BREAST CANCER PROGNOSIS PREDICTION USING NOVEL SHRUNK KERNEL KNN METHOD WITH MLO AND CC FEATURES

Dr. V. Sridevi
Assistant Professor
Department of Computer Science
PSG College of Arts & Science, Coimbatore
vissridevi@gmail.com

Abstract

The most continual malignancy found in women is breast cancer. The good news is that if detected early, it is one of the most treatable forms of cancer. High-dimensional data results in large number of computation redundancy but also advances diagnostic techniques. As a result significant information must be extracted and the feature dimension must be reduced for good prediction and a precise treatment decision. However, past studies for diagnosing breast cancer have relied mostly on labelled data that is difficult to get. To solve this problem, two different sorts of perspectives, such as CC and MLO are employed to improve diagnostic effectiveness. This chapter comprises segmentation, feature extraction and classification of images. The two perspectives from a mammography picture are segmented using the adaptive K-means clustering approach. The Gabor filter is used in conjunction with the traditional k-means clustering method during the feature extraction stage to extract the features of the CC and MLO perspectives. The mammography image is finally classified into benign and malignant using a unique Shrunk Kernel K-Nearest Neighbor (SKKNN) classifier.

Keywords: SKKNN, MLO view, CC View, adaptive K-means segmentation, Gabor filter, Mammogram

1. INTRODUCTION

Modern society faces a cancer pandemic. Breast Cancer, which primarily affects women, spreads from the breast due to uncontrolled growth in cells. Only skin cancer is more frequent in women than in BC. Over-50-year-old women are the ones most likely to be affected. Though it's uncommon, BC may also affect males. Men make up less than 1% of instances of male BC, which affects around 2,600 men annually in the US. Transgender women are more likely to acquire BC than cisgender males. Transgender males are also less likely to acquire BC than cisgender women. 6,85,000 people died worldwide and 2.3 million women were diagnosed with BC in 2020 (Lei et al., 2021). By 2020, 7.8 million women had been diagnosed with BC,

making it the most frequent cancer worldwide. BC occurs in every country of the world in women at any age after puberty but with increasing rates in later life. Over the last several decades, researchers have shown that it is possible to automate the basic level of tumor classification and identification.

Kulkarni (2019) sought to categorize mammographic lesions using Pixel N-gram features and a variety of classifiers like MLP, KNN and SVM, and it was discovered that the performance showed the effect of improving N. The obtained results were evaluated and it was observed that performance attained with MLP classifier excels that achieved with KNN or SVM or classifiers.

A Radial Basis Function Network (RBF)-based CAD system for diagnosis of breast was introduced by Ibrahim et al (2020). Using RBF network classifiers, the decision-making system aids in the categorization of malignancies. The MLP method obtained 54.1667% accuracy overall, whereas the RBF neural network achieved a classification accuracy of 79.166%, demonstrating the efficacy of the RBF neural network in categorizing the mammography images.

Hough transform was used by Vijayarajeswari (2019) to identify characteristics in the mammography picture. These characteristics are used as inputs to the SVM classifier. The accuracy range of the SVM classifier is 94%, which is higher than that of the LDA classifier (86%).

The back propagation neural network (BPNN) classification model was used by Mughal (2019). The method correctly recognises the tumour in its early stages with 99.0% classification accuracy on MIAS and DDSM datasets.

Lakshmi (2019) provided the implementation of employing a KNN classifier to diagnose breast cancer, which largely distinguishes between breast masses and cancer, using the supplied inputs, a portion of the breast region's cancer-affected cells is located. For the aid of their collected characteristics and the corresponding labels for each picture generated from the breast image dataset, the KNN classifier is trained. The proposed approach achieves a total accuracy close to >90%.

Using the Mini-MIAS dataset of 322 images, Prabh Kaur (2019) put forward a study that discusses about an inventive scheme that includes pre-processing stage and K-mean clustering feature extraction for Speed-Up Robust Features (SURF) selection. According to the findings, a decision tree model is outperformed by the automated DL strategy that is

suggested, which applies K-mean clustering with MSVM. The results of the studies demonstrate that the suggested methodologies achieve average accuracy (ACC) rates of 95%, 94%, and 98% for the three types of cancer, namely benign, malignant, and normal. When SVM is utilized in contrast to MLP and K-mean clustering WEKA manual approach, the increased sensitivity rate is realized at 3%, specificity is at 2%, and ROC area is at 0.99. After using 10-fold cross validation, the SVM, LDA, K-nearest neighbour (KNN), and Decision Tree produced results of 96.9%, 93.8%, 89.7%, and 88.7%, respectively.

Viswanath (2019) proposed three significant classifiers, including Support Vector Machine (SVM), Random Forest (RF), and k-Nearest Neighbors (K-NN) with the goal of determining how accurate they are at generating decisions. Moreover, the influence of pre-processed mammogram images prior to its submission into the classifier is discussed, yielding efficient classification.

Kayode et al (2019) utilized a modified version of SVM classification for an automated CAD System for mammogram classification. The technology could prove to be a valuable tool for radiologists to support their decision-making when interpreting mammograms, according to the findings of performance analysis. The effectiveness of the radiologists' diagnostic proficiency could be increased by using this to make a precise and proactive judgement.

Arooj (2022) proposed effective machine learning technique for the quick identification of breast cancer. The suggested method employed a stack made up of the three algorithms decision tree, SVM and KNN, and adopts the CRISP-DM procedure to develop a collaborative model. The meta-classifier performance is compared using the three collaboration model to that of the individual works of DT, SVM, and KNN. Chi-square analysis is used to evaluate the top characteristics. The recommended model produced the best accuracy of 78% and least log loss of 0.56. The result of one-tailed t-test at $\alpha = 0.05$ gave P-value of 0.014.

Dewangan et al (2022) exhibited the reduced sensitivity and efficacy during breast cancer identification. An innovative Back Propagation Boosting Recurrent Widening Model (BPBRW) with a Hybrid Krill Herd African Buffalo Optimization (HKH-ABO) technique was generated in this research for the diagnosis of breast cancer at an early step that uses breast MRI data. The initial training of the system is achieved with the help of MRI breast images. Python is helpful in the model simulation. It is shown that the accuracy rate of this model is about 99.6%.

Lin et al (2022) developed the model to increase detection precision and decrease breast cancer misdiagnosis. The dataset was put into ANN and SVM for breast cancer tumor classification and comparison. The best classification accuracy is 91.6% for the SVM using radial features and 76.6% for the ANN. Consequently, SVM was utilized for demonstrating the significance of breast screening. The second phase is about the application of transfer learning for training the AlexNet, inceptionV3, and ResNet101. AlexNet scored 81.16%, ResNet101 scored 85.51%, and InceptionV3 scored 91.3 %, as per the data.

Alruwaili & Gouda (2022) introduced a model that depends on the concept of transfer learning. To avoid fitting problem and provide constant results through the enlargement of numerous mammography images, various augmentation methods were used, including rotation, scale, and shifting. On the MIAS database, ResNet50 achieved 89.5% accuracy and NASNet-Mobile 70%. The efficiency and effectiveness of Pre-trained categorization networks are quite remarkable, rendering them highly desirable for diagnostic performance.

Alshammari (2021) recommended a machine learning-based algorithm that is useful to the radiologist in reading the mammography images and classifying the tumor within a reasonable amount of time in this research. Several features have been extracted from the considered region of the mammogram', whose manual labelling was done by the physician. Consequently, it was found from this research that multivarious conditions can influence the outcomes, which was not considered after the examination. This research emphasizes the usage of the optimized SVM or Nave Bayes that yields an accuracy of 100%.

In Jayandhi et al, (2022), the Deep Residual learning model is combined with a Decision Tree Machine Learning mechanism to accomplish breast cancer prediction with efficiency. Hence, the proposed study was referred to be RDT model. Recall, accuracy, accuracy, and specificity are used to develop and verify the RDT model. HDL model offers better prediction accuracy in terms of the traditional mechanisms.

In Diaz et al., 2019, introduced the first order procedure of feature determination which is baseline for the classification using KNN. The highest accuracy levels were obtained with K = 5 for classification with cross-validation and K = 15 for classification without cross-validation, both of which produced accuracy values of 91.8%.

The examination of single views has served as the foundation for the majority of CAD system development. The development of CAD methods that make use of information from several perspectives such as bilateral views of the same breast is of great interest to reduce the

amount of false positives and increase consistency (Bandyopadhyay, 2010). In order to find abnormal asymmetry densities, radiologists are trained to compare the left and right breasts. Earlier screening views are utilized to spot developing density. Screening with two mammograms medio lateral oblique MLO and craniocaudal (CC) is also recommended (Li et al., 2022) which increases the detection accuracy of breast abnormalities. Two projections may show lesions hidden by glandular tissue in one projection (Gennaro et al., (2013). Because the breast is compressed under a different perspective, it is more challenging to establish the correspondence between MLO and CC views than for temporal image pairs. Additionally, there aren't many constant landmarks.

2. PROPOSED METHODOLOGY

The CBIS-DDSM dataset was used in this study (Kouser, 2018) which contains the biopsy-proven annotated mammograms. Pre-processing, segmentation, and feature extraction were performed when the training images are supplied to the computer. The suspicious regions are segmented from the mammography region using the adaptive K-Means segmentation technique. The segmented regions are passed to the feature extraction step, where the KMC-GF algorithm is used. The KMC-GF procedure consists of two steps: the first involves applying the adaptive K-Means clustering approach to a segmented region to cluster it, and the second involves using the Gabor filter to extract features from the cluster region. In order to differentiate between benign and malignant lesions in the mammography area, the SKKNN classifier is used. Figure 2.1 describes the proposed architecture using SKKNN in BC Detection.

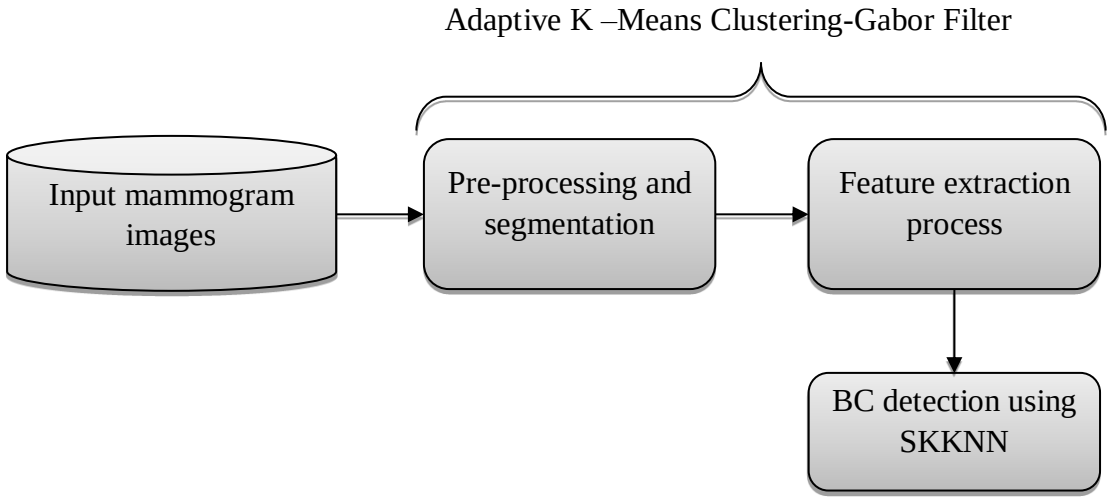


Figure 2.1 Proposed Architecture using SKKNN in BC Detection

2.1 Data Set

In this study, we make use of the biopsy-proven annotated mammograms from the CBIS-DDSM dataset (Kouser, 2018). The collection includes bilateral breast scans taken from the cranio-caudal (CC) and medio-lateral (MLO) perspectives. Based on the diagnosis of the radiologists, we extract 1418 normal mammograms, 852 benign mammograms, and 897 malignant mammograms from the database. Each of these breasts has two mammograms from the CC and MLO views, for a total of 6334 mammograms from 3167 breasts. Training and testing data sets are created from the data. The training includes 1240 mammograms with normal results and 1507 mammograms with abnormal results (738 benign and 769 malignant). In the test set, there are 178 normal mammograms and 242 abnormal ones (114 benign and 128 malignant). The specification for the datasets divided is shown in Table 2.1. We may evaluate the model's performance using ten-fold cross-validation, we proportionately divide the training data into 10 subsets.

Table 2.1 Partitioning Specification of Dataset

Data / Type	Normal	Benign	Malignant	Total
Train	1240	738	769	2747
Test	205	205	205	615
Total	1445	943	974	3362

2.2 Image Pre-Processing

The quality of the image is increased through image preprocessing, making it more acceptable for the visual features employed by computer and human recognition systems.Background noise, tape artifacts, high-intensity rectangle labels, edge shadowing effects, and low-intensity labels in mammography images are the types of noise that have been noticed. Due to the complexity and variability of breast tissue, the mammographic image has low contrast and therefore the doctor only extracts a limited amount of information from the image. Misdiagnoses occur as a consequence of even skilled doctors being unable to identify hidden MC. We used several preprocessing methods in this work to smooth, brighten, denoise, and identify edges in breast pictures to enhance their look. Sharp edges were preserved while noise was reduced using adaptive median, Gaussian, and bilateral filtering. Artifact and

pectoral muscle were eliminated by denoising. Tissue density in the non-breast region was substantially correlated, which might affect future mammography analyses.

To increase the image quality and smoothness, this research used image enhancement techniques based on wavelet analysis, CLAHE, and adaptive unsharp masking, as illustrated in Figure 2.2 The suggested method reduces superfluous background data, highlights the image's weak borders and calcification spots, and accentuates tiny calcification sites. In order to accurately assess the impact of image edge enhancement and to measure the image's denoising to confirm the efficacy of the procedure, this research employed the contrast improvement index (CII). Additionally, in order to assess the improved performance of mammography images, we assessed the CII and PSNR. A greater denoising effect is shown by higher PSNR and CII values

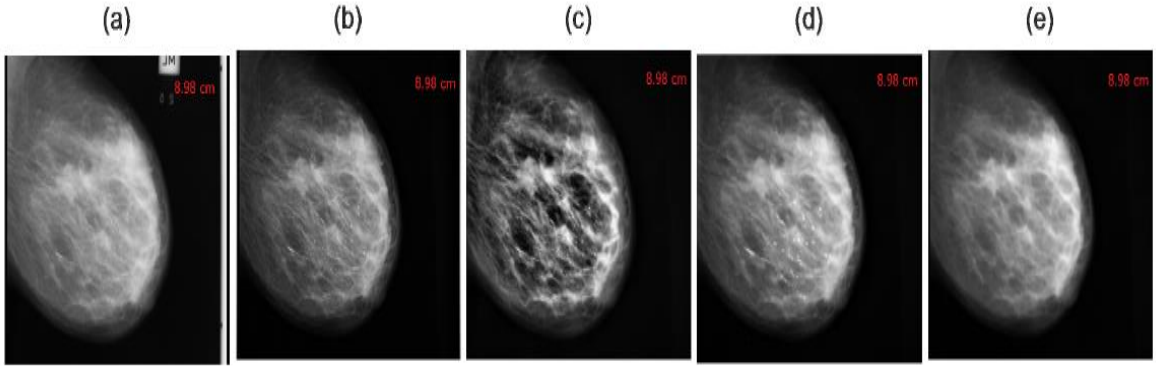


Figure 2.2 Pre-processing - (a) original Mammogram Image, (b) Adaptive Unsharp Masking, (c) Image Enhancement (CLAHE), (d) Dilation, (e) Erosion

2.3 Adaptive K Means Segmentation

Starting with a selection of k inputs from the provided database, the adaptive k means clustering method is run. The K randomly chosen elements are used to create the clusters. Each K element that makes up an element has a set of qualities that combine to generate the cluster properties. Figure 2.3 shows the flowchart of the adaptive K-Means segmentation procedures.

Based on the aforementioned algorithm, the distance between the provided element and the clusters is calculated. The distance should be taken into consideration depending on the qualities, which is a crucial factor that is also normalized. Consequently, none of the qualities either dominate or are neglected by the result. The most common use of the Euclidean distance is given by Equation 2.1.

$$Euclidean\ Distance = \sqrt{(E_{11} - E_{12})^2 + (E_{12} - E_{22})^2 + \cdots + (E_{1n} - E_{2n})^2} \quad (2.1$$

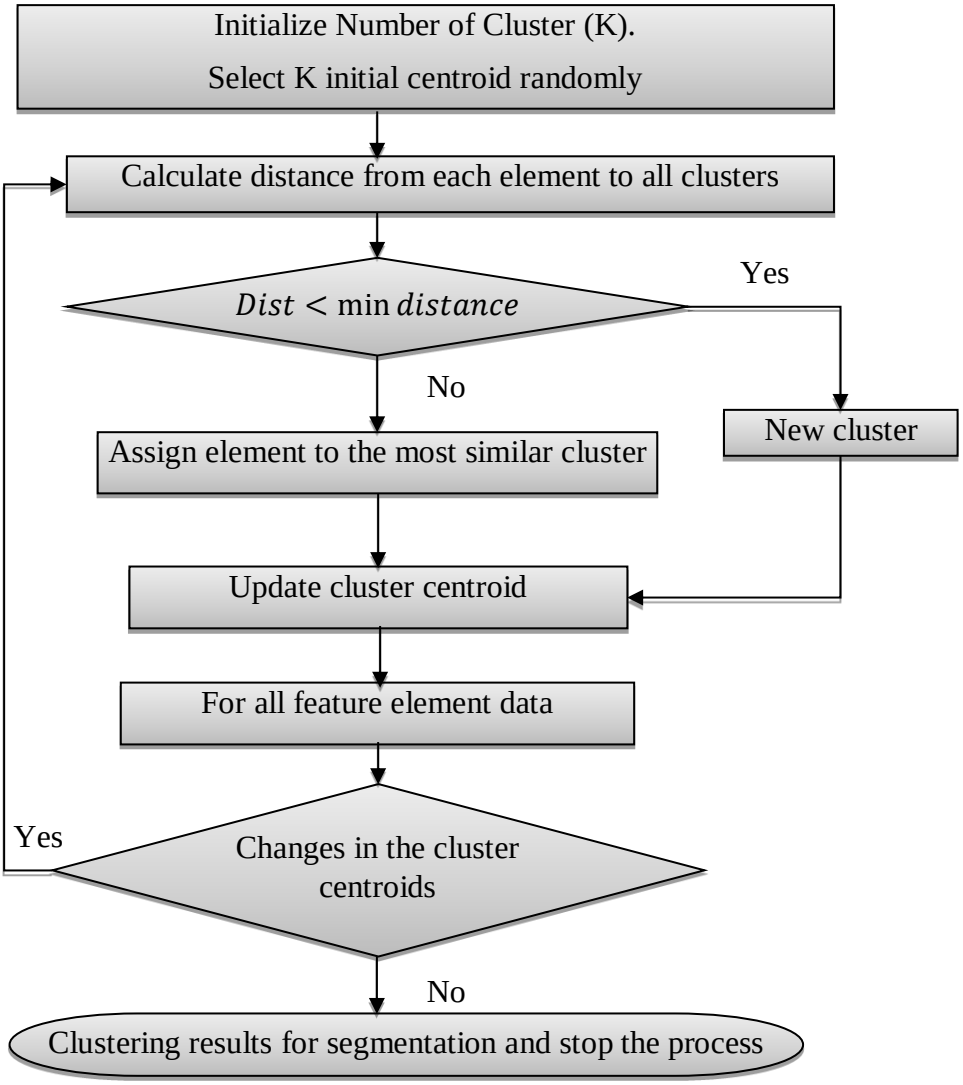


Figure 2.3 Flowchart of AKM Segmentation

For the purpose of dropping the square root function, the derived distance function must be adjusted. To compare properties in this method, several weights are needed for each property. It is determined and stored as a triangular matrix how far apart each cluster is from the others. For every element that isn't in a cluster, the distance is determined. For this element, there are three possible scenarios.

- If the distance is zero, place the appropriate element in that cluster and then begin the procedure from that element.
- If the distance is less than the minimum distance, the element should be assigned to the cluster that is nearest to it if. As a result, the cluster or centroid's representation may

alter. The centroid is recalculated using mean values of new clustered group. Recalculate the minimum distance between each cluster and the distance to the damaged cluster.

- The last case is occurring when the minimum distance is less than the distance of the element from the nearest cluster. In this case, two clusters are merged and cluster is destroyed by removal of all elements from the cluster.
- Finally, new cluster is established by the addition of new members into the empty cluster. At last, a new estimate of the distance between each cluster identifies the two nearest clusters.
- The above three processes are repeated until all the elements have the clustered.
- The obtained k values are given in Gabor filter for the extraction of features from mammogram image.

2.4 Proposed Gabor Filter for Feature Extraction

Gabor filters have been used for image coding, image representation, texture segmentation, target recognition, edge detection, retina identification, and more. The Gabor family of filters has gained popularity in recent years because they can imitate the properties of some cells in the visual brain of animals. The primary visual cortex is thought to undertake identical orientation and Fourier space decomposition tasks, according to biological studies, making them appear reasonable for a technological vision system. Additionally, it has been shown that these 2D band-pass filters offer the best localization capabilities in the spatial and frequency domains, making them ideal for extracting image edges or features that are oriented in a particular frequency range.

As a sinusoidal plane modulated by a GE, a Gabor filter may be thought of as having a certain frequency and direction. Equation 2.2 is a possible representation.

$$\mathrm{GE}(x,y) = e^{\frac{1}{2}\left[\frac{x_2}{\Sigma_x^2}+\frac{y_2}{\Sigma_y^2}\right]}{}_{e-j2\Phi(u_0x+v_0y)} \quad (2.2)$$

Two 2D Gaussian functions GF from Equation 2.3 make up the filter's response in the Fourier frequency domain.

$$\mathrm{GF}(u,v) = \mathrm{GE}_1 + \mathrm{GE}_2 \quad (2.3)$$

where $\Sigma_u = \frac{1}{2\pi\Sigma_x}$ and $\Sigma_v = \frac{1}{2\pi\Sigma_y}$assuming that the Fourier transform's origin has been centered and are the standard deviation along two orthogonal directions. To create the edge

histogram descriptors that will serve as the classification features, a collection of edge histogram descriptors for each alarm segment is produced with its n counterparts present in n Gabor-filtered images. After clustering the EHD features with Adaptive k-means clustering method, a SKKNN classifier is used to reduce the number of false alarms (Sundara Vadivel, 2019). The histogram that is produced by EHD indicates the local frequency of four distinct kinds of edges at each band. These corners have angles of 90^o vertical, 0 degrees horizontal, 45^o diagonal, and 135^o diagonal. The vertical histogram frequency for a particular segment at each band is defined as the proportion of pixels in the vertical edge-extracted image with the highest intensity values in comparison to the pixel values in the other three directional edge-extracted images (horizontal, 45° diagonals, and 135° diagonal). Four directional frequencies may be combined to generate a four-dimensional EHD signature, and the remaining three directional frequencies can be calculated in a similar manner. The EHD characteristics are expressed as in Equation 2.4.

$$\mathrm{EHD}\ (\mathrm{m},\mathrm{n}) = \frac{\text{No.of max intensity pixels at direction n}}{\text{Alarm segment area}} \tag{2.4}$$

$$\mathrm{m} = 1,2,3,4,5;\ \mathrm{n} = 1,2,3,4$$

The EHD computation is identical to counting the quantities of pixels with maximum intensity in each band's direction. For instance, if band 1's EHD signature has the biggest vertical frequency, band 1 will be dominated by vertical edges. Such a cancer segment EHD characteristic displays information on both directional edges and frequency scales from low to high frequencies. Regardless of the absolute intensity settings, the statistical properties of the EHD are consistent and dependable. Both the MLO and CC views of the mammography are used to extract features. So we can have two feature vectors.

The GLCM matrix is created by applying a Gabor filter on the cropped ROI after it has been trimmed. Co-occurrence matrix characteristics such as contrast, energy, entropy, mean, standard deviation, homogeneity, correlation, entropy, cluster shade, and cluster prominence are among the second order statistical features that may be derived from this information and computed with the use of characteristics mentioned in Figure 2.4.

$$Energy = \sum_{i}^{M}\sum_{i}^{N} i^2\,[i,j]$$

$$Entropy = -\sum_{i}^{M}\sum_{j}^{N} I\,[i,j]\log I\,[i,j]$$

$$Mean = \frac{1}{M * N} \sum_{i}^{M} \sum_{j}^{N} I\,[i,j]$$

$$Homogeneity = \sum_{i}^{M} \sum_{j}^{N} \frac{I\,[i,j]}{1 + \,[i+j]^2}$$

$$Correlation = \sum_{i}^{M} \sum_{j}^{N} \frac{(i,\mu)(j,\mu)}{\sigma} i\,[i,j]$$

$$Inverse\ Difference\ Moment = \sum_{i}^{M} \sum_{j}^{N} \frac{I[i,j]}{[i,j]^2}\ \ where\ i\ \neq j$$

$$Contrast = \sum_{i}^{M} \sum_{j}^{N} [i,j]^2\ I^2[i,j]$$

$$STD = \sqrt{\left(I[i,j] - \overline{I[i,j]}\right)^2}$$

$$Cluster\ Shade = \sum_{i}^{M} \sum_{j}^{N} \left(i\ M_x + j\ M_y\right)^3\ I[i,j]$$

$$Cluster\ Prominance = \sum_{i}^{M} \sum_{j}^{N} \left(i\ M_x + i\ M_y\right)^4 I[i,j]$$

$$Where$$

$$M_x = \sum_{i}^{M} \sum_{j}^{N} i\ I[i,j]\ \ and\ \ M_y = \sum_{i}^{M} \sum_{j}^{N} j\ I[i,j]$$

$$\mu - Mean \qquad \sigma = variance\ of\ co - occurrence\ matrix$$

Figure 2.4 Texture Features Extracted from Each View Mammogram in Gabor Filter

The following explanations provide the previous attribute's physical significance.

Contrast: The intensity contrast between a pixel and its neighbour over the source image.

Correlation: Measures the degree to which pixels in the whole image are statistically associated with one another; range = [-1 1]. A fully positively or negatively correlated image has a correlation of either 1 or -1.

Energy: Squared element GLCM summation; range = [0 1]. For an unchanging image, energy equals 1.

Homogeneity: Range = [0 1] indicates how nearly the GLCM element distribution follows the GLCM diagonal. For a diagonal GLCM, homogeneity is 1.

Entropy: It is a randomness measure which is used to describe the texture of the input image.

Cluster shade and cluster prominence: Indicators of the skewness, or lack of symmetry, of the matrix.

After the features are determined for each view, Feature fusion is applied. To create a single feature vector that is more discriminative than either of the input feature vectors, two feature vectors are combined in a process known as feature fusion. Feature level fusion is done by DCAFUSE employing a Discriminant Correlation Analysis (DCA)-based methodology. It gathers the train and test data matrices and accompanying class labels from two modalities X and Y and merges them into a single feature set Z.

2.5 Proposed SKKNN for BC Detection

Although KNN is effective at classifying data, when the training set grows large, the computational cost often prevents its use in practical projects. This section introduces an effective and quick model based on the KNN baseline model for enhancing classification performance and improving training effectiveness. In this part, the kernel method is presented together with the KNN model. It employs the characteristic of kernel method for expanding the dimension of features and enhances the performance of classification in the baseline model KNN. The Shrunk Kernel technique is then introduced, which primarily aims to cut down on kernel computation.

The fundamental principle of the reduced kernel technique is to compute the kernel matrix using just a portion of the training data from each class. According to this research, the training training samples are used to construct the kernel matrix. This kind of feature representation is equivalent to or superior to the traditional kernel matrix. The following equation 2.5 illustrates how the Shrunk kernel mathematical works.

$$\overline{\mathrm{SK}}_{\mathrm{tr}} = \mathrm{G}(\mathcal{U}_{\mathrm{s}}, \mathcal{V}, \sigma) \tag{2.5}$$

The low-dimensional characteristics are converted to high-dimensional features through the kernel approach. Use the Gaussian Kernel technique to process the characteristics in this research. The following equation 2.6 displays its mathematical formula:

$$\mathrm{G}(\mathcal{U}_{\mathrm{s}}, \mathcal{V}, \sigma) = \mathrm{e}^{\left(-\frac{(\mathcal{U}-\mu_{\mathcal{U}})^2+(\mathcal{V}-\mu_{\mathcal{V}})^2}{2\sigma^2}\right)} \tag{2.6}$$

where $\mu_{\mathcal{U}}$and $\mu_{\mathcal{V}}$ are the average of the input data $\mathcal{U}$ and $\mathcal{V}$, respectively, and σ denotes the user-defined kernel parameter. $\overline{\mathrm{SK}}$ is the reduced kernel matrix, s is the chosen kernel matrix computation percentage, and $\mathcal{U}_{\mathrm{s}}$ denotes the chosen s input data sample units.

Following equations 2.7 and 2.8 may be used to generate the Shrunk kernel matrix of training and testing data.

$$\overline{\mathrm{SK}}_{\mathrm{tr}} = \mathrm{G}\left(\mathcal{U}_{\mathrm{tr}_s}, \mathcal{U}_{\mathrm{tr}}, \sigma\right) \tag{2.7}$$

$$\overline{\mathrm{SK}}_{\mathrm{te}} = \mathrm{G}\left(\mathcal{U}_{\mathrm{te}_s}, \mathcal{U}_{\mathrm{te}}, \sigma\right) \tag{2.8}$$

where $\mathcal{U}_{\mathrm{tr}_s}$indicates the training observation class percentages $\mathcal{U}_{\mathrm{tr}}$. Choosing a certain proportion of training samples, along with speeding up the calculation of the kernel matrix, the reduced kernel methodology maintains the high-dimensional features of the original kernel method. The pseudocode of SKKNN is depicted in algorithm 2.1.

Algorithm 2.1 Pseudocode of SKKNN

Input: Training data matrix $\mathcal{U}_{\mathrm{tr}}$; Testing data matrix $\mathcal{U}_{\mathrm{te}}$; Number of training data $\mathbb{L}$; Parameter of SKKNN SK; Kernel parameter σ; The percentage of selected data for Shrunk kernel matrix s; Ensure: Prediction class Y.

Step 1 : Choose s percentage samples from the training data for each class as X_{tr} P;

Step 2 : Using equation $\overline{\mathrm{SK}}_{\mathrm{tr}}$, determine the reduced kernel matrix for the training features.

Step 3: Create the reduced kernel matrix using equation $\overline{\mathrm{SK}}_{\mathrm{te}}$ to test features.

Step 4 : Loop: for i $\in 1, \ldots, \mathbb{L}$ do

Find the distance between training and testing samples

end for

Step 5 : Sort the distance in the ascending order;

Step 6 : Selecting the top K vectors from the sorted collection to serve as an index Step 7 : Set the forecasting the class label Y based on the most frequent class of processed index.

Step 8 : end for return Y

3. RESULTS AND DISCUSSION

The working platform of MATLAB was used to construct the suggested breast cancer detection system. The confusion matrix is used to estimate the proposed framework's performance indicators. For both perspectives, the performance metrics are calculated. Performance metrics include the following terms: TP (True Positive): The abnormal region is appropriately classified as abnormal by the classifiers. False Positive (FP): The classifiers misclassify a

region that is normally occurring as abnormal. TN (True Negative): The normal zone is appropriately classified as normal by the classifiers. FN (False Negative): The aberrant region is misclassified as normal by the classifiers. Here used a variety of measures, including precision, recall, F1-score, accuracy, sensitivity, false positive rate (FPR), false negative rate (FNR), and specificity, to evaluate the transfer learning models. Each metric's detail and mathematical expression are provided below:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad \text{Sensitivity or Recall} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \qquad \text{AUC} = \frac{\text{sensitivity} + \text{Specificity}}{2}$$

$$\text{Precision} = \frac{TP}{TP+FP} \qquad \text{Prediction Error} = \frac{FP+FN}{FP+FN+TP+TN}$$

$$\text{F1 Score} = 2\,\frac{\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \qquad \text{FPR} = \frac{FP}{FP+TN}$$

$$\text{FNR} = \frac{FN}{FN+TP}$$

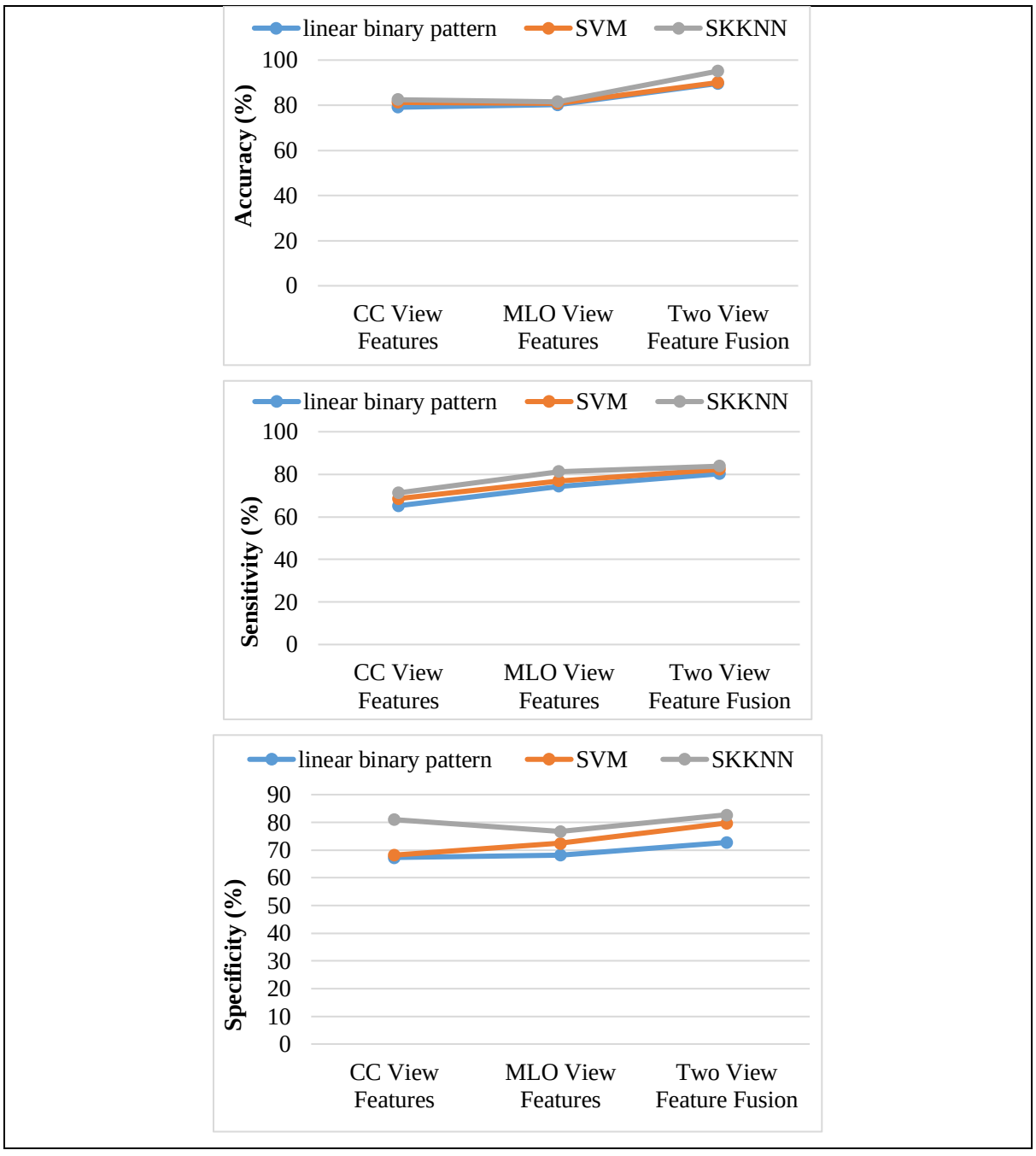
$$\text{Younden}'\text{sindex} = \text{Sensiivity} - (1 - \text{Specificity})$$

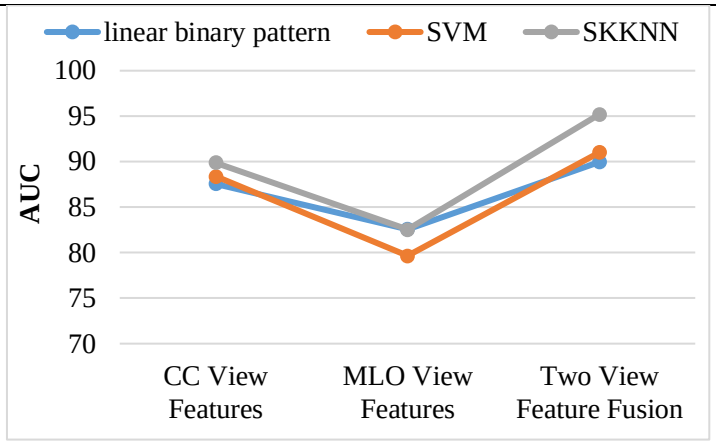
The comparison of the suggested two-view feature SKKNN's performance with that of the single view feature is shown as given in Table 3.1

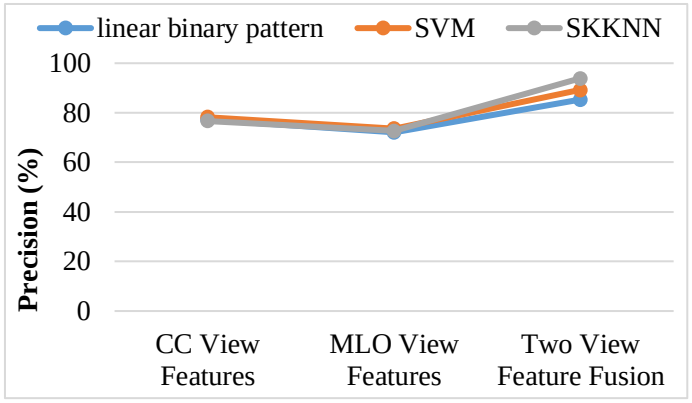
Table 3.1 Numerical Results of BC Detection

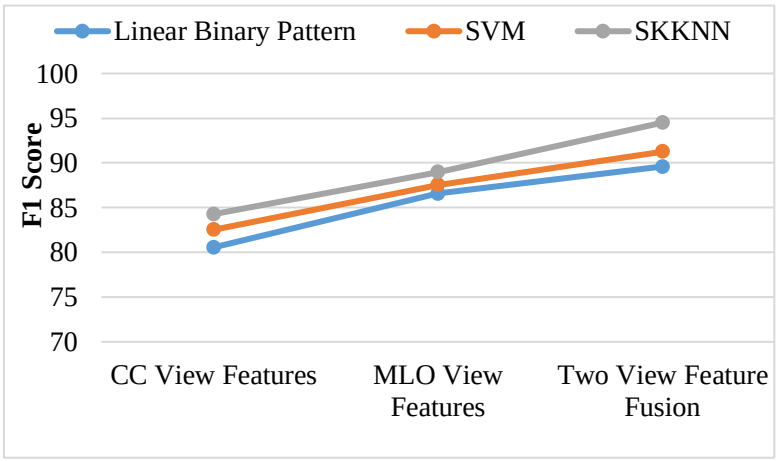
SKKNN	Accuracy (%)	Sensitivity	Specificity	AUC	Precision	Prediction Error	F1 Score	FPR	FNR	Youden's Index
CC View Features	81.8	71.2	81.09	89.9	76.7	12.0	84.2	18.9	17.6	179.8
MLO View Features	81.9	81.3	76.72	82.5	72.7	18.	88.9	23.2	11.6	165.1
Two View Feature Fusion	82.5	83.8	82.75	95.2	93.9	4.5	94.5	6.44	3.1	190.4

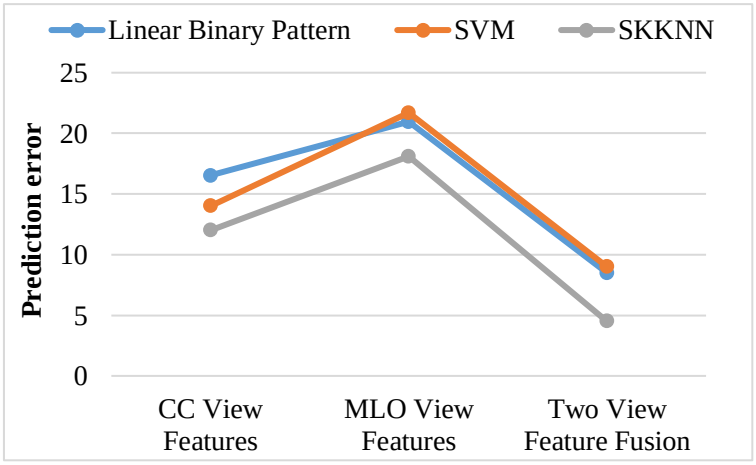
The comparison graph of accuracy, sensitivity, specificity, AUC, precision and f1-score for each is appeared in Figure 3.1




linear binary pattern
SVM
SKKNN
AUC
100
95
90
85
80
75
70
CC View Features
MLO View Features
Two View Feature Fusion


linear binary pattern
SVM
SKKNN
Precision (%)
100
80
60
40
20
0
CC View Features
MLO View Features
Two View Feature Fusion


Linear Binary Pattern
SVM
SKKNN
F1 Score
100
95
90
85
80
75
70
CC View Features
MLO View Features
Two View Feature Fusion


Linear Binary Pattern
SVM
SKKNN
Prediction error
25
20
15
10
5
0
CC View Features
MLO View Features
Two View Feature Fusion

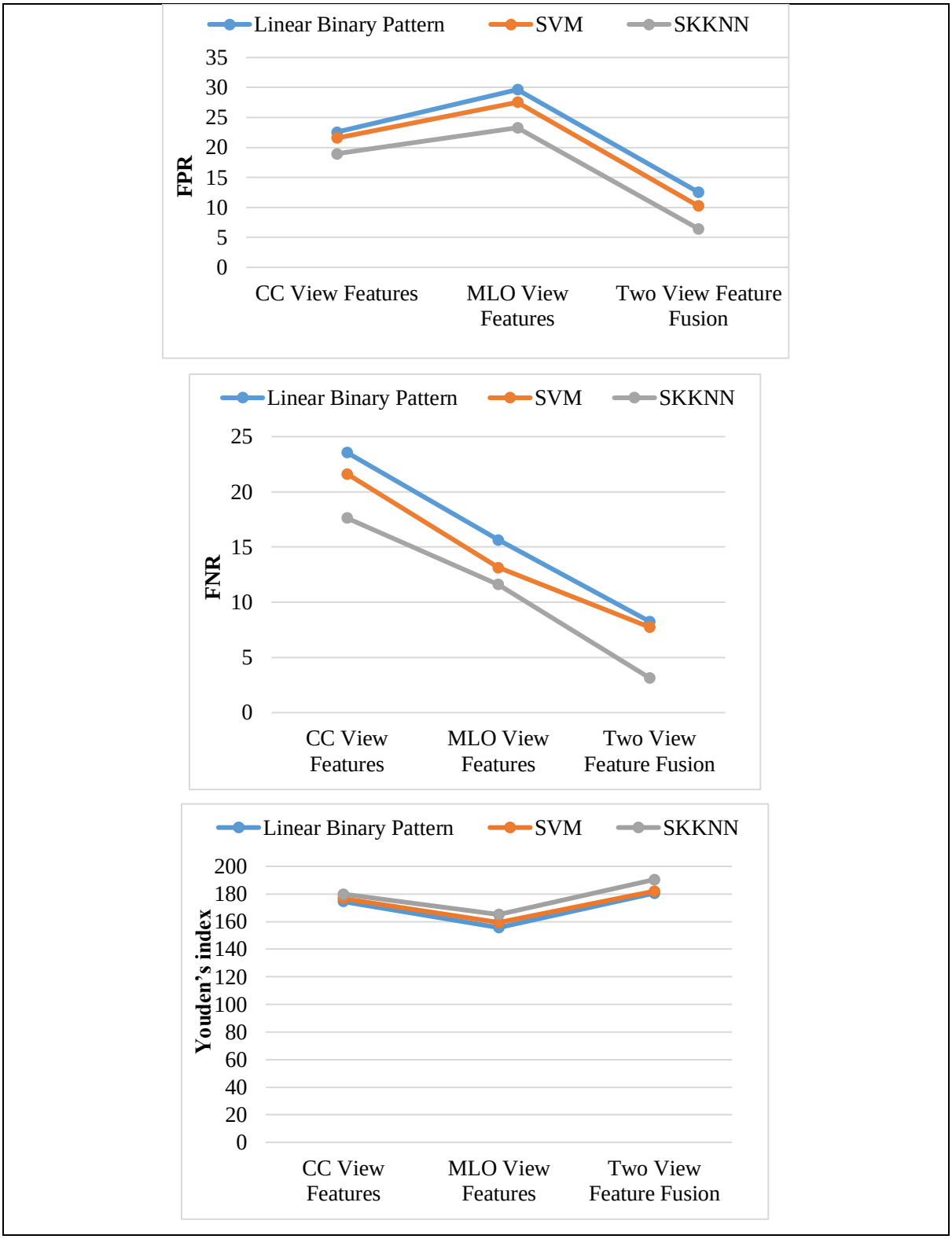


Figure 3.1 Chart Comparison of Linear Binary Pattern, SVM and SKKNN

The accuracy, senstitivity , specificity, AUC, precision, F1-measure for SKKNN was 82.56%, 83.8%, 82.75%, 95.2%, 93.93%, and 94.5% which are higher comparing single view

features. The FPR, FNR and prediction error for SKKNN was 6.44, 3.15 and 4.5 which is lower comparing single view features. The Youden's Index for SKKNN was 190.41 which is higher comparing singe view features. In spite of the fact that the execution of the fundamental show is marginally distinctive on size of information, the proposed system has moved forward the execution of the essential show.

4. CONCLUSION

In comparison to single view, using a double view yields superior results. Breast cancer detection in this chapter is done using MLO and CC view mammography images. The segmentation of regions is done using the Adaptive K-means approach. In this study, the KMC-GF feature extraction approach is used. The benefit of the KMC-GF approach is that it gives shape and orientation characteristics while also extracting fine texture features from the clustered area. KMC-GF characteristics provide a good result for breast cancer early diagnosis when compared to GLCM features. The view of mammography images is then classified as normal or malignant using the SKKNN classifier. As a result, this technique could be employed in the medical field to diagnose breast cancer and also produce few false positive results. The main benefit of this strategy is that it lightens the burden for radiologists.

References

1. P. Kulkarni, (2019). Fine grained classification of mammographic lesions using pixel NGRAMS, AJCT.
2. A.O. Ibrahim, (2020). Classification of mammogram images using radial basis function neural network, Emerging Trends in Intelligent Computing and Informatics. Springer.
3. R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, A. Alavudeen Basha, (2019), Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform, Measurement 146, 800–805, https:// doi.org/10.1016/j.measurement.2019.05.083, ISSN 0263-2241.
4. Bushra Mughal, (2019). Early Detection and Classification of Breast Tumor From Mammography, COMSATS Institute of Information Technology, Islamabad.
5. N.S. Lakshmi, M. Chitra, E. Christina, (2019), Breast cancer detection using KNN classifier, IRE J. 2 (10), 76–79.

6. P. Kaur, G. Singh, P. Kaur, (2019). Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification, Inf. Med. Unlocked 100239, https://doi.org/10.1016/j. imu.2019.100239.

7. H. Viswanath, L. Guachi-Guachi, and S. P. Thirumuruganandham, (2019). Breast Cancer Detection Using Image Processing Techniques and Classification Algorithms Breast Cancer Detection Using Image Processing Techniques and Classification Algorithms, 1(1).

8. A.A. Kayode, N.O. Akande, A.A. Adegun, M.O. Adebiyi, (2019). An automated mammogram classification system using modified support vector machine, Med Devices (Auckl) 12, 275–284, https://doi.org/10.2147/MDER. S206973.

9. Arooj, S., Zubair, M., Khan, M. F., Alissa, K., Khan, M. A., & Mosavi, A. (2022). Breast cancer detection and classification empowered with transfer learning. Frontiers in Public Health, 10.

10. F. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, (2016). A dataset for breast cancer histopathological image classification," IEEE Transactions of Biomedical Engineering, 20(16).

11. Dewangan KK, Dewangan DK, Sahu SP, Janghel R. (2022). Breast cancer diagnosis in an early stage using novel deep learning with hybrid optimization technique. Multimed Tools App, 81,13935–60. 10.1007/s11042-022-12385-2.

12. Lin RH, Kujabi BK, Chuang CL, Lin CS, Chiu CJ. (2022). Application of deep learning to construct breast cancer diagnosis model. Appl Sci, 12,1957. 10.3390/ app12041957.

13. Alruwaili M, Gouda W. (2022). Automated breast cancer detection models based on transfer learning. Sensors.

14. Alshammari MM, Almuhanna A, Alhiyafi J. (2021). Mammography image-based diagnosis of breast cancer using machine learning: a pilot study. Sensors, 22,203. 10.3390/s22010203.

15. Nasiri, Y., Hariri, M., & Afzali, M. (2015). Breast cancer detection in mammograms using wavelet and contourlet transformations. 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI), IEEE, 923-926).

16. Qayyum, A., & Basit, A. (2016). Automatic breast segmentation and cancer detection via SVM in mammograms. nternational conference on emerging technologies (ICET), IEEE, 1-6.

17. Jayandhi, G., Jasmine, J. L., Seetharaman, R., & Joans, S. M. (2022). Breast Cancer Prediction Based on Mammographic data by Hybrid Resnet and Decision Tree. International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), IEEE, 234-238.

18. Diaz, R. A. N., Swandewi, N. N. T., & Novianti, K. D. P. (2019). Malignancy determination breast cancer based on mammogram image with k-nearest neighbor. 1st International Conference on Cybernetics and Intelligent System (ICORIS), IEEE, 1, 233-237.

19. Ghongade, R. D., & Wakde, D. G. (2017). Detection and classification of breast cancer from digital mammograms using RF and RF-ELM algorithm. 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech), IEEE, 1-6.

20. Jothilakshmi, G. R., & Raaza, A. (2017). Effective detection of mass abnormalities and its classification using multi-SVM classifier with digital mammogram images. International Conference on Computer, Communication and Signal Processing (ICCCSP), IEEE, 1-6.

21. Al-Ayyoub, M., AlZu'bi, S. M., Jararweh, Y., & Alsmirat, M. A. (2016). A gpu-based breast cancer detection system using single pass fuzzy c-means clustering algorithm. 5th International Conference on Multimedia Computing and Systems (ICMCS), IEEE, 650-654.

22. Nedra, A., Shoaib, M., & Gattoufi, S. (2018). Detection and classification of the breast abnormalities in Digital Mammograms via Linear Support Vector Machine. IEEE 4th Middle East Conference on Biomedical Engineering (MECBME), IEEE, 141-146.

23. Chitradevi, B., & Srimathi, P. (2014). An overview on image processing techniques. International Journal of Innovative Research in Computer and Communication Engineering, 2(11), 6466-6472.

24. Acharya, T., & Ray, A. K. (2005). Image processing: principles and applications. John Wiley & Sons.

25. Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., ... & Lungren, M. P. (2020). Preparing medical imaging data for machine learning. Radiology, 295(1), 4-15.

26. Cheng, H. D., Jiang, X. H., Sun, Y., & Wang, J. (2001). Color image segmentation: advances and prospects. Pattern recognition, 34(12), 2259-2281.

27. Cloete, I., & Zurada, J. M. (Eds.). (2000). Knowledge-based neurocomputing. MIT press.

28. Islam, M. J., Wu, Q. J., Ahmadi, M., & Sid-Ahmed, M. A. (2007). Investigating the performance of naive-bayes classifiers and k-nearest neighbor classifiers. In 2007 international conference on convergence information technology (ICCIT 2007), IEEE, 1541-1546.

29. Ahmad, H. A., Yu, H. J., & Miller, C. G. (2014). Medical imaging modalities. In Medical imaging in clinical trials , Springer, London, 3-26.

30. Mattila, J., Koikkalainen, J., Virkki, A., Simonsen, A., Van Gils, M., Waldemar, G., ... & Alzheimer's Disease Neuroimaging Initiative. (2011). A disease state fingerprint for evaluation of Alzheimer's disease. Journal of Alzheimer's Disease, 27(1), 163-176.