International Journal of Engineering & Technology, 7 (4) (2018) 4813-4818

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet. v7i4.24797
Research paper

Multi feature malignant weight-based mammogram classification with ANN using fuzzy rules

Ramya B [1] *, Dr. K. Arthi [2]

[1] *Ph. D Research Scholar, Bharathiar University, Assistant Professor, Department of Computer Applications, Dr. N. G. P. Arts and Science College (Autonomous), Coimbatore-641046, India*
[2] *Assistant Professor, Department of Computer Application, Government Arts College (Autonomous), Coimbatore – 641 018*
**Corresponding author E-mail: ramyab@drngpasc.ac.in*

Abstract

Towards the development of disease prediction, detection and classification accuracy, the features of mammogram can be used to estimate the similarity. In this paper, we motivate to adapt the fuzzy logic with ANN to perform classification. With the motivation, a multi feature fuzzy logic analysis with ANN Based Mammogram Classification algorithm is presented. First, it is necessary to obtain the gray features of mammogram. The method considers the brightness of nodule, lobulation, granularity and nodule size. Using the features considered, the method generates the fuzzy rules. The artificial neural network has been generated with number of neurons where each layer of the neuron has been initialized with specific feature. At the testing phase, the input features have been used to estimate different weight measures using the fuzzy rule generated. The method estimates multi feature malignant weight (MFMW) using the fuzzy rules at each layer of the neurons. Finally, the classification is performed based on the value of MFMW.

Keywords: *ANN; Fuzzy Rules; Mammogram Classification; MFMW.*

1. Introduction

The breast cancer has been identified as most threatening disease in women. The presence of cancer can be identified by the physician by monitoring the mammogram of the women. But the accuracy of the classification requires different automatic classification tools to support the medical practitioner. Diagnosing the presence of cancer would help to eradicate it in an easier way and is highly curable at the initial stage. The mammogram images are captured through the MRI scan devices and other capturing devices. By monitoring the captured images, the medical practitioner would identify the presence of disease. But, the accuracy of the classification is highly questionable because, they vary in shape, size, texture, granularity, and so on. So, it needs some efficient classification algorithms and scientific tools in the classification of the images.

Image processing techniques has higher impact in the medical domain to support many decision making processes. There are number of classification algorithms proposed earlier to support the classification of mammograms. The k-nearest neighbor (KNN) algorithm has been used towards the classification of mammogram images. The method considers a single feature like the white mass value, for the classification by computing distance measures. The accuracy of classification is not up to the mark being expected. Similarly, there are number of algorithms been used in this problem. The most algorithms use either shape, texture or other features. There is no algorithm which combines multiple features in the problem of mammogram classification. To achieve higher performance in the classification, it is necessary to consider more features. This paper, aimed to combine such multiple features in the classification of mammogram images.

Fuzzy rules set have been used in this approach of mammogram classification. Because of the features of the mammogram varies in their size, shape, mass value, brightness, granularity and so on. They vary between different values, so it is necessary to consider the varying feature values. The fuzzy rules can be generated from the train set of mammogram images, and that can be used for the classification at the test phase. Similarly, the most algorithms miss many values of mammogram instances. By adapting the artificial neural network for this problem, the hidden values can be used. The neural network has been generated at different layers where each layer of neuron estimates the weight measure at specific feature according to the fuzzy rules available and the feature values of the neuron at specific feature considered. Based on the weight measures, the problem of mammogram classification can be performed effectively.

2. Related works

The problem of mammogram classification has been approached with several techniques. This section presents a set of methods and identifies few problems from the methods in detail.

To detect the tumors in breast a combination of RF and RF-ELM based approach is presented in [1]. The classifier uses digital mammogram pictures available from MIAS dataset. The method improves the image quality and identifies the interested region. The segmentation results have been used to extract the GLCM features. Based on the GLCM features extracted, the method performs classification using CBF approach.

The density of breast tissue has been used for the classification of mammogram images in [2]. In the preprocessing stage, the noise has been removed and the presence of label and artifacts are removed by applying morphological operations. The method extracts the texture of tissue and estimates different measures like standard deviation, skewness, mean and correlation. Based on the extracted features, the classification is performed using support vector machines.

A region based approach for the classification of mammogram is presented in [3], which uses convolution neural network. The method adapts the neural network with you only look once technique for classification. The mass feature obtained from the interested region has been applied with fully connected neural network in the classification process.

In [4], the propagation uncertainty in breast mass classification has been approached. The method uses the characterization of metrology with CAD systems. The result of classification has been validated ROC and AUC. In [5], the author presented a classification algorithm for tomosynthesis and mammogram images. The classification is performed with connected neural network and has been evaluated with huge set of mammogram images.

An morphological wavelet based mammogram lesion classification is performed in [6], with the support of ANN. The performance of various classifiers has been validated with the IRMA data set. For evaluation, SVM, MLP and RBF are considered. In [7], the digital mammograms are classified according to the mass value. The method extracts different features and feature selection is performed using T-Test algorithm. The classification is performed using SVM, KNN and ANN.

In [8], the gestalt psychology based approach is presented to identify the breast mass. The method incorporated with integration of sensation, semantic and their verification. The method uses visual rules generated from the characteristics of morphology breast mass. The method is more efficient in both top down and bottom up techniques.

A texture based classification approach is presented in [9], which extracts the breast density values and uses to train them. The classification is performed using hidden markov random field. In [10], an object based mammogram classification algorithm is presented using deep learning technique. The method performs the classification in an automatic manner using INbreast dataset.

A transfer learning based lesion classification is presented in [11]. The method uses neural network with convolution for the classification. The transfer learning algorithm is used for the classification task. In [12], the dimensionality of mammogram has been reduced towards the development of mammogram classification. The method reduces the dimensionality by using the continuous wavelet transform. The features extracted have been passed through number of classifiers. The performance of the classifiers has been evaluated with different data sets.

All the methods discussed in this section has the problem of higher false classification ratio as they consider only limited features in mammogram classification.

3. Proposed work

3.1. MFMW based mammogram using fuzzy logic/ANN

The proposed multi feature malignant weight based mammogram classification algorithm first takes the mammogram images and performs noise removal by applying gabor filters. In the second stage, the method performs segmentation based on the gray scale values. Third, the method extracts the brightness of the nodule, lobulation, granularity and nodule size. Using the features extracted, the method generates artificial neural network and initializes the neurons with the features. Using the same, the method generates the fuzzy rules for each feature considered. At the test phase, the features are extracted from input image and estimates different weight measures to produce MFMW measure. The same has been used for classification.

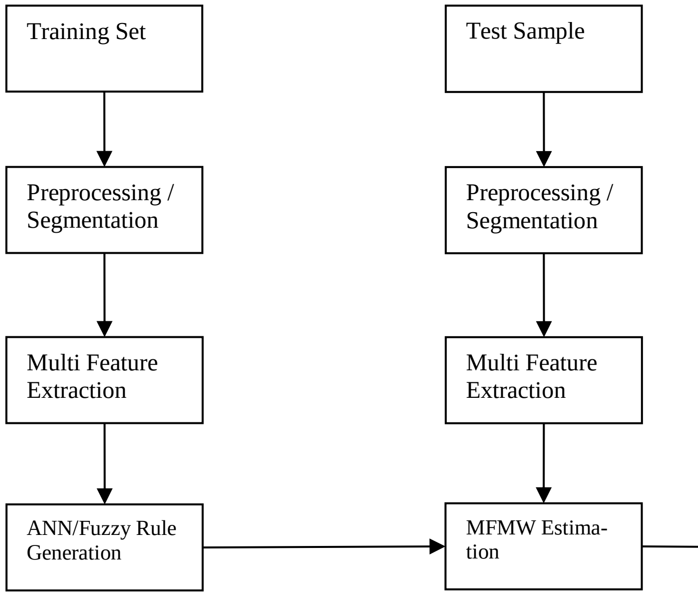

Fig. 1: Architecture of MFMW Based Mammogram Classification with ANN/Fuzzy Logic.

The Fig 1, shows the block diagram of the proposed MFMW based mammogram classification algorithm using the artificial neural network and fuzzy logic. The detailed approach has been presented in this section.

3.2. Preprocessing/segmentation

The mammogram image captured through the device has noise intruded by the device being used. First, the Gabor filter has been applied to remove the noise. Second, the method generates the histogram of gray scale values and identifies two distinct bins with different boundary. Based on the lower bin value selected, the method performs segmentation. The method groups the lower bin values with gray scale as the interested region and leaves other as uninterested. Based on this, the segmentation is performed. The segmented image has been used for further processing.

The gabor filter has been applied in different level as follows:

Initialize the filter Gf = $\int_{i=1}^{\mathrm{No.of\ Level}} \mathrm{GF}(i, \mathrm{covariants})$ (1)

Apply the image with GF to perform noise removal as follows:

Noise removed image NRI = $\int_{i=1}^{\mathrm{No\ of\ Level}} \mathrm{GF}(\mathrm{MI}, i)$ // MI – mammogram Image. (2)

Generate Histograms of Gray scale values as follows.

$$\mathrm{Hist} = \int_{i=1}^{256} \sum \mathrm{NRI}(\mathrm{Pi} == i) \quad (3)$$

Now, select the bins with higher gray scale as interested region.

$$\mathrm{Sbin} = \int_{i=1}^{\mathrm{Size(Hist)}} \sum \mathrm{Max}(\mathrm{Hist}(i)) \,\&\&\, size(\mathrm{Hist}(i)) > 10 \quad (4)$$

Perform segmentation with the selected bin on the NRI.

If there exist more than one region of interest, then it has been returned as set of bins as bin set Bins.

(A) **(B)**

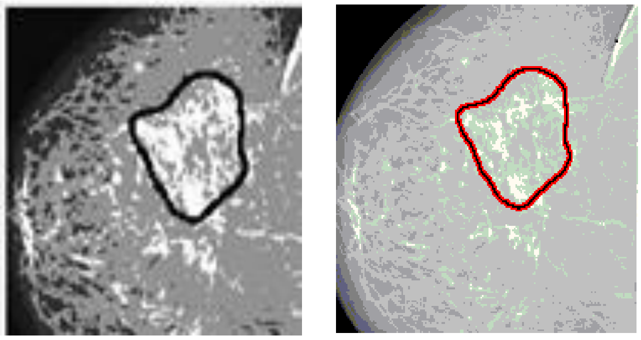

Fig. 2: A) and B) Original Image and Preprocessed Image.

The Fig 2a, shows the original image being taken and 2b shows the result of preprocessing which removes the noise and applies the histogram equalization to produce the segmented image.

3.3. Multi feature extractions

The method considers size and brightness of nodule, margin and granularity. The brightness of the nodule has been extracted from the ROI and its gray scale values of the pixels of ROI. Similarly, the size of the nodule has been counted from identified ROI. The margin feature has been extracted by computing the area being covered. Similarly, if there is more than one ROI then it has been counted for lobulations. The granularity feature has been extracted using the gray scale value itself. The extracted features are converted into a feature vector which has been used to perform classification.

First, the input noise is removed and segmented image has been read as SI.

From the segmented image SI, and the bin set given as Bins, this method computes the size of nodule

$$\text{Son} = \int_{i=1}^{size(SI)} \sum SI(i).value == Bins.Value \tag{5}$$

The brightness of the nodule is estimated as follows:

$$\text{Nodule brightness Nbr} = \int_{i=1}^{size(SI)} \frac{\sum_{j=1}^{Son} Dist(ROI(j),SI(i))}{Son} \tag{6}$$

The margin value has been estimated as follows:

$$\text{Mar} = {Son}/{size(SI)} \tag{7}$$

Similarly, the granularity has been extracted as follows:

$$\text{Gr} = \frac{\sum_{i=1}^{size(ROI)} ROI(i).GrayScale}{Size(ROI)} \tag{8}$$

Now the above extracted features are converted into feature vector as follows:

$$F = \{Son, Nbr, Mar, Gr\} \tag{9}$$

The generated feature Vector has been used for the classification of mammogram images.

3.4. ANN/fuzzy rule generation

The method consider number of mammogram features for the classification of mammogram images and to perform cancer detection. The feature vectors generated from the mammogram images are used for the fuzzy rule generation. The size of nodule, brightness of the nodule, margin value and the granularity has been used for the rule generation. At the training phase, the method extracts the above mentioned features from all the mammogram images. The extracted features have been used for the rule generation. Finally, the method generates number of neurons and initializes them with feature vector. The generated rules have been used for weight estimation in the classification stage or in the testing phase.

Algorithm
Input: Mammogram Image Set Mis
Output: Fuzzy Rule Fr
Start
    Read data set Mis
    For each input image Im
        SI = Preprocessing (Im)
        Feature Vector Fv = Multi-Feature-Extraction (SI)
    End
    For each feature f
        Compute Min feature Value Fmin {Son,Nbr,Mar,Gr} = $\int_{i=1}^{size(Mis)} Min(\forall Fv(Son, Nbr, Mar, Gr))$
        Compute Max Feature values Fmax {Son,Nbr,Mar,Gr} = $\int_{i=1}^{size(Mis)} Max(\forall Fv(Son, Nbr, Mar, Gr))$
    End
    Generate Fuzzy Rule Fr = Merge (Fmin,Fmax)
    Initialize neural network.
    For each neuron
        Initialize with feature vector f
    End
Stop

The fuzzy rule generation algorithm extracts various features and generates minimum and maximum values of all the features to produce the fuzzy rule.

3.5. Mammogram classification

The input mammogram image has been taken for classification. First, the method performs preprocessing to produce segmented image. In the second stage, the method extracts the features of multi variant and generates a feature vector. The feature vector has been given for ANN classification where the neuron estimates different weight measures. The method estimates nodule weight, brightness weight, margin weight and granularity weight. Each weight has been measured based on the values of the feature and the fuzzy values. Finally, the method estimates the MFMW measure which has been used for the classification.

Algorithm
Input: Mammogram Image MI, Fuzzy Rule Fr, ANN
Output: Class C
Start
    Read mammogram image MI
    SI = Preprocessing (MI)
    Fv = Feature Extraction (SI)
    {Nw,Bw,Mw,Gw} = Apply-Neural-Net (Fv,ANN)
    Compute MFMW
    If MFMW>Th   Then
        Malignant
    Else
        Benign
    End
Stop

The above discussed algorithm estimates different weight measures and computes MFMW value which has been used for the classification.

3.6. MFMW estimation

In this stage, the method reads the feature vector given. From the feature vector and the neuron reads the fuzzy rule available. Using these two, for each feature considered, the neuron estimates the measures as follows:

Nodule weight Nw has been measured based on the size of the nodule. The size of the nodule may differ with different images. First it checks whether the size of the nodule falls within the range and if so, it computes the distance with min and max values. Based on the distance measure, the method estimates the nodule weight.

Nodule weight Nw =
$$\frac{Dist(Fv(Son),FR(Son).Min)+Dist(Fv(Son).Fr(Son.Min))}{2} \quad (10)$$

Similarly, the weights of brightness has been measured as follows:

Brightness weight Bw =
$$\frac{Dist(Fv(Nbr),FR(Nbr).Min)+Dist(Fv(Nbr).Fr(Nbr.Min))}{2} \quad (11)$$

The margin weight has been measured as follows:

Margin weight Mw =
$$\frac{Dist(Fv(Mar),FR(Mar).Min)+Dist(Fv(Mar).Fr(Mar.Min))}{2} \quad (12)$$

Finally, the granularity weight has been measured as follows:

Granularity weight Gw =
$$\frac{Dist(Fv(Gr),FR(Gr).Min)+Dist(Fv(Gr).Fr(Gr.Min))}{2} \quad (13)$$

Using these weights, the MFMW weight has been measured as follows:

$$\text{Compute MFMW} = \frac{Nw}{size(SI)} \times \frac{BW}{Avg(ROI)} \times \frac{MW}{size(ROI)} \times \frac{GW}{Avg(SI)} \quad (14)$$

4. Results and discussion

The proposed multi feature malignant weight based mammogram classification algorithm has been implemented using Matlab. The method has been evaluated for its performance in mammogram classification using different data set. The data set from open source has been used to evaluate the performance of the algorithm in cancer detection.

Table 1: Details of Dataset

Evaluation Key	Value
Data Set	DDSM
Number of Classes	Normal, Malignant, Benign (3)
Number of Training Images	2000
Number of Test Images	620

The Table 1, shows the details of data set being used for the evaluation of proposed mammogram classification algorithm. The Digital Data set for Screening Mammogram (DDSM) has been published by the United states army medical community. The data set has been obtained and used for the evaluation of the proposed algorithm.

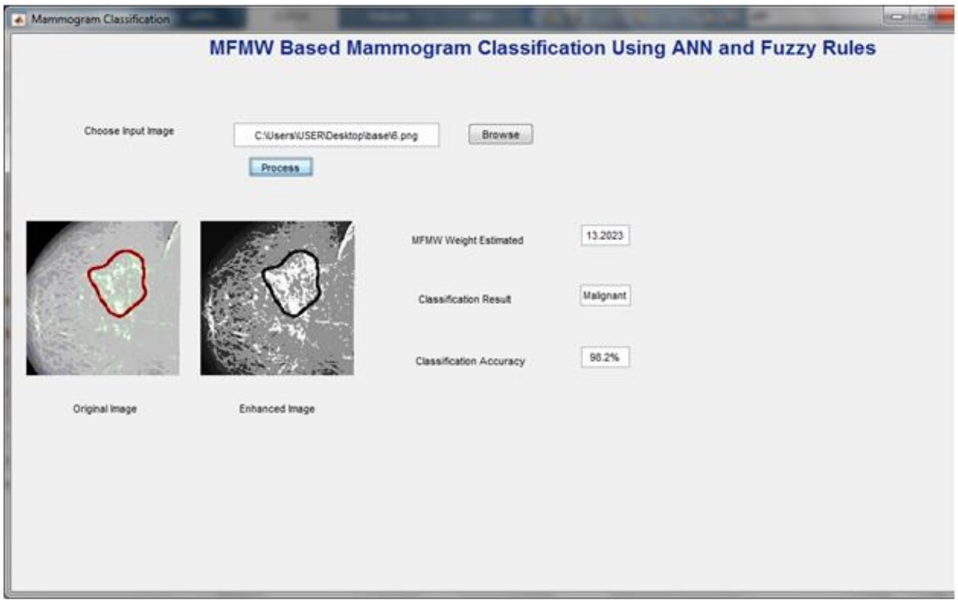

Fig. 3: Snapshot of Mammogram Classification.

Fig 3, shows the snapshot of mammogram classification and shows the result produced by the proposed algorithm.

Table 2: Benign vs. Malignant vs. Normal Classification

Mass Type	Classified as Benign	Classified as malignant	Classified as Normal	Total	Classification Rate
Benign	611	4	0	615	0.97
Malignant	25	560	1	586	0.955
Normal	0	1	560	560	0.99
Overall				1761	0.98

The Table 2, shows the comparative result of classification on various class of mammogram images. The overall classification accuracy has been measured and presented in Table 2.

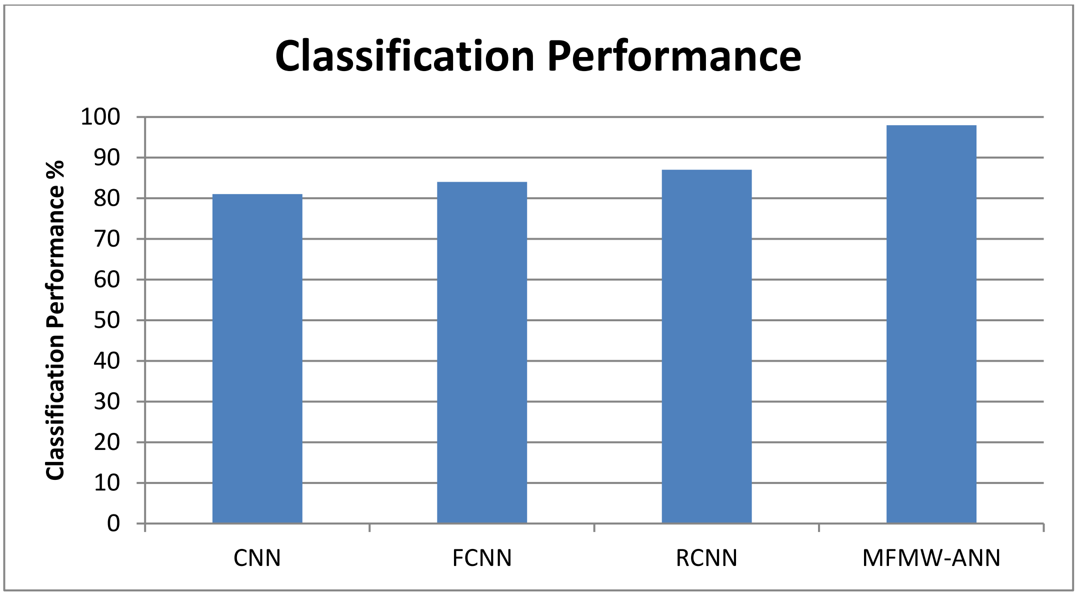

Fig. 4: Comparison on Mammogram Classification Performance.

The performance of classification accuracy of mammograms has been measured with different methods. According to the results obtained, the proposed MFMW based ANN for mammogram classification has achieved higher performance than other methods.

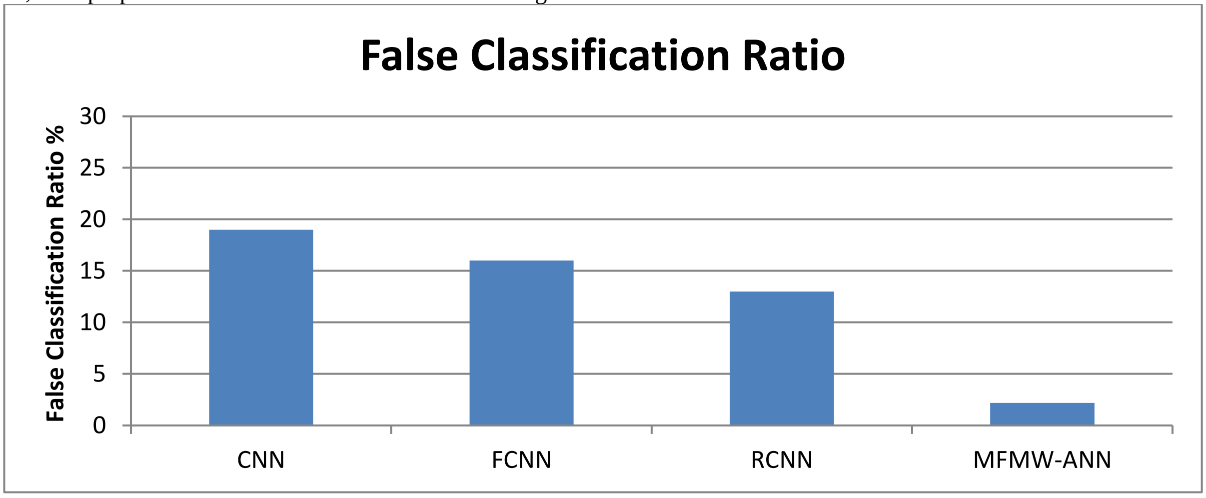

Fig. 5: Comparison on False Ratio.

The ratio of false results produced in the classification by different methods has been measured and compared. The result received has shown that the proposed multi feature malignant weight based ANN for mammogram classification has reduced the false ratio than other approaches.

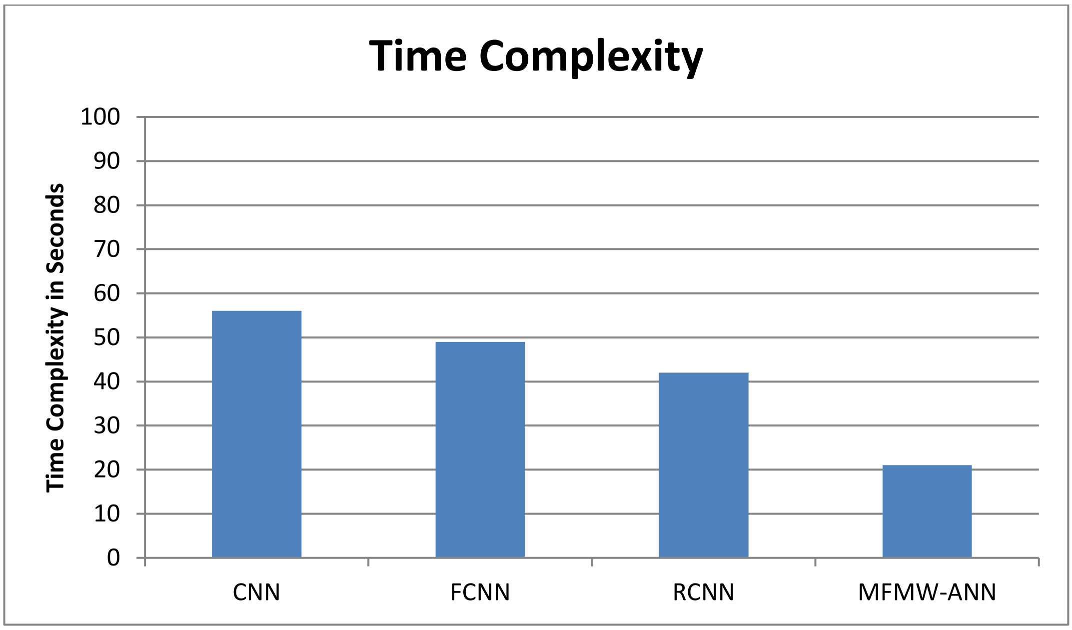

Fig. 6: Comparison on Time Complexity.

The time taken for the classification of mammogram image by different method has been recorded and compared. The result shows, the proposed approach has produced less time than other methods for the classification of the mammogram image.

5. Conclusion

In this paper, an efficient multi feature malignant weight based mammogram classification algorithm with Fuzzy rules and artificial neural network has been presented. The method performs preprocessing and segmentation of mammogram images initially. From the segmented image, the features like size of nodules, brightness, granularity has been extracted. Using the extracted features, the fuzzy rules are generated and neural network has been initialized with number of neurons. At the testing phase, the segmentation is performed to extract the features of the input mammogram image. Then, the method estimates different weights through the neurons and ANN performs different weight estimation. Finally, the MFMW measure has been computed which has been used for the classification. The proposed algorithm has produced classification accuracy up to 98.2% and the false ratio has been reduced up to 2% with 22 second time complexity.

References

[1] R. D. Ghongade ; D. G. Wakde, Detection and classification of breast cancer from digital mammograms using RF and RF-ELM algorithm, IEEE International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech), 2017. https://doi.org/10.1109/IEMENTECH.2017.8076982.

[2] S. Sathiya Devi; S. Vidivelli, Classification of breast tissue density in digital mammograms, International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

[3] M. A. Al-masni ; M. A. Al-antari ; J. M. Park ; G. Gi ; T. Y. Kim ; P. Rivera ; E. Valarezo ; S.-M. Han; T.-S. Kim Detection and classification of the breast abnormalities in digital mammograms via regional Convolutional Neural Network, Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),2017.

[4] Arianna Mencattini ; Marcello Salmeri ; Giulia Rabottino ; Simona Salicone, Metrological Characterization of a CADx System for the Classification of Breast Masses in Mammograms, IEEE Transactions on Instrumentation and Measurement (Volume: 59 , Issue: 11 , Nov. 2010)

[5] Xiaofei Zhang ; Yi Zhang ; Erik Y. Han ; Nathan Jacobs ; Qiong Han ; Xiaoqin Wang ; Jinze Liu, Classification of Whole Mammogram and Tomosynthesis Images Using Deep Convolutional Neural Networks, IEEE Transactions on NanoBioscience (Volume: 17 , Issue: 3 , July 2018).

[6] Thais Nayara da Cruz ; Thamyris Mayara da Cruz ; Wellington Pinheiro dos Santos , Detection and classification of mammary lesions using artificial neural networks and morphological wavelets, IEEE Latin America Transactions (Volume: 16 , Issue: 3 , March 2018)

[7] Basma A. Mohamed; Nancy M. Salem, Automatic classification of masses from digital mammograms, 2018 35th National Radio Science Conference (NRSC). https://doi.org/10.1109/NRSC.2018.8354408.

[8] Hongyu Wang, Jun Feng, Qirong Bu, et al., “Breast Mass Detection in Digital Mammogram Based on Gestalt Psychology,” Journal of Healthcare Engineering, vol. 2018, Article ID 4015613, 13 pages, 2018.

[9] Yang Can Gon, gMichael Brady, Styliani Petroudi, Texture Based Mammogram Classification and Segmentation, International Workshop on Digital Mammography, pp 616-625, 2006.

[10] Anna Horváth, Zsuzsa Unger, Péter Pollner, and István Csabai, Detecting and classifying lesions in mammograms with Deep Learning, scientific reports,2018.

[11] Fan Jiang, Breast mass lesion classification in mammograms by transfer learning, International Conference on Bioinformatics and Computational Biology, Pages 59-62, 2017. https://doi.org/10.1145/3035012.3035022.

[12] Mohammad Kazem Ebrahimpour, Hamid Mirvaziri & Vahid Sattari-Naeini, Improving breast cancer classification by dimensional reduction on mammograms, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization Volume 6, 2018 - Issue 6, 2018.