AUTOMATIC BREAST CANCER PREDICTION USING NOVEL CONVOLUTIONAL NEURAL NETWORK COMPARED TO RELIEF ALGORITHM

S. Vishnu[1], S. S. Arumugam[2*]

Article History: **Received:** 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The objective of the work is to predict the Accuracy of Breast Cancer Prediction through Novel Convolutional Neural networks (NCNN) Comparative with the Relief Algorithm (RA). **Material And Methods:** The dataset used for Accuracy and Loss is from the GitHub library. The total sample size is 48. The two groups Novel Convolutional Neural Network (N=24), and Relief Algorithm (N=24) proposed by predicting the accuracy (97.10%) of Breast Cancer Prediction compared with Novel Convolutional Neural Network. **Results:** The Result proved that the Relief algorithm with Better accuracy than the Novel Convolutional Neural Network. The Novel Convolutional Neural Network appears significantly better than Relief Algorithm ($p<0.05$). **Discussion and Conclusion:** The Prediction of breast cancer is better in novel convolutional neural networks when compared with the relief algorithm (RA).

Keywords: Breast Cancer, Relief Algorithm, Mammography, Novel Convolutional Neural Network, Machine Learning Algorithm, Accuracy.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.
[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.

1. Introduction

The most frequent malignancy in women is breast cancer. For breast cancer screening, full-field digital mammography (FFDM) is frequently employed ((Clinical and Technical Evaluation of Full Field Digital Mammography 2000)). However, mammography has an inherent limitation when tissue overlaps, especially in dense breasts, which causes mammography to miss some suspicious cancerous lesions ((Teoh et al. 2021)). A brand-new, developing method for diagnosing and screening breast cancer that produces images that resemble three dimensions is called digital breast tomosynthesis (DBT). It can be used to give a precise evaluation of the breast's dense tissue. By enhancing the identification, characterization, and diagnosis of lesions, DBT has reduced overlapping tissue on screening. It elevates this method over digital mammography (DM).

The most common type of cancer among women is breast cancer. The impact of this cancer may be lessened with prevention. Therefore, the survival rate can be improved by an early cancer diagnosis. Particularly developing nations have a very low ability to diagnose and diagnose in late stages due to a lack of knowledge and inadequate medical facilities. if they are identified at an early stage using reasonably priced machine vision-based medical facilities. For screening, diagnosis, and tracking the course of tumors (Hande et al. 2021), many cutting-edge approaches are used. Convolutional mammography is the method of screening that is most frequently used. This led to the development of other testing procedures like ultrasonography and biopsy. Second, when mammography detects cancer, overdiagnosis and therapy may occur. Third, Fatty tissue appears darker on mammography, which is the most significant drawback of x-ray mammograms ((Holbrook 2020)) whereas fibroglandular tissue appears as white areas. The reason behind this is that fibro glandular tissue and tumors have a similar density which is why tumors are harder to detect in women with dense breasts ((Kuhl and Braun 2008)).

Our team has extensive knowledge and research experience that has translated into high quality publications (K. Mohan et al. 2022; Vivek et al. 2022; Sathish et al. 2022; Kotteeswaran et al. 2022; Yaashikaa, Keerthana Devi, and Senthil Kumar 2022; Yaashikaa, Senthil Kumar, and Karishma 2022; Saravanan et al. 2022; Jayabal et al. 2022; Krishnan et al. 2022; Jayakodi et al. 2022; H. Mohan et al. 2022)

In a previous study, the efficiency improvement of the Relief Algorithm (RF) in Data retrieval was not properly considered to increase accuracy. To overcome this issue a Novel Convolutional Neural Network is implemented to improve secured data retrieval with reduced time in machine learning.

2. Materials and Methods

The research is conducted in the Saveetha School of Engineering's Machine Learning Laboratory at the Saveetha Institute of Medical and Technical Sciences in Chennai. The sample size has been calculated using the GPower software by comparing both of the controllers in Supervised learning. For the purpose of comparing the procedure and the outcome, two numbers of groups are chosen. For this work, 24 sets of samples from each group, for a total of 48 samples, are chosen. The pre-test power value is calculated using GPower 3.1 software (g power setting parameters: =0.05, power=0.80, statistical test difference between two independent means, t-test value=252.538 Two algorithms (RF and Novel Convolutional Neural Network Algorithm) are implemented using Technical Analysis software. Since no human or animal samples were used in this study, there is no need for ethical approval.

Convolutional Neural Network

The paper suggests a convolutional neural network (CNN) strategy to improve the whole-slide imaging of aggressive ductal carcinoma tissue zones for breast cancer detection (WSIs). The study suggests a method for automatically detecting breast cancer using several convolutional neural networks (CNNs), and compares its findings to those from machine learning (ML).

Relief Algorithm

Graphic depiction of various connections between random variables in a specific collection. It functions as a classifier without relying on qualities. Unsupervised training on a set of instances enables a DBN to develop the ability to probabilistically recreate its inputs. After that, the layers serve as feature detectors. A DBN can be further taught under supervision to perform categorization after this learning phase.

DBNs can be thought of as a compilation of networks, such as autoencoders or restricted Boltzmann machines (RBMs), where the hidden layer of each sub-network acts as the layer for the layer below it. An RBM is a generative energy-based model having connections between but not within layers, a "visible" input layer, a

hidden layer, and both. Contrastive divergence is applied to each sub-network in turn, starting from the "lowest" pair of layers, in a quick, layer-by-layer unsupervised training method as a result of this composition (the lowest visible layer is a training set). The collaborative Python programming language is used to create the software application for assessing CNN and FS algorithms. The system configuration consists of an Intel i7 processor and 8GB of RAM. Windows 10 64-bit was the operating system utilized.

Statistical Analysis

For statistical implementation, the software used is IBM SPSS V26. Statistical package for social sciences is for calculating the statistical calculations such as mean, standard deviation also to plot the graphs, etc. The independent variables are UR Label and the dependent variable is 'accuracy'.

3. Results

Below Table shows the simulation result of the proposed algorithm Novel Convolutional Neural Algorithm and the existing system in RA was run at different times in the google collab with a sample size of 24. From the table, it was observed that the mean accuracy of the Machine learning Algorithms like the Novel Convolutional Neural algorithms was 91.81% and the RA was 76.48%. In order to compute the Mean, Standard Deviation, and Standard Error Mean, a T-test with an independent variable was performed among the research groups. A value of 1.320 and an effect size of 1.712 show a considerable departure from the RA for the Novel Convolutional Neural Algorithm. Table 2 represents the Mean of the Novel Convolutional Neural Algorithm which is better compared with the RA with a standard deviation of 0.71499 and 0.63395 respectively. From the algorithm and RA in terms of mean and accuracy. The mean results, the Novel Convolutional Neural algorithm (91.81%) gives better accuracy than the RA (76.48%). Figure 1 gives the comparison chart of the Novel Convolutional Neural Algorithm accuracy of the algorithm is better than RA. It is, therefore, conclusive that a novel convolutional neural network performs better than RA. The resultant plots are shown below in the figure. The figure has been placed at the end of the paper.

4. Discussion

To increase the precision of cancer prediction, novel convolutional neural networks and RA are applied and compared for Brust Prediction. Based on the results, it can be said that the Novel Convolutional Neural algorithm performs more accurately than the RA. In the recent survey, the proposed NCNA is a promising option for predicting the future values of stocks with a root mean square value of 0.04 ((Rueda 2015)). proposed a semi-based model for different companies belonging to the banking sector based on historical data and observed that the error level comes down drastically with the data for longer periods ((Kim, Do Yeun, and Park 2020)). implemented six machine learning techniques ((Zhuang et al. 2021)) i.e., Prophet, LR, SVM, Decision Tree, and Naive Bayes, and by comparing them concluded that NCNA works better with an accuracy of 80% ((Lin et al. 2021)). Major research contribution supports Implementation and comparative analysis of WTA to optimize prediction gain of drive with reduced efficiency improvement ((McCann 2000)). Even though few articles listed the disadvantages of the proposed RA. Further, the RA is not suitable for improving the accuracy of Breast Cancer prediction ((Yan et al. 2021)). From the above discussion, ((Gupta et al. 2021)) only a few articles ensure that they provide better performance than the proposed Novel Convolutional Neural Network and watershed transform algorithm for improving the accuracy of breast cancer prediction ((Phan et al. 2021)). Also, the present cancer prediction demands no additional expense and has consequently attracted a lot of interest recently (Wu, Feng, and Li 2021; Olsen et al. 2021). So, we can infer that the proposed NCNA and RA can be used to improve the accuracy of cancer disease prediction ((Wu, Feng, and Li 2021)).

5. Conclusion

When compared to the accuracy of the watershed transform algorithm, which is 76.48%, the work uses a Convolutional Neural Network (CNN) algorithm to locate the breast cancer prediction to retrieval with reduced time to be demonstrated with greater accuracy of 91.80%.

Declaration

Conflicts of interest

No conflicts of interest.

Author's Contributions

Author SV was involved in data collection, data analysis, and manuscript writing along with him Author SSA was in conceptualization, data validation, and critical review manuscript.

Acknowledgments
The authors would like to express their gratitude to Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly Known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding
We thank the following organization for providing financial support that enabled us to complete the study.

1. Cyclotron Technologies, Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Science.
4. Saveetha School of Engineering.

6. Reference

Clinical and Technical Evaluation of Full Field Digital Mammography. 2000.

Gupta, Meenu, Rachna Jain, Arun Solanki, and Fadi Al-Turjman. 2021. Cancer Prediction for Industrial IoT 4.0: A Machine Learning Perspective. CRC Press.

Hande, Pradipta C., Sabita S. Desai, Sarabjeet K. Arneja, and Sreedevi Sathian. 2021. "Utility of Digital Breast Tomosynthesis with Two-Dimensional Synthesized Mammography Images: A Pictorial Essay." The Indian Journal of Radiology & Imaging 31 (3): 678–88.

Holbrook, Anna I. 2020. "Breast Pain, A Common Grievance: Guidance to Radiologists." AJR. American Journal of Roentgenology 214 (2): 259–64.

Jayabal, Ravikumar, Sekar Subramani, Damodharan Dillikannan, Yuvarajan Devarajan, Lakshmanan Thangavelu, Mukilarasan Nedunchezhiyan, Gopal Kaliyaperumal, and Melvin Victor De Poures. 2022. "Multi-Objective Optimization of Performance and Emission Characteristics of a CRDI Diesel Engine Fueled with Sapota Methyl Ester/diesel Blends." Energy. https://doi.org/10.1016/j.energy.2022.123709.

Jayakodi, Santhoshkumar, Rajeshkumar Shanmugam, Bader O. Almutairi, Mikhlid H. Almutairi, Shahid Mahboob, M. R. Kavipriya, Ramesh Gandusekar, Marcello Nicoletti, and Marimuthu Govindarajan. 2022. "Azadirachta Indica-Wrapped Copper Oxide Nanoparticles as a Novel Functional Material in Cardiomyocyte Cells: An Ecotoxicity Assessment on the Embryonic Development of Danio Rerio." Environmental Research 212 (Pt A): 113153.

Kim, Do Yeun, Kim Do Yeun, and Hannah Lui Park. 2020. "Breast Cancer Risk Prediction in Korean Women: Review and Perspectives on Personalized Breast Cancer Screening." Journal of Breast Cancer. https://doi.org/10.4048/jbc.2020.23.e40.

Kotteeswaran, C., Indrajit Patra, Regonda Nagaraju, D. Sungeetha, Bapayya Naidu Kommula, Yousef Methkal Abd Algani, S. Murugavalli, and B. Kiran Bala. 2022. "Autonomous Detection of Malevolent Nodes Using Secure Heterogeneous Cluster Protocol." Computers and Electrical Engineering. https://doi.org/10.1016/j.compeleceng.2022.107902.

Krishnan, Anbarasu, Duraisami Dhamodharan, Thanigaivel Sundaram, Vickram Sundaram, and Hun-Soo Byun. 2022. "Computational Discovery of Novel Human LMTK3 Inhibitors by High Throughput Virtual Screening Using NCI Database." Korean Journal of Chemical Engineering. https://doi.org/10.1007/s11814-022-1120-5.

Kuhl, C. K., and M. Braun. 2008. "[Magnetic resonance imaging in preoperative staging for breast cancer: pros and contras]." Der Radiologe 48 (4): 358–66.

Lin, Fei, Hang Sun, Lu Han, Jing Li, Nan Bao, Hong Li, Jing Chen, Shi Zhou, and Tao Yu. 2021. "An Effective Fine Grading Method of BI-RADS Classification in Mammography." International Journal of Computer Assisted Radiology and Surgery, December. https://doi.org/10.1007/s11548-021-02541-8.

McCann, Jenny. 2000. "Mammographic Prediction of Early Breast Cancer Survival." Breast Cancer Research. https://doi.org/10.1186/bcr-2000-66658.

Mohan, Harshavardhan, Sethumathavan Vadivel, Se-Won Lee, Jeong-Muk Lim, Nanh Lovanh, Yool-Jin Park, Taeho Shin, Kamala-Kannan Seralathan, and Byung-Taek Oh. 2022. "Improved Visible-Light-Driven Photocatalytic Removal of Bisphenol A Using V2O5/WO3 Decorated over Zeolite: Degradation Mechanism and Toxicity." Environmental Research. https://doi.org/10.1016/j.envres.2022.113136.

Mohan, Kannan, Abirami Ramu Ganesan, P. N. Ezhilarasi, Kiran Kumar Kondamareddy, Durairaj Karthick Rajan, Palanivel Sathishkumar, Jayakumar Rajarajeswaran, and Lorenza Conterno. 2022. "Green and Eco-

Friendly Approaches for the Extraction of Chitin and Chitosan: A Review." Carbohydrate Polymers 287 (July): 119349.

Olsen, Maria, Krista Fischer, Patrick M. Bossuyt, and Els Goetghebeur. 2021. "Evaluating the Prognostic Performance of a Polygenic Risk Score for Breast Cancer Risk Stratification." BMC Cancer 21 (1): 1351.

Phan, Nam Nhut, Chi-Cheng Huang, Ling-Ming Tseng, and Eric Y. Chuang. 2021. "Predicting Breast Cancer Gene Expression Signature by Applying Deep Convolutional Neural Networks From Unannotated Pathological Images." Frontiers in Oncology 11 (December): 769447.

Rueda, Hugo Gómez. 2015. Comparison of Diverse Genomic Data for Outcome Prediction in Cancer: A Study of TCGA Data and Breast Cancer Patients from Northeast Mexico.

Saravanan, A., P. Senthil Kumar, B. Ramesh, and S. Srinivasan. 2022. "Removal of Toxic Heavy Metals Using Genetically Engineered Microbes: Molecular Tools, Risk Assessment and Management Strategies." Chemosphere 298 (July): 134341.

Sathish, T., R. Saravanan, V. Vijayan, and S. Dinesh Kumar. 2022. "Investigations on Influences of MWCNT Composite Membranes in Oil Refineries Waste Water Treatment with Taguchi Route." Chemosphere 298 (July): 134265.

Teoh, Kar Choon, Hanani Abdul Manan, Norhashimah Mohd Norsuddin, and Iqbal Hussain Rizuana. 2021. "Comparison of Mean Glandular Dose between Full-Field Digital Mammography and Digital Breast Tomosynthesis." Healthcare (Basel, Switzerland) 9 (12). https://doi.org/10.3390/healthcare9121758.

Vivek, J., T. Maridurai, K. Anton Savio Lewise, R. Pandiyarajan, and K. Chandrasekaran. 2022. "Recast Layer Thickness and Residual Stress Analysis for EDD AA8011/h-BN/B4C Composites Using Cryogenically Treated SiC and CFRP Powder-Added Kerosene." Arabian Journal for Science and Engineering. https://doi.org/10.1007/s13369-022-06636-5.

Wu, Yunyun, Yuqing Feng, and Xi Li. 2021. "Classification of Breast Cancer by a Gold Nanoparticle Based Multicolor Fluorescent Aptasensor." Journal of Colloid and Interface Science 611 (December): 287–93.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Algal Biofuels: Technological Perspective on Cultivation, Fuel Extraction and Engineering Genetic Pathway for Enhancing Productivity." Fuel. https://doi.org/10.1016/j.fuel.2022.123814.

Yaashikaa, P. R., P. Senthil Kumar, and S. Karishma. 2022. "Review on Biopolymers and Composites – Evolving Material as Adsorbents in Removal of Environmental Pollutants." Environmental Research. https://doi.org/10.1016/j.envres.2022.113114.

Yan X. X., Li Y. J., Cao M. D., Wang H., Liu C. C., Wang X., Ran J. C., et al. 2021. "[DALYs for breast cancer in China, 2000-2050: trend analysis and prediction based on GBD 2019]." Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi 42 (12): 2156–63.

Zhuang, Hongkai, Shujie Wang, Bo Chen, Zedan Zhang, Zuyi Ma, Zhenchong Li, Chunsheng Liu, et al. 2021. "Prognostic Stratification Based on HIF-1 Signaling for Evaluating Hypoxic Status and Immune Infiltration in Pancreatic Ductal Adenocarcinomas." Frontiers in Immunology 12 (December): 790661.

Tables and Figures

Table 1. Predicted Accuracy of Breast Cancer Prediction for 48 different sample sizes.

GROUP	Algorithms	Average accuracy
1	Convolutional Neural Network	91.81%
2	Relief Algorithm	76.48%

Table 2. Statistical analysis of Convolutional Neural Network and Relief Algorithm. Mean accuracy, Standard deviation, and standard error values are obtained for 48 sample datasets.

	Algorithms	N	Mean	Std.Deviation	Std.Error Mean

Accuracy	Convolutional Neural Network	24	91.4362	025510	.05207
	Relief Algorithm	24	76.2929	.14568	.02974

Table 3. Independent sample T-test with a confidence interval at 95% and level of significance as 0.05. It shows the statistical significance of P<0.05 2-tailed.

	Levene's Test for Equality of Variances		**T-test for Equality of Means**						
	F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	**Mean Differences**	**Std.Error Differences**	**95% Confidence Interval of the Differences**	
								Lower	**Upper**
Equal Variances assumed	9.79	.003	252.538	46	.000	15.14333	.05996	15.02263	15.26404
Equal Variances not assumed			252.538	36.559	.000	15.14333	.05996	15.02178	15.26488

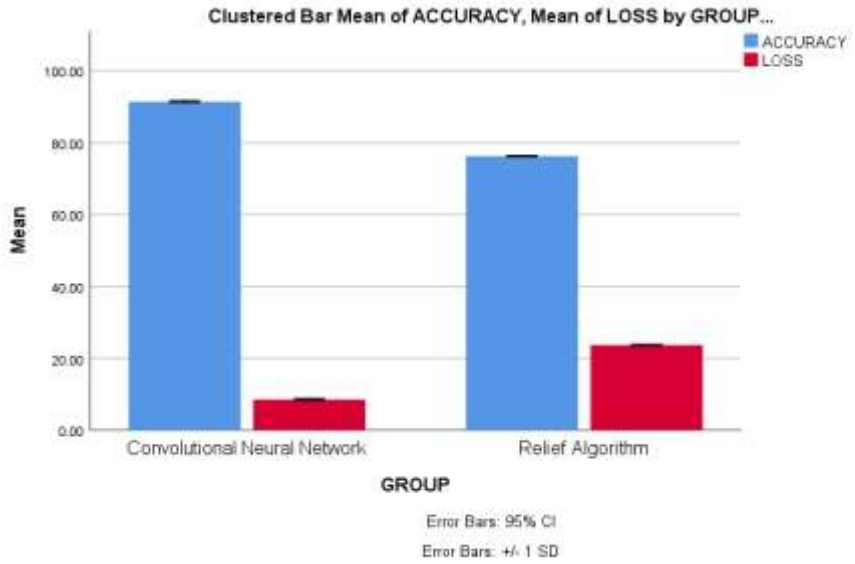

Fig. 1. In terms of means and accuracy, Novel Convolutional Neural Network and Relief Algorithm are compared. The Novel Convolutional Neural Network outperforms the Relief Algorithm in terms of mean accuracy. Y-axis: Mean accuracy; X-axis: Novel Convolutional Neural Network vs Relief Algorithm ±1SD error bar.