



JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Research on Breast Cancer Detection Using Image Processing

Mammography image enhancement with image processing techniques

Pranali A. Shinde, Prof. Nitin N. Mandaogade

Department Of Electronics and Telecommunications, G.H. RAISONI University, Amravati 444701, Maharashtra, India

Guided, Prof. Department of Electronics and Telecommunications, G.H. RAISONI University, Amravati 444701, Maharashtra, India

Abstract: Breast cancer is very common disease in women, all among the patient of cancer breast cancer patient have larger percentage. Where in India situation is very bad because of improper medical facilities the process of breast cancer diagnosis is get delayed. breast cancer accounts for 14% of cancers in Indian women. It is reported that with every four minutes, an Indian woman is diagnosed with breast cancer. Breast cancer is on the rise, both in rural and urban India. A 2018 report of Breast Cancer statistics recorded 1,62,468 new registered cases and 87,090 reported deaths. Because of poverty rate Indian women can't afford medical facilities. Manual physical examination can't detect cancer exactly, this process only helps to find out lumps and tumors present inside breast. The common and cheaper way to diagnose breast cancer is mammography. But mammography image quality is not that much good. This disease was subject to a lot of research to improve the accuracy of its diagnosis and detection. However, the condition is still crowned as one of the most life-threatening diseases posing a threat to the life of one out of six women. The ambiguity of the cause of this disease increases the difficulty of handling such a disease since it makes the preventive approach impossible. Therefore, the last resort that the targets of this disease find themselves with is early detection. Using image processing techniques, this paper will help give an accurate method to detect the disease at its infant stage. In this paper, a detailed explanation will be given regarding the steps followed in image processing: image enhancement, segmentation, and feature extraction that is done using a Convolutional Neural Network (CNN).

Keywords: Breast Cancer, Mammography, Image processing, Mammography image, Image enhancement, Radiology, Cancer

I. INTRODUCTION

Breast cancer is one of the most frequent diagnoses of diseases among women. It can be detected by clinical breast examination, yet the detection rate endures to be very low. Additionally, the abnormal areas that cannot be felt can be quite challenging to check using traditional techniques but can be easily seen on a conventional mammogram or with ultrasound. Mammography is currently the best method for detecting breast cancer at its early stage. The problem with mammography images is they are complex. Thus, image processing and feature extraction techniques are used to assist radiologists in detecting tumors. Features extracted from suspicious regions in mammography images can help doctors to discover the existence of the tumor in real-time thus speeding up the treatment process. Detecting breast cancer can be quite a challenging job. Especially, as cancer is not a single disease but a collection of multiple diseases. Thus, every cancer is different from every other cancer that exists. Also, the same drug may have different reactions to similar types of cancer. Thus, cancer varies from person to person. Depending on only one technique or one algorithm to detect breast cancer may not provide us with the best possible result. As one cancer differs from another, similarly every breast appears differently from another. The mammography image can also be compromised if the patient has undergone some breast surgery. Breast Cancer has been a big topic in the research field for the last two decades. It has been well funded medical research topic across the globe. Many people have been cured of it, due to early detection. Still, the progress in diagnosis and treatment for it remains expensive and time-consuming. Automated detection of mass remains a difficult task, this might be because every cancer is different like its host and each requires customized medication to be cured. So, a lot of work is still left to be done.

II. LITERATURE REVIEW

Detecting macrocalcification in dense breast tissue can be a difficult task as both tends to depict white pixel on the mammogram. The number of false positive cases on dense breast tissue are higher. Indicators of cancer symptom are generally, masses and microcalcifications. Detecting masses are more challenging task than detection of microcalcifications. As their size and shape varies in large variation and they often exhibit poor image contrast. The utilization of grouping frameworks in classification and pattern recognition system, in medical diagnosis, especially cancer diagnosis is growing rapidly. Evaluation and decision making

based on machine learning for medical diagnosis is a key factor. Intelligent classification algorithm may help doctor in identifying symptoms that may not be possible through traditional approaches⁸. Any Image processing and analysis applications would require a unique function for alignment of feature for classification and segmentation. Mainly texture features and statistical features are of more suitable in pattern recognition area to find this alignment. Screening Mammography is the easiest and affordable way to diagnosis for breast cancer. The mammography image is checked through several techniques like finding edges, smoothing border, finding structures & shapes among matrixes. Finally finding the size distribution of tissues in an Image without explicitly segmenting each object. Digital mammography is the standard procedure for breast cancer diagnosis, various classification problem is applied on the digital mammography image. Various features are extracted as per standard procedure for breast cancer diagnosis. These features are calculated from the sensitive part of the breast to avoid any unwanted features to affect the classification problem. Area of tumor is calculated by the Maximum Likelihood Estimation (MLE). All the features extraction techniques are applied on the stored database image¹³. This paper mainly studies the multiple image processing algorithms which can be extensively used for finding cancerous cells. The techniques in computer aided mammography includes image.

pre-processing, image segmentation, feature extraction, feature selection and classification. Further developments are required to extract more features to find pattern in tumor to have a better understanding on them. Texture analysis method can be used to classify between benign and malignant masses by means to identify the micro-calcification in the mammography. Research In the Field of Cancer Many research has been done in the field of image processing to find the cancer. Yet, the accuracy rate lies between 75% - 92%. Thus, there is still a gap of 8% to 25% of accuracy to be achieved. The new research analysis and techniques to find the cancerous cells and eradication methodology to cure the cancer from any person. However, even cancer cells have evolved them to hide from drugs and medications. As cancer cells are immortal, they are not affected by the immune system. There is research for curing the cancer tumor, the methods are as follows. CRISPR Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) also known as Cas9 is a simple and powerful gene editing tool. Genetic engineering has allowed cancer researchers to screen the drug to target the cancer cells in efficient manner. There is also a vast door for direct treatment of cancer through gene interference or activation. Artificial Immune System Artificial Immune Systems resembles the natural properties of our biological immune system. Natural immune system has the property to pattern matching which is used to distinguish between normal and abnormal cell¹⁹. Nano Technology Nano technology can give fast and delicate location of cancer cells in the breast tissues. Empowering researchers to identify molecular changes not withstanding when they happen just in a smaller number of cells. The nanodevices can be programmed to annihilate infected cells and kill those infected cells¹⁷.

Radiology today is a diagnostic and therapeutic support specialty since it includes from conventional radiology to mammography, computed tomography, ultrasound, magnetic resonance, nuclear medicine, positron emission tomography, and radiotherapy.

In radiology, technologies have been advancing very fast and are becoming increasingly indispensable in the continuous development of this area. The various procedures require the radiologist an intersection of his technical-scientific knowledge with others, such as administrative ones, for the adequate performance of his function. It must have an “integrated and complementary multidisciplinary action, incorporating complex processes and cutting-edge technologies, with large investments in equipment, techniques and inputs”.

The early breast lesions are detected, the more mastectomy is reduced, and the chances of survival increase. Even considered the “gold standard” method in detecting breast cancer, mammography can present a false positive that leads to unnecessary biopsies because it has low specificity despite its high sensitivity. Wild and Neal were the pioneers in proposing the use of ultrasound in the examination of the breasts. Since then, it has been proven that ultrasound imaging techniques can identify many of the cancers missed by mammographic techniques, especially those that occur in women with dense breasts. Furthermore, ultrasound imaging has the advantage of being noninvasive, low-cost, and there is no need for ionizing addition. Despite the above advantages, it largely depends on developing efficient segmentation algorithms.

There is a frequency distribution/frequency diagram graphically represented in columns/rectangles from a dataset that, beforehand, is tabulated and segmented into uniform classes, known as a histogram. In an image, this “indicates the number or percentage of pixels that the image has in a certain level of grey or color”. And so, it generates an image quality indicator referring to contrast and light intensity. Significant data are obtained from the highlighted regions/objects with binary images, segmented into objects and background. Through digitization, the image relates to the truth with a virtual potency. A simulation with an infinite property makes it an image imagination. With a mouse click, the real is configured.

In the binary vision system, thresholding stands out as an efficient and simple-to-implement strategy that uses pixel intensity as a separator; it is a type of segmentation (local threshold, global threshold, and multiple thresholds), which is the process that fragments the image into distinct regions, each with pixels with similar attributes. “A system created to perform digital image processing is usually composed of five elements: image acquisition, storage, processing itself, communication and display of the final result of the process”. Particularly in healthcare, the computer diagnostic system automatically or semiautomatically detects anomalies in imaging exams. The image is segmented, subdividing it, to distinguish the “object of interest from the image’s background”, which occurs through the segmentation that subdivides the image. Using thresholding, we define “a threshold value capable of separating the object of interest from the background and then an image in gray levels becomes a binary image”. The information that is obtained from these techniques helps the doctor’s work. These are just some concepts and techniques for creating digital medical images.

Wang et al. (2019) state that the development of equipment quality control techniques reached “conventional radiology, dental, conventional mammography services, processors that should be controlled daily, fluoroscopy equipment with image intensifier and some initiative in computed tomography”. It was imperative to establish mastery over the irradiation of patients. The equipment began to be calibrated and operated in more controlled environments, which triggered a more significant number of shielding calculation services and the performance of radiometric surveys. With time and the advancement of digital technology, the

detectors capture radiological images with properties capable of leading “to the evaluation of different magnitudes from those used in screen-film systems”.

In this sense, Fiorica considers that in a context of more restrained budgets, it is essential that the acquisition of equipment “be subjected to a rigorous cost-benefit evaluation process”. In addition, he comments that the incessantly growing interventional radiology enables minimally invasive therapies based on more accurate imaging means, unlike past decades in which “radiology was an exclusively diagnostic activity and without clinical contact”.

Yousif M.Y Abdallah et.al Enhancement of mammography images is considered as powerful methods in categorization of breast normal tissues and pathologies. The digital image software gives a chance to improve the mammography and increase their illustration value. The image processing methods in this paper used contrast improvement, noise lessening, texture scrutiny and portioning algorithm. The mammography images were kept in high quality to conserve the quality. Those methods aim to augment and hone the image intensity and eliminate noise from the images. The assortment factor of augmentation depends on the backdrop tissues and type of the breast lesions; hence, some lesions gave better improvement than the rest due to their density. The computation speed examined used correspondence and matching ratio. The results were 96.3 ± 8.5 ($p > 0.05$). The results showed that the breast lesions could be improved by using the proposed image improvement and segmentation methods.

Prannoy Giri et.al Breast Cancer is one of the significant reasons for death among ladies. Much research has been done on the diagnosis and detection of breast cancer using various image processing and classification techniques. Nonetheless, the disease remains as one of the deadliest diseases. Having conceived one out of six women in her lifetime. Since the cause of breast cancer stays obscure, prevention becomes impossible. Thus, early detection of tumours in the breast is the only way to cure breast cancer. Using CAD (Computer Aided Diagnosis) on the mammographic images is the most efficient and easiest way to diagnose breast cancer. Accurate discovery can effectively reduce the mortality rate brought about by using mamma cancer. Masses and microcalcifications clusters are important early symptoms of possible breast cancers. They can help predict breast cancer in its infant state. The image for this work is being used from the DDSM Database (Digital Database for Screening Mammography) which contains approximately 3000 cases and is being used worldwide for cancer research. This paper quantitatively depicts the analysis methods used for texture features for the detection of cancer. These texture features are extracted from the ROI of the mammogram to characterize the microcalcifications as harmless, ordinary, or threatening. These features are further decreased using Principal Component Analysis (PCA) for better identification of Masses. These features are further compared and passed through the Back Propagation algorithm (Neural Network) for a better understanding of the cancer pattern in the mammography image.

Arpita Joshi and Dr.Ashish Mehta compared the classification results obtained from the techniques i.e. KNN, SVM, Random Forest, Decision Tree (Recursive Partitioning and Conditional Inference Tree). The dataset used was Wisconsin Breast Cancer dataset obtained from UCI repository. Simulation results showed that KNN was the best classifier followed by SVM, Random Forest and Decision Tree.

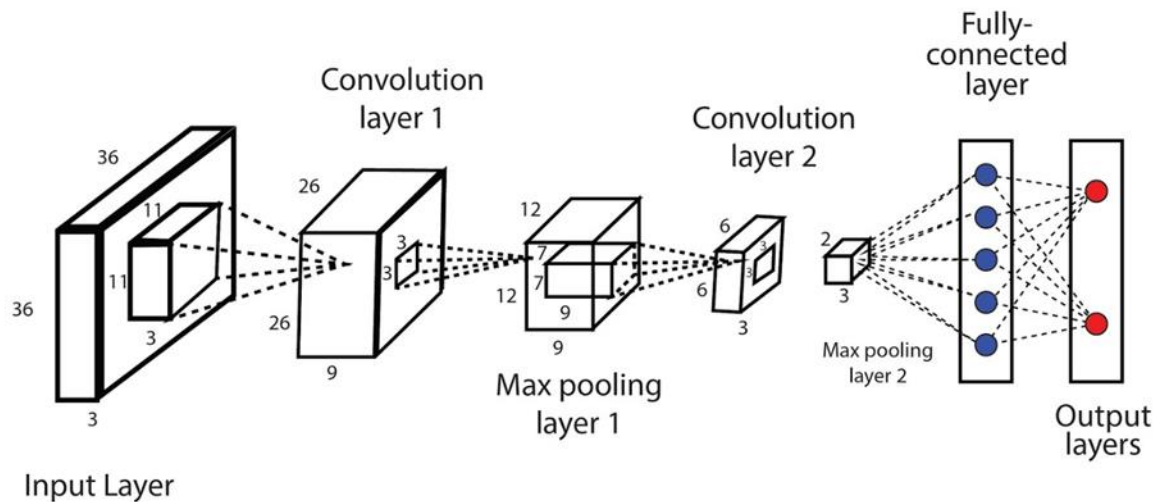
David A. Omondiagbe, Shanmugam Veeramani, Amandeep S. Sidhu, investigated the performance of Support Vector Machine, Artificial Neural Network and Naïve Bayes using the Wisconsin Diagnostic Breast Cancer (WDBC) Dataset by integrating these machine learning techniques with feature selection/feature extraction methods to obtain the most suitable one. The simulation results showed that SVM-LDA was chosen over all the other methods due to their longer computational time.

Kalyani Wadkar, Prashant Pathak and Nikhil Wagh, did a comparative study on ANN and SVM and integrated various classifiers like CNN, KNN and Inception V3 for better processing of the dataset. The experimental results and performance analysis concluded that ANN was a better classifier than SVM as ANN proved to have a higher efficiency rate.

For R. Thyagarajan and S. Murugavalli (2012), “Radiology has been undergoing profound changes since the end of the 20th century. In the coming years, it is expected the emergence of multiple technologies that will certainly be disruptive, but that will also create new opportunities”.

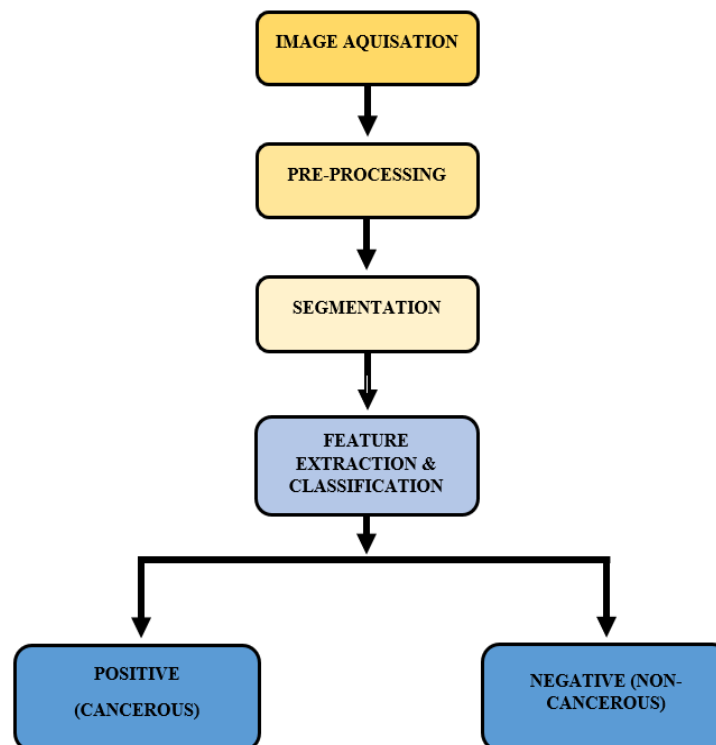
III. METHODOLOGY

The proposed methodology will help us to detect breast cancer at a faster rate. CNN being a complex and complicated classifier can extract vital features automatically without depending on preprocessing. It is more proficient because it filters the important parameters and also is flexible, being capable of working exceptionally well on image data. The main focus of our paper is to differentiate between malignant and benign tumor using Convolution Neural Network with Kera’s in the backend and then analyses the result to see how the model can be useful in a practical scenario.

Block Diagram**Figure1. Block Diagram of CNN**

The following steps are performed for model building and evaluation

- Importing all the essential libraries.
- Making a dictionary of images and labels.
- Labels are based on image category.
- Normalization of the image set.
- Splitting data into training and testing sets.
- Building Architecture of the model (CNN).
- Testing Model.

ARCHITECTURE DIAGRAM OF BREAST CANCER DETECTION**Figure 2: Flow Chart****IV. SYSTEM REQUIREMENT**

1. CPU I3 processor
2. RAM 4GB
3. OS window 8
4. ROM 250 GB

SOFTWARE REQUIREMENT

1. Python version 3.8
2. Machine Learning

IMPLEMENTATION

To implement the process of breast cancer detection using image processing first turn on python version 3.8 Select the original Image from dataset

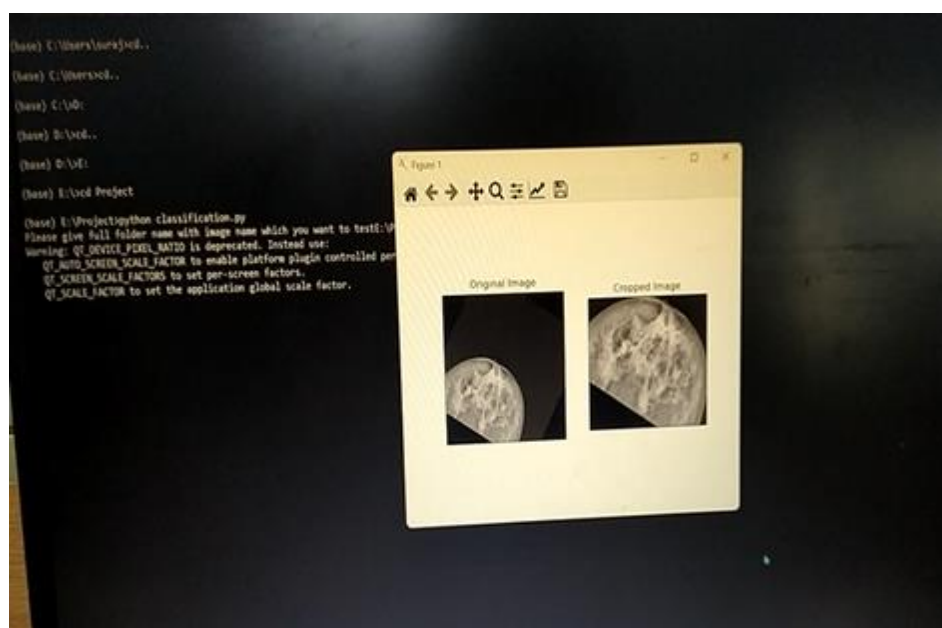


Figure 3. Original image from data set

RESULT

The below Images show that Benign and Malignant breast cancer



Figure 4: Shows that masses of Benign Breast Cancer

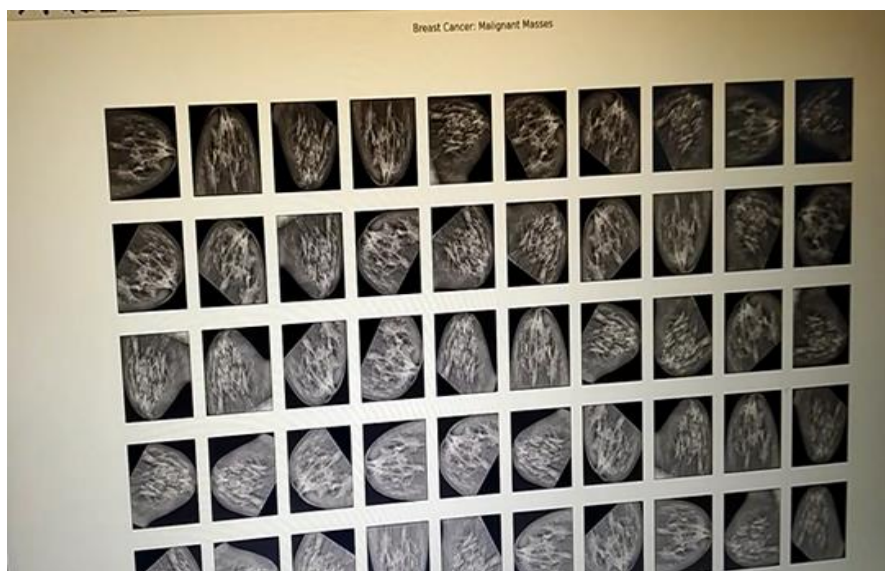


Figure 5. shows that masses of Malignant Breast Cancer

The below Images show that Benign and Malignant breast cancer

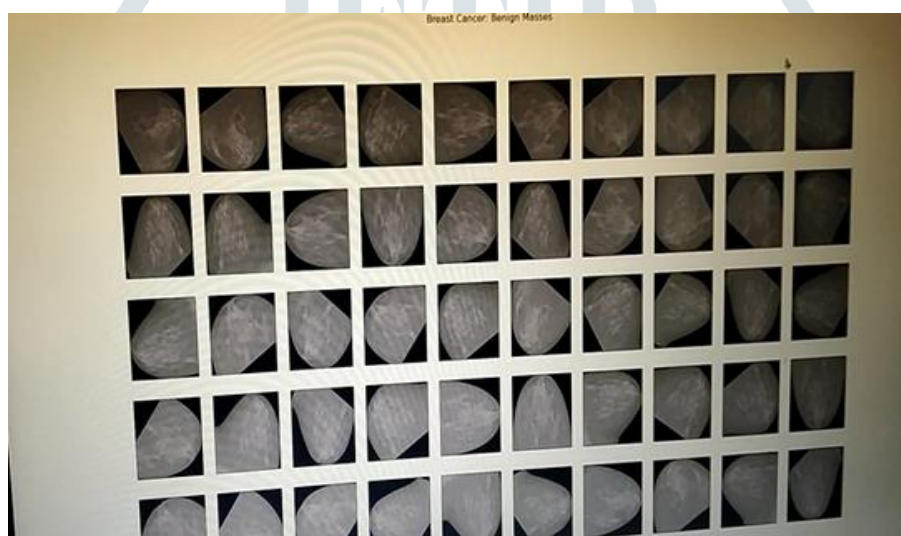


Figure 6. Shows that masses of Benign Breast Cancer

V. CONCLUSION

As we know that in medical diagnostic techniques radiology and mammography is used to diagnose breast cancer. But the denoises of Breast cancer using uncleaned mammographically image is not possible. The directional view of breast helps but still the image quality gets compromised. But according to our research we proudly can say that we are now able to improve this image quality by using image processing. By doing this research we can conclude that the existing system can get improved by using image processing. In the diagnosis process, due to the wide range of features associated with breast abnormalities, some abnormalities may be missed or misinterpreted. There are also a number of false positive findings and therefore a lot of unnecessary biopsies may be required. Computer-aided detection and diagnosis algorithms have been developed to help radiologists give an accurate diagnosis and to reduce the number of false positives. In this study, typical steps in image processing algorithms have been extensively studied. The techniques in the field of computer-aided mammography include image pre-processing, image segmentation techniques, feature extraction, feature selection, classification techniques, and features for mammograms. Texture features are obtained to distinguish between normal cells and cancerous cells. Cancer is one of the oldest diseases, and much research has been carried out in this field. Cancer is not a single disease but rather a collection of multiple diseases thus a single medicine to cure cancer is not possible.

FUTURE SCOPE

The future plan is to validate the model with other datasets that include new ultrasound images.

VI. ACKNOWLEDGMENT

We are deeply grateful to all those who contributed to the success of this review research paper. First and foremost, we would like to thank our primary supervisor **PROF. NITIN N. MANDAOGADE** for their guidance, support, and encouragement

throughout the entire process. Their mentorship and expertise were invaluable in helping us to shape the direction of our review research and to bring our ideas to fruition. I would also like to thank the organizations and individuals who provided me a support for this review research, including G.H. RAISONI University Amravati, Maharashtra, India. Without their generous contributions, this review research would not have been possible. Overall, this research project would not have been possible without the support and contributions of so many people. We are deeply grateful to all of those who helped to make this project a reality, and we hope that our findings will make a meaningful contribution to the field.

REFERENCES

- [1] Zahra abdolali kazemi, “Diagnosis of breast cancer using image processing with SVM and KNN” Vol 13 No 1 Mar-Apr 2021
- [2] Saif Ali et.al “Identification of Cancer Disease Using Image Processing Approaches” Volume 9, Issue 2, April 2020, Pages: 6-15.
- [3] Mutiullah et.al “Lung Cancer Detection Using Digital Image Processing Techniques: A Review” Vol. 38, No. 2, 351-360 April 2020.
- [4] Yousif M.Y Abdallah et.al “Breast cancer detection using image enhancement and segmentation algorithms” 2021 Volume 29, Issue 20
- [5] Prannoy Giri et.al “Breast Cancer Detection using Image Processing Techniques” DOI: <http://dx.doi.org/10.13005/ojcs/10.02.19>
- [6] Arpita Joshi and Dr. Ashish Mehta “Comparative Analysis of Various Machine Learning Techniques for Diagnosis of Breast Cancer” (2020).
- [7] David A. Omondiagbe, Shanmugam Veeramani and Amandeep S. Sidhu “Machine Learning Classification Techniques for Breast CancerDiagnosis” (2021).
- [8] Kalyani Wadkar, Prashant Pathak and Nikhil Wagh “Breast Cancer Detection Using ANN Network and Performance Analysis with SVM” (2020).
- [9] Anji Reddy Vaka, Badal Soni and Sudheer Reddy “Breast Cancer Detection by Leveraging Machine Learning” (2020).
- [10] Monika Tiwari, Rashi Bharuka, Praditi Shah and Reena Lokare “Breast Cancer Prediction using Deep Learning and Machine Learning Techniques”

