iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** V **Month of publication:** May 2019

DOI: https://doi.org/10.22214/ijraset.2019.5443

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Breast Cancer Detection using Deep Learning

Dr. S.Y. Amdani[1], Prof. S.T. Khandare[2], G. G. Kale[3]
[1,2,3]Department of Computer Science and Engineering. B.N.C.O.E Pusad, S.G.B.U. Amaravati.

Abstract: Accurate diagnosis of breast cancer in histopathology images is challenging due to the heterogeneity of cancer cell growth as well as of a variety of benign breast tissue proliferative lesions. In this work, we propose a practical and self interpretable invasive cancer diagnosis solution. With minimum annotation information, the proposed method mines contrast patterns between normal and malignant images in unsupervised manner and generates a probability map of abnormalities to verify its reasoning. Particularly, a fully convolutional autoencoder is used to learn the dominant structural patterns among normal image patches. Patches that do not share the characteristics of this normal population are detected and analyzed by one-class support vector machine and 1-layer neural network. We apply the proposed method to a public breast cancer image set. Our results, in consultation with a senior pathologist, demonstrate that the proposed method outperforms existing methods. The obtained probability map could benefit the pathology practice by providing visualized verification data and potentially leads to a better understanding of data-driven diagnosis solutions.
Index Terms: Breast cancer diagnosis, abnormality detection, convolutional auto encoder, discriminative pattern learning, histopathology image analysis

I. INTRODUCTION

Artificial Neural Networks (ANN) is one of the best Artificial Intelligence techniques for common Data Mining tasks, such as classification and regression problems (Supervised Learning). A lot of research already showed that ANN delivered good accuracy in breast cancer diagnosis. In this method, first it involves in building the network for the model, parameters to be tuned in the beginning of the training process such as number of input nodes, hidden nodes, output nodes and initial weights, learning rates, and activation function. We know that it takes long time for training process due to complex architecture and parameter update process in each iteration and that has more computation cost. So, to overcome this I am using the Gradient based Back Propagation techniques and to speed up the computation I am leveraging the Deep Learning platform (Cloud GPU) called FloydHub. In this project, I implement the Artificial Neural Network from scratch for efficiently detecting the Breast Cancer. Here to calculate the performance of the model I am using various performance metrics like Accuracy, Specificity, Sensitivity, Recall, Precision and visualizing it by using Confusion matrix.

II. RELATED WORK

Nastac, P. Jalava, M. Collan, Y. Collan, T. Kuopio and B. Back et-al proposed a Gradient based Back Propagation Artificial Neural Networks (GBP-ANN). Along with ANN, they also using Stochastic Gradient Descent Classifier (SGD) and SVM with Linear and Gaussian Kernel and they compared the performances and trade-off's between models. The development of this technique is promising as intelligent component in medical decision support systems.

Abien Fred M. Agarap et-al presents a comparison of six machine learning (ML) algorithms: GRU-SVM Linear Regression, Multilayer Perceptron (MLP), Nearest Neighbor (NN) search, Softmax Regression, and Support Vector Machine (SVM) on the Wisconsin Diagnostic Breast Cancer (WDBC) dataset by measuring their classification test accuracy, and their sensitivity and specificity values. The said dataset consists of features which were computed from digitized images of FNA tests on a breast mass. For the implementation of the ML algorithms, the dataset was partitioned in the following fashion: 70% for training phase, and 30% for the testing phase.The hyper-parameters used for all the classifiers were manually assigned. Results show that all the presented ML algorithms performed well (all exceeded 90% test accuracy) on the classification task. The MLP algorithm stands out among the implemented algorithms with a test accuracy of ≈99.

Araújo et-al said Breast cancer is one of the main causes of cancer death worldwide. The diagnosis of biopsy tissue with hematoxylin and eosin stained images is non-trivial and specialists often disagree on the final diagnosis. Computer-aided Diagnosis systems contribute to reduce the cost and increase the efficiency of this process. Conventional classification approaches rely on feature extraction methods designed for a specific problem based on field-knowledge. To overcome the many difficulties of the feature-based approaches, deep learning methods are becoming important alternatives. A method for the classification of hematoxylin and eosin stained breast biopsy images using Convolutional Neural Networks (CNNs) is proposed. Images are classified in four classes, normal tissue, benign lesion, in situ carcinoma and invasive carcinoma, and in two classes, carcinoma and non-carcinoma. The architecture of the network is designed to retrieve information at different scales, including both nuclei and overall tissue organization. This design allows the extension of the proposed system to whole-slide histology images. The features extracted by the CNN are also used for training a Support Vector Machine classifier. Accuracies of 77.8% for four class and 83.3% for carcinoma/non-carcinoma are achieved. The sensitivity of our method for cancer cases is 95.6%.

III. METHODOLOGY

Artificial Neural Networks (ANN) is a part of Artificial Intelligence that uses various optimization techniques to train and learn from the data and predict for the new unseen dataset.

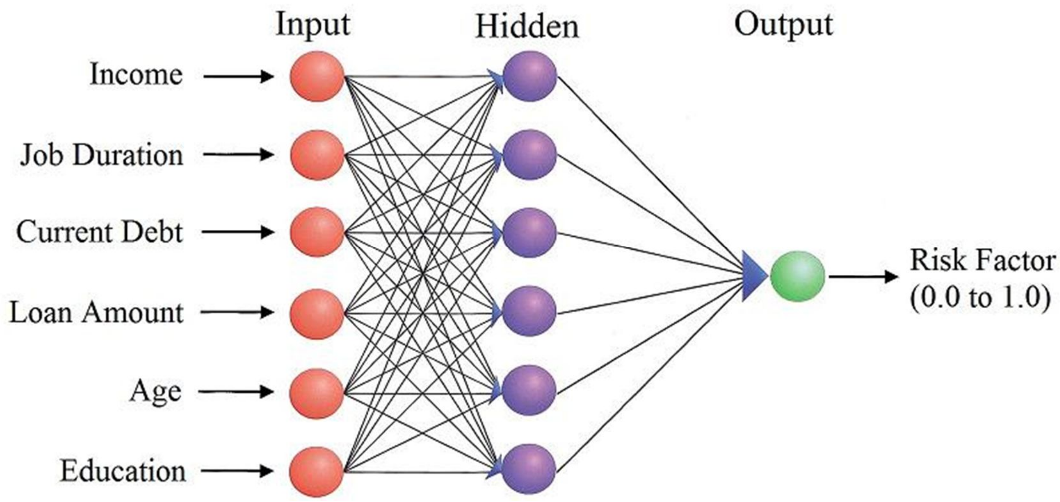


Fig. 3.1 : Architecture of proposed system

Steps:
1) Network Architecture
2) Forward Propagation
3) Backward Propagation
4) Updating Network Weights
5) Testing, Performance calculation

Network Architecture involves defining parameters for Network (Input, Output, Hidden, learning rate, number of epochs). Multi-level thresholding for image segmentation Forward Propagation is first step in training the network, that multiplies the input node with weights and it is given the activation function to generate the value between [0, 1]. The values will be input for next layer and then it gets multiplied with the next layer weights. All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

IV.RESULTS

Algorithm	Time (s)	Accuracy (%)	Mean Cross-validation Score (%)
Linear Regression	0.074	89.69	87.21
Decision-Tree	0.049	95.98	65.85
k-NN	0.187	92.21	90.03
Support Vector Machine	0.177	94.47	91.23
Artificial Neural Network	**146**	**99.24**	**97.07**

Fig. Comparison of classifier performances

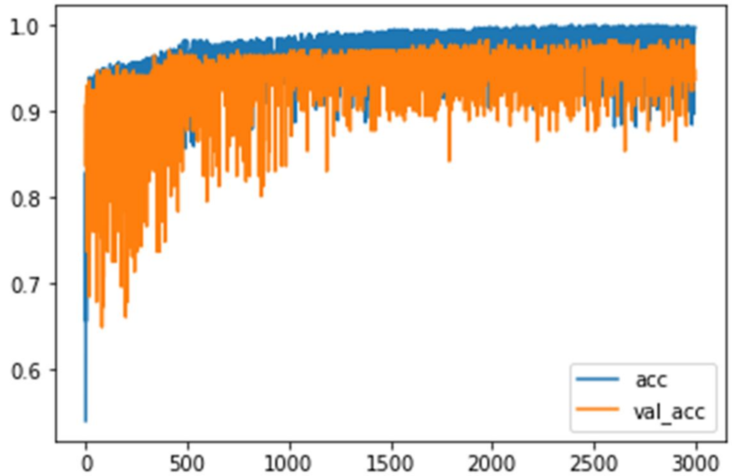


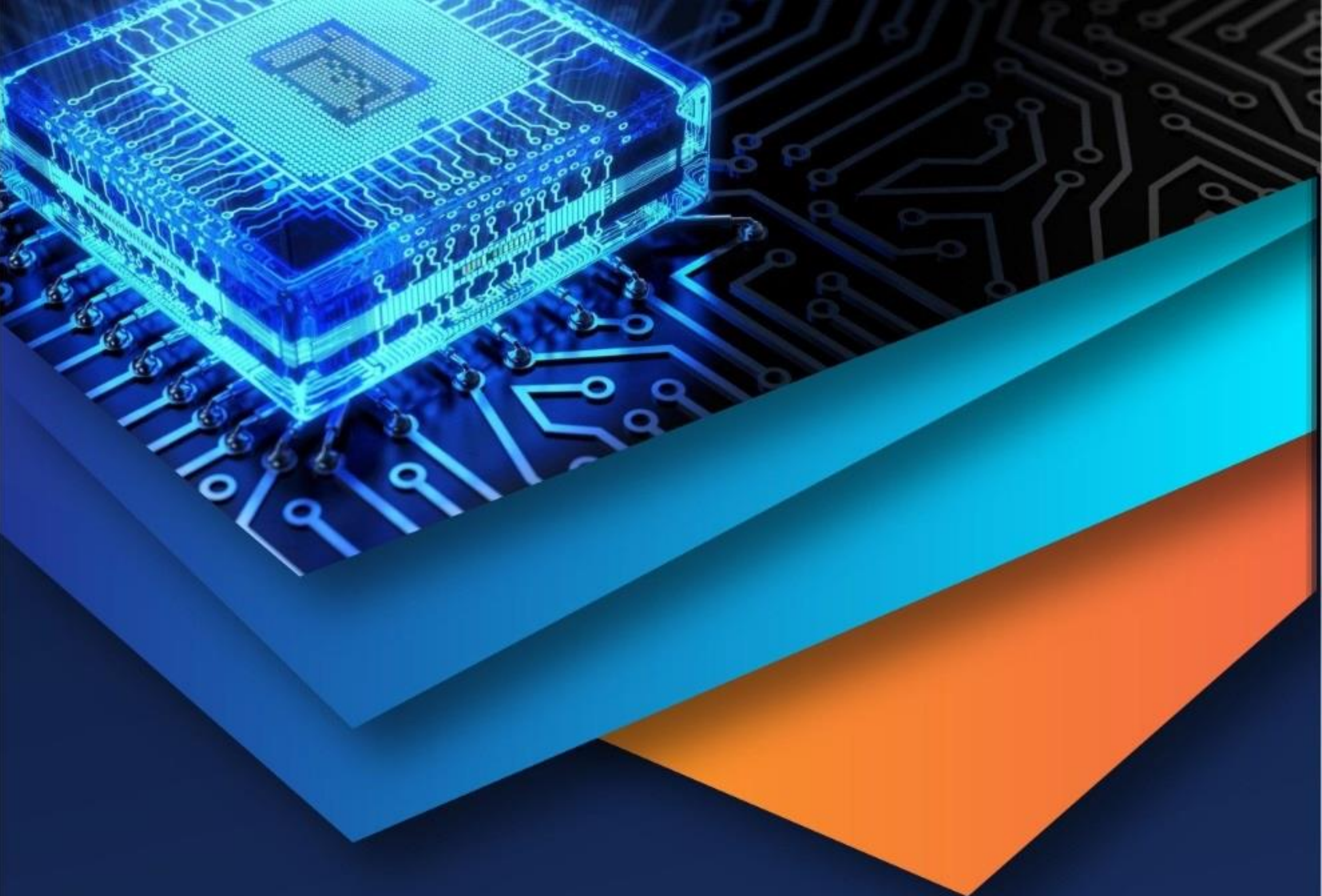
Fig ANN Accuracy and loss graph.

V. CONCLUSION

Classifiers works differently on different nature of data. There is trade-off between the accuracy of prediction on test data, and higher computation time and power. Simpler model will have low variance, high bias whereas complex models can perform well in prediction and can have high accuracy and sometimes it tries to generalize every single data point in the data which can lead to overfitting. ANN performs well on the data, but it takes high computation power and time to train it. So, optimization should be used in order to find the predictions in less time and iterations, so I have used to gradient descent optimization while training. For 5000 epochs for training and testing ANN took 30.772 sec. The results show that this model works well in predicting the cancer. It can be used as intelligent component in medical decision making.

REFERENCES

[1] Yuan and Chen, 2009, “Image Segmentation Method Using Thresholds Automatically Determined from Picture Contents”, EURASIP Journal on Image and Video Processing Volume 2009, Article ID 140492, 15 pages doi:10.1155/2009/140492.
[2] M Veta, JPW Pluim, PJ Van Diest, and MA Viergever, “Breast cancer histopathology image analysis: A review,” IEEE Transactions on Biomedical Engineering, vol. 61, no. 5, pp. 1400 – 1411, May 2014.
[3] P Filipczuk, T Fevens, A Krzyzak, and R Monczak, “Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies,” IEEE Transactions on Medical Imaging, vol. 32, no. 12, pp. 2169 – 2178, Dec. 2013.
[4] YM George, HH Zayed, MI Roushdy, and BM Elbagoury, “Remote computer-aided breast cancer detection and diagnosis system based on cytological images,” IEEE Systems Journal, vol. 8, no. 3, pp. 949 – 964, Sept. 2014.
[5] M Kandemir, C Zhang, and FA Hamprecht, “Empowering multiple instance histopathology cancer diagnosis by cell graphs,” in Proc. International Conference on Medical Image Computing and Computer Assisted Intervention, 2014.
[6] SH Bhandari, “A bad of features approach for malignancy detection in breast histopathology images,” in Proc. IEEE International Conference on Image Processing, 2015.
[7] W Li, J Zhang, and SJ McKenna, “Multiple instance cancer detection by boosting regularised trees,” in Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.
[8] X Li and KN Plataniotis, “Toward breast cancer histopathology image diagnosis using local color binary pattern,” in Proc. 14th Annual Imaging Network Ontario Symposium, 2016.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY