Quantitative and qualitative analyses of Dimension Reduction Methods effect on the classification of mammographic images

N. Hamdi[1], S. H. Bouazza[1], K. Auhmani1[2*], M. M. Hassani[1]

**[1] Cadi Ayyad University
Faculty of Sciences, Semlalia,
Marrakech 40000
Morocco**

**[2] Cadi Ayyad University
National School of Apllied sciences
Safi, BP 63, 46000, Morocco**

Abstract

This paper presents a comparative study of dimension reduction methods combined with wavelet transform. This study is carried out for mammographic image classification. It is performed in three stages: extraction of features characterizing the tissue areas then a dimension reduction was achieved by four different methods of discrimination and finally the classification phase was carried. We have experimented for this purpose K-Nearest neighbours classifier.

Results show the classification accuracy in some cases has reached 100%. We also found that generally the classification accuracy increases with the dimension but stabilizes after a certain value which is approximately d=60. We also present the results as a projection of onto a two-dimensional space. In some cases we observed a clear separation between normal images and abnormal ones.

Keywords: *Dimension reduction, Classification, Feature extraction, Mammographic images.*

1. Introduction

The feature selection problem often arises when it comes to consider a very large number of variables. In recent years the need has evolved with the manipulation of very large databases and especially in areas such as genetic field and image processing field [1]. Hence the need to reduce the number of variables (features) exploited for classifying an image or an object.

Dimension reduction algorithms used for this purpose are therefore trying to find a projection of the data in a space of reduced dimension, while preserving the information contained therein. This projection may be linear or nonlinear.

Our goal in this paper is to perform a comparative study of the effect of dimension reduction methods on the classification accuracies of breast cancer (mammographic image). We will present the performance of the following methods: locality-preserving projection (LPP), locally linear embedding (LLE), Isometric Mapping (ISOMAP) and spectral regression (SR). We will present the results of the classification versus the reduced space dimension, the method used for the reduction and the type of the wavelet transform used for the extraction of features.

The classification task of mammographic image has been carried out through different stages. The first step is to prepare the vector containing the variables describing the image. The second phase aims to reduce the size (dimension) of this vector. The last step is to find the class of the incoming picture. This last step begins with a learning phase of the classifier and then returns the final result.

2. MATERIALS AND METHODS

2.1 Features extraction

Initially, the features are calculated to form the feature vector for subsequent learning step. These features were calculated on a set of two classes labelled images (normal and abnormal). These images are firstly pre-processed and transformed into the frequency domain by three types of wavelet transforms.

The following features are extracted:
- A vector of 24 texture descriptors is formed from a multi-level histogram of 3, 5, 7, and 9 bins [2].
- A vector of descriptors is calculated from the first order of statistical moments which are based on one-dimensional distribution of the gray levels of an image. This distribution can be described by the moments of order 1 to 4: Mean, Variance, and Fisher coefficients, Skewness and Kurtosis. However, such indicators do not take into account the spatial dependencies, which are inseparable from the concept of texture. They are therefore not sufficient to fully characterize the texture and are often associated with dynamic statistical descriptors [3]. We calculate the mean, the variance, the Skewness and the kurtosis of the image, in four different directions (0, 45, 90 and 135 degrees). Then an histogram is calculated for each of the three bins resulting series, which gives 4 * 4 * 3 = 48 descriptors.
- Three of the six parameters introduced by Tamura are used to characterize textures by Tamura method, namely, coarseness, contrast and directionality. We also calculate an histogram of 3 bins on the coarseness. In total, this group of descriptors forms a vector of six texture features [4,5].
- Radon's characteristics are calculated for angles 0°, 45°, 90 ° and 135°. We also calculate the histogram of 3 bins for the four series, which gives a vector of 12 features [6, 7].
- Zernike's moments of order n = 12 are calculated corresponding to 49 features [8].
In order to exploit the advantage of each feature, we merge them into a single feature vector. Depending on the used wavelet transform, there will be a different number of features. If the used wavelet transform is the double density discrete wavelet transform (3DWT), the image is divided into nine sub-images, so the number of features is 9 * 139=1251. If the transformation is based on the stationary wavelet transform (SWT) or the discrete wavelet transform (DWT), the original image is decomposed into four sub-images, and the number of calculated features is 4*139= 556. Each original image will be represented by a group of 556 or 1251 genes. Our original database consisted of 107 patients divided into two classes: 56 tumoral samples and 51 normal samples.

2.2 Dimension reduction

The number of features appears to be very large, which may affect the results of the classification. For this reason, we thought to make a reduction of the number of features to be used by eliminating redundant and irrelevant ones.
We applied four approaches of discriminant analysis, introduced above, on the feature vector to extract a sub-optimal space improving learning and classification. We examine the effectiveness of the reduced space in relation with the considered classification techniques. We will also perform a comparison between the wavelet transforms in order to form the final combination of best compromise: best wavelet transform / best reduced space.

2.3 Benchmark

Our benchmark (Figure 1) is divided in four parts. First, the original mammographic image will undergo preprocessing to reduce noise, enhance the features and enhance the presentation. The next part deals with the extraction of features as described above. In the third part, the studied dimension reduction methods will be applied to the complete dataset. The low-dimensional datasets will then be classified by two different classifiers, namely KNN and decision tree. To evaluate and compare the performance of each method, the classification accuracies of KNN on decision tree will be presented versus reduced dimension and wavelet transform type.

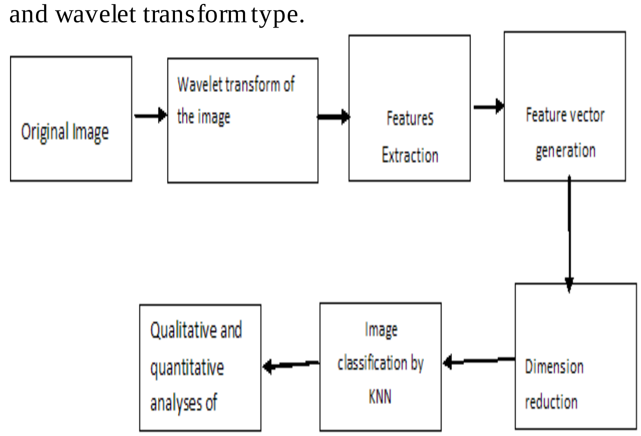


Fig 1 Benchmark of the studied methods

2.4 Dataset

The methods were tested on the original mammographic images of MIAS database [9]. These images of original size 1024*1024 are accompanied by the opinion of the radiologists that specify the class of the image, the type of microcalcifications, the center of clusters and other informations. We decomposed the images into regions of interest (ROI) of size 256 * 256 and applied our benchmark on the ROI.

IJCSI
www.IJCSI.org

3. RESULTS AND DISCUSSION

3.1 Quantitative analyses of Dimension influence

Figures 2, 3 and 4 show the effect on classification performance for different wavelet transforms and different dimension reduction methods. We observe an interesting general result: the classification accuracy increases with the reduced space dimension.

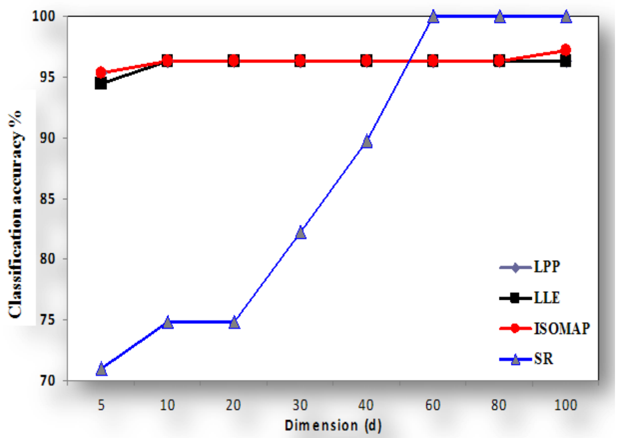


Fig 2: Effect of different dimension reduction methods on classification accuracy; features extracted from transformed image by the double-density discrete wavelet transform 3DWT.

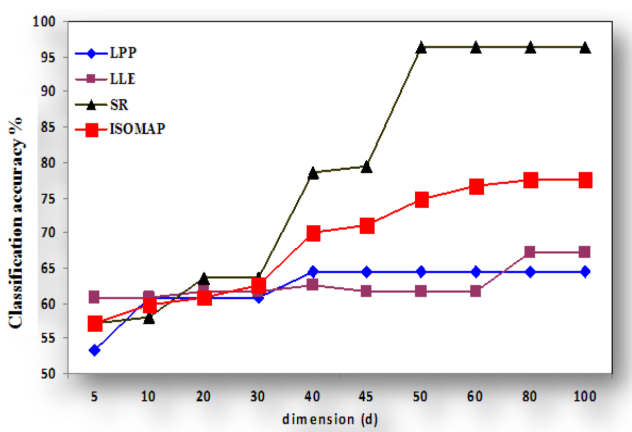


Fig 3: Effect of different dimension reduction methods on classification accuracy, features extracted from transformed image by the discrete wavelet transform DWT

Figure 2.presents the classification accuracy versus space dimension (d). Features are extracted from image transformed by the double-density wavelet transform 3DWT. We remark that the best performance is 100% (all the 107 samples are correctly classified) obtained with Spectral regression method (SR) for space dimension d=60. However, the comparison of the accuracies corresponding to LLP, LLE and ISOMAP, shows that reducing the space for small size (d=5) the classification accuracy of 95.3% is held.

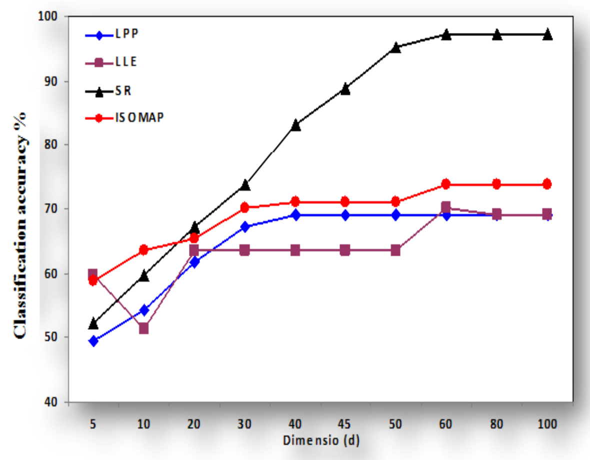


Fig 4 Effect of different dimension reduction methods on classification accuracy: features extracted from transformed image by the stationary wavelet transform SWT.

Figures 3 and 4 show the accuracy versus dimension d, features are extracted from mammographic images transformed by discrete wavelet transform (DWT) and stationary wavelet transform (SWT) respectively. The classification accuracy is less performant than the results corresponding to the 3DWT. The accuracy reaches 96,2% and 97,2 % for DWT and SWT respectively in a reduced space with SR method

Figures show clearly that for small dimensions, i.e. d=10, 20, 25, the accuracy with SR method increases with dimension up stabilization for d=60. This means that classification task is easier with high dimensions than low dimensions. Moreover, we find this behavior for the other methods: classification accuracy increases slightly as the data size increases.

3.2 Qualitative analyses

Figures 5, 6, 7 and 8 show the projections of all the classified samples on the first two dimensions plan. Normal images are represented by the "*" symbol and tumor images represented with a "+" symbol. The parameters used to generate the projections are as follows: 5, 8, 15 nearest neighbors to calculate graph for LLE, ISOMAP and LPP methods and α= 0.2 for spectral regression method.

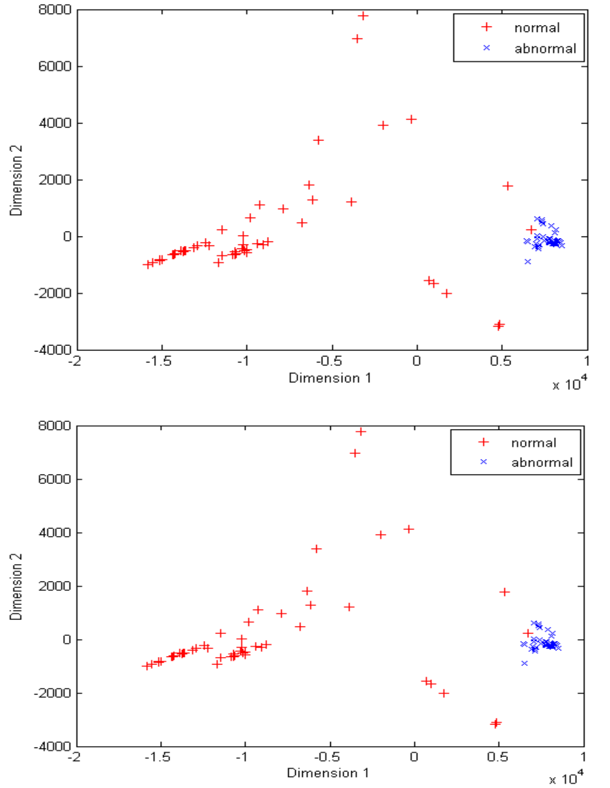
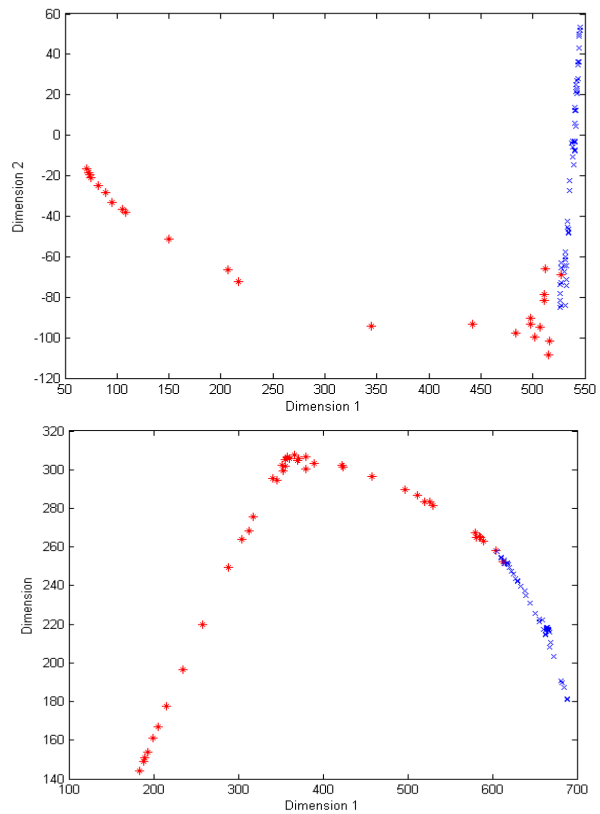


Fig 5 Classified data Projection by ISOMAP method in bidimensional space for k=5, k=15



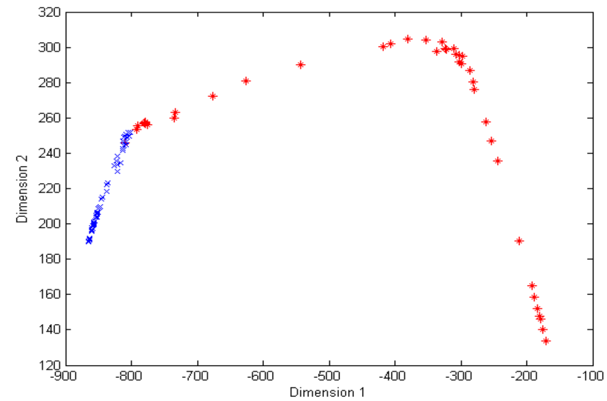
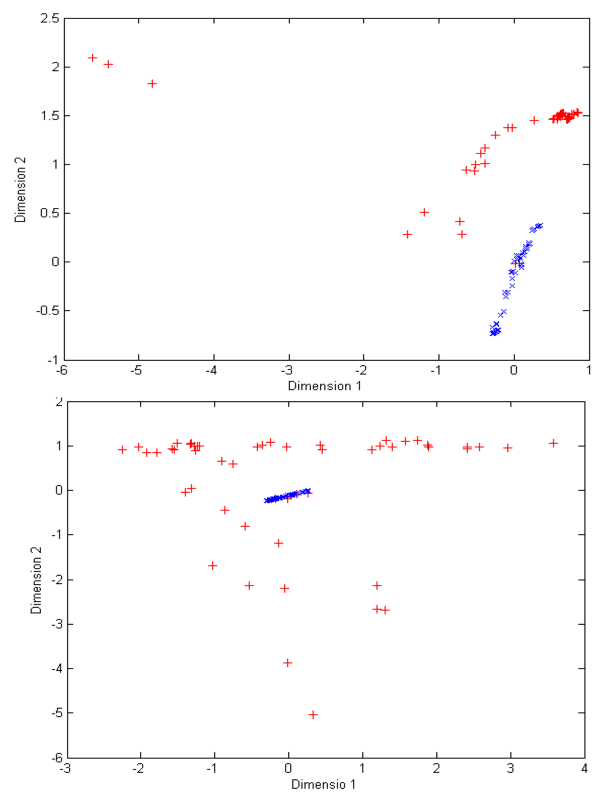


Fig 6 Projection Classified data Projection by LPP method in bidimensional space for k=5, 8, 15 respectively.

Several observations can be made on the results: We see in the projections that the separation of normal images and tumor ones is sharper with nonlinear treatments such as LLE, LPP and Isomap. For most of the points; classification seems obvious except for some of them located in the separating zone of the two classes. We also note that for the Spectral Regression method the classification is impossible for two dimensions (Figure 8) which confirms the result announced in the previous section. Our experiences had shown us that the value of k that gives a better separation of classes is k = 8. This result is in agreement with other authors [10].



IJCSI
www.IJCSI.org

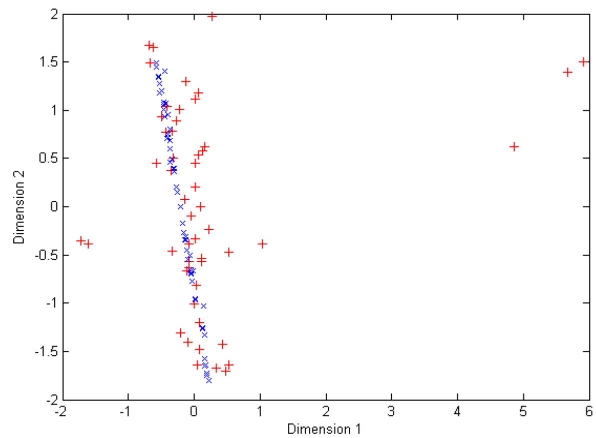


Fig 7 Classified data Projection by LLE method in bidimensional space for k= 5, k=8, k=15

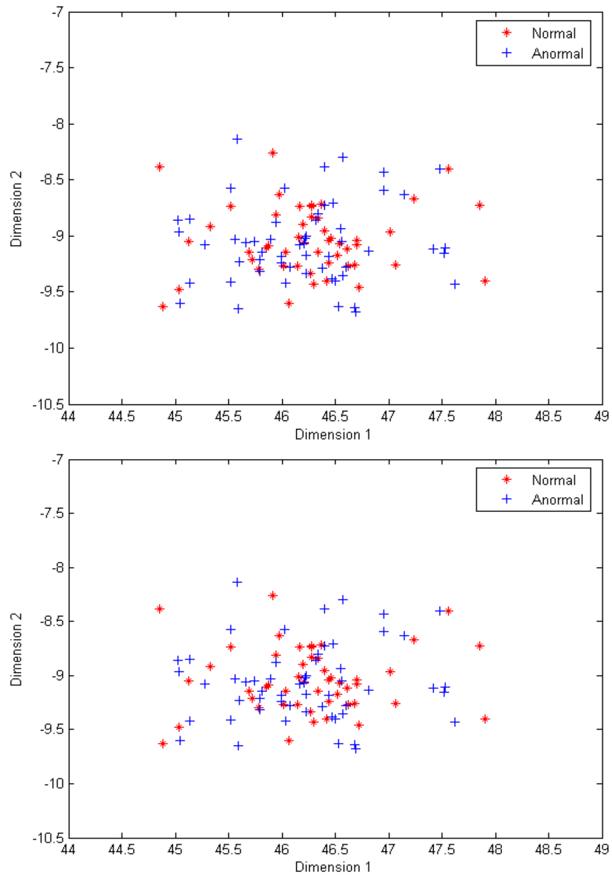


Fig 8 Classified data Projection by SR method in bidimensional space for k= 5 et k=15

3.3 Interpretation

All the four algorithms (of dimension reduction) that we have used in the experiment try to optimize an objective function. The spectral regression method performs a discriminant analysis of variables maximizing inter-class dispersion (dispersion of cluster centers) and minimizing intra-class dispersion (dispersion in a class about its center), whereas the other methods depend on the number of k nearest neighbors used to construct the graph.

Spectral regression method achieves better classification than the other methods, which verifies the effectiveness of regulation for classification [11]. This is due to the fact that it structures data to make different classes disjoint. However, this transformation associated with a classification model that operates its own data processing (e.g SVM) can cause a problem of over-learning and decreased performance.

The presented results allow us to determine the optimal dimensions as a compromise between the size of the feature vector and the classification performance: from the 1251 features the spectral regression retains an average of 60 features to describe each class.

4. CONCLUSION

The dimensionality reduction algorithms attempt to find a projection of the data in a space of smaller dimension, while preserving the information contained in those. In this paper, we presented a process of tissue classification applied to the evaluation of the pathologic state of breasts. This process was performed in three stages: extraction of features characterizing the tissue areas then a dimension reduction was achieved by four different methods of discrimination and finally the classification phase was carried out

During the experimental phase, we compared the different techniques of dimension reduction associated with a wavelet transform applied to the image before the features extraction process. We have late compared the performance of two classifiers KNN and decision tree.

We have reached a classification accuracy of 100% for some combinations. We also found that generally the classification accuracy increases with the dimension but stabilizes after a certain value which is d=60.

We did the projection of results on a two-dimensional space for each dimension reduction method and found that the LPP, LLE and Isomap methods can discriminate many samples according to their classes for low dimensions (d=2). While for the SR method images are indistinguishable confirming that for this method gives good results you have to work with higher dimensions.

REFERENCES

[1] Ali El Akadi, « Contribution à la sélection de variables pertinentes en classification supervisée: application à la

sélection des gènes pour les puces à ADN et des caractéristiques faciales » ; these de doctorat, Faculté des sciences Rabat, 2012.

[2] Hadjidemitriou, E., Grossberg, M. D., & Nayar, S. K.. Multiresolution histograms and their use for recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26 (7), 831-847. (2004)

[3] S. HERLIDOU, Caractérisation tissulaire en IRM par l'analyse de texture. Étude du tissu musculaire et de tumeurs intracrâniennes, université de Renne1, 1999.

[4] H. TAMURA, S. MORI et T. YAMAWAKI. Texture features corresponding to visual perception, IEEE Transactions on Systems, Man and Cybernetics, SMC-8(6):460–473, 1978.

[5] P. HOWARTH et S. RÜGER. Evaluation of texture features for content-based image retrieval. In Proceedings of the International Conference on Image and Video Retrieval (CIVR'04), volume LNCS 3115, pages 326–334, Dublin, Ireland, jul 2004.

[6] Deans, S.R. Hough Transform from the Radon Transform_, IEEE Trans. On Patt. Anal. and Mach. Intell., Vol. PAMI-3, No. 2, pp. 185_188, 1981.

[7] Murphy, L.M. Linear feature detection and enhancement in noisy images via the Radon transform, Pattern Recognition Letters, No. 4, pp. 279_284, 1986.

[8] C.-W. Chong, P. Raveendran, and R. Mukun-dan. A comparative analysis of algorithms for fast computation of zernike moment. Pattern Recognition, 36:731-742, 2003.

[9] http://www.wiau.man.ac.uk/services/mias/miasweb.html

[10] Deng Cai, Xiaofei He, and Jiawei Han. "Spectral Regression for Efficient Regularized Subspace Learning", IEEE International Conference on Computer Vision (ICCV'07), Rio de Janeiro, Brazil, Oct. 2007

[11] T. Hastie, A. Buja, and R. Tibshirani. Penalized discriminant analysis. Annals of Statistics, 23:73–102, 1995.