International Journal of Research Studies in Computer Science and Engineering (IJRSCSE)
Volume. 1, Issue 4, August 2014, PP 27-31
ISSN 2349-4840 (Print) & ISSN 2349-4859 (Online)
www.arcjournals.org

Evaluation of Breast Percent Density in Digital Mammography Images via Fuzzy C-Means Clustering and Support Vector Machine

Sangeetha. N.M
M.Phil Research Scholar, Dept of Computer Science, Quaid-E- Millath Govt. College for Women (A)
sangeethasainath@gmail.com

Dr.(Mrs.).Pugazhenthi
Asst.Professor- p.g & Research Dept of Computer Science, Quaid–E- Millath Govt.College for Women
Anna salai, Chennai

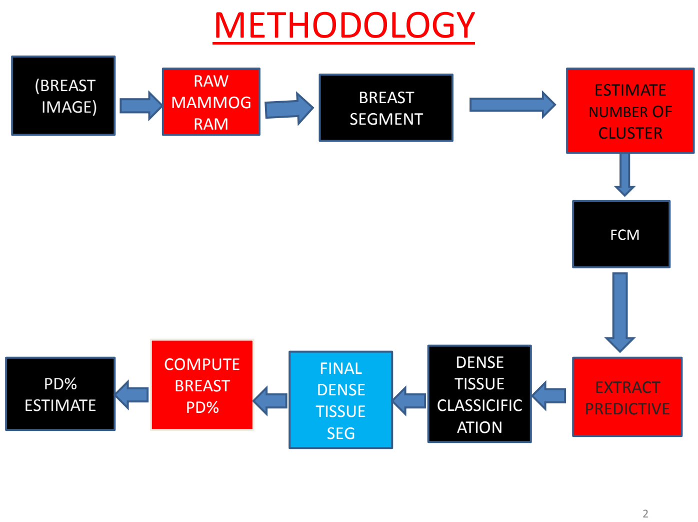
Abstract: *Breast cancer is detected and prevention is also adopted and care by the technological solutions. The human health care processes are advanced with the modern technology and computation. Estimation of breast percent density and stages of cancer predicated based on the level of affected density using mammography density of the breast predicated based on the fuzzy c means clustering and Support vector machine*

Keywords: *Digital mammography, Density determination, clustering*

1. Introduction

Mammogram is the special type of x-ray. It is effective, low cost and efficient method to detect breast cancer early. Digital image consists of discrete picture elements called pixels which can be associated with digital number represented as DM that depicts the average radiance of relatively small area within a scene. The quantity of fibro glandular tissue content in the breast as estimated mammographic ally commonly referred to as breast percent density (PD %), is one of the most significant risk factor for developing breast cancer.[1].Furthermore, most studies Published to date investigating computer-aided assessment of breast PD% have been Performed using digitized screen-film mammograms, while digital mammography is increasingly replacing screen film mammography in breast cancer screening protocols. Digital mammography imaging generates two types of images for analysis, raw (i.e., “FOR PROCESSING”) and vendor post processed (i.e., “FOR PRESENTATION”), of which post processed images are commonly used in clinical practice.[2]. Development of an algorithm which effectively estimates breast PD% in mammography images.

2. Methodology



3. Data Set

The data for estimating the breast percent density is digital mammogram images. The mammogram images should be 12-to- 14 gray levels depth.

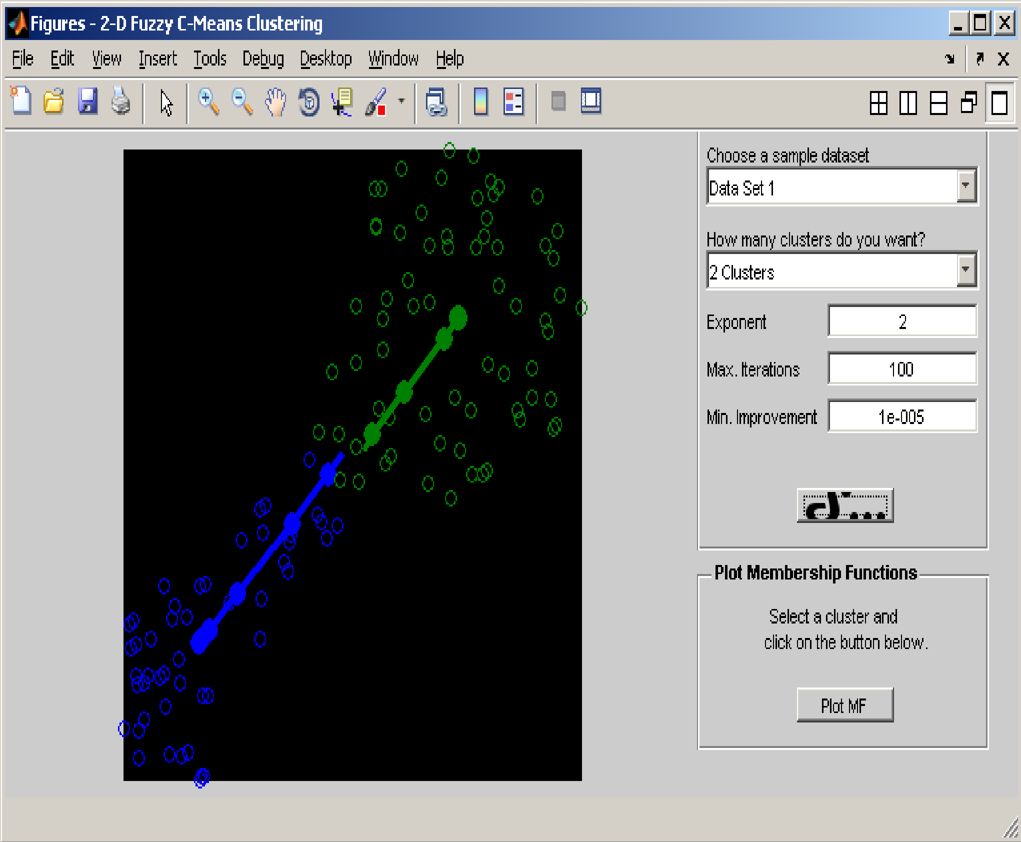
4. Image Preprocessing

Image is collected from on line and then preprocessing to be done. Preprocessing is always a necessity whenever the data to be mind in noisy and incomplete and preprocessing improves the effectiveness of the data mining techniques. The applied the techniques to the images, which is called by the name cropping.

The cropping operation removes the unwanted parts of the digital image (i.e) this proposed method using initial level of digital mammogram (mammogram with cancer tissues).

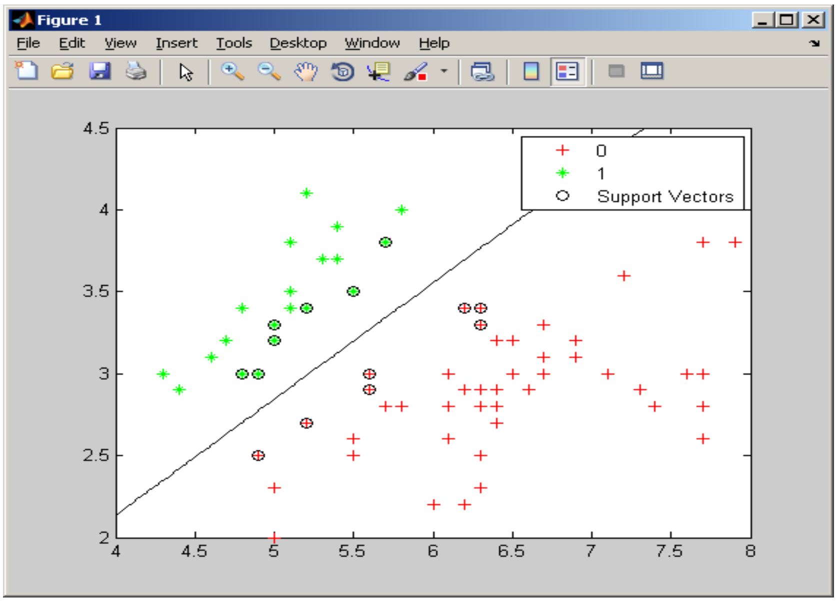
5. Fuzzy Clustering

Clustering of numerical data forms the basis of many classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior. The Fuzzy Logic is equipped with some tools that allow you to find clusters in input-output training data. You can use the cluster information to generate a Sugeno-type fuzzy inference system that best models the data behavior using a minimum number of rules. The rules partition themselves according to the fuzzy qualities associated with each of the data clusters. This type of FIS generation can be accomplished automatically using the command line function, fcm.



6. SVM Classifiers

The SVM is predictor variable is called an attribute, and a transformed attribute that is used to define the hyper plane is called a feature. The task of choosing the most suitable representation is known as feature selection. A set of features that describes one case (i.e., a row of predictor values) is called a vector. So the goal of SVM modeling is to find the optimal hyper plane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyper plane are the support vectors.



7. Analysis and Implementation

7.1. Standard Dataset

The standard dataset which is used for evaluating the breast percent density is shown below:

a) Calculate the total no of pixels in each image.

b) Selection of affected image from the segmented image which is high density pixels and then identify the non-zero element combination of set value to determine the percentile of affected particles in particular image.

C) Calculate the sum in corresponding to non-zero elements which is occurred in the segments sub image for each and every high resolution images.

Breast Density %				Fibrogland. Tissue Volume (CM3)				Breast Tissue Volume (CM3)				Compression Thickness			
LCC	RCC	LMLO	RMLO	LCC	RCC	LMLO	RMLO	LCC	RCC	LMLO	RMLO	LCC	RCC	LMLO	RMLO
24.4	27.0	14.1	26.1	83.0	79.9	39.5	66.7	340.8	295.7	280.2	256.0	55.0	50.0	48.0	47.0
25.9	35.0	16.1	26.4	81.2	96.0	38.2	70.2	313.6	274.5	237.4	265.6	56.0	54.0	46.0	47.0
30.3	33.5	14.6	23.8	85.5	83.5	33.7	57.6	282.4	249.7	231.0	242.3	47.0	49.0	41.0	42.0
16.7	13.1	7.5	15.3	41.2	29.8	16.3	34.6	246.4	228.1	218.4	226.3	38.0	38.0	38.0	35.0
15.6	17.7	13.5	12.6	28.6	30.5	25.6	27.4	182.8	172.6	189.8	217.2	27.0	28.0	29.0	34.0
12.0	12.3	9.8	12.1	21.8	29.0	19.4	25.6	181.5	235.9	196.6	212.0	30.0	36.0	33.0	33.0
14.0	12.9	12.3	11.5	92.7	77.6	91.7	73.5	661.7	603.0	747.1	641.1	67.0	69.0	69.0	65.0
14.9	13.0	13.5	13.9	95.4	74.2	99.5	87.9	640.3	569.5	736.4	630.3	64.0	66.0	64.0	65.0
10.8	10.3	12.9	11.7	89.6	68.6	102.3	81.6	832.1	667.9	790.6	697.2	73.0	67.0	66.0	66.0
10.9	20.4	13.3	11.1	41.1	61.7	57.7	43.2	378.6	302.8	434.1	389.7	48.0	44.0	42.0	42.0
15.7	16.6	12.7	13.2	65.5	58.8	49.5	47.7	417.0	354.6	389.2	362.6	44.0	42.0	40.0	41.0
12.8	17.4	13.5	16.5	49.4	60.4	52.3	55.9	385.6	347.3	386.8	339.0	45.0	42.0	41.0	39.0
24.0	24.2	12.6	12.3	87.9	80.2	52.0	46.3	366.4	331.1	412.1	376.4	38.0	40.0	37.0	37.0
26.6	27.6	14.9	16.8	105.3	95.2	64.3	61.6	396.2	344.8	432.0	366.2	37.0	36.0	34.0	33.0
23.2	19.9	17.5	24.0	117.1	83.8	85.8	97.0	503.7	422.2	489.9	404.4	44.0	41.0	38.0	35.0
15.3	14.5	25.0	15.3	53.1	53.6	91.9	66.4	347.2	369.5	368.1	432.7	64.0	61.0	59.0	60.0
33.2	17.9	25.9	23.9	84.6	62.8	70.9	99.9	254.8	351.9	274.3	417.7	44.0	49.0	46.0	52.0
22.8	24.9	28.6	23.8	80.3	85.5	105.9	103.0	352.4	343.5	369.9	432.7	55.0	52.0	53.0	55.0
7.5	8.3	6.8	8.9	116.3	115.6	113.3	138.7	1548.8	1390.8	1674.3	1564.3	73.0	69.0	76.0	75.0
6.5	9.1	7.2	10.5	94.7	116.9	113.9	151.4	1452.1	1285.0	1591.3	1445.2	73.0	65.0	73.0	70.0
6.0	7.3	7.6	9.0	99.1	108.2	129.1	144.3	1657.8	1483.4	1709.1	1605.7	85.0	77.0	78.0	76.0
5.2	5.4	5.8	5.4	64.4	46.6	69.5	50.2	1242.1	857.0	1189.1	928.7	54.0	48.0	57.0	51.0
3.1	3.5	4.3	4.1	50.3	41.0	80.7	54.1	1624.2	1161.8	1875.5	1325.3	68.0	66.0	75.0	64.0

7.2. Calculation Process

The formula for calculating the breast percent density is

a) For each image which is having high density for each image property.
b) Find how many times the same clustered images occurs (i.e) frequency.
c) Calculate total density and average density for each clustered image.
d) Calculate level of affected using the following formula:

$$PD\% = MD/MB*100\%$$

Where MD=Dense Tissue, MB=Breast Tissue

7.3. Stages of Breast Cancer

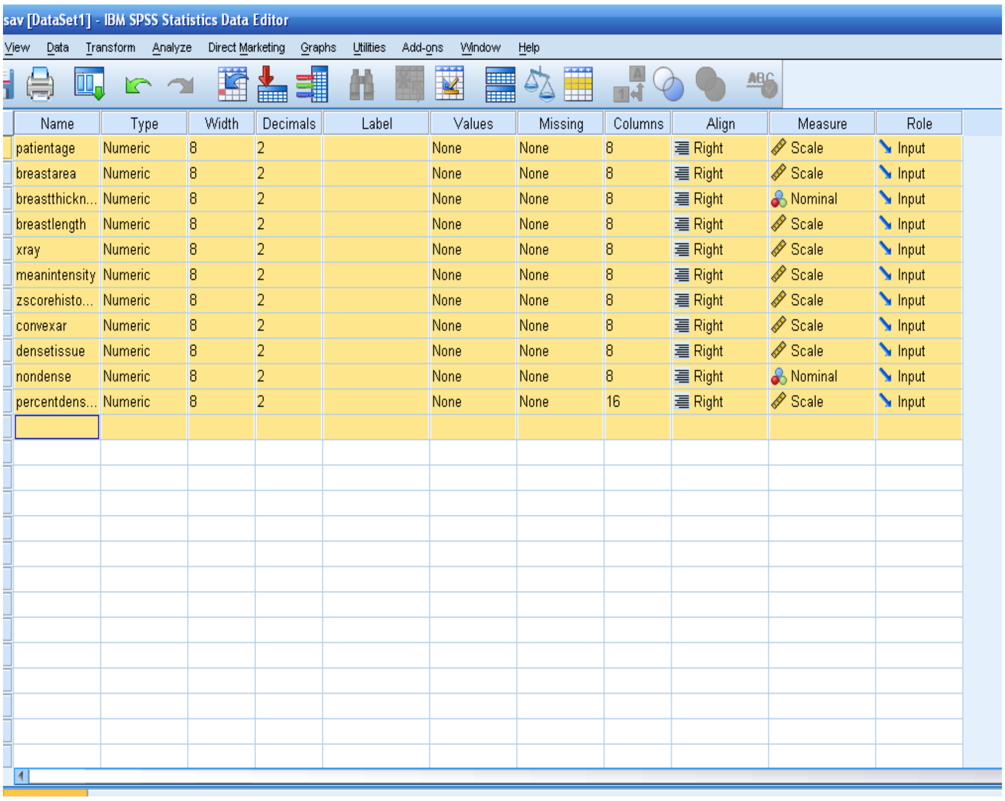
After calculating the common index value can find the stages of the cancer. There are 5 stages, which from 0-4.

a) If the pd=0-20 then stage=0.
b) If the pd=21-40 then stage=1.
c) If the pd=41-60 then stage=2.
d) If the pd=61-100 then stage=3.

The stage 0 is the type of vasive breast cancer. The remaining stages such as stage1, stage2, stage3 and stage 4 are invasive breast cancer.

7.4. Parameter Declaration

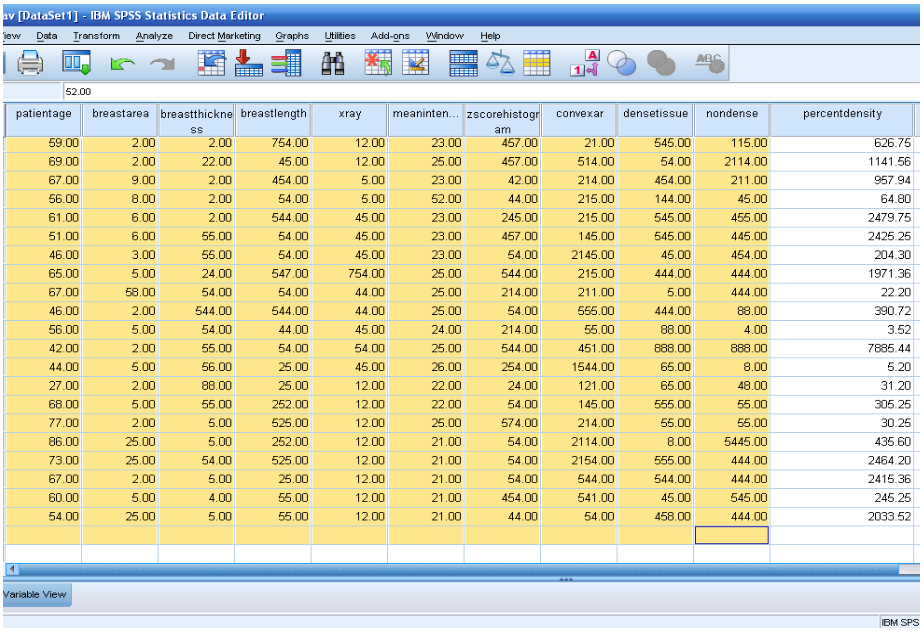
The parameter also called as attributes which involves for calculating the breast percent density method is:



Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
patientage	Numeric	8	2		None	None	8	Right	Scale	Input
breastarea	Numeric	8	2		None	None	8	Right	Scale	Input
breastthickn...	Numeric	8	2		None	None	8	Right	Nominal	Input
breastlength	Numeric	8	2		None	None	8	Right	Scale	Input
xray	Numeric	8	2		None	None	8	Right	Scale	Input
meanintensity	Numeric	8	2		None	None	8	Right	Scale	Input
zscorehisto...	Numeric	8	2		None	None	8	Right	Scale	Input
convexar	Numeric	8	2		None	None	8	Right	Scale	Input
densetissue	Numeric	8	2		None	None	8	Right	Scale	Input
nondense	Numeric	8	2		None	None	8	Right	Nominal	Input
percentdens...	Numeric	8	2		None	None	16	Right	Scale	Input

8. Result

The process of estimating breast Percent density as shown below: The rotated image can be adopted by the property rotate-90, rotate+90 and so on. From that average values can be calculated for each clustered image. The following table shows for each image property, average values affected density pixels. The following table shows for each image property, which clustered image is having high density pixels among the five clustered images.



patientage	breastarea	breastthickness	breastlength	xray	meaninten...	zscorehistogram	convexar	densetissue	nondense	percentdensity
59.00	2.00	2.00	754.00	12.00	23.00	457.00	21.00	545.00	115.00	626.75
69.00	2.00	22.00	45.00	12.00	25.00	457.00	514.00	54.00	2114.00	1141.56
67.00	9.00	2.00	454.00	5.00	23.00	42.00	214.00	454.00	211.00	957.94
56.00	8.00	2.00	54.00	5.00	52.00	44.00	215.00	144.00	45.00	64.80
61.00	6.00	2.00	544.00	45.00	23.00	245.00	215.00	545.00	455.00	2479.75
51.00	6.00	55.00	54.00	45.00	23.00	457.00	145.00	545.00	445.00	2425.25
46.00	3.00	55.00	54.00	45.00	23.00	54.00	2145.00	45.00	454.00	204.30
65.00	5.00	24.00	547.00	754.00	25.00	544.00	215.00	444.00	444.00	1971.36
67.00	58.00	54.00	54.00	44.00	25.00	214.00	211.00	5.00	444.00	22.20
46.00	2.00	544.00	544.00	44.00	25.00	54.00	555.00	444.00	88.00	390.72
56.00	5.00	54.00	44.00	45.00	24.00	214.00	55.00	88.00	4.00	3.52
42.00	2.00	55.00	54.00	54.00	25.00	544.00	451.00	888.00	888.00	7885.44
44.00	5.00	56.00	25.00	45.00	26.00	254.00	1544.00	65.00	8.00	5.20
27.00	2.00	88.00	25.00	12.00	22.00	24.00	121.00	65.00	48.00	31.20
68.00	5.00	55.00	252.00	12.00	22.00	54.00	145.00	555.00	55.00	305.25
77.00	2.00	5.00	525.00	12.00	25.00	574.00	214.00	55.00	55.00	30.25
86.00	25.00	5.00	252.00	12.00	21.00	54.00	2114.00	8.00	5445.00	435.60
73.00	25.00	54.00	525.00	12.00	21.00	54.00	2154.00	555.00	444.00	2464.20
67.00	2.00	5.00	25.00	12.00	21.00	54.00	544.00	544.00	444.00	2415.36
60.00	5.00	4.00	55.00	12.00	21.00	454.00	541.00	45.00	545.00	245.25
54.00	25.00	5.00	55.00	12.00	21.00	44.00	54.00	458.00	444.00	2033.52

9. Conclusion

In this paper, for preprocessing, cropping operation performed to remove unwanted parts and make all images in equal size. The c-means clustering algorithm is used to cluster the images. The presented a svm approach to find the percentage of affected area digital mammogram. This method can predict breast percent density based on patient age and breast tissue thickness.

References

[1] C. D. Lehman, J. D. Blume, P. Weatherall, D. Thickman, N. Hylton, E.Warner, E. Pisano, Gatsonis, M. Schnall, G. A. DeAngelis,P. Stomper, E. L. Rosen,M. O'Loughlin, S. Harms, and D. A. Bluemke, "Screening women at high risk for breast cancer with mammography and magnetic resonance imaging," Cancer 103, 1898–1905 (2005).Medical Physics, Vol. 39, No. 8, August 2012.

[2] M. G. Kallenberg, M. Lokate, C. H. van Gils, and N. Karssemeijer, "Automatic breast density segmentation: An integration of different approaches,"Phys. Med. Biol. 56, 2715–2729 (2011).

[3] J. A. Shepherd, K. Kerlikowske, L. Ma, F. Duewer, B. Fan, J. Wang,S. Malkov, E. Vittinghoff, and S. R. Cummings, "Volume of mammographic density and risk of breast cancer," Cancer Epidemiol. Biomarkers Prev. 20, 1473–1482 (2011).

[4] O. Alonzo-Proulx, N. Packard, J. M. Boone, A. Al-Mayah, K. K. Brock,S. Z. Shen, and M. J. Yaffe, "Validation of a method for measuring thevolumetric breast density from digital mammograms," Phys. Med. Biol.55, 3027–3044 (2010).

[5] Jemal, R. Siegel, J. Xu, and E. Ward, "Cancer statistics, 2010," Ca–Cancer J. Clin. 60, 277–300 (2010).

[6] N. Boyd, L. Martin, A. Gunasekara, O.Melnichouk, G. Maudsley, C. Peressotti,M. Yaffe, and S. Minkin, "Mammographic density and breast cancer risk: Evaluation of a novel method of measuring breast tissue volumes," Cancer Epidemiol. Biomarkers Prev. 18, 1754–1762 (2009).