

A Novel Approach for Mass Classification in Digital Mammogram Using Multiresolution Analysis and Adaptive Dimension Reduction

Veena U K¹, Catherine Jose², Jayakrishna Vijayakumar³

PG Scholar, Department of Computer Science, AJCE, Kanjirappally, India¹

PG Scholar, Department of Computer Science, AJCE, Kanjirappally, India²

Assistant Professor, Department of Computer Science, AJCE, Kanjirappally, India³

ABSTRACT: A novel approach for mass classification of breast cancer in digital mammogram based on Discrete Wavelet Transform (DWT) and Adaptive Dimension Reduction (ADR) technique is presented here. The mass classification is obtained by using DWT at various levels for decomposition. Then reduction of the high dimensional wavelet coefficients is done by using ADR. Dimension reduction and unsupervised learning (clustering) are combined in ADR. Classification of the mammogram image into normal, benign or malignant is done using this reduced wavelet coefficients as features. In the proposed system KNN (K-Nearest Neighbor) and SVM (Support Vector Machine) classifiers are used to classify the mammograms.

KEYWORDS: Mass classification, Adaptive Dimension Reduction, Digital Mammogram, K Nearest Neighbor, KNN, Support Vector Machine (SVM)

I. INTRODUCTION

Mammography is used for the detection of breast cancer, by the detection of characteristic lesions and micro calcifications [3]. Since breast cancer has been full-fledged and of more occurrence than in the early days, the need for mammography increases. The early detection of breast cancer is required to reduce the rate of criticality and death. The automation of detection and the classification of the same is becoming a requirement.

The currently available technique used for dimensionality reduction is the Principal Component Analysis (PCA)[3]. The number of principal components is comparatively less than or equal to that of original variables. This transformation is in such a way defined that the first principal component has the largest possible variance and each succeeding component has the highest variance possible under the constraint that it is orthogonal to the preceding components. Since PCA based subspace is fixed, the ADR subspace can be considered to have equal or better results as it is an adaptive learning technique.

Here we propose the use of ADR [2], for dimensionality reduction where the subspace is adaptively adjusted and integrated with the clustering process.

II. METHODOLOGIES

A. Discrete Wavelet Transform

DWT is used for decomposing the image into different subbands. In DWT [5], wavelets are sampled discretely. Four sub-band images referred to as low–low (LL), low–high (LH), high–low (HL), and high–high (HH) are resulted from 2D-DWT decomposition. The multi-resolution analysis [6], is done in different number of levels which can be chosen from 2 to 6.

B. Dimensionality Reduction

Dimensionality reduction or dimension reduction is the process of reducing the number of random variables under consideration, which can be divided into feature selection and feature extraction [3]. Feature extraction and dimension reduction can be combined in one step using principal component analysis (PCA), linear discriminant analysis (LDA), or canonical correlation analysis (CCA) techniques as a pre-processing step followed by clustering by K-NN on feature vectors in reduced-dimension space.

Feature Extraction

Feature Extraction transforms the data in the high-dimensional space to a space of fewer dimensions. The data transformation may be linear or non-linear.

Adaptive Dimension Reduction Technique (ADR)

Adaptive dimension reduction (ADR)[2], does combine dimension reduction and unsupervised learning (clustering) to improve the reduced data (subspace) adaptively. K-means clustering is used to generate class labels and LDA is used to do subspace selection. The integration of clustering process with the subspace selection process is done and then simultaneously clustering is done while the feature subspaces are selected. The clusters are discovered in low dimensional subspace which is then adaptively adjusted to global level.

Adaptive subspace selection using LDA and K-means

LDA is a supervised learning method where we need to know the class label for each data point previously. Since minimizing and maximizing is done in both LDA and K-means clustering [2], ways are required to combine them into a single framework. K-means clustering is used to generate class labels and LDA is used to do subspace selection. The final result of this learning process is that simultaneously the data are clustered while the feature subspaces are selected.

The initially chosen subspace may not be the optimal subspace, which we defined as the subspace spanned by the cluster centroids. ADM starts with the observation that the PCA subspace is necessarily not the best subspace to perform data clustering (we consider the best subspace to be the subspace spanned by cluster centroids). Iteratively it proceeds to find the best subspace.

C. K-NN classifier

The classifier input is set of k closest training samples in the feature space [4]. The class membership for the classifier is the output. Majority vote of the neighbors is used for classifying an object. Neighbors are correctly classified object sets.

D. SVM classifier

Support vector machines (SVMs) [1] are a supervised learning model which is used for classification and regression

analysis. Given a set of training samples as input, each of which is marked as belonging to one of two types, an SVM training algorithm builds a model that assigns new samples into one category or the other. For each given input it predicts, which of two possible classes the image is a member of. A classification task usually consists of training and testing data each of which consists of some data instances [6]. Each data instance in the training set consists of one target value (class labels) and several attributes (features).

III. PROPOSED METHOD

The proposed method is depicted in the figure below. The various phases are training and testing. In the training phase the features are extracted using ADR and stored in the database. In the testing phase, using the features stored in the database, classify the image as normal or abnormal.

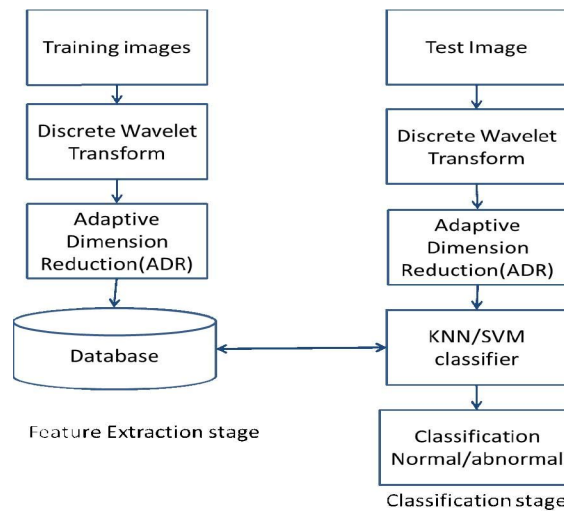


Figure 1: Figure 1: Block diagram of the proposed method.

A. Feature Extraction Stage

Feature extraction reduces the amount of resources that is required to define the large set of data correctly [1]. The extracted features are stored in a database. The input of the classification stage is the set of extracted features. In the proposed system, features are extracted by applying DWT and reduced by using ADR. Figure 1 shows the block diagram of feature extraction and classification stage of the proposed method.

The area which consists of the mass is extracted from the image that is taken from the MIAS database [7]. To train the classifier, the proposed features are extracted from this known Region of Interest (ROI). Initially, the decomposition of the extracted ROI image is done using DWT at various scales. In the proposed system, DWT decomposition level varies from 2 to 6. After the decomposition, ADR is used for dimension reduction. The result of the reduction is stored in the database which is taken as a feature. The initial phase of the database is created using the training images that includes normal and abnormal ones. The final phase database is created using the benign and malignant images.

B. Classification Stage

In the classification stage, the given test image is first tested for normal or abnormal category. At a predefined level the test image is decomposed using DWT, where the level of decomposition is the same for initial and final stage database. Then the high dimensional data of test image is reduced into a relatively low dimensional dataset by ADR. Testing the reduced dataset with the trained classifier, the result is used as a initial phase database. The abnormal image from the initial phase is then classified into benign or malignant. The reduced dataset is again tested with the trained classifier which uses the final phase classifier.

IV. CONCLUSION

In this paper, a novel approach for the mass classification in digital mammogram based on DWT and ADR is presented. The high dimensional data of wavelet decomposed mammogram image is reduced into low dimensional data set by using ADR. Combination of LDA and K-means clustering is used into the LDA-Km algorithm for adaptive dimension reduction. In this algorithm, class labels are generated using K-means and subspace selection is performed using LDA [2]. The Integration of clustering process with the subspace selection process is done followed by performing the learning algorithm. The reduced data set is used as features to classify the given mammogram images into normal or abnormal as well as benign or malignant.

REFERENCES

- [1] S Mohan Kumar and G. Balakrishnan "Multi Resolution Analysis for Mass Classification in digital Mammogram using Stochastic Neighbor Embedding". In ICCSP, 2013, pp101-105.
- [2] C. Ding and T. Li. "Adaptive dimension reduction using discriminant analysis and k-means clustering". In ICML, 2007.
- [3] <http://en.wikipedia.org/>
- [4] S. Mohan Kumar and G. Balakrishnan, "Classification of Microcalcification in Digital Mammogram using Stochastic Neighbor Embedding and KNN Classifier", In IJCA journal 2012.
- [5] Ibrahim Faye and Brahim Belhaouari Samir, "Digital Mammograms Classification Using a Wavelet Based Feature Extraction Method", IEEE conference on Computer and Electrical Engineering, 2009, pp 318-322.
- [6] S. Mohan Kumar and G. Balakrishnan, "The performance Evaluation of the Breast Mass Classification CAD system based on DWT, SNE and SVM", In IJETAE Oct 2013, pp 581-587.
- [7] J. Suckling, S. Astley, D. Betal, N. Cerneaz, D. R. Dance, S.-L. Kok, J. Parker, I. Ricketts, J. Savage, E. Stamatakis, and P. Taylor, Mammographic Image Analysis Society MiniMammographic Database, 2005. [Online]. Available: <http://peipa.essex.ac.uk/ipa/pix/mias/>