A Hybrid Algorithm with Modified SVM and KNN for Classification of Mammogram Images using Medical Image Processing with Data Mining Techniques

V. Punithavathi[1]
Ph.D Research Scholar, Government Arts College,
Coimbatore, India

Dr. D. Devakumari[2]
Assistant Professor, PG and Research Department of Computer Science, Government Arts College, Coimbatore, Tamil Nadu, India

Abstract:
Digital mammography is most reliable and effective technique for early and accurate identification of Breast cancer. Image processing plays a significant role in diagnosis and classification of breast cancer in medical field. In this paper, a system is created to classify the mammogram images into three classes, namely Benign, Malignant and Normal. Mammogram images are pre-processed and the features are extracted from the segmented region. These features are used to train modified SVM and KNN classifier. The proposed Hybrid algorithm with modified SVM and KNN classifier helps to classify the mammogram images. This latest technique improves the SVM algorithm with introducing multi class for classification of breast cancer. It exploits the KNN algorithm according to the distribution of test images in a feature space. This study also evaluates the accuracy with the SVM and KNN classifier. The modified SVM and KNN hybrid algorithm produces higher prognosis accuracy than the KNN method and SVM technique. This method is tested for 10 test images with 20 trained. This methodology achieves an overall mean accuracy of 99.3406% in classification of mammogram images.
Keywords: Classification, KNN, MIAS, Proposed KNN with SVM.

I. INTRODUCTION

Breast cancer is a tumor that forms in the cells of the breast. It is the most familiar non skin tumor in women and the second leading disease caused in female [1]. Breast cancer endurance rates have greater than before, and the number of deaths associated with this disease is gradually waning, largely due to factors such as earlier detection of tumor. Breast tumors and masses usually appear in the form of dense regions in mammograms. A typical benign mass has a round, smooth and well constrained boundary; on the other hand, a malignant tumor usually has a postulate, rough, and blurred boundary [2], [3].

Early identification of cancer is important for a fast reply and better chances of treatment. Unfortunately, early identification of cancer is often hard since the symptoms of the disease at the beginning are not present. Thus, cancer remains one of the topics of health research, where many researchers have provided with the aspire of creating proof that can develop treatment, precaution and diagnostics.

In machine learning there are two types: the supervised and unsupervised learning. Declare the classes first to classify the data that are known in earlier and the second, the classes are not known. Among the machine learning methods, there are: Support Vector Machines, Decision Tree, Neural Network, Bayesian networks, k-nearest neighbors and so on.,

The k-nearest neighbor algorithm is widely used in data classification [4]. The KNN allows the classification of a new component by calculating its distance from the other entire component. The suitable functioning of the system depends

s

on the choice of the parameter k which represents the number of neighbors chosen to assign the class to the new element and the choice of the distance.

Support Vector Machine is a machine learning technique used as a tool for classification of data, function estimation, and so on due to its generalization ability and has found success in many applications [5]. Feature of Support Vector Machine is that it minimizes and upper bound of generalization error through maximizing the margin between separating hyper plane and dataset. Support Vector Machine has an extra benefit of usual model selection in the logic that both the finest number and locations of the basic functions are mechanically obtained through training. The performance of SVM mainly depends on the kernel [6].

The digital mammogram image classification method is discussed in this paper. Figure 1 demonstrates the image classification method. Generally mammogram images undergo image Pre-processing, Segmentation and so on.

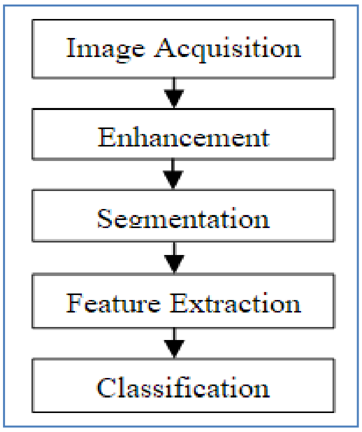

Fig. 1: Image Categorization procedure

In this proposed paper a modified SVM with KNN classification algorithm is implemented with KNN and SVM classification technique. In this paper detailed related work is presented in Section. 2. Proposed Methodology is offered in Section 3; Experimental Results are discussed in Section 4 and in Section 5 Conclusion is mentioned.

II. RELATED WORK

Marcano-Cedeno et.al., **2011** has discussed a new improvement for pattern classification using neural network training. The author proposed training algorithm that has inspired by the biological metaplasticity property of neurons and Shannon's information theory. During the training phase the Artificial Metaplasticity Multilayer Perceptron (AMMLP) algorithm gives precedence to updating the weights for the less frequent stimulation over the more frequent ones [7]. In this way metaplasticity is modeled artificially. AMMLP has achieved most efficient training, while maintaining MLP performance. The author used Wisconsin Breast Cancer Database (WBCD) to test the proposed algorithm.

Buciu, I. and Gacsadi, A, **2011** has been proposed an approach to deal with the classification of digital mammograms. Blotches around tumors were manually taken out to divide the abnormal areas from the remaining of the image, well thought-out as background. Gabor wavelets are used to filter the mammogram images. The features are extracted at different direction and frequencies [8]. Principal Component Analysis (PCA) has engaged to reduce the dimension of filtered and unfiltered high-dimensional data. Finally, Proximal Support Vector Machines were used to categorize the data. Superior mammogram image classification performance was attained when Gabor features were extracted instead of using original mammogram images.

Wang, et.al., **2016** has been improved the diagnostic accuracy of micro-calcification. This paper calculates the performance of deep learning-based models on large datasets for its distinction. All micro-calcifications are characterized using a semi-automated segmentation method. A distinction classifier model was designed to evaluate the accuracies of micro-calcifications and breast masses. It can moreover used in segregation or integration, for breast cancer classification. Performances were compared to benchmark models. Their deep learning model was achieved a discriminative accuracy of 87.3% when compared to 85.8% with a support vector machine.

Gardezi SJS, et.al., **2019** has been discussed a Machine Learning (ML) has become an important part of research in medical image processing. An ML technique has evolved over the days from manual seeded inputs to initialize automatically [9]. The benefit in the field of ML have led to more intelligent and self-reliant Computer-Aided Diagnosis (CAD) method. There was constant improvement in the learning ability of ML techniques. The authors presented an outline of ML and DL methods with exacting function for breast cancer.

Dr. D. Devakumari and V. Punithavathi, **2019** has been proposed a HDF algorithm for removing noise in the mammogram images. This was achieved by combination of Median Filter and Applied Median Filter [11]. It was recognized out by using Hybrid Denoising Filter (HDF) to remove the unwanted noises. The various de-noising algorithms were discussed in this paper. The experimental consequences of Hybrid Denoising Filter algorithm mainly focused on removal inappropriate noise and on the duration of implementation time using MATLAB R2013a software.

Dr. D. Devakumari and V. Punithavathi, **2020** has proposed an Optimized Kernel Fuzzy Clustering Algorithm was developed (OKFCA) and to find out the cancer portions in mammogram images. The OKFCA algorithm has been described to identify the segmented regions in MIAS database. The proposed segmentation algorithm was carried out with pre-processed mammogram images and the proposed OKFCA has been a significant approach to discover out the cancer segment of mammogram image. Data clustering facilitates to place data of similar types in one collection and of dissimilar types in different group [12]. The results from the experiments which were carried on the MIAS data confirm the efficiency of the proposed system in terms of accuracy when compared to that of the famous K-Means, OKFCA and Otsu methods.

Dr. D. Devakumari and V. Punithavathi, **2020** has been discussed about feature extraction technique for shrinking the dimension of image data by finding out an essential detail as of the segmented image. The visual substance of a segmented image portions can be confined using this procedure. Feature extraction is one of the Image retrieval methods for achieving higher accuracy. In this paper an Improved Gray Level Co-occurrence Matrix (GLCM) Feature based Extraction with shape and Tamura features were discussed [13]. Tamura features are extracted from Optimized Kernel Fuzzy Clustering Algorithm (OKFCA) which is one of the segmentation methods. Experimental results of 322 images were evaluated from MIAS database and used to estimate the feature extraction process.

III. PROPOSED METHODOLOGY

In the proposed methodology, all the testing is carried out by using the Mammographic Image Analysis Society (MIAS) **-** a benchmarked dataset [17]. A Hybrid algorithm with modified SVM with Multi class and KNN classification algorithm is proposed in this paper. The proposed algorithm is an expansion of KNN and SVM classification model. This classification starts with initializing with a number of k neighbors, Training and Test image features. This classification requirement is provided for all available classes and classifies the breast cancer category derived from similarity measure with correlation distance function. The complete proposed flow diagram is illustrated in Figure 2.

s

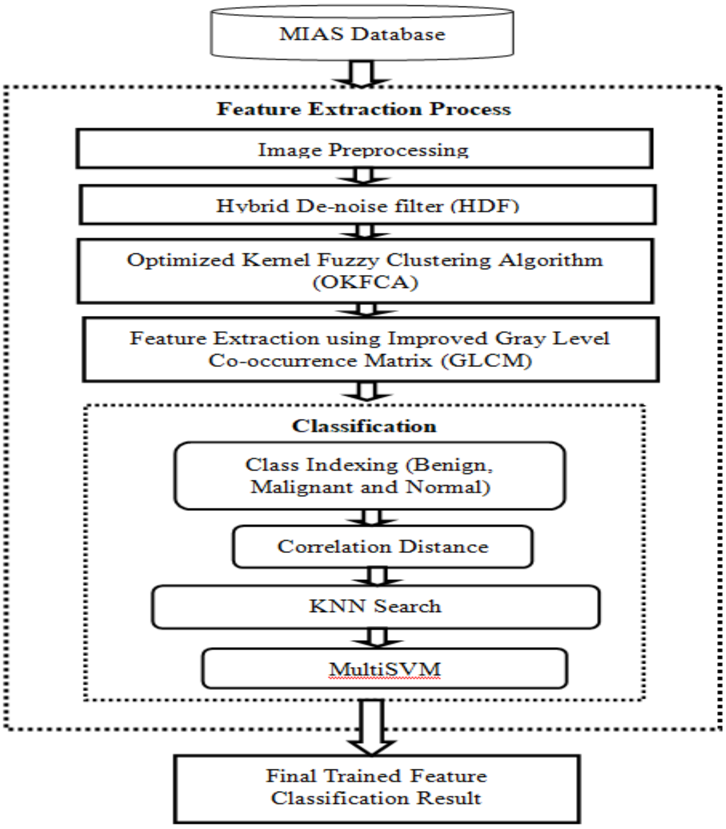

Fig. 2: Proposed Flow Diagram

The proposed flow diagram shows vividly the flow of the classification of breast cancer. The classification of mammogram images are carried out by class indexing then finding out the correlation distance by applying KNN and multi class is applied with SVM to obtain the classification result.

A. Image Preprocessing, HDF, OKFCA and GLCM

Dr. *D. Devakumari and V. Punithavathi*, 2019, already processed the following methods: Image preprocessing, HDF, Optimized Kernel Fuzzy Clustering Algorithm (OKFCA) for Segmentation and Improved GLCM Feature extraction. The mammogram images may contain noises thus it reduces the unwanted noises to detect the lesions. So, Hybrid Denoising Filter (HDF) algorithm is executed to get better the accuracy and to increase the excellence of the image. After de-noising process, segmentation process is done with Optimized Kernel Fuzzy Clustering Algorithm (OKFCA) which segments the cancer portions from the preprocessed image results that are shown in Fig 3 and 4. The above three process are done before extracting the features of the mammogram images.

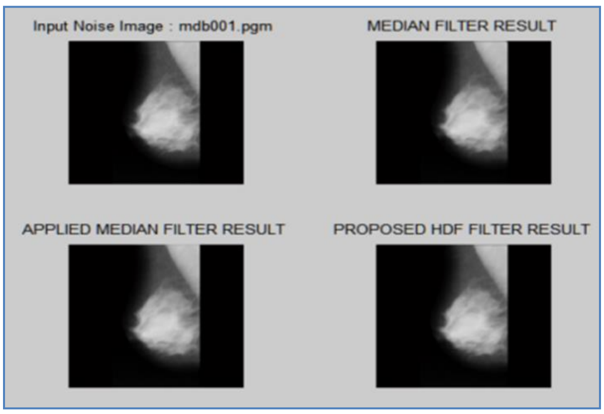

Fig. 3: Proposed Hybrid De-noising Filter Result

s

Fig. 4: OKFCA Segmentation Result

After the segmentation process an Improved Gray Level Co-occurrence Matrix (GLCM) Feature based Extraction with shape and Tamura features are executed. Tamura features are extracted from Optimized Kernel Fuzzy Clustering Algorithm (OKFCA) which is one of the segmentation methods. This paper proposed Improved GLCM Feature Based Extraction algorithm to obtain the trained features which are used for Classification process. The feature extraction results are shown in Figure 5.

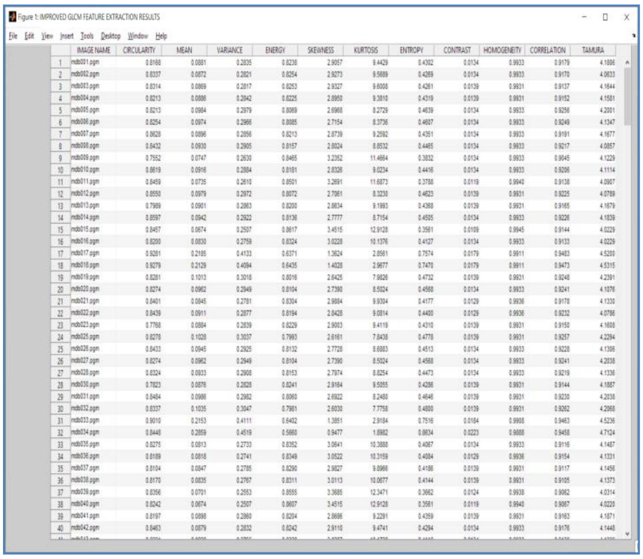

Figure 1: IMPROVED GLCM FEATURE EXTRACTION RESULTS

File Edit View Insert Tools Desktop Window Help

	IMAGE NAME	CIRCULARITY	MEAN	VARIANCE	ENERGY	SKEWNESS	KURTOSIS	ENTROPY	CONTRAST	HOMOGENEITY	CORRELATION	TAMURA
1	mdb001.pgm	0.8168	0.0881	0.2835	0.8238	2.9057	9.4429	0.4302	0.0134	0.9933	0.9179	4.1806
2	mdb002.pgm	0.8337	0.0872	0.2821	0.8254	2.9273	9.5689	0.4289	0.0134	0.9933	0.9170	4.0633
3	mdb003.pgm	0.8314	0.0869	0.2817	0.8253	2.9327	9.6008	0.4261	0.0139	0.9931	0.9137	4.1644
4	mdb004.pgm	0.8213	0.0886	0.2842	0.8225	2.8950	9.3810	0.4319	0.0139	0.9931	0.9152	4.1581
5	mdb005.pgm	0.8213	0.0984	0.2979	0.8069	2.6968	8.2729	0.4639	0.0134	0.9933	0.9256	4.2001
6	mdb006.pgm	0.8254	0.0974	0.2966	0.8085	2.7154	8.3736	0.4607	0.0134	0.9933	0.9249	4.1347
7	mdb007.pgm	0.8628	0.0896	0.2856	0.8213	2.8739	9.2592	0.4351	0.0134	0.9933	0.9191	4.1677
8	mdb008.pgm	0.8432	0.0930	0.2905	0.8157	2.8024	8.8532	0.4465	0.0134	0.9933	0.9217	4.0857
9	mdb009.pgm	0.7552	0.0747	0.2630	0.8465	3.2352	11.4664	0.3832	0.0134	0.9933	0.9045	4.1229
10	mdb010.pgm	0.8619	0.0916	0.2884	0.8181	2.8326	9.0234	0.4416	0.0134	0.9933	0.9206	4.1114
11	mdb011.pgm	0.8459	0.0735	0.2610	0.8501	3.2691	11.6873	0.3788	0.0119	0.9940	0.9138	4.0907
12	mdb012.pgm	0.8550	0.0979	0.2972	0.8072	2.7061	8.3230	0.4623	0.0139	0.9931	0.9225	4.0789
13	mdb013.pgm	0.7989	0.0901	0.2863	0.8200	2.8634	9.1993	0.4368	0.0139	0.9931	0.9165	4.1679
14	mdb014.pgm	0.8597	0.0942	0.2922	0.8136	2.7777	8.7154	0.4505	0.0134	0.9933	0.9226	4.1839
15	mdb015.pgm	0.8457	0.0674	0.2507	0.8617	3.4515	12.9128	0.3561	0.0109	0.9945	0.9144	4.0229
16	mdb016.pgm	0.8200	0.0830	0.2759	0.8324	3.0228	10.1376	0.4127	0.0134	0.9933	0.9133	4.0229
17	mdb017.pgm	0.9281	0.2185	0.4133	0.6371	1.3624	2.8561	0.7574	0.0179	0.9911	0.9483	4.5200
18	mdb018.pgm	0.9279	0.2129	0.4094	0.6435	1.4028	2.9677	0.7470	0.0179	0.9911	0.9473	4.5315
19	mdb019.pgm	0.8281	0.1013	0.3018	0.8016	2.6425	7.9826	0.4732	0.0139	0.9931	0.9248	4.2391
20	mdb020.pgm	0.8274	0.0962	0.2949	0.8104	2.7390	8.5024	0.4568	0.0134	0.9933	0.9241	4.1076
21	mdb021.pgm	0.8401	0.0845	0.2781	0.8304	2.9884	9.9304	0.4177	0.0129	0.9936	0.9178	4.1330
22	mdb022.pgm	0.8439	0.0911	0.2877	0.8194	2.8428	9.0814	0.4400	0.0129	0.9936	0.9232	4.0786
23	mdb023.pgm	0.7768	0.0884	0.2839	0.8229	2.9003	9.4119	0.4310	0.0139	0.9931	0.9150	4.1608
24	mdb025.pgm	0.8278	0.1028	0.3037	0.7993	2.6161	7.8438	0.4778	0.0139	0.9931	0.9257	4.2294
25	mdb026.pgm	0.8433	0.0945	0.2925	0.8132	2.7728	8.6883	0.4513	0.0134	0.9933	0.9228	4.1306
26	mdb027.pgm	0.8274	0.0962	0.2949	0.8104	2.7390	8.5024	0.4568	0.0134	0.9933	0.9241	4.2038
27	mdb028.pgm	0.8324	0.0933	0.2908	0.8153	2.7974	8.8254	0.4473	0.0134	0.9933	0.9219	4.1336
28	mdb030.pgm	0.7823	0.0876	0.2828	0.8241	2.9164	9.5055	0.4286	0.0139	0.9931	0.9144	4.1887
29	mdb031.pgm	0.8484	0.0986	0.2982	0.8060	2.6922	8.2480	0.4646	0.0139	0.9931	0.9230	4.2038
30	mdb032.pgm	0.8337	0.1035	0.3047	0.7981	2.6030	7.7758	0.4800	0.0139	0.9931	0.9262	4.2068
31	mdb033.pgm	0.9010	0.2153	0.4111	0.6402	1.3851	2.9184	0.7516	0.0184	0.9908	0.9463	4.5236
32	mdb034.pgm	0.8448	0.2859	0.4519	0.5660	0.9477	1.8982	0.8634	0.0223	0.9888	0.9458	4.7124
33	mdb035.pgm	0.8275	0.0813	0.2733	0.8352	3.0641	10.3888	0.4067	0.0134	0.9933	0.9116	4.1487
34	mdb036.pgm	0.8189	0.0818	0.2741	0.8349	3.0522	10.3159	0.4084	0.0129	0.9936	0.9154	4.1331
35	mdb037.pgm	0.8104	0.0847	0.2785	0.8290	2.9827	9.8966	0.4186	0.0139	0.9931	0.9117	4.1456
36	mdb038.pgm	0.8170	0.0835	0.2767	0.8311	3.0113	10.0677	0.4144	0.0139	0.9931	0.9105	4.1373
37	mdb039.pgm	0.8356	0.0701	0.2553	0.8555	3.3685	12.3471	0.3662	0.0124	0.9938	0.9062	4.0314
38	mdb040.pgm	0.8242	0.0674	0.2507	0.8607	3.4515	12.9128	0.3561	0.0119	0.9940	0.9067	4.0220
39	mdb041.pgm	0.8197	0.0898	0.2860	0.8204	2.8686	9.2291	0.4359	0.0139	0.9931	0.9163	4.1871
40	mdb042.pgm	0.8463	0.0879	0.2832	0.8242	2.9110	9.4741	0.4294	0.0134	0.9933	0.9176	4.1448
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Fig. 5: Improved GLCM Feature Extraction Results

B. Modified SVM and KNN Classification

This paper presents a new method of Modified SVM and KNN Classification Algorithm. It is an extension of KNN and SVM with Multi class Classification model. Both classification and regression predictive problems can be used by these methods. This is a supervised learning technique as well as a geometric method for classification. This classifier collects all existing instance and categorizes new instances developed on similarity rate (Example: Correlation or Cityblock distance functions). To classify the extracted improved GLCM features in Feature Extraction process the KNN method is used. The classification is designed using Correlation or Cityblock distance as a metric between the features of the testing data and the reference data as shown in Equation 1.

The correlation between vectors A and B are represented as follows:

s

$$dist\_Corr\ (A, B) = \frac{\sum_{a=1}^{m}(x_a - \mu_a)(y_a - \mu_a)}{\sqrt{\sum_{a=1}^{m}(x_a - \mu_a)^2 \sum_{a=1}^{m}(y_a - \mu_a)^2}} \text{ eqn. (1)}$$

Where *A* and *B* are feature space, μ_a and μ_b are the means of A and B respectively. The KNN classification estimates class attributes based upon the *k* nearest training instances in the feature space. A mammogram image trained feature extraction dataset is given; it chooses the k nearest samples from the classified training data and determines the class considering the most representative samples. The option of the constraint k ($k \in N$) is find out by the user, this option depends on the data of the image. The consequence of sound on the classification is reduced when the values are chosen for k is greater, but this makes the limitations between classes of less distinct. A better option of the value of k can be selected by different probing methods such as cross-validation. The value of k is selected which minimizes the classification error.

A SVM (Support Vector Machine) with Multi Class Classification is a machine learning technique which works on the principle of structural risk minimization in order to find out the three classes (normal, Benign and malignant). A KNN is combined with SVM with Multi class data which are used for training and testing purpose. In this proposed work, the testing data are divided into three groups. The testing data were taken inside from training data in the first group. The testing data were taken outside from training data in the second group. Testing data were taken inside and outside from training data in the third group. Grouping is done to see the accuracy from each group. The process of classification is performed to classify the category of mammogram images as normal and abnormal (Benign and Malignant) images.

$$Modified\ SVM\ with\ KNN = \sum_{n=1}^{outside} \sum_{m=1}^{inside} dist\_corr\big(a, class(a, b)\big) \quad eqn.(2)$$

The Modified SVM with KNN is modified by introducing multi classes such as a and b. The variable ‘a’ represents trained feature and ‘b’ represents test image feature (Benign, Malignant, and Normal).
The proposed algorithm takes OKFCA segmented image and GLCM features as Input and several steps are followed to obtain classification process. The proposed algorithm is depicted in Figure 6.

Algorithm 1: A Hybrid Algorithm with Modified SVM and KNN for Classification
Input: OKFCA Segmented Image, X: training data, Y: class labels of X.
Output: Class *y* (Benign, Malignant and Normal)
Process
Step 1: Choose a value for the parameter *k*.
Step 2: Calculate Modified SVM with KNN
for j = 1 to *m* **do**
Compute distance dist_Corr (A,B) using eqn. (1)
end for
Step 3: Apply Modified SVM with KNN using eqn. (2)
Step 4: Integrate the classes of these Y samples in one
class c.
Step 5: The class of y is c(Y) = c

Fig. 6 A Hybrid Algorithm with Modified SVM and KNN for Classification

IV. EXPERIMENTAL RESULTS

The experimentation consequence has been executed and demonstrated the performance of Modified SVM with KNN Classification. The experimental results are performed on Intel I5 processor with 3.20 GHz 4 speed, 8GB RAM, with Windows 10 operating system, on MATALB R2013a software. The experimental process is carried out with MIAS

s

image database and are analyzed with OKFCA segmentation and evaluation of Improved GLCM Feature extraction method. In this experimental results, the MIAS image database are applied to the Modified SVM with KNN Classification algorithm with existing KNN and SVM algorithms which are pliably constructed to forecast the exactness or accuracy of database to meet up the requirements' of various test requirements.

Quantifying this proposed classification algorithm the performance is normally done by combining True/False Positives/Negatives in order to measure the accuracy. They are defined by:

$$Modified\ SVM\ with\ KNN\ Accuracy = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \quad eqn. (3)$$

Table 1 shows the evaluation of MIAS database of sample seventeen images of accuracy values for various existing classification algorithms. Table 1 shows the evaluation of accuracy of ten test images with existing algorithms (SVM and KNN) and Proposed KNN with SVM. It also classifies the Disease categories : Normal (N),Benign (B) and Malignant(M). The Figure 7 shows the performance chart which depicts the accuracy for the classification of Ten Test mammogram Images.

Table 1: Comparison of Classification Accuracy Measures for Ten Test mammogram Images

Image Name	Disease Type	SVM	KNN	Proposed KNN With SVM
mdb001.png	**B**	**82**	**80**	**99.747**
mdb002.png	**M**	**75**	**77.273**	**99.964**
mdb003.png	**N**	**73**	**83.33**	**98.247**
mdb004.png	**M**	**79**	**80**	**98.954**
mdb005.png	**M**	**82**	**72.727**	**98.749**
mdb006.png	**M**	**89**	**87.5**	**98.315**
mdb007.png	**M**	**82**	**76.190**	**98.806**
mdb008.png	**B**	**78**	**68.182**	**99.169**
mdb009.png	**M**	**77**	**68.182**	**98.941**
mdb010.png	**M**	**84**	**90**	**99.612**

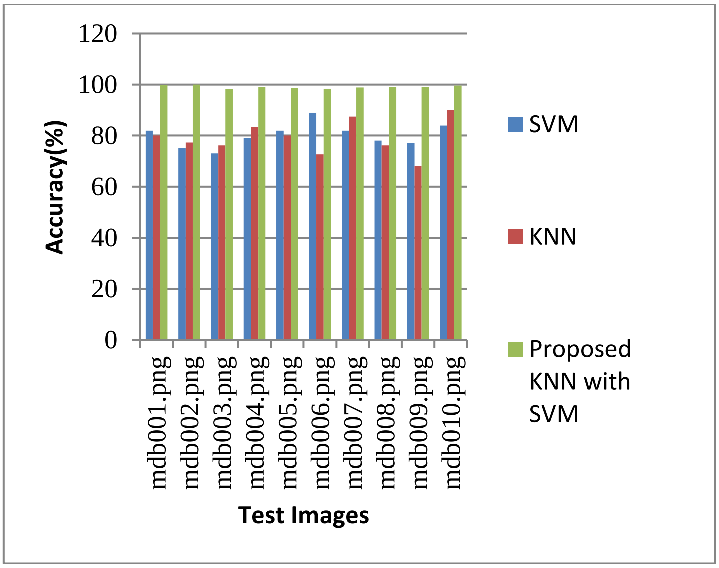

Fig. 7: Performance Chart: Accuracy for Classification of Ten Test mammogram Images

Table 2 shows the average Classification Accuracy for Ten test mammogram images from 20 trained dataset which are applied with existing KNN with the proposed KNN with SVM Classification algorithm.

s

Table 2: Average Classification Accuracy for Ten Test Images with 20 Trained Mammogram Images

Measure	SVM	KNN	Proposed KNN with SVM
ACCURACY	80.1	79.139	99.050

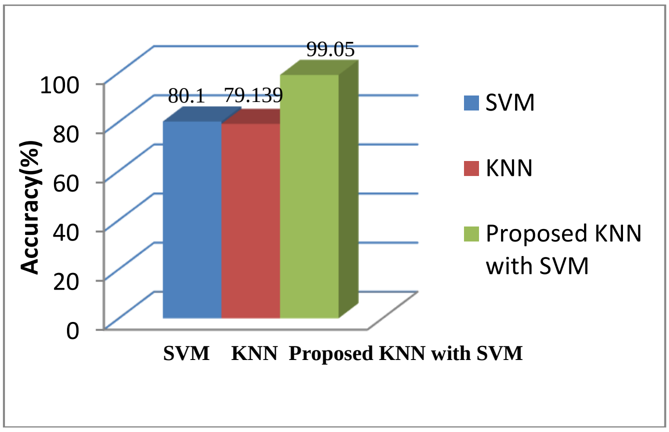

Fig. 8. Overall Classification Accuracy chart for Ten Test Mammogram Images

Table 3 shows the computation of classification of Mammogram image Accuracy on MIAS database. This work is also carried out for 322 images which are available in MIAS database and the overall accuracy for 322 images are evaluated. The average Classification accuracy of 322 images and the ratio are defined by,

$$OverAll\ Accuracy = abs(mean(Classifcaiton\_Accuracy))\ eqn. (4)$$

Table 3: Overall Classification Accuracy for 322 MIAS sample Images

Measure	SVM	KNN	Proposed KNN with SVM
ACCURACY	87.69	88.5401	99.3406

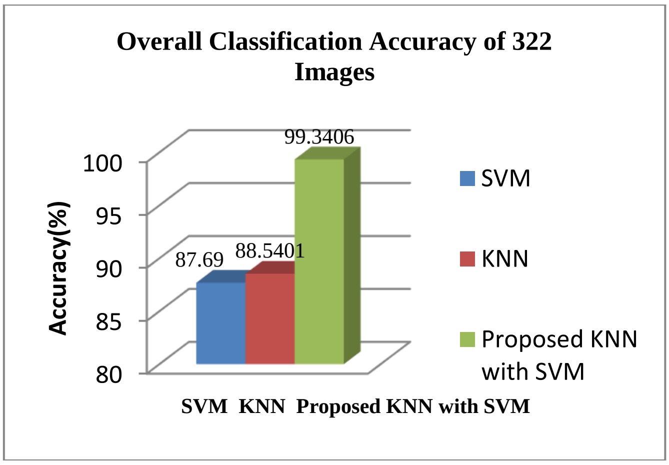

Fig. 9. Gross Classification Accuracy Graph of 322 Images

s

The Figure 9 depicts the overall classification accuracy chart of 322 Mammogram images which are taken from MIAS database. This performance chart shows the accuracy of Proposed KNN with SVM as 99.3406 which are proved best with the experimental results when compared to the existing algorithms (KNN and SVM).

IV. CONCLUSION

The proposed KNN with SVM Classification algorithm is done by combining the KNN and SVM algorithm. This paper finds out the performance for ten test mammogram images with 20 trained mammogram images and also finds out the overall classification accuracy for 322 mammogram images from MIAS database. The experimental results proves the classification accuracy for ten test images with trained data as 99.050 whereas the existing algorithm 79.139 and 80.1. The Proposed KNN with

SVM also shows the high accuracy for 322 mammogram images in MIAS database when compared with KNN and SVM. SVM classifier suffers from the thorough statistical problem due to the complex training and classifying processes, mainly when the training samples number and more number of features for each sample are huge. Thus Hybridized or Improved models gives number of advantages for classification of Breast Cancer by using multi class datasets in training as well as testing to provide a better accuracy with the existing algorithms.

REFERENCES

[1] E.C.Fear, P.M.Meaney, and M.A.Stuchly,"Microwaves for breast cancer detection", IEEE potentials, vol.22, pp.12-18, February-March 2003.
[2] Homer MJ.Mammographic Interpretation: A practical Approach. McGraw hill, Boston, MA, second edition, 1997.
[3] American college of radiology, Reston VA, Illustrated Breast imaging Reporting and Data system (BI-RADSTM) , third edition, 1998.
[4] P. Shi, S. Ray, Q. Zhu, and M. A Kon. Top scoring pairs for feature selection in machine learning and applications to cancer outcome prediction. BMC Bioinformatics, 12, 2011.
[5] C. Cortes, V. N. Vapnik, "Support vector networks", Machine learning Boston, vol.3, Pg.273-297, September 1995
[6] Girosi f. Jones M. and Poqqio T., "Regularization theory and neural network architectures", Neural computation Cambridge, vol.7, pg.217-269, July 1995.
[7] A. Marcano-Cedeno, J. Quintanilla-Domnguez, and D. Andina. Wbcd breast cancer database classification applying artificial metaplasticity neural network. Expert Systems with Applications, (38), 2011.
[8] Buciu, I. & Gacsadi, A. Directional features for automatic tumor classification of mammogram images. Biomed. Signal Process. Control. 6, 370–378 (2011).
[9] Wang, J. et al. Discrimination of breast cancer with microcalcifications on mammography by deep learning. Sci. reports 6, 1–9 (2016).
[10] Gardezi SJS, Elazab A, Lei B, Wang T. Breast cancer detection and diagnosis using mammographic data: Systematic Review. Journal of Medical Internet Research. 2019; 21(7).
[11] Devakumari D and Punithavathi V, "Noise Removal in Breast Cancer using Hybrid De-Noising Filter for Mammogram Images", published by International Conference on Computational Vision and Bio-Inspired Computing (ICCVBIC 2019), January,2020,pp. 109-119.
[12] Punithavathi V and Devakumari D, "A New Proposal for the Segmentation of Breast Lesion in Mammogram Images using Optimized Kernel Fuzzy Clustering Algorithm", accepted by Materials Today: Proceedings - Elsevier, 2020.
[13] Punithavathi V and Devakumari D, "Detection of Breast Lesion Using Improved GLCM Feature Based Extraction in Mammogram Images", published by SSRN eLibrary, July 30, 2020.