International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 - 5161

Available at http://inpressco.com/category/ijcet

Research Article

Bladder Cancer Recognition: A Comparative Study

Shaymaa M. AlKashef [Á*], Abdelhameed Ibrahim[Á] , Hesham Arafat[Á] and Tarek A. El-Diasty[Á]

[Á]Computer Engineering & Systems Department, Mansoura University , Mansoura, Egypt.
[Á]Urology & Nephrology Center, Mansoura University, Mansoura, Egypt.

Accepted 01 April 2014, Available online 15 April 2014, **Vol.4, No.2 (April 2014)**

Abstract

Medical Imaging technologies, such as magnetic resonance imaging (MRI), have been widely applied to various medical procedures. Daily growth of medical data volume leads to human mistakes in the manual analysis and increases the need of automatic analysis. Therefore, applying some tools to collect, classify, and analyze the medical data automatically is a must. Medical imaging issues are so complex owing to high importance of correct diagnosis and treatment of diseases in healthcare systems. For these reasons, algorithms of automatic medical image analysis are used to help in increasing reliability and accurate understanding of the medical images. AI methods such as digital image processing and its combinations with other techniques like machine learning, fuzzy logic, neural networks, and pattern recognition are so valuable in visualization and analysis of medical images. The objective of this paper is to investigate the use of artificial intelligence techniques like artificial neural networks (ANN) algorithms such as (multilayer perceptron (MLP), Jordan /Eleman, Self Organizing Feature Map and support vector machine (SVM)) to early detect bladder cancer (diagnosis), to determine tumor staging (for sake of prognosis), and to assess the accuracy of MRI in T staging bladder cancer. A set of functional images taken by magnetic resonance (MR) is to be used. It was found that, multilayer perceptron neural network (MLP) gives better result than other algorithms.

Keywords: *Bladder Cancer Recognition (BCR), Artificial Neural Networks, Magnetic Resonance Imaging, MLP, SVM.*

1. Introduction

Bladder cancer starts in the cells of the bladder. The bladder, a part of the urinary system, is in the lower part of the abdomen. It is a hollow, balloon-shaped organ with a flexible, muscular wall. Urine is made by the kidneys. It is then passed to the bladder through 2 tubes called ureters. When the bladder is full, the muscles in the bladder wall tighten to force the urine out of the bladder. The urine empties out of the bladder and passes out of the body through a tube called the urethra.

Nearly all bladder cancers start in the lining of the bladder. Cancer that is only in the lining is called superficial bladder cancer. If the cancer spreads into the muscle wall of the bladder, it is called invasive bladder cancer (O'Brien *et al,* 2010). Bladder cancer is any of several types of malignancy arising from the epithelial lining (i.e., the urothelium) of the urinary bladder. Rarely the bladder is involved by non-epithelial cancers, such as lymphoma or sarcoma, but these are not ordinarily included in the colloquial term bladder cancer. It is a disease in which abnormal cells multiply without control in the bladder. The most common type of bladder cancer recapitulates the normal histology of the urothelium and is known as transitional cell carcinoma or more properly urothelial cell carcinoma. It is estimated that there are 383,000 cases of bladder cancer worldwide (GLOBOCAN, 2010 ; Filip Velickovski, 2010).

2. Problem definition

The MRI scans in each case contain different types of information relating to specific 3D voxels in the bladder. Bladder cancer is the fifth most common malignancy in the Western world after prostate, breast, lung and colorectal cancer. Its incidence directly increases with age reaching a maximum between 60 and 70 years. Bladder cancer is 2.5 times more common in men than in women. The most common bladder carcinogens are aromatic amines. A number of etiological factors is tied to the development of bladder tumors . However, cigarette smoking is by far the most important risk factor nowadays. More than 90% of bladder cancers are urothelial cell carcinomas (UCC). The other histological types concern squamous cell carcinoma (6-8%) and adenocarcinoma (1-2%). In Egypt and other countries with endemic spread of schistosomiasis, squamous cell carcinoma used to be the most frequent histological subtype. The majority of UCCs is found in the bladder. UCCs of the renal pelvis, ureter and urethra together account for less than 10% of the carcinomas. The most important symptom that raises suspicion for a bladder tumor is micro- or macroscopic hematuria. The hematuria is mostly painless and often

*Corresponding author: **Shaymaa M. AlKashef**

intermittent which frequently causes a delay in the diagnosis (Thomas K. *et al*, 2000).

The problem the paper addresses is to determine which of the four AI approaches (MLP, Jordan /Eleman, Self Organizing Feature Map and SVM) networks leads to the best accuracy in determining bladder cancer.

The main challenges foreseen are:

1. *Alignment*: All the modalities have different resolution and orientation. Furthermore, the patient may have moved slightly between the acquisition of one modality and the next.
2. *Feature extraction*: A large number of information can be present in a specific cell; however which information is useful for detecting cancer and separating it from normal tissue and tissue from other disorders. DWT is used as a feature extraction technique.
3. *Normalization of features*: Once we have the useful features identified, how should they be normalized?
4. *Acquiring ground truth*: What should we consider as reliable ground truth so that we can measure the performance of the system.
5. *Recognition*: in this paper a comparative study is made between four techniques (MLP, Jordan /Eleman, Self Organizing Feature Map and SVM) networks (Voor moeder,2005) to capture the most effective approach for detecting bladder cancer.

3. Bladder Cancer Recognition Constraints

Due to the limited facilities; a set of constraints have been placed on the system to make the algorithm more manageable. These constraints are: (i) use jpeg 256*256 MRI images only; (ii) images of the bladder are taken from one direction; (iii) about 40 images were taken to every patient (iv) we process data for about 40 patients, and (v) only Egypt urine center images were used.

4. Proposed Algorithm

The proposed algorithm for bladder cancer recognition proceeds as follows:

1. Data Collection;
2. Image Preprocessing;
3. Image Enhancement;
4. Feature Extraction using DWT features with different mother functions;
5. Dimensionality Reduction (PCA);
6. Bladder Cancer Detection using the four approaches. See Fig.1

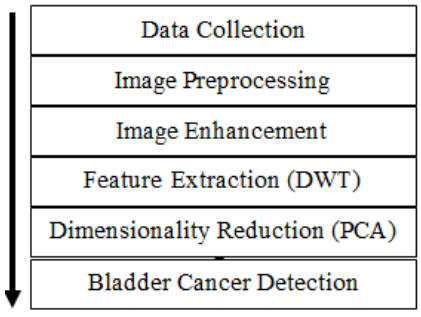


Fig 1. Flowchart of the proposed algorithm for bladder cancer detection

A. Data Collection

The images of bladder cancer were taken with a MRI device. On average, there are about 40 images for every patient. They were stored in color JPEG format. Matlab is used to convert the color JPEG images into gray scale raw format on the PC.

MRI images were taken from one angle for 40 patients, 40 images for each patient in different levels. Images are divided into 4 levels or layers .Each level contains images from different type of diseases (free,T1 stage,T2 stage,T3 stage,T4 stage).

Figures [2, 3, 4 and 5] show free images for levels 1, 2, 3 and 4.

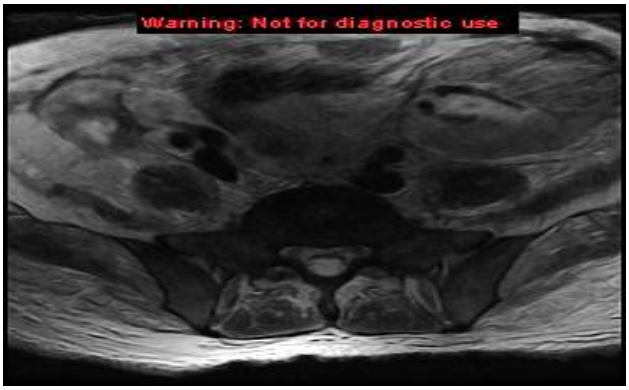


Fig 2. Free image level 1

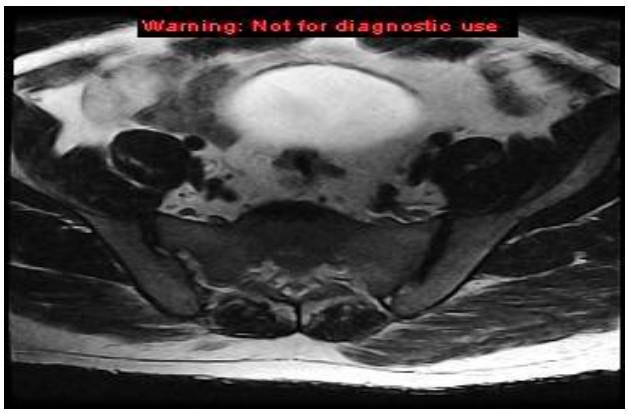


Fig 3. Free image level 2

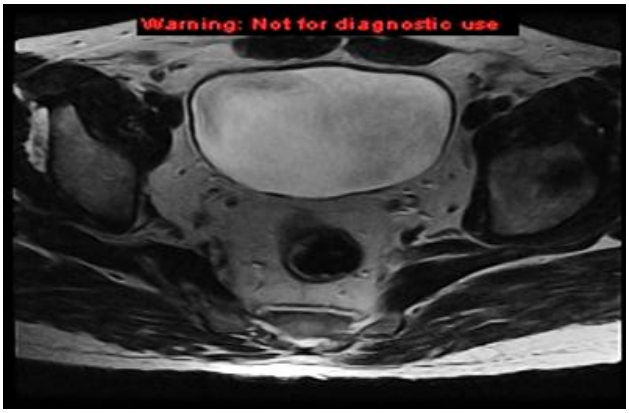


Fig 4. Free image level 3

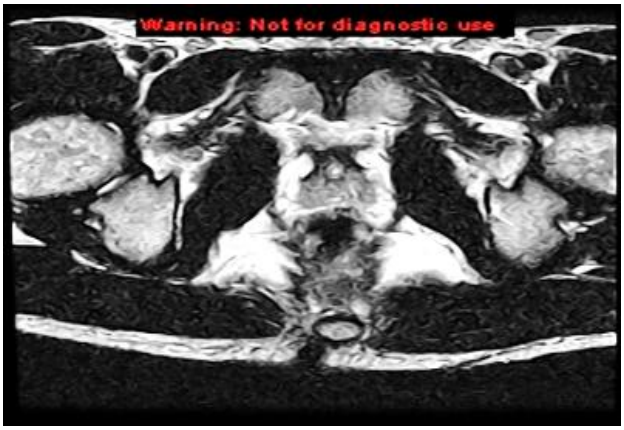


Fig 5. Free image level 4

B. Image Eenhancement

Wiener filter is used to filter out noise that has corrupted the signal. Wiener filters are characterized by the following:

Signal and (additive) noise are stationary linear stochastic processes with known spectral characteristics or known autocorrelation and cross-correlation. Performance criterion is the minimum mean-square error (MMSE) (Andrew R. *et al* ,2002). A noise-filtered image is shown in Fig.6.

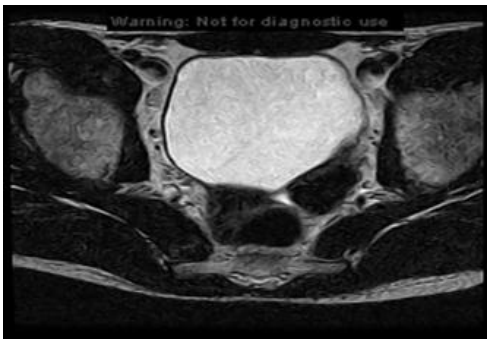


Fig 6-a. Image before wiener filter

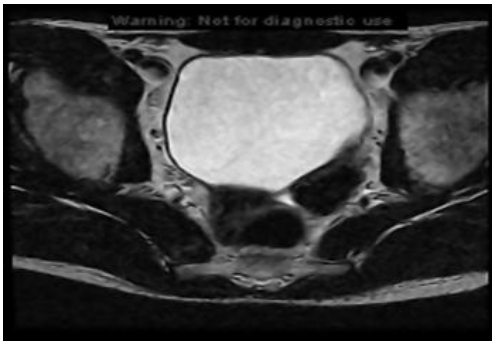


Fig 6-b. Image after wiener filter

C. Removal of Border and Background

The proposed system uses a matlab imclearborder built-in function for border and background removal, see result in Fig.7.

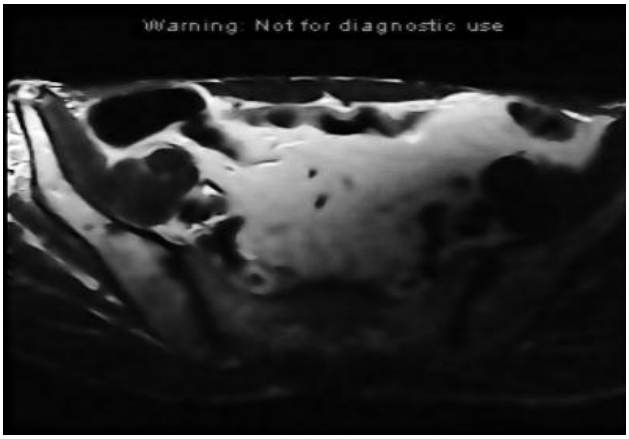


Fig 7-a. Image after removing border and background

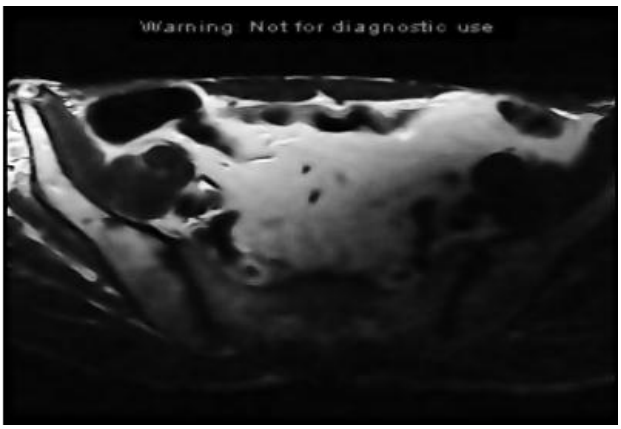


Fig 7-b. Image after removing border and background

D. Bladder Cancer Recognition

1. Features Extraction

Feature extraction is the transformation of the original data (using all variables) to a data Set with a reduced number of variables. In the problem of feature selection, the aim is to select those variables that contain the most discriminatory information. Alternatively, we may wish to limit the number of measurements we make, perhaps on grounds of cost, or we may want to remove redundant or irrelevant information to obtain a less complex classifier.

In feature extraction, all variables are used and the data are transformed (using a linear or nonlinear transformation) to a reduced dimension space. Thus, the aim is to replace the original variables by a smaller set of underlying variables. There are several reasons for performing feature extraction: (i) to reduce the bandwidth of the input data (with the resulting improvements in speed and reductions in data requirements) ; (ii) to provide a relevant set of features for a classifier, resulting in improved performance, particularly from simple classifiers; (iii) to reduce redundancy; (v) to recover new meaningful underlying variables or features that the data may easily be viewed and relationships and structure in the data identified (Ince *et al* , 2009; Julie, I. *et al* ,2012; Martin, E., 2011; A.N. Akansu, *et al*, 2010) .

Wavelets have been demonstrated to give quality representations of images. The discrete wavelet transform (DWT) presents a multi-resolution analysis in the form of coefficient matrices which can be used in a manner similar to Fourier series coefficients. This DWT representation can be thought of as a form of feature extraction on the original image, For the proposed approach Haar-like features are used, where sums of pixel intensities are computed over rectangular sub-windows (Ince *et al* , 2009; Julie, I. *et al* ,2012; Martin, E., 2011; A.N. Akansu, *et al*, 2010) . See Figures [8, 9].

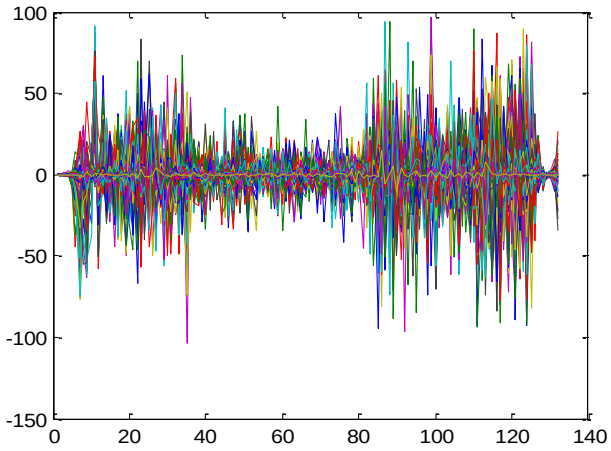


Fig 8. Horizontal features for all images

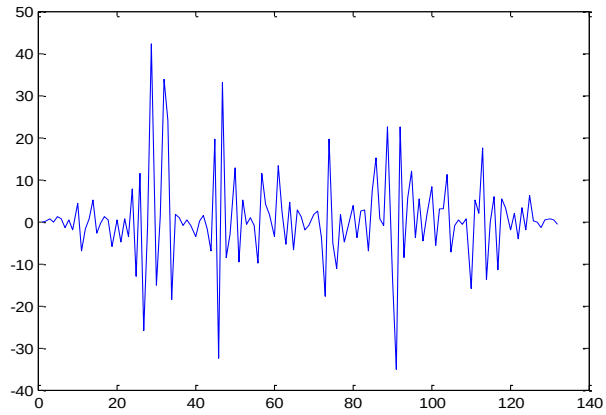


Fig 9. Horizontal features for one image

2. Principal Component Analysis

Geometrically, principal components analysis can be thought of as a rotation of the axes of the original coordinate system to a new set of orthogonal axes that are ordered in terms of the amount of variation of the original data they account for.

This technique has three effects: (i) it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other), (ii) It orders the resulting orthogonal components (principal components) so that those with the largest variation come first, (iii) and it eliminates those components that contribute the least to the variation in the data set (Abdi. H. *et al*, 2010*)*, see Figures [10,11].

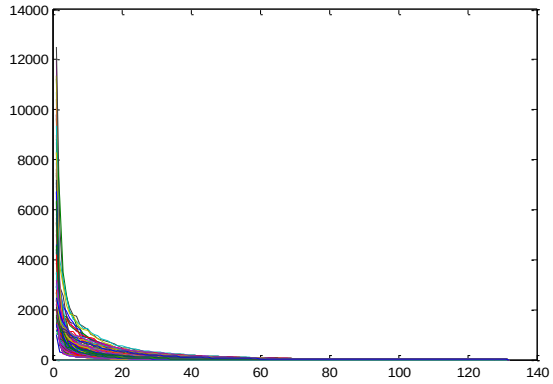


Fig10. PCA for all images

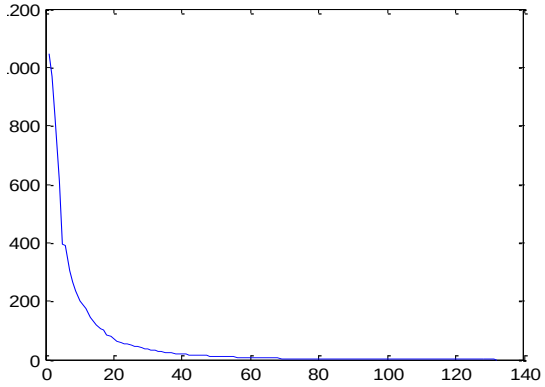


Fig 11. PCA for one image

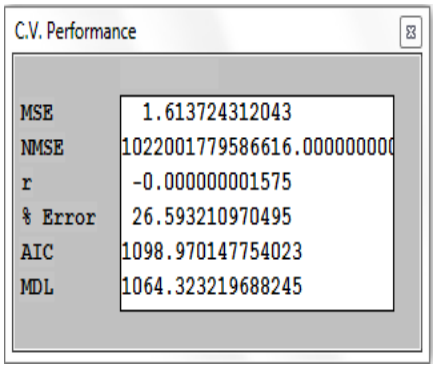
5. Artificial Neural Network (ANN) Algorithms

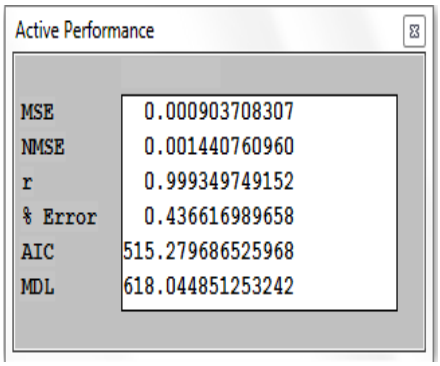
1) Multilayer Perceptron (MLP)

A multilayer perceptron (MLP) is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a learning technique called back propagation for training the network. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable (*http://www.kurzweilai.net/*, 2012; R. Collobert *et al*, 2004).

The active confusion matrix in Fig. 12(e) tallies the results of all exemplars of the last epoch and computes the classification percentages for every output vs. desired combination. For example, in Fig. 12(e), 100% of the T2 exemplars were correctly classified while 0% of the exemplars were classified incorrectly as T3. Similarly, 100% of the T3 exemplars were correctly classified while 0% of the T3 exemplars were classified as T2.

Fig. 12(c) shows the active cost of the MLP when using DWT features; Fig. 12(d) shows the performance of neural network and Fig. 12(e) shows the active confusion matrix of the network. One can see that the cost curve of the MLP approaches to zero, which means that MLP has understood the problem.





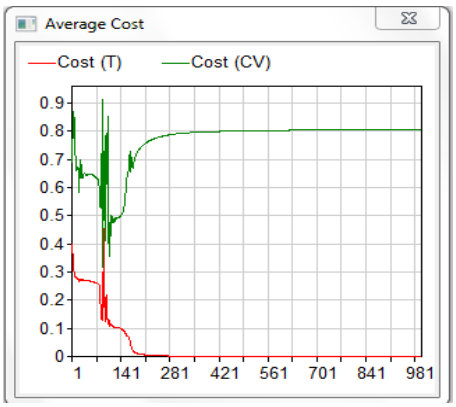
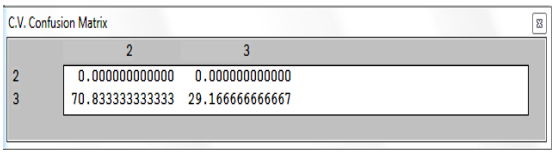


Fig 12. (a) C.V Performance .(b) Active Performance .(c) Average Cost.



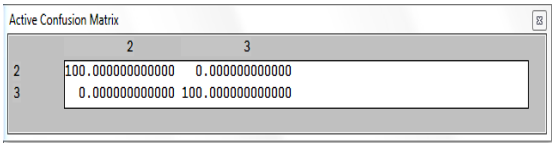
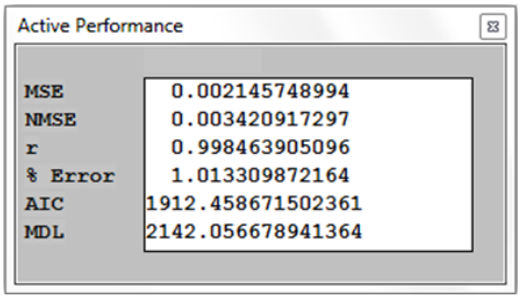
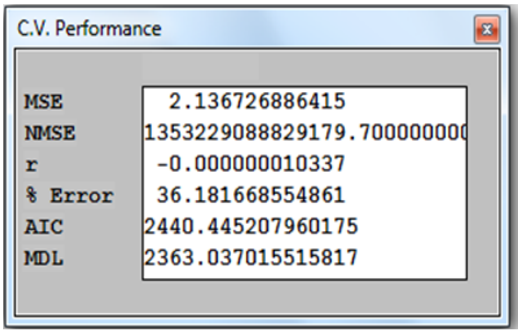


Fig 12. (d) C.V. Confusion Matrix. (e)Active Confusion Matrix

2) Jordan/ Eleman Network

Jordan and Elman networks extend the multilayer perceptron with context units, which are processing elements (PEs) that remember past activity. Context units provide the network with the ability to extract temporal information from the data. In the Elman network, the ctivity of the first hidden PEs is copied to the context units, while the Jordan network copies the output of the network. (D.T Pham *et al*, 1999).





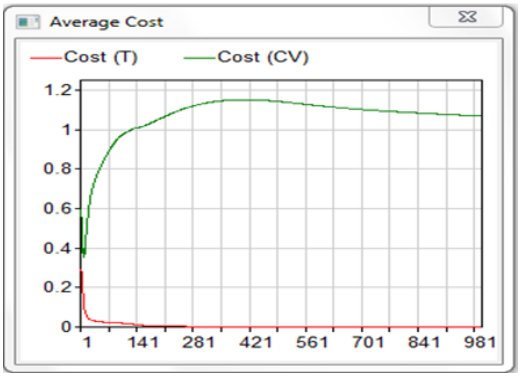
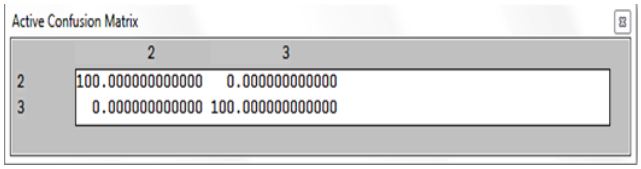


Fig 13. (a) C.V Performance. (b) Active Performance. (c) Average Cost.

Fig. 13(c) shows the active cost of the Jordan/Eleman network; Fig. 13(d) shows the performance of neural network and Fig. 13(e) shows the active confusion matrix of the network.



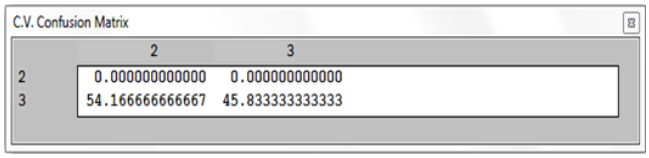


Fig 13. (d) C.V. Confusion Matrix. (e)Active Confusion Matrix

3) Self-Organizing Feature Map Network

A self-organizing map (SOM) or self-organizing feature map (SOFM) is a type of artificial neural network (ANN) that is trained using unsupervised learning to produce a low-dimensional (typically two- dimensional), discretized representation of the input space of the training samples, called a map. Self-organizing maps are different from other artificial neural networks in the sense that they use a neighborhood function to preserve the topological properties of the input space. This makes SOMs useful for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling. The model was first described as an artificial neural network by the Finnish professor Teuvo Kohonen, and is sometimes called a Kohonen map or network (Kohonen *et al* 2007).Fig. 14(c) shows the active cost of the SOM network ; Fig. 14(d) shows the performance of neural network and the other Fig. 14(e) shows the active confusion matrix of the network.

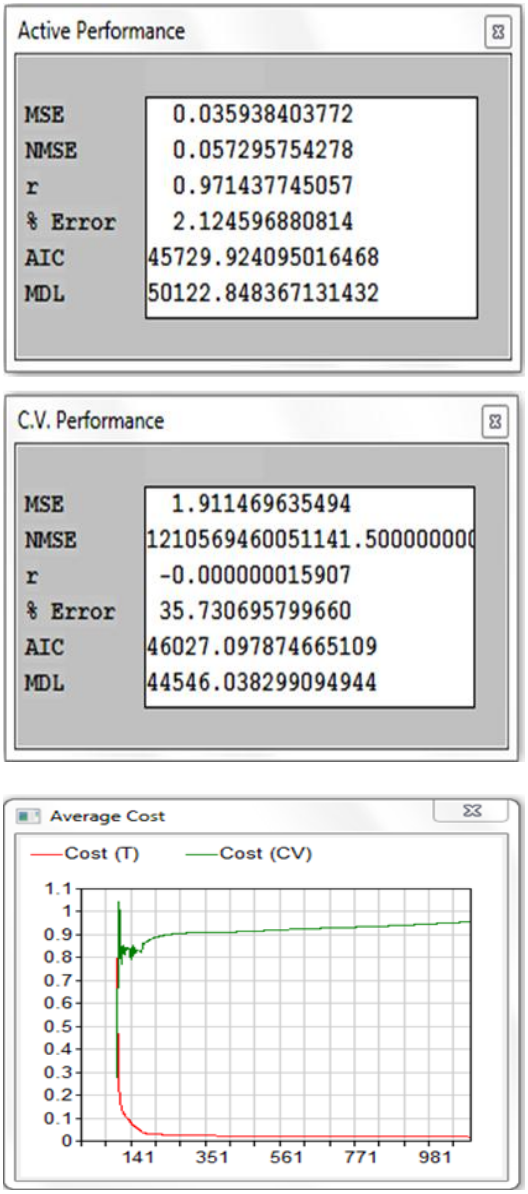


Fig 14. (a) C.V Performance. (b) Active Performance. (c) average cost

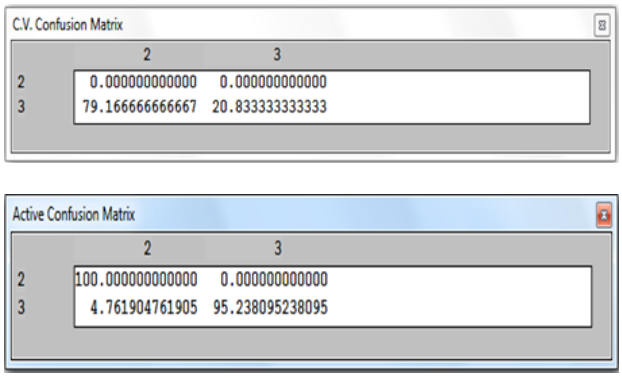


Fig 14. (d) c.v. confusion matrix. (e) active confusion

4) Support Vector Machine (SVM)

The Support Vector Machine (SVM) is implemented using the kernel Adatron algorithm. The kernel Adatron maps inputs to a high-dimensional feature space, and then optimally separates data into their respective classes by isolating those inputs which fall close to the data boundaries. Therefore, the kernel Adatron is especially effective in separating sets of data which share complex boundaries. SVMs can only be used for classification, not for function approximation (Sudhir D.*et al*, 2006).

Support vector machines are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output, making it a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, a SVM training algorithm builds a model that assigns new examples into one category or the other. A SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on (William H. *et al*, 2007).

Fig. 15(c) shows the active cost of the SVM network; Fig. 15(d) shows the performance of neural network and the other Fig. 15(e) shows the active confusion matrix of the network.

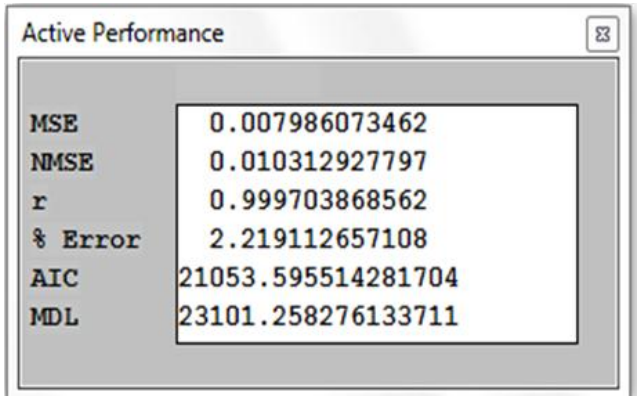
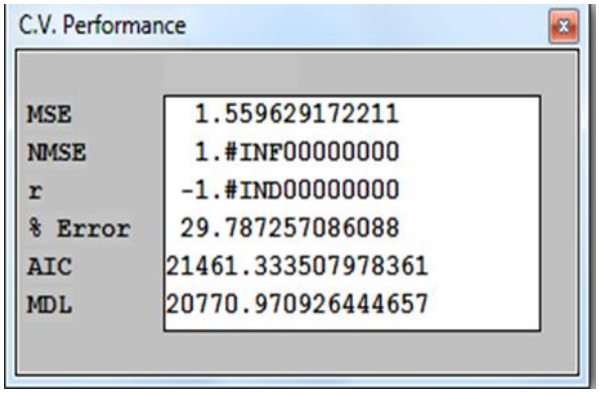
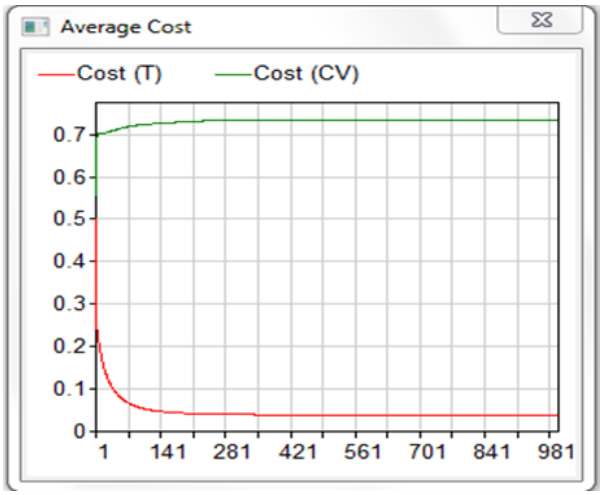
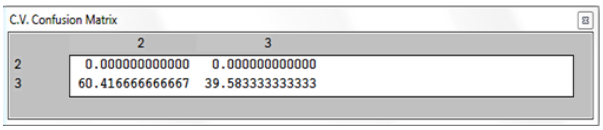


Fig 15. (a) C.V Performance







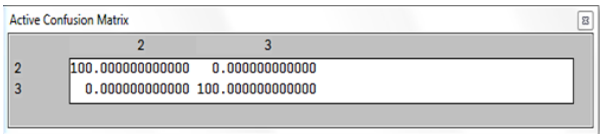


Fig 15. (b) Active Performance. (c) Average Cost (d) C.V. Confusion Matrix. (e) Active Confusion

Conclusion

The purpose of this study is to develop a method of classifying bladder cancers to specific diagnostic Categories based on cancer stage (FREE, T1, T2, T3, T4) using artificial neural networks (ANNs). We trained the ANNs using the 40 patient MRI images, each having about 40 images, and then we divide these images into four categories based on position of view (layer). The ANNs correctly classified all samples and identified the images most relevant to the classification. Sequential weight/bias training algorithm was chosen as learning function because of its speed and accuracy. DWT was chosen as feature extraction technique with only horizontal features which gives us best results. About 130 PCA features were introduced as input to ANN. We achieved accuracy rate about 99% for our database using MLP neural network which gives us better result than any other used classifier. To test the ability of the trained ANN models to recognize MRI bladder cancer images, we analyzed additional blind samples that were not previously used for the training procedure, and correctly classified them in most cases. This study demonstrates the potential applications of these methods for tumor diagnosis and the identification of candidate targets for therapy. This comparative study describes the efficiency and accuracy of MLP.

References

O'Brien T, Thomas K. (November 2010).Bladder cancer: Photodynamic diagnosis can improve surgical outcome. Nature Reviews Urology 7 (11).

Globocan (2010).Bladder — Estimated incidence, all ages: both sexes. *International Agency for Research on Cancer* (IARC), World Health Organization.

Filip Velickovski, (June 2010). Prostate Cancer Detection Via Classification of Features From Different MRI Modalities, *MSc Erasmus Mundus in Vision and Robotics* (VIBOT).

Thomas Kailath, Ali H. Sayed, and Babak Hassibi, (2000). Linear Estimation, *Prentice-Hall, NJ*, ISBN 978-0-13-022464-4.

Voor mijn moeder. (2005). Molecular Diagnosis and Prognosis of Bladder Cancer Towards the implementation of molecular markers in clinical practice, *B.W.G. van Rhi7jn*.

Andrew R. Webb, Malvern,(2002) . Statistical Pattern Recognition, *John Wiley & Sons Ltd*.

Ince, Kiranyaz, Gabbouj , (2009). A generic and robust system for automated patient-specific classification of ECG signals, *IEEE Trans Biomed Eng.* , 56(5):1415-26.

Julie, I. ; Kirubakaran, E. (2012). MFE-HC: The maximizing feature elimination technique based hybrid classifier for cancer molecular pattern discovery, *IEEE International Conference on Pattern Recognition*, Informatics and Medical Engineering (PRIME),pp(s): 376 – 380.

Martin, E., (2011). Novel method for stride length estimation with body area network accelerometers, *IEEE BioWireless*, pp. 79-82.

A.N. Akansu, W.A. Serdijn, and I.W. Selesnick,(March 2010) . Wavelet Transforms in Signal Processing: A Review of Emerging Applications, *Physical Communication, Elsevier*, vol. 3, issue 1, pp. 1-18.

Abdi. H., & Williams, L.J. (2010). Principal component analysis., *Wiley Interdisciplinary Reviews*, Computational Statistics, pp. 433–459,.

http://www.kurzweilai.net/how-bio-inspired-deep-learning-keeps-winning-competitions, (2012).

R. Collobert and S. Bengio, (2004) . Links between Perceptrons, MLPs and SVMs . *Proc. Int'l Conf. on Machine Learning* (ICML).

D.T Pham, D Karaboga, (April 1999). Training Elman and Jordan networks for system identification using genetic algorithms, *science direct*, Volume 13, Issue 2, pp. 107–117.

Kohonen, Teuvo; Honkela, Timo, (2007). Self- Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics* 43 (1): pp. 59–69.

Sudhir D. Sawarkar, Ashok A. Ghatol, Amol P. Pande, (June 12-14, 2006). Neural Network Aided Breast Cancer Detection and Diagnosis Using Support Vector Machine, *Proceedings of the 7th WSEAS International Conference on Neural Networks*, Cavtat, Croatia, pp.158-163.

William H.; Teukolsky, Saul A.; Vetterling, William T.; Flannery, B. P. (2007) .Section 16.5. Support Vector Machines. Numerical Recipes: The Art of Scientific Computing (3rd ed.). , *Cambridge University Press*. ISBN 978-0-521-88068-8.