JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Machine Learning Approach for Breast Cancer Detection

Raza Mohsin [1], Ramesh Kumar (HOD) [2], Pradip Kumar Yadava[3]

[1] M.Tech Scholar, [2,3]Assistant Professor

[1,2] Department of Computer Science and Engineering, K. K. University Nalanda (Bihar)

Abstract : Indian ladies are most at risk for breast cancer, as it accounts for 27% of all females cancers. One woman out of 28 will be infected with breast cancer during their lifetime. the occurrence of cancer is 1 out of 22, with the incidence in urban regions being 1 out of 60. It's one of the most widely recognized diseases among ladies and also one of the leading causes for malignant growth in women after lung cancer. In this proposed work, the goal is to improve the detection and classification of breast cancer. Breast cancer occurs when cells start developing too much, which leads to a malignant growth. A Median Filter can be used to clean an image and scrutinize it more closely. The Median Filter method is a way to clear the noise in images during analysis. This helps with better examination of the image. After Image segmentation, the image is classified using a k-Means approach to determine breast cancer. The results are better than existing approaches and promising for the future.

***IndexTerms* – Breast Cancer Detection , Machine Learning, Median Filter , K-Means, Gaussian mixture model.**

I. INTRODUCTION

Breast cancer is the most prevalent type of cancer in women. It's more often diagnosed in women than any other type of cancer, and it was the world's most prevalent cancer in 2020 with over 685 000 deaths globally. Early detection and new modes of treatment have increased survivals rates from one breast cancer diagnosis to death.[1]

Breast cancer can't be caught like other types of cancer due to an infection. There is no evidence that a virus or bacterial infection is linked to breast cancer. The factors that increase the risk of breast cancer include being female, having a family history of the disease, being over forty years old and obese.[1]

Some hormonal factors can also influence the risk of developing breast cancer as well. Factors that increase your risk are using alcohol or tobacco, low levels of vitamin D, starting menstruation late or having children at an older age, taking hormone replacement therapy (HRT) after menopause, and radiation exposure.[2]

Risk factors for breast cancer include age, obesity, harmful use of alcohol, family history of each cancer, reproductive history, tobacco use, and postmenopausal hormone therapy.[2]

1% of all breast cancers happen in men, who treat their cancer like women do. If a person has modifiable risk factors and can reduce them, they'll significantly reduce their risk of developing breast cancer. [3]

For some, a family history of breast cancer increases the risk of developing breast cancer, but only a minority of women are affected. Western countries have higher rates of cancer than India, but rates in India are rising because they're increasing migration from rural to urban areas and changes in lifestyles.[3]

Artificial intelligence algorithms could be a solution to overdiagnosis by detecting cancers, including interval cancers. These AI algorithms are capable of analyzing mammograms and can improve population breast cancer screening programs- if they can detect cancer without already classifying it as cancerous. [4]

Machine learning has been popular in the realm of healthcare. It helps with early diagnosis and analysis of breast cancer by looking at size and shape of tumors.[4]

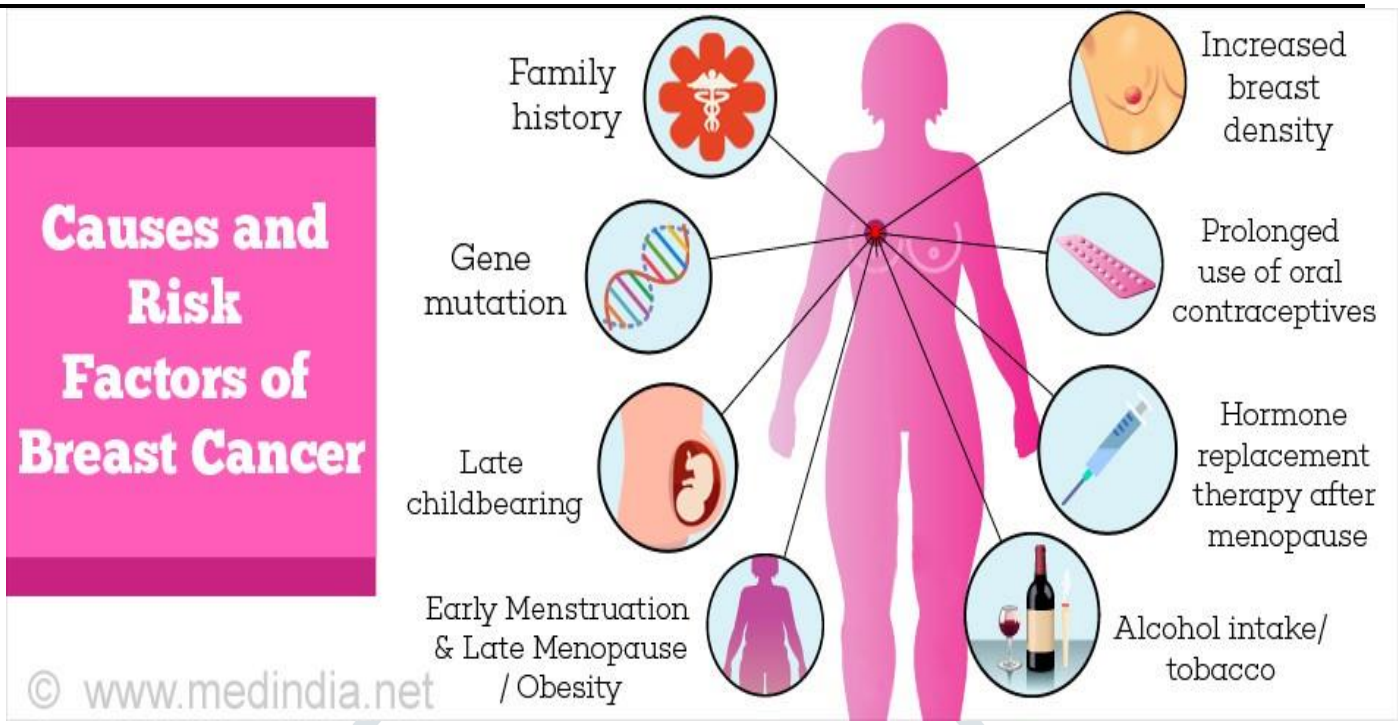

Fig 1. Breast Cancer Causes

Breast cells sometimes grow abnormally, and healthy cells may accumulate. These cancerous cells divide rapidly and can form a mass in your breast. If these cancerous cells have spread to other parts of the body, it means that the disease has metastasized. Breast cancer can begin in the ducts, lobules or any other area. [5]

It's is likely that a persons genetic makeup and the environment have an impact on their likelihood to develop cancer. Scientists are still investigating why some people with risk factors never develop cancer, whereas others who don't have any risk factors do. Scientists have found a number of genes that can increase one's likelihood of getting breast cancer, such as the BRCA1 and BRCA2 genes. These two genes also increase the risk of ovarian cancer. [5]

II. Literature Review

S. Sayed, S. Ahmed and R. Poonia [6] The vast majority of breast cancer research is currently aimed at finding a cure rather than detection. Supports exist for almost all aspects of the disease. Even though there are limitations to machine learning, professionals have found new uses for it in many fields. Some professionals have developed forecasts that use machine learning and have the goal of dealing with cases where data on certain conditions is not available or where the data has already been collected.

T. Liu, L. Zhou and W. Rong [7] To investigate the correlation between the expression of PETA-3 and clinicopathological features in breast cancer. Results: Of forty cases with cancer, they determined that there was a strong correlation between PETA-3 expression and clinicopathological features.

T.Kao et. al [18] T. Kao et. al [8] discuss the properties of tumors and how they affect healthy tissue, as well as other types of cancers. Malignant cells can be detected by testing for Electric Impedance Spectroscopy and most types of cancer can be found by DBT mammography or 3D mammograms.

H. Melody, H. Watanabe, X. Xiao and T. Kikkawa [9] This study explores the impact of air holes between a bosom's surface and recieving wires of drive radar-put together bosom malignancy locator with respect to confocal imaging. The device can identify a cancerous growth and identify it more precisely than a human can under difficult conditions.

Umesh D R and B. Ramachandra [10] In the created countries, breast cancer is the most common type of malignant growth among women. It can occur again in the few years after the treatment of breast cancer and mainly it happens in women who have suffered from breast disease before.

S. Ara, A. Das and A. Dey [11] have discovered there are two types of breast tumors: benign and malignant. Malignant tumors grow rapidly and can be fatal, while benign ones do not pose any real risk of death. It is essential for doctors to diagnose breast cancer as early as possible. Machine learning algorithms have been shown to be more efficient than our brains when it comes to diagnosing breast cancer. One study found that their best performance was achieved by logistic regression, and another found the Random Forest and Support Vector Machine classifiers had higher accuracy than any other algorithm.

III. Proposed Concept

The proposed algorithm will explain the steps which are followed for the task for the breast cancer detection.

Step 1: Read the Selected Input Image.

Step 2: Browse for the Image and select the image related to breast cancer.

Step 3: Image gets Displayed as Figure in the form.

Step 4: Click on Median Filter button, it will clear the noise in order for the better examination of the images.

Step 5: The Image without the noise will get displayed and the next click on the Segmentation of the image.

Step 6: The image segmentation follows the Gaussian mixture model approach for the image segmentation.

Step 7: The result of that, will be sent for the classification, which performed the classification of the cancer using the k-Means classification approach, which classify the images on the basis of the feature extraction related to the shape, thickness etc. and classify into the Benign breast and Malignant tumors.

IV. Implementation and Result Analysis

The Implementation of the proposed model is done in MATLAB.

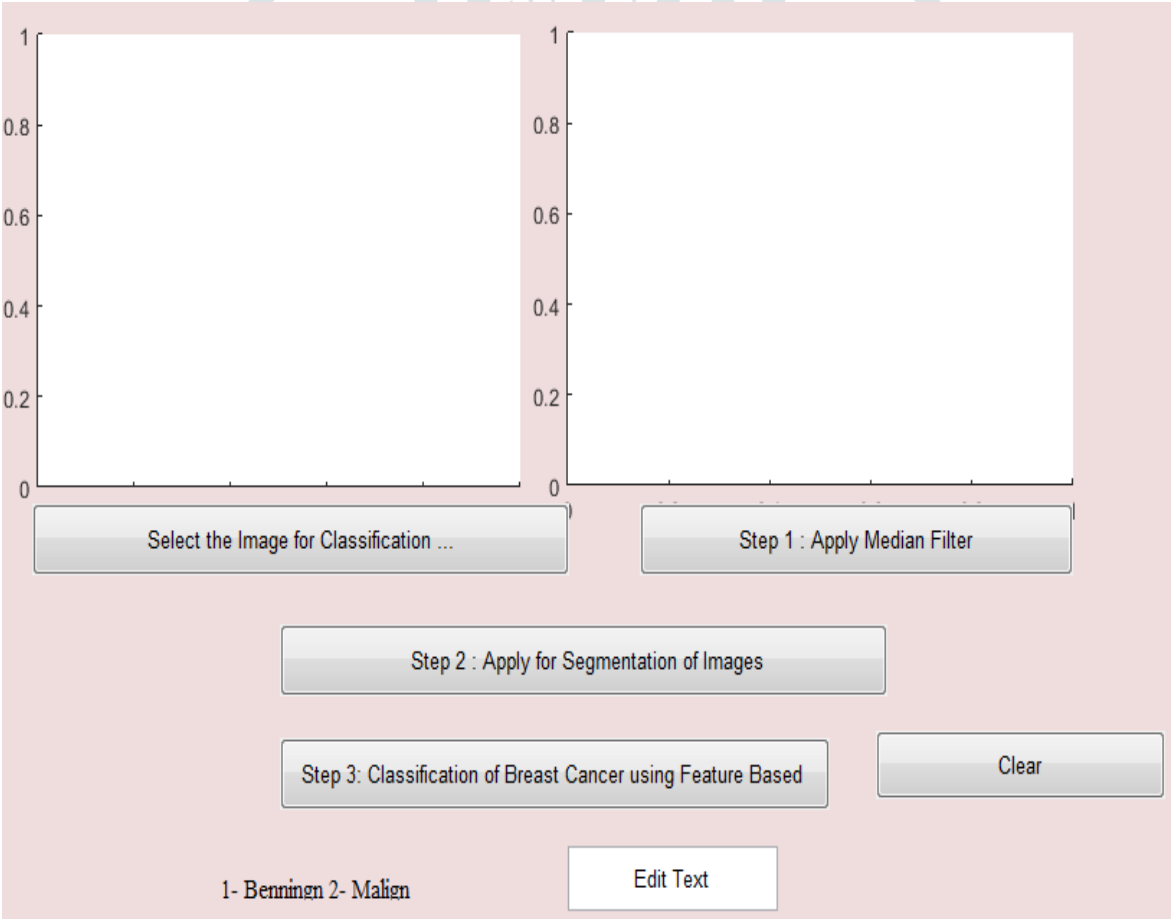

Fig 2. Proposed Implementation

The concept of Classification of the Cancer images,

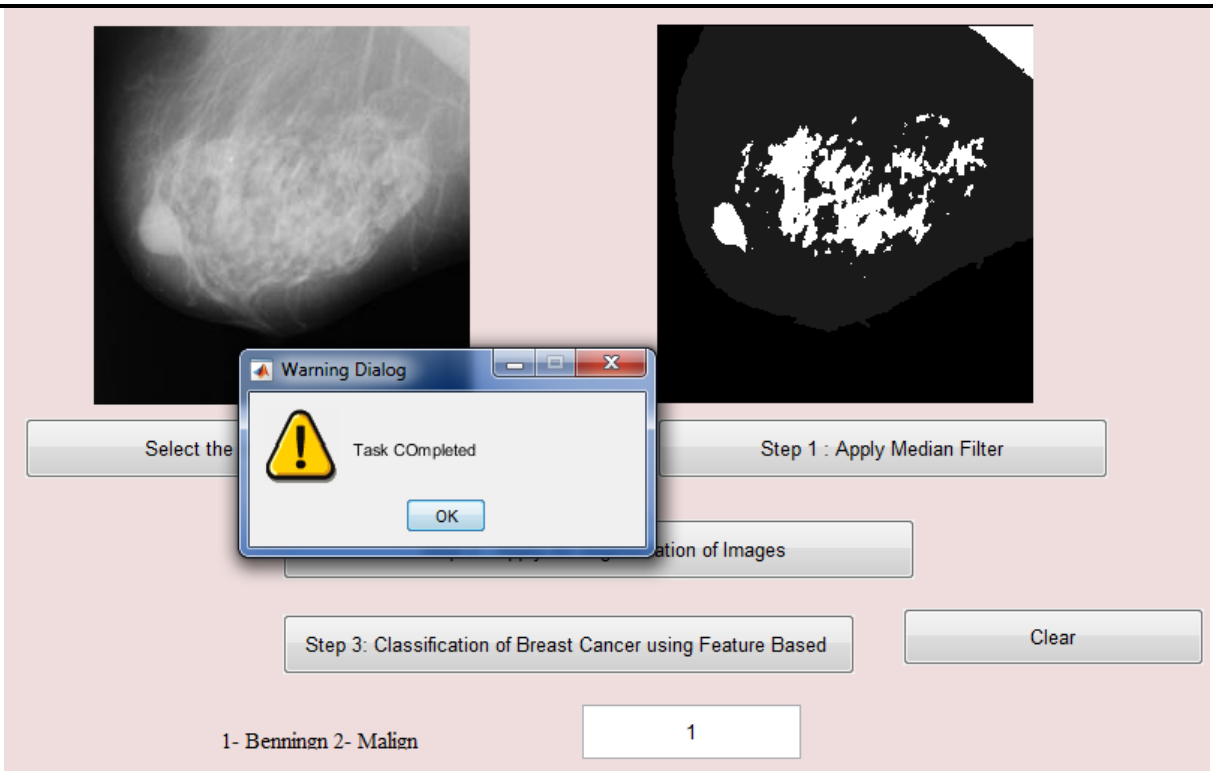

Fig 3. Cancer Classification

4.1 Benign Breast

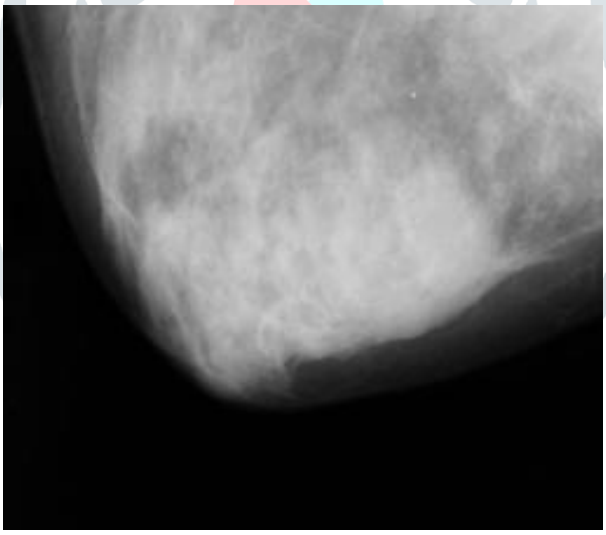

Fig 4. Sample 1

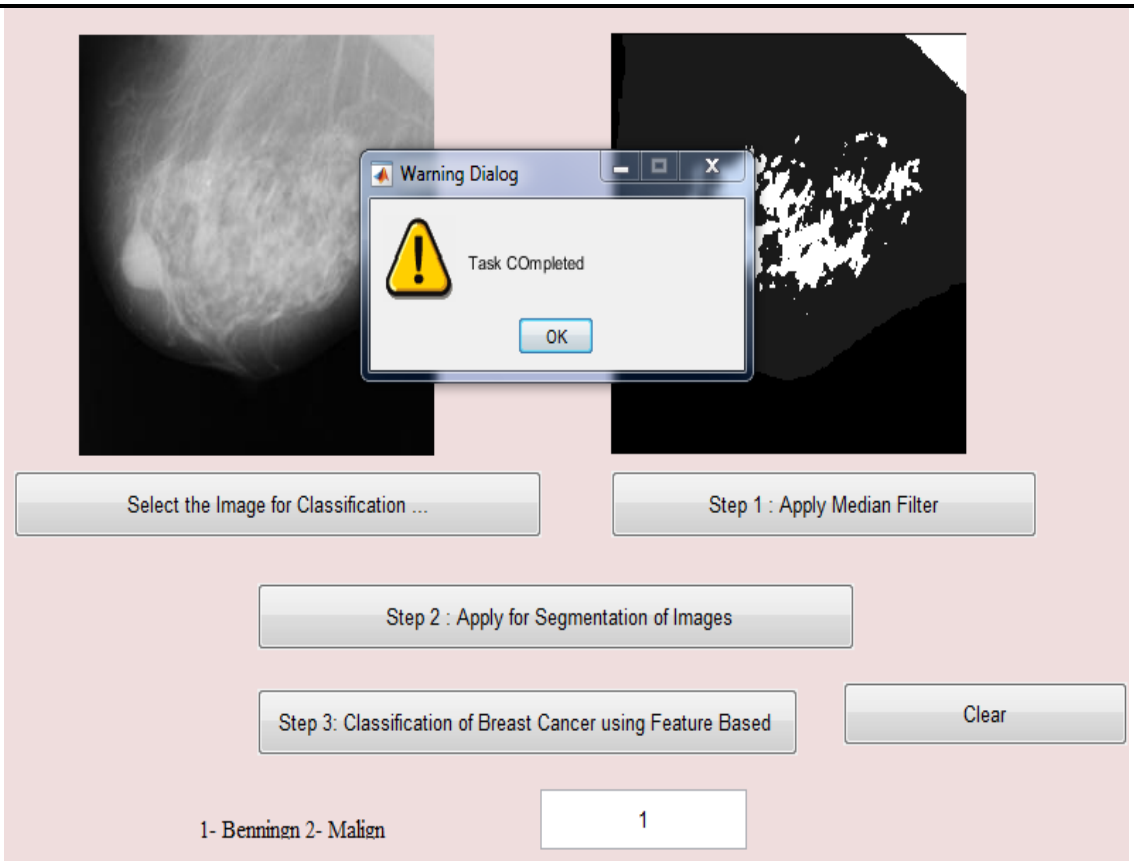

Fig 5. Test Outcome Sample 1

Table 1 Benign Breast Cancer Detection Accuracy

Classification Result	Test Accuracy %	False Positive
Benign Breast	97%	3%

4.2 Malignant Tumours

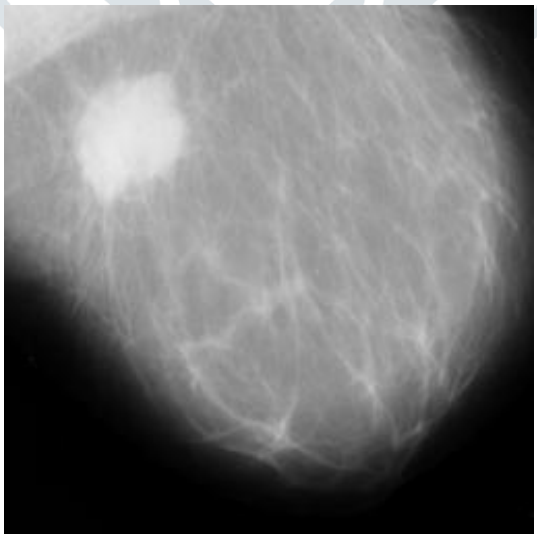

Fig 6. Sample 2

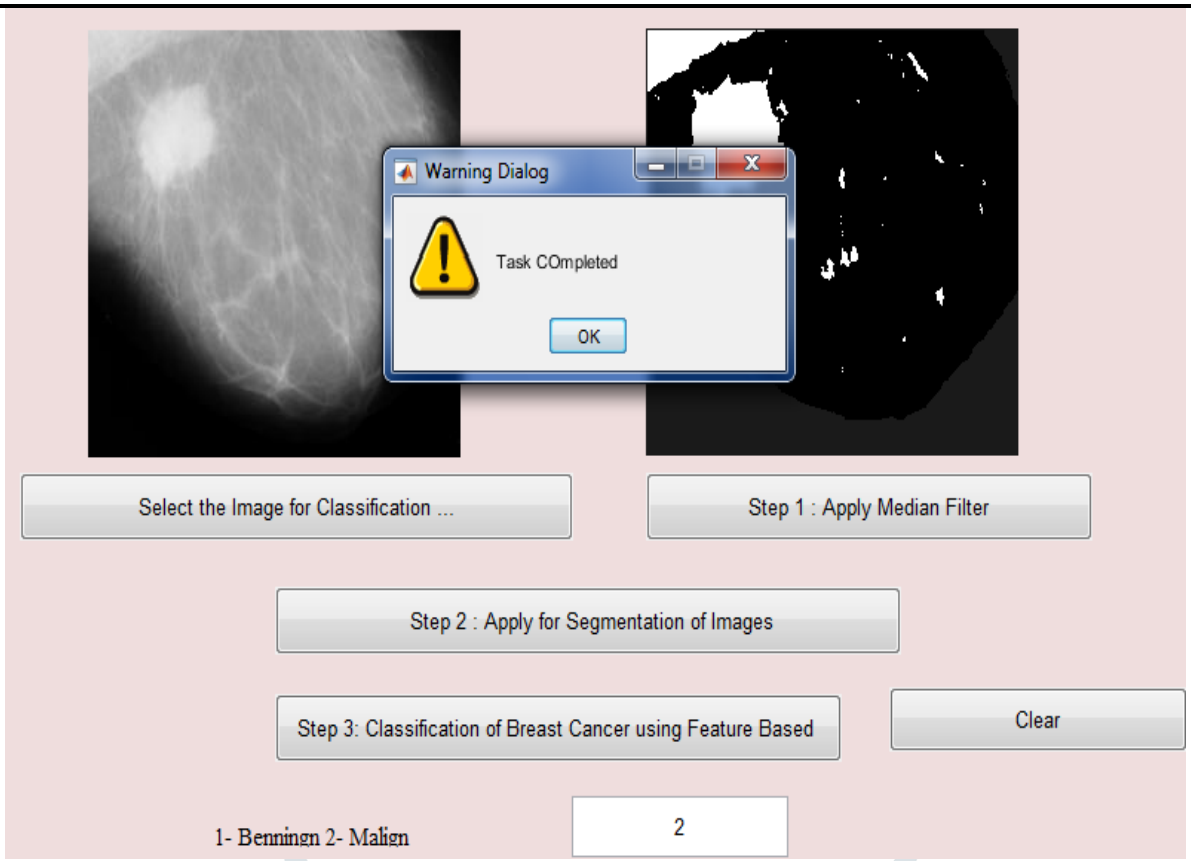

Fig 7. Test Outcome Sample 2

Table 4.2 Accuracy Table for Malignant Tumours

Classification Result	Test Accuracy %	False Positive
Malignant Tumors	98%	2%

S. Ara, A. Das and A. Dey, "Malignant and Benign Breast Cancer Classification using Machine Learning Algorithms,2021, applied various machine learning algorithms and they have achieved the accuracy of about 96.5% for detection of Malignant and Benign Breast Cancer detection.

V. Conclusion

It's one of the most widely recognized diseases among ladies and also one of the leading causes for malignant growth in women after lung cancer. In this proposed work, the goal is to improve the detection and classification of breast cancer. Breast cancer occurs when cells start developing too much, which leads to a malignant growth. A Median Filter can be used to clean an image and scrutinize it more closely. The Median Filter method is a way to clear the noise in images during analysis. This helps with better examination of the image. After Image segmentation, the image is classified using a k-Means approach to determine breast cancer. The results are better than existing approaches and promising for the future.

References

1. S. Reis *et al.*, "Automated Classification of Breast Cancer Stroma Maturity From Histological Images," in *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 10, pp. 2344-2352, Oct. 2017.
2. T. Botterill, T. Lotz, A. Kashif and J. G. Chase, "Reconstructing 3-D Skin Surface Motion for the DIET Breast Cancer Screening System," in *IEEE Transactions on Medical Imaging*, vol. 33, no. 5, pp. 1109-1118, May 2014.
3. N. M. Zain, V. P. Seriramulu, K. K. Chelliah, N. Soin and S. Arasaratnam, "Association of bone mineral density and mammographic breast density in premenopausal and postmenopausal Malaysian women: Study on women with and without breast cancer," *2016 International Conference on Bio-engineering for Smart Technologies (BioSMART)*, Dubai, 2016, pp. 1-4.

4. A. Portieri, M. Grootendorst and T. Fitzgerald, "Intra-operative terahertz probe for detection of breast cancer," *2015 8th UK, Europe, China Millimeter Waves and THz Technology Workshop (UCMMT)*, Cardiff, 2015, pp. 1-2.
5. M. Atif, M. S. AlSalhi, S. Devanesan, V. Masilamani, K. Farhat and D. Rabah, "Spectral characterization of breast cancer," *2014 11th Annual High Capacity Optical Networks and Emerging/Enabling Technologies (Photonics for Energy)*, Charlotte, NC, 2014, pp. 199-201.
6. S. Sayed, S. Ahmed and R. Poonia, "Holo entropy enabled decision tree classifier for breast cancer diagnosis using wisconsin (prognostic) data set," *2017 7th International Conference on Communication Systems and Network Technologies (CSNT)*, Nagpur, 2017, pp. 172-176.
7. T. Liu, L. Zhou and W. Rong, "Real-time PCR and Immunohistochemistry Detecting the Expression of PETA-3 in Breast Cancer and the Clinical Significance," *2012 International Conference on Biomedical Engineering and Biotechnology*, Macao, 2012, pp. 864-867.
8. T. Kao *et al*., "Regional Admittivity Spectra With Tomosynthesis Images for Breast Cancer Detection: Preliminary Patient Study," in *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1762-1768, Dec. 2008.
9. H. Song, H. Watanabe, X. Xiao and T. Kikkawa, "Influence of Air-gaps between Antennas and Breast on Impulse-Radar-Based Breast Cancer Detection," *2019 13th European Conference on Antennas and Propagation (EuCAP)*, Krakow, Poland, 2019, pp. 1-2.
10. Umesh D R and B. Ramachandra, "Association rule mining based predicting breast cancer recurrence on SEER breast cancer data," *2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, Mandya, 2015, pp. 376-380.
11. S. Ara, A. Das and A. Dey, "Malignant and Benign Breast Cancer Classification using Machine Learning Algorithms," 2021 International Conference on Artificial Intelligence (ICAI), 2021, pp. 97-101.