International Journal of Emerging Science and Engineering (IJESE)
ISSN: 2319–6378 (Online), Volume-11 Issue-2, January 2023

Machine Learning: A Literature Review for Breast Cancer

Vaishnavi Karma, Prateek Nahar

Abstract: ***Breast cancer is, after lung cancer, the most prevalent form of the disease in the globe. Women are the demographic most likely to be affected by this condition. Breast cancer is the most common kind of cancer to result in a woman's death if she is of childbearing age. Because there is always more to learn and there is room for improvement in every line of work, medical imaging is not an exception to this rule. It is expected that the death rate associated with cancer would decrease if it is discovered early and effectively treated. The diagnosis accuracy of persons working in the health care profession may be improved via the use of machine learning techniques. The technique known as deep learning has the potential to differentiate between breasts that are healthy and those that have cancer also known as neural networking. This method might be used to differentiate between healthy breast tissue and breast tissue affected by illness. Long-term research on the topic aimed, among other things, to examine breast cancer and screening practises among Indian women. This was one of the primary goals of the inquiry. A literature study was carried out with the assistance of several databases along with additional sources. Participants in the study were instructed to use phrases linked to breast cancer such as "breast carcinoma" and "breast cancer awareness," in addition to terms such as "knowledge" and "attitude," as well as the gender neutral term "women." In addition, India had a role in the study that was done. This search does not look for articles that have been published in the English language in the last 12 years.***

Keywords**: Breast Cancer, CNN, Mammograms-MINI-DDSM, Machine Learning**

I. INTRODUCTION

As the population ages, more individuals are diagnosed with cancer, which raises the cancer rate. In poorer nations, cancer-causing smoking is common. It's the most frequent cancer in women and overall. Breast cancer accounts for 23% of all female cancer diagnoses and 14% of female cancer deaths. Eight in 10 U.S. women will be diagnosed with breast cancer. In nations with poorer incomes and higher life expectancies, more women are diagnosed with breast cancer. [2] In oncology, a disease has aberrant cell development and fast spread. This includes cancer. Cell accumulation in the body, manifesting as a bump or lump, is a common indication of sickness. This ailment is usually caused by the creation of tumours on the body; the spot where they originally emerged may help identify them. Because breast cancer's early symptoms are rare, women must undergo frequent screenings to boost their chances of an early diagnosis. Many moderately intense lumps were identified, however they caused little harm. that its predecessors before it. Breast cancer has 16 to 21 subgroups. Even though most breast cancers are non-invasive, there are 21 forms. Men had 2470 new invasive cancer cases in 2017 [1]. As individuals age, their breast cancer death risk rises. Breast cancer patients were 62 on average between 2010 and 2014. WHO says Pakistan has Asia's highest breast cancer incidence rate. There are 90,000 new cases and 40,000 deaths annually. Now, the individual has a good chance of recovering from breast cancer. Because cancer's early stages are painless, it's not diagnosed until its later, more severe stages. Cancer patients in Pakistan are, on average, 40 years old when diagnosed. Patient endurance is the proportion of patients who continue to live after discovering they will recover soon. The tumor's stage affects how fast a patient loses stamina throughout treatment. [1,] Recent studies suggest breast cancer patients have tremendous stamina. Researchers formerly thought just 89% of Americans used social media. The proportion rose thereafter. After 15 years, 80% of the population remains. According to the American Cancer Society, having a close relative with breast cancer enhances the individual's risk. Both sexes are affected. Mammography is a low-dose x-ray used to examine breasts. Breast cancer may be identified in many ways. Methods include wavelet, curvelet, restricted scale, and multiscale images. Third and fourth steps: Teachers use neuro-flooding frames and fuzzy thinking to discourage students from taking harmful subjects. [3] Deep learning methods may be used to simulate an information architecture. This paradigm combines text, graphics, and sound. Various datasets and CNN architectures are needed to build abstract models. Clinical imaging employing sophisticated learning approaches may spot cancer cells. To build an extended convergence network, acquire and analyse a lot of data. Using the aforesaid approach, calibrate a well-prepared group. Deep learning is used in bioinformatics, Alzheimer's diagnosis, and subatomic imaging. Imaging molecules combines both [4]. This industry includes physical image readers and subatomic data collectors. Taking many courses at once isn't the best method to learn. Information collecting and pre-tests would be easier. Due to lack of medical knowledge, many impoverished individuals avoid the doctor. Breast cancer early detection and treatment are poorly understood.

Manuscript received on 27 December 2022 | Revised Manuscript received on 30 January 2023 | Manuscript Accepted on 15 January 2023 | Manuscript published on 30 January 2023.
*****Correspondence Author(s)
Vaishnavi Karma*, Research Scholar, Department of Computer Science and Engineering, IES IPS Academy, Indore (MP), India. Email: vsnvi4286karma@gmail.com, ORCID ID: https://orcid.org/0000-0002-8643-8278
Prateek Nahar, Assistant Professor, Department of Computer Science and Engineering, IES IPS Academy, Indore (MP), India. Email: prateeknahar@ipsacademy.org, ORCID ID: https://orcid.org/0000-0002-7839-7728

Random breast cancer screenings should end. Community-centered, integrated programmes should be implemented as soon as practicable.
10 Several methods exist for early cancer diagnosis. Mammography, breast self-examination, and clinical breast examination aren't the only ways to detect breast cancer.
The national health care system doesn't include breast cancer prevention or treatment, despite public awareness campaigns. One study found that women have positive attitudes about breast cancer despite being unaware of it. Breast cancer screening workers are still learning. Because breast cancer can be fatal, Indian women need information fast. This study examined Indian women's breast cancer screening knowledge, attitudes, and practices. [5].

II. RELATED WORK

Mammography developments in the 1990s made breast calcifications detectable. CNN has now become ubiquitous in clinical imaging. CNN's continuing success depends on the network's ability to adapt. This research found two kinds of motion learning in clinical imaging. Using highlights from a single organization's layer demonstrates the process in action. When the pre-made structure has been entirely exploited, a new strategic layer is developed. SVM, wavelet analysis, cosine modifications, and CNN lighting may erase this dataset's highlights. This dataset was tested using large-scale picture comparisons and extraction classifiers. Used with SIFT. Components were categorised as "good," "bad," "benevolent," "childlike," and "destructive" Mammography patches increased data collection. We might assess its effectiveness by comparing it to another dataset. 2D-DWT was utilised to separate advanced mammography into four groups, which were then used to generate CNNs using SVMs with softmax layers and DCT (Discrete curve altering). Before DCT was utilised to build CNNs using SVM and softmax layers, 3D-DWT was used to categorise advanced mammography into four groups. Based on IRMA knowledge base data, 81.83 percent of DCTs and 83.74 percent of CTs were right. Also, 81.83 percent of CTs were correct. [6].
High-pass and low-pass channels are employed for transmission. After producing a wavelet, it may be measured and moved. Changes in curvature help identify more difficult wedge sections and shorter lines. [7]. [8] Fuzzy reasoning may exhibit conjecture and logical reasoning [3–4]. Fuzzy logic can be employed even if there's no right method to present a number. Adjustments to a fluffy model require recreating it. "Neurofluffy thinking" includes starting a research with erroneous information and concepts. [3] Multi-scale curve's mass identification accuracy of 98.59% has never been exceeded. C-mean bunching is more effective than the division's impact area extraction and placement skills for legacy estimations. 3D ultrasound images were used to distinguish grassy and non-fatty breast tissue. To minimise breast cancer risk. K-bunching breast thermography may detect breast cancer in its early stages. [9] Undercover investigations of the illness region are needed. [10] Researchers found that ribosome-sized tumour zones were cleared. Tumor division options included configurable criteria and multiple categories [10]. For fact-checking, the same sources were used. In addition, the tumor's division has led to new regulations, flexible criteria, and other perspectives [10]. Researchers employed a new approach to find thick areas in recurrence models [11]. This research examined 32 categories of visual quality. In the next paragraphs, we'll explain this mammography enhancement and denoising approach. Microcalcification and slight colour changes, like masses, were also identified. We utilised weighted misfortune so our finder wouldn't be lucky. This ensured it wouldn't land in a good spot [12]. ESTD and Surface Examination were utilised for part of the audit. We used these to illustrate how mammograms are taken throughout the presentation. [13] Lead component examination and classification identified breast cancer risk factors. [14].

III. DATASET

A significant amount of training data is required for a CNN in order for it to provide satisfactory results. Due to the dearth of large datasets, training and testing were carried out using the largest dataset that could be accessed by the general public over the internet. The vast bulk of the data needed for this inquiry was gathered with the assistance of the MINI-DDSM. It took a grand total of 5358 pictures to accomplish what needed to be done here. To put it into perspective, the width of each picture is 1372 pixels and the height is 2340 pixels [15]. There were about 2474 images of cancerous growths and around 1940 shots of healthy people. Throughout the whole of the course, students were taught by and tested using CNN. Before being utilised, images that have been taken in grayscale need to be transformed into one of the available colour formats [16].

Table 1: Dataset for MINI-DDSM.

		Class	
		Benign	malignant
Images	Training Samples (80%)	1940	2474
	Test Samples (20%)	420	524

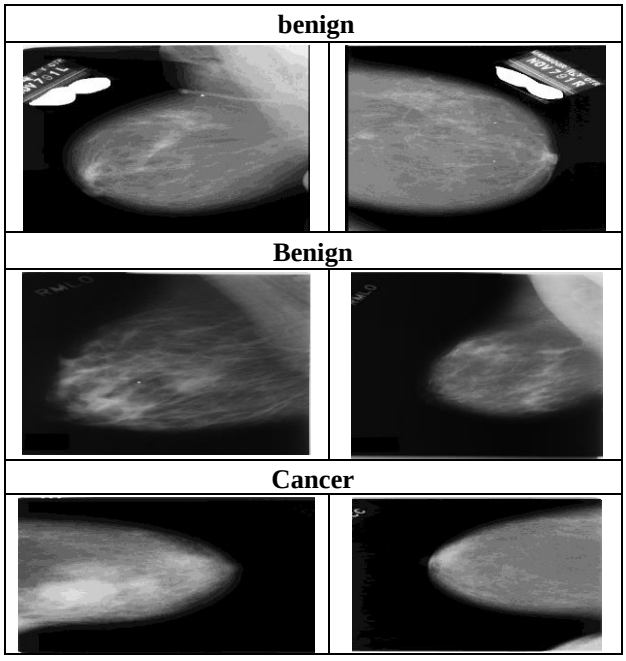

Figure 1 Show a Mammogram with the MINI-DDSM Dataset.

Retrieval Number: 100.1/ijese.B25430111223
DOI: 10.35940/ijese.B2543.0111223
Journal Website: www.ijese.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

IV. LIMITATIONS

This study has certain limitations, such as its reliance on previous studies' findings [17]. This may have skewed the results in favour of research that had already been published in peer-reviewed publications. We searched far and low for relevant articles and included as many as possible, although some ones were left out due to our inclusion criteria [18]. The review's results may not apply uniformly to all of India due to the country's enormous socioeconomic and cultural diversity, as well as its many geographical regions and environmental conditions [19]. Some figures may be off because rural residents may not have the same linguistic proficiency as urbanites [20]. It is possible that women were reluctant to disclose their breast cancer diagnosis to others or that they accidentally spread false information due to social or cultural pressures [21].

V. CONCLUSION

There is a substantial gap in knowledge and attitudes that must be addressed in practise, even though the systematic review indicated that more than half of the study group had sufficient information and a favourable attitude toward breast cancer screening and early detection treatments. Because of the prevalence of breast cancer in India, it is crucial that women there be made aware of the risk factors that might lead to the disease. Partnerships with community organisations and the medical community, as well as national and state-level initiatives to educate the public about cancer, are urgently needed. Governments in low-income countries like India should prioritise breast cancer early detection measures in hospitals and clinics as soon as possible to further reduce already low incidence and death rates. Early diagnosis and screening methods such as breast self-examination (BSE), cervical cancer screening (CBE), and mammography might increase the likelihood of detecting breast cancer at its earliest stages, when it is most treatable.

DECLARATION

Funding/ Grants/ Financial Support	No, I did not receive.
Conflicts of Interest/ Competing Interests	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material/ Data Access Statement	Not relevant.
Authors Contributions	All authors have equal participation in this article.

REFERENCES

1. "Breast cancer statistics." [Online]. Available: http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics
2. "Over 40,000 die of breast cancer every year in pakistan." [Online]. Available: http://www.google.com.pk/amp/s/www.dawn.com/news/amp/ 1319675
3. S. Bagchi and A. Huong, "Signal processing techniques and computer-aided detection systems for diagnosis of breast cancer–a review paper," Indian Journal of Science and Technology, vol. 10, no. 3, 2017. [CrossRef]
4. F. A. Jaffer and R. Weissleder, "Molecular imaging in the clinical arena," Jama, vol. 293, no. 7, 2005. [CrossRef]
5. S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, "Early diagnosis of alzheimer's disease with deep learning," IEEE, 2014 [CrossRef]
6. J. C. Tobias Charistian Cahoon, Melanie A. Sutton, "Three-class mammogram classification based on descriptive cnn features," 2000.
7. S. Sharma, M. Kharbanda, and G. Kaushal, "Brain tumor and breast cancer detection using medical images," International Journal of Engi-neering Technology Science and Research, vol. 2, 2015.
8. P. Gu, W.-M. Lee, M. A. Roubidoux, J. Yuan, X. Wang, and P. L. Carson, "Automated 3d ultrasound image segmentation to aid breast cancer image interpretation," Ultrasonics, vol. 65, 2016. [CrossRef]
9. P. Hankare, K. Shah, D. Nair, and D. Nair, "Breast cancer detection using thermography," Int. Res. J. Eng. Technol, vol. 4, 2016.
10. W. K. Moon, I.-L. Chen, J. M. Chang, S. U. Shin, C.-M. Lo, and R.-F. Chang, "The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound," Ultrasonics, vol. 76, 2017. [CrossRef]
11. F. Strand, K. Humphreys, A. Cheddad, S. Törnberg, E. Azavedo, J. Shepherd, P. Hall, and K. Czene, "Novel mammographic image features differentiate between interval and screen-detected breast cancer: a case-case study," Breast Cancer Research, vol. 18, no. 1, 2016. [CrossRef]
12. A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, "Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing," IEEE transactions on instrumentation and measurement, vol. 57, no. 7, 2008. [CrossRef]
13. S. Hwang and H.-E. Kim, "Self-transfer learning for fully weakly supervised object localization," arXiv preprint arXiv:1602.01625, 2016.
14. Z. Mohammadzadeh, R. Safdari, M. Ghazisaeidi, S. Davoodi, and Z. Azadmanjir, "Advances in optimal detection of cancer by image processing; experience with lung and breast cancers," Asian Pacific journal of cancer prevention: APJCP, vol. 16, no. 14, 2015. [CrossRef]
15. "Mammographic image analysis society digital mammogram database." [Online]. Available: http://peipa.essex.ac.uk/info/MINI-DDSM.html
16. M. M. Jadoon, Q. Zhang, I. U. Haq, S. Butt, and A. Jadoon, "Threeclass mammogram classification based on descriptive cnn features," BioMed research international, 2017. [CrossRef]
17. "Breast cancer facts & figures 2017-2018." [Online]. Available: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2017-2018.pdf
18. K. Vennila, K. Sivakami, and R. Padmapriya, "Detection of mass in digital mammograms," International Journal of Computer Applications, vol. 104, no. 5, 2014. [CrossRef]
19. K. Djaroudib, A. T. Ahmed, and A. Zidani, "Textural approach for mass malignantity segmentation in mammographic images," arXiv preprint arXiv:1412.1506, 2014.
20. S. Deepa and V. S. Bharathi, "Efficient roi segmentation of digital mammogram images using otsu's n thresholding method," Indian Journal of Automation and Artificial Intelligence, vol. 1, no. 2, pp. 51–56, 2013.
21. "Filters in the context of convolutional neural networks." [Online]. Available: https://www.quora.com/What-is-afilter-in-the-context-of-Convolutional-Neural-Networks.

AUTHORS PROFILE

Vaishnavi Karma has received B.E. (CSE) from SPITM, Mandleshwer (RGPV) in year 2019 and M.E. (CSE) from IES IPS Academy, Indore in year 2022. His areas of interest are Breast Cancer using Deep Learning , A Literature Review for Breast Cancer , Machine Learning. She has published 02 paper in various international and national conference, Research Scholar, Department of Computer Science Engineering, IES IPS Academy, Indore (MP), India. E-mail: vsnvi4286karma@gmail.com

Mr. Prateek Nahar has received B.E. (C.S.E.) from Rishiraj Institute of Technology, Indore in year 2008 and M.Tech. (C.S.) from School of Computer Science, D.A.V.V., Indore in year 2010. Ph.D. in Soft Computing from SCSIT, D.A.V.V., Indore (M.P.) in year 2020. He is having more than 12 years of teaching and research experience. He has guided more than 05 PG dissertation and Project. He has published 09 papers in various international and national conferences and journals. His areas of interests are Theory of Computation, Principle of Programming Language, Machine Learning, Data Science, Compiler and Image Processing. Assistant Professor, Department of Computer Science Engineering IES IPS Academy, Indore (MP), India. E-mail: prateeknahar@ipsacademy.org

Retrieval Number: 100.1/ijese.B25430111223
DOI: 10.35940/ijese.B2543.0111223
Journal Website: www.ijese.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)