Identification of Breast Cancer using Deep Learning Algorithm with Mammography Images

K.Sreenath[1]*Nannuri Suresh[2]Dr.G.LakshmiVara Prasad [3]K. Hari Krishna[4] T.Rajendran[5] Dr.R.Vijayarangan[6]

[1, 2, 3, 4, 6]Department of Information Technology, QIS College of Engineering and Technology, Ongole, India

[4]Department of ECE, QIS College of Engineering and Technology, Ongole, India

*Corresponding Author: K.Sreenath&srinathits@gmail.com

Abstract: Women are drawn to cancer, the world's most dangerous disease. Thus, our practical goal should be to cure cancer through scientific research, followed by early cancer detection to help eliminate it. We found few cancer detection methods in 41 papers. This research proposes a Deep Learning method convolutional brain organization for diagnosing breast cancer using Mammography MIAS data. Using MIAS Dataset, the research shows how deep learning can diagnose breast cancer. First, we gathered the dataset and used pre-processing technique for scaled and channel information. Then, we divided the dataset into preparation and testing and created a few charts for representation. Execute model on the dataset again and achieve 98. This database has 200 photos and 12 highlights. This report uses 12 breast cancer diagnosis highlights from pre-processing. We employed Watershed Segmentation, Color-based segmentation, and Adaptive Mean Filters to scaled datasets before applying the model and achieving exactness. In this research, we contrast profound learning calculations with different man-made intelligence calculations and find that our system performs well.

1. INTRODUCTION

At a projected 115,251 new cases and 53,592 fatalities in 2008, breast cancer is the second most frequent disease in India, after cervical cancer. India has a high death rate from breast cancer [3, 4], with an age-standardized incidence rate of 2.27 per 10,000, which is one-third that of Western countries. This is also true for other malignancies, such as those of the lungs, prostate, and colon. Breast cancer accounts for 22.4% of all cancer diagnoses and 17.6% of all cancer deaths in India. India's urban breast cancer rate is three times greater than its rural rates.CAD has been used to identify and classify cell abnormalities in histology images.

However, little study has been done on applying ADS to breast histology images or improving a quantitative analytic estimation to identify and grade them.ADS for breast cancer using a quantitative symptomatic framework that recognizes and evaluates breast histological photos dispassionately reduces subjectivity and makes accurate, fast, and consistent diagnostic decisions. Because doctors base breast cancer treatment on tumor grading, it is crucial to grade the tumor accurately. Thus, accurate grading is crucial for proper therapy. Thus, grading goals are essential.

2. RELATED WORK

Many deep learning and machine learning methods can identify and predict cancer. Convolutional and recurrent neural networks, as well as pre-trained models Alex Net, Google Net, VGG16, VGG19, and ResNet, are employed for breast

cancer diagnosis. Mammogram images, SEER, UCI, WBCD, and others are popular teaching and testing datasets.

2.1 DataCollection

This section describes the data-collection methods. In 2012, this research began to develop an objective computer-aided breast cancer grading system. To do this, the search for labelled breast cancer histology digital images began. The search for a digital breast tumor biopsy data set and laboratories that could provide it failed. Thus, labelled datasets were not allowed. Thus, biopsy slides were used to capture images for therapeutic professionals.

2.2 HistologySlidePreparation

This study uses breast biopsy histology images. Breast biopsies remove suspected cancerous tissue. Surgeons perform it before surgery. Tissue samples provide histology slides. Tissue samples must be handled as follows.

- Fixation/Processing
- Embedding
- Sectioning
- Staining
- RegularStaining
- Coverslipping
- TheMountingReagent

2.3 ImageAcquisition

In this examination, an example of chest biopsy slides were assembled somewhere near 2010 and 2015 was used. The information includes kind, grade1, grad2 and gread3 cells inside the bosom biopsy. A dubious return on initial capital investment catches the picture. The chest histology slides were furnished with a 10x, 20x, and 40x amplifying focal point with a spatial goal of 4928 x 3264 pixels.

2.4 Imagedataexplanatorynote

Breast histology slides were used after security. Pathologists partitioned and rated all pictures. An expert pathologist named the photos based on the distribution of aberrant cells and marked some as typical and weird. Expert pathologists evaluated the slides as benign, Grad1, Grad2, and Grad3. This research used pathologist-scored photos as "ground truth" for training and testing. Table 1 shows all benign Grad1, Grad2, and Grad3 images.

Table1:Total number of images collected

Type of Image	Total Images	Zoom
Benign	51	40x
Grad1	108	40x
Grad2	121	40x
Grad3	125	40x

We now have 404 JPG photos in our database, each with a spatial resolution of 4928 x 3264 pixels.

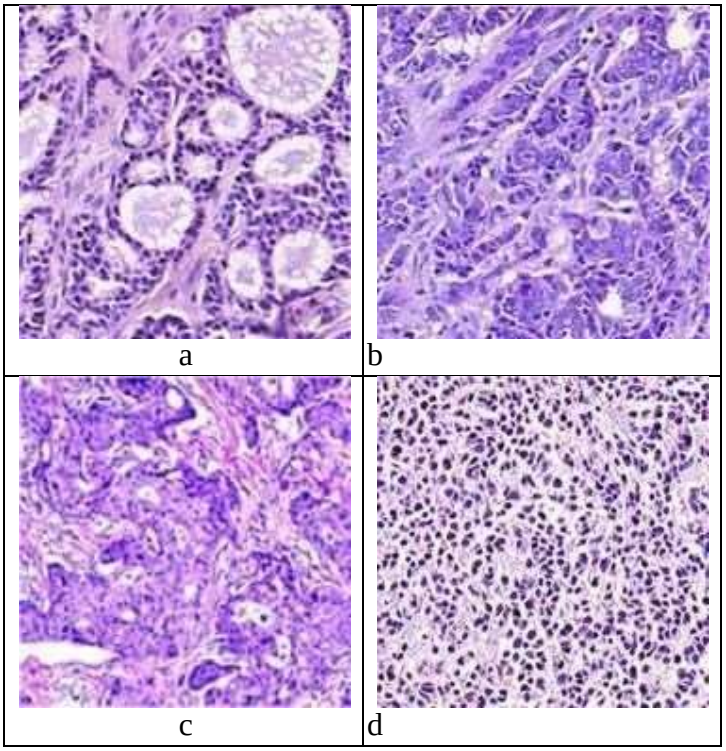

Figure 1: Examples of breast images:(a)normal(b)G1(c)G2(d)G3

2.5 Cellular–Level Feature Extraction

we address the methodology to focus the cell - level elements. To get cell - level elements we have used the district developing using multi scale space seed distinguishing proof division to piece the cells. Multi scale space mass identification is a gadget for dividing a bosom tissue histology picture into its part locales. District developing using multi scale space seed distinguishing proof division is a neighborhood approach that examines a bosom tissue histology picture from a tissue level and chooses the locale of interest. Copying this system, to portray the kind of cores, morphological component extraction is associated.

It exhibits the means in remembered for cell qualities extraction of bosom tissue histology. The overall method apportioned into three areas, seed revelation, cell division and highlights extraction. Cultivated locale developing (SRG) [2] is strong, fast and liberated from tuning boundaries. These perspectives license execution of incredible calculation. On the other hand, the SRG calculations moreover experience the evil impacts of the issues of programmed seed age and pixel arranging orders for marking.

2.6 Laplacian Filtering

Remembering the ultimate objective to recognize cell in made scale-space, we channel every one set aside picture using an unbounded 2D Lapalacian channel. In a picture that has been convolved with a Gaussian, this is the channel that is most effective at identifying symmetric mass-like objects.We can kind of take the Laplaican of a Gaussian and convolve that with the converted BR image to explain the properties of convolution.

Figure 2:LaplacianofGaussian.

3. RESULTS AND DISCUSSION

We employed convolutional neural network to diagnose breast cancer and implemented the same dataset on additional machine learning algorithms like neural network, SVM, and random forest. Predict models and preprocessing data comprise the study.

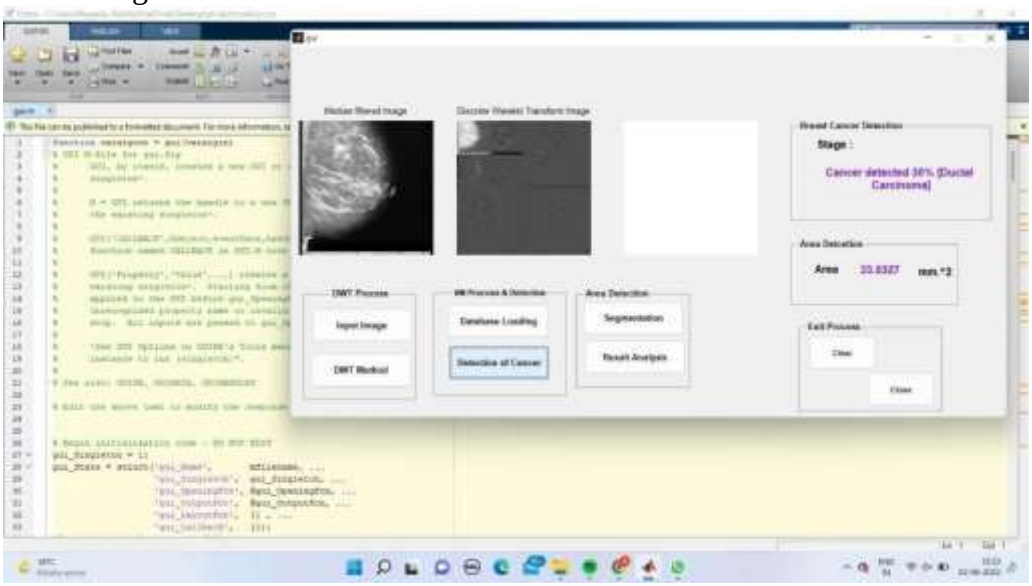

Figure 3: Cancer area detection

This paper uses the freely available MIAS dataset of 200 histopathology images. In the experiment, we pre-processed the dataset, retrieved useful features using histogram and watershed model, implemented encoders, and discovered the final training dataset. Deep learning model convolutional neural network approach obtained 98% accuracy for the trained dataset.

4. Conclusion and Future Work

India's rising breast cancer risk, especially in younger women, confronts pathologists. Indian habits increase breast cancer rates. The expansion of pollutants and hormone-infused food, the rise in cigarette and alcohol use, unpleasant lifestyles, and other factors create more cancer cases worldwide, forcing disease authorities to make faster, more accurate decisions. Since algorithmic methods are more consistent and quantitative, pathologists must use automated leaders like computers to support them.Automated breast cancer diagnosis investigations are mostly cellular or tissue-level. Cell and nuclei segmentation is essential for cellular-level analysis of digital microscope pictures.This investigation will examine distinct

tumour tissues with commanding structures. This idea of using non-nuclei pieces with other structural approaches could be a future shift. Because cancer tissue has aberrant nuclei-to-cytoplasm ratios, it should be included in cellular- level feature extraction. The cell segmentation algorithm segments nuclei and cytoplasm.

References

1. Sharma, P. Khanna, "ROI Segmentation using Local Binary Image," IEEE International conference on System, Computing and Engineering, pp 136-141, 2013.

2. W. P. Jr Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for speculated lesions," Radiology, vol. 191, no. 2, pp. 331–337, 1994.

3. R. Campanini, D. Dongiovanni, E. Iampieri, N. Lanconelli, M. Masotti, G. Palermo, A. Riccardi, and M. Roffilli, "A novel featureless approach to mass detection in digital mammograms based on support vector machines," Phys. Med. Biol., vol. 49, no. 6, pp.961– 975, 2004.

4. B. Sahiner, H. P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," IEEE Trans. Med. Imag., vol. no. 15, pp. 598–610, 1996.

5. R. Bellotti, et.al., ''A completely automated CAD system for mass detection in a large mammographic database,'' Med. Phys., vol. no. 33, pp. 3066--3075, 2006.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon,1892, pp.68–73.

6. R. C. Gonzalez., "Digital image processing," 2nd Edition. India: Prentice Hall, 2003.

7. J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," Plenum Press, New York, 1981.

8. R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, second ed., John Wiley andSons, New York 2001.

9. H. Tamura, S Mori , T. Yamavaki " Textural features corresponding to visual perception," IEEE Trans On Systems, Man and Cybernetics, vol8, pp 460-472, 1978.

10. M. K. Hu, "Visual pattern recognition by moment invariants," IRE Transactions on Information Theory, vol. 8, pp 179-187, 1962.

11. T. M. Nguyen and R. M. Rangayyan, ''Shape Analysis of Breast Masses in Mammograms via the Fractal Dimension," in Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China,2005.

12. Ibrahim, S. Jafar Ali, and M. Thangamani. "Prediction of Novel Drugs and Diseases for Hepatocellular Carcinoma Based on Multi- Source Simulated Annealing Based Random Walk." *Journal of medical systems* 42, no. 10 (2018): 188. https://doi.org/10.1007/s10916- 018-1038-y

13. Thangamani, M., and S. Jafar Ali Ibrahim. "Ensemble Based Fuzzy with Particle Swarm Optimization Based Weighted Clustering (Efpso-Wc) and Gene Ontology for Microarray Gene Expression."*In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing, pp. 48-55. 2018. https://dl.acm.org/doi/abs/10.1145/3299852.3299866*

14. S. Jafar Ali Ibrahim and M. Thangamani. 2018. Momentous Innovations in the Prospective Method of Drug Development. In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing (DMIP '18). *Association for Computing Machinery, New York, NY, USA,* 37–41. https://doi.org/10.1145/3299852.3299854

15. Ibrahim, S. Jafar Ali, and M. Thangamani. "Enhanced singular value decomposition for prediction of drugs and diseases with hepatocellular carcinoma based on multi- source bat algorithm based random walk." *Measurement* 141 (2019): 176-183. https://doi.org/10.1016/j.measurement.2019.02.056

16. Dr. M. Thangamani, Jafar Ali Ibrahim, Information Technology E-Service Management System, International Scientific Global Journal in Engineering Science and Applied Research (ISGJESAR). Vol.1. Issue 4, pp. 13-18, 2017. http://isgjesar.com/Papers/Volume1,Issue4/paper2.pdf \

17. **Dr**.R.Chinnaiyan , M.S.Nidhya (2018), " Reliability Evaluation of Wireless Sensor Networks using EERN Algorithm" , Lecture Notes on Data Engineering and Communications Technologies, Springer International conference on Computer Networks and Inventive Communication Technologies (ICCNCT - 2018), August 2018 (Online)

18. Dr.R.Chinnaiyan , R.Divya (2018), " Reliable AI Based Smart Sensors for Managing Irrigation Resources in Agriculture" , Lecture Notes on Data Engineering and Communications Technologies, Springer International conference on Computer Networks and Inventive Communication Technologies (ICCNCT - 2018), August 2018 (Online)

19.

20. Dr.R.Chinnaiyan , S.Balachandar (2018) , " Reliable Digital Twin for Connected Footballer" , Lecture Notes on Data Engineering and Communications Technologies, Springer International conference on Computer Networks and Inventive Communication Technologies (ICCNCT - 2018), August 2018 (Online)

21. Dr.R.Chinnaiyan , S.Balachandar (2018) , " Centralized Reliability and Security Management of Data in Internet of Things (IoT) with Rule Builder" , Lecture Notes on Data Engineering and Communications Technologies, Springer International conference on Computer Networks and Inventive Communication Technologies (ICCNCT - 2018), August 2018 (Online)

22. Dr.R.Chinnaiyan, Abishek Kumar (2017) " Reliability Assessment of Component Based Software Systems using Basis Path Testing" , IEEE International Conference on Intelligent Computing and Control Systems, ICICCS 2017, 512 – 517

23. Dr.R.Chinnaiyan, Abishek Kumar (2017) " Reliability Assessment of Component Based Software Systems using Basis Path Testing" , IEEE International Conference on Intelligent Computing and Control Systems, ICICCS 2017, 512 – 517

24. Dr.R.Chinnaiyan, Abishek Kumar(2017) ,"Construction of Estimated Level Based Balanced Binary Search Tree", 2017 IEEE International Conference on Electronics,Communication, and Aerospace Technology (ICECA 2017), 344 - 348, 978-1-5090-5686-6.

25. Dr.R.Chinnaiyan, Abishek Kumar(2017), Estimation of Optimal Path in Wireless Sensor Networks based on Adjancy List, 2017 IEEE International Conference on Telecommunication,Power Analysis and Computing Techniques (ICTPACT2017) ,6,7,8th April 2017,IEEE 978-1-5090-3381

26. Dr.R.Chinnaiyan, AbishekKumar(2017) ,"Construction of Estimated Level Based Balanced Binary Search Tree", 2017 IEEE International Conference on Electronics,Communication, and Aerospace Technology (ICECA 2017), 344 - 348, 978-1-5090-5686-6.

27. Dr.R.Chinnaiyan, AbishekKumar(2017), Estimation of Optimal Path in Wireless Sensor Networks based on Adjancy List, 2017 IEEE International Conference on Telecommunication,Power Analysis and Computing Techniques (ICTPACT2017) ,6,7,8th April 2017,IEEE 978-1-5090-3381-2.

28. Dr.R.Chinnaiyan, R.Divya (2017)," Reliability Evaluation of Wireless Sensor Networks", IEEE International Conference on Intelligent Computing and Control Systems, ICICCS 2017, 847 – 852

29. Dr.R.Chinnaiyan, Sabarmathi.G (2017)," Investigations on Big Data Features , Research Challenges and Applications", IEEE International Conference on Intelligent Computing and Control Systems, ICICCS 2017, 782 – 786

30. G.Sabarmathi , Dr.R.Chinnaiyan (2018), "Envisagation and Analysis of Mosquito Borne Fevers – A Health Monitoring System by Envisagative Computing using Big Data Analytics" in ICCBI 2018 – Springer on 19.12.2018 to 20.12.2018 (Recommended for Scopus Indexed Publication IEEE Xplore digital library)

31. G.Sabarmathi , Dr.R.Chinnaiyan, Reliable Data Mining Tasks and Techniques for Industrial Applications, IAETSD JOURNAL FOR ADVANCED RESEARCH IN APPLIED SCIENCES, VOLUME 4, ISSUE 7, DEC/2017,PP- 138-142, ISSN NO: 2394-8442

32. Ibrahim, Mr S. Jafar Ali, K. Singaraj, P. Jebaroopan, and S. A. Sheikfareed. "Android Based Robot for Industrial Application." International Journal of Engineering Research & Technology 3, no. 3 (2014).

33. Ibrahim, S. Jafar Ali, and M. Thangamani. "Momentous Innovations in the Prospective Method of Drug Development." In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing, pp. 37-41. 2018.

34. Ibrahim, S. Jafar Ali, and M. Thangamani. "Prediction of Novel Drugs and Diseases for Hepatocellular Carcinoma Based on Multi-Source Simulated Annealing Based Random Walk." Journal of medical systems 42, no. 10 (2018): 188. https://doi.org/10.1007/s10916-018-1038-y ISSN 1311-8080, https://acadpubl.eu/hub/2018-119-16/1/94.pdf

35. Jafar Ali Ibrahim. S, Mohamed Affir. A "Effective Scheduling of Jobs Using Reallocation of Resources Along With Best Fit Strategy and Priority", International Journal of Science Engineering and Advanced Technology(IJSEAT) – ISSN No: 2321- 6905, Vol.2, Issue.2, Feb-2014, http://www.ijseat.com/index.php/ijseat/article/view/62

36. M. Thangamani, and Jafar Ali Ibrahim. S, "Knowledge Exploration in Image Text Data using Data Hiding Scheme," Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2018, 14-16 March, 2018, Hong Kong, pp352-357 http://www.iaeng.org/publication/IMECS2018/IMECS2018_pp352-357.pdf

37. M. Thangamani, and Jafar Ali Ibrahim. S, "Knowledge Exploration in Image Text Data using Data Hiding Scheme," Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2018, 14-16 March, 2018, Hong Kong, pp352-357 http://www.iaeng.org/publication/IMECS2018/IMECS2018_pp352-357.pdf

38. R.Chinnaiyan, S. Somasundaram (2011) ,"An SMS based Failure Maintenance and Reliability Management of Component Based Software Systems", European Journal of Scientific

Research, Vol. 59 Issue 1, 9/1/2011, pp.123 (cited in EBSCO, Impact Factor: 0.045)

39. R.Chinnaiyan, S.Somasundaram (2012) , Reliability Estimation Model for Software Components using CEP", International Journal of Mechanical and Industrial Engineering (IJMIE) , ISSN No.2231-6477, Volume-2, Issue-2, 2012, pp.89-93.

40. R.Chinnaiyan, S.Somasundaram(2010) "Evaluating the Reliability of Component Based Software Systems " ,International Journal of Quality and Reliability Management , Vol. 27, No. 1., pp. 78-88 (Impact Factor: 0.406)

41. R.Chinnaiyan, S.Somasundaram(2011), "An Experimental Study on Reliability Estimation of GNU Compiler Components - A Review", International Journal of Computer Applications, Vol.25, No.3, July 2011, pp.13-16. (Impact Factor: 0.814)

42. S. Jafar Ali Ibrahim and M. Thangamani. 2018. Momentous Innovations in the Prospective Method of Drug Development. In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing (DMIP '18). Association for Computing Machinery, New York, NY, USA, 37–41. https://doi.org/10.1145/3299852.3299854

43. S. Jafar Ali Ibrahim and Thangamani, M "Proliferators and Inhibitors Of Hepatocellular Carcinoma", International Journal of Pure and Applied Mathematics (IJPAM) Special Issue of Mathematical Modelling of Engineering Problems Vol 119 Issue. 15. July 2018

44. Thangamani, M., and S. Jafar Ali Ibrahim. "Ensemble Based Fuzzy with Particle Swarm Optimization Based Weighted Clustering (Efpso-Wc) and Gene Ontology for Microarray Gene Expression." In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing, pp. 48-55. 2018. https://dl.acm.org/doi/abs/10.1145/3299852.3299866