ISSN: 2581-3404 (Online) IF: 5.445 (SJIF)
International Journal of Innovative Research in Technology and Management, Volume-6, Issue-2, 2022.

A Survey of Machine Learning Based Healthcare Applications

Kumari Deepshikha[1], Chetan Agrawal[2], Nidhi Ruthia[3]
Dept. of CSE, Radharaman Institute of Technology & Science, Bhopal, India[1, 2, 3]
kdlk752@gmail.com[1], chetan.agrawal12@gmail.com[2], nruthia11@gmail.com[3]

Abstract: *In recent years, breast cancer has overtaken lung cancer as the second biggest cause of mortality due to the disease. Breast cancer is one of the most frequent malignancies found in women around the world. It is possible for radiologists to have difficulties in both diagnosing and treating breast cancer. As a result, primary care is beneficial for reducing the risk of disease and death. The primary goal of this research is to improve treatment options and maybe save lives through earlier detection of disease. This research demonstrates the adaptability of the methodology by combining modern segmentation methods with machine learning techniques, both of which are burgeoning fields of research. Imaging modalities are of utmost importance for image-based CAD systems due to the fact that the speed, accuracy, and consequently the overall performance of CAD systems are directly determined by the quality of the images that are produced by these imaging modalities. In this article, we discuss the various computer-assisted procedures for the early identification of breast cancer, as well as their effectiveness and the effect that they have on the health of patients.*

Keywords: Diseases diagnosis, Classification, Machine learning, Image segmentation, CAD.

Introduction

According to figures compiled from around the world in 2018, breast cancer was the most prevalent kind of cancer found in females. In addition, it is predicted that the number of new instances of breast cancer reached 2 million in 2018, while the number of fatalities from the disease was over 600,000. The efficiency with which newly discovered cases of the disease are categorized needs to be significantly improved. Breast cancer mortality can be significantly reduced through the use of early detection and treatment, which is one of the most effective strategies to do so. The word "cancer" originates from an Ancient Greek word that was spelled karkinos. Hippocrates, a Greek physician, is credited with being the first to employ it. He used it to explain carcinoma, a condition in which aberrant cells divide uncontrollably and in an unregulated manner[1]. Evidence of cancer in humans was discovered in ancient writings dating back to 1600 B.C. and in the mummies of ancient Egyptians who had died centuries earlier. Egypt was the location of one of the earliest cases of breast cancer ever documented, dating back to 1500 BC. Breast cancer is one of the leading causes of death among women, but if detected in its early stages, it is possible to lessen the disease's aggressiveness. Mammography and ultrasound are two diagnostic tools that are frequently utilized in the process of diagnosing breast cancer. In spite of the fact that mammography is routinely used in clinical practice for the evaluation of breast masses and that it improves the visibility of nonpalpable and small tumors, ultrasound is the method of choice because it is less invasive for patients, it is more cost-effective, and it takes significantly less time[2]. In most cases, the predictive value of mammography is not very high. Patients who have tumors that are not cancerous are therefore subjected to biopsies that are not essential. This shortcoming of mammography has been remedied by the addition of ultra-sonography to the screening process [3], which serves as a supplementary tool.

The advancement of medical imaging technology paves the way for new lines of inquiry in the field of computer-assisted diagnostics [4]. The examination of cytological pictures is also a crucial part of the progress. Recently, it has become possible to scan the material that is acquired as a consequence of FNB into a digital form and analyse it using a variety of computer approaches. The examination of cytological pictures for the purpose of diagnosing cancer places a primary emphasis on the cell nuclei. Therefore, accurate and efficient automatic segmentation of cell nuclei is the first step toward accurate and successful classification of cytological images. A collection of test images that have had their cell nuclei manually segmented should be generated so that the viability of the proposed method of segmentation may be evaluated. Unhappily, the collection of test photographs that was generated contains examples of a wide variety of quality levels. There are photos that have nuclei that are clearly separated from one another, but there are also examples that have very little space between the nuclei of nearby cells. Because of the range of samples, it is possible to prove that the offered approaches are successful for both straightforward and complex scenarios [5]. The unchecked reproduction of normally occurring cells in the body can lead to the development of cancer. When the DNA in a particular portion of the body is damaged, cancer cells can arise from normal cells. The body is usually able to repair broken DNA, but unfortunately, damaged DNA in cancer cells is not repaired. This is because cancer cells lack the ability to repair damaged DNA. It is also possible for people to inherit the damaged DNA from their parents, which can lead to tumors being passed down through generations. And in many instances, a person's DNA becomes damaged as a result of the individual's prolonged exposure to unfavorable circumstances, such as smoking. The typical appearance of cancer is that of a solid tumor. Some tumors are not cancerous, often known as malignant, while others are noncancerous (benign). Malignant tumors, as opposed to benign ones, pose a risk to the patient's life. The most important stage is getting a correct diagnosis of the tumor, so that medical professionals may move on to the subsequent treatment steps [6]. The use of several imaging modalities is the first step in the process of detecting breast abnormalities. The quality of the images obtained through the use of various techniques may have a significant bearing on how well the detection process works. The following imaging methods are frequently utilized in the process of breast cancer screening: The performance of various imaging methods may be evaluated using a variety of metrics, the most notable of which are sensitivity, specificity, recall rates, positive projected value, F-score, accuracy, and area under the curve (AUC) [7].

It is possible to exercise some degree of control over the diagnosis of breast cancer in its earlier stages. Micro-calcifications and masses, both of which are rather common anomalies, are the root cause of breast cancer. In the connective tissues and epithelia of the breast region, micro-calcifications and breast lumps can develop. Breast tumors develop locally within the breast and can range in both size and form. The degree to which they have progressed determines whether or not they are considered benign or malignant. Breast lumps that are benign are not malignant and do not spread to other parts of the body, but when they grow, they can press on other organs and cause additional issues [8]. Breast cancers that are malignant are very dangerous and aggressive. In order to prevent their death, treatment needs to begin as soon as feasible. Tumors that are malignant are irregular in shape, whereas benign masses have limited and smooth boundaries and are either oval or circular in shape. Lumps in the breasts that are described as fuzzy, scratchy, or ambiguous are the examples of malignant breast masses [9]. In addition to this, the cancerous tumor is noticeably whiter than any of the tissues that surround it. An automatic method that will assist expert radiologists in assuring improved interpretation and accuracy has been developed as a result of the difficulties and benefits of prior breast tumor categorization and detection methods.

II. Literature Review

Computer-assisted cytology is typically made up of a collection of procedures that include (a) semantic segmentation, (b) cell nuclei detection, (c) instance segmentation, (d) feature extraction and selection, and (e) classification. These procedures are listed in the order that they are typically performed.

[1] Breast cancer is one of the primary reasons for the rise in the overall death rate among women. When it comes to diagnosing geological diseases like breast cancer, the ultrasound exam is by far the most common and popular procedure. Extraction of the region of interest is the first stage in determining the abnormality of the breast cancer (malignant versus benign), which is the first step in the process (ROI). In order to accomplish this goal, a novel method for breast ROI extraction has been presented with the intention of lowering the number of instances of false positive results (FP). The information from the surrounding pixels and a neural network were used in the construction of the proposed model. Training and examination are the two stages that are included in this process. During the training phase, the number of batches was extracted from the ROI as well as the background in order to construct a trained model. During the testing phase, the image was scanned with a window of a set size so that the ROI could be distinguished from the backdrop. Following that, a distance transform was applied in order to locate the ROI and eliminate the non-ROI. The on-data set contained 250 ultrasound images, of which 150 were benign and 100, were malignant. The preliminary findings reveal that the proposed method achieves a success rate of approximately 95.4 percent for breast contour extraction.

This research report [2] outlines an innovative method for locating micro-calcifications inside mammography pictures. An automated method has been proposed for the identification of breast micro calcifications (MCs). This method makes use of two-level segmentation processes: first, crop the breast area from the image utilizing k-means clustering; then, an optimized region growing (ORG) approach has been used, where multi-seed points and thresholds are generated optimally depending on the color values of the image pixels. The features of the texture are then retrieved using Haralick's definitions of what constitutes texture analysis. In addition, three features are extracted from the segmented image. These features are the cross-correlation coefficient, the Pearson correlation, and the average area of the segmented spots. A support vector machine (SVM) classifier was used to assess how effectively the system made use of the digital pictures contained within the digital database for screening mammography (DDSM) dataset.

. A fine-needle biopsy that does not involve aspiration is one of the most common approaches taken in the process of diagnosing breast cancer. The examination of cytological pictures for the purpose of diagnosing cancer places a primary emphasis on the cell nuclei. Therefore, good automatic segmentation of cell nuclei is the first step toward accurate classification of cytological images. The objectives of their study [3] are as follows: (a) the development of methods for the segmentation of cell nuclei based on deep learning techniques; (b) the extraction of some morphometric, colorimetric, and textural features of individual segmented nuclei; and (c) the construction of effective classifiers for detecting malignant or benign cases based on the extracted features. Both fully convolutional neural networks and the marker-controlled watershed technique were used as the basis for the segmentation approaches that were utilized in this particular piece of research. Seven distinct classification approaches are utilized in order to accomplish the goal of classification. According to the F-score, the accuracy of cell nuclei segmentation can reach 90 percent when applied to benign nuclei and 86 percent when applied to malignant nuclei.

As a result of their ability to enhance the accuracy of mammograms in the detection of cancerous masses, the methods of machine learning are frequently utilized in the classification of breast lesions. Cancer of the breast continues to be one of the leading causes of death among women. A timely diagnosis can improve the likelihood of receiving adequate therapy, hence lowering the risk of morbidity and mortality. Mammograms, which are used in cancer screening, have the potential to be an effective approach for locating breast tumors at an early stage. The most challenging part of treating cancer is determining whether or not a patient has a malignant or benign form of the disease. This can be a challenge both during the detecting process and when differentiating between possible diagnoses. The use of machine learning techniques such as the K-Nearest Neighbors classifier, which provides a high level of accuracy performance, helps to overcome this problem. In order to improve the diagnosis accuracy of the mammography, one of the machine learning techniques known as K-Nearest Neighbors is utilized. This article [4] provides a summary of some recent research that highlight the accuracy of the K-Nearest Neighbors method, which is a machine learning technique, in identifying breast cancer.

The agonizing therapy for breast cancer, which only offers a temporary relief from the disease's symptoms but ultimately results in the patient's death, makes it one of the most dangerous diseases affecting women in the world. It is a difficult-to-treat medical condition that is associated with a high death rate. The good news is that if it is caught in its early stages, it can be treated successfully and totally cured. The unchecked reproduction of aberrant breast cells, which results in the growth of a tumor in the breast region, is what leads to the development of breast cancer. It's possible that the tumor is benign, in which case it wouldn't cause cancer, or malignant, in which case it would. It is possible for cancerous cells to move from the site of the primary tumor to lymph nodes located nearby or even to other sections of the body that are located a great distance away; this process is referred to as metastasis. Pain in the breast, swelling of the breast, thickening of the breast tissues, dimpling of the breast, or a lump in the breast or armpit are some of the possible symptoms of breast cancer. There have been many different kinds of studies conducted to improve the accuracy of breast cancer predictions. The SVM (Support Vector Machine) method is one that has been attempted and tested to detect the existence of breast cancer with improved accuracy. This paper [5] examines and compares three distinct models that use this strategy. In the later stages of this investigation, a unique method for the identification of breast cancer is utilized. This method makes use of Artificial Neural Networks in conjunction with Multi-Level Support Vector Machine models.

The community as a whole has shown a significant amount of interest in the research and development of automated breast cancer detection systems that make use of deep learning. Breast cancer, a leading cause of death among women that is responsible for millions of deaths each year, is a disease that may be managed and even cured provided that it is discovered at an early stage using advanced techniques. In this research, the authors [6] investigated computer-aided methods for the diagnosis, detection, and segmentation of breast cancer based on cutting-edge deep convolutional neural networks (CNN). Because the datasets that are available also play an indirect role in determining the performance of CAD systems, we introduced and explored the specifics of public datasets. At the end of this paper, we will go through some of the problems that still need to be solved in CAD systems for breast cancer. The three aspects listed below are primarily responsible for this survey's most interesting findings. They presented the key elements in deep learning to form the compactness for methods that were mentioned in reviewed papers. First, we covered a wide range of the fundamentals of breast cancer, from imaging modalities to popular databases in the community. Second, they presented the key elements in deep learning. Thirdly, and last, the summative features of each reviewed study are included in order to give interested readers with an updated version of these works that does not require them to refer to the original publications themselves.

III. Machine Learning

Machine Learning (ML), is a kind of Artificial Intelligence (AI) which allows a machine to learn by providing a set of data in order to get information via experience without excessive programming. The objectives of ML are to allow machines to make predictions, do clustering, extract association rules, or make decisions from a data set. Classification strategies are most useful for the selection of approaching instances, based on certain patterns and constraints [10]. ML methods include absolute conditionality, Boolean logic, and unconventional optimization approaches to make construct prediction models and patterns. ML consists of four stages: data collection, model selection, model training and model testing. In the process of machine learning, a machine is trained with data to make a decision for similar cases. It is applied in many sections e: g; object recognition, network, security, and healthcare. In the field of oncology, machine learning can be used effectively to differentiate a malignant lesion from a benign one [11]. One of the most common malignancies worldwide is breast cancer (BC), which is a leading reason of death among women. BC is the common disease in women in the real world. It is a cancer that develops in female breast cells by spreading into the body's surrounding tissues. Earlier diagnosis can improve the survival rates in BC patients. There is no recent association between ML and cancerous diseases, since it was used for decades to classify many types of malignancies, including breast cancer. ML is widely accepted as the system of choice for BC pattern recognition and prediction modelling. After being suspected, BC can be diagnosed using the mammogram which is a very effective tool as it can detect a breast lesion (a mass or a micro-calcification) even two years before being felt by the patient [12, 13, 14]. The diagnosis of such lesions, however, may be missed owing to distraction or fatigue while interpreting the mammogram report. As such, automatic classification via Computer-Aided Diagnosis (CAD) is important. This can help health care professionals to properly describe a mass or a micro-calcification. It has been found that CAD can improve cancer diagnosis by up to 77%. This can be done with the use of advanced techniques, including AI. An algorithm was developed using K-NN, which analyzes the mammographic dataset to predict breast tissue malignancy using predefined features. A set of data is reserved for the algorithm to train, allowing the remaining values to test for accuracy. K-NN is one of the easiest types of classification method, it is a supervised learning method used to diagnose cancer, heart disease, and hepatitis. K-NN is relatively simple and effective classification algorithm when compared to other algorithms, it is non-pragmatic algorithm I: e; does not need the assumption for distributing data. It classifies the case study directly by the samples in the data set and thus, does not require a training process. It's an easy supervised learning algorithm for pattern recognition [15,16]. K-NN algorithm stores all cases and categorizes new cases dependent on similarity measures; searches the pattern space for the k training tuples nearest to the unknown tuples. Performance relies on the optimum number of neighbours (k) chosen, which varies from one data sample to another. After the above-mentioned points, this review will focus on the usage of the K-NN algorithm as a classifier for mammogram images and how accurate is it in identifying sinister breast lesions [17].

The main types of ML are supervised learning (SL) and unsupervised learning (USL). Briefly, SL needs training with labelled data that has inputs and target output. In comparison to SL, USL does not need labelled training data, and the environment only contains inputs without desirable goals [18]. In SL, this method can be described as a problem of classification and regression. The role of classification leads to a learning process which categorizes the data into a number of finite classes. In the field of regression problems, the learning function maps the data to a real-value variable. Consequently, the value of the predictive variable may be calculated for each new sample based on this method. The classification model is used for discrete value problems, although the regression pattern is used to make choices on continuous value problems.

IV. Conclusion

In today's modern medicine, specialists are concentrating increasingly on the use of technology treatments for a wide range of chronic illnesses. In spite of the fact that there is currently no cure for a number of diseases, including cancer, stroke, heart attack, chronic liver diseases, viral hepatitis, and coronary artery disease, the death rate attributable to breast cancer continues to rise each year. Over the past decade, there has been a rise in the amount of research conducted on breast cancer, which occurs when anomalies and uncontrollability in the breast cell tissues of women evolve into significant breast cancer. This report only summarized previous research on computer-aided detection (CAD) systems for breast cancer. This paper discusses many methods that have been developed recently with the goal of detecting, segmenting, and classifying bulk and micro-calcification. These topics are considered to be current hot topics. Because they were an essential component of the examinations of CAD systems, the well-known mammographic databases were presented in a clear and concise manner right at the beginning of this piece of writing.

References:

[1] Diyar Qader Zeebaree, Habibollah Haron, Adnan Mohsin Abdulazeez, Dilovan Asaad Zebari, “Machine learning and Region Growing for Breast Cancer Segmentation”, International Conference on Advanced Science and Engineering, IEEE, 2019, pp. 88-93.

[2] Srwa Hasan Abdulla, Ali Makki Sagheer, Hadi Veisi, “Breast cancer segmentation using K-means clustering and optimized region-growing technique”, Bulletin of Electrical Engineering and Informatics, 2022, pp. 158-167.

[3] Marek Kowal, Marcin Skobel, Artur Gramacki, Józef Korbicz, “Breast Cancer Nuclei Segmentation and Classification Based on A Deep Learning Approach”, Int. J. Appl. Math. Comput. Sci., 2021, pp. 85-106.

[4] Shler Farhad Khorshid, Adnan Mohsin Abdulazeez, “ Breast Cancer Diagnosis Based on K-Nearest Neighbors: A Review”, Palarch’s Journal of Archaeology of Egypt/Egyptology, 2021, pp. 1927-1953.

[5] S. Leena Nesamani, S. Nirmala Sugirtha Rajini, M. S. Josphine, J. Jacinth Salome, “Deep Learning-Based Mammogram Classification for Breast Cancer Diagnosis Using Multi-level Support Vector Machine”, Springer, 2021,pp. 371-384.

[6] Juan Vizcarra, Ryan Place, Li Tong, David Gutman, May Dongmei Wang, “Fusion in Breast Cancer Histology Classification” ACM-BCB ’19, September 7–10, 2019, Niagara Falls, NY, USA, pp 485- 493.

[7] Mutiullah, Mehwish Bari, Adeel Ahmed, Muhammad Sabir, And Sajid Naveed, “Lung Cancer Detection Using Digital Image Processing Techniques: A Review”, Mehran University Research Journal of Engineering & Technology Vol. 38, No. 2, April 2019, pp 351-360.

[8] Poonam Jaglan, Rajeshwar Dass, Manoj Duhan, “Detection of Breast Cancer using MRI: A Pictorial Essay of the Image Processing Techniques”, IJCERT 2019, Volume-6, Issue-01 ,2019, pp 238-245.

[9] Kalyani Wadkar, Prashant Pathak and Nikhil Wagh, "Breast Cancer Detection using Ann Network and Performance Analysis with SVM" , International Journal of Computer Engineering and Technology, 10(3), 2019, pp. 75-86.

[10] Marwan Abo Zanona, "Image Processing & Neural Network Based Breast Cancer Detection", Computer and Information Science, Vol. 12, No. 2; 2019, pp 146-154.

[11] Mohamed Abdel-Nasser, Antonio Moreno and Domenec Puig, "Breast Cancer Detection in Thermal Infrared Images Using Representation Learning and Texture Analysis Methods", Electronics 2019, 8, 100, MPDI, pp 1-18.

[12] Jwan Najeeb Saeed, " A Survey Of Ultrasonography Breast Cancer Image Segmentation Techniques", Academic Journal of Nawroz University (AJNU) Volume 9, No 1 (2019). pp 1-14.

[13] Ahmed Abdullah Farid, Gamal Ibrahim Selim, and Hatem A. Khater, "A Composite Hybrid Feature Selection Learning-Based Optimization of Genetic Algorithm For Breast Cancer Detection", preprints March 2019, pp 1-21.

[14] R.Dhivya, R.Dharani, " Survey On Breast Cancer Detection Using Neural Networks" , International Research Journal of Engineering and Technology, 2019, Volume: 06 Issue: 03, pp 3943-3945.

[15] Shubham Bhardwaj, Shalini Bhadola, Kirti Bhatia, "Lung Cancer Detection Using Digital Image Processing and Artificial Neural Networks", International Journal of Scientific Research & Engineering Trends. 2019 Volume 5, Issue 2, pp 582-587.

[16] Nazir Jan, Shahid Khan, Hazrat Ali, "Breast Cancer Detection and Classification based on Multilevel Wavelet Transformation", IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications 18-21 September, 2019, Metz, France.

[17] Ümit Budak, Zafer Cömert, Zryan Najat Rashid, Abdulkadir Şengür, Musa Çıbuk, "Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images", Applied Soft Computing Journal, 2019 pp 1-7.

[18] Hafiz Mughees Ahmad, Sajid Ghuffar, Khurram Khurshid, "Classification of Breast Cancer Histology Images Using Transfer Learning", IEEE conference paper IBCAST 2019.