ISSN: 2581-3404 (Online) IF: 5.445 (SJIF)
IJIRTM, Volume-7, Issue-2, March-2023.

Breast Cancer Diagnosis using Supervised Machine Learning Classification Techniques: Survey & Discussion

Neelesh Gour[1], Dr. Neelesh Jain[2], Prof. Prateek SInghal[3]

[1]M. Tech Scholar, Department of Computer Science & Engineering

[2]Professor, Department of Computer Science & Engineering

[3]Assistant Professor, Department of Computer Science & Engineering

[1,2,3] SAM College of Engineering and Technology, Bhopal (M.P.)

Abstract: Early detection of the disease has become an important issue in the medical field due to the increasing population of the world. With the rapid growth of the population, the risk of dying from cancer is progressive each day. "Cancer" is a scary disease for most people. But in women, "breast cancer" can be even more frightening, because it is directly related to the woman's body. However, men can also get breast cancer, but men are much less likely to get the disease than women. In this paper we present the different techniques used for breast cancer diagnosis using machine learning.

Keywords:- Machine Learning, Deep Learning, Cancer Diagnosis, World health organization, Classification techniques.

Introduction

Early detection of the disease has become an important issue in the medical field due to the increasing population of the world. With the rapid growth of the population, the risk of dying from cancer is progressive each day. "Cancer" is a scary disease for most people. But in women, "breast cancer" can be even more frightening, because it is directly related to the woman's body [13]. However, men can also get breast cancer, but men are much less likely to get the disease than women. Breast cancer is a major concern in Bangladesh. According to the World Health Organization, an exceeding 15,000 populace are diagnosed with breast cancer in Bangladesh every year, of which more than 98 percent are women, and a very small amount of men are also diagnosed with breast cancer. According to expert doctors, the number of people suffering from breast cancer is rapidly growing. According to medical science, when some breast cells grow abnormally, those irregular and extra cells divide into tumors or lumps and expansion to diverse virtue of the organism care of lymphatic and other blood vessels. The reasons for the increasing number of breast cancer cases in Bangladesh and around the world are given below: i) there have been many changes in our way of life and eating habits. It can happen if someone has a family history of breast cancer. ii) If someone menstruates before the age of twelve or has late menopause, they are also at risk. iii) Radioactivity also increases the risk of breast cancer. It can be caused by late adoption or those who do not have children, or not breastfeeding the child. iv) If you have more fat and animal meat than vegetables or fruits in your diet,

eating more processed foods increases your risk of breast cancer. v) Those who are overweight also have a risk of breast cancer. vi) Prolonged use of birth control pills or hormone injections can lead to breast cancer. vii) Breast cancer also increases with age. After the age of 50, this risk increases a lot. Breast cancer can be detected through various tests. These are- 1. Mammography 2.Ultrasonography 3.MRI 4.FNAC 5. Biopsy/meat examination Treatment The treatment of breast cancer is mainly divided into several parts. These are- 1. Surgery 2.Chemotherapy 3.Radiotherapy 4.Hormone therapy 5.Targeted therapy. Breast cancer is the most diagnosed cancer worldwide. According to [11], that more than a million women have breast cancer each year and 400,000 of those cases lead to death. This fatal disease not only affects emerging countries but also developed countries. To find a cure, it is important to correctly diagnose the disease and very quickly and treat it based on the type of symptoms that have arisen. Generally, the most effective approach suggested by radiologists for this disease is medical images.

The rest of this paper is organized as follows in the first section we describe an introduction of disease detection and cancer disease diagnosis system. In section II we discuss cancer disease diagnosis using machine learning techniques. In section III we discuss cancer disease diagnosis using deep learning techniques. In section IV we discuss the computer aided diagnosis system for health care sector, finally in section V we conclude the about our paper.

II. Machine Learning

Breast Cancer is a most common form of cancer among women and life taking disease around the globe. Histopathological imaging is one of the methods for cancer diagnosis where Pathologists examine tissue cells under different microscopic standards but disagree on the final decision. This is a tiresome task and for that reason, Deep Neural Networks are being used for the supervised classification. Breast Histology dataset having 240 training and 20 test images for classification of the histology images among four classes, i.e. Normal, Benign, In-situ carcinoma and Invasive carcinoma. The dataset was preprocessed for proper classification. They have applied transfer learning based on AlexNet, GoogleNet, and ResNet that can classify images at multiple cellular and nuclei configurations. This approach has resulted in 85% accuracy in case of ResNet as the highest among others and further research is being done to increase its efficiency and reduce the human dependency. Breast cancer is one of the most common disease from which most of the females are suffering. Histopathological images play remarkable notch in the medical domain. Segmentation of breast cancer images for cell analysis is the utmost thought-provoking task because of uncertainties present in these images. Identifying cancerous cells effectively in histopathological images may help in early diagnosis of breast cancer.

Machine Learning is a branch of Artificial Intelligence that has become very popular, and useful, in the last 10 years. One definition of Machine Learning is that it is the semi-automated extraction of knowledge from data. Broadly speaking, machine learning (ML) deals with the question of how to build computer programs that learn from data and, as a result, can generate programs that generalize from that data in the form of a program that reflects concepts implicit in the underlying data. In effect, with machine learning we have programs using data to create new programs. This is in contrast to the traditional way that programs have been generated by human programmers in which they encode the rules that the computer follows in a programming language in order to produce a solution to a specified problem. Traditional or conventional writing of programs for a computer can be summarized as automating the procedures to be performed on input data in order to create output artifacts. Almost always, they are linear, procedural and logical.

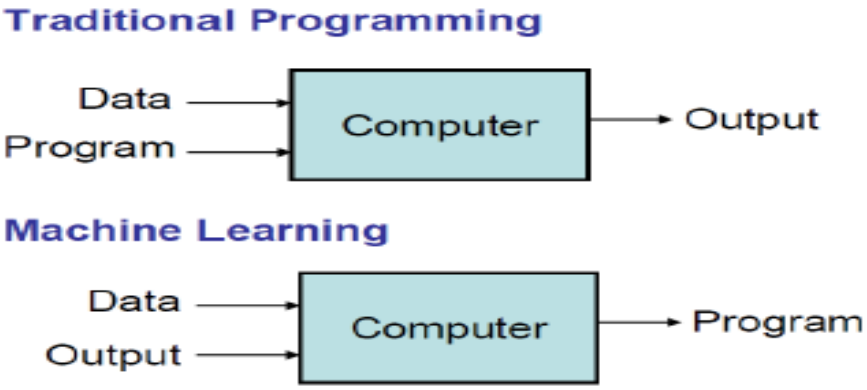

Figure 1: Traditional programming vs. machine learning.

This is not always the case as sometimes, however, there are problems that you can represent in a computer that you cannot write a traditional program to solve. There are two broad classes of machine learning techniques, supervised learning and unsupervised learning.

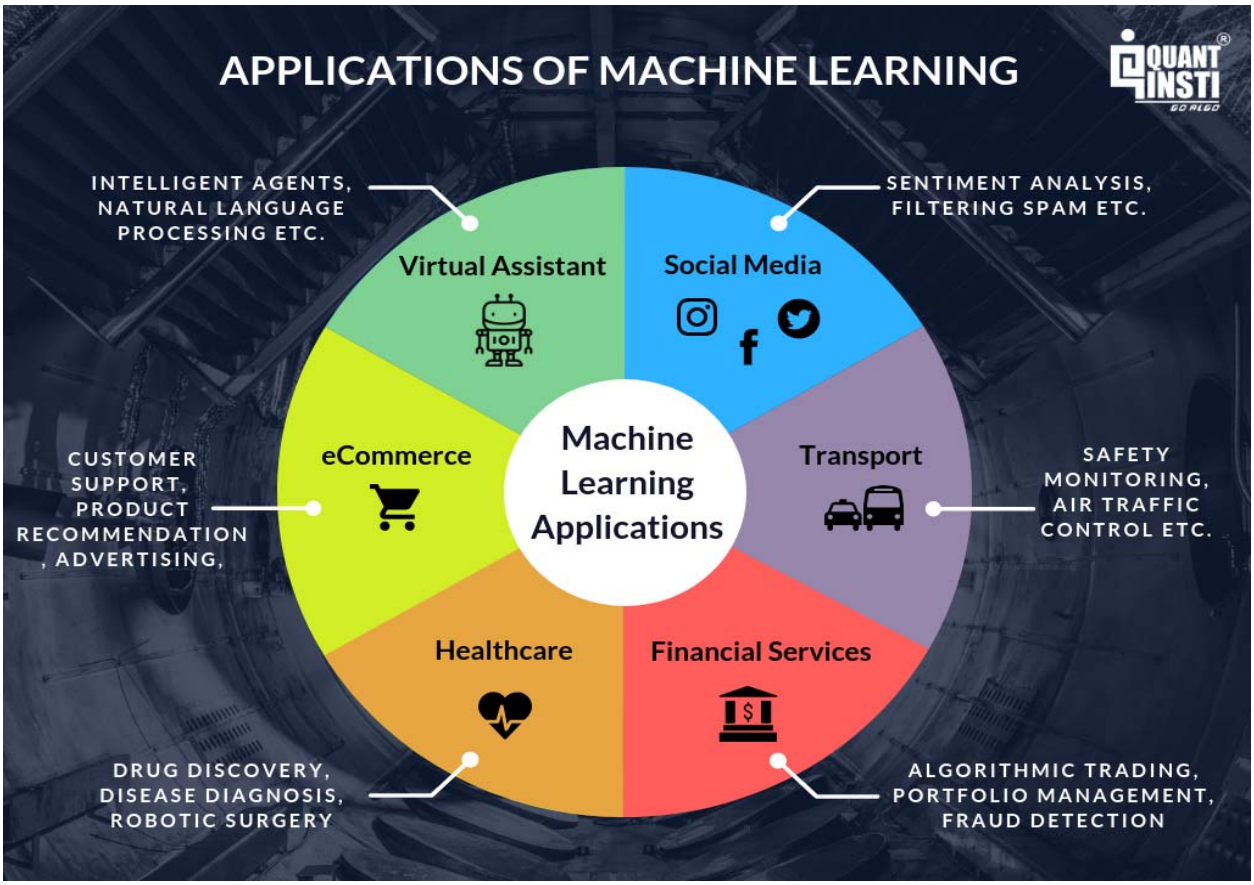

Figure 2: Applications of machine learning.

III. Deep Learning

Recently, artificial intelligence (AI) techniques have been applied in spectrum sensing analysis. The AI technique is able to learn significant features from original historical data and can make a decision based on on-line data. Deep learning has been proved to be an advanced technology for big data analysis with a large number of successful cases in image processing, speech recognition, object detection, and so on. CNN including a set of components (convolutional layers, pooling layers, fully connected layers, and so on) is currently considered as one of the most popular machine intelligence models for big data analysis in various research areas. A typical architecture of CNN model for classification problems is displayed in below figure. Convolution operations are implemented by traversing input matrices with convolution kernels that can be understood as filters for feature extraction. Different from filters used in conventional image processing method whose parameters need to be set manually, the parameters inside the kernel can be learned automatically by deep learning method. Convolutional layers are built by a set of convolution kernels, whose parameters (channels, kernel size, strides, padding, activation, and so on) should be set and optimized according to the practical problem. The computed output from convolutional layer is then sub-sampled by pooling layers. A group of chained convolutional layers and pooling layers can learn high level features representing the original input. The fully connected network (FNN) block, composed by fully connected neural units, is usually placed at the end as the classifier or used to generate numerical output for regression problems exploiting the learned feature map.

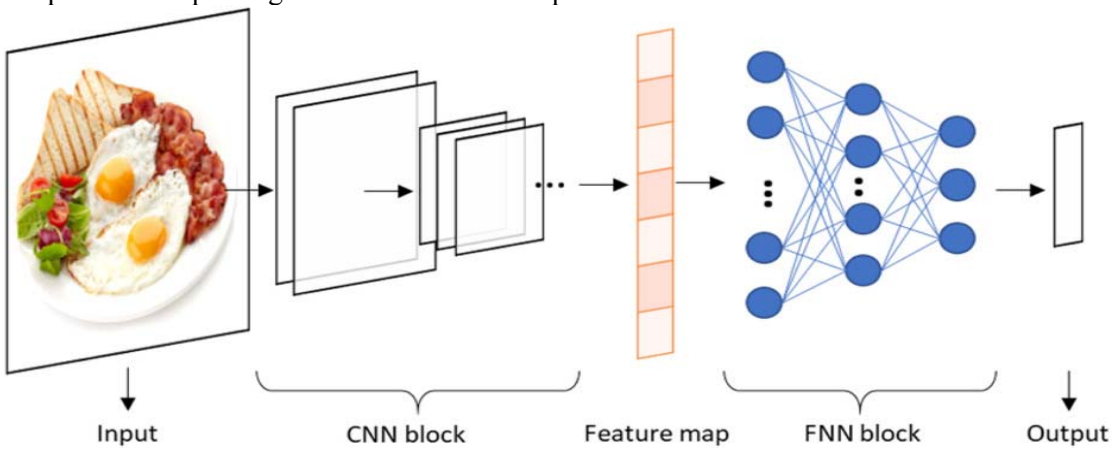

Figure 3: A typical CNN structure for image classification.

Deep learning, as an effective machine learning algorithm, has been widely studied and now attracts more attentions from various fields such as remote sensing, wireless communication, agriculture production, medical science, robotics, healthcare, human action recognition, speech recognition and so on. The structure of a CNN used in this work is presented in below figure. This CNN network consists of an input layer, several composite computing units (CCUs), a fully connected layer and a logistic regression layer (the first sub-figure of below figure). The input layer receives the WP decomposition of a data segment to be processed. A CCU consists of a convolutional layer, a pooling layer, and an activation layer (the second subfigure of below figure). This work used a CNN with two CCUs. There are two key concepts, convolution kernel (CK) and pooling for CNN, which represent the computations in the convolutional layer and the pooling layer respectively.

IV. Computer Aided Detection

A Computer Aided Detection (CAD) system can be described as the use of computers to evaluate the medical images in an automatic or semi-automatic manner. In the field of breast imaging, availability of accurate CAD methods can make a real impact in improving the current breast screening procedures. Reading and understanding breast images requires a well trained and experienced radiologist, and the CAD systems can be effectively used as the "second opinion" and assist radiologists in screening programs. CAD employs computer vision techniques and/or artificial intelligence to deal with radiological and pathology images. In recent years, with the advancement in technology there has been an increase in the use of different imaging modalities for detection and diagnosis of cancers/tumours. Reading and understanding the images from these different modalities require highly experienced and well-trained doctors/radiologists. Moreover, even with well trained experts, there exists high possibilities of intra- and inter-variabilities between the readers. This motivated the use of computers to support radiologists to make accurate diagnosis. The advantage of using a CAD systems is that it can reduce the diagnostic time, reduce inter-observer variations, and can act as a supplementary tool for the radiologists. In recent years, with the advancements in computer technology and data sciences, there has been a lot of interest in exploring deep learning methods for various tasks. In this regard, deep learning methods based on Convolutional Neural Network (CNN) have also gained importance in the field of medical image analysis and the efforts are laid to develop modern CAD systems based on the these newly developed CNN algorithms. This thesis is also a effort in this direction, exploring recent advancements in CNN to facilitate the development of an automated CAD system to assist radiologist in fast and accurate detection of lesions during breast cancer screening.

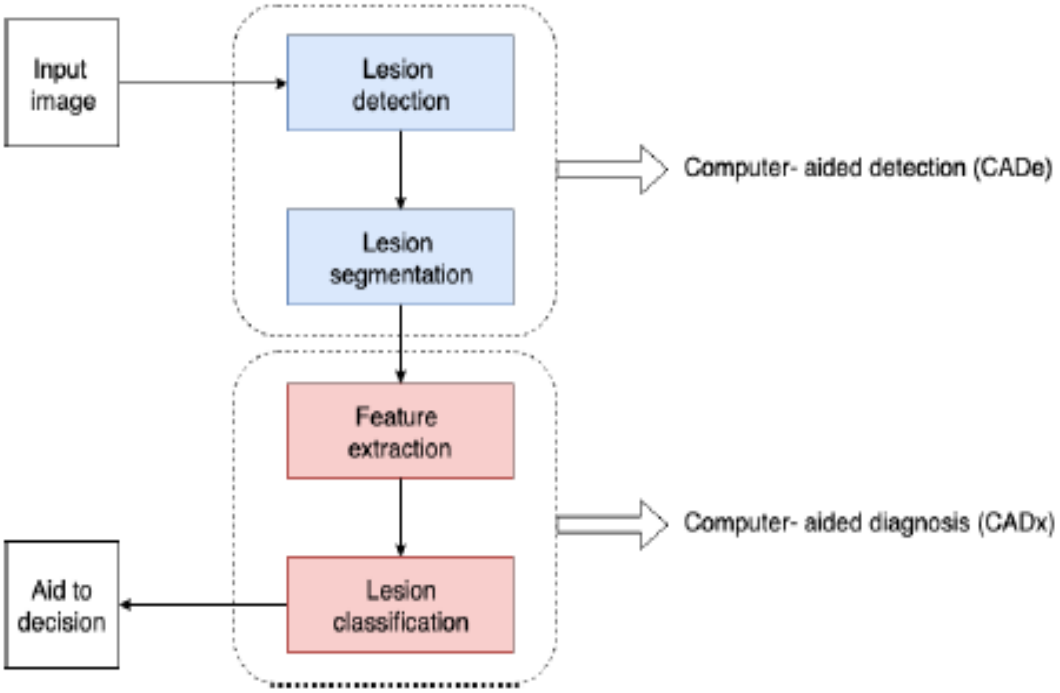

Figure 4: Flow chart showing different steps involved in CAD system and their characterization into CADe and CADx.

V. Conclusion

Cancer is a term used for a group of disorders associated with abnormal cell growth. The abnormal cells have the potential to spread to other parts of the body (metastasis). Cancer can develop in any part of the body; the most common types are prostate, lung and bronchus, breast and colorectal. Cancer is staged according to its extent at the time of diagnosis. Breast cancer is the second leading cause of death for women, so accurate early detection can help decrease breast cancer mortality rates. Computer-aided detection allows radiologists to detect abnormalities efficiently.

References

[1] 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", Advances in Automation, Signal Processing, Instrumentation, and Control, Springer, 2021, pp. 371-384.

[2] P. Esther Jebarani, N. Umadevi, Hien Dang, Marc Pomplun, "A Novel Hybrid K-Means and GMM Machine Learning Model for Breast Cancer Detection", IEEE Access 2021, pp. 146153-146162.

[3] Jian-xue Tian, Jue Zhang, "Breast cancer diagnosis using feature extraction and boosted C5.0 decision tree algorithm with penalty factor", MBE, 2021, pp. 2193–2205.

[4] S. Nirmala Sugirtha Rajini, S. Leena Nesamani, P.Abirami, "Brain Tumor Segmentation from MRI Images using Deep Learning-based CNN with SVM Classifier", International Journal of Grid and Distributed Computing, 2021, pp. 1557-1564.

[5] Deepa B G, S. Senthil, "Predicting Invasive Ductal Carcinoma Tissues in Whole Slide Images of Breast Cancer by Using Convolutional Neural Network Model and Multiple Classifiers in Google Colab", 2021, pp. 1-23.

[6] R Dubey, D Rathore, D Kushwaha, JP Maurya, "An empirical study of intrusion detection system using feature reduction based on evolutionary algorithms and swarm intelligence methods", International Journal of Applied Engineering Research 12 (19), 2017. pp. 8884-8889.

[7] Jignesh Chowdary, Suganya, Premalatha, Asnath Victy Phamila, Karunamurthy, "Machine Learning and Deep Learning Methods for Building Intelligent Systems in Medicine and Drug Discovery: A Comprehensive Survey", 2021, pp. 1-27.

[8] Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Yaman Afadar, Omar Elgendy, " Breast cancer detection using artificial intelligence techniques: A systematic literature review", Artificial Intelligence in Medicine, Elsevier, 2022, pp. 1-26.

[9] Zakaria Senousy, Mohammed M. Abdelsamea, Mona Mostafa Mohamed, Mohamed Medhat Gaber, "3E-Net: Entropy-Based Elastic Ensemble of Deep Convolutional Neural Networks for Grading of Invasive Breast Carcinoma Histopathological Microscopic Images", Entropy 2021, pp. 1-21.

[10] Deepak Kumar Rathore, Dr. Praveen Kumar Mannepalli, "A Review of Machine Learning Techniques and Applications for Health Care ", International Conference on Advances in Technology, Management & Education, 2021, IEEE proceeding, 978-1-7281-8586-6/21.

[11] Mahsa Arabahmadi, Reza Farahbakhsh, Javad Rezazadeh, "Deep Learning for Smart Healthcare-A Survey on Brain Tumor Detection from Medical Imaging", Sensors 2022, pp. 1-27.

[12] Noor Salah Hassan, Adnan Mohsin Abdulazeez, Diyar Qader Zeebaree, Dathar A. Hasan, "Medical Images Breast Cancer Segmentation Based on K-Means Clustering Algorithm: A Review", Asian Journal of Research in Computer Science, 2021, pp. 23-38.

[13] S. Leena Nesamani, S. Nirmala Sugirtha Rajini, "Breast Cancer Detection with Transfer Learning Technique in Convolutional Neural Networks", Design Engineering, 2021, pp. 11102-11109.

[14] Omneya Attallah, Fatma Anwar, Nagia M. Ghanem, Mohamed A. Ismail, "Histo-CADx: duo cascaded fusion stages for breast cancer diagnosis from histopathological images", 2021, pp. 1-39.

[15] Deepak Kumar Rathore, Praveen Kumar Mannepalli, "Diseases Prediction and Classification Using Machine Learning Techniques", AIP Conference Proceedings 2424, 070001 (2022); https://doi.org/10.1063/5.0076768.

[16] Tania Khatun, Md. Ashiqul Islam, "Performance Analysis of Breast Cancer: A Machine Learning Approach", 2021, pp. 1-10.

[17] Yassir Benhammou, Boujemaa Achchab, Francisco Herrera, Siham Tabik," BreakHis based Breast Cancer Automatic Diagnosis using Deep Learning: Taxonomy, Survey and Insights, Neurocomputing September 2019, pp 1-40.

[18] 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.

[19] 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.

[20] 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.

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

[22] Chetna Kaushal, Deepika Koundal, Anshu Singla, "Comparative Analysis of Segmentation Techniques using Histopathological Images of Breast Cancer", Proceedings of the Third International

Conference on Computing Methodologies and Communication (ICCMC 2019) , IEEE explorer . Pp 261-266.

[23] Deepak Kumar Rathore, Dr. Praveen Kumar Mannepalli, "Recent Trends in Machine Learning for Health Care Sector ", International Journal of Innovative Research in Technology and Management, Vol-5, Issue-2, 2021.

[24] Nadeem Tariq, Beenish Abid, Khawaja Ali Qadeer, Imran Hashim, Zulfiqar Ali, Ikramullah Khosa, "Breast Cancer Classification using Global Discriminate Features in Mammographic Images", International Journal of Advanced Computer Science and Applications, Vol 10 No. 2 , 2019. pp 381-387.