Outcomes: Development and Execution – Hybrid Segmentation and Mammogram Classification for Detection and Prevention of Cancer

[1]Neeti Arora,[2]Dr.Gaurav Aggarwal

[1]Research Scholar, [2] Professor and Head Department of Computer Science and Engineering, Jagannath University, NCR Haryana.

Abstract: This fragment offering the obtained results ofthe input dataset, flow/system model, image pre-processing techniques, image feature extraction technique, machine learning, and performance evaluation metrics. The primary goal is to be classifying the breast cancer from the huge medical database based on the density level then the severity level has been observed like normal, moderate and severe. In order to identify the non-cancerous and cancerous image, we used texture and edge based segmentation approach

Keywords: Image processing techniques, extraction technique, segmentation approach.

Introduction

The purpose is to detect and diagnosis the breast cancer in early stage. Initially, we have pre-processed the input image and normalized it with high resolution which describes the complex texture and geometric structure of an image. Then the image features are extracted for histopathological images of breast cancer which improves the system classification performance. Furthermore, the extracted image features are imported into classification module which differentiate the projection and direction of breast then remove the pectoral muscle which might influence the output results. With this regards, one can differentiate with injuries like masses and normal breast. The proposed system performance has been tested and validated using MATLAB simulation software

Input dataset

In this learning, the image technique is used to study the metabolism in the body for detecting the disease. The images were collected from cancer imaging archive dataset which is online free resources. MRI images of 30 patients were collected. Specifically, the breast cancer images have been considered in three axial slices which comprises the one with the largest tumour diameter.

Evaluation Criteria

The performance indices can be considered both for image classification and image retrieval. Accordingly, they are referred as classification performance indices and retrieval performance indices.

Retrieval Performance Indices

For testing the accurateness and competence of a query algorithmcompletely, some evaluation indices are used and helpful for comparing the performance of the algorithms. Though there are many indices for assessing the performance, the most significant assessing metrics for the analysis of image retrieval performance are precision and recall indices which can be well-defined as under (Fadaei et al., 2017; Ji et al., 2017)

Precision

It is the ratio of the number of related retrieved images to the total number of images retrieved.

$$\text{Precision} = \frac{\text{number of related image retrieved}}{\text{Total number of images retrieved}}$$

It is evident from the equation that the precision is more if a greater number of related images are retrieved. Precision is the measurement that expresses the capability of the system to retrieve only the most relevant image.

Recall

It is the ratio of total retrieved images to the total images in the dataset.

$$\text{Recall} = \frac{\text{number of related image retrieved}}{\text{Total number of related images in the database}}$$

Since recall is directly proportional to the number of related images retrieved, the more the number of related retrieved images, the more will be the recall. Recall is the measurement that shows the capacity of the system to retrieve all the replicas that are relevant.

Classification performance Indices

There are various metrics/indices available for evaluating the performance of retrieval classification. Average Precision, Average Recall, Accuracy, F-measure, Error Rate, Correct Rate, Specificity and Sensitivity are few among them.

Average Precision

The mathematical expression for Average Precision is given by,

$$Average\ \mathrm{Pr}ecision = \frac{TP}{TP + FP}$$

Average Recall

Mathematically, the average Recall can be expressed as,

$$Average\ \mathrm{Re}call = \frac{TP}{TP + FN}$$

Accuracy

The accuracy of the classifier is calculated mathematically from the following equation.

$$Accuracy = \frac{TN + TP}{TP + FP + FN + TN}$$

F-measure

Precision and recall values are utilized to estimate the F-measure and is known as as balanced F-score, the equation for which is given by,

$$F - measure = 2 * \frac{\mathrm{Pr}ecision * \mathrm{Re}call}{\mathrm{Pr}ecision + \mathrm{Re}call}$$

Error Rate and Correct Rate

Error rate is nothing but the ratio between incorrectly classified data and the classified samples. The ratio of correctly classified data to the total classified data is known as Correct Rate. They are mathematically defined by the following equations.

$$Error\ Rate = \frac{incorrectly\ classified\ samples}{classified\ samples} = \frac{FP + FN}{TP + FP + FN + TN}$$

$$Correct\ Rate = \frac{correctly\ classified\ samples}{classified\ samples} = \frac{TP + TN}{TP + FP + FN + TN}$$

Results

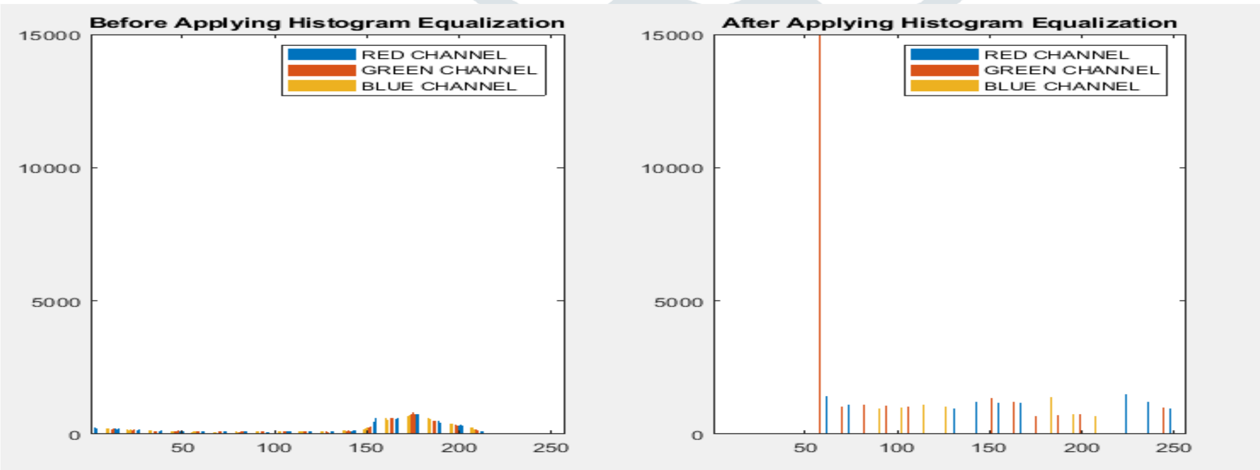
This section presents the performance evaluation of our proposed approach on publicly available medical image datasets. The performance evaluation compared the results with traditional techniques are discussed. In our study, breast cancer image dataset were utilized,and it has been randomly split into a test set of 50 images, validation 20 images and training set of 80 images. In order to fine-tune the system performance, the validation stage has been utilized, whereas each value computes the validations set and obtain the highest mean average precision.

The input image is pre-processed using image contrast enhancement approach.

Image equalization/ Contrast enhancement/ Brightness or illumination variation

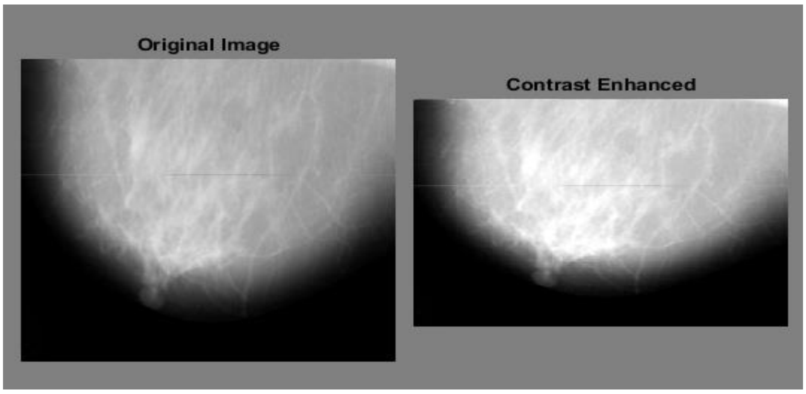
Histogram Equalization (HE) is an image processing approach used to expand contrast in images. It accomplished through spreading out the most frequent intensity values, i.e. stretching out the intensity range of the image. On the other hand, this method is primarily used for sharping the edge portion of image and gather detailed features data that belongs to the high frequency components. Some cases, the high frequency components may consist noises. The suggested model is utilized to remove the noises from the image in effective manner and extracted the edge portion of an image.

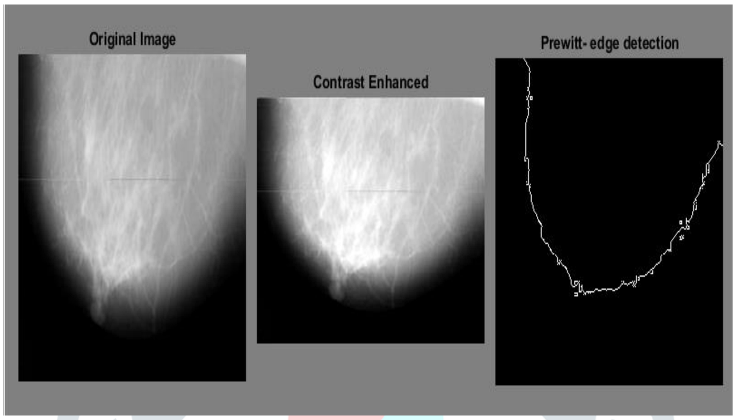
Image histogram analysis



Generally, image histogram is a graphical representation of the intensity distribution of an image. The RGB (Red, green, blue) channel intensity distribution for both input and histogram equalized images are shown in the figure.

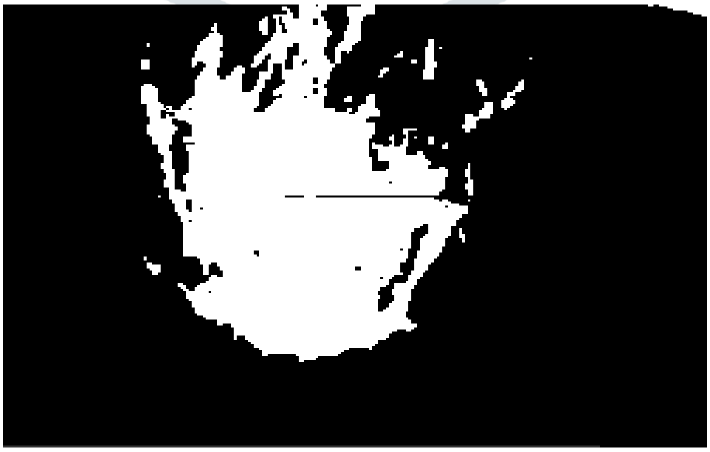
Contrast enhancement

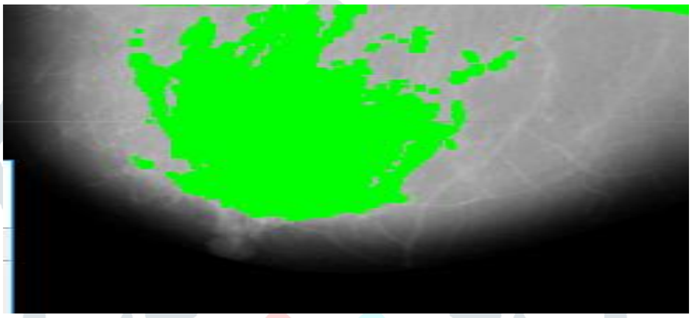




The image contrast enhancement is performed to the variation in luminance or color which makes an object. A higher magnitude of contrast designates additional apparent discrepancies among colours. The picture contrast or intensity ratio is measured by adopting 3x3 block nearby every pixel (x,y). the picture of contrast could be illustrated by means of 0-255 gray levels when regularize the image contrast level.

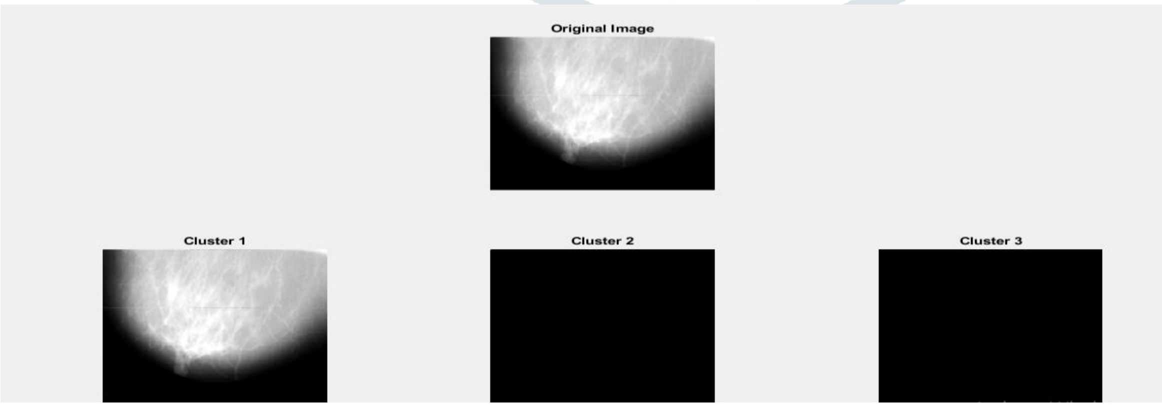
Shape and texture extraction





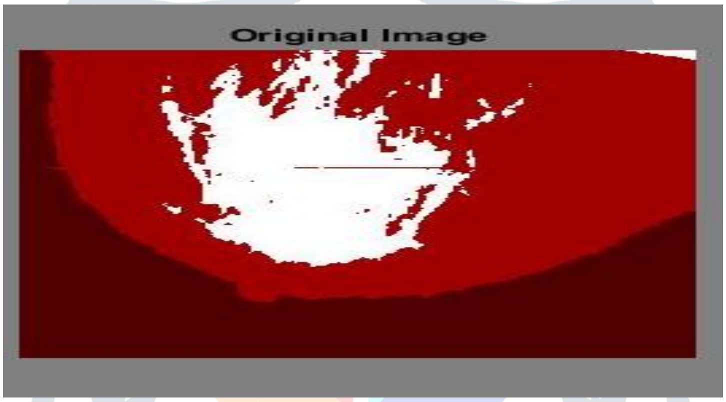
The extracted image features that means shape and texture region are shown in the above figure. We have used prewitt method for detection and extraction of edge portion in an image. Initially, we have removed the noise using Gaussian filter. Then we have calculated the edge direction and edge gradient of a given image. The gradient direction is perpendicular to edges which is rounded to one of four angles that represents vertical, horizontal and two diagonal directions.

CNN classifier



In our study the image is segmented on the basis of classification with least squares CNN. Initially, we have extracted the color and texture features then this has been imported into SVM classifier. This extraction has been performed by filtering approach with local spatial similarity measure framework. Then its has been trained into fuzzy classifier. The simulated results shows that the obtained segmented results gives better performance with respect to faster processing and increased quality of segmented image as compared with other segmented method. Generally, CNN based learning approach uses regression and classification based supervision learning approach. In our study, the learning is based on the linear classification where it cluster/segment the pixel information both color and texture value. The performance of suggested model has been evaluated with respect to mean square error ratio and standard deviation. For obtaining better results, the mean square error ratio should be less. Also, the standard deviation should be high which is used to estimate the extracted results contrast ratio. From this, the obtained results provided better results with respect to error ratio and contrast level.

Image contrast enhancement

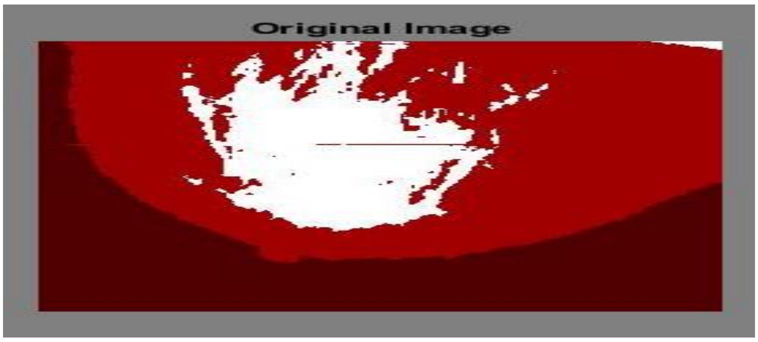


(a)

(b)

The Contrast adjustment is not only used for the image segmentation, but it also helps to view and understand the image data in an efficient manner. It also supports low contrast images. The contrast, of an image, is adjusted by automatic contrast adjustment technique. The edge detection is a part of image segmentation. The effectiveness of many image processing; also, computer vision tasks depends on the perfection of detecting meaningful edges. It is one of the techniques for detecting intensity discontinuities in a digital image. The process of classifying and placing sharp discontinuities in an image is called edge detection.

Image noise reduction

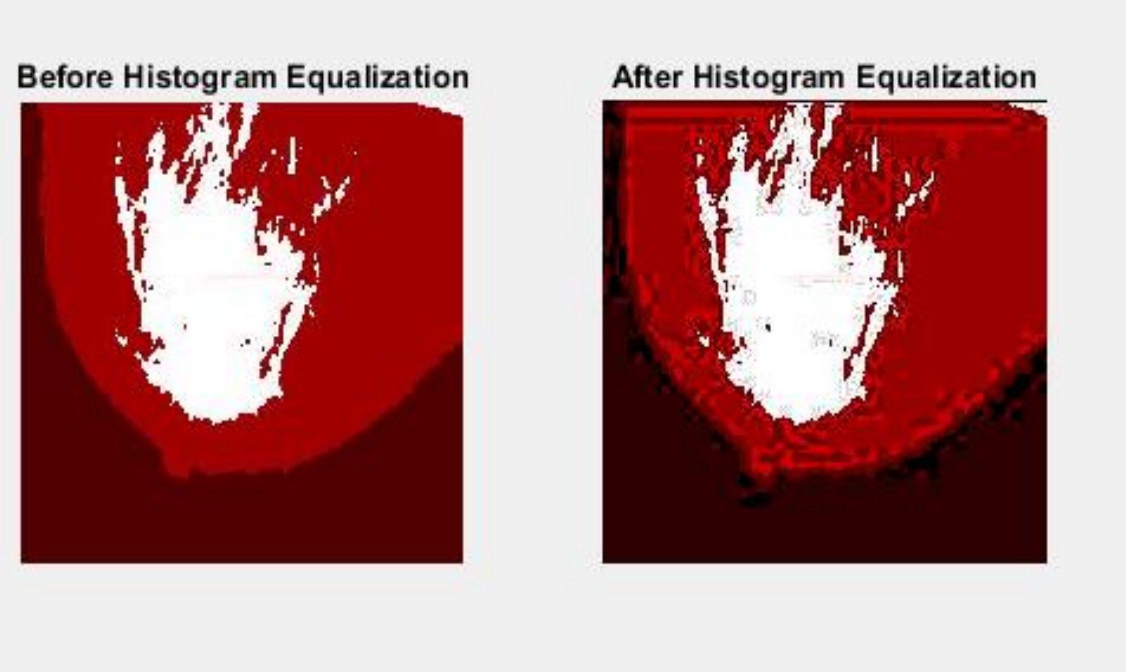


(a)

(b)

(c)

Image histogram equalization



Histogram analysis

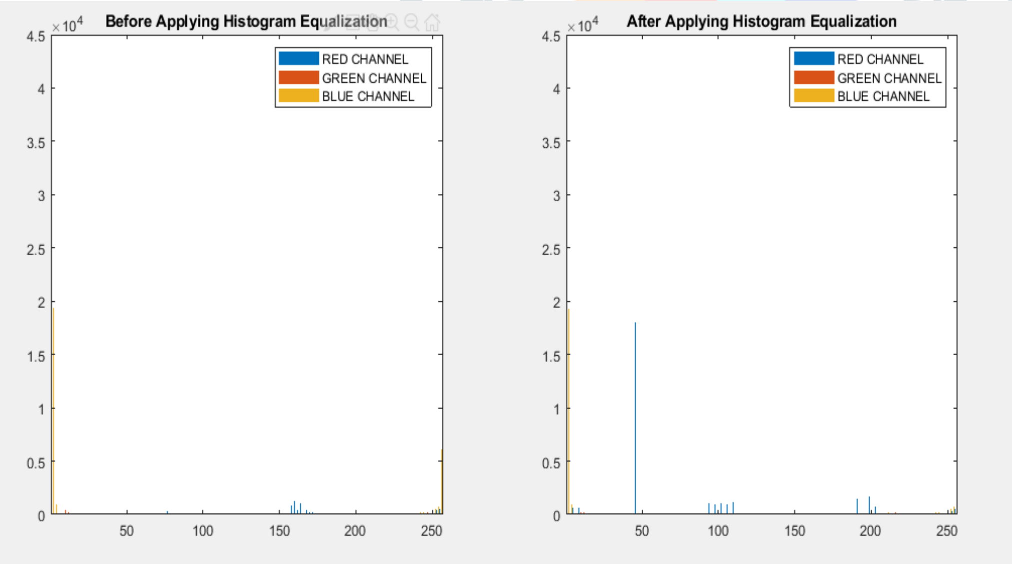
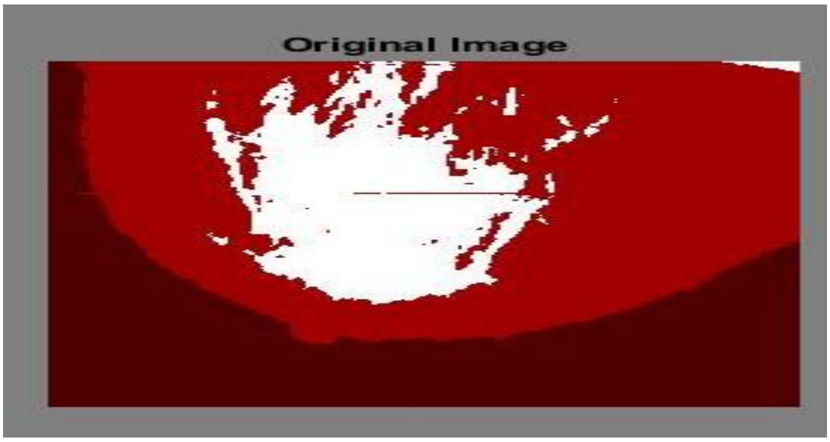
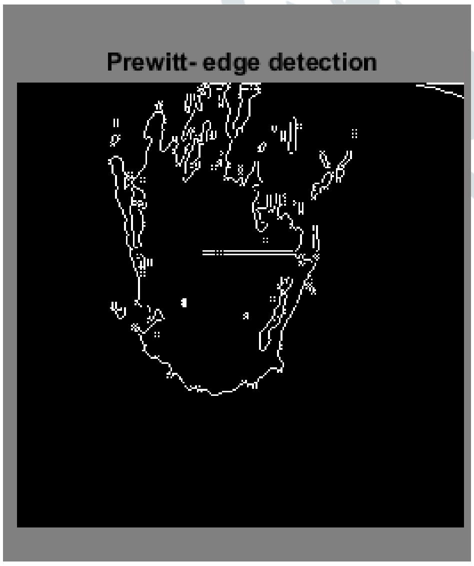


Image edge detection



(a)

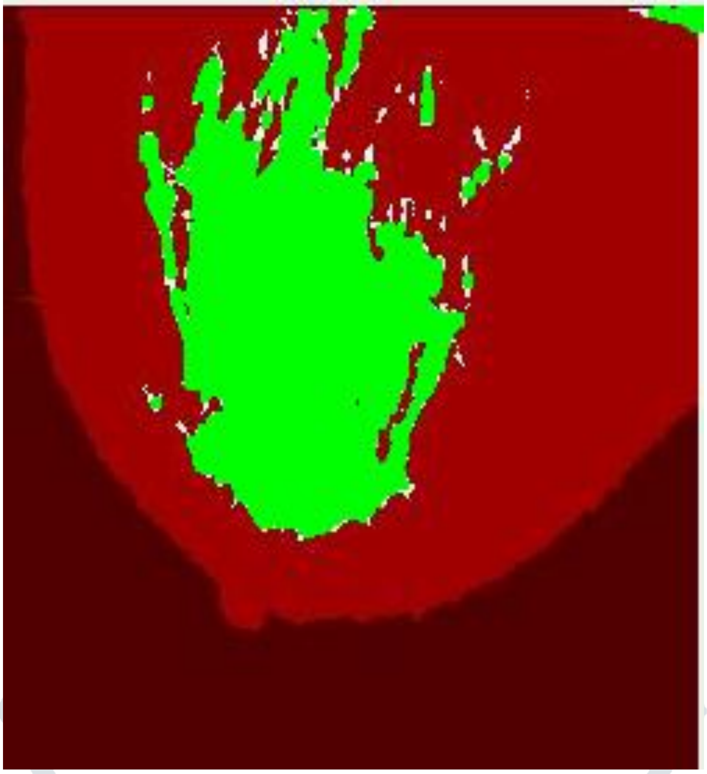
(b)



(c)

The image texture feature has been extracted based on the gray level co-occurrence matrix (GLCM). Then the shape features are extracted using connected regions. The texture features are measured based on the image contrast ratio, dissimilarity, energy, mean, variance, and standard deviation. GLCM describes relationship amongst neighbour pixel and reference pixel in various orientation. Here the value is measured based on the co-occuracnce of pixels together. Shape and Texture extraction

Segmented region



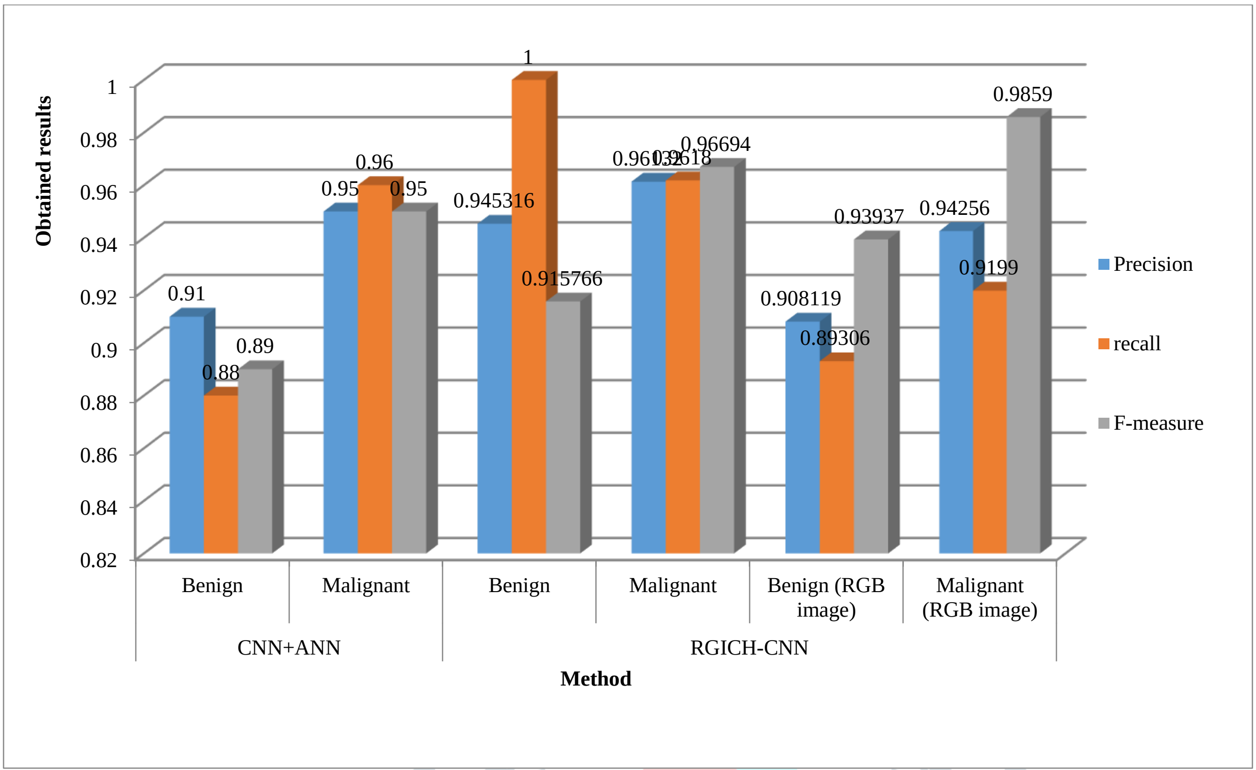
In our study, the segmentation is performance to split-up of single or many areas or objects in a metaphordepending on a similarity or discontinuity measure. An area of the image or metaphor could be stated by its boundary (border) or by the inner region contained by it. The techniques of segmenting the metaphors can usually be classified into two groups namely, "edge"-based and "region"-based approaches.

Summary of obtained result (RGICH-CNN)

Predicted results	Precision	Recall	F-measure	Sensitivity	Specificity
Benign	0.945316	1	0.915766	1	0.943076
malignant	0.96132	0.9618	0.96694	0.9962	0.94481
Benign (RGB image)	0.908119	0.89306	0.93937	0.98428	0.96846
Malignant (RGB image)	0.94256	0.9199	0.9859	0.9086	0.96891
Average/total	**0.939329**	**0.94369**	**0.951994**	**0.97227**	**0.956314**

The performance of suggested model (by combining region growing, feature extraction method and CNN classifier) has been validated in terms of precision, recall, f-measure, sensitivity and specificity. Also, we have verified performance using breast cancer disease diagnosis from patient outcomes from tumour tissue images in both benign and malignant level. The average performance ratio for precision, recall, f-measure, sensitivity and specificity is 0.939329, 0.94369, 0.951994, 0.97227, and 0.956314 respectively.

Performance evaluation with traditional method for benign and malignant of breast cancer

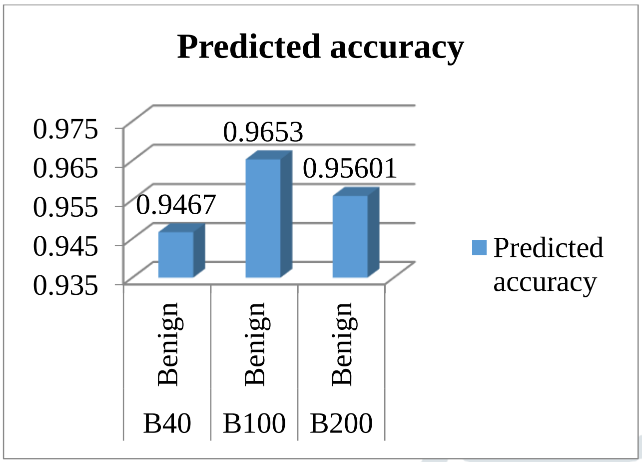


We have used 5000 images of benign and malignant breast tissue in this model with a train-test split of 80:20 and obtained the desired result. The predictive output of three benign and three malignant images from the testing set are displayed in figures.

Predicted results using RGICH-CNN

Label	**Actual class**	**Predicted accuracy**	**Predicted class**
B40	Benign	0.9467	Benign
B100	Benign	0.9653	Benign
B200	Benign	0.95601	Benign
M40	Malignant	0.97105	Malignant
M100	Malignant	0.9806	Malignant
M200	Malignant	0.9963	Malignant

Simulated results of accuracy ratio using benign breast cancer image



The performance of proposed model as RGICH-CNN has been validated in terms of extracting the features namely tumor size, number of nodes, extension of tumor. By the simulation, the attained accuracy is 96.93%.

Existing methods and respective Accuracy

Year	**Author**	**Method**	**Accuracy**	**Error rate**
2017	Sun and Binder,(2017)	Pre-Trained Networks	89	4.74
2017	Samah et al.,(2017)	K-Nearest Neighbor	86	19.28
2017	Song et al., (2017)	Feature Extracted Using CNN	90	4.28
2018	Adeshina et al., (2018)	Deep Convolution Neural Network	91.54	8.54
2019	Dabeer et al.,(Dabeer et al., 2019)	convolutional neural network	93.4458	6.55
2019	Our proposed model	RGICH-CNN	96.93	3.07

Outcome

The suggested model is utilized toward minimize or avoid the noises from the image in effective manner and extracted the edge portion of an image by incorporating feature extraction, segmentation and classification method. With regards, we able to suppress the enlarged noise and minimized the computation complexity of the system. The resulted outcome from the histogram equalization is efficiently enhanced the edges, contrast, and local features of an image. This method usually increases the global contrast of images when its usable data is represented by close contrast values. This allows for areas of lower local contrast to gain a higher contrast.

References

Adeshina, S.A., Adedigba, A.P., Adeniyi, A.A. & Aibinu, A.M. (2018). Breast Cancer Histopathology Image Classification with Deep Convolutional Neural Networks. In: *2018 14th International Conference on Electronics Computer and Computation (ICECCO)*. November 2018, IEEE, pp. 206–212.

Akila, K., Jayashree, L.S. & Vasuki, A. (2015). Mammographic Image Enhancement Using Indirect Contrast Enhancement Techniques – A Comparative Study. *Procedia Computer Science*. 47. pp. 255–261.

Berry, D.A., Cronin, K.A., Plevritis, S.K., Fryback, D.G., Clarke, L., Zelen, M., Mandelblatt, J.S., Yakovlev, A.Y., Habbema, J.D.F. & Feuer, E.J. (2005). Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer. *New England Journal of Medicine*. 353 (17). pp. 1784–1792.

Bevilacqua, V., Pannarale, P., Abbrescia, M., Cava, C., Paradiso, A. & Tommasi, S. (2012). Comparison of data-merging methods with SVM attribute selection and classification in breast cancer gene expression. *BMC Bioinformatics*. 13 (S7). pp. S9.

Carney, P.A., Miglioretti, D.L., Yankaskas, B.C., Kerlikowske, K., Rosenberg, R., Rutter, C.M., Geller, B.M., Abraham, L.A., Taplin, S.H., Dignan, M., Cutter, G. & Ballard-Barbash, R. (2003). Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Annals of internal medicine*. 138 (3). pp. 168–75.

Casti, P., Mencattini, A., Salmeri, M., Ancona, A., Mangeri, F., Pepe, M.L. & Rangayyan, R.M. (2016). Contour-independent detection and classification of mammographic lesions. *Biomedical Signal Processing and Control*. 25. pp. 165–177.

Cheng, H.D., Shan, J., Ju, W., Guo, Y. & Zhang, L. (2010). Automated breast cancer detection and classification using ultrasound images: A survey. *Pattern Recognition*. 43 (1). pp. 299–317.

Emory Winship Cancer Institute (2019). *Advanced Radiation Therapy for Your Breast Cancer*. 2019.

Gandhi, K.R. & Karnan, M. (2010). Mammogram image enhancement and segmentation. In: *2010 IEEE International Conference on Computational Intelligence and Computing Research*. 2010, IEEE, pp. 1–4.

Gevaert, O., De Smet, F., Timmerman, D., Moreau, Y. & De Moor, B. (2006). Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks. *Bioinformatics (Oxford, England)*. 22 (14). pp. e184-90.

Giri, P. & Saravanakumar, K. (2017). Breast Cancer Detection using Image Processing Techniques. *Oriental journal of computer science and technology*. 10 (2). pp. 391–399.

Huang, G.-B., Ding, X. & Zhou, H. (2010). Optimization method based extreme learning machine for classification. *Neurocomputing*. 74 (1–3). pp. 155–163.

Ibrahim, A.M., Baharudin, B., Md Said, A. & Hashimah, P.N. (2016). Classification of breast tumor in mammogram images using unsupervised feature learning. *American Journal of Applied Sciences*. 13 (5).

pp. 552–561.

Ikeda, D.M., Baker, D.R. & Daniel, B.L. (2000). Magnetic resonance imaging of breast cancer: Clinical indications and breast MRI reporting system. *Journal of Magnetic Resonance Imaging*. 12 (6). pp. 975–983.

Iswardani, A. & Hidayat, W. (2018). Mammographic Image Enhancement using Digital Image Processing Technique. *Electrical Engineering and Systems Science*. 16 (5). pp. 1–5.

Karthiga, B. & Rekha, M. (2020). Feature extraction and I-NB classification of CT images for early lung cancer detection. *Materials Today: Proceedings*.

Lateef, F. & Ruichek, Y. (2019). Survey on semantic segmentation using deep learning techniques. *Neurocomputing*.

Li, H., Meng, X., Wang, T., Tang, Y. & Yin, Y. (2017). Breast masses in mammography classification with local contour features. *BioMedical Engineering OnLine*. 16 (1). pp. 44.

Li, Q. & Nishikawa, R.M. (2015). *Computer-Aided Detection and Diagnosis in Medical Imaging*. 1st Ed. CRC Press.

Lima, S.M.L. de, Silva-Filho, A.G. da & Santos, W.P. dos (2016). Detection and classification of masses in mammographic images in a multi-kernel approach. *Computer Methods and Programs in Biomedicine*. 134. pp. 11–29.

Mert, A., Kılıç, N., Bilgili, E. & Akan, A. (2015). Breast Cancer Detection with Reduced Feature Set. *Computational and Mathematical Methods in Medicine*. 2015. pp. 1–11.

Mesleh, A.M. (2017). Lung Cancer Detection Using Multi-Layer Neural Networks with Independent Component Analysis: A Comparative Study of Training Algorithms. *Jordan Journal of Biological Sciences*. 10 (4).

NCBI (2020). *National Center for Biotechnology Information*. 2020.

Pandey, S. (2010). *Study and implementation of morphology for image segmentation*. Thapar University.

Park, K., Ali, A., Kim, D., An, Y., Kim, M. & Shin, H. (2013). Robust predictive model for evaluating breast cancer survivability. *Engineering Applications of Artificial Intelligence*. 26 (9). pp. 2194–2205.

Rankin, S.C. (2000). MRI of the breast. *The British Journal of Radiology*. 73 (872). pp. 806–818.

Rouhi, R., Jafari, M., Kasaei, S. & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*. 42 (3). pp. 990–1002.

Samah, A.A., Fauzi, M.F.A. & Mansor, S. (2017). Classification of benign and malignant tumors in histopathology images. In: *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. September 2017, IEEE, pp. 102–106.

Sarno, A., Mettivier, G. & Russo, P. (2015). Dedicated breast computed tomography: Basic aspects. *Medical Physics*. 42 (6Part1). pp. 2786–2804.

Information Sciences. 187. pp. 277–290.

Sehgal, C.M., Weinstein, S.P., Arger, P.H. & Conant, E.F. (2006). A Review of Breast Ultrasound. *Journal of Mammary Gland Biology and Neoplasia*. 11 (2). pp. 113–123.

Sharma, M., Dubey, R., Sujata & Gupta, S.K. (2012). *Feature Extraction of Mammograms*.

Singh, N. & Asuntha, A. (2016). Image Processing used for Lung Cancer Detection in Medical Imaging. *Journal of Chemical and Pharmaceutical Research*. 8 (4). pp. 1044–1049.

Tamura, H., Mori, S. & Yamawaki, T. (1978). Textural Features Corresponding to Visual Perception. *IEEE Transactions on Systems, Man, and Cybernetics*. 8 (6). pp. 460–473.

Tomar, R.S., Singh, T., Wadhwani, S. & Bhadoria, S.S. (2009). Analysis of Breast Cancer Using Image Processing Techniques. In: *2009 Third UKSim European Symposium on Computer Modeling and Simulation*. 2009, IEEE, pp. 251–256.

Valarmathie, P., Sivakrithika, V. & Dinakaran, K. (2016). Classification of mammogram masses using selected texture, shape and margin features with multilayer perceptron classifier. *Biomedical Research (India)*. (2). pp. S310–S313.

Veldkamp, W.J.H. & Karssemeijer, N. (2000). Normalization of local contrast in mammograms. *IEEE Transactions on Medical Imaging*. 19 (7). pp. 731–738.

Xian, M., Zhang, Y. & Cheng, H.D. (2015). Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains. *Pattern Recognition*. 48 (2). pp. 485–497.

Yang, S.-N., Li, F.-J., Liao, Y.-H., Chen, Y.-S., Shen, W.-C. & Huang, T.-C. (2015). Identification of Breast Cancer Using Integrated Information from MRI and Mammography F. Santanelli (ed.). *PLOS ONE*. 10 (6). pp. e0128404.

Zhou, Z., Qutaish, M., Han, Z., Schur, R.M., Liu, Y., Wilson, D.L. & Lu, Z.-R. (2015). MRI detection of breast cancer micrometastases with a fibronectin-targeting contrast agent. *Nature Communications*. 6 (1). pp. 7984.