each of which would be connected by their edges and would not overlap with any other areas. This would ensure that each pixel in the image would only belong to one particular region. The concept of segmentation as provided idealizes the real world in a simple way for numerous uses.

However, it is worth remembering that it is an invention of the human mind and that although it is a useful method to materialize solutions for image analysis tasks, it might in some instances lead to disappointing outcomes. It may be argued that the creation of algorithms targeted at automating segmentation operations is one of the most difficult challenges within the image processing sector. For image segmentation is primarily an issue of psychophysical perception, and hence not accessible to a merely analytical solution. Thus, to produce a good result, any segmentation method, regardless of the mathematical model employed, must be complemented with heuristics that involve both information and semantics as a priori knowledge about the images being processed. The construction of the algorithms takes into account two essential aspects of pixels that entail two concepts: discontinuity and similarity.

Through the usage of these principles, as structured by, segmentation is done via edge detection and similarity detection. In turn, the most popular methods employed in segmentation may be grouped into three main groups, based to the way they see the image: Segmentation methods in the spatial domain; Segmentation methods in the measurement space; and, Frequency domain segmentation methods. It is worth mentioning that the quality of picture segmentation directly effects the result of the analysis phase and, thus, the choice of an adequate segmentation technique entails the knowledge of the details of the existing techniques.

3.3.1 K-means Algorithm

One method that can be utilized in the clustering of data is called the K-Means algorithm, and it is founded on the concept of division. The processes of data mining and pattern recognition typically make use of unsupervised approaches just like the one that is explained in this article. This tactic makes the cluster's performance index lower by applying square-error and error criteria, which together drive the value of the index down. This technique looks for K divisions that meet a certain set of parameters in the order that you provide. After that, group the remaining sample points to their respective cluster focal points based on the minimum distance

criterion; finally, obtain the first classification; if the classification is unreasonable, modify it (recalculate each cluster focal point); and finally, obtain the final classification.

The K-Means method divides the data into distinct groups or clusters. It is also rapid and effective, in addition to being easy to use, which are all perks it delivers. Because of the nature of the operation itself, the results of the process are frequently unanticipated the majority of the time. It is also crucial to note that the gradient method is utilized in this approach in order to locate the extremums of the target function.

This fact is vital to notice because it is related to the previous point. Because gradient approaches operate to decrease energy in an ascending direction, a defective initial cluster focus point might lead to a local minimum in the energy reduction. This is because gradient approaches work to reduce energy in an upward way. The search is at an end at this point. "Data clustering" refers to a method that groups data sets that are extremely similar to one another with the purpose of making the process of managing them more efficient. Data clustering is one method that engineers may do in order to find component families for cellular manufacturing.

When it comes to clustering data, one of the most frequent ways is called the K-means clustering method. Before it can be used, the number of clusters that are present in the data must first be specified. This is a prerequisite for using the tool. Because clustering is an essentially subjective process, it can be difficult to establish the ideal number of clusters to apply to a certain data set. This contributes to the process being a considerable amount more complicated. Utilizing the K-means clustering method is the best way to get at the "K" number that represents the ideal number of clusters.

However, it is worth remembering that it is an invention of the human mind and that although it is a useful method to materialize solutions for image analysis tasks, it might in some instances lead to disappointing outcomes. It may be argued that the creation of algorithms targeted at automating segmentation operations is one of the most difficult challenges within the image processing sector. For image segmentation is primarily an issue of psychophysical perception, and hence not accessible to a merely analytical solution. Thus, to produce a good result, any segmentation method, regardless of the mathematical model employed, must be complemented with heuristics that involve both information and semantics as a priori knowledge about the

images being processed. The construction of the algorithms takes into account two essential aspects of pixels that entail two concepts: discontinuity and similarity.

Through the usage of these principles, as structured by, segmentation is done via edge detection and similarity detection. In turn, the most popular methods employed in segmentation may be grouped into three main groups, based to the way they see the image: Segmentation methods in the spatial domain; Segmentation methods in the measurement space; and, Frequency domain segmentation methods. It is worth mentioning that the quality of picture segmentation directly effects the result of the analysis phase and, thus, the choice of an adequate segmentation technique entails the knowledge of the details of the existing techniques

By employing a set of objective evaluation criteria to find the ideal values for K , we can do away with the necessity of employing a trial-and-error method to discover what those values should be. Image segmentation methods that are utilized in a variety of settings and serve a variety of functions are all dependent on threshold values, borders, and regions. However, when compared to other clustering algorithms, the k means clustering technique is the most effective way to detect any pattern's dominance in a picture. This is because it prioritizes the patterns that appear most frequently. This is the case since the approach known as k -means clustering was developed with the intention of serving this function. When conducting research on behaviors and patterns, it is essential to make use of cluster analysis.

It's likely that using the K means strategy will assist the subconscious grouping pattern in becoming more adept at differentiating between different types of things. With the use of the k means technique, continually evolving clustering algorithms for medical image analysis are being developed in an effort to improve upon those that are already available. Both the central processor unit and the graphics processing unit are put to use in these ways (GPU). The k -means data clustering method takes a considerable amount of time, and the amount of time required to complete the task grows proportionally with the size of the dataset. Because it is such an excellent method, graphics processors typically make use of the k -means algorithm for the purpose of grouping massive datasets (GPU). The method of k -means clustering is currently being implemented in a wide variety of various algorithms, one of which is the genetic algorithm. Reduce the amount of background noise in adaptive k means by lowering the

brightness, which is symbolized by the letter l , before beginning segmentation in the LAB color space. This is done in order to cut down on the quantity of background noise.

One aspect of research was the use of sequential k means clustering and machine learning to make a prediction about the process of asphalt coating stripping. It is important to keep in mind that the process of training a machine learning system could take a considerable amount of time. The prediction of sample stripping in asphalt mixed samples might be aided by the use of a categorization that is derived from machine learning.

It is possible that this will cut down on the amount of time necessary for the execution of the method, while simultaneously improving its structure over multiple processor cores, making it a great candidate for parallelization. On a single photo dataset, an estimate of the typical number of pixels that are associated with an area of interest (ROI) was computed by utilizing the Map-Reduce architecture. This allowed for more accurate results to be obtained. Akhtar examined the patterns of batch processing for thousands of photographs, taking attention to the length of time the CPU was active as well as the accuracy of its operations. According to the results of their research, increasing the number of compute nodes or cores resulted in the production of a more comprehensive data set in the end.

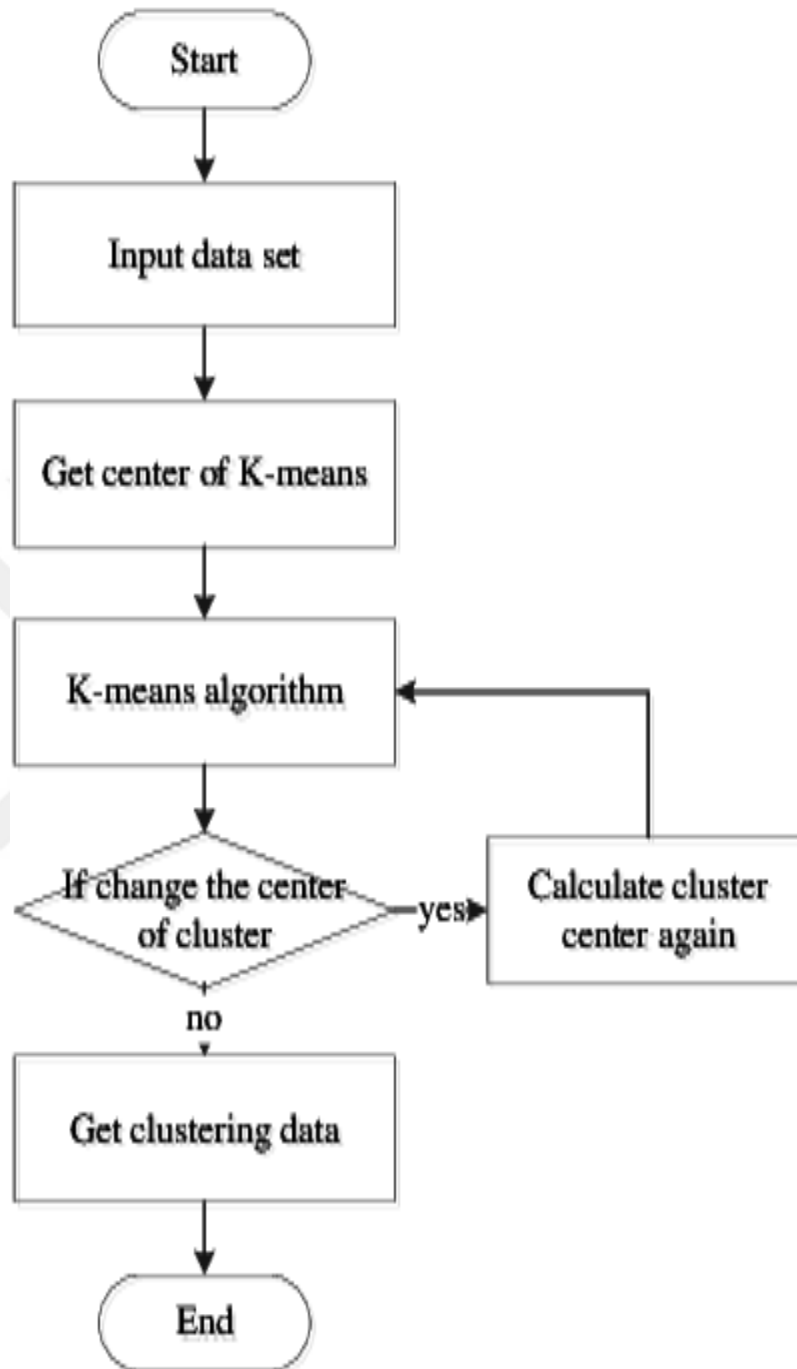


Figure 3.4: K-means algorithm the clustering

3.3.2 Fuzzy C-Means FCM

You may construct fuzzy clusters using the Fuzzy C-Means method independently. The "quality of partition" is decided by supplied criteria, and the dataset is "hard partitioned" accordingly.

Hard partitioning means that each data point belongs to a single cluster. However, soft clustering uses a statistical model to find "soft partitions" inside a dataset. In a "soft partition," a single data point might belong to many clusters. The input space may be larger than the dataset, which distinguishes a soft partition from a fuzzy partition. Generally, soft partitioning methods are used to generate a soft fuzzy partition. Point x has a membership degree of one in all clusters, suggesting a unique soft clustering.

A restricted soft partition satisfies the additional limitation criteria. It provides a limited number of soft divisions when using fuzzy C-Means. Constrained soft partitioning was achieved by improving the objective function J_1 of hard cmeans. The computation includes a cluster's degree of fuzzy membership. Fuzzy membership now has a weight exponent m . J_m is a fuzzy partition of the dataset X formed by the clusters $C_1, C_2, \dots, C_k$ and the cluster count k . The fuzzy partition of the dataset X formed by the clusters C_1, C_2, C_k and their number k .

The parameter m affects the clustering findings since it represents the weight of each cluster member in the final conclusion. Finding a good partition in fuzzy c-means is like finding the best prototype V_i in hard C-means. For fuzzy c-means, unlike hard c-means, a membership function c_i must be established to reduce J_m . To be a local minimum of the objective function J_m , a limited fuzzy partition must meet the following criteria: Using the FCM algorithm requires certain critical considerations: If m is greater than 1, it will converge. At a local minimum, the objective function J_m is minimized. The parameters m and c used in the FCM analysis, as well as the original prototype used, all influence the results.

$$\sum_j \mu_{c_j}(x_i) = 1 \quad \forall x_i \in X \quad \dots\dots\dots 3.1$$

$$J_m(P, V) = \sum_{i=1}^k \sum_{x_k \in X} (\mu_{C_i}(x_k))^m \|x_k - v_i\|^2 \quad \dots\dots\dots 3.2$$

The operation known as image segmentation is one that plays a very important role in the overall processing of images. In order to get the healing process started and to move it forward more quickly, it is essential to make decisions based on medical imaging both before and after surgery. These decisions should be made in order to move the healing process along more quickly. When

attempting to diagnose aberrant tissue growth with the help of a computer, one should strive for precision to the fullest extent that is practically possible. In other words, one should be as precise as they can get away with being. The high-speed computer procedures that are currently accessible are far more accurate than the manual segmentation of these aberrant tissues, which is not even close to being as accurate as the human method.

However, it is worth remembering that it is an invention of the human mind and that although it is a useful method to materialize solutions for image analysis tasks, it might in some instances lead to disappointing outcomes. It may be argued that the creation of algorithms targeted at automating segmentation operations is one of the most difficult challenges within the image processing sector. For image segmentation is primarily an issue of psychophysical perception, and hence not accessible to a merely analytical solution. Thus, to produce a good result, any segmentation method, regardless of the mathematical model employed, must be complemented with heuristics that involve both information and semantics as a priori knowledge about the images being processed. The construction of the algorithms takes into account two essential aspects of pixels that entail two concepts: discontinuity and similarity.

Through the usage of these principles, as structured by, segmentation is done via edge detection and similarity detection. In turn, the most popular methods employed in segmentation may be grouped into three main groups, based to the way they see the image: Segmentation methods in the spatial domain; Segmentation methods in the measurement space; and, Frequency domain segmentation methods. It is worth mentioning that the quality of picture segmentation directly effects the result of the analysis phase and, thus, the choice of an adequate segmentation technique entails the knowledge of the details of the existing techniques

In clinical settings, the imaging modalities that are utilized in the majority of cases are ones such as ultrasound, computed tomography (CT) with X-rays, digital mammography, and magnetic resonance imaging (MRI) (MRI). The contrast of the images obtained through MRI is substantially higher than the contrast of the images obtained through CT scans (CT). The magnetic resonance imaging (MRI) system is used in the great majority of studies on medical image processing (MRI). Because it requires differentiating between white matter, grey matter, CSF, and pathological entities like tumors, labeling voxels in MRIs is a well-known segmentation challenge. This is because the task at hand involves labeling voxels. This is a

problem that has been present for a considerable amount of time. Utilization of an MRI for the examination of a human body (magnetic resonance imaging) The act of breaking down an image into its component elements is referred to as thresholding. You can construct basic or complex segmentations from images that have intensities of varying degrees by using this method.

The images must first be converted into grayscale. Pixels are divided into many categories using this technique, and the classifications are based on whether or not the pixel meets a certain criteria or threshold. The precedent for this norm was set in advance. There is a significant cause for concern due to the fact that utilizing this method of segmentation only results in the production of two separate groups. As a direct result of this fact, this method is unable to process images that contain a high number of channels, which is a major limitation of the methodology. In addition, the spatial qualities that might be anticipated from MRI images are not taken into consideration. Some examples of these properties are noise and intensity inhomogeneity. If either of these characteristics is present and functioning, there is a possibility that the resulting histogram will be extremely skewed. Researchers have developed thresholding algorithms that detect medical photographs based on the local intensities and connectivity data that are present in the images. These algorithms were established by the researchers.

The development of these algorithms was motivated by the desire to find solutions to these problems. It is possible for the thresholding process to be made more difficult by non-stationary items and noise, ambient light, a significant number of gray levels in both the item and its backdrop, an insufficient amount of contrast, and an object that is out of proportion with the scene. When particular conditions are met, the target function is produced by employing a function that conducts an examination of the distinctions that can be made between classes, the object, and the environment that is all around it. Find the local minimum of the divergence function first, and then divide that local minimum by two in order to determine the threshold. Finding the local minimum of the divergence function is the first step in determining the threshold. After that, only then will one have reached the limit.

Imaging techniques based on magnetic resonance, more generally known as MRI, are increasingly being put to use in the process of diagnosing and treating conditions that affect the brain. This is mostly owing to the fact that magnetic resonance imaging (MRI) is capable of detecting both the size and location of tumors with a very high level of accuracy and precision.

Research strategies such as region-growing and shape-based procedures have traditionally been employed relatively extensively in brain investigations. [Case in point:] [citation needed] [Case in point:] The first thing that has to be done when calculating the Bayes factor is to make sure that the features of the environment that the subject is in are taken into account. This method does not produce photographs that have a natural look, despite the fact that it is not an entirely automated process.

The results of using this method are not satisfactory (human input is required to choose a seed). It will only function correctly under conditions that are stable throughout time. Because of the inherent noise in the process, the recovered zones may on occasion have gaps or even be completely discontinuous. This is due to the fact that the process itself generates noise. This is because the approach itself creates noise, which is the reason for this result. One additional method that can be utilized to segment a brain tumor is to make use of a technology that is based on the shape of the tumor. This can be done in a number of different ways. Having said that, there are a number of constraints that need to be taken into consideration when employing this method. It is necessary to first develop a strategy for the area's contours in order to be able to extract the region of interest that is of interest.

This strategy is also semi-automatic as a direct result of this, very much like the plan for the growth of the region. Both the initial design description and the procedure for selecting seed pictures include human input, which means that there is a risk of making errors in either one of these areas. It would appear that the most dependable alternatives to the procedures that were previously discussed are the applications of statistical analysis and fuzzy logic. Utilizing both supervised and unsupervised processing techniques at the same time is considered to be industry standard practice in the field of image processing.

When KNN and both hard and fuzzy c-means clustering are applied, there appears to be a limitation that prevents successful volume segmentation. This constraint is caused by the dispersion of data that exists between slices. Due to this constraint, accurate volume segmentation cannot be achieved. Through the utilization of supervised classification methods, we are able to construct representative parameters for each of the classes that are of interest to us. This is made possible because we have access to a very large number of pixels whose values are

already known to us. When utilizing the unsupervised categorization method, prior knowledge of the categories being categorized is not required in order to be successful.

Clustering is a process that is used in order to determine how the visual input ought to be categorized, and it is one of the methods that are applied. The application of a KNN, machine learning, or Parzen window classifier requires the involvement of a human operator and cannot be carried out independently. The K-Means and FCM procedures, in addition to hierarchical clustering, minimum distance/maximum distance, and a number of other methods, are examples of the kind of classification that is known as unsupervised classification.

3.4 DEEP LEARNING

There are four major eras to consider: the Enlightenment (1890–1969), the trough (1969–1982), the recovery (1982–1986), and the current new age (1986 to now). Artificial intelligence has been around since the early twentieth century, when academics first started looking at it. When McCulloch and Pitts initially proposed propositional logic as a means of analyzing brain processes and their relationships in 1943, it was deemed innovative. Although not widely acknowledged, the significance of this finding for artificial neural networks has been widely acknowledged (ANNs).

When it came to the building of the first artificial neural network, it was widely assumed that the perceptron was the first of its kind. However, as Minsky points out, the perceptron's inability to cope with nonlinear issues has led to the surge's downfall. In the years that followed, no substantial advances in the area of ANNs were achieved, and the field of ANNs entered its winter. The Hopfield neural network, a recursive artificial neural network (ANN), was developed in 1982, and the notion of deep learning was introduced to the general public at the same time. Deep learning models, both unsupervised and supervised, may be classified as either supervised or unsupervised. Deep learning models that are unsupervised are more challenging to categorize. There are several kinds of supervised models available, the most common of which are the classic ANN, CNN, and RNN models. CNNs are often used to deal with array-like data, however the research suggests that RNNs are better suited to cope with sequential data.

In addition to being the ANN counterpart, CNN components contain features like as pooling and flattening, which are used to reduce the dimensionality of information acquired by CNN blocks. RNNs utilize long-term memory (LSTM) to extract temporal information rather than ANNs and CNNs, which varies from the three types of model architectures that are often used. Before being used to train supervised models, the data was thoroughly annotated. However, the utility of certain annotated data sets is very restricted, and not all models must be trained in a supervised way. Deep learning also employs unsupervised models, which are in line with the current trend in unsupervised learning. There are several unsupervised models accessible, such as self-organizing maps (SOMs), Boltzmann machines, and autoencoders.

SOMs are often employed in feature dimension reduction, in part because the output dimension is always reduced to two. Boltzmann machines are less often used for feature reduction than restricted Boltzmann machines (RBMs), which are the foundation of deep belief networks (DBNs). Restricted Boltzmann machines (RBMs) are more popular than Boltzmann machines for feature reduction. Another prominent deep learning approach that is becoming increasingly popular is the use of auto encoders to minimize the amount of features. Because CNNs are widely utilized in the field, we'll focus our efforts on CNNs with deep learning capabilities. Those who want to learn more about it are urged to take part in the work. There are semi supervised models that include both supervised and unsupervised learning models, as well as supervised and unsupervised learning models, in the region between generative networks and attention processes.

A variety of auto encoders (VAEs) and more complicated generative adversarial networks (GANs) are being used to supplement mammography data. With the addition of the attention mechanism, classifiers may now concentrate their attention on the most relevant elements of each picture rather than on everything else. This section will provide an overview of CNNs and fully convolutional neural networks (FCNNs), as well as their applications (FCNNs). While CNNs are often used as classifiers in classification, FCNNs are mostly used as semantic segmentation models in semantic segmentation. We felt it was vital to include information on these two types of networks in our results since this is a systematic survey article. Furthermore, since both of them play critical roles in breast disease computer-aided design systems, a clear description would be beneficial to readers from many areas of life.

Finally, we offer a variety of cutting-edge deep learning applications to illustrate a wider range of its capabilities rather than limiting it to breast cancer alone. It is difficult to prevent breast cancer in the contemporary world, particularly in industrialized nations, due to a variety of factors including genetics, lifestyle, and environmental exposures. Mammography is one of the most used screening technologies because of its easy screening technique and low cost of study. However, a skilled expert radiologist is needed to correctly detect abnormalities in a mammography. Micro calcifications, which are aberrant formations smaller than one millimeter in diameter, might be difficult to see in certain circumstances. Another problem is that dense breast tissues may be difficult to differentiate from lesions that have a similar shape and contrast. In order to improve the survival rate of cancer patients, it was necessary to design a smart tool that could give expert-level analysis of mammograms.

Health care in the community is undergoing a technological revolution thanks to two new developments. With big data comes an abundance of information that can be shared, analyzed, and integrated in ways that accommodate a broad variety of differences. In the realm of medical imaging, a possible example is the Medical Image Archive, which provides access to thousands of pictures in a variety of modalities. For the second time, artificial intelligence has surpassed human intellect, thanks to deep learning (or beyond). To extract previously unreachable information, deep neural networks may now be built with many more layers and more computing capacity (mostly based on graphics processing units, or GPUs).

The results of using this method are not satisfactory (human input is required to choose a seed). It will only function correctly under conditions that are stable throughout time. Because of the inherent noise in the process, the recovered zones may on occasion have gaps or even be completely discontinuous. This is due to the fact that the process itself generates noise. This is because the approach itself creates noise, which is the reason for this result. One additional method that can be utilized to segment a brain tumor is to make use of a technology that is based on the shape of the tumor. This can be done in a number of different ways. Having said that, there are a number of constraints that need to be taken into consideration when employing this method. It is necessary to first develop a strategy for the area's contours in order to be able to extract the region of interest that is of interest.

This strategy is also semi-automatic as a direct result of this, very much like the plan for the growth of the region. Both the initial design description and the procedure for selecting seed pictures include human input, which means that there is a risk of making errors in either one of these areas. It would appear that the most dependable alternatives to the procedures that were previously discussed are the applications of statistical analysis and fuzzy logic. Utilizing both supervised and unsupervised processing techniques at the same time is considered to be industry standard practice in the field of image processing.

When KNN and both hard and fuzzy c-means clustering are applied, there appears to be a limitation that prevents successful volume segmentation. This constraint is caused by the dispersion of data that exists between slices. Due to this constraint, accurate volume segmentation cannot be achieved. Through the utilization of supervised classification methods, we are able to construct representative parameters for each of the classes that are of interest to us. This is made possible because we have access to a very large number of pixels whose values are already known to us. When utilizing the unsupervised categorization method, prior knowledge of the categories being categorized is not required in order to be successful.

Clustering is a process that is used in order to determine how the visual input ought to be categorized, and it is one of the methods that are applied. The application of a KNN, machine learning, or Parzen window classifier requires the involvement of a human operator and cannot be carried out independently. The K-Means and FCM procedures, in addition to hierarchical clustering, minimum distance/maximum distance, and a number of other methods, are examples of the kind of classification that is known as unsupervised classification

Images may be segmented, analyzed, and classified using convolutional neural networks, which have a significant influence. Deep learning models have already been built for the diagnosis and detection of breast cancer by analyzing digital mammograms. An innovative convolutional neural network (CNN) design for breast cancer detection is presented in this research. Various track of convolution/deconvolution layers are used to extract picture characteristics in the proposed architecture. The suggested structure is meant to address the issue of abnormalities variability in digital mammography, and it is projected to enhance the detection of lesions with a better quality than current approaches.

3.4.1 CNN Algorithm

CNN is a well-known example of a deep learning architecture that was developed with the visual perception mechanism of living creatures as its primary source of inspiration. This resulted in the creation of CNN as an example of a deep learning architecture. This design objective was met by basing CNN on the visual perception process of live things, which is how living things see the world. Beginning in the late 1960s, a number of major advancements were made in a variety of academic fields, which opened the way for it to become conceivable to conceive of the idea. Beginning in the late 1960s, a number of significant advances were made in a variety of academic fields. The concept of a hierarchical structure among the neurons that make up the visual system was proposed for the very first time in 1968 as a consequence of research that was conducted on the visual cortex of monkeys.

These investigations were carried out in the year 1968. According to this theory, there are certain cells in the visual system that are responsible for establishing a "receptive field" for other cells that are situated at a higher level. These cells can be found at lower levels. Receptive fields that are relatively simple can be combined in this way to produce receptive fields that are more complicated and have a wider range of use. 1980 marked the year that a recognition neural network was initially put on display publicly for the first time. This network makes use of the theories behind in the form of a network that is based on layers of neurons that utilize sparse connections, which is comparable to what we refer to as convolution in the modern day. The recognition was utilized with the intention of accomplishing the goal of recognizing handwritten numbers, with the primary qualities being the capacity to learn in an unsupervised manner and the requirement that it be invariant to the need that it be invariant to the requirement that it be invariant to the displacement of inputs. In 1989, LeCunn presented a proposal for a network design that incorporated both convolutional and fully connected layers. This idea was made public in the year 1989. The community of network administrators was given this proposition to consider. Backpropagation was used as a method for training such a network in order to improve its performance. Following training, the network was put to use to recognize handwritten numerals, and it did so with only a 5% error rate.

The network is comprised of four layers: two convolutional layers, two completely linked layers, and two layers that are fully connected to each other. An application is built in the year 1994 that makes use of a network that contains convolutional layers in order to tackle the problem of recognizing minute calcifications in digital mammograms. This was a problem that had been around since the year 1994. In addition to this, the writers had also been working simultaneously on the development of a computer software that is capable of recognizing digits. The conventional model of CNNs was initially created in the year 1998. This model is comprised of networks of interleaved convolutional layers, max pooling operations for dimensionality reduction, and some fully connected layers for the final classification. This model was constructed using convolutional layer networks that were layered atop one another and interleaved. This article also dives into the famous LeNet-5 neural network as a topic of discussion, and it does so in some detail. It is judged against other handwritten digit recognition systems that have a higher degree of performance, and the winner is chosen based on the results of this evaluation. It was proposed in 2005 to use Graphics Processing Units (also known as GPUs) to perform computational calculations for applications related to machine learning. This came about as a result of the continued development of GPUs, also known as Graphics Processing Units, as well as the demand for graphics processing skills in 3D games.

This demand is a direct consequence of the high demand for people with skills in graphics processing in 3D games. This is because graphics processing units (GPUs) carry out highly optimized multiplications of matrices in order to carry out computations that are linked to three-dimensional geometry. The objective of these computations is to create images. In addition, graphics processing units, sometimes known as GPUs, have the ability to simultaneously manage a sizeable number of matrices. After making a few changes to the way the machine learning algorithms were implemented on GPUs, it was discovered that they were highly effective once these changes were done. In the year 2012, it was proposed that CNN AlexNet should be put into action. This network participated in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), and it ended up winning the competition by a very significant margin. Because of this, the error rate at the current state of the art dropped from 25.8 percent to 16.4 percent, which is a direct result of this.

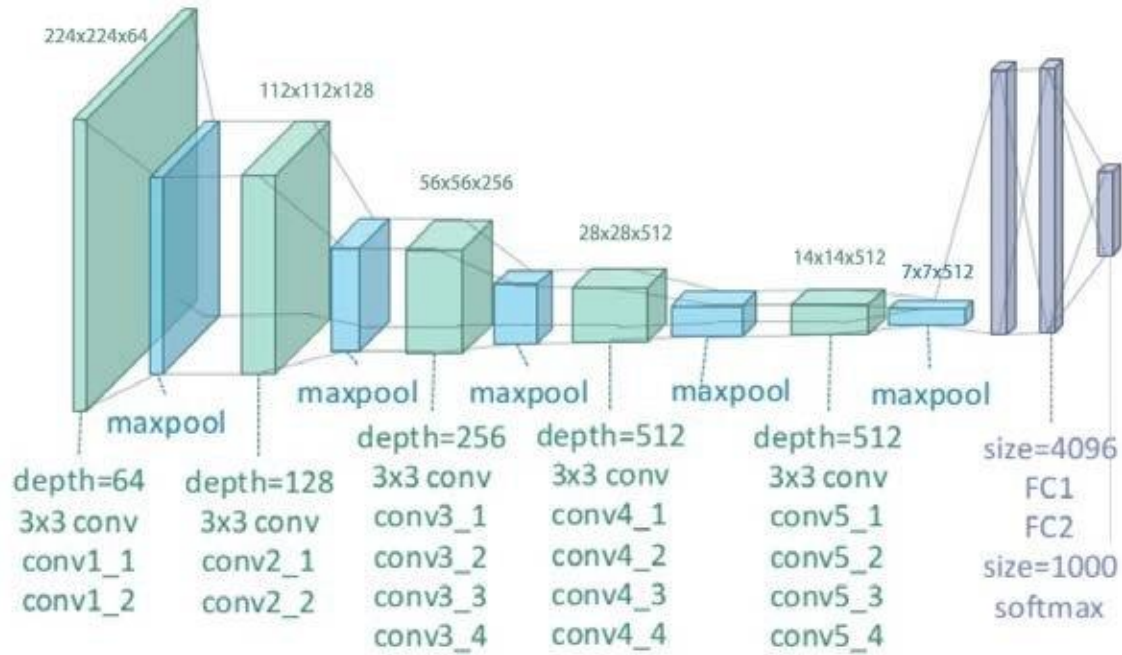


Figure 3.5: Illustration of The Network Architecture of Vgg-19 Model

This result popularized the use of CNNs networks in the area of computer vision. In 2014, the VGG-19 consisting of 19 layers was presented. The main contribution of VGG-19 was to empirically demonstrate that the number of layers (the depth of the network) is a critical component for good performance. In 2015, CNN Google Net was proposed. This network had as main objective to use the smallest number of parameters without losing depth, proposing the Inception module that significantly reduced the number of network parameters. Briefly, it can be seen that CNNs have been applied to computer vision tasks since the late 1980s. However, there were only a few scattered applications, and research was effectively inactive until the mid-2000s, when there was an increase in amount of labeled data, complemented by improved algorithms and the increasing computing power of GPUs that have contributed to its advancement and rapid progress since 2012. Currently, CNNs continue to be a relevant research subject, looking for new architectures and applications. A current summary on the topic can be found.

CNN is a well-known example of a deep learning architecture that was developed with the visual perception mechanism of living creatures as its primary source of inspiration. This resulted in the

creation of CNN as an example of a deep learning architecture. This design objective was met by basing CNN on the visual perception process of live things, which is how living things see the world. Beginning in the late 1960s, a number of major advancements were made in a variety of academic fields, which opened the way for it to become conceivable to conceive of the idea. Beginning in the late 1960s, a number of significant advances were made in a variety of academic fields. The concept of a hierarchical structure among the neurons that make up the visual system was proposed for the very first time in 1968 as a consequence of research that was conducted on the visual cortex of monkeys.

These investigations were carried out in the year 1968. According to this theory, there are certain cells in the visual system that are responsible for establishing a "receptive field" for other cells that are situated at a higher level. These cells can be found at lower levels. Receptive fields that are relatively simple can be combined in this way to produce receptive fields that are more complicated and have a wider range of use. 1980 marked the year that a recognition neural network was initially put on display publicly for the first time. This network makes use of the theories behind in the form of a network that is based on layers of neurons that utilize sparse connections, which is comparable to what we refer to as convolution in the modern day. The recognition was utilized with the intention of accomplishing the goal of recognizing handwritten numbers, with the primary qualities being the capacity to learn in an unsupervised manner and the requirement that it be invariant to the need that it be invariant to the requirement that it be invariant to the displacement of inputs. In 1989, LeCunn presented a proposal for a network design that incorporated both convolutional and fully connected layers. This idea was made public in the year 1989. The community of network administrators was given this proposition to consider. Backpropagation was used as a method for training such a network in order to improve its performance. Following training, the network was put to use to recognize handwritten numerals, and it did so with only a 5% error rate.

The network is comprised of four layers: two convolutional layers, two completely linked layers, and two layers that are fully connected to each other. An application is built in the year 1994 that makes use of a network that contains convolutional layers in order to tackle the problem of recognizing minute calcifications in digital mammograms. This was a problem that had been around since the year 1994. In addition to this, the writers had also been working simultaneously

on the development of a computer software that is capable of recognizing digits. The conventional model of CNNs was initially created in the year 1998. This model is comprised of networks of interleaved convolutional layers, max pooling operations for dimensionality reduction, and some fully connected layers for the final classification. This model was constructed using convolutional layer networks that were layered atop one another and interleaved. This article also dives into the famous LeNet-5 neural network as a topic of discussion, and it does so in some detail. It is judged against other handwritten digit recognition systems that have a higher degree of performance, and the winner is chosen based on the results of this evaluation. It was proposed in 2005 to use Graphics Processing Units (also known as GPUs) to perform computational calculations for applications related to machine learning. This came about as a result of the continued development of GPUs, also known as Graphics Processing Units, as well as the demand for graphics processing skills in 3D games.

This demand is a direct consequence of the high demand for people with skills in graphics processing in 3D games. This is because graphics processing units (GPUs) carry out highly optimized multiplications of matrices in order to carry out computations that are linked to three-dimensional geometry. The objective of these computations is to create images. In addition, graphics processing units, sometimes known as GPUs, have the ability to simultaneously manage a sizeable number of matrices. After making a few changes to the way the machine learning algorithms were implemented on GPUs, it was discovered that they were highly effective once these changes were done. In the year 2012, it was proposed that CNN AlexNet should be put into action. This network participated in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), and it ended up winning the competition by a very significant margin. Because of this, the error rate at the current state of the art dropped from 25.8 percent to 16.4 percent, which is a direct result of this.

4. PROPOSED METHOD

4.1 PREPROCESSING THE IMAGES

The pre-processing of the images was carried out through the use of a MATLAB program, the product of previous research [25] which reads the annotations of the overlays made by specialists on the lesions and cuts these lesions from the images to have images only of the area of interest. The operation of this program can be seen in figure below. This program also improves the contrast of the images through adaptive equalization [30] The images were also zero-centered, subtracting the mean of the pixels. These images were divided into disjoint subsets of training and validation with 85% and 15% of the total images respectively:

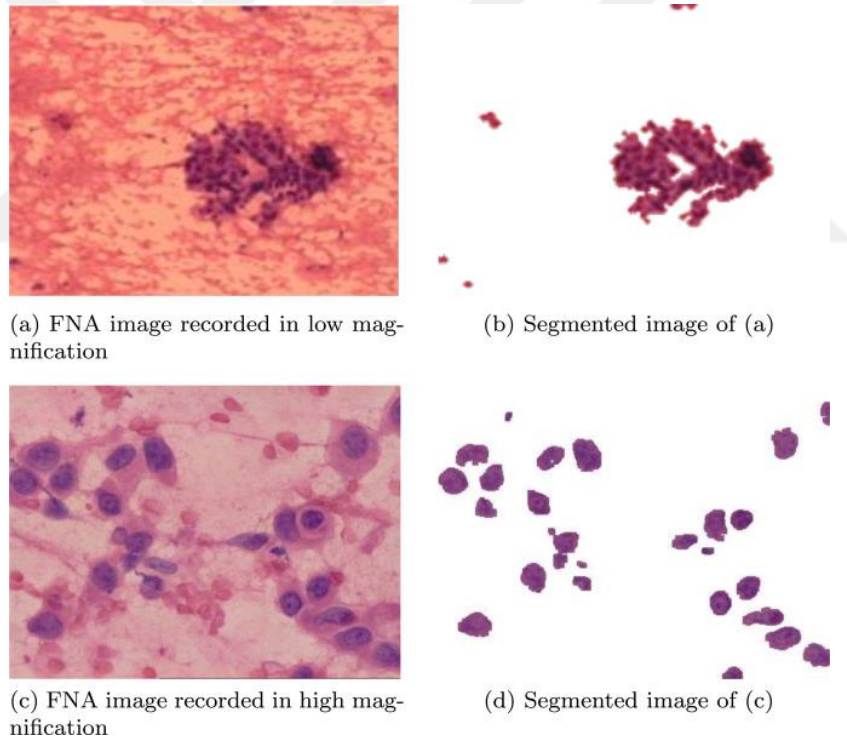


Figure 4.1: Choosing region of interest ROI in the pre-processing phase

4.2 SEGMENTATION

The process that has been suggested can be broken down into a number of steps, each of which contains technology for image processing at a different level than the one that came before it in the sequence. The first step of the process is called "image capture," and it involves retrieving

normal and abnormal mammograms from a database that is kept up to date by the Marathon Center for Statistical Education and Research

(<http://marathon.csee.usf.edu/Mammography/Database.html>).

This phase of the procedure is referred to as the initial phase of the procedure. Taking a picture is the first thing that you need to do in this process. Mammograms are going to be evaluated as part of the process's image analysis stage, which comes after the first step and is the second step overall. The name of a database that is utilized for the process of screening mammography is referred to as the Digital Database for Screening Mammography, or DDSM for short (DDSM). Following this step, the digitized mammograms undergo preprocessing with Gaussian filters so that the amount of noise that is visible in the pictures can be reduced. After the noise has been eliminated from the photographs, the following step is to treat them in such a way that the region of interest (ROI) that corresponds to the breast masses may be retrieved from the images that have been processed to remove the noise. The utilization of the strategy described above allows for the optimal growth of the region to be accomplished. After the regions of interest (ROIs) have been processed for the goal of feature extraction, the Gray level co-occurrence Matrix (GLCM) and the Gray level Run Length Matrix (GLRLM) are applied in order to recover a collection of textural features from the ROIs. These matrices are used concurrently with one another in a great number of different applications (GLRLM). Following this stage, the information is entered into a Feed Forward Neural Network (FFNN), which is then trained using a back propagation approach called LevenbergeMarquardt to classify mammograms as normal, benign, or malignant based on the textural features gathered. This step concludes the process. In order to achieve this objective, the mammograms that have been gathered are evaluated. A diagrammatic depiction of the methodology that has been provided may be obtained at this place in the form of Figure 1, which can be accessed at this address. The breast images that were obtained from patients and used in this recommended approach are digital images that were taken from), which was used in the research [14,15] and is available at <http://marathon.csee.usf.edu/Mammography/Database/Html.html>. The breast images that were obtained from patients and used in this recommended approach are digital images that were taken from. To be more specific, the breast photos that were received from patients and used in this recommended approach were digital photographs that were taken from). These are the digital photographs that were used to apply the method that was proposed to be used. It was

suggested that this method be implemented. The development of the DDSM data base was a collaborative effort between the Sandia National Laboratories, the University of South Florida, and the Massachusetts General Hospital. They collaborated in order to produce this database as a result of their efforts. The database now has a total of 2500 digitized mammograms for breast cancer screening purposes. These mammograms all came from various folks who went to the doctor to get them done. The Breast Imaging Reporting and Data System was used to collect patient information, and this information, together with the image data that was obtained from breast imaging procedures, is maintained in the database (BI-RADS). There are two categories of situations that have a severity degree: those that involve breast masses that are either ill-defined or well-defined. Spiculated masses, microcalcifications, and architectural faults are some of the other types of cases that might occur. For the objectives of teaching and testing from both the MLO and CC points of view, respectively, this study makes use of a set of three hundred digital mammograms with varied degrees of severity of normal, benign, and malignant lesions. The mammograms were taken using a digital imaging system. These mammograms show normal breast tissue in addition to benign lesions and cancerous lesions in the breast. There were 154 cases of breast cancer that were found to be benign, and there were 146 cases of breast cancer that were confirmed to be malignant in this collection. These tumors were found in the patient's left and right breasts, and they were found in both of the patient's breasts. In mammography, the background tissues can be broken down into three distinct types: dense tissues, granular tissues, and fatty tissues. These categories are given in the order that they are shown.

CNN is a well-known example of a deep learning architecture that was developed with the visual perception mechanism of living creatures as its primary source of inspiration. This resulted in the creation of CNN as an example of a deep learning architecture. This design objective was met by basing CNN on the visual perception process of live things, which is how living things see the world. Beginning in the late 1960s, a number of major advancements were made in a variety of academic fields, which opened the way for it to become conceivable to conceive of the idea. Beginning in the late 1960s, a number of significant advances were made in a variety of academic fields. The concept of a hierarchical structure among the neurons that make up the visual system was proposed for the very first time in 1968 as a consequence of research that was conducted on the visual cortex of monkeys.

These investigations were carried out in the year 1968. According to this theory, there are certain cells in the visual system that are responsible for establishing a "receptive field" for other cells that are situated at a higher level. These cells can be found at lower levels. Receptive fields that are relatively simple can be combined in this way to produce receptive fields that are more complicated and have a wider range of use. 1980 marked the year that a recognition neural network was initially put on display publicly for the first time. This network makes use of the theories behind in the form of a network that is based on layers of neurons that utilize sparse connections, which is comparable to what we refer to as convolution in the modern day. The recognition was utilized with the intention of accomplishing the goal of recognizing handwritten numbers, with the primary qualities being the capacity to learn in an unsupervised manner and the requirement that it be invariant to the need that it be invariant to the requirement that it be invariant to the displacement of inputs. In 1989, LeCunn presented a proposal for a network design that incorporated both convolutional and fully connected layers. This idea was made public in the year 1989. The community of network administrators was given this proposition to consider. Backpropagation was used as a method for training such a network in order to improve its performance. Following training, the network was put to use to recognize handwritten numerals, and it did so with only a 5% error rate.

The network is comprised of four layers: two convolutional layers, two completely linked layers, and two layers that are fully connected to each other. An application is built in the year 1994 that makes use of a network that contains convolutional layers in order to tackle the problem of

recognizing minute calcifications in digital mammograms. This was a problem that had been around since the year 1994. In addition to this, the writers had also been working simultaneously on the development of a computer software that is capable of recognizing digits. The conventional model of CNNs was initially created in the year 1998. This model is comprised of networks of interleaved convolutional layers, max pooling operations for dimensionality reduction, and some fully connected layers for the final classification. This model was constructed using convolutional layer networks that were layered atop one another and interleaved. This article also dives into the famous LeNet-5 neural network as a topic of discussion, and it does so in some detail. It is judged against other handwritten digit recognition systems that have a higher degree of performance, and the winner is chosen based on the results of this evaluation. It was proposed in 2005 to use Graphics Processing Units (also known as GPUs) to perform computational calculations for applications related to machine learning. This came about as a result of the continued development of GPUs, also known as Graphics Processing Units, as well as the demand for graphics processing skills in 3D games.

This demand is a direct consequence of the high demand for people with skills in graphics processing in 3D games. This is because graphics processing units (GPUs) carry out highly optimized multiplications of matrices in order to carry out computations that are linked to three-dimensional geometry. The objective of these computations is to create images. In addition, graphics processing units, sometimes known as GPUs, have the ability to simultaneously manage a sizeable number of matrices. After making a few changes to the way the machine learning algorithms were implemented on GPUs, it was discovered that they were highly effective once these changes were done. In the year 2012, it was proposed that CNN AlexNet should be put into action. This network participated in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), and it ended up winning the competition by a very significant margin. Because of this, the error rate at the current state of the art dropped from 25.8 percent to 16.4 percent, which is a direct result of this.

One of the solutions to the clustering problem that sees the most use is the partitioning method, which involves bringing together distinct components in order to form clusters. The majority of people who have attempted to solve the clustering problem have used this method at some point. Right off the outset, a decision will be made on the total number of unique groupings that will

eventually emerge from this process (n). After this step, the objects that will be used as the first set of n centers of groups that do not yet have any members will be selected. These groups have not yet been filled out with members. These organizations presently have no members in their ranks. After that, each of the m different things will be given a location within one of the groups, and once the group has been finished being assembled, one of the criteria that will be used to evaluate the quality of the group will be the degree of similarity that exists between the m different things that are a part of the group. It does not present the available sum distance adjustment work, consistently finding a model image of defined pixels, with the shape of one of the square-shaped groups, that is, of the differences between each pixel and the center of the group for which it was assigned are minimized (see Equation (1)). In other words, it does not find an image with defined pixels that have the shape of one of the square-shaped groups. To put it another way, it does not locate a picture with defined pixels that have the shape of one of the groups that are square-shaped. To express this another way, it does not locate a model image that has defined pixels that have the shape of one of the groups that are square-shaped. Where p_i is the vector of pixel i and g_j is the vector of the center of the group j ; p_{iv} is the value of the v attribute of pixel i ; g_{jv} is the value of the attribute v of the center of the group j , which means that this value is the average of the values of the attributes v of the pixels inserted in the group; w_{ij} is the weight associated with the value of p_i with g_j , and w_{ij} will receive the value that is determined by If the value that is obtained by the function that is given in Equation (1) is decreased, then the end goal of the segmented image will be improved. This may be done by lowering the value that is supplied in the equation. As a consequence of this, the pixel clusters will have a higher degree of similarity, which will result in an improvement to the segmented image. It is of the utmost importance to keep in mind that they are respected during the process of grouping, as every element that is going to be grouped must belong to a group, and every group must have at least one element in it. Keeping this in mind throughout the process of grouping is crucial. Keeping this in mind is necessary, as it is essential to remember that they are respected, and it is important to keep this in mind.

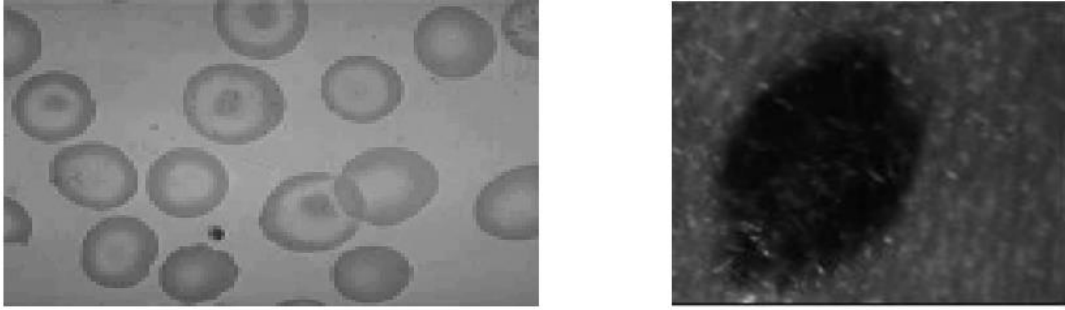


Figure 4.2: Lesion detection under the microscope

The ACO algorithm used in the present work to segment images is called ACOSI and it is based on the algorithm proposed by Kao and Chen (2006) for data clustering. A local search technique is applied together with ACOSI at the end of each solution built by an ant in an attempt to improve its quality. The search space in ACOSI contains p pixels and g groups. As in the work of Kao and Cheng (2006), at each step the artificial ants select a pixel not yet clustered and add it to a partial solution considering the pheromone intensity and the heuristic information. The latter indicates the desire to assign a pixel to a particular group and its value is obtained by calculating the inverse of the Euclidean distance between the current pixel and each existing group center. Each ant carries an array of cluster centers and updates it after each pixel is clustered. To select a group for a pixel i , one of two strategies can be adopted depending on the result.

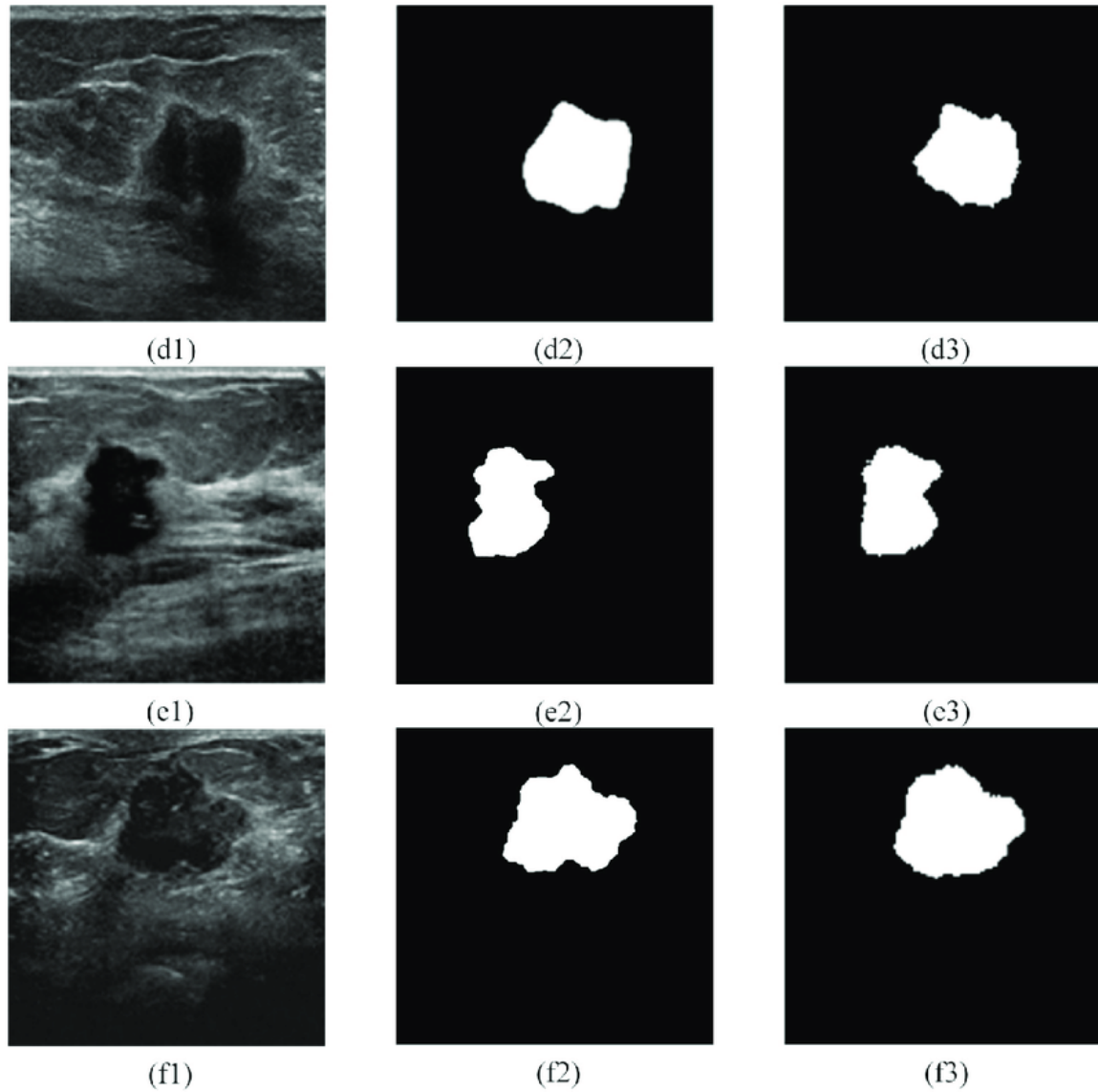


Figure 4.3: Segmentation of the lesions by the proposed algorithm.

4.3 DEEP LEARNING DETECTION

In this experiment, three different models of convolutional networks are utilized in order to meet the difficulty of cancer cell categorization in photos obtained from very fine needles. These photographs were taken using very small needles. The first model makes use of a genuine ResNet50 architecture and is made up of two neuron layers that are fully coupled with each other and with each other as well. These two layers have a total of 512 neurons shared between them, yet each layer only has four neurons of its own. A drop-out with a value of 0.5 was implemented in the region that formed the boundary between the third and fourth levels. The apprenticeship

rate, which was 5×10^{-4} , was characterized by a fall of 5×10^{-4} , which corresponded to the nesting velocity of 0.99. In other words, the apprenticeship rate fell at the same rate as the nesting velocity. The RGB color space must be used in any photos contained inside the files that are accepted as input by this model. These files must have a dimension of $250 \times 250 \times 3$. The training of the model took place over the course of 20 epochs with a total of 7892, and the batch size that was utilized was 258. In the beginning, we used a method that was completely at random in order to determine the weights. There are a total of four unique types of production, and they are referred to as benign mass, malignant mass, benign calcification, and malignant calcification, respectively. The second structure is composed of a total of seven levels, three of which are neuronal layers, four of which are coevolutionary layers, and one of which is an external layer. There are sixteen filters included in the first coevolutionary sheet, thirty-two filters included in the second sheet, sixty-four filters included in the third sheet, and one hundred and twenty-eight filters included in the fourth sheet. In order to provide output values that are consistent across layers, the 3×3 format must be applied by each layer, the PreLu features must be activated, and batch standardization must be utilized. This is done in order to maintain a steady level of output values throughout the process. The third, fourth, and fifth layers of fully-connected neural layers, respectively, each have a total of 64 neurons, 16 neurons, and two neurons. Throughout the course of the previous three levels, there was a drop-out that was equivalent to 0.5. The Nester impulse was configured to have a value of 0.99, the decay rate was set to 5×10^{-4} , and the learning rate that was employed was also 5×10^{-4} . This model's input consists of photographs with grayscale images that are 250 pixels on each side and 250 pixels tall. The dimensions of the pictures are as follows: In order to construct this model, 40,000 photographs were taken, and the batch size was multiplied by 258 times. The HeNormal strategy was the one that was applied when it came time to complete the initialization of the weights.

The methodology selected for evaluation is described by the following points:

- i. Creation of the dataset
- ii. Image preprocessing
- iii. Validation of results based on the accuracy metric

4.3.1 Creating the Dataset

This collection was put together using the FNA picture Digital Database serving as a point of departure for ideas and concepts. This collection consists of 3,946 pictures (Diagnostic Data Set: [https://archive.ics.uci.edu/ml/datasets/sheath%2BCancer%2BWisconsin%2B\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/sheath%2BCancer%2BWisconsin%2B(Diagnostic))), all of which were extracted from 2,620 distinct patients. The overlay file that goes along with each picture contains information that relates to the severity of the injury as well as its present state. There are 894 photographs that are benign, and 901 photographs that are malignant. There are 1,081 images that are benign, and there are 1,070 photographs that are malignant. The picture presents several illustrations that were derived from the database.:

object	Attributes									Class A ₁₀
	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉	
x ₁	10	5	5	6	8	8	7	1	1	Malignant
x ₂	7	5	10	10	10	10	4	10	3	Malignant
x ₃	10	10	10	10	3	10	10	6	1	Malignant
x ₄	7	2	4	1	6	10	5	4	3	Malignant
x ₅	8	10	10	1	3	6	3	9	1	Malignant
x ₆	8	10	5	3	8	4	4	10	3	Malignant
x ₇	6	10	2	8	10	2	7	8	10	Malignant
x ₈	10	8	10	10	6	1	3	1	10	Malignant
x ₉	7	8	7	2	4	8	3	8	2	Malignant
x ₁₀	8	8	9	6	6	3	10	10	1	Malignant
x ₁₁	1	1	1	1	2	1	2	1	1	Benign
x ₁₂	1	1	1	1	1	1	1	1	1	Benign
x ₁₃	3	2	2	1	4	3	2	1	1	Benign
x ₁₄	4	1	1	1	2	1	2	1	1	Benign
x ₁₅	5	1	1	1	2	1	3	1	1	Benign
x ₁₆	5	1	1	1	2	1	3	1	2	Benign
x ₁₇	3	1	1	1	2	1	1	1	1	Benign
x ₁₈	4	1	1	2	2	1	2	1	1	Benign
x ₁₉	1	1	1	1	2	1	1	1	1	Benign
x ₂₀	3	1	1	1	1	1	2	1	1	Benign
x ₂₁	1	1	1	1	2	1	2	1	1	Benign
x ₂₂	3	1	1	1	2	1	2	1	1	Benign
x ₂₃	2	1	1	1	1	1	2	1	1	Benign
x ₂₄	6	1	1	1	2	1	3	1	1	Benign
x ₂₅	2	1	2	1	2	1	3	1	1	Benign

Figure 4.4: Breast FNA images dataset from the university of Wisconsin

4.3.2 Data Validation

The accuracy measure of the model was applied to the validation data set so that the data could be checked for accuracy before being used. Because of this, the model was able to be applied to the data in order to validate them. While one examines model 1, they will notice in picture 4.5 that the curves that represent the training and validation sets get better as they move along the loss plot. This can be seen when moving along the plot from left to right. This is evident when one examines the curves in conjunction with the loss plot. In contrast to the loss training curve, which continues to exhibit a downward trend as indicated in the graph, the validation curve starts to show an upward trend at epoch 8, whilst the loss training curve continues to show an upward trend. The model is able to maintain its dependability due to the fact that it can correctly classify the photos that it was trained on; nevertheless, it is ineffective due to the fact that it cannot correctly identify the instances that are included in the validation set. Because of this, obtaining reliable results from the model is impossible. A clear illustration of the problem known as overfitting can be seen in this situation in the form of a discrepancy between the training curves and the validation curves.

4.3.3 Image Classification

In order for the network to learn, the back propagation learning approach is being applied to it. The network that is being employed is a multi-layered perceptual network. In addition to that, the network contains linkages that can travel in both the forward and backward directions, forward and backward. One layer of input neurons, one layer in the center with a mysterious function, and one layer of output neurons are all included in this architecture. There is a correlation between the total amount of weight and the number of neurons that extend their connections into the neurons that are adjacent to them. The value that is sent down from neuron i to neuron j is referred to as the weight W_{ij} . It has been proposed that the quantity obtained by subtracting w_{ij} from x_j , which stands for the value of neuron j in this context, should be utilized to represent the value of neuron j . It has been asserted that the weight that a neuron j takes in as sustenance is denoted by the symbol W_{ij} . In this approach, in order to train the network with the aid of the back propagation algorithm, it is required to identify the relative significance of the contributions made by each neuron in the network. This is done so that the network may be properly trained.

To get things rolling, a mechanism that is entirely based on chance is used to determine each of its weights. After the output layer has been finished and a network error measurement for the data that is being reviewed has been collected, the outcomes that were achieved are compared to the outcome that was anticipated to take place. This step takes place after the data has been reviewed. Since this error keeps occurring, the weights of everything will need to be recalculated to reflect the correct values. In a training set, the technique is only carried out once for each piece of data; after that, it is iterated an extraordinarily high number of times to ensure that the findings are getting gradually less significant in comparison to the data set.

In order to build upon the outcomes of our prior research, we worked on a variety of different classification algorithms in an effort to determine which ones were the most accurate and required the least amount of time. This was done with the intention of enhancing the results of our past research on the topic. As can be seen in the table that follows, we conducted an analysis to determine the parallels and differences that exist between the SVM algorithm, the Bayes algorithm, and the Bayes naive approach.

Table 4.1: Comparing SVM algorithm performance to other algorithms

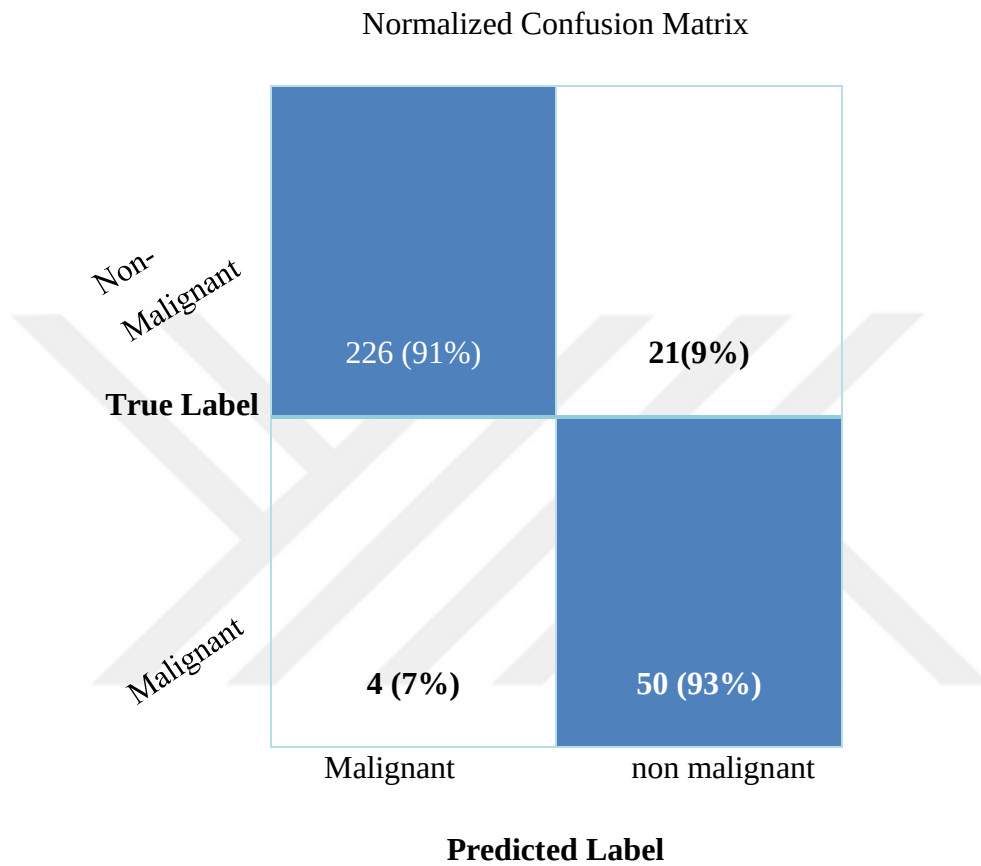
Type	SVM		BNT			LVQ	LP	BPN
Kernel	C-SVM	Nu-SVM	Naïve	K2	Bdeu			
Polynomial	97.54%	92.79%	89.55 %	88.41 %	91.31%	91.04 %	97.50 %	95.33% [17]
RBF	98.24%	95.08%						
Sigmoid	97.72%	93.85%						

Table 4.2: Time consumption for the execution by the SVM and NB algorithms.

Type of Network	Training/Prediction Time (seconds)
SVM	(2.11s)/0.94s
Bayesian Network Classifier	(4.27+1.51)s/0.07s
Learning Vector Quantization	67.18s/1.45

As seen in the figures above the SVM algorithm performed best in both terms of accuracy and time.

Table 4.3: Confusion matrix of true positive labels



5. CONCLUSION

This thesis is part of a larger goal of using bio-inspired technologies to improve breast cancer detection. At this point of a Computer-Aided Diagnosis (CAD) system, we concentrated on the categorization of breast masses into Benign and Malignant, using Artificial Immune Systems (AIS). There are various uses for this family of algorithms, and it has been able to establish itself within the family of evolutionary algorithms. We began by explaining the natural immune system and elucidating some of the key aspects that make computer modeling so useful. A state of the art was then tracked for three major AIS aspects: negative selection, immunological networks, and artificial clonal selection. Due to their ability to directly simulate the process of an adaptive immune response employing two crucial operators: cloning and mutation, we have concentrated our research on artificial clonal selection algorithms for breast cancer detection. An objective of this thesis was to contribute to pattern identification and optimization by giving professionals with a rapid and efficient tool to aid in the final choice. In the subject of clonal selection, the CLONALG algorithm is the most often cited method for pattern recognition. These observations pertain to the initialization process (which is similar to most clonal selection algorithms) and the stage in which memory cells are discarded and replaced with new ones produced randomly in order to preserve algorithmic variety. As a first step in dealing with the initialization issue, we offered a basic way for creating memory cells that are representative of all the learning data by averaging small subgroups of the latter. As a result, we've presented three methods to ensure that the algorithm doesn't reject potentially useful cells and introduces unpredictability into the algorithm (which risks presenting new prototypes that are worthless to the learning process). First, the Median Filter Clonal Algorithm (MF-CLONALG) generates potential median cells through the square matrix of the best cells closest to the example being trained, while the two other approaches. A validity interval is utilized in the VI-CS technique, which is an upgrade on ACCLONALG, to pick the cells that will join the final set of memory cells used in the categorization. One of our findings is that these strategies have improved CLONALG performance significantly, albeit at a significant computational expense. If increasing the number and efficiency of memory cells allows for more accuracy and variety to be maintained, then this comes at the cost of an extended learning period. As a result of this, we've developed a new method for speeding up training time while yet maintaining the clonal selection

algorithms' accuracy. When dealing with time-consuming test functions, this technique employs a simple mechanism that limits the number of such assessments (done by each antigen to be learnt) while still maintaining the precision of the AIS. Our goal was to build a speedy and efficient method for pattern identification by clonal selection for the categorization of breast masses, and to assist in the detection of breast cancer, with this effort. Second, multimodal optimization is addressed in the thesis. Artificial clonal selection (AIS) is often recommended for optimizing activities, and this is especially true when it comes to AIS. As a result, we're eager to learn more about these algorithms and the potential benefits of cloning and mutation operators for optimizing functions.

These files must have a dimension of 250x250x3. The training of the model took place over the course of 20 epochs with a total of 7892, and the batch size that was utilized was 258. In the beginning, we used a method that was completely at random in order to determine the weights. There are a total of four unique types of production, and they are referred to as benign mass, malignant mass, benign calcification, and malignant calcification, respectively. The second structure is composed of a total of seven levels, three of which are neuronal layers, four of which are coevolutionary layers, and one of which is an external layer. There are sixteen filters included in the first coevolutionary sheet, thirty-two filters included in the second sheet, sixty-four filters included in the third sheet, and one hundred and twenty-eight filters included in the fourth sheet. In order to provide output values that are consistent across layers, the 3x3 format must be applied by each layer, the PreLu features must be activated, and batch standardization must be utilized. This is done in order to maintain a steady level of output values throughout the process. The third, fourth, and fifth layers of fully-connected neural layers, respectively, each have a total of 64 neurons, 16 neurons, and two neurons. Throughout the course of the previous three levels, there was a drop-out that was equivalent to 0.5. The Nester impulse was configured to have a value of 0.99, the decay rate was set to 5×10^{-4} , and the learning rate that was employed was also 5×10^{-4} . This model's input consists of photographs with grayscale images that are 250 pixels on each side and 250 pixels tall. The dimensions of the pictures are as follows: In order to construct this model, 40,000 photographs were taken, and the batch size was multiplied by 258 times. The HeNormal strategy was the one that was applied when it came time to complete the initialization of the weights

REFERENCES

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA Cancer J Clin*, vol. 66, pp.7-30, 2016.
- [2] Poorolajal, J., et al., Breast cancer screening (BCS) chart: a basic and preliminary model for making screening mammography more productive and efficient. *J Public Health (Oxf)*, 2017: p. 1-8.
- [3] Z. Jiao, X. GAO, Y. Wang, and J. Li, "A deep feature based framework for breast masses classification," *Journal of Neurocomputing*, vol. 197, pp. 221-231, 2016.
- [4] J. Arevalo, F. A. González, R. Ramos-Pollán, J. Oliveira, and M. A. Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Comput. Methods Program Biomed*, vol. 127, pp. 248-257, 2016.
- [5] Deng, L., Yu, D.: *Deep Learning: Methods and Applications*. Now publishers, Boston (2014).
- [6] Krizhevsky, A., I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 2012.
- [7] Zhang, Y., et al. DeepSplice: Deep classification of novel splice junctions revealed by RNA-seq. in *Bioinformatics and Biomedicine (BIBM)*, 2016 IEEE International Conference on. 2016. IEEE
- [8] Hua, K.-L., et al., Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and therapy*, 2015.
- [9] Pravin S. Mane, Indra Gupta, M. K. Vasantha, "Implementation of RISC Processor on FPGA", *Electrical Engineering Department, Indian Institute of Technology Roorkee*, 2006 IEEE.
- [10] Kui YI, Yue-Hua DING, "32-bit RISC CPU Based on MIPS", *International Joint Conference on Artificial Intelligence* 2019.
- [11] Mrs. Rupali S. Balpande, Mrs. Rashmi S. Keote, "Design of FPGA based Instruction Fetch and Decode Module of 32-bit RISC (MIPS) Processor", *International Conference on Communication Systems and Network echnologies*, 2011.

- [12] Mr. S. P. Ritpurkar, Prof. M. N. Thakare, Prof. G. D. Korde, "Synthesis and Simulation of a 32Bit MIPS RISC Processor using VHDL", IEEE International Conference on Advances in Engineering & Technology Research, 2014.
- [13] David A. Patterson, John L. Hennessy, "Computer Organization and Design", the Hardware and Software Interface, Third Edition.
- [14] J. C. Tobias Charistian Cahoon, Melanie A.Sutton, "Three-class mam mogram classification based on descriptive cnn features," 2020.
- [15] S. Sharma, M. Kharbanda, and G. Kaushal, "Brain tumor and breast cancer detection using medical images," International Journal of Engi-neering Technology Science and Research, vol. 2, 2015.
- [16] P. Hankare, K. Shah, D. Nair, and D. Nair, "Breast cancer detection using thermography," Int. Res. J. Eng. Technol, vol. 4, 2016.
- [17] S. Hwang and H.-E. Kim, "Self-transfer learning for fully weakly supervised object localization," arXiv preprint arXiv: 1602.01625, 2016.
- [18] Z. Mohammadzadeh, R. Safdari, M. Ghazisaeidi, S. Davoodi, and Z. Azadmanjir, "Advances in optimal detection of cancer by image processing; experience with lung and breast cancers," Asian Pacific journal of cancer prevention: APJCP, vol. 16, no. 14, 2015.
- [19] Ding, J.; Cheng, H.-D.; Huang, J.; Liu, J.; Zhang, Y. Breast ultrasound image classification based on multiple-instance learning. J. Digit. Imaging 2012, 25, 620–627. [CrossRef] [PubMed]
- [20] Bing, L.; Wang, W. Sparse representation based multi-instance learning for breast ultrasound image classification. Comput. Math. Methods Med. 2017, 2017, 7894705. [CrossRef]
- [21] Prabhakar, T.; Poonguzhali, S. Automatic detection and classification of benign and malignant lesions in breast ultrasound images using texture morphological and fractal features. In Proceedings of the 2017 10th Biomedical Engineering International Conference (BMEiCON), Hokkaido, Japan, 31 August–2 September 2017; IEEE: New York, NY, USA, 2017; pp. 1–5.
- [22] Zhang, Q.; Suo, J.; Chang, W.; Shi, J.; Chen, M. Dual-modal computer-assisted evaluation of axillary lymph node metastasis in breast cancer patients on both real-time elastography and B-mode ultrasound. Eur. J. Radiol. 2017, 95, 66–74. [CrossRef]

- [23] GAO, Y.; Geras, K.J.; Lewin, A.A.; Moy, L. New frontiers: An update on computer-aided diagnosis for breast imaging in the age of artificial intelligence. *Am. J. Roentgenol.* 2019, 212, 300–307. [CrossRef]
- [24] Geras, K.J.; Mann, R.M.; Moy, L. Artificial intelligence for mammography and digital breast tomosynthesis: Current concepts and future perspectives. *Radiology* 2019, 293, 246–259. [CrossRef]
- [25] Fujioka, T.; Mori, M.; Kubota, K.; Oyama, J.; Yamaga, E.; Yashima, Y.; Katsuta, L.; Nomura, K.; Nara, M.; Oda, G. The utility of deep learning in breast ultrasonic imaging: A review. *Diagnostics* 2020, 10, 1055. [CrossRef] [PubMed]
- [26] Zahoor, S.; Lali, I.U.; Javed, K.; Mehmood, W. Breast cancer detection and classification using traditional computer vision techniques: A comprehensive review. *Curr. Med. Imaging* 2020, 16, 1187–1200. [CrossRef] [PubMed]
- [27] Kadry, S.; Rajinikanth, V.; Taniar, D.; Damaševičius, R.; Valencia, X.P.B. Automated segmentation of leukocyte from hematological images—A study using various CNN schemes. *J. Supercomput.* 2021, 1–21. [CrossRef]
- [28] Abayomi-Alli, O.O.; Damaševičius, R.; Misra, S.; Maskeliunas, R.; Abayomi-Alli, A. Malignant skin melanoma detection using γ image augmentation by oversampling in nonlinear lower-dimensional embedding manifold. *Turk. J. Electr. Eng. Comput. Sci.* 2021, 29, 2600–2614. [CrossRef]
- [29] Maqsood, S.; Damaševičius, R.; Maskeliunas, R. Hemorrhage detection based on 3d cnn deep learning framework and feature γ fusion for evaluating retinal abnormality in diabetic patients. *Sensors* 2021, 21, 3865. [CrossRef]
- [30] Hussain, N.; Kadry, S.; Tariq, U.; Mostafa, R.R.; Choi, J.-I.; Nam, Y. Intelligent Deep Learning and Improved Whale Optimization Algorithm Based Framework for Object Recognition. *Hum. Cent. Comput. Inf. Sci.* 2021, 11, 34.
- [31] . Kadry, S.; Parwekar, P.; Damaševičius, R.; Mehmood, A.; Khan, J.A.; Naqvi, S.R.; Khan, M.A. Human gait analysis for osteoarthritis prediction: A framework of deep learning and kernel extreme learning machine. *Complex Intell. Syst.* 2021, 1–19. [CrossRef]
- [32] Dhungel, N.; Carneiro, G.; Bradley, A.P. The automated learning of deep features for breast mass classification from mammograms. In *Proceedings of the International*

- Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; Springer: Cham, Switzerland, 2016; pp. 106–114.
- [33] Alhaisoni, M.; Tariq, U.; Hussain, N.; Majid, A.; Damaševičius, R.; Maskeliunas, R. COVID-19 Case Recognition from Chest CT – Images by Deep Learning, Entropy-Controlled Firefly Optimization, and Parallel Feature Fusion. *Sensors* 2021, 21, 7286.
- [34] Odusami, M.; Maskeliunas, R.; Damaševičius, R.; Krilavičius, T. Analysis of features of alzheimer’s disease: Detection of early – stage from functional brain changes in magnetic resonance images using a finetuned resnet18 network. *Diagnostics* 2021, 11, 1071. [CrossRef]
- [35] Nawaz, M.; Nazir, T.; Masood, M.; Mehmood, A.; Mahum, R.; Kadry, S.; Thinnukool, O. Analysis of Brain MRI Images Using Improved CornerNet Approach. *Diagnostics* 2021, 11, 1856. [CrossRef]
- [36] Farzaneh, N.; Williamson, C.A.; Jiang, C.; Srinivasan, A.; Bapuraj, J.R.; Gryak, J.; Najarian, K.; Soroushmehr, S. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries. *Diagnostics* 2020, 10, 773. [CrossRef] [PubMed]
- [37] Meng, L.; Zhang, Q.; Bu, S. Two-Stage Liver and Tumor Segmentation Algorithm Based on Convolutional Neural Network. *Diagnostics* 2021, 11, 1806. [CrossRef] [PubMed]
- [38] Khaldi, Y.; Benzaoui, A.; Ouahabi, A.; Jacques, S.; Taleb-Ahmed, A. Ear recognition based on deep unsupervised active learning. *IEEE Sens. J.* 2021, 21, 20704–20713. [CrossRef]
- [39] Majid, A.; Nam, Y.; Tariq, U.; Roy, S.; Mostafa, R.R.; Sakr, R.H. COVID19 classification using CT images via ensembles of deep learning models. *Comput. Mater. Contin.* 2021, 69, 319–337. [CrossRef]
- [40] Sharif, M.I.; Alhussein, M.; Aurangzeb, K.; Raza, M. A decision support system for multimodal brain tumor classification using deep learning. *Complex Intell. Syst.* 2021, 1–14. [CrossRef]
- [41] Liu, D.; Liu, Y.; Li, S.; Li, W.; Wang, L. Fusion of handcrafted and deep features for medical image classification. *J. Phys. Conf. Ser.* 2019, 1345, 022052. [CrossRef]

- [42] Alinsaif, S.; Lang, J.; Alzheimer's Disease Neuroimaging Initiative. 3D shearlet-based descriptors combined with deep features for the classification of Alzheimer's disease based on MRI data. *Comput. Biol. Med.* 2021, 138, 104879. [CrossRef]
- [43] Khan, M.A.; Muhammad, K.; Sharif, M.; Akram, T.; de Albuquerque, V.H.C. Multi-Class Skin Lesion Detection and Classification via Teledermatology. *IEEE J. Biomed. Health Inform.* 2021, 25, 4267–4275. [CrossRef].
- [44] Masud, M.; Rashed, A.E.E.; Hossain, M.S. Convolutional neural network-based models for diagnosis of breast cancer. *Neural Comput. Appl.* 2020, 1–12. [CrossRef]
- [45] Jiménez-Gaona, Y.; Rodríguez-Álvarez, M.J.; Lakshminarayanan, V. Deep-Learning-Based Computer-Aided Systems for Breast Cancer Imaging: A Critical Review. *Appl. Sci.* 2020, 10, 8298. [CrossRef]
- [46] Zeebaree, D.Q. A Review on Region of Interest Segmentation Based on Clustering Techniques for Breast Cancer Ultrasound Images. *J. Appl. Sci. Technol. Trends* 2020, 1, 78–91.
- [47] Huang, K.; Zhang, Y.; Cheng, H.; Xing, P. Shape-adaptive convolutional operator for breast ultrasound image segmentation. In *Proceedings of the 2021 IEEE International Conference on Multimedia and Expo (ICME)*, Shenzhen, China, 5–9 July 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.
- [48] Sadad, T.; Hussain, A.; Munir, A.; Habib, M.; Ali Khan, S.; Hussain, S.; Yang, S.; Alawairdhi, M. Identification of breast malignancy by marker-controlled watershed transformation and hybrid feature set for healthcare. *Appl. Sci.* 2020, 10, 1900. [CrossRef]
- [49] Mishra, A.K.; Roy, P.; Bandyopadhyay, S.; Das, S.K. Breast ultrasound tumour classification: A Machine Learning—Radiomics based approach. *Expert Syst.* 2021, 38, e12713. [CrossRef]
- [50] Hussain, S.; Xi, X.; Ullah, I.; Wu, Y.; Ren, C.; Lianzheng, Z.; Tian, C.; Yin, Y. Contextual level-set method for breast tumor segmentation. *IEEE Access* 2020, 8, 189343–189353. [CrossRef]
- [51] Xiangmin, H.; Jun, W.; Weijun, Z.; Cai, C.; Shihui, Y.; Jun, S. Deep Doubly Supervised Transfer Network for Diagnosis of Breast Cancer with Imbalanced Ultrasound Imaging Modalities. *arXiv* 2020, arXiv:2007.066342020.

- [52] Moon, W.K.; Lee, Y.-W.; Ke, H.-H.; Lee, S.H.; Huang, C.-S.; Chang, R.-F. Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. *Comput. Methods Programs Biomed.* 2020, 190, 105361. [CrossRef]
- [53] Byra, M.; Jarosik, P.; Szubert, A.; Galperin, M.; Ojeda-Fournier, H.; Olson, L.; O’Boyle, M.; Comstock, C.; Andre, M. Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network. *Biomed. Signal Process. Control* 2020, 61, 102027. [CrossRef]
- [54] Kadry, S.; Damaševičius, R.; Taniar, D.; Rajinikanth, V.; Lawal, I.A. Extraction of tumour in breast MRI using joint thresholding and segmentation—A study. In *Proceedings of the 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, 25–27 March 2021; IEEE: New York, NY, USA, 2021; pp. 1–5.
- [55] Lahoura, V.; Singh, H.; Aggarwal, A.; Sharma, B.; Mohammed, M.; Damaševičius, R.; Kadry, S.; Cengiz, K. Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. *Diagnostics* 2021, 11, 241. [CrossRef].
- [56] Maqsood, S.; Damasevicius, R.; Shah, F.M. An Efficient Approach for the Detection of Brain Tumor Using Fuzzy Logic and U-NET CNN Classification, *International Conference on Computational Science and Its Applications*; Springer: Cham, Switzerland, 2021; pp. 105–118.
- [57] Rajinikanth, V.; Kadry, S.; Taniar, D.; Damasevicius, R.; Rauf, H.T. Breast-cancer detection using thermal images with marinepredators-algorithm selected features. In *Proceedings of the 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Noida, India, 26–27 August 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.
- [58] P. Hankare, K. Shah, D. Nair, and D. Nair, “Breast cancer detection using thermography,” *Int. Res. J. Eng. Technol*, vol. 4, 2016.
- [59] S. Hwang and H.-E. Kim, “Self-transfer learning for fully weakly supervised object localization,” *arXiv preprint arXiv: 1602.01625*, 2016.
- [60] Z. Mohammadzadeh, R. Safdari, M. Ghazisaeidi, S. Davoodi, and Z. Azadmanjir, “Advances in optimal detection of cancer by image processing; experience with lung and breast cancers,” *Asian Pacific journal of cancer prevention: APJCP*, vol. 16, no. 14, 2015.

- [61] Ding, J.; Cheng, H.-D.; Huang, J.; Liu, J.; Zhang, Y. Breast ultrasound image classification based on multiple-instance learning. *J. Digit. Imaging* 2012, 25, 620–627. [CrossRef] [PubMed]
- [62] Bing, L.; Wang, W. Sparse representation based multi-instance learning for breast ultrasound image classification. *Comput. Math. Methods Med.* 2017, 2017, 7894705. [CrossRef]
- [63] Prabhakar, T.; Poonguzhali, S. Automatic detection and classification of benign and malignant lesions in breast ultrasound images using texture morphological and fractal features. In *Proceedings of the 2017 10th Biomedical Engineering International Conference (BMEiCON)*, Hokkaido, Japan, 31 August–2 September 2017; IEEE: New York, NY, USA, 2017; pp. 1–5.
- [64] Zhang, Q.; Suo, J.; Chang, W.; Shi, J.; Chen, M. Dual-modal computer-assisted evaluation of axillary lymph node metastasis in breast cancer patients on both real-time elastography and B-mode ultrasound. *Eur. J. Radiol.* 2017, 95, 66–74. [CrossRef]
- [65] GAO, Y.; Geras, K.J.; Lewin, A.A.; Moy, L. New frontiers: An update on computer-aided diagnosis for breast imaging in the age of artificial intelligence. *Am. J. Roentgenol.* 2019, 212, 300–307. [CrossRef]
- [66] Geras, K.J.; Mann, R.M.; Moy, L. Artificial intelligence for mammography and digital breast tomosynthesis: Current concepts and future perspectives. *Radiology* 2019, 293, 246–259. [CrossRef]
- [67] Fujioka, T.; Mori, M.; Kubota, K.; Oyama, J.; Yamaga, E.; Yashima, Y.; Katsuta, L.; Nomura, K.; Nara, M.; Oda, G. The utility of deep learning in breast ultrasonic imaging: A review. *Diagnostics* 2020, 10, 1055. [CrossRef] [PubMed]
- [68] Zahoor, S.; Lali, I.U.; Javed, K.; Mehmood, W. Breast cancer detection and classification using traditional computer vision techniques: A comprehensive review. *Curr. Med. Imaging* 2020, 16, 1187–1200. [CrossRef] [PubMed]
- [69] Kadry, S.; Rajinikanth, V.; Taniar, D.; Damaševičius, R.; Valencia, X.P.B. Automated segmentation of leukocyte from hematological images—A study using various CNN schemes. *J. Supercomput.* 2021, 1–21. [CrossRef]
- [70] Abayomi-Alli, O.O.; Damaševičius, R.; Misra, S.; Maskeliunas, R.; Abayomi-Alli, A. Malignant skin melanoma detection using γ image augmentation by oversampling in

- nonlinear lower-dimensional embedding manifold. *Turk. J. Electr. Eng. Comput. Sci.* 2021, 29, 2600–2614. [CrossRef]
- [71] Maqsood, S.; Damaševičius, R.; Maskeliunas, R. Hemorrhage detection based on 3d cnn deep learning framework and feature fusion for evaluating retinal abnormality in diabetic patients. *Sensors* 2021, 21, 3865. [CrossRef]
- [72] Hussain, N.; Kadry, S.; Tariq, U.; Mostafa, R.R.; Choi, J.-I.; Nam, Y. Intelligent Deep Learning and Improved Whale Optimization Algorithm Based Framework for Object Recognition. *Hum. Cent. Comput. Inf. Sci.* 2021, 11, 34.
- [73] . Kadry, S.; Parwekar, P.; Damaševičius, R.; Mehmood, A.; Khan, J.A.; Naqvi, S.R.; Khan, M.A. Human gait analysis for osteoarthritis prediction: A framework of deep learning and kernel extreme learning machine. *Complex Intell. Syst.* 2021, 1–19. [CrossRef]
- [74] Dhungel, N.; Carneiro, G.; Bradley, A.P. The automated learning of deep features for breast mass classification from mammograms. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Athens, Greece, 17–21 October 2016; Springer: Cham, Switzerland, 2016; pp. 106–114.
- [75] Alhaisoni, M.; Tariq, U.; Hussain, N.; Majid, A.; Damaševičius, R.; Maskeliunas, R. COVID-19 Case Recognition from Chest CT Images by Deep Learning, Entropy-Controlled Firefly Optimization, and Parallel Feature Fusion. *Sensors* 2021, 21, 7286.
- [76] Odusami, M.; Maskeliunas, R.; Damaševičius, R.; Krilavičius, T. Analysis of features of alzheimer’s disease: Detection of early stage from functional brain changes in magnetic resonance images using a finetuned resnet18 network. *Diagnostics* 2021, 11, 1071. [CrossRef]
- [77] Nawaz, M.; Nazir, T.; Masood, M.; Mehmood, A.; Mahum, R.; Kadry, S.; Thinnukool, O. Analysis of Brain MRI Images Using Improved CornerNet Approach. *Diagnostics* 2021, 11, 1856. [CrossRef]
- [78] Farzaneh, N.; Williamson, C.A.; Jiang, C.; Srinivasan, A.; Bapuraj, J.R.; Gryak, J.; Najarian, K.; Soroushmehr, S. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries. *Diagnostics* 2020, 10, 773. [CrossRef] [PubMed]

- [79] Meng, L.; Zhang, Q.; Bu, S. Two-Stage Liver and Tumor Segmentation Algorithm Based on Convolutional Neural Network. *Diagnostics* 2021, 11, 1806. [CrossRef] [PubMed]
- [80] Khaldi, Y.; Benzaoui, A.; Ouahabi, A.; Jacques, S.; Taleb-Ahmed, A. Ear recognition based on deep unsupervised active learning. *IEEE Sens. J.* 2021, 21, 20704–20713. [CrossRef]
- [81] Majid, A.; Nam, Y.; Tariq, U.; Roy, S.; Mostafa, R.R.; Sakr, R.H. COVID19 classification using CT images via ensembles of deep learning models. *Comput. Mater. Contin.* 2021, 69, 319–337. [CrossRef]
- [82] Sharif, M.I.; Alhussein, M.; Aurangzeb, K.; Raza, M. A decision support system for multimodal brain tumor classification using deep learning. *Complex Intell. Syst.* 2021, 1–14. [CrossRef]