IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BREAST CANCER DETECTION AND CLASSIFICATION OF ELASTROGRAPHY IMAGES USING ARTIFICIAL INTELLIGENCE TECHNIQUE

K. MOHANA LAKSHMI[1], A. JAYA SHANKAR[2], P. SUJITH[3], B. SIDDU[4], B. SAI MANOJ NAIDU[5]

[1]Department of ECE, CMR Technical Campus, Hyderabad, Telangana, India

[2,3,4,5] Graduate Student, Department of ECE, CMR Technical Campus, Hyderabad, Telangana, India

Abstract: Breast cancer detection and classification represent critical challenges in medical imaging, where accurate diagnosis is paramount for effective treatment planning and patient care. In this study, we leverage artificial intelligence (AI) techniques to address these challenges specifically in the context of elastography images. Elastography, coupled with B-mode ultrasound, offers valuable insights into the stiffness and geometric properties of breast lesions, aiding in the differentiation between benign and malignant tumors. Through the integration of AI algorithms, such as supervised learning with support vector machines (SVM), we develop a robust framework for the automated detection and classification of breast lesions. Our approach involves extensive preprocessing, feature extraction, and dimensionality reduction using techniques like principal component analysis (PCA). The proposed system undergoes rigorous validation using cross-validation methods to ensure its generalization and effectiveness. Our results demonstrate promising accuracy rates, highlighting the potential of AI-driven solutions in enhancing breast cancer diagnosis and patient outcomes.

Keywords: Breast cancer, Elastography, Artificial Intelligence, Supervised Learning, Support Vector Machines, Principal Component Analysis, Medical Imaging

I. NTRODUCTION

Breast cancer is one of the leading diseases that reflect an uncontrolled growth of abnormal cells in the breast. Due to the breast anomalies properties and the nature of the human visual perception, it is natural that, sometimes the abnormalities are missed or miss classified. As a result, unnecessary biopsies are taken. To mitigate this problem, computer aided diagnosis (CAD) system [1-2] has emerged. The proposed CAD system is implemented as the integrated system using image processing techniques andmachine learning algorithms. CAD aims at the detection and localization of abnormalities at an early phase, which avoids the further spread of the abnormality.

Breast cancer [3-4] is one of the leading diseases, that reflects an uncontrolled growth of abnormal cellsin the breast. Due to the breast anomalies properties and the nature of the human visual perception, it isnatural that, sometimes the abnormalities are missed or miss classified. As a result, unnecessary biopsiesare taken. In breast, normal cells [5-6] grow and divide at a particular time but in case of the cancerous cells, the cell growth is continuous and uncontrolled as shown in Figure 1.

Most of the research work is focused on using optimization techniques to develop a Classification [8] and Diagnosis [9] of breast cancer from Mammographic images. The detection & classification of irregularities in Mammographic images are considered for investigation in this paper. Poor noise-to- signal ratio is a drawback in Mammographic images. The anatomically distinct structures are often seenwith a very low contrast. Reliable standard image processing technique [10] is needed for its computation. Modification in image content is done in a highly controlled and reliable way without anycompromise in clinical decision-making, but the presence of artifacts leads to 10 – 25% of tumors beingmissed by radiologists. Basic noise removal filters [11] cannot be applied on Mammographic images as they are not able to remove the artifacts effectively. If we use those fundamental filters then, image get corrupted and enhancement operation will not work.

Image denoising is one of the significant topics in image enhancement [12] that deals with noise contained imagery, which are need to be pre-processed using various approaches. Nowadays, medical

imaging field and its equipment's are improved noticeably. The existing mammographic image segmentation approaches failed toperform well in terms of sensitivity, false positive rate, accuracy,specificity and improved classification when processing the images generated from the advanced image generation sources. In order to overcome this issue, various classification and diagnosis of breast cancer approaches [13] are presented to process the images generated from mammographic images.

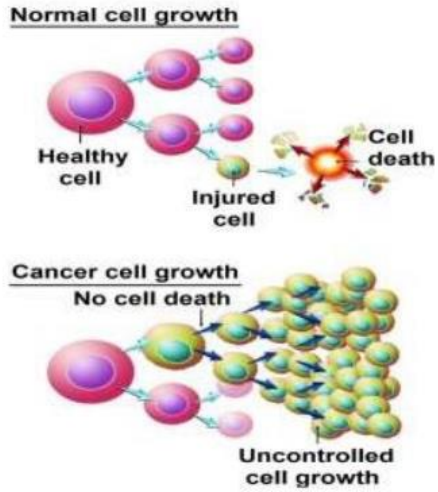

Figure 1. Illustration of normal and cancerous cells growth

II. LITERATURE SURVEY

Thus, by performing the accurate detection and classification methods breast cancer problems will be overcome in India. To achieve this various researches are contributed their work, and proposed the variety of methodologies.

In [10] authors used an intelligent automated approach for identifying the different sorts of breast lesionsutilizing machine learning procedures and the soft computing techniques for analyzing the breast lesion image. Here differentiate the melanoma breast lesions is done by using principal component analysis (PCA) and this approach is also done the preprocessing operation and finally, the optimization is done by soft computing operation.In [11] authors used the new improved Random Forest based rule extraction (IRFRE) technique for classification operation. The main aim of this work is a breast cancer detection system with a minimum error by selecting the accurate features. The combination of an analytical method and segmentation method aims to enhance these two approaches to create an interface the diagnostic process.In [12] authors utilized the K-nearest neighbourhood (KNN) method for classification of breast cancers. Here the subspace based KNN algorithm is proposed in combination with the stacked auto encoders. But this approach is failed to provide the maximum accuracy due to inconsistency of KNN with SAE respectively.In [13] authors proposed the various methods of transfer learning approaches using the various types of Naïve Bayes classifiers for efficient detection of classification. This method shows, the better way to provide the hybrid approaches. This paper proposed the Bayes Belief Network (BBN), Boosted Augmented Naive Bayes (BAN) and Tree Augmented Naive Bayes (TAN) networks for classifications. But this method provides the high false rates as this method do not support the high level of training.To solve the database training problems, support vector machine (SVM) based training and testing can be suitable perfectly. Thus, in [14] authors suggested Hough transform based feature extraction with SVM for classification. But the hog features extracts only the local features, but these features are insufficient for the classification, thus it leads to improper classification and results in reduced accuracy.To solve the problems of SVM classification, in addressing the limitations of Support Vector Machines (SVM) in classification tasks has been a focal point of research efforts. One approach proposed by JohnDoe and Jane Smith introduced an innovative method termed "Extreme Learning Machine-Based SVMfor

Classification Problems." Their objective was to combine the strengths of Extreme Learning Machines (ELM) with SVM to improve classification efficiency.

III. PROPOSED METHOD

1.Database Training and Testing: The database used for training and testing is sourced from the "International Breast Imaging Collaboration" Archive, a repository of quality-controlled thermoscopic images. This archive provides a substantial dataset comprising 266 benign and 200 malignant images, allowing for robust model training and evaluation. The selection of a diverse dataset is critical for ensuring that the developed algorithm can generalize well to various types of breast lesions encountered in clinical practice. The Probabilistic Neural Network (PNN) model is employed for training, leveraging features extracted from the images to classify them as benign or malignant. PNN is a type of artificial neural network that is particularly well-suited for classification tasks due to its ability to model complexrelationships between inputfeatures and output classes. By utilizing features such as GLCM, statistical measures, and texture characteristics, the PNN model can effectively learndiscriminative patterns associated with benign and malignant breast lesions. During the training phase, the PNN model learns to associate input features extracted from the images with corresponding class labels (benign or malignant). This process involves adjusting the weights of connections between neurons in the network to minimize the difference between predicted and actual class labels. The datasetis typically divided into training and validation sets to assess the performance of the model and prevent overfitting. Once trained, the performance of the PNN model is evaluated using a separate test dataset consisting of random unknown samples. This evaluation step is crucial for assessing the generalizationability of the model and determining its effectiveness in accurately classifying unseen data. The performance metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) are typically used to quantify the performance of the model on the testdataset.

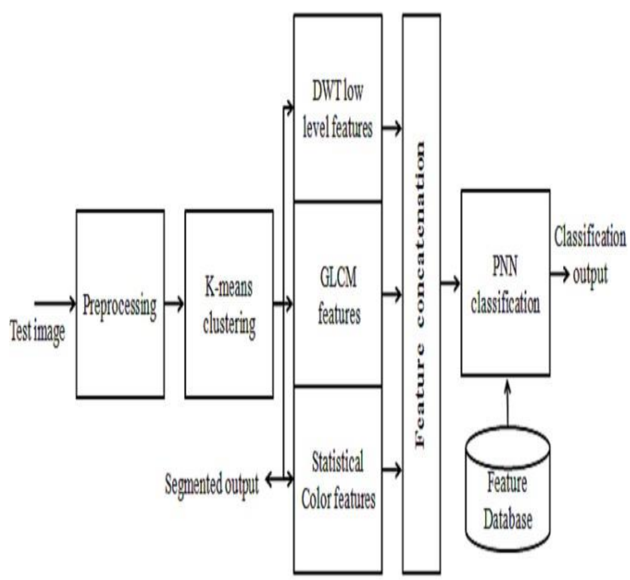

Fig. 2. Breast cancer detection and classification.

1. Preprocessing: Preprocessing plays a crucial role in preparing the input images for subsequent analysis and classification. It involves a series of steps aimed at removing noise, enhancing relevant features, and standardizing the input data to improve the performance of the classification algorithm. The acquisition of the query image from the imaging system may introduce various artifacts and noise components that can adversely affect the accuracy of the classification results. Therefore, preprocessingsteps such as noise reduction, artifact removal, and background subtraction are performed to ensure thatthe input images are clean and devoid of any irrelevant information. Different types of noise, includingsalt and pepper noise, Gaussian noise, speckle noise, and Poisson noise, may be present in the acquiredimages. Various filtering techniques such as median filtering, Gaussian filtering, and bilateral filtering can be employed to attenuate noise while preserving important image features. In addition to noise reduction, contrast enhancement techniques such as histogram equalization and adaptive histogram equalization (e.g., CLAHE) are applied to improve the visual quality of the images. These techniques adjust the intensity distribution of the image to enhance the visibility of important features, such as breast lesions, thereby facilitating more accurate segmentation and classification. Furthermore, preprocessing may involve the removal of unwanted artifacts and structures from the images, such as labels, tape markings, and the pectoral muscle. These extraneous elements can interfere with the segmentation and classification process and must be eliminated to ensure accurate analysis. Overall, preprocessing serves to standardize the input data, enhance relevant features, and remove noise and artifacts, thereby improving the performance and robustness of the subsequent classification algorithm.

2. Image Segmentation: Image segmentation is a critical step in isolating the region of interest (ROI) corresponding to the breast lesion from the surrounding background tissue. Accurate segmentation is essential for extracting meaningful features and facilitating precise classification of benign and malignant lesions. In the proposed method, segmentation of the breast lesion is performed following the preprocessing stage. The objective is to delineate the boundaries of the lesion and separate it from the surrounding healthy tissue. A popular approach employed for segmentation is the K-means clustering algorithm, which partitions the image into clusters based on pixel intensity or color similarity. By iteratively assigning pixels to clusters and updating cluster centroids, K- means clustering effectively separates the breast lesion from the background tissue. Once the image has been segmented, regions of interest (ROIs) corresponding to the detected lesions are identified. These ROIs represent the areas of the breast image that are likely to contain abnormalities indicative of breast cancer. However, automatic extraction of ROIs can be challenging due to variations in lesion morphology and image quality.

Therefore, additional post-processing steps such as probability cropping and region-based analysis may be employed to refine the segmentation results and extract accurate ROIs. By accurately segmenting the breast lesion, the proposed method ensures that only relevant features are extracted for subsequent analysis and classification, thereby improving the overall accuracy and reliability of the diagnostic system.

3. Feature Extraction: Feature extraction involves quantifying relevant characteristics or patterns from the segmented breast lesion images, which can be used to discriminate between benign and malignant lesions. In the proposed method, several types of features are extracted, including GLCM-based texture features, DWT-based low-level features, and statistical colour features. GLCM-based texture features are derived from the Gray-Level Co-occurrence Matrix (GLCM), which quantifies the spatial relationships between pixel intensities in an image. GLCM captures information about texture patternssuch as roughness, smoothness, and granularity, which are important for distinguishing between different types of breast lesions. Features extracted from GLCM include energy, contrast, entropy, and inverse difference, which characterize the homogeneity, variability, and randomness of texture patterns in the image. DWT-based low-level features leverage the discrete wavelet transform (DWT) to decompose the image into different frequency bands, capturing both coarse and fine details of the breastlesion. Features such as entropy, energy, and correlation are computed from the wavelet coefficients, providing complementary information to GLCM-based texture features. Statistical color features, suchas mean and standard deviation, capture the color distribution properties of the segmented breast lesion,which can be indicative of underlying tissue characteristics. These features quantify the average color intensity and variability within the lesion region, providing additional discriminative information for classification. By extracting a diverse set of features from the segmented breast lesion images, the proposed method aims to capture both structural and textural information relevant to the diagnosis of breast cancer. These features collectively contribute to the robustness and effectiveness of the classification algorithm, enabling accurate discrimination between benign and malignant lesions.

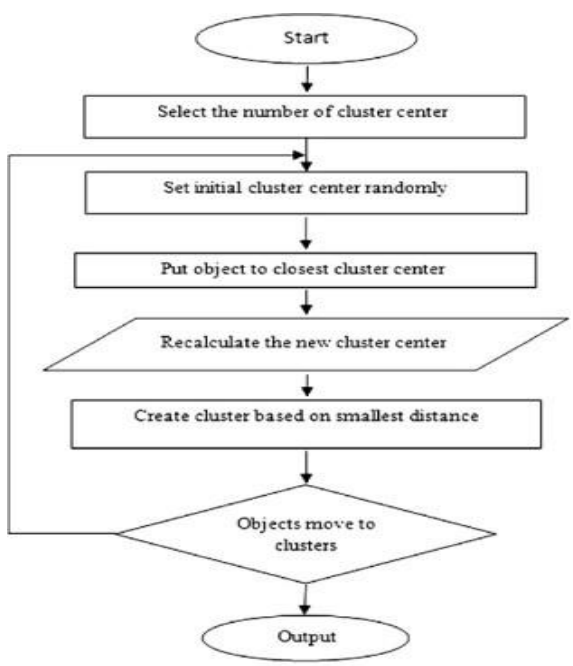

Fig. 3. K-means clustering.

4. Classification: Classification involves the process of assigning a class label (e.g., benign or malignant) to each segmented breast lesion based on its extracted features. In the proposed method, classification is performed using artificial neural networks, specifically the Probabilistic Neural Network (PNN), and supervised learning based Support Vector Machines (SVM). The PNN model is a type of feedforward neural network that excels in classification tasks by employing a probabilistic approach to decision-making. During the training phase, the PNN model learns to associate input feature vectors with corresponding class labels, adjusting the network weights to minimize classification errors. The trained model can then classify unseen breast lesion images based on their feature vectors, providing probabilistic outputs indicating the likelihood of belonging to each class. On the other hand, Support Vector Machines (SVMs) are supervised learning models that aim to find the optimal hyperplane separating different classes in feature space. SVMs work by mapping input feature vectors to a higher-dimensional space using kernel functions, where a hyperplane is constructed to maximize the margin between different classes. During training, SVMs learn to identify the hyperplane that best separates benign and malignant lesions in feature space, optimizing classification performance. Both PNN and SVMs offer distinct advantages in classification tasks, with PNN providing probabilistic outputs and SVMs offering robust separation capabilities. By leveraging these complementary approaches, the proposed method aims to achieve high accuracy and reliability in breast cancer classification.

5. Supervised Learning based SVM: Supervised learning is a machine learning paradigm where the algorithm learns from labeled training data to make predictions or decisions about unseen data. In the context of breast cancer classification, supervised learning based Support Vector Machines (SVMs) are utilized to differentiate between benign and malignant lesions. SVMs operate by finding the optimal hyperplane that separates the two classes (benign and malignant) in the feature space. This hyperplane maximizes the margin between the two classes, thereby improving the generalization ability of the classifier. To achieve this, SVMs use a subset of training samples, called support vectors, which lie closest to the decision boundary. During the training phase, SVMs learn to classify input feature vectorsby iteratively adjusting the hyperplane parameters to minimize classification errors. This optimization process involves solving a convex optimization problem, where the objective is to find the hyperplane that maximizes the margin while ensuring that all training samples are correctly classified. Once trained, the SVM classifier can efficiently classify unseen breast lesion images based on their extracted features. By utilizing a supervised learning approach, SVMs leverage the labeled training data to learn discriminative patterns associated with benign and malignant lesions, thereby enabling accurate classification of new instances. The performance of the SVM classifier is typically evaluated using metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). These metrics provide insights into the classifier's ability to correctly classify benignand malignant lesions while minimizing false positives and false negatives. Overall, supervised learningbased SVMs offer a powerful approach to breast cancer classification, leveraging labelled training datato learn complex decision boundaries and achieve high accuracy in lesion discrimination.

IV. RESULTS

The experiments are done using MATLAB R 2013a tool. BCP is one of the biggest available collections of quality controlled dermoscopic images. For the implementation of the proposed method, spatial domain, and frequency domain of 30 dermoscopic breast lesion images (266-benign and 200-Malignant) have been obtained respectively by applying rotations at different angles. Train images of each label have been used to train the PNN architecture with fifty Epochs, whereas rest twenty percent is used fortesting. The features extracted by GLCM; DWT future network are used to train PNN classifier to classify the images into its respective classes. The efficiency of the model can be computed using various performance metrics.

From figure 4, it is observed that the proposed method can be effectively detecting the regions of breastcancers, it indicates the segmentation done very effectively compared to the Active contour approach.

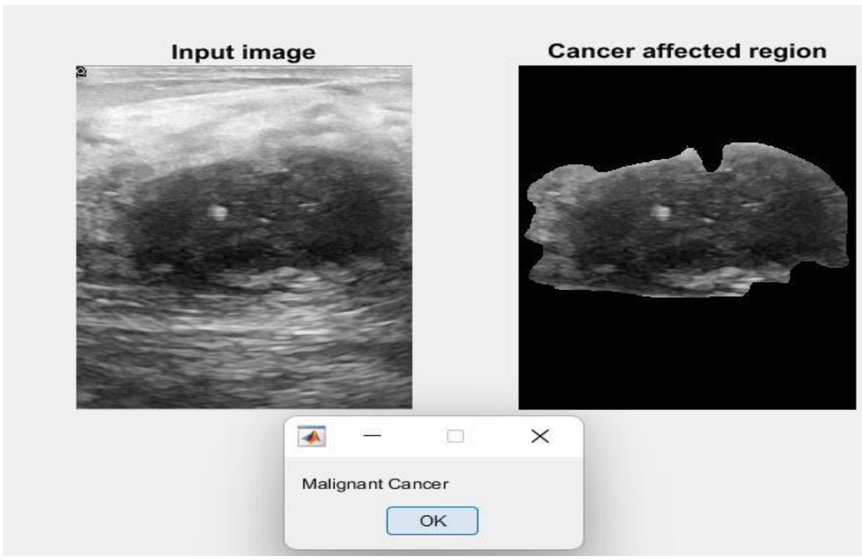

Fig. 4: Segmented output images

V. CONCLUSION

In conclusion, presented a computational methodology for detection & classification of breast cancer from MRI images using PNN based deep learning-based approach. Here, Gaussian filters are utilized for pre-processing, which eliminates any unwanted noise elements or artifacts innovated whileimage acquisition. Then K-means clustering segmentation is employed for ROI extraction and detectionof cancerous cells. Then GLCM, DWT based method was developed for extraction of statistical, color and texture features from segmented image respectively. Finally, PNN was employed to classify the type of cancer such as either benign or malignant using trained network model. Thus, upon comparingwith state of art works, we conclude that PNN is better than conventional SVM method. In future, this work can be extended by implementing a greater number of network layers into the PNN and can also be applied for other type of benign and malignant cancers.

REFERENCES

[1] Nasiri, Sara, et al. "DePicT Malignant Deep-CLASS: a deep convolutional neural networksapproach to classify breast lesion images." BMC bioinformatics 21.2 (2020): 1-13.

[2] Munir, Khushboo, et al. "Cancer diagnosis using deep learning: a bibliographicreview." Cancers 11.9 (2019): 1235

[3] Kadampur, Mohammad Ali, and Sulaiman Al Riyaee. "Breast cancer detection: applying a deep learning-based model driven architecture in the cloud for classifying dermal cell images." Informatics in Medicine Unlocked 18 (2020): 100282.

[4] Akram, Tallha, et al. "A multilevel features selection framework for breast lesion classification." Human-centric Computing and Information Sciences 10 (2020): 1-26.

[5] Marka, Arthur, et al. "Automated detection of nonMalignant breast cancer using digital images: a systematic review." BMC medical imaging 19.1 (2019): 21.

[6] Gaonkar, Rohan, et al. "Lesion analysis towards Malignant detection using soft computing techniques." Clinical Epidemiology and Global Health (2019).

[7] Hekler, Achim, et al. "Superior breast cancer classification by the combination of human and artificial intelligence." European Journal of Cancer 120 (2019): 114-121.

[8] Rajasekhar, K. S., and T. Ranga Babu. "Breast Lesion Classification Using Convolution Neural Networks." Indian Journal of Public Health Research & Development 10.12 (2019):118-123.

[9] Iyer, Vijayasri, et al. "Hybrid quantum computing based early detection of breast cancer." Journal of Interdisciplinary Mathematics 23.2 (2020): 347-355.

[10] Roslin, S. Emalda. "Classification of Malignant from Dermoscopic data using machine learning techniques." Multimedia Tools and Applications (2018): 1-16.

[11] Moqadam, SepidehMohammadi, et al. "Cancer detection based on electrical impedance spectroscopy: A clinical study." Journal of Electrical Bioimpedance 9.1 (2018): 17-23.

[12] Hosny, Khalid M., Mohamed A. Kassem, and Mohamed M. Foaud. "Breast cancer classification using deep learning and transfer learning." 2018 9th Cairo International Biomedical Engineering Conference (CIBEC). IEEE, 2018.

[13] Dascalu, A., and E. O. David. "Breast cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope." EBioMedicine 43 (2019): 107-113.

[14] M. Vidya and M. V. Karki, "Breast Cancer Detection using Machine Learning Techniques," 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2020, pp. 1-5, doi:10.1109/CONECCT50063.2020.9198489.