A Hybrid Multifarious Clustering Algorithm for the Analysis of Memmogram Images

T. Velmurugan[1], E. Venkatesan[2]

[1]PG and Research, Department of Computer Science, D. G. Vaishnav College, Chennai, India
[2]PG Department of Computer Science, R. V. Government College, Chengalpattu, India
Email: velmurugan_dgvc@yahoo.co.in, venkatelumalai12@yahoo.co.in

Abstract

A number of clustering algorithms were used to analyze many databases in the field of image clustering. The main objective of this research work was to perform a comparative analysis of the two of the existing partitions based clustering algorithms and a hybrid clustering algorithm. The results verification done by using classification algorithms via its accuracy. The performance of clustering and classification algorithms were carried out in this work based on the tumor identification, cluster quality and other parameters like run time and volume complexity. Some of the well known classification algorithms were used to find the accuracy of produced results of the clustering algorithms. The performance of the clustering algorithms proved meaningful in many domains, particularly k-Means, FCM. In addition, the proposed multifarious clustering technique has revealed their efficiency in terms of performance in predicting tumor affected regions in mammogram images. The color images are converted in to gray scale images and then it is processed. Finally, it is identified the best method for the analysis of finding tumor in breast images. This research would be immensely useful to physicians and radiologist to identify cancer affected area in the breast.

Keywords

Medical Images, Hybrid Clusteing Algorithm, K-Means Algorithm, Fuzzy C Means Algorithm, Classification Algorithms

1. Introduction

The breast cancer segmentation process is still a challenging issue in the field of medicine. The problem of breast malignant neoplastic disease is one of the ha-

zardous cancers for women around the world. It has been rated as the second most common disease that causes death in adult females. The highest incidence of breast cancer in women has increased significantly in the last few years. The breast cancer is a malignant tumor that grows in or around the breast tissue, mainly in the milk ducts and glands. A tumor usually starts as a lump or calcium deposit that develops as a result of abnormal cell growth. Most breast lumps are benign but can be premalignant (may become cancer). Breast cancer is classified as either primary or metastatic. The initial malignant tumor that develops within the breast tissue is known as primary breast cancer. Sometimes, primary breast cancer can also be found when it is spread to lymph nodes that are close by in the arm pit. Metastatic breast cancer, or advanced cancer, is formed when cancer cells located in the breast break away and travel to another organ or part of the body [1] [2] [3]. Detecting cancer at advanced stage leads to very complicated surgeries and the chances of death are also very high. Early detection of breast cancer helps in less complicated operations and early recuperation and many such tests have been initiated in a successful manner. Some of those tests are mammography, ultrasound, etc.

Mammography is a method that helps in early detection of breast cancer [4] [5] [6]. Though mammography has been identified as the best method, finding the mass (or Classifications) and spreading of cancer in the female body from mammography images has proved to be very difficult. Expertise radiologists are needed for accurate reading of a mammogram image for the prediction of mass and other types of diseases in the breast tissue. Frequently used partition based data mining algorithms, namely k-means and Fuzzy C-Means (FCM) have been used for analyzing the mammogram images in this research work and compared to the proposed hybrid algorithm named as Multifarious Clustering Algorithm (MCA). The mammography image helps to provide some measures in society to help the physicians decide whether a certain disease is abnormal or normal [7].

The purpose of this research work is to identify the tumor, in and around the breast and find its affected region by partitioning the images into clusters based on its intensity and color contrast. The tumor area has been identified in some suitable stage of the clustering process [8]. The clustered results are then verified by the classification algorithms for its performance accuracy [9]. With these small introductions, the structure of the research paper is organized as follows. Section 2 explores the materials and methods used in this research work. Section 3 gives the methods of image clustering. The results and discussion is given in Section 4. Finally, Section 5 concluded the research work via its findings.

2. Materials and Methods

A number of clustering algorithms were used to analyze many databases in the field of image clustering. The main objective of this research work was to perform a comparative analysis of the three partitions based clustering algorithms and verification of the accuracy denoted by classification algorithms. The performance of clustering and classification algorithms were carried out in this

work based on the finding of tumor, cluster quality and other parameters like running time and volume complexity. The classification algorithms were used to find the accuracy of produced results of the clustering algorithms. The k-Means, fuzzy C-Means (FCM) and Multifarious clustering algorithms performance proved meaningful in many domains, particularly k-Means, FCM and multifarious clustering technique have revealed their efficiency in terms of performance in predicting tumor affected regions in mammogram images.

2.1. Description of Data Set

This research work uses mammogram images of three types such as normal, benign and malignant. The mammogram images collected from Swamy Vivekananda Diagnostic Centre (SVDC) Hospital in Chennai at D.G. Vaishnav College campus. Symptoms of abnormality in some of the mammogram images were marked by SVDC head. The mammogram breast cancer images in DICOM (Digital Imaging and Communications in Medicine) format were taken for analysis. The DICOM file format supports the encapsulation of object type data. In this research work, the attributes of mammogram images of the patients data like age, gender, modality, study description, date of the image taken, image size and type etc. were considered for analysis. One of the example experimental data is shown in **Figure 1**.
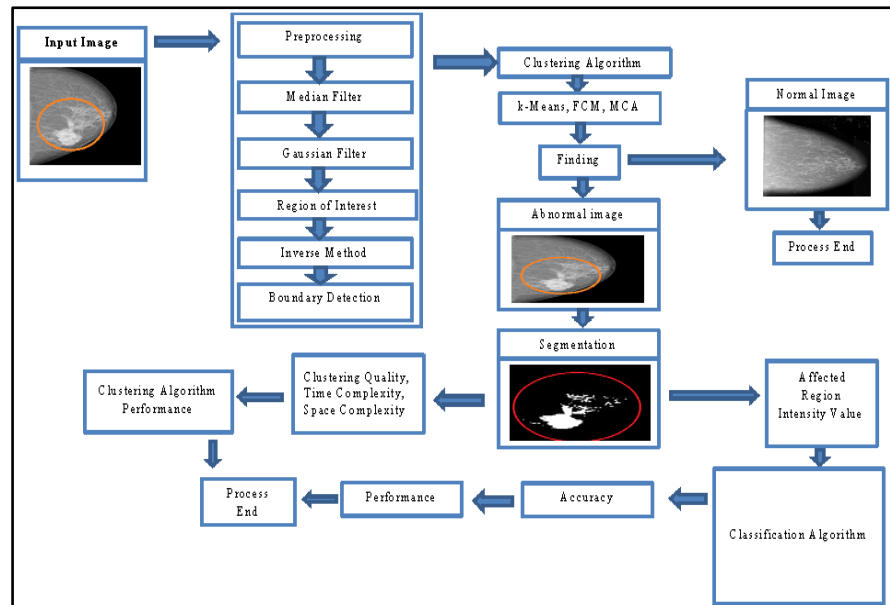
2.2. Proposed Method

Many methods used by various researchers for the analysis and findings of breast cancer in mammogram images. This research uses 310 images for the analysis, which includes three types: normal 10, benign 116 and malignant 184 images to find the affected and unaffected images of mammograms. The proposed method has three stages; pre-processing, image segmentation and classification. Image pre-processing techniques are necessary, in order to find the orientation of the mammogram, to remove the noise and to enhance the quality of the image. Since all the images are extracted by using the clustering and classification to find the tumor area and to find the accuracy respectively. MATLAB software was used to write the source code for the entire work in this research. The various methods used for this work are discussed as follows. The steps involved in the proposed Architecture in **Figure 2**. The steps involved in the proposed method.

Step 1: Convert the mammogram image DICOM format into JPG format.

Step 2: Input the images for preprocessing.

| (a) | (b) | (c) |
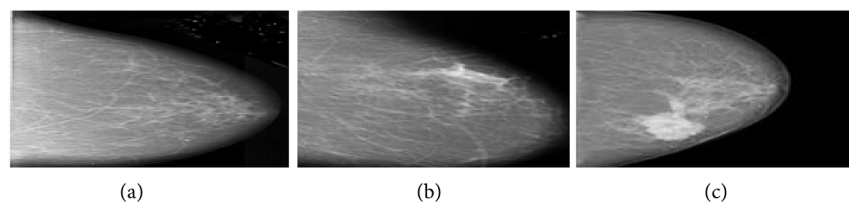
Figure 1. Original images (a) Normal; (b) Benign; (c) Malignant.

Figure 2. Proposed architecture.

Step 3: Preprocessing the images using median filter, Gaussian method, region of interest, inverse method and boundary detection methods to remove the noise.

Step 4: Apply the k-Means, FCM and multifarious algorithm to find the affected region based on the intensity of the images.

Step 5: Enter the number of clusters.

Step 6: Display the tumor affected area by k-Means, FCM and multifarious algorithm via its output images.

Step 7: Find the number of pixels in each and every output of the k-Means, FCM and multifarious algorithm.

Step 8: The k-Means, FCM and multifarious algorithms run time, volume complexity and clustering quality in comparison of the best algorithms.

Step 9: Input the number of pixel values into the input for classification algorithms.

Step 10: Find the accuracy using the classification algorithms J48, JRIP, SVM, Naive Bayes and CART.

Step 11: Find the performance of classification algorithms based on its accuracy.

2.3. Preprocessing Techniques

The main objective of the preprocessing is to develop the image quality to make it ready for further processing by removing or reducing the non related and surplus parts in the background of the medical images. These methods are separated into following categories data cleaning, data integration, data transformation and data reduction. The steps involve the process are Region of Interest (ROI), Inverse method and boundary detection method [10]. The preprocessing

methods used in this work are Median filter [11], Gaussian filter [12] [13], Regions of Interest (ROI) [14] [15], inverse method [14] and boundary dedection method [16] [17]. These methods are exactly utilized for the preprocessing of mammogram images.

3. Image Clustering

The main purpose of clustering is to divide a set of objects into significant groups. The clustering of objects is based on measuring of correspondence between the pair of objects using distance function. Thus, result of clustering is a set of clusters, where object within one cluster is further similar to each other, than to object in another cluster. The Cluster analysis has been broadly applied in numerous applications, including segmentation of medical images, information analysis, and image processing. Clustering is also called segmentation in images form some applications The process of grouping a set of physical or abstract objects into classes of similar objects is called clustering. By clustering, one can identify dense and sparse regions and therefore, discover overall distribution patterns and interesting correlations between data attributes. Thus, clustering in measurement space may be an pointer of similarity of image regions, and may be used for segmentation purposes [18] [19] [20].

3.1. The k-Means Algorithm

The k-Means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. Since, k-mean clustering is normally introduced to group a set of data points $\{x1, x2, ..., xn\}$ into k clusters. It has high computational efficiency and can support multidimensional vectors. So it reduces the distortion measurement by minimizing a cost function as:

Where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point ( )j i x and the cluster center cj , is an indicator of the distance of the n data points from their respective cluster centers. The algorithm is composed of the following steps:

Step 1: Place k points into the space represented by the objects that are being clustered. These points represent initial group centroids.

Step 2: Assign each object to the group that has the closest centroid.

Step 3: When all objects have been assigned, recalculate the positions of the k centroids.

Step 4: Repeat steps 2 and 3 until the centroids no longer move.

This produces a separation of the objects into groups from which the metric to be minimized can be calculated. The k-means is simple clustering algorithm that has been improved to several problem domains [15].

3.2. Fuzzy C-Means Algorithm

This method is widely used in pattern recognition. It is based on minimization of the following objective function. Where m is any real number greater than 1,

x_i is the ith of d-dimensional measured data, c_j is the d-dimension center of the cluster, u_{ij} is the degree of membership of x_i in the cluster j, and $\|*\|$ is any norm expressing the similarity between any measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers , c_j. This iteration will stop when, $\max y \left\{ \left| u_{ij}^{(k+1)} - u_{ij}^{(k)} \right| \right\} \leq \xi$ where ξ is a termination criterion between 0 and 1, whereas k is the iteration steps. This procedure converges to a local minimum or a saddle point of Fm. The algorithm is composed of the following steps:

Step 1: Initialize $U = \left[u_{ij} \right]$ matrix, $U(0)$.

Step 2: At k-step: calculate the centers vectors $C(K) = \left[c_j \right]$ with $U(K)$.

Step 3: Update $U(K), U(K+1)$ $U(k)$,

Step 4: if $\|U(K+1) - U(K)\| < \xi$ then STOP; otherwise return to step 2 [18].

3.3. Multifarious Clustering Algorithm

A hybrid method called "Multifarious Clustering Algorithm (MCA)" is proposed in this research work, which incorporates the advantages of both k-Means and FCM algorithms. The MCA is a method of clustering which allows one part of data to belong to one or more clusters. This algorithm is newly developed with the combination of k-Means and FCM algorithms for this work. It is based on minimization of an objective function:

where m is any real number greater than 1, W_{IJ} is the degree of membership of x_i in the cluster j, x_i is the ith of z-dimensional measured data, c_j is the z-dimension center of the cluster, and $||*||$ is any norm expressing the similarity between any measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership W_{IJ} and the cluster centers Cj by:

This iteration will stop when $\max_{ij} \left\{ \left| W_{ij}^{(k+1)} - W_{ij}^{(k)} \right| \right\} < \xi$, where ξ is a termination criterion between 0 and 1, whereas k is the iteration steps. This procedure converges to a local minimum or a saddle point of V_M The algorithm is composed of the following steps:

Input: x_j is the cluster j, C_J is the z dimension center, W_{IJ} is the degree of membership. Let $C_J, X_j \in w_{ij}$

Output

$w_{ij} = \left[\arg \right]$ Matrix

While $k = x_i$ do

$$C^K = \left[C_J \right] \quad U^K \leq 1, \quad M_I = X_I - C_J, \quad M_K = X_I = c_k, \quad W_{IJ} = \frac{M_J}{M_K} \qquad (12)$$

$$MAX_{IJ} = \left[{}_w (K+1) - {}_{W_{iij}} (k) \right]$$

Return C^K;

End;

In this algorithm, data are bound to each cluster by means of a membership function, which represents the fuzzy behavior of the algorithm. To answer that,

the algorithm has to construct an appropriate matrix named W_{IJ} whose factors are numbers between 0 and 1, and represent the degree of membership between data and the cores of clusters. The operation of the three clustering algorithms has been examined on the basis of clustering quality and efficiency of the algorithms.

4. Experimental Results

The experiments carried out in this work divided into three parts via preprocessing, clustering and classification as discussed. Based on this notion, the experimental results are analysed.

4.1. Result of Preprocessing Techniques

The main objective of the preprocessing is to develop the image quality to make it ready for further processing by removing or reducing the not related and surplus parts in the background of the mammogram image's pixels. The research work analysis three type input mammogram breast cancer images. The common characteristics of the breast cancer images like as unknown noise, poor image contrast, weak boundaries and unrelated parts will affect the content of the breast cancer images. In the preprocessing, first the noise can be removed using the median filter method in **Figure 3** shows the result normal, benign and malignant breast cancer images [21] [22].

The results of Gaussian filter shows in **Figure 4** which shows normal images and identify no affected cancer, benign image is affected cancer early stages and malignant image is cancer, abnormal stage in tumor human body spread any part of organs. Mammogram image enhancement is the process of manipulation of images by reducing noise and increase the image contrast in order to detect the abnormalities [22]. The aim of enhancing mammograms is to eliminate the background noise and improve the image quality for the purpose of determining the region of Interest (ROIs) in the image thereby making it easier for the Radiologist to read and interpret. The underlying principle in the enhancement of mammogram is to enlarge the intensity difference between objects and background and to produce reliable representations of breast tissue structures [23].

| (a) | (b) | (c) |
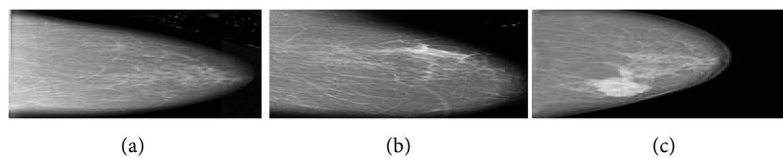
Figure 3. Result of Median filter (a) Normal; (b) Benign; (c) Malignant.


| (a) | (b) | (c) |
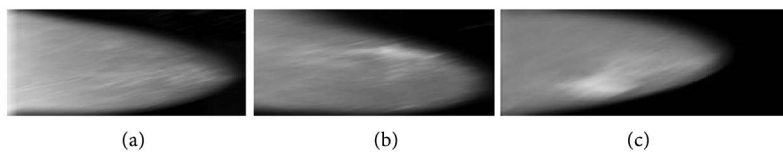
Figure 4. Result of Gaussian filter (a) Normal; (b) Benign; (c) Malignant.

The main objective of this procedure is to develop the quality of the image, to make it ready for further processing. This process was done by using ROI, inverse method and the boundary detection method respectively. The ROI method finds the areas of images based its intensity. The detection of the ROI consists in finding a region of the image which appears different from the background with respect to low-level features such as contrast, color, region size and shape, distribution of contours or texture pattern. Different methods have been proposed to detect regions of interest in an image. The pixel having highest intensity value in the digital image is chosen, then that pixel is compared to the neighboring pixels. The comparison goes on till there is a modification in the pixel value [10]. The inverse method uses to abnormal area inverse to image. The boundary detection method is used to remove unwanted areas and taking only breast regions. ROI, inverse method and boundary method in **Figure 5** shows the result of normal images, benign image and malignant breast cancer. The preprocessing is carried out by before image pixels and after preprocessing image pixel difference, then preprocessing after original image memory space and before preprocessing original image memory space difference analysis. **Table 1** shows the result of normal image, result of benign breast cancer image and result of malignant image. The BP means before preprocessing image pixels, AP after preprocessing image pixels, BM means before preprocessing original image memory space and AM means after preprocessing original image memory space [24].

Table 1. Result of preprocessing breast normal images.

Image No	BP	AP	Difference	Average (BP, AP)	BM (KB)	AM (KB)	Difference (KB)	Average (KB)
1	4200	4086	114	4143	15.8	14.8	1.0	15.3
2	7130	6973	157	7051.5	9.27	8.2	1.25	8.7
3	9745	9619	126	9682	9.13	8.12	1.01	8.6

Result of Benign Breast Cancer Images

Image No	BP	AP	Difference	Average (BP, AP)	BM (KB)	AM (KB)	Difference (KB)	Average (KB)
4	13,800	13,694	106	13,747	28.0	22.01	6.04	14.025
5	11,330	11,179	151	11,254.5	24.04	18.02	6.02	12.02
6	13,208	13,060	148	13,134	20.06	18.02	2.04	10.03

Result of Malignant Breast Cancer Images

Image No	BP	AP	Difference	Average (BP,AP)	BM (KB)	AM (KB)	Difference (KB)	Average (KB)
7	13,748	13,548	200	13,648	33.06	30.1	3.05	31.58
8	24,720	24,620	100	24,670	38.0	34.1	4.0	36.05
9	14,700	14,505	195	14,602.5	32.07	30.1	2.06	31.085
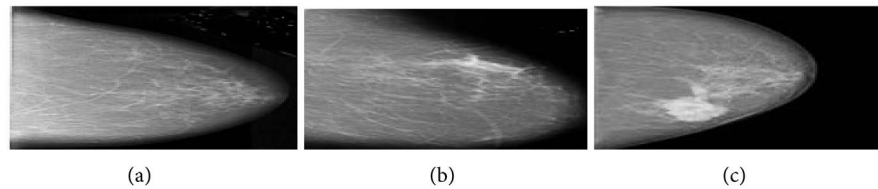
Figure 5. Result of ROI, Inverse method and boundary detection method (a) Normal; (b) Benign; (c) Malignant.

4.2. Results and Discussion

The segmentation of images by the k-Means, FCM and MCA clustering algorithms were carried out in this work to find the tumor affected regions in the mammogram images. During the process of clustering the images, **Figures 6-8** normal images were identified without any abnormal portions in the image. Single color (black color only) is found in the normal images after the clustering process. **Figures 6-8** shows the result of benign and the malignant images, the white color pixels were identified by both k-Means and FCM algorithms in the 5th cluster and by MCA algorithm in the 4th cluster itself. The performance of three algorithms were measured by using the parameters **Tables 2-5** like run time, volume complexity and clustering quality. **Table 5** the processing time (given in milli seconds) taken for clustering normal image with k-Means algorithm was 1544 ms, FCM algorithm was 1244 ms and MCA was 1044 ms. The processing time taken for clustering benign image with k-Means algorithm was 2540 ms, FCM was 2040 ms and MCA was 1040 ms. Finally, the processing time taken for clustering malignant image with k-Means algorithm was 2068 ms, FCM was 2028 ms and MCA was 1468 ms.

Therefore, it is evident that the time taken for analyzing the images by MCA was less than the k-Means and FCM algorithms. **Table 5** shows result memory space utilized for clustering normal image with k-Means algorithm was 7.02 KB, FCM algorithm was 2.02 KB and MCA was 1.02 KB. The memory space utilized for clustering benign image with k-Means algorithm was 16.01 KB, FCM was 10.01 KB and MCA was 9.01 KB. The memory space utilized for clustering malignant image with k-Means algorithm was 24.01 KB, FCM was 14.01 KB and MCA was 09.01 KB. Therefore the space utilized for analyzing the images by using MCA was less than the k-Means and FCM algorithms.

The proposed MCA produced better results in the clustering process with high performance, less execution time and occupied less space. For verification of the results, this work used classification algorithms such as CART, J48, JRIP, SVM and Naive Bayes. **Table 6** shows various performance measures like FP rate, TP rate, Recall, Precision, ROC Area and F-measure were used in this work to measure accuracy of the algorithms. In the implementation process, numerical values of some attributes in the breast cancer data were considered. Error report was generated for all the five classification algorithms considering errors such as kappa statistic, mean absolute error, root mean squared error, relative absolute error and root relative squared error in percentage using breast cancer
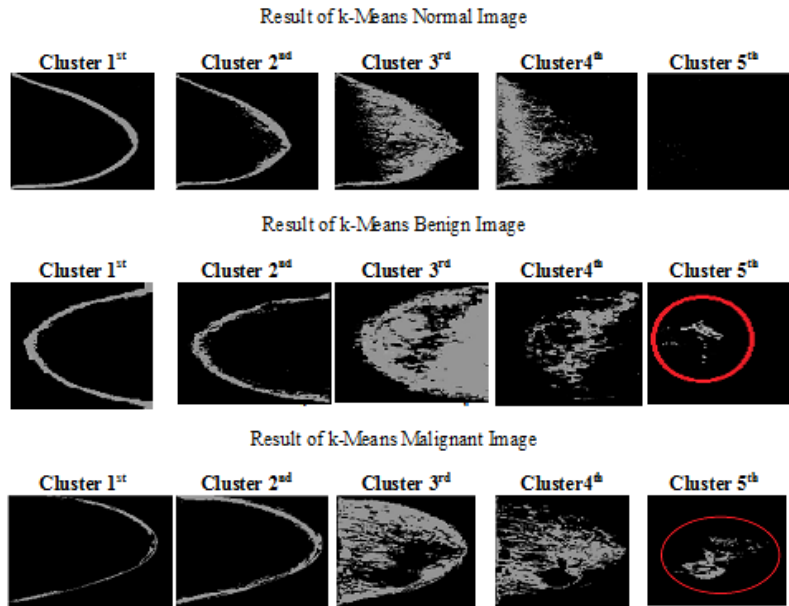
Result of k-Means Normal Image

Result of k-Means Benign Image

Result of k-Means Malignant Image

Figure 6. Result of k-means clustering algorithm.

Result of FCM Normal Image

Result of FCM Benign Image
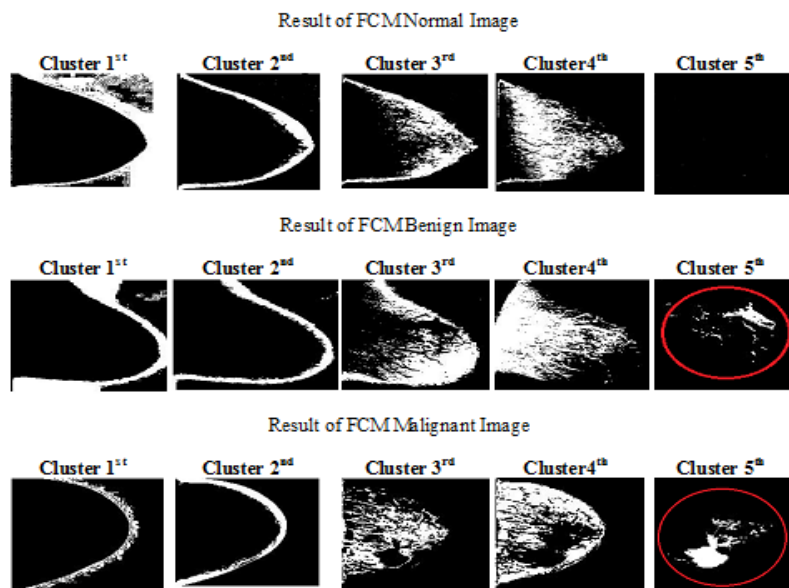

Result of FCM Malignant Image

Figure 7. Result of FCM clustering algorithm.

Table 2. Results of k-means algorithm in normal breast images.

Image No	C1		C2		C3		C4		C5		Total Pixels
	W	B	W	B	W	B	W	B	W	B	
1	1234	2852	1400	2686	1200	2886	244	3842	8	4078	4086
Results of k-Means Algorithm in Benign breast images											
2	2345	11,349	1345	12,349	8456	5238	1300	12,394	248	13,446	13,694
Results of k-Means Algorithm in Malignant breast images											
3	891	12,657	2449	11,099	8762	4786	1219	12,529	327	13,221	13,548

T. Velmurugan, E. Venkatesan
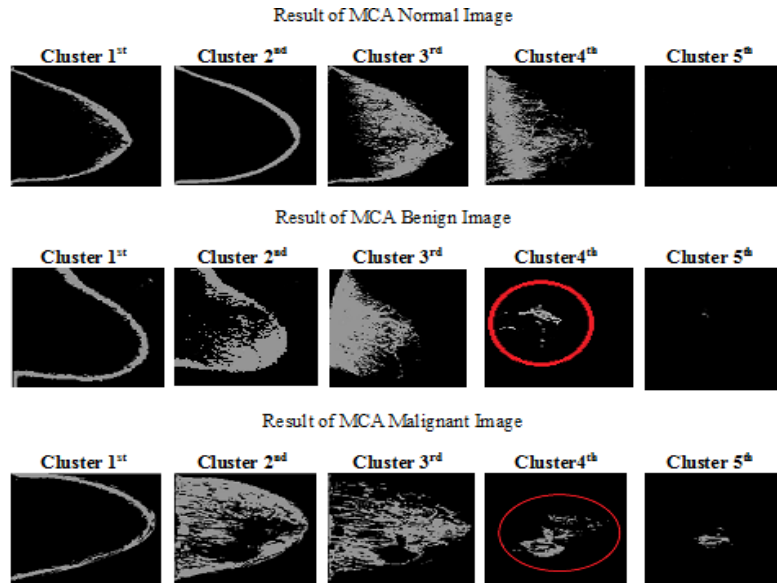
Result of MCA Normal Image

Figure 8. Result of MCA clustering algorithm.

Table 3. Results of FCM Algorithm in normal breast images.

Image No	C1		C2		C3		C4		C5		Total Pixels
	W	B	W	B	W	B	W	B	W	B	
1	1334	2752	1500	2586	1200	2886	244	3842	8	4078	4086
Results of FCM Algorithm in Benign breast images											
2	2145	11,449	1245	12,449	8156	5338	1200	1394	248	13,446	13,694
Results of FCM Algorithm in Malignant breast images											
3	991	12,557	2349	11,199	8862	4686	1119	12,429	227	13,321	13,548

Table 4. Results of MCAAlgorithm in Normal breast images.

Image No	C1		C2		C3		C4		C5		Total Pixels
	W	B	W	B	W	B	W	B	W	B	
1	1534	2552	1500	2586	1200	2886	244	3842	8	4078	4086
Results of MCA Algorithm in Benign breast images											
2	3694	10,000	1694	12,000	8694	5000	200	13,494	148	13,546	13,694
Results of MCA Algorithm in Malignant breast images											
3	2991	10,557	1349	12,199	7862	5686	2119	11,429	327	13,221	13,548

Table 5. Result of clustering algorithm performances.

Result of k-means Algorithm					
Normal Image		Benign Image		Malignant Image	
RunTime(ms)	Memory(KB)	RunTime(ms)	Memory(KB)	RunTime(ms)	Memory(KB)
1544	7.2	2540	16.1	2068	24.1
Result of FCM Algorithm					
1244	2.2	2040	10.1	2028	14.1
Result of MCA Algorithm					
1044	1.2	1010	4.1	1168	7.1

Table 6. Result of classification algorithm (performance).

Predictive Model	Accuracy	Sensitivity	Specificity	Time in ms
J48	75.58	75.5853	24.4147	135
CART	92.30	92.3077	7.6923	186
JRIP	82.60	82.6087	17.3913	201
SVM	87.95	87.9599	12.0401	225
Naïve Bayes	84.28	84.2809	15.7191	308

data by varying the statistic rate. It is observed that kappa statistic, mean absolute error and root mean squared error are almost negligible for all the five classification algorithms. For SVM classification algorithm alone relative absolute error is negligible, whereas for the other four algorithms the relative absolute error is greater than 96% in which Naive Bayes algorithm has the highest relative absolute error rate of 152%. The root relative squared error is greater than 99% for all the five classification algorithms, in which Naive Bayes algorithm has the highest root relative squared error rate of 121%.

The performance of all the five classification algorithms were analysed based on its Sensitivity, Specificity and Accuracy. CART classifier has the highest sensitivity of 92%, J48 has the lowest Sensitivity of 75%, wheres JRIP, SVM and Naïve Bayes have sensitivity of 82%, 87% and 84%, respectively. J48 has highest the specificity of 24%, CART classifier has the lowest Specificity of 7%, whereas JRIP, SVM and Naïve Bayes have specificity of 17%, 12% and 15% respectively. The outcomes indicate that CART classifier has highest accuracy of 92.30%, SVM algorithm has second highest accuracy of 87.95%, Naive Bayes algorithm has accuracy of 84.28%, JRIP algorithm has accuracy of 82.60%, whereas the J48 algorithm has least accuracy of 75.58%. **Figure 9** shows that the performance of algorithms based on time in milliseconds. **Figure 10** depicts the performance by memory.

Figure 11 shows that the predicted values for accuracy, sensitivity, specificity and time with respect to the classification algorithms using the breast cancer data. **Figure 12** shows the run time in milliseconds. Among the choice of classification algorithms, the performance of CART is better than the other algorithms best for the selected data and also ensures that quality of the clustering algorithms. Totally, 310 mammogram images are taken for the analysis and segmented by the three clustering algorithms based on its pixel intensity. Initially, the number of clusters was given as 5 or more. Before applying the three clustering algorithms, the images were preprocessed by the preprocessing methods median filter, Gaussian method, ROI, inverse method and boundary detection method. The cancer affected region was identified in the 5th cluster by k-Means algorithm and by FCM also. But, in the 4th cluster itself the MCA algorithms were successfully identified the region very effectively. The time complexity and volume complexity was also very less in MCA compared with the other two algorithms.
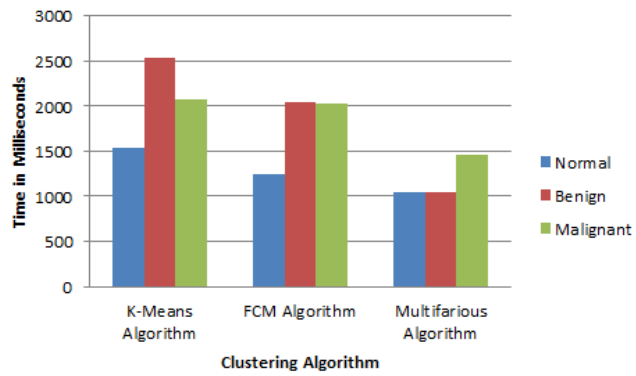
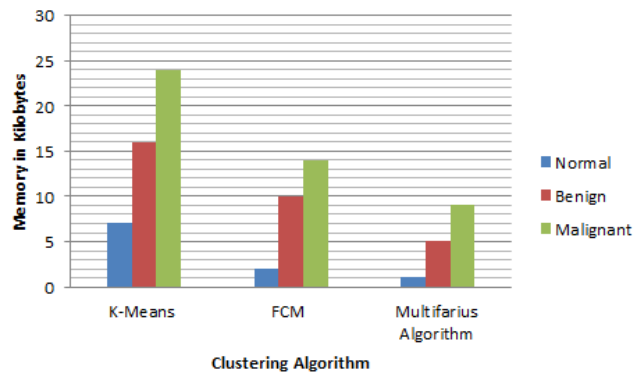**Figure 9.** Performance of clustering algorithms (Run Time based).

Figure 10. Performance of clustering algorithms (Memory based).
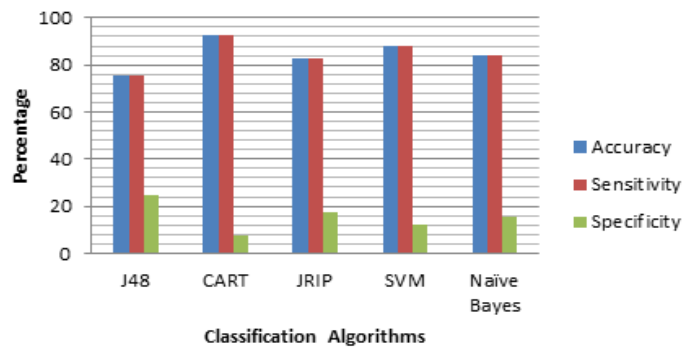

Figure 11. Performance of classification algorithms.
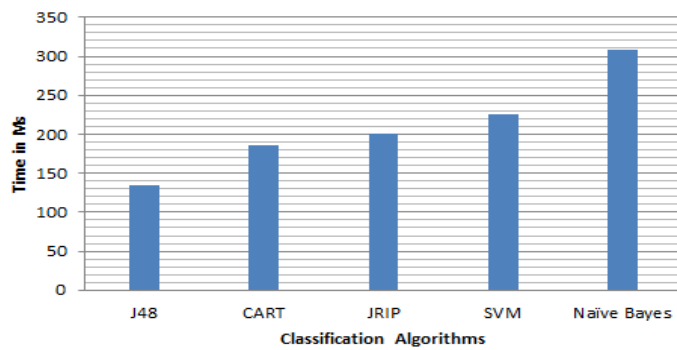

Figure 12. Run time in milliseconds.

5. Conclusion

Generally, no one can say that a particular algorithm is the best algorithm for the prediction purposes using any kind of real world data set for some applications. But, it is possible to suggest the performance of algorithms for the chosen data. Based on this clear notion, the performance of two partitioning based clustering algorithms k-Means and FCM were compared with the proposed method MCA. The results were analyzed by several executions of the programs. After enormous comparisons of all three algorithms, it is concluded that the performance of MCA algorithm was better than the other two algorithms. The novel, the multifarious clustering algorithm identifies the cancer affected regions very effectively and efficiently. The accuracy was verified by classification algorithms J48, JRIP, SVM, Naive Bayes and CART algorithms by its various performance measures. Within the classification algorithms, the accuracy of CART was found to be better than the remaining algorithms. The images analyzed by k-Means, FCM and MCA helped to detect the breast cancer affected area by detecting tumor in the images. The method of analysis and the result of this research work was accepted after due consultation with a physician. The future work in this area involves the use of image segmentation methods and other types of clustering algorithms.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Abdin, Z., Zaheeruddin, J. and Singh, L. (2013) Performance Analysis of Image Segmentation Methods for the Detection of Masses in Mammograms. *Performance Analysis of Image Segmentation Methods for the Detection of Masses in Mammograms*, **82**, 44-50. https://doi.org/10.5120/14092-2100

[2] Karmilasari, Widodo, S., Hermita, M. and Lussiana, E.T.P. (2014) Sample K-Means Clustering Method for Determining the Stage of Breast Cancer Malignancy Based on Cancer Size on Mammogram Image Basis. **5**, 86-90. https://doi.org/10.14569/IJACSA.2014.050312

[3] Pokar, H.N. and Patel, P.H. (2011) Survey on Different Techniques Used for Detection of Malignancy in Mammograms of Breast Cancer. *International Journal of Advance Engineering and Research Development*, **2**, 656-664.

[4] Laffont, V., Durupt, F., Birgen, M.A., Bauduin, S. and Laine, A.F. (2001) Detection of Masses in Mammography through Redundant Expansion of Scales. *The Proceedings of the 23rd Annual International Conference EMBS (Engineering in Medicine and Biology Society)*, 2797-2800. https://doi.org/10.1109/IEMBS.2001.1017366

[5] Akila, K. and Sumathy, P. (2015) Early Breast Cancer Tumor Detection on Mammogram Images. *International Journal Computer Science, Engineering and Technology*, **5**, 334-336.

[6] Joshi, S. and Priyanka Shetty, S.R. (2015) Performance Analysis of Different Classification Methods in Data Mining for Diabetes Dataset Using WEKA Tool. *International Journal on Recent and Innovation Trends in Computing and Communication*, **3**, 1168-1173. https://doi.org/10.17762/ijritcc2321-8169.150361

[7]   Mei, Y.C. (2011) High-Speed Volumetric *in Vivo* Medical Imaging for Morphometric Analysis of the Human Optic Nerve Head. PhD Dissertation, Applied Science: School of Engineering Science and Simonfraser University.

[8]   Senthilkumar, B. and Umamaheswari, G. (2013) Combination of Novel Enhancement Technique and Fuzzy C Means Clustering Technique in Breast Cancer Detection. *Biomedical Research*, **24**, 252-256.

[9]   Ramani, R., Valarmathy, S. and Suthanthira Vanitha, N. (2013) Breast Cancer Detection in Mammograms Based on Clustering Techniques: A Survey. *International Journal of Computer Applications*, **62**, 17-21. https://doi.org/10.5120/10123-4885

[10] Naveen, A. and Velmurugan, T. (2016) A Novel Layer Based Logical Approach (LLA) Clustering Method for Performance Analysis in Medical Images. *International Journal of Control Theory and Applications*, **9**, 4647-4660.

[11] Aarth, S.L. and Prabu, S. (2016) A Computerized Approach on Breast Cancer Detection and Classification. *Institute of Integrative Omics and Applied Biotechnology Journal*, **7**, 157-169.

[12] Wahdan, P., Saad, A. and Shoukry, A. (2016) Automated Breast Tumor Detection in Ultrasound Images Using Support Vector Machine and Ensemble Classification. *Journal of Biomedical Engineering and Biosciences*, **3**, 4-11.

[13] Kambo, R. and Amit, Y. (2014) Principal Component Analysis Based Classification Technique for Basmati Rice Grain Analysis. *International Journal of Science and Research*, **3**, 1764-1769.

[14] Kaur, G. and Chhabra, A. (2014) Improved J48 Classification Algorithm for the Prediction of Diabetes. *International Journal of Computer Applications*, **98**, 13-17. https://doi.org/10.5120/17314-7433

[15] Velmurugan, T. and Santhanam, T. (2010) Computational Complexity between K-Means and K-Medoids Clustering Algorithms for Normal and Uniform Distributions of Data Points. *Journal of Computer Science*, **6**, 363-368. https://doi.org/10.3844/jcssp.2010.363.368

[16] Kambo, R. and Amit, Y. (2014) Classification of Basmati Rice Grain Variety Using Image Processing and Principal Component Analysis. *International Journal of Computer Trends and Technology*, **11**, 80-85. https://doi.org/10.14445/22312803/IJCTT-V11P117

[17] Wun, L.M., Merril, R.M. and Feuer, E.J. (1998) Estimating Lifetime and Age-Conditional Probabilities of Developing Cancer. *Lifetime Data Analysis*, **4**, 169-186. https://doi.org/10.1023/A:1009685507602

[18] Velmurugan, T. (2014) Performance Based Analysis between K-Means and Fuzzy C-Means Clustering Algorithms for Connection Oriented Telecommunication Data. *Applied Soft Computing*, **19**, 134-146. https://doi.org/10.1016/j.asoc.2014.02.011

[19] Amygdalos, I. (2014) Detection and Classification of Gastrointestinal Cancer and Other Pathologies through Quantitative Analysis of Optical Coherence Tomography Data and Goniophotometry. PhD Dissertation, Department of Medicine, Imperial College, London.

[20] Velmurugan, T. and Venkatesan, E. (2016) Effective Fuzzy C Means Algorithm for the Segmentation of Mammogram Images of Identify Breast Cancer. *International Journal of Control Theory and Applications*, **9**, 4647-4660.

[21] Velmurugan, T. (2014) Performance Based Analysis between K-Means and Fuzzy C-Means Clustering Algorithms for Connection Oriented Telecommunication Data. *Applied Soft Computing*, **19**, 134-146. https://doi.org/10.1016/j.asoc.2014.02.011

[22] Cholavendhan, S., Siva, K. and Karnan (2014) A Research on Various Filtering

Techniques in Enhancing Mammogram Image Segmentation. *International Journal of Engineering Trends and Technology*, **9**, 451-453.

[23] Kayode, A.A., Afolabi, B.S. and Ibitoye, B.O. (2015) An Explorative Survey of Image Enhancement Techniques Used in Mammography. *International Journal of Computer Science Issues*, **12**, 72-79.

[24] Kashyap, K.L., Bajpai, M.K. and Khanna, P. (2015) Breast Cancer Detection in Digital Mammograms. *IEEE International Conference on Imaging Systems and Techniques*, 1-6. https://doi.org/10.1109/IST.2015.7294523