Classification of Mammograms Using Morphological Segmentation

Jagreet Kaur[1], Dr. Kulwinder Singh[2]
[1]Research Scholar, IKGPTU, Jalandhar, India.
[2]Professor and HOD(IT), GNDEC, Ludhiana, India

Abstract- In some medical applications where a tissue of interestcovers a large fraction of the image or a prior knowledge on the Region of Interest (ROI) is available, extracting features from the image is sufficient. However in the general case, one would like to identify features for each tissue in the image. This would require prior image segmentation. Medical image segmentation is one of the most challenging problems in medical image analysis and a very active research topic. Therefore, there is no algorithm available in the general case for isolating medical image regions [1]. This paper presents an accurate method for extracting texture features from mammographic images for classification. As a first step, segmentation is used which is based on morphological and filtering operations. Extraction of 26 texture features from those regions. Classification of the images is done into two classes: normal and abnormal. SVM classifier is used for classification. Two kernels named Gaussian RBFkernel and Polynomial kernel are used.Experiments held on MIAS (Mammographic Images Analysis Society) dataset.

Keywords – MIAS, SVM, gamma, complexity, accuracy, ROI.

I. INTRODUCTION

BREAST cancer is a major health risk for women, often leading to death that can be avoided by early detection. Prevention of the disease is not yet possible since the cause is not understood, and the success of treatment is dependent on how early in its development the cancer is detected [2].

Among the different imaging modalities used for breast cancer detection, mammography remains the key screening tool for the detection of breast abnormalities [3]. In [4], it has been shown that lesion detection is improved by up to 15% through double reading. A Computer Aided Diagnosis (CAD) system is ideally suited in assisting the radiologist for this task [2].

CAD systems use a digitized mammographic image that can be obtained from either a conventional film mammogram or a digitally acquired mammogram. The computer software then searches for abnormal areas of density, mass, or calcification that may indicate the presence of cancer. The CAD system highlights these areas on the images, alerting the radiologist to the need for further analysis [5].

This paper concentrates on classification of mammograms. Classification of abnormal from normalmammograms is treated as a two-class pattern classification problem, i.e., a mammogram under consideration is either abnormal or normal. Malignant or Benign class is considered as '**abnormal**' and Normal class is considered as '**normal**'. Classification involves three steps: segmentation, feature extraction, assignment of class labels. Segmentation is based on morphological operations. Feature extraction step involves extraction of 13 texture features. Assigning class labels is the classification of mammograms into normal and abnormal classes.

II. DATA COLLECTION

The data collection that has been used in our experiments was taken from MIAS [5]. It consists of 322 images, which belongto three big categories: normal, benign and malign. There are 208 normal images, 63 benign and 51 malign, which are considered abnormal. In addition, the abnormal cases are further divided in six categories: micro calcification, circumscribed masses, speculated masses, ill-defined masses, architectural distortion and asymmetry. In our experiments, 104 fatty images have been taken. Out of those 104 images, 38 images are abnormal and 66 images are normal.

Digital image mammogram

Segment the image
(ROI selection)

Extraction of features
from ROI

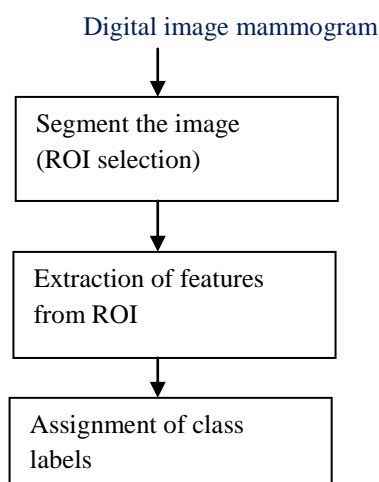Assignment of class
labels

Fig. 1 Steps for mammogram classification

III.   SEGMENTATION

The very first step is to segment an image into regions. Segmentation algorithm is based on morphological operations. Morphology is used as a tool for extracting image components that are useful in the representation and description of region shape, such as boundaries, skeletons etc [7]. Morphological transformation involves two images, rather than a single image. The first image is on which the morphological operations are performed whereas the second image is a structuring element. Segmentation involves:

1. Filtering.
2. Structuring element is defined.
3. Morphological Operations.

Various morphological operations that have been used for segmentation are: open, close, erosion, reconstruction, dilation. The image is segmented into a number of regions (two or more). Taking advantage of ground truth information, region can be viewed as normal tissue or containing mass. In the MIAS database, most of the images have atmost one region which contain mass if it is malignant or benign image. If the image is abnormal, by taking the help of ground truth, crop that portion that contain the mass and then that cropped portion has been segmented. All the regions that come after segmentation of that cropped portion would be abnormal regions. Normal image has been segmented into a number of regions. Each region is considered as normal.

(a)   (b)   (c)
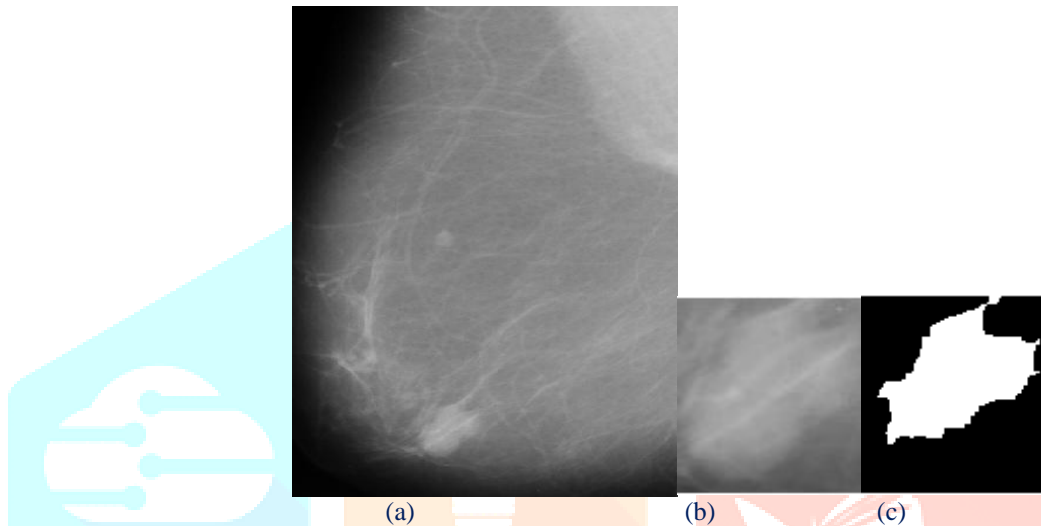
Fig. 2 Segmentation of a mammogram from MIAS (mdb005):
(a) Original Image (b) Cropped image (c) segmented image.

IV FEATURE EXTRACTION

After segmentation, features will be extracted from each region of the segmented image. There are 13 features which will be extracted from each region. Those features are:

1. Angular Second Moment (ASM)
2. Contrast
3. Variance
4. Inverse Difference Moment (IDM)
5. Sum Average
6. Sum Entropy
7. Sum Variance
8. Entropy
9. Difference Variance
10. Difference Entropy
11. Info. Measure of
12. Info. Measure of Correlation1
13. Info. Measure of Correlation2

All the 13 features values were taken in 4 directions 0^0, 45^0, 90^0 and 135^0 for each region. After extracting all the 13 features for all the regions of a number of images, minimum and maximum values were taken for each region, making total 26 features that were used for classification.

IV CLASSIFICATION

After extracting 26 features (13 minimum and 13 maximum for each region), classification has been performed. For classification, SVM classifier is used. Two kernels named Gaussian RBFkernel and Polynomial kernel are used.

Classification of images has been done into normal and abnormal categories. After checking the actual classes and predicted classes of the images, the classifier generates the results which include

1. Correctly classified images
2. Incorrectly classified images
3. Confusion Matrix
4. TP (True Positive) rate
5. FP (False Positive) rate

6. Accuracy

V EXPERIMENTAL RESULTS

In our experiment, 104 fatty images are taken. Out of 104 images, 38 images are abnormal and these 38 images gave only one region after segmentation and remaining 66 images are normal and these 66 images gave more than one region after segmentation, some images gave two regions, some gave three or four regions. After segmenting all the 104 images, total 237 regions have been extracted and used for classification.

SVM classifier is used, such that 4-fold cross-validation process is used. In Gaussian RBF kernel, two factors vary: *gamma* and *complexity* whereas in Polynomial kernel, two different factors vary named *degree* and *complexity*. Values for *gamma* taken are .13, .25, .50, 1.00, 1.25, 1.50, 2.00 and 2.5. For each value of gamma, *complexity* values taken are 1, 10, 100, 1000 and 10000. Values of *degree* taken are 2, 3, 4, 5, 6, 7, 8 and 9. For each value of *degree*, different values of *complexity* are taken (same as in Gaussian RBFkernel).
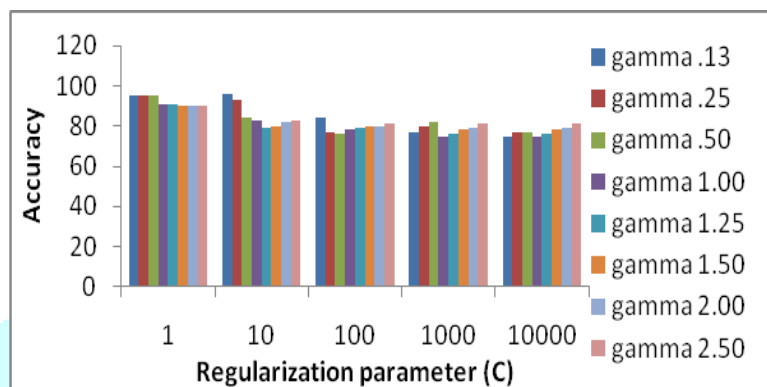

Fig. 2 Plot of Accuracy values versus regularization parameter C achieved by trained SVM classifier using a Gaussian RBF kernel with different gamma.

For Gaussian RBF kernel, it has been observed that gamma = .13 and C = 10 gave highest accuracy of 96% as compared to other values of gamma and regularization parameter. So gamma = .13 and C = 10 are considered to be optimal parameters for Gaussian RBF kernel.
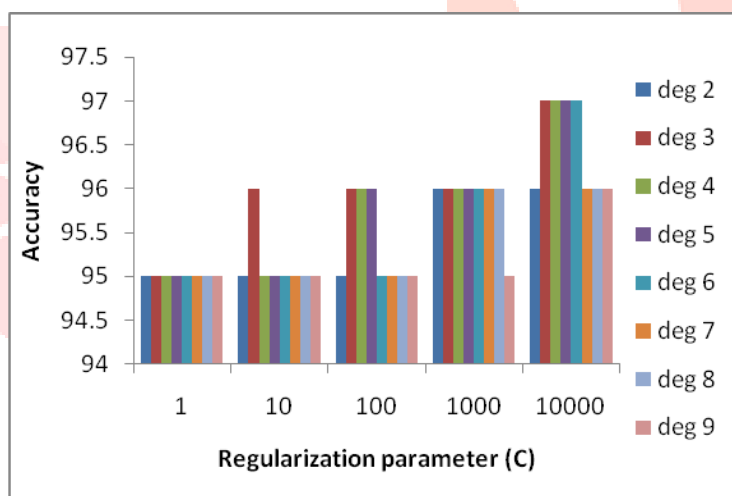

Fig. 3 Plot of Accuracy values versus regularization parameter C achieved by trained SVM classifier using a polynomial kernel with different degrees.

For Polynomial kernel, it is found that degree = 3, 4, 5 and 6, and C = 10000 gave highest accuracy of 97%. As clear from Fig. 3, it can be concluded that for deg = 3 and C = 10000 are the optimal parameters for Polynomial kernel.

Classifier	Sensitivity	Specificity	Accuracy
RBFkernel	100%	95.79%	96%
Polykernel	100%	96.8%	97%

Table I Classification rate comparison

From Table I, we can say that Polynomial kernel is more accurate than RBF kernel for classification of image.

VI CONCLUSIONS

Morphological based segmentation algorithm is used for segment images. 13 texture features (called Haralick texture features) have been extracted from each region for four directions: 0^0, 45^0, 90^0 and 135^0. Minimum and maximum values for each region have been extracted, so total 26 features are used for classification. Classification is done into two classes: normal and abnormal. For classification, SVM classifier has been used. Two kernels of SVM are used: Gaussian RBFkernel and Polynomial kernel and their success rates for classification have been compared. After comparison, it has been found that polynomial kernel gives better accuracy than RBF kernel.

ACKNOWLEDGEMENT

REFERNCES

[1] L. Tsochatzidis, K. Zagoris, N. Arikidis, A. Karahaliou, L. Costaridou, and I. Pratikakis, ''Computer-aided diagnosisof mammographic masses based on a supervised content-based image retrieval approach,'' PatternRecognit., vol. 71, pp. 106–117, Nov. 2017.

[2]D. Nie et al., ''Medical image synthesis with context-aware generative adversarial networks,'' in Proc. Int. Conf. Med. Image Comput. Comput.- Assist. Intervent., 2017, pp. 417–425.

[3] P. Costa et al. (Jan. 2017). ''Towards adversarial retinal image synthesis.''

[4]N. Coudray, A. L. Moreira, T. Sakellaropoulos, D. Fenyo, N. Razavian, and A. Tsirigos,''Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning,'' bioRxiv, p. 197574, 2017.

[5] http://www.radiologyinfo.org/en/info.cfm/

[6] http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html

[7] http://imaginis.com/breast-health/

[8]http://www.breastcancer.org/symtoms/

[9]http://nlp.stamford.edu/IR-book/html/htmledition/feature-selection/

[10]http://www.expresshealthcare.in/inimaging2010/

[11] http://www.cs.waikato.ac.nz/ml/weka/