

Design and implementation of Breast Cancer Detection Using Hybrid Machine Learning

Nidhi Mangoriya (Student), Vinod Patel (Ass. Professor)
Dept. of CSE, LNCTS Bhopal

Abstract

The early detection of breast cancer is critical, as therapeutic actions are likely to be successful in the early stage of the disease. The different best radiological techniques are currently available for early detection of breast cancer. However, some of the breast lesions are missed during screening by radiologist, as they need to interpret large number of screening programs. Medical image processing techniques have been developed to help radiologists to improve detection process. The need for the thesis is to exploit recent developments of image processing techniques to improve breast cancer detection and reduce diagnosis errors. To proposed a new algorithm using hybrid machine learning for Breast Cancer Detection.

Keywords: Machine learning , breast cancer , Medical image processing.

I. INTRODUCTION

They used dissimilar algorithms to detect and diagnose breast cancer; accomplished high values in accurateness. Furthermost of the proposed algorithms individual were concentrating in just detection and diagnosis of breast cancer although a suitable treatment for respectively case of breast cancer didn't occupied into their accounts. To perceive and diagnose breast cancer, physicians or radiologists have been used dissimilar images that are distributed by expending a unusual device called modality accountable to screen somewhat organ of human body like brain and breast. In Digital Tomosynthesis, images of multiple angles are acquired at a stationary trampled breast with a short duration of scan. It is a 3D imaging technique. The main advantage of this screening method is that the abnormal portion is fetched when other regions are blurred and also it provides improved accuracy, reduced breast compression, localization of 3D lesion and gives 3D enhanced imaging. The drawback of this modality is longer duration of image acquisition, high radiation dose and takes more processing time. These modalities are Digital respectively device issues only one type of images. Practically a big number of the studied

papers just contract with one kind of image. Few papers transaction with two type of images but not completely type images that are used in breast cancer. Dissimilar methods already exist to detect breast cancer like soft computing and simulation techniques but the greatest precise results were

increased by machine learning techniques. According to the proposed works at the composed research ; numerous approaches are used like logistic regression, k-Nearest-Neighbor (KNN), Support vector machine (SVM), attention on how to decrease the computational expense in the diagnosis of breast cancer, consequently; a innovative ensembling process is proposed, namely (Breast Cancer).

II. RELATED WORK

the main purpose of this review is to highlight[1] all the previous studies of machine learning algorithms that are being used for breast cancer prediction and this article provides the all necessary information to the beginners who want to analyze the machine learning algorithms to gain the base of deep learning.

[2] used machine learning to train a model by using the predicted features of the nuclei of cells. As well as, A comparative study of two different algorithms KNN and SVM are conducted where the accuracy of each classifier is measured. KNN was achieved a better accuracy 97.48 compare to 97.27 with SVM. Besides, image processing was used to analyze a digital image of a fine needle aspirate (FNA) of breast tissue to discover the features of nuclei of the cells, and then, the feature values were applied to the trained model to find whether the tumor developed is benign or malignant.

[3] General performance was estimated by using trained models through kfold cross-validation and confounder-based cross-validations methods. In a colorectal cancer cohort, a mean AUC of 0.92 with a mean sensitivity at 85% specificity was achieved. In tumor stage and increasing tumor fraction, Sensitivity generally was raised. The validation of the machine learning methods used is underway

[4], suggested a way in Artificial Neural Network (ANN) to enhance the accuracy of estimation of breast cancer cases to be malignant or benign by selecting the number of hidden layers, the number of neurons in the hidden layer and the type of activation functions in hidden layers. Wisconsin Breast Cancer Database (WBCD) was used as a sample to train the network. Three transfer functions: LOGSIG, TANSIG, and PURELINE were applied with mean square error (MSE). The performance for the different numbers of neurons in the hidden; layer was conducted. The best network design was achieved using 21 neurons in the hidden layer, and TANSIG as the activation function.

[5] develop a scheme to detect breast cancer using deep belief network unsupervised path followed by back-propagation supervised path. The suggested network uses a back-propagation neural network with Liebenberg Marquardt learning function, while, the weights of the network was initialized by the deep belief network path (DBN-NN). The Wisconsin Breast Cancer Dataset (WBCD) used as input to the network. The proposed work gives an accuracy of 99.68% indicating promising results over previously published studies.

III. Proposed methodology

Breast cancer is the significant cancer in the worldwide amongst women in order that every eight women have been suffered due this disease throughout their lifetime. Mammography is the one of the most efficient imaging modality to detect breast cancer earlier. Various computer- aided detection algorithms had been widened to assist radiologists for precise diagnosis. Literature survey is focused on:

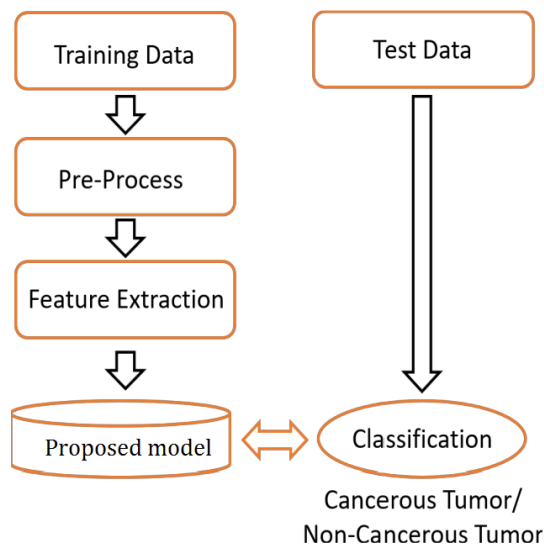


Figure 1: proposed methodology

- Microcalcification detection,

- Mass classification,
- Detection based on feature extraction and feature selection,
- Computer aided mammography,
- SVM mammograms and
- Other classifying algorithms in digital mammography.

Hybrid classifier for Breast Cancer Classification

In the proposed research work the extracted features such as area, perimeter and diameters are used for the classification of input breast images which belong to one of two classes such as normal and abnormal. The classification is performed by the supervised machine learning algorithm called Support Vector Machine (SVM). The input images are viewed as i dimensional vector and it is created as $i - 1$ dimensional hyperplane. The better training data is achieved by choosing the maximum hyperplane distance from the nearest data point.

Proposed algorithm

Step 1: Load training and testing breast image dataset.

Step 2: Applied hybrid (SVM and SVD) Algorithm

Step 3: Dependent on the class labels(SVD) (Normal and abnormal) classify the extracted features.

Step 3: Calculate candidate support value.

Step 4: if instances = null, the support value = Similarity among each instance in the attribute. Compute the total error.

Step 5: if instances < 0, calculate Decision value = support value & total error.

Step 6: Repeat until empty.

Results analysis

It will not be a problematic to control the best operating system that can be recycled even if it is Windows or Linux platform since the recommended submission is a website and it be contingent on the selected programming language, in short, it can work on any platform. Conversely, the suitable platform will be determined. To perform the simulation using anaconda software on Windows or Linux platform according to the selected programming language like Python . However, it conceivable to use further one programming language like by means of to analysis the outputs of some of procedures and draws some of figures using machine

learning and deep learning library for each case.

Table 1: Comparative Analysis different Algorithm

| Model / Parameter | Adaboost | Singular Value Decomposition(SVD) | SVM | Hybrid Algorithm |
|-------------------|----------|-----------------------------------|-------|------------------|
| Precision | 85.67 | 88.91 | 92.81 | 97.90 |
| Recall | 83.32 | 88.81 | 92.89 | 94.88 |
| F- Measure | 84.23 | 89.32 | 92.56 | 94.44 |
| Accuracy | 85.32 | 89.34 | 92.43 | 97.00 |

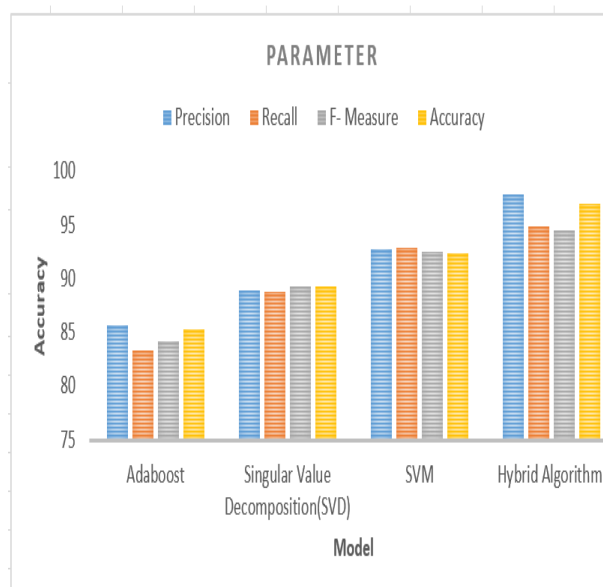


Figure 2: Comparative Analysis different Algorithm

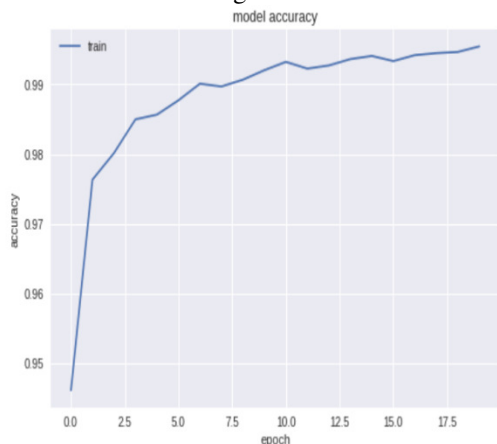


Figure 3: Model Accuracy

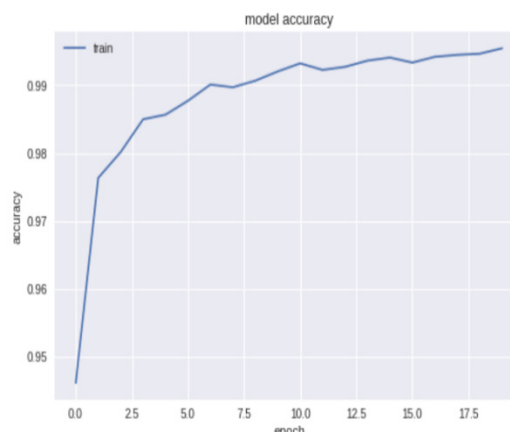


Figure 4: Model Accuracy

IV. Conclusion and Future work

In early detection of breast cancer, mammograms are the most cost effective, low cost and are reliable. The MIAS database provides the preliminary experiments that are carried out. Mammography can otherwise be termed as the low dose x-ray procedure for visualization of internal structure of the breast. Feature extraction of mammograms is done through Symlet wavelets. For feature reduction Singular Value Decomposition is used. Classification of the features that are extracted takes place through Naïve Bayes, SVM and hybrid algorithms. hybrid algorithms is considered as efficient algorithm as a set of weak classifiers are selected and are combined into a final strong classifier. Results show that the classification accuracy of hybrid algorithms classifier performs better than naïve bayes

Reference

- [1]. S. Mojrian et al., "Hybrid Machine Learning Model of Extreme Learning Machine Radial basis function for Breast Cancer Detection and Diagnosis; a Multilayer Fuzzy Expert System," 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), 2020, pp. 1-7, doi: 10.1109/RIVF48685.2020.9140744.
- [2]. K. Polat and U. Sentürk, "A Novel ML Approach to Prediction of Breast Cancer: Combining of mad normalization, KMC based feature weighting and AdaBoostM1 classifier," 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2018, pp. 1-4, doi: 10.1109/ISMSIT.2018.8567245.

- [3]. M. A. S. A. Husaini, M. H. Habaebi, S. A. Hameed, M. R. Islam and T. S. Gunawan, "A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks," in IEEE Access, vol. 8, pp. 208922-208937, 2020, doi: 10.1109/ACCESS.2020.3038817.
- [4]. F. A. Tania and P. C. Shill, "A Modified Support Vector Machine with Hybrid Kernel Function for Diagnosis of Diseases," 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), 2019, pp. 42-46, doi: 10.1109/BECITHCON48839.2019.9063176.
- [5]. Y. Mate and N. Somai, "Hybrid Feature Selection and Bayesian Optimization with Machine Learning for Breast Cancer Prediction," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 612-619, doi: 10.1109/ICACCS51430.2021.9441914.
- [6]. F. A. Tania and P. Chandra Shill, "An Improved Support Vector Machine for Disease Diagnosis," 2019 22nd International Conference on Computer and Information Technology (ICCIT), 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038369.
- [7]. F. A. Tania and P. Chandra Shill, "An Improved Support Vector Machine for Disease Diagnosis," 2019 22nd International Conference on Computer and Information Technology (ICCIT), 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038369.
- [8]. P. Sonar, U. Bhosle and C. Choudhury, "Mammography classification using modified hybrid SVM-KNN," 2017 International Conference on Signal Processing and Communication (ICSPC), 2017, pp. 305-311, doi: 10.1109/CSPC.2017.8305858.
- [9]. V. V. de Melo, "Breast cancer detection with logistic regression improved by features constructed by Kaizen programming in a hybrid approach," 2016 IEEE Congress on Evolutionary Computation (CEC), 2016, pp. 16-23, doi: 10.1109/CEC.2016.7743773.
- [10]. M. S. Munir Prince, A. Hasan and F. M. Shah, "An Efficient Ensemble Method for Cancer Detection," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934817.
- [11]. F. A. Tania and P. C. Shill, "A Hybrid Kernel Support Vector Machine with Feature Selection for the Diagnosis of Diseases," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-6, doi: 10.1109/EICT48899.2019.9068804.
- [12]. Rim, S.H., Allaire, B.T., Ekwueme, D.U. *et al.* Cost-effectiveness of breast cancer screening in the National Breast and Cervical Cancer Early Detection Program. *Cancer Causes Control* **30**, 819–826 (2019). <https://doi.org/10.1007/s10552-019-01178-y>
- [13]. Hoppe, E.J., Hussain, L.R., Grannan, K.J. *et al.* Racial disparities in breast cancer persist despite early detection: analysis of treatment of stage 1 breast cancer and effect of insurance status on disparities. *Breast Cancer Res Treat* **173**, 597–602 (2019).
- [14]. Yeom, Y.K., Chae, E.Y., Kim, H.H. *et al.* Screening mammography for second breast cancers in women with history of early-stage breast cancer: factors and causes associated with non-detection. *BMC Med Imaging* **19**, 2 (2019). <https://doi.org/10.1186/s12880-018-0303-3>
- [15]. Alvarado, M.D., Mittendorf, E.A., Teshome, M. *et al.* Correction to: SentimagIC: A Non-inferiority Trial Comparing Superparamagnetic Iron Oxide Versus Technetium-99m and Blue Dye in the Detection of Axillary Sentinel Nodes in Patients with Early-Stage Breast Cancer. *Ann Surg Oncol* **27**, 979 (2020). <https://doi.org/10.1245/s10434-020-08688-z>
- [16]. K. Barhanpurkar, A. S. Rajawat, P. Bedi and O. Mohammed, "Detection of Sleep Apnea & Cancer Mutual Symptoms Using Deep Learning Techniques," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 821-828, doi: 10.1109/I-SMAC49090.2020.9243488.
- [17]. Page, K., Hava, N., Slade, M. *et al.* Detection of gene amplification in matched tumour and plasma DNA from breast cancer patients by quantitative

- PCR. Breast Cancer Res 10, P70 (2008).
<https://doi.org/10.1186/bcr1954>
- [18]. Allaire, B.T., Ekweme, D., Hoerger, T.J. et al. Cost-effectiveness of patient navigation for breast cancer screening in the National Breast and Cervical Cancer Early Detection Program. Cancer Causes Control 30, 923–929 (2019). <https://doi.org/10.1007/s10552-019-01200-3>
- [19]. Krithiga, R., Geetha, P. Deep learning based breast cancer detection and classification using fuzzy merging techniques. Machine Vision and Applications 31, 63 (2020). <https://doi.org/10.1007/s00138-020-01122-0>
- [20]. Qu, J., Zhao, X., Chen, P. et al. Deep learning on digital mammography for expert-level diagnosis accuracy in breast cancer detection. Multimedia Systems (2021). <https://doi.org/10.1007/s00530-021-00823-4>
- [21]. Debelee, T.G., Schwenker, F., Ibenthal, A. et al. Survey of deep learning in breast cancer image analysis. Evolving Systems 11, 143–163 (2020). <https://doi.org/10.1007/s12530-019-09297-2>
- [22]. A. S. Rajawat, O. Mohammed and P. Bedi, "FDLM: Fusion Deep Learning Model for Classifying Obstructive Sleep Apnea and Type 2 Diabetes," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 835-839, doi: 10.1109/I-SMAC49090.2020.9243553.
- [23]. Wang, Y., Yang, F., Zhang, J. et al. Application of artificial intelligence based on deep learning in breast cancer screening and imaging diagnosis. Neural Comput & Applic (2021). <https://doi.org/10.1007/s00521-021-05728-x>
- [24]. Durda, K., Jaworska, K., Jakubowska, A. et al. Rapid test for detection of high risk of breast cancer. Hered Cancer Clin Pract 10, A7 (2012). <https://doi.org/10.1186/1897-4287-10-S1-A7>