Adaptive Gaussian Filtering (AGF) and Segmentation using Deep Neural Network (DNN) for Breast Tumour Classification

[1]Nateshia.G, Research scholar, Department of Computer Science and Applications, Providence college for women,coonoor.
Email –id: nateshia89@gmail.com
[2]K.Deepa M.Sc., M.Phil., (Ph.D), Assistant Professor, Department of computer science and applications, Providence College for women,coonoor.

ABSTRACT: Breast cancer is one of the major causes of death among women. An improvement of early diagnostic techniques is critical for women's quality of life. Mammography is the main test used for screening and early diagnosis. Many researchers worked in the area of breast cancer detection and proposed segmentation methods. However, no solution given by researchers is best promising and has limitations and it is still a challenging problem because of noises presented in the samples. So the early diagnosis of breast cancer plays critical roles in increasing the chance of tumour segmentation results. A few recent works also investigated the digital implementations of Cellular Neural Network (CeNN), they are still computationally redundant and sensitive to input image variations without fully utilizing the advantages of hardware acceleration. This research addresses the problem of segmenting the breast cancer from the Mammography images. Proposed an Adaptive Gaussian Filtering (AGF) algorithm in which the filter variance is adapted to both the noise characteristics and the local variance of the signal. Using an AGF for noise suppression, the noise is smoothed out; at the same time the signal is also distorted. The use of an AGF is proposed as pre-processing for noise removal. Deep Neural Network (DNN) based approach is presented for segmentation. The approach is featured with quantization to significantly reduce the computational complexity and non-linear template for robustness. Compared with other prior works, the proposed work is able to reduce error value and increase segmentation results.

Index terms: Adaptive Gaussian Filtering (AGF), Noise Removal, classification, tumour segmentation, Deep Neural Network (DNN), Cellular Neural Network (CeNN), segmentation.

1. INTRODUCTION

Breast tumour segmentation is most significant stages in breast cancer recognition and analysis. Many techniques have been anticipated to progress segmentation efficiency (Zhang et al 2012). One generally utilized technique is region growing, which segments masses grouping related neighbouring pixels of seed points and is capable to attain promising precision (Wei et al 2012). Nevertheless, it is computationally intensive and therefore incompetent for high-resolution mammographic images, which attains raising popularity in classical medical imaging. To attain both efficiency and accuracy, Cellular Neural Network (CeNN) based method has been examined (Chua et al 1988). Still, owing to huge amount of multiplications in CeNNs, conventional analogue implementations experience from I/O resource and noise interference (Roska et al 1992). Some current works also examine digital implementations of CeNN, however still computationally outmoded and sensitive to input image variations devoid of completely using benefits of hardware acceleration.

A new digital CeNN-based method to address abovementioned issues in tumour segmentation, with subsequent contributions: (i) Quantisation system to understand multiplications by logic shift, thus diminishing hardware resource overheads; (ii) Non-linear templates deployment by sequence of linear templates to progress accurateness; and (iii) Parallel hardware architecture to progress complete efficiency. It performs widespread experiments on Digital Database for Screening Mammography (DDSM) (Heath et al 1998). The foremost contribution of this investigation is described as trails,

- Anticipated an Adaptive Gaussian Filtering (AGF) algorithm where filter variance is adapted to both noise features and local variance of signal.
- Deep Neural Network (DNN) sourced approach is offered for segmentation.

2. LITERATURE REVIEW

Marcano-Cedeno et al. (2011) proposed another "Artificial Metaplasticity Multilayer Perceptron algorithm" which performed superior to BPNN on a similar breast malignant growth datasets. By evacuating the example with missing quality, remaining 683 example cases are taken for

further examination. The proposed system is prepared by 60% of dataset, that is, 410 examples and testing is performed with remaining 273 examples. This proposed AMMLP calculation has single hidden layer with eight neurons. They utilized two criteria to pause training procedure. The first is that if the mean squared error is drawn closer to 0.01 and the other criteria are by settling training period to 2000. This technique gives classification precision of 99.26% when contrasted with BPNN which has precision 94.51% for 60/40 samples for training and testing.

Senapati et al. (2013) utilized neural system with "Local Linear Wavelet For Detection of Breast Cancer Recognitionas typical or unusual. In this methodology the association weights between input layer and output layer are traded by central linear model. Recursive Least Square technique is utilized to refresh its parameter for preparing to upgrade execution. The nature of result has been evaluated and contrasted and past methodologies. The displayed technique is exceptionally powerful, proficient and accomplished higher classification exactness 97.2%.

Nahato et al (2015) displayed, "Knowledge Mining From Clinical Datasets Using Rough Sets and Backpropagation Neural Network" have structured therapeutic classification replica which incorporates rough set ambiguity connection strategy and backpropagation neural system. This model works in two phases. In initial phase missing qualities is taken care of to get a smooth informational collection and suitable characteristics are chosen from the clinical dataset by indiscernibility relation technique. In second stage arrangement is finished utilizing backpropagation neural system on chosen redacts of dataset. They revealed a characterization precision of 98.60% for breast malignant growth analysis.

Bhardwaj and Tiwari (2015) introduced "Genetically Optimized Neural Network (GONN)" show which all the while advanced structure and weight of neural system to classify breast malignancy database as deadly or typical. This extends mutation and cross over operators to decrease damaging idea of these operators. In this plan, every one of the people left after multiplication are taken for cross over task and people that can't create posterity with better fitness value from their fitness are transformed. The model gets a normal accuracy rate of 99.26% for 10-fold cross validation approach.

Aytug Onan (2015) displayed an intelligent hybrid classification system for breast malignant growth diagnosis which comprises of three stages: instance selection, feature selection and classification. In first stage fuzzy rough set instance selection method dependent on frail gamma evaluator is used to expel useless or incorrect examples. In feature selection stage, consistency-based component choice technique is utilized related to re-ranking calculation, inferable from its productivity in finding the conceivable lists in the inquiry space. In classification of framework, fuzzy rough nearest neighbour calculation is used. Since this classifier does not require ideal incentive for K neighbours and has more extravagant class certainty values, this methodology is used for classification task. They acquired order precision of 99.71%.

Nguyen et al (2015) proposed a medicinal breast disease classification replica on fuzzy standard additivemodel and genetic algorithm. In this plan, adaptive vector quantization clustering handles rule instatement. A genetic calculation is used for standard optimization and parameter tuning is finished utilizing gradient descent calculation. With the end goal to deal with high dimensional datasets, wavelet transform is used. This model gives classification accuracy of 97.40%.

Tomar and Agarwal (2015) gives a "Hybrid Feature Selection (HFS)" based proficient disease diagnostic model for breast malignant growth. A HFS consolidates positive parts of both Filter and Wrapper FS approaches. This model receives weighted least squares twin support vector machine (WLSTSVM) as classification approach, successive forward selection (SFS) as search strategy, and correlation feature selection (CFS) to assess significance of each component. This model chooses related feature subset as well as proficiently manages information unevenness issue reported classification precision of 98.55%.

Chen (2014) planned hybrid intelligent model for recognition of breast disease that can work without labelled training data. Feature extraction strategies in unsupervised learning models, model coordinate clustering and feature selection approach. That choosing a subset of pertinent features as opposed to utilizing every one of features in first informational data set can expand interpretability of clustering results.

3. PROPOSED METHODOLOGY

Proposed an Adaptive Gaussian Filtering (AGF) algorithm in which the filter variance is adapted to both the noise characteristics and the local variance of the signal. Using an AGF for noise suppression, the noise is smoothed out; at the same time the signal is also distorted. The use of an AGF is proposed as pre processing for noise removal. Deep Neural Network (DNN) based approach is presented for segmentation. The approach is featured with quantisation to significantly reduce the computational complexity and non-linear template for robustness. Compared with other prior works, the proposed work is able to reduce error value and increase segmentation results (See figure 1).

3.1.ADAPTIVE GAUSSIAN _FILTERING

Many existing algorithms in image processing assume a scale for feature detection, such as zero-crossings of second derivative, local curvature, corner and edge detection, to name a few. This scale represents a compromise between localization accuracy and sensitivity to noise. Obviously, there is seldom a single scale that is appropriate for a complete image. The local variance calculated must be useful for, at least, two basic and

necessary operations in image processing: noise removal, and facilitating the localisation of features. Rewriting the last statement, we can say that the local blur, as in anisotropic diffusion or adaptive smoothing, should smooth directly within regions and not across discontinuities. More formally, in terms of Bayesian estimation, maximizing the likelihood of the observed image given the local scale, and at the same time, minimizing the residual . To see this more clearly, consider a stack of smoothed images or scale-space, I obtained by convolving the initial image with a set of Gaussian kernels, G σ where $I_\sigma = I_0 * G_\sigma$. Thus, our objective is to choosethe appropriate scale (σ), from the scale-space generated, at each pair x; y. Thecriterion for such selection is the optimal coding, which must be the less complexor minimal, and, at the same time, effective for describing the parameter. In practice, it is straightforward to estimate the optimal code with tools like

the minimal description length principle (MDL) . An image due a Gaussian Filtering can be seen as:

$$(x, y) = I_\sigma(x, y) + \in_\sigma (x \quad) \quad (1)$$

Where the residual, called , is the original image minus the smoothed image. The minimal description length states that the optimal coding deal with the minimum number of bits that are necessary for maximising the information of both parameters. That is, the maximum smoothness with the minimal residual.

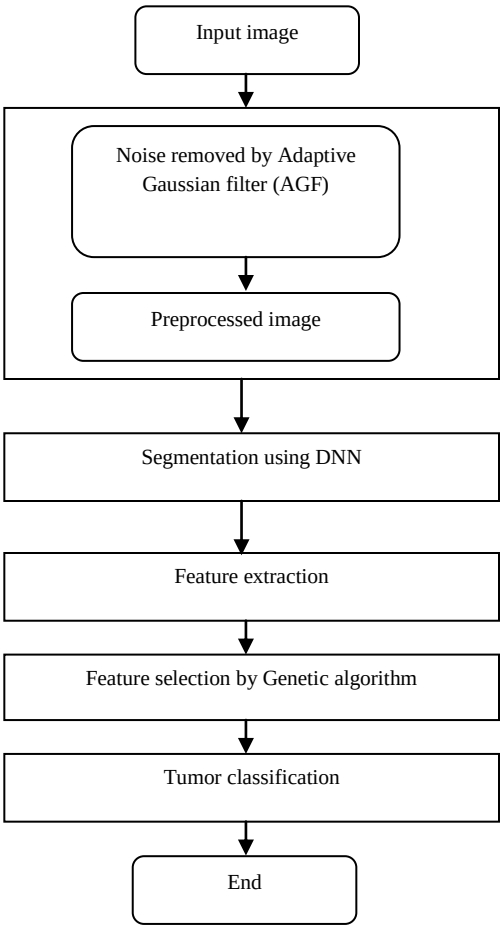


Figure 1.Flow chart for proposed system

Let us rewrite equation (1) as description length (dl):

$$dl_{(X,Y)} = dl_{I\sigma(x,y)} + dl_{\in\sigma(x,y)} \quad (2)$$

3.1.1 DESCRIPTION LENGTH OF

It is not difficult to see that the measure of information, in bits, for I_0 is greaterthan , I , ∞ the measure of information over $I_\sigma$must be minimal. The reason for this fact is that the variance σin the Gaussian filter controls the amount of information in the scale-space.

In order to calculate a fast approximation of the amount of information in ,I

We start with the ideal model of low-pass. Allow us to consider the behaviour of signal amplitude (a) in space (x) and frequency f through the Scaling-Uncertainty Principle:

$$f\{s(ax)\} = \frac{1}{-} s\left(\frac{f}{-}\right); \quad a \neq \quad (3)$$

Then, the Gaussian distribution in space and frequency takes the form:

$$\left(\frac{-}{-}\right) = e^{\left(\frac{-}{-^2}\right)} \quad ; \qquad 0(4)$$

Hence, it is clear that the relation between the spatial domain $\left(G_{\sigma^2}\right)$ is inversely proportional to the frequency domain$\left(G \quad\right)$The relation (2) states that .σ —Further, the sampling theorem states that for any signal of frequency f, the numberof samples s (Nyquist rate) needed for reconstructing accurately the original signals, at least, 2f. As we know, $f \propto \sigma_\omega^2$, that is the frequency f is proportional, given constant , $(\propto)$ to the Gaussian filter band-width, and it is controlled by its varianceσ_ω^2. Thus, our sampling rate could be rewritten as

$$\eta(\alpha\sigma_\omega^2) \quad ; \eta \geq 2 \qquad (5)$$

Then, as a consequence of the inverse relation (4), we present the next equation, showing that the amount of samples needed for describing a smoothed image is controlled by the spatial variance of the Gaussian kernel. That is:

$$S = \eta\left(\frac{}{-^2}\right) \; ; \qquad (6)$$

Although we cannot know in advance how many bits represents each s, we effectively see that it is proportional $(s$ to the amount of information, given a constant β Whereβ means the precision, in bits, used to represent I . Summaries the above relations in the following equation:

$$= \eta\left(\frac{}{-^2}\right); \qquad 2(4.7)$$

Equation (6) estimates the description length, in bits, of a smoothed image given the spatial variance of the Gaussian filter. In addition, equation (7) shows that correct description length of $I_\sigma$should be computed in a range. The

fact that λisunknown, does not, of course, means that is not possible to compute.

Even if we do not know the noise distribution, it will approximate to Gaussian distribution by the Central Limit Theorem (zero-mean Gaussian density). Thus, the probability distribution of noise can be written as:

Pr (ϵ (8)

Where σ_ϵ^2 " is the variance of the noise, and ϵ means the local quadratic residual1between the original image,I_0 and the filtered image.I_σAs works, the measure of information is close related to the probability. Compute the measure of information in bits of equation (8) applying logarithm (log2). Thus, the description length of the residual takes the forms of:

= k$\left(\frac{}{_2}\right)$ (9)

Where k= —. Once a $ld_{I\sigma} and ld_{\epsilon\sigma}$ have been denoted, we may, therefore, write the local description length as the sum of both terms, (8) and (9):

$\eta\left(\frac{}{_2}\right)$ + k$\left(\frac{}{_2}\right)$(10.)

——————————(11)

$\left(\frac{}{_2}\right)$ +ϵ^2 (12)

Where ———. It is positive since that 2 from equation (12), and the other proportional variables are also positives. It is convenient to distinguish inequation (12), that λ represents, principally, the assumed noise variance σ_ϵ^2 " andprecision used to represent $.I_\sigma$ If we empirically fix the precision, we can recognize the range where λ becomes useful.

Note that in the definition of equation (12) does not appear an edge stopping function nor gradient threshold. The minimal and maximal amount of smooth-ness are controlled by the $\sigma_{min} and\ \sigma_{max}$in the scale space. Moreover, applyingequation (12) the filter adapts its shape (σ) in the next way: if the filter is close to the discontinuities, the MDL selects a ; σ otherwise, if the filter islocated in a uniform structure, the MDL selects a variance close to $.\sigma_{max}$ Noticethat, the Gaussian shape (σ)changes as a function of MDL. The reason this work is that in the neighbourhood of a discontinuity the absolute value of the residual ϵ^2 grows much faster with respect t σthan in a relatively uniform region of animage.

As a result, the local scale will be small close to the discontinuities and large over the uniform regions, which permits effective noise elimination andan accurate discontinuity preservation at the same time.

Once equation (12) is denied, one may use it to compute directly the local amount of smoothing in terms of variance, and, in consequence, perform a Gaussian filtering. To be specific, compute the local description length at each pair x; y through2. I Then, based in the MDL principle, we choose the minimal value ofdl at each location x; y. This min (dl returns the optimal smoothnes(σ y. This value represents the maximum smoothing, and, at the same time,theminimum residual

At this step, we have the local variance $\left(\sigma^*(x,y)\right)$ which is by itself, the appropriate smoothing for regions and not acrossDiscontinuitiesFinally, the output smoothed image in location x; y is the intensity $(x$)atthe selected scale $\sigma^*(x,y)$.

The local variance $\sigma$of the Gaussian at x; y, allows us the possibility of makingan adaptive Gaussian filter. Each point x; y is blurring as $\sigma^*(x$)That is,

$_{(x,y)}=I$ $\left(-\frac{}{*(x\ )}\right)$(13)
$_0\,(x$

Breast tissue segmentation with DNN

The ability to process large data-sets allowed DNNs to produce state-of-the-art results when applied to computer vision, text understanding and speech recognition benchmarks.

DNNs can provide discriminative image representations, sometimes referred to as features, by successive application of linear filters, non-linear activation functions, normalization, and pooling operations, thus, avoiding the need to design such features manually.

The proposed DNN classifier, applied to raw image pixels, provides the probability of each pixel to belong to one of the four classes described above: Pr(class(p) = k), k = 0, 1, 2, 3. The DNN is applied in a patch-wise manner: to classify the pixel p,the DNN is fed with a square image patch of size w × w, centre data. The patches are pre-processed prior to training and classification, to have a zero mean, by subtracting from them the mean of all patches in the training set. The segmentation accuracy is acquired by multinomial logistic loss function.

The motivation for using segmentation is twofold. In medical imaging applications in general, data, manually annotated by experts, is rarely available. This is in contrast to general computer vision tasks, such as natural image classification, segmentation and object boundary detection, for which there exist data-sets with thousands and even millions of annotated.

Since a typical DNN has between tens of thousands to millions of parameters, it requires a large amount of manually annotated example properly train all the parameters and avoid over fitting to an insufficient training data. Therefore, by working with separate, possibly overlapping image patches, we can obtain a training set large enough to overcome these limitations – in our experiments, used training sets of approximately 8 · 105 training examples. In addition, expect that large enough patches would capture a sufficient part of the local information required to correctly classify pixels belonging to different breast tissues.

By applying the DNN to separate image patches, loosethe spatial dependency between neighbouring patch labels, as well as information about the spatial pixel location. Clearly ,Pr(class(p)) and Pr(class(q)) are dependent for neighbouring patches p and q.

In contrast, the spatial pixel location information is easier to include into the proposed network, for instance, in the form ofx and y pixel coordinates, given that the images are identically aligned, similarly .To normalise the location information across images of breasts of different sizes, the x pixel coordinate is divided by the x coordinate of the rightmost image pixel which does not belong to the background. The coordinate is normalised to the range [0,1]. The coordinates can then serve as an additional input for the proposed network, and added after the first fully connected layer by concatenating themto the first fully connected layer output, and then transferring the results to the next fully connected layer. While this simple approach allows us to incorporate the spatial information in learning scheme, it produces equivalent or inferior results than only intensity-based classification – see the results in Table 3. Thus, more research is required to devise a correct way to use the spatial information.

DNN architecture

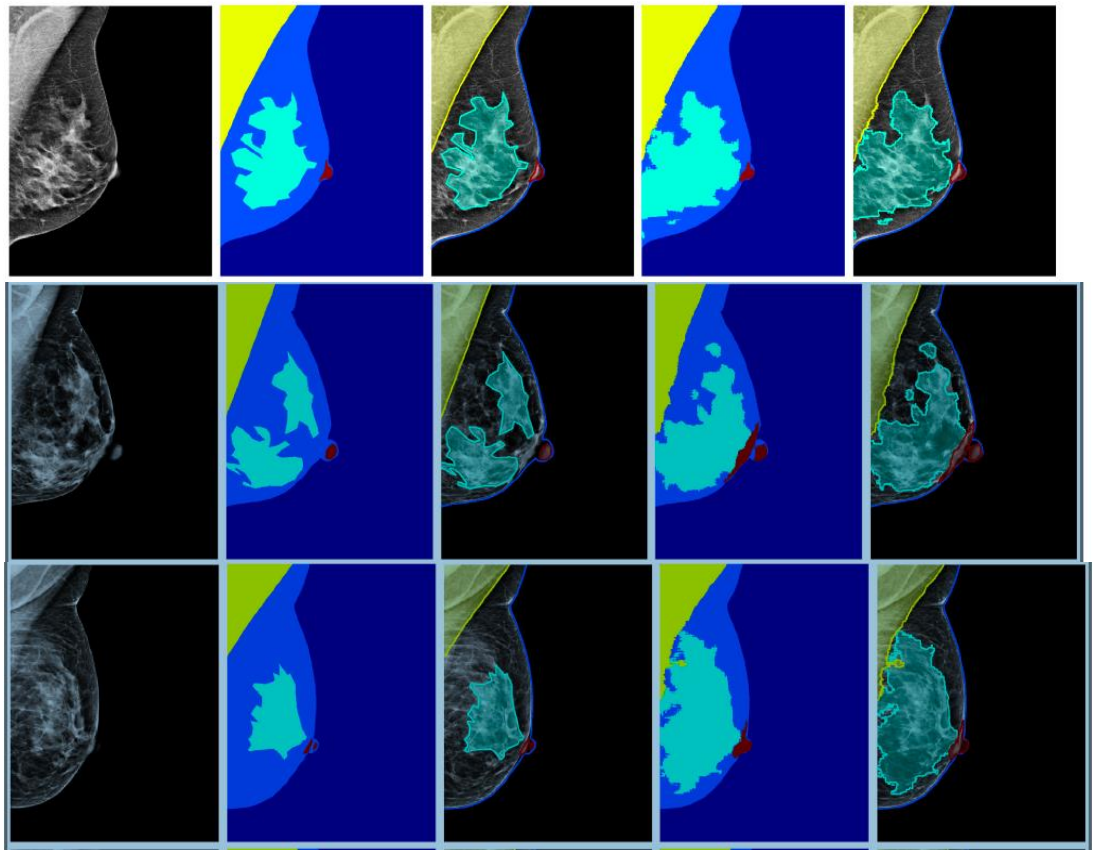
The architecture of the proposed network is summarized in Table 1. It consists of three stages of convolutional layers, ReLU(Rectified Linear Unit) activation layers, and max pooling layers, followed by three fully connected layers. To prevent overfitting, a dropout layer) with dropout factor of 0.5 was added between the convolutional and the fully connected layers. The image intensity and the normalize coordinate information can be combined after the first fully connected layer in the following manner: the two normalized patch centre coordinates and the 128-dimensional fifth layer output are concatenated into a 130-dimensional vector, passed to the second fully connected layer.

Layer	1	2	3	4	5	6	7 – Output
Stage	conv+relu+max	conv+relu+max	conv+relu+max	dropout	full+relu	full+relu	full
# Channels	16	16	16	16	128	16	4
Filter size	7 x 7	5 x 5	5 x 5	-	-	-	-
Pooling size	3 x 3	3 x 3	3 x 3	-	-	-	-
Pooling stride	2	2	2	-	-	-	-
Dropout factor	-	-	-	0.5	-	-	-
Spatial input size	61 x 61	27 x 27	11 x 11	3 x 3	3 x 3	1 x 1	1 x 1

Table 1.Initial deep neural network architecture

Segmentation results

The DNN was trained by stochastic gradient descend with momentum used mini batches of 256 image patches, and learning rate of 10−3, reduced by a factor of 10twice, after 30,000 and 60,000 training iterations (approximately 10 and 20 epochs).



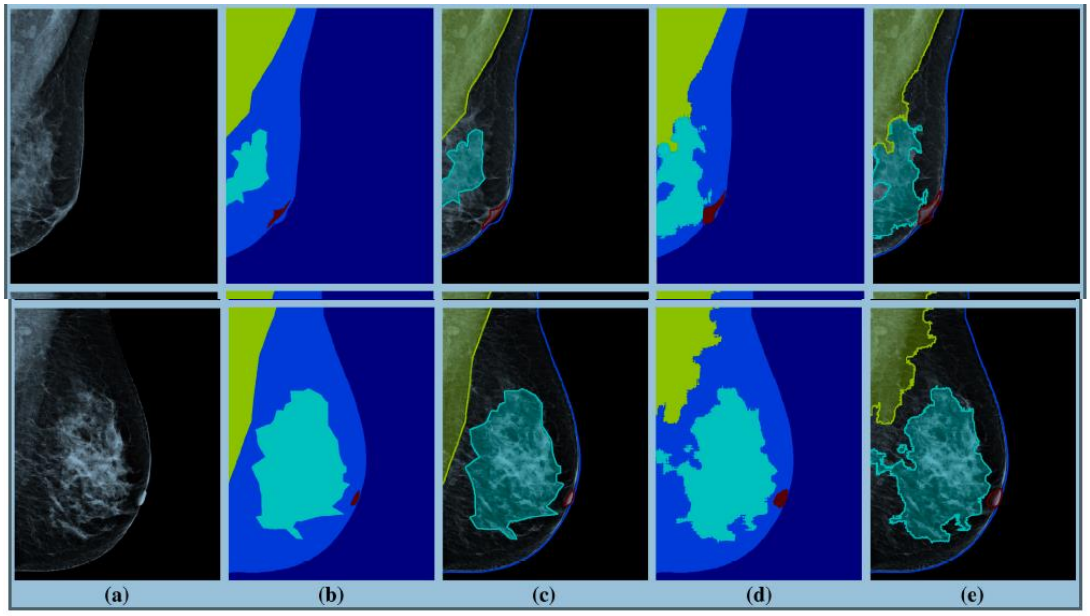


Figure 2.Segmentation results. (a) Original image. (b), (c) Manual labels. (d), (e) Labels obtained using the proposed method. Colour coding is as in Figure 1.

Feature Extraction

The efficient classification of nuclei cells from the total segmented regions requires the generation of meaningful features of very good discriminative ability. Having found the areas of the nuclei enclosed by the detected boundaries, features concerning the shape and the texture of the detected regions can be easily determined. In our work, we use ten shape-based features and two textural features. The values obtained for these features yield a good differentiation between cancerous and healthy cells. These features are proposed as input data for the classification phase.The extracted features consist of Shape features. The detected boundaries for the nuclei are expected to present an ellipse-like shape and several features to describe this characteristic are chosen. Ten features are calculated from the extracted shape of the detected region boundary namely perimeter, compactness, smoothness, eccentricity, solidity, equivalent diameter, extent, major axis length and minor axis length for these features. Textural features. Two features are calculated from the texture of the cell nucleus that is the standard deviation for the intensities of the region in both grayscale and Y-level.

Multilayer Perceptron Classifier

Feed forward networks consist of a series of layers. The first layer has a connection from the network input. Each subsequent layer has a connection from the previous layer. The final layer produces the network's output. Feed forward networks can be used for any kind of input to output mapping. Each network input-to-unit and unit-to-unit connection is modified by a weight. In addition, each unit has an extra input that is assumed to have a constant value of one. The weight that modifies this extra input is called the bias. All data propagate along the connections in the direction from the network inputs to the network outputs. When the network is run, each hidden layer unit performs the calculation on its inputs and transfers the result (Oc) to the next layer of units. While each output layer unit performs the calculation on its inputs and transfers the result (Oc) to a network output. Oc is the output of the current layer unit c, P is either the number of units in the previous hidden layer or number of network inputs, ic,p is an input to unit c from either the previous hidden layer unit p or network input p, wc,p is the weight modifying the connection from either unit p to unit c or from input p to unit c, and bc is the bias. hHidden(x) is the sigmoid activation function for hidden units and hOutput(x) is the linear activation function for output units. Other types of activation functions exist, but these functions were performed in this research.

$$(\Sigma \qquad) \qquad (13)$$

where$h$ $(X$ ——

$$(\Sigma \qquad)(14)$$

Where $(x$

The neural network has to be trained on an appropriate data series to make meaningful forecasts. We use back propagation training algorithm for this purpose. Backpropagation training consists of three steps:

1. Present an example's input vector to the network inputs and run the network: compute activation

functions sequentially forward from the first hidden layer to the output layer.

2. Compute the difference between the desired output for that example, output, and the actual network output (output of unit(s) in the output layer). Propagate the error sequentially backward from the output layer to the first hidden layer.

3. An error term is computed for each unit in the output

$$(x)(D \quad (15)$$

$$(x)\sum \quad (16)$$

Where $(x)=o_k(1$ (17)

Dc is the desired network output (from the output vector) corresponding to the current output layer unit, Oc is the actual network output corresponding to the current output layer unit, and is the derivative of the output unit linear activation function i.e. 1. Also the error term is computed for each unit in the hidden layers . N is the number of units in the next layer (either another hidden layer or the output layer), n is the error term for a unit in the next layer, and wn,c is the weight modifying the connection from unit c to unit n.

For every connection, change the weight modifying that connection in proportion to the error . where wc,p is the weight modifying the connection from unit p to unit c , is the learning rate which controls how quickly and how finely the network converges to a particular solution, and Op is the output of unit p or the network input p. When these three steps have been performed for every example from the data series, one epoch has occurred. Training usually lasts until a predetermined maximum number of epochs (epochs limit) is reached or the network output error (error limit) falls below an acceptable threshold. Training can be time-consuming, depending on the network size, number of examples, epochs limit, and error limit.

ALGORITHM

Input : Breast cancer tumour image

Output : Breast cancer tumour classification

Step 1: Start

Step 2: Image Preprocessing

Step 2.1 : Noise removed by Adaptive GaussianFilter (AGF)

Step 3: Segmentation by DNN

Step 4: Feature extraction

Step 4.1: shape feature extracted by Zernike moments

Step 4.2: Texture feature are extracted by ripley's function.

Step 5 : Feature selection

Step 5.1 : Initial population

Step 5.2: Genetic operations

Step 5.3: Evaluate fitness function

Step 5.4 :Feature selected result

Step 6: Classification using MLP

Step 7: Benign and maligant masses classification results

Step 8: Performance comparison

4. RESULTS AND DISCUSSION

In this work describes about the use of an AGF is proposed as pre-processing for noise removal. Deep Neural Network (DNN) based approach is presented for segmentation. The figure 3 shows the source image for segmentation. The figure 4 shows the pre-processed image. The figure 5 shows the segment image using CeNN algorithm. The figure 6 shows the test image. The figure 7 shows the classification using MLP. The figure 8 shows the DNN Segmented Image. The figure 9 shows the feature extraction. The experimental results shows that the trained Deep Neural Network is featured with quantisation to significantly reduce the computational complexity and non-linear template for robustness.

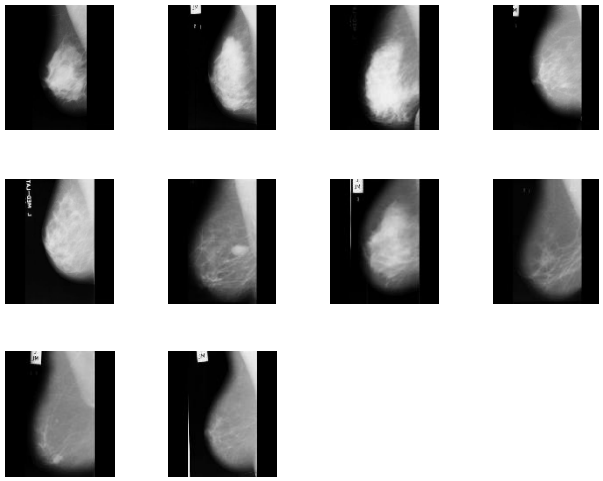


Figure 3.Input image

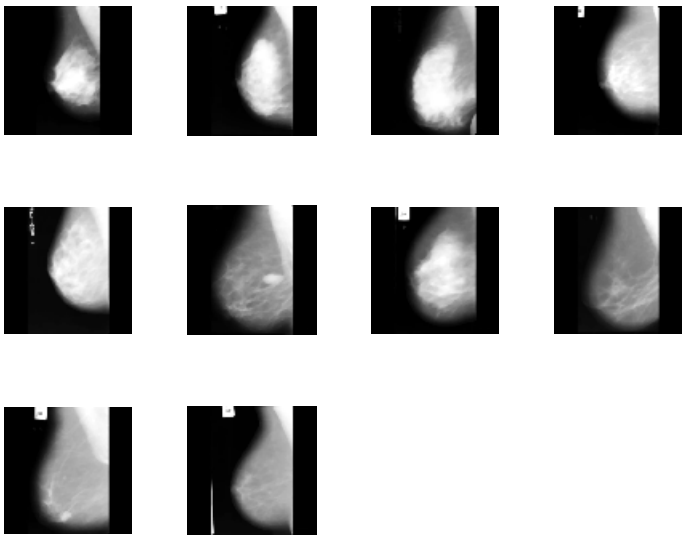


Figure 4.Preprocessed image

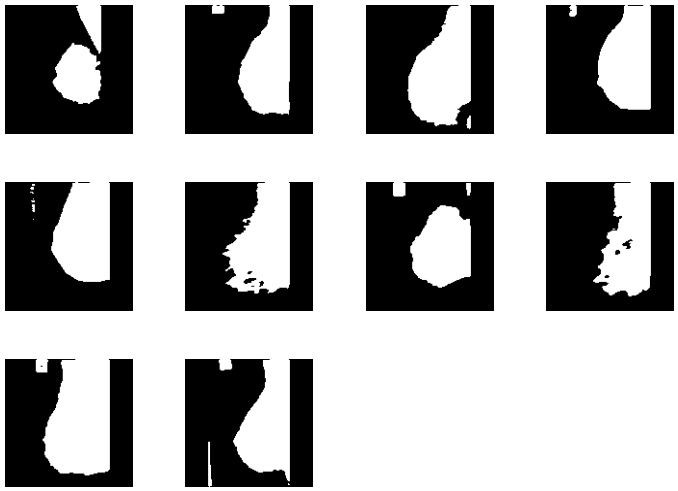


Figure 5.CeNN Segmented Image

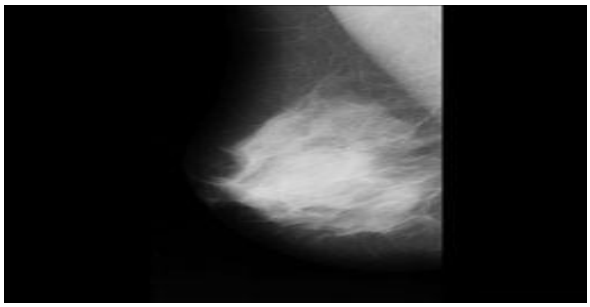


Figure 6.Test image

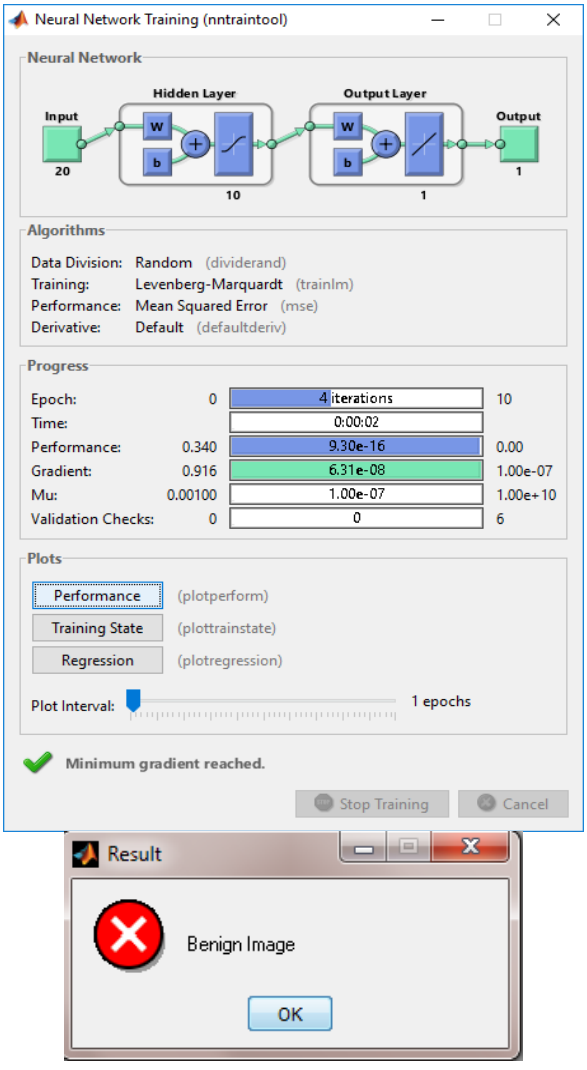


Figure 7. Classification using MLP

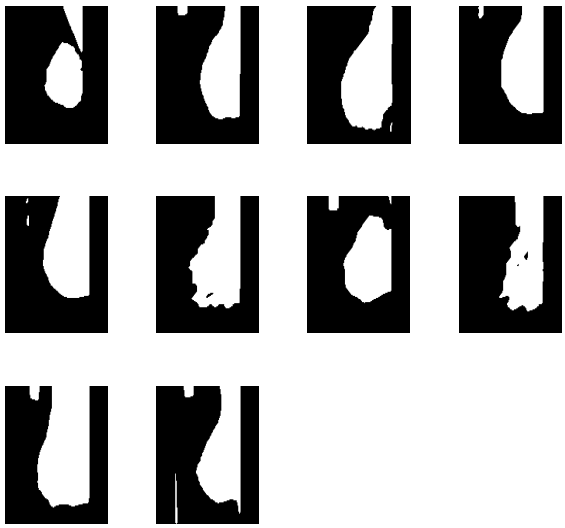


Figure 8 .DNN Segmented Image

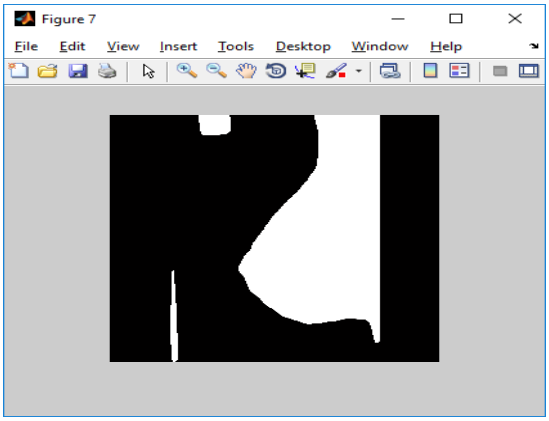


Figure 9. Feature Extraction

Accuracy

Accuracy is one metric for evaluating classification models. Informally, **accuracy** is the fraction of predictions our model got right. Formally, accuracy has the following definition.

————————(18)

Where *TP* = True Positives, *TN* = True Negatives, *FP* = False Positives, and *FN* = False Negatives.

Error Rate

Error Rate is one metric for evaluating classification models. Informally, Error Rate is the fraction of irrelevant predictions results. Formally, Error Rate has the following definition.

cy(19.)

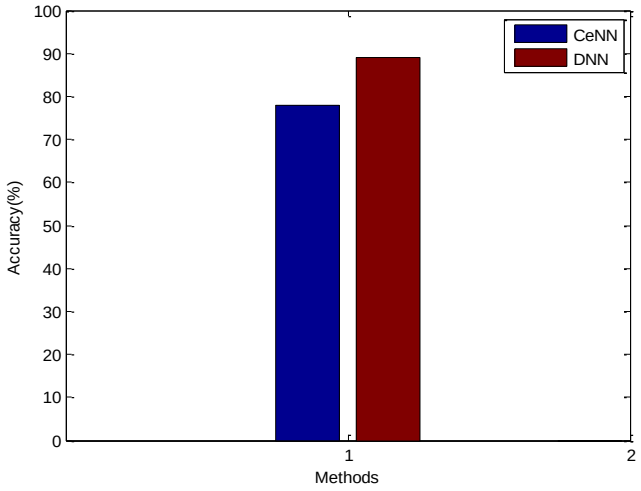


Figure 10. Comparison results of the Breast Tumour Classification Accuracy Between proposed and existing method

Figure 10. shows that the breast tumour classification accuracy comparison results of the proposed and existing system. The Accuracy(%) results are shown in y-axis and methods are represented in the X-axis. So the proposed DNN produces higher accuracy results of the 89% whereas the existing CeNN produces only 78%. It concludes that DNN algorithms produce higher Accuracy results when compared to existing CeNN method.

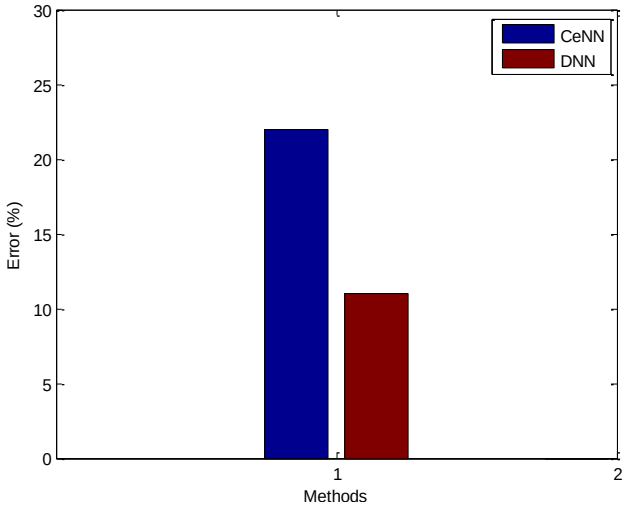


Figure 11.Error rate comparison results of the Breast Tumour Classification between proposed and existing method

Figure 11.shows thatthe breast tumour classification error rate comparison results of the proposed and existing system. The Error(%) results are shown in y-axis and methods are represented in the X-axis. So the proposed DNN produces lesser error results of the 11% whereas the existing CeNN produces22%. It concludes that DNN algorithms produce lesser error results when compared to existing CeNN method.

5. CONCLUSION AND FUTURE WORK

Proposed one new method using an Adaptive Gaussian Filtering (AGF) algorithm in which the filter variance is adapted to both the noise characteristics and the local variance of the signal. Deep Neural Network (DNN) based approach is presented for segmentation. The approach is featured with quantisation to significantly reduce the computational complexity and non-linear template for robustness.MLP is used for classification. At last performance is analyzed by means of using various lists of parameters such as Accuracy and Error rate . In the future we plan on integrating spatial smoothness into the classification process, and further improve the classification rates. To that end, we plan to extend the proposed framework and explore the use of recurrent neural networks instead of the fully convolutional setup.

REFERENCES

1. Bhardwaj A. and A. Tiwari, "Breast cancer diagnosis using genetically optimized neural network model," Expert Systems with Appli-cations, vol. 42, no. 10, pp. 4611–4620, 2015.
2. Chen C.-H., "A hybrid intelligent model of analyzing clinical breast cancer data using clustering techniques with feature selection," Applied Soft Computing, vol. 20, pp. 4–14, 2014.
3. Chua, L.O., and Yang, L.: 'Cellular neural networks: applications', Trans. Circuits Syst., 1988, 35, pp. 1257–1272.
4. Heath, M., Bowyer, K., and Kopans, D.: 'Current status of the digital database for screening mammography', in Krupinski, E. (Ed.): 'Digital mammography' (Kluwer Academic Publishers, Tucson, AZ, USA, 1998), pp. 457–460.
5. Marcano-Cedeno,˜ J. Quintanilla-Dom´ınguez, and D. Andina, "Wbcd breast cancer database classification applying artificial metaplasticity neural network," Expert Systems with Applications, vol. 38, no. 8, pp. 9573–9579, 2011.
6. Nahato K. B., K. N. Harichandran, and K. Arputharaj, "Knowledge mining from clinical datasets using rough sets and backpropagation neural network," Computational and mathematical methods in medicine, Volume 2015, Article ID 460189, pp.1-13, 2015.
7. Nguyen T., A. Khosravi, D. Creighton, and S. Nahavandi,, "Classification of healthcare data using genetic fuzzy logic system and wavelets," Expert Systems with Applications, vol. 42, no. 4, pp. 2184–2197, 2015.
8. Onan, A., 2015. A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer. Expert Systems with Applications, 42(20), pp.6844-6852.
9. Roska, T., and Chua, L.O.: 'Cellular neural networks with nonlinear and delay-type template', Int. J. Circuit Theory Appl., 1992, 20, pp. 469–481.
10. Senapati M. R., A. K. Mohanty, S. Dash, and P. K. Dash, "Local linear wavelet neural network for breast cancer recognition," Neural Computing and Applications, vol. 22, no. 1, pp. 125–131, 2013.
11. Tomar D. and S. Agarwal, "Hybrid feature selection based weighted least squares twin support vector machine approach for diagnosing breast cancer, hepatitis, and diabetes," Advances in Artificial Neural Systems, Volume 2015, Article ID 265637 , pp.1-10, 2015.
12. Wei, C.H., Chen, S.Y., and Liu, X.: 'Mammogram retrieval on similar mass lesions', Comput. Methods Programs Biomed.., 2012, 106, (3), pp. 234–248.
13. Zhang, Y., Tonuro, N., Furst, J., et al.: 'Building an ensemble system for diagnosing masses in mammograms', Int. Comput. Assist. Radiol., 2012, 7, pp. 323–329.