Diagnosis of breast cancer using image processing with SVM and KNN

Zahra abdolali kazemi,

Department of (Electrical engineering),
Islamic Azad University, South Tehran Branch

Abstract

Objective:

Breast cancer accounts for 19% of deaths caused by cancers and 24% of all cases afflicted to cancers in European countries. Approximately 25% of breast cancer deaths happen in women between the ages of 40 and 49. Anomaly detection is done by the separation of the initial steps in computer detection systems. In this chapter, while reviewing the different techniques, a qualitative comparison between them will be provided. In this study, two approaches for the presentation of mammography (comparison of previous and current mammography images) are evaluated: together (simultaneously) and alternately on the same screen.

Methods:

In this study, MATLAB software is used.In this study, image processing algorithms of support vector machine (SVM), genetic algorithm (GA), convolutional neural networks (CNN), and K-nearest neighbours (KNN) are exploited. In this regard, the performance of these algorithms will be explained in this section.

Results:

In this method, it is first essential to conduct training. Training means that a number of features related to class one and class two are given to the function, and the algorithm updates its parameters based on the labeling done. Then, the unlabeled data are given to the algorithm for the classification, and it automatically specifies the corresponding class. the Segmentation is the simplification or modification of image view for more meaningful and easier analysis. this is the process of labeling each pixel in each image, which results in a set of segments that together cover the whole image. By analyzing the resulting images, the physicians can identify cancer cells and offer their diagnostic results.

Conclusion:

It is possible to expand the MATLAB environment by adding a toolbox for various purposes.For simulation, training and classification need to be done with the classification method.

Introduction:

Currently, cancer is one of the major and essential health issues in Iran and worldwide. The incidence of breast cancer among women has increased during the past four decades in Asian countries such as Iran, and it is known as one of the most common malignancies among Iranian women (1). Breast cancer is rare in women younger than 33 years, and its prevalence is higher after the age of 33. In Iran, breast cancer occurs in women a decade earlier than in developed countries (2). The age of affliction to the disease is low in Iran, and the mean age of patients in the lower range relative to the whole world exhibits a lower value (3). Breast cancer is a malignancy of the main breast tissue. Indeed, breast tissue cells, which make milk for feeding babies, become malignant. According to the National Cancer Institute (NCI) of the United States, one out of every eight women will be diagnosed with breast cancer in their lifetime (4). Despite breast cancer at an earlier age in recent years and the presence of dense breast tissue in this group and the likelihood to remain hidden the lesion in this type of tissue, a supplemental screening method to increase the detection sensitivity seems to be crucial. It is found that the sensitivity of mammography is strongly influenced by breast tissue density, and the sensitivity of mammography decreases with increasing breast density such that mammography sensitivity in women with dense breast may be reduced by 30% (5). Breast tissue is dense in more than half of women younger than 50 years (51 to 75) and at least one-third of women over 50 years; this is while this group is at the exposure of risk for a longer period and has a worse clinical prognosis (6).

Unfortunately, the incidence of cancer is rising sharply in some countries (5). In 2003, studies by Smith et al. revealed that early detection of breast cancer would prevent premature death in patients. Early detection of this cancer is one of the most significant challenges for public health (8). Hence, many studies have been carried out for the diagnosis and treatment of the disease as soon as possible. However, in many cases, breast cancer is not diagnosed until reaching an advanced stage. It is therefore associated with the regrettable statistics of survival of people with breast cancer. The ever-rising trend of cancer in the country, particularly breast cancer, raises the necessity for information to promote public awareness, early detection, prevention strategies, and treatment management of the illness. One of the most important and influential ways to diagnose this cancer, especially in the early stages of the disease, is mammography. A mammogram is an X-ray examination of the breast in which low-energy X-rays are used. Interest in breast imaging has grown with the recognition that approximately one-eighth of women will experience breast cancer during their lifetime. Compared to other diagnostic methods such as ultrasound and Magnetic Resonance Imaging (MRI1), mammography is the most accurate method in the early detection of cancer. Since dense breasts per se are a risk factor for breast cancer due to the density of glandular tissue, so ultrasound in addition to mammography, especially in high-risk women and younger ages (less than 49 years) seems to be crucial.

Statement of the Problem:

Advanced breast cancer is among the most common cancers among women worldwide. In the Western world, 10% of women with breast cancer live. A study in 2003 by the American Cancer Society estimated that one out of eight people in the United States is afflicted to breast cancer during their lifetime. In this chapter, an overview of different mammography mass separation methods, description of the main features, and the differences between them will be discussed, the goal of which is the advantages and disadvantages of these approaches.Mammograms were displayed at the dedicated mammography workstation. Stimulation arises using the extraction of lesions that are normal on mammography. The desired image is selected for each stimulus, which contains the image of the largest lesion. The threshold for detecting lesion growth is determined in both conditions. Analysis interval changes are crucial for the diagnosis of lesions and classification of benign and malignant masses. It was demonstrated that observers were more accurate in diagnosing the lesion when the lesion growth was alternately displayed on a screen.

Although the difference in performance between the display methods is significant, the change in threshold detection size may seem relatively small. However, even a tiny change in size can be critical in detection. Since mammography screening has been beneficial in detecting growing tumors, image processing is part of the field of electrical and computer engineering, which has impressive growth with the computer revolution.

One of the dangers that severely threatens women is breast cancer. The chances of people for a healthy and normal life will be higher in the case the tumor can be diagnosed in the breast. Hence, the construction of systems to detect this type of cancer, like detection systems in other medical areas, is considered crucial.

A) Genetic Algorithm (GA)

The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that are based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution. You can apply the genetic algorithm to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, nondifferentiable, stochastic, or highly nonlinear. The genetic algorithm can address problems of mixed-integer programming, where some components are restricted to be integer-valued.

The genetic algorithm uses three main types of rules at each step to create the next generation from the current population:

* Selection rules select the individuals, called parents that contribute to the population at the next generation.

* Crossover rules combine two parents to form children for the next generation.

* Mutation rules apply random changes to individual parents to form children.

The genetic algorithm differs from a classical, derivative-based, optimization algorithm in two main ways, as summarized in the following table.

Classical Algorithm	Genetic Algorithm
Generates a single point at each iteration. The sequence of points approaches an optimal solution.	Generates a population of points at each iteration. The best point in the population approaches an optimal solution.
Selects the next point in the sequence by a deterministic computation.	Selects the next population by computation which uses random number generators.

1-1-1- History of Genetic Algorithm:

Genetic algorithms were developed to mimic some of the processes observed in the natural evolution of living organisms. The fact that life, along with its complexities, has evolved over a relatively short period of time, estimated from fossils, has puzzled many people, including biologists. The main idea of Genetic Algorithm (GA) is to utilize the power of the evolutionary process to solve optimization problems. The father of the original genetic algorithm was John Holland, who invented it in the early 1970s.

The notion of Genetic Algorithm (GA) is an adaptive heuristic search algorithm designed based on the evolutionary ideas of natural selection and genetics. Thus, this algorithm represents the intelligent use of a random search algorithm for solving optimization problems. Although GAs exploit random phenomena, they are no random at all; instead, they employ the available historical information to direct search operations to an area with better performance in the search space. The main techniques of GA are designed such that they could simulate the processes critical for the evolution of natural systems. Among the most important of these processes are those that follow the rules first introduced by Charles Darwin's "survival of the fittest". This is since, in nature, competition between living organisms for attaining scarce resources leads to the dominance of the fittest beings over weaker beings.

Why Genetic Algorithm (GA)?

Because of its greater strength and durability, Genetic Algorithm (GA) is better than other methods based on Artificial Intelligence (AI). Dissimilar to previous artificial intelligence systems, GA is not easily interrupted with little change in input values or significant amounts of noise in the system. Moreover, the application of GA enjoys many more advantages compared to conventional search methods in other optimization techniques such as linear programming (LP), random search (RS) and with Depth-first search (DFS), Breadth-first search (BFS) or Fraxis methods in search with a large state space, Multime state space, or a multidimensional approach.

An overview of the genetic algorithm to solve the problem simulates the principle of survival of the fittest among members of a population over successive generations. Each generation contains a population of character strings that are the same as the chromosomes found in DNA. Each one represents a point in the search space and a possible solution. The members of each generation then enter into a process similar to the process of evolution of living organisms. Genetic algorithms operate similar to the genetic structures and behaviour of chromosomes in a population of individuals using specific principles. These principles include:

Members of a population compete for resources and mating. Members who are more successful in any competition produce more offspring than members who do not have good performance. Genes from well-functioning members are released into the population. Therefore, good parents sometimes have a child that is better than either of the parents. Thus, each successive generation would be more suitable for living in their environment.

In order to employ a genetic algorithm, a population of people in search space is maintained. Any of these individuals is, indeed, a possible solution to the problem raised. Each member is coded with a limited length vector of components or variables with the help of particular letters of the alphabet. The letters are typically binary alphabets. To continue the analogy of this approach with genetic principles, each member is interpreted as a chromosome, and the variables are considered identical to genes. So a chromosome consists of a solution to a problem of several variables. To demonstrate each member's ability to compete with the rest of the population, each possible solution is assigned a fitness score. The member that has the optimal fitness score is generally closer to the desired optimum. The genetic algorithm intends to use selective breeding of solutions to produce better offspring with the help incorporate the chromosome information.

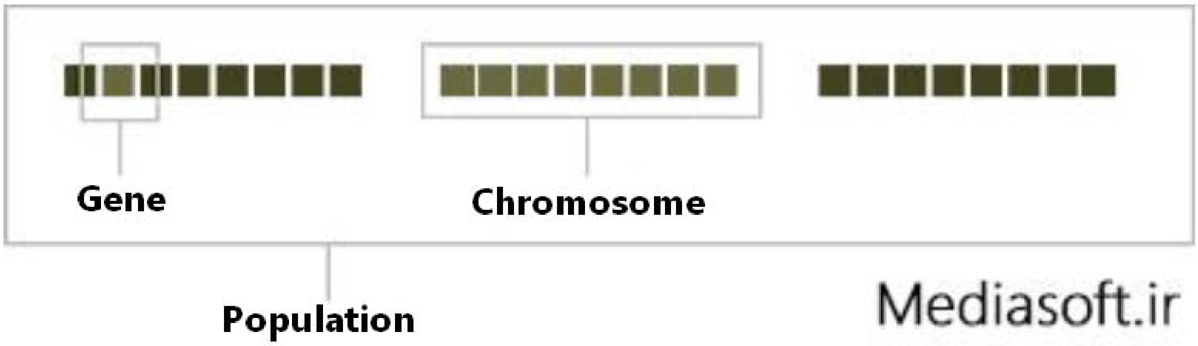

Figure (1-1)

GA constantly maintains a population of 7 chromosomes (solutions) with their fitness values. Parents are selected based on their fitness values for mating and create offspring using a specific program. As a result, solutions that have high fitness will have more opportunities for rehabilitation and having children. Thus, children can inherit attributes from each parent. Taking into account that the size of the population is always kept constant, at the same time of mating parents and producing offspring, a vacancy must be created for them. To this end, members of the population die and are replaced by new solutions. After all mating situations have been used in the initial population, the new generation gradually replaces the previous generation. It can be hoped that better solutions will be developed over successive generations with this method, while solutions that have lower performance will be gradually omitted.

New generations of solutions, on average, have more good genes than a solution from the previous generation. Each successive generation will have relatively better solutions than the previous generations. Ultimately, when the population converges and produces offspring that are not very different from those of previous generations, it is argued that the algorithm itself converges to a set of solutions to the intended problem.

Implementation of genetic algorithm based on natural selection: After the initial generation is randomly created, the genetic algorithm will start the evolution of this initial generation using three operators.

Selection: which is equivalent to the principle of survival of the fittest

Crossover: which represents mating between members of the population

Mutation: Inserts correction code randomly into the problem.

Selection Operator:

The main idea: Giving priority to better members that allow them to transfer their genes to the next generation.

• The well-being of each member relies on the fitness of that member.

Fitness can be achieved with the aid of an objective function or with subjective judgment.

Crossover Operator:

The main factor characterizing the genetic algorithm (GA) compared to other optimization techniques.

• Two members of the population are selected using the selection operator as parents.

• A site is randomly chosen between strings of bits.

• The values of the two strings cross over to the specified site.

If 51000000 and 52111111 are the crossover sites, then 51=110000 and 52001111.

• The two offspring created through this process will enter the next generation of the population.

• By recombining good members of the community, this process is likely to lead to the construction of children better than parents.

Mutation Operator:

• Some bits of part of the new generation members of the previous generation offspring will be altered with very low probability.

• The goal is to maintain diversity among members of a population and prevent premature and incomplete convergence.

• The mutation operator alone causes a generation of a random search in the search space.

• Mutation and selection without crossover operator result in a hill-climbing (HC) algorithm parallel and resistant to noise.

Investigation in conjunction with cancer has become one of the most common endeavours in the 21st century. According to the dataset in 2002, the International Agency for Research on Cancer (IARC) estimates the number of people with cancer at 25 million. The American Cancer Society (ACS) also announced in 2004 that cancer had formally replaced heart disease as the leading cause of mortality so that the survival rate for any type of cancer is 63 per cent. Timely diagnosis and appropriate treatment of cancers increase the patients' safety and improve their condition. Therefore, while cancer detection is on the basis of interventional methods such as surgery, radiotherapy, and chemotherapy, studies demonstrate that the application of novel computerized technologies such as image processing mechanisms was successful in the processes associated with the diagnosis and classification of cancers. In the early 1960s, NASA's Ranger 7 spacecraft began sending important television images from the moon surface to Earth. Extracting image details to find a place to land the Apollo spacecraft required applying decisions on the images. This important task was entrusted to Jet Propulsion Laboratory. Thus, the specialized field of digital image processing was started and, like all other technologies, quickly gained numerous uses so that the issue of image processing dramatically promoted from 1964 until now. In addition to the space research program, image processing techniques are now extensively employed in multiple fields such as astronomy, biology, archaeological geography and in various industries such as aerospace, packaging and printing, automotive, pharmaceuticals and medicine, electronics, food industry, steel, aluminium, copper, and other industries. Image processing is a method to convert an image in digital form and perform some operations on it, in order to obtain an improved image or to extract some useful information from it. Specifically, image processing refers to any type of signal processing that is the input of an image, and the output of the image processor can be an image with a set of particular symbols with image-related variables. Most image processing techniques involve dealing with the image as a two-dimensional signal and applying standard signal processing techniques to them. Image processing often refers to digital image processing, but there are optical and analog image processing as well. The goal of image processing is image sharpening, improvement, retrieval, pattern measurement, and recognition. Image processing could be seen as decisions making tool that facilitates the detection of cancer in the early stages. Early detection of cancer through image screening is the most crucial contribution to reducing deaths from specific cancer. Medical imaging plays an essential role in all phases of prediction, screening, detection, staging, prognosis, treatment planning, treatment response, relapse, recurrence, and relief of cancer. In other words, images make up an important part of cancer clinical protocols and have the potential to provide functional, metabolic, structural, and morphological information and to be helpful along with other diagnostic tools in clinical decision making. Overall, medical images are one of the most important diagnostic tools for physicians, who have always devoted a considerable part of their research, since they show the body's condition in two-dimensional and even three-dimensional forms (by computer). Imaging systems can be divided into groups such as X-ray, radiography, fluoroscopy, Computed Tomography scan (CT scan), magnetic resonance imaging (MRI), nuclear medicine, and ultrasound methods. The resulting images in the above methods cannot be used as a raw material; thus, extensive processes are performed on them, which generally includes image processing and extraction of effective information in detecting and finding the intended positions, reconstruction of images on the computer in three dimensions, and interpolation of data to produce the necessary cuts of the organ under imaging, noise removal (denoising), color assignment, and overall image quality improvement. Taking into account that medical imaging techniques are employed in all phases associated with screening, diagnosis, treatment, monitoring, and follow-up of cancer patients and the image processing related to each of these stages can provide complete information for treatment planning and classification of the patients to improve the final condition of the patient in this way, so image processing and the significance of its application in cancer detection will be discussed in this study.

Results:

Essentially, image processing contains image capture with optical scanners or digital cameras and sensors, image analysis, including data compression, image enhancement, pattern recognition, and the output stage. This step can be an image or a report that is achieved by the result of image analysis. The two types of methods used for image processing are analog and digital image processing. Analog techniques employ image processing for hard copies such as printing and photos. Digital image processing, better known today, has numerous applications of satellite image analysis with dimensional control of microscopic segments. In the image processing process of cancer, after taking the desired image, the image enhancement will start in the pre-processing stage. This step is done with the goal of improving the ability to interpret and understand the information in images for individuals or provide better inputs for other automated image processing techniques. The next step is image segmentation, which is an essential step in image analysis, and the current techniques for capturing and describing the image largely depend on the results of this phase. The purpose of the Segmentation step is the simplification or modification of image view for more meaningful and easier analysis. At this point, the objects and the margins of the image, such as its lines and curves, are specified. In other words, this step is the process of labeling each pixel in each image, which results in a set of segments that together cover the whole image. Image feature extraction is another important step in the image processing process that employs algorithms to identify and isolate various image features that must have sufficient information about the image. This stage represents the final results to determine normal or abnormal images in this way. By analyzing the resulting images, the physicians can identify cancer cells and offer their diagnostic results.

In recent years, image processing mechanisms have been extensively employed in various areas of medicine to improve and accelerate the timely detection and treatment of diseases and conditions such as lung and breast cancer tumors in which time is a critical factor in the diagnosis of the disease.

Literature Review

In a study entitled "Using Different Imaging Methods in Breast Cancer Detection", Yavarian expressed that an isolation algorithm in mammography is an algorithm used to determine the breast region or specific types of abnormalities such as microcalcifications and masses. This fact is generally accepted that determining masses are technically much more complicated than microcalcifications and are stimulated or disappeared by the main breast tissue. The purpose of this section is the review and classification of different algorithms to determine the masses. The remarkable thing is that algorithms are usually exploited to detect specific masses such as comparisons between breast views, comparison between a person's left and right breasts as well as the appearance comparison of the same mammograms.

Azizi carried out a study entitled mammography image processing for the detection of malignant tumors using image processing techniques. He argued that one of the most important and effective ways to diagnose this cancer, especially in the early stages of the disease, is mammography. Based on the mammography images, the masses in the images are detected. In this method, the software thoroughly and accurately identifies the areas around the tumor without the need to detect the correspondence and shows them separately.

In this study, the analysis of mammography images was addressed using image processing techniques such as edge detection based noise removal and separation of areas with higher density. High speed and human error constraints are the most influential factors in the proposed method.

Masoudi conducted a study entitled "Diagnosis and classification of breast cancer masses by feature extraction of mammography images using image processing techniques and employing a neural network." In this article, geometry and new features based on the geometric shapes and edges for the classification of lesions based on the Breast Imaging Reporting and Data System (BIRADS) standard are presented.

Rezaei and Haddadnia, in 2011, examined the exact position of breast cancer in mammography images using image processing and discrete wavelet transform. They proposed a new method based on image processing techniques that has increased the accuracy, sensitivity, and specificity in the detection of a cancerous mass. By applying the discrete wavelet transform (DWT) to the input image and constructing the approximate coefficients of the scaling component, different areas of the image were classified, and the mass suspicious of the cancer was isolated by selecting the appropriate threshold.

Najafabadi et al. in 2016 investigated the detection of breast cancer using image processing techniques of mammography images. The goal of the performed study was the image processing and obtaining a tool so that a mean can declare with high accuracy whether the output image contains a cancerous mass (tumor) or not. A set of standardized images containing healthy and cancerous breast tissue was prepared and processed to meet the need raised.

Mahmoudabad et al. in 2015 evaluated the mammography image processing for the detection of malignant tumors using image processing techniques in MATLAB software. On the basis of this investigation, breast cancer occurs earlier compared to other cancers and is hence the most significant cause of loss of life in women and the major problem for their health. Despite the spread of the disease, it enjoys the potential of timely detection and definitively treatment. High speed and lack of human error are the most crucial factors in the proposed method.

In 2016, Nikbakht and Sambrani studied the application of image processing techniques to detect the position of breast cancer tumors in mammographic images. Breast cancer is still considered a significant public health problem in the world and the second leading cause of death caused by cancer among women. The number of breast cancer patients in Iran is also increasing, and the age of affliction to the disease is ten years less than in Western countries. Nevertheless, it is the most treatable cancer if diagnosed early. Mammography is among the most reliable methods for early detection of breast cancer. However, providing an accurate and uniform assessment for massive mammography produced in extensive screening will be difficult for a radiologist. Hence, an intelligent system with high accuracy in detecting the cancerous mass site will be essential.

In 2015, Effatparvar and Yavarian explored the use of different image processing techniques in breast cancer detection. They announced that an isolation algorithm in mammography is an algorithm employed to determine the breast region or specific types of abnormalities such as microcalcifications and masses. This fact is generally accepted that determining masses are technically much more complicated than microcalcifications and are stimulated or disappeared by the main breast tissue. The purpose of this section is the review and classification of different algorithms to determine the masses. It is worth mentioning that algorithms are usually exploited to detect specific masses such as comparisons between breast views, comparison between a person's left and right breasts as well as the appearance comparison of the same mammograms.

Vanakumar and Pranovi, in 2017, examined breast cancer detection using mammography image processing. In this study, the typical steps in image processing algorithms were widely investigated. Computerised mammography techniques include image pre-processing, image segmentation techniques, feature extraction, feature selection, classification techniques, and mammogram features. Tissue specificity must be achieved to detect between normal cells and cancer cells. Cancer is one of the oldest diseases that many research has been carried out in this area. The obtained results revealed that cancer is not a disease but a collection of numerous diseases. Thus, a single drug for the treatment of cancer is not possible. A key input for cancer treatment is to customize the drug based on the type of cancer and may be treated if detected early.

Guzhman, Torres, et al., in 2013, investigated digital image processing for breast cancer detection. They suggested that the analysis of digital mammograms based on tissue segmentation effectively be provided to detect early-stage tumors. The proposed algorithm was tested on several images taken from a digital database for mammography screening for cancer research and diagnosis. It was found that the algorithm was quite suitable to detect masses and microcalcification from background tissue with morphological operators and then classify them using machine learning techniques and a clustering algorithm for intensity-based segmentation (Thresholding).

Kabbon et al. addressed the examination of breast cancer detection using image processing techniques. They described the use of fuzzy model segmentation and KNN algorithm classification to help detect breast cancer in digital mammograms. Using images from a Digital Database for Screening Mammography (DDSM), they demonstrated that supervised methods such as k-nearest neighbour (KNN) and fuzzy C-means (FCM) in digital mammograms have high classification error rates, when only intensity is exploited as a diagnostic feature.

In a study in 2013 conducted by Nisthula and Yadhu at University Center of Calicut in India entitled a novel method to detect bone cancer using image fusion and edge detection, a fast, easy, and reliable method was employed to detect cancerous tissue in bone. Different image processing techniques, such as contrast enhancement, edge detection, and image fusion, were exploited in this approach. The results gained from this study represented that the proposed method can provide a smooth image with edges and show the effect of the disease without spatial and spectral noise.

Another study performed in 2011 by Shikha and Dhandra entitled Abnormality Detection in Endoscopic Images of Throat Cancer by Morphological Operations used the mathematical morphology method provided by image processing algorithms to detect throat cancer. The algorithms were employed to segment, classify, and analyze endoscopic medical images of the throat to demonstrate abnormalities of this organ. The results indicated that the application of the proposed segmentation method creates more areas in the abnormal images compared to the normal images. Due to the low detection rate of lung tumors, to diagnose such rare long-term tumors, there must be techniques that correctly detect small changes in the images.

The study of Sharma and Jindal in 2011 entitled Identifying Lung Cancer Using Image Processing Techniques was performed using computer-aided automatic detection system to detect lung tumors. In this system, which was able to identify cancerous lesions by analyzing CT scan images of the lungs, the area under study was initially segmented and then each of the segments was analyzed to identify the cancerous mass. Using this system, a precision of 80% was provided for surgeons and radiologists when viewing CT scan images of 2.5 to 7 cm to search for cancerous lesions.

Al-Naami and Al-Bashir, in their article entitled "A Fusion Technique Based on Image-Statistical Analysis for Detection of Throat Cancer Types", in 2010, presented a simple approach for the classification of laryngeal tumors using statistical analysis techniques without the need for biopsy with further testing. In this study, MRI images of 35 patients were processed for the purpose of classification and diagnosis of laryngeal cancer and changed to a number of pixels that covered the tumor area. The data extracted from these pixels were analyzed using traditional statistical methods such as descriptive analysis, box plots, and hypothesis tests to determine the accuracy in differentiating between tumor types. The results of this study indicated that the application of statistical approaches to analyze the data extracted from MRI images to detect and determine the type of laryngeal cancer with a 95% confidence interval is reliable and effective.

In another study in 2007 entitled Mammogram Breast Cancer Image Detection using Image Processing Functions carried out by Alhadidi et al., they aimed to develop algorithms and software to help physicians detect breast cancer from mammogram images. In this study, image processing techniques such as threshold, watershed, and edge-based were exploited. The threshold technique, meanwhile, was simpler and faster in helping physicians make faster diagnoses.

Research Methodology:

In the project, two automated methods are provided to detect the type of breast tumor mass that is benign or malignant. In the first proposed method, segmentation is achieved by automatic regional growth, the threshold of which is trained by the CNN-trained Circular Neural Network. In the second proposed method, segmentation is conducted by a cellular neural network (CNN), whose parameters are specified by the genetic algorithm (GA).

Some of the things that are very important when extracting a tumor segmentation include light intensity, texture, and shape features. A genetic algorithm is employed for proper feature selection of image tissue at the time of tumor segmentation. Next, the neural network is used as a classifier that is a malignant tumor or a benign tumor. In order to assess the efficiency of the proposed method, various classifiers such as Naive Bayesian, Nearest Neighbor Algorithm (KNN), and support vector machine (SVM) have been employed.

An isolation algorithm in a mammography field is an algorithm used to determine the environment of the breast or specific types of abnormalities such as microcalcifications and masses. It is generally argued that the determination of the masses is technically much more difficult than microcalcifications since masses can be stimulated or disappeared by the main breast tissue. The aim of this section is the review and classification of different algorithms to determine the masses. The remarkable thing is that algorithms are commonly used for the detection of certain masses, such as circular or needle-shaped masses. On the other hand, algorithms are usually exploited to detect specific masses, such as comparisons between breast views, comparison between a person's left and right breasts as well as the appearance comparison of the same mammograms.

Analysis of engineering problems:

Engineering problem analysis method

Overall, there are three methods for solving physical problems:

Exact Solution

Numerical Solution

Experimental Solution

In the exact solution, as its name suggests, the parameters of differential equations governing physical fields such as stress field, thermal field, or electric field, etc., are calculated accurately. However, the approximate and numerical solution of these problems will be discussed in the second method. The experimental or laboratory method is also a suitable approach but costly and time-consuming method, considering that it is based on and taken from the realities themselves.

Meanwhile, numerical methods are among the most practical methods used in solving engineering problems. Some advantages of numerical solution over other methods are as follows:

The major weakness of laboratory methods is that they are costly and time-consuming, while this is not the case in the numerical solution method.

The exact solution method fails to analyze models with complex geometry, and only numerical methods, especially the finite element method, are applicable in this field.

The exact solution is also incapable of solving problems where boundary conditions are slightly complicated, and only conventional numerical methods are exploited to solve such problems.

Conclusion:

In this study, MATLAB software is used. MATLAB software allows easy operation of the matrix, computational and functional operations, the use of different algorithms, as well as the ability to communicate with different programming languages easily, and MATLAB has a wide range of applications. These include signal and image processing, communications, controller design, testing and measurement, financial modeling and analysis, and computational biology. It is also possible to expand the MATLAB environment by adding a toolbox for various purposes.

For simulation, training and classification need to be done with the classification method. Here, the Support Vector Machine (SVM) classifier is exploited.

In this method, it is first essential to conduct training. Training means that a number of features related to class one and class two are given to the function, and the algorithm updates its parameters based on the labeling done. Then, the unlabeled data are given to the algorithm for the classification, and it automatically specifies the corresponding class.

The following function is employed for training SVM in MATLAB:

SVMSTRUCT = svmtrain (TRAINING, Y)

Thus, TRAINING is input data for training in which each row represents data (an observation) and the columns represent the dimension of the data. The Y, which is a vertical vector, specifies the class number.

In this section, some classified images related to the train classifier are provided.

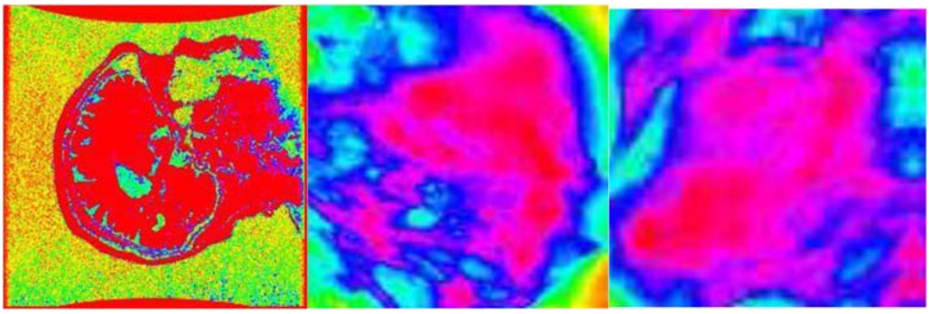

Figure (5-1): Classified images of the train classifier

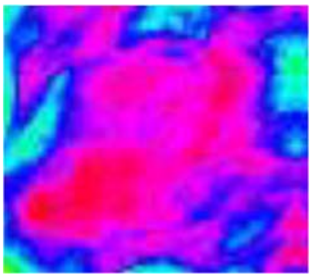

Figure (5-1 a): Output from the program (test section)

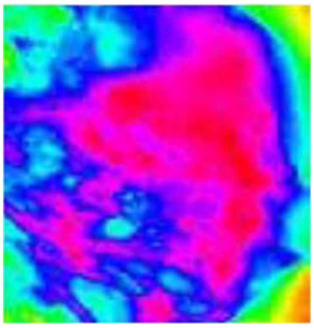

Figure (5-1 b): Output from the program (test section)

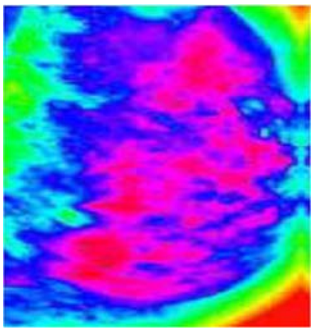

Figure (5-1 c): Output from the program (test section)

The approach used in this article has better performance. The system introduced to classify mammograms into normal, benign, and malignant patterns have been developed to assist radiologists in visual detection. In this study, a new strategy is offered for the detection of breast cancer. Providing an SVM algorithm for mass classification and tumor detection with image processing done through a neural network has been an innovative aspect of this work. The results with higher accuracy have been achieved by using this solution.

Reference:

[1] Alavi Majd, Hamid, Vahedi, Mohsen, Mehrabi, Yadollah and Naqavi, Bahar, 2012. Application of Clustering Methods in DNA Microarray, Research Journal of Shahid Beheshti University of Medical Sciences and Health Services, Volume 31, Number 1, Spring.
[2] Harirchi I, Karbakhsh M, Kashefi A and Momtahen AJ. (2000). Breast Cancer in Iran: results of a multicenter study. Asian Pac j Cancer Prev. Vol. 0. No. 1. 20-20.
[3] Oloumi, Mana, Bouzari, Saeed, Rasaeian, Afsaneh and Mohaqqeqi, Mohammad Ali, 2012. Study of CK13 biomarker in the blood of patients with breast cancer, Iranian Journal of Surgery, Volume 13, Number 12.
[4] Noone AM, Krapcho M and Howlader N. (2012). SEER Cancer Statistics Review, 1320-2003 (Vintage 2003 Populations), National Cancer Institute. Bethesda, MD2012.
[5] Fathi Najafi, Tahereh, Jabarzadeh Ganjeh, Sepideh, Mojahedrezaian, Sima and Mazlum, Seyyed Reza, 2014. Evaluation of some risk factors for breast cancer in women of childbearing age in the city of Mashhad, Journal of Iran University of Medical Sciences, Eleventh Year, No. 11
[6] Smith RA, Cokkinides V and Eyre HJ.) 2003). American Cancer Society guidelines for the early detection of cancer, American Cancer Society. CA Cancer J Clin. No. 03. 22-03.
[7] Qayyoumizadeh, Hossein, Abbaspour Kazerouni, Iman, Haddadnia, Javad, and Hashemian, Maryam. 2011. Detection of breast cancer based on thermal pattern in infrared images, Iranian Breast Diseases Quarterly, Fourth Year, First and Second Issues.
[8] Yavarian, Maria and Mehdi Effatparvar, 2015, Using Different Image Processing Methods in Breast Cancer Detection, Third International Conference on Applied Research in Computer Engineering and Information Technology, Tehran, Malek Ashtar University of Technology.
[9] Azizi Mahmoudabad, Roya; Mehran Azizi Mahmoudabad and Sima Azizi Mahmoudabad, 2015, Mammographic image processing for detection of malignant tumors using image processing techniques in MATLAB software, International Conference on Applied Research in Information Technology, Computer and Telecommunications, Torbat Heydariyeh, Khorasan Razavi Telecommunication Company.
[10] Masoudi, Pouyan; Naser Safdarian and Behzad Kalantar, 2015, Diagnosis and classification of breast cancer masses by feature extraction from mammographic images using image processing techniques and applying neural network, International Conference on Nonlinear Systems and Electrical and Computer Engineering Optimization, Dubai, Pendar Andish Rahpoo Company.
[11] Masoudi, Pouyan; Naser Safdarian and Behzad Kalantar, 2015, Diagnosis and classification of breast cancer masses by feature extraction from mammographic images using image processing techniques and applying neural network, International Conference on Nonlinear Systems and Electrical and Computer Engineering Optimization.
[12] Azizi Mahmoudabad, Roya; Mehran Azizi Mahmoudabad and Sima Azizi Mahmoudabad, 2015, Mammographic image processing for detection of malignant tumors using image processing techniques in MATLAB software, International Conference on Applied Research in Information Technology, Computer and Telecommunications, Torbat Heydariyeh, Khorasan Razavi Telecommunication Company.
[13] Sambrani, Nasim and Elham Nikbakht Molla Mahmoud, 2016, Using image processing methods to detect the location of breast cancer tumors in mammographic images, the first national conference on applications of mechatronic and robotic systems, Ilkhchi, Islamic Azad University, Ilkhchi branch
[14] Yavarian, Maria and Mehdi Effatparvar, 2015, Using Different Image Processing Methods in Breast Cancer Detection, Third International Conference on Applied Research in Computer Engineering and Information Technology, Tehran, Malek Ashtar University of Technology.
[15] Field RW, Withers BL. Occupational and Environmental Causes of Lung Cancer. Clinics in Chest Medicine. 2012; 33 (4): 681-703.
[16] Ghaderzadeh M, Sadoughi F, Ketabat A. Designing a Clinical Decision Support System Based on Artificial Neural Network for Early Detection of Prostate Cancer and Differentiation from Benign Prostatic Hyperplasia. Health Information Management. 2012; 9 (4): 464.
[17] Glotsos D, Kalatzis I, Theocharakis P, Georgiadis P, Daskalakis A, Ninos K, et al. A multi-classifier system for the characterization of normal, infectious, and cancerous prostate tissues employing transrectal ultrasound images. Computer methods and programs in biomedicine. 2010; 97 (1): 53-61.
[18] Plaza A, Valencia D, Plaza J, Martinez P. Commodity cluster-based parallel processing of hyperspectral imagery. Journal of Parallel and Distributed Computing. 2006; 66 (3): 345-58.
[19] Kruse FA, Lefkoff AB, Boardman JW, Heidebrecht KB, Shapiro AT, Barloon PJ, et al. The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment. 1993; 44 (2-3): 145-63.
[20] Hosseinpour S, Rafiee S, Aghbashlo M, Mohtasebi SS. A novel image processing approach for in-line monitoring of visual texture during shrimp drying.Journal of Food Engineering. 2014; 143 (0): 154-66.
[21] Zalevsky Z, Livshits P, Gur E. Chapter 2 - New Image Processing Methods for Advanced Metallization in Micro- and Nano-Electronics. In: Zalevsky Z, Livshits P, Gur E, editors. New Approaches to Image Processing Based Failure Analysis of Nano-Scale ULSI Devices. Oxford: William Andrew Publishing; 2014. p. 29-44.
[22] Yao Q, Xian D-x, Liu Q-j, Yang B-j, Diao G-q, Tang J. Automated Counting of Rice Planthoppers in Paddy Fields Based on Image Processing. Journal of Integrative Agriculture. 2014; 13 (8): 1736-45.
[23] Fass L. Imaging and cancer: A review. Molecular Oncology. 2008; 2 (2): 115-52.
[24] Gemmeke H, Ruiter NV. 3D ultrasound computer tomography for medical imaging. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2007; 580 (2): 1057-65.
[25] Wayne R. Chapter 14 - Image Processing and Analysis. In: Wayne R, editor. Light and Video Microscopy (Second Edition). San Diego: Academic Press; 2014. p. 255-69.
[26] Borase S and Deshmukh M. Medical Imaging Advancement International Journal on Recent and Innovation Trends in Computing and Communication, 2014. 2 (3): p. 638-640.
[27] Härkönen S, Lehtonen A, Eerikäinen K, Peltoniemi M, Mäkelä A. Estimating forest carbon fluxes for large regions based on process-based modelling, NFI data and Landsat satellite images. Forest Ecology and Management. 2011; 262 (12): 2364-77.
[28] Chaudhary A, Singh SS. Lung cancer detection using digital image processing. Volume. 2012; 2: 1351-9.
[29] Nisthula P, Yadhu .R.B. A Novel Method to Detect Bone Cancer Using Image Fusion and Edge Detection. International Journal of Engineering and Computer Science. 2013; 2 (6): 2012-8.

[30] Shikha D, Dhandra B. Abnormality Detection in Endoscopic Images of Throat Cancer by Morphological Operations. Indian Streams Research Journal. 2011; 1 (4).
[31] Sharma D and Jindal G. Identifying Lung Cancer Using Image Processing Techniques, in International Conference on Computational Techniques and Artificial Intelligence. 2011: Landran, India. p. 115-120.
[32] Al-Bashir A, Al-Naami B. A Fusion Technique Based on Image-Statistical Analysis for Detection of Throat Cancer Types. EDITORIAL BOARD.2010; 4 (6): 677.
[33] Alhadidi B, Zu'bi MH, Suleiman HN. Mammogram breast cancer image detection using image processing functions. Information Technology Journal. 2007; 6 (2): 217-21.
[34] Chandra P, Noh H-B, Shim Y-B. Cancer cell detection based on the interaction between an anticancer drug and cell membrane components. Chemical Communications. 2013 49 (19): 1900-2.