Cardiac Problem Diagnosis with Statistical Neural Networks and Performance Evaluation by ROC Analysis

Gökhan Bilgin$^1$, Öğuz Altun$^2$

$^1$Yildiz Technical University, Istanbul, Electrical and Electronics Faculty - Computer Engineering Department
gokhanb@ce.yildiz.edu.tr
oguz@ce.yildiz.edu.tr

Abstract — Electronic medical imaging technologies are growing rapidly and simplifying diagnosis in medical area. The proper use of this technology requires a better understanding, interpretation, and development of new, efficient algorithms. Processing and recognition techniques of patterns related to these medical devices are becoming more important. Among these techniques artificial neural network structures are very promising in the diagnosis decision support mechanisms. In this paper, it is aimed to present the performance of statistical neural network structures on classifying cardiac problems which are obtained from SPECT (Single Photon Emission Computed Tomography) images. Principal component analysis has been used to overcome excessive dimensionality of data. After classification we used Receiver Operation Characteristics (ROC) analysis to evaluate system performance. Results show that proper neural network based statistical pattern recognition models will play a fundamental role in medical signal processing and image analysis.

INTRODUCTION

Modern electronic medical imaging systems generate huge amounts of image data that can be analyzed and processed only with the use of specialized computer software. The computerized analysis of this data is easier, less time consuming and relatively cheaper than human analysis.

In this paper it is intended to create a neural network based diagnosis system that is able to semi-automate SPECT myocardial perfusion diagnosis process. For this reason, performance of statistical neural network structures has been examined. It’s not enough to evaluate real performances of networks so we used ROC analysis to overcome this problem.

Kurgan et al. [1] describes a computerized process of myocardial perfusion diagnosis from cardiac SPECT images using a data mining and knowledge discovery approach. A database consisting of 267 cleaned patient SPECT images accompanied by clinical information and physician’s interpretations was created first. The system is able to provide a set of computer diagnoses for cardiac SPECT studies and can be used as a diagnostic tool by a cardiologist.

CARDIAC SPECT DIAGNOSIS

First we need to understand the medical problem domain. The patient is injected with radioactive tracer. Then two studies are performed, one 10-15 min. after injection during maxima stress - called stress study (stress image) -, and one 2-5 hour after injection - called rest study (rest image) -. The studies are collected as two sets of three-dimensional images. Cardiologists compare stress time and rest time studies in order to detect abnormalities in the perfusion [1].

The main data mining goal is to identify key features from cardiac SPECT images. This is done using image analysis and processing algorithms. For partial diagnoses there are seven classes (normal, reversible, partially reversible, artifact, fixed, equivocal, reverse distribution) which are used in a heuristics approach. For the overall diagnosis there are eight classes (normal, ischemia, infarct, artifact, ischemia & artifact, equivocal, LV dysfunction). This overall diagnosis also uses partial diagnoses as input. Various combinations of these classes define the overall diagnosis result (normal or abnormal) [1]. Our system will be designed as a computer based decision support tool for the diagnosis of myocardial perfusion to help cardiologists.

UNDERSTANDING THE CARDIAC SPECT DATA

Cardiac SPECT images represent myocardial muscle perfusion that is proportional to distribution of radioactive counts within the myocardium. Typical 2D image resolution is 64x64, all the images are black and white, 8 bits per pixel with 256 shades of grey. Brighter areas on the figure1 correspond to well perfused areas of myocardium.[1]

![Figure1](attachment:image1.png)

Figure 1. a.) Normal Perfusion b.) Abnormal Perfusion
The SPECT database consists of images and clinical patient records. The processed new database was created based on the initial SPECT database (after cleaning). This database is available as part of UCI-ML Databases [2]. The database of 267 SPECT images was processed to extract features that summarize the original SPECT images. There are 22 different sensors which get information from body in two phases: stress and rest times. As a result, 44 continuous feature patterns were created for each patient. Further information can be found in simulation results section.

RADIAL BASIS FUNCTION NETWORKS

Basic idea of the radial basis function networks (RBFN) is to make a weighted sum of radial basis functions to approximate the desired function. RBFN is a three layers structure. The input layer is related to input vector space and the output layer is related to the pattern classes. The activation functions of the hidden layer neurons are determined by the function centers \( c_j \) and bandwidths \( \sigma_j \). The activation function is a Gaussian given by

\[
\phi_j = \exp\left(-\frac{\|x - c_j\|^2}{2\sigma_j^2}\right)
\]  

(1)

The \( j \)th output neuron has the output given by

\[
s_j(x) = \sum_{i=1}^{K} w_{ij} a_i(x) + b_j
\]  

(2)

where \( w_{ij} \) is the weight between hidden layer neuron \( i \) and output layer neuron \( j \) [4].

GENERAL REGRESSION NEURAL NETWORKS

A general regression neural network (GRNN) is a special case of a RBFN. The difference is that in GRNN the centers and bandwidths of activation functions are deterministically obtained from the training data. Assume that we are given training observations \( (x_p, y_p), p = 1, \ldots, P \). Each input vector \( x_p \) is assigned as the center of one of the Gauss kernels. The output of the \( p \)th RBF unit for an input \( x \) is given as

\[
\beta_p = \exp\left(-\frac{(x - x_p)^T (x - x_p)}{2\sigma^2}\right),
\]  

(3)

where \( \sigma \) is the user-specified smoothing parameter. The output of the network is obtained by (4):

\[
y = \sum_{p=1}^{P} \alpha_p y_p, \quad 0 \leq \alpha_p \leq 1, \quad \sum_{p=1}^{P} \alpha_p = 1
\]  

(4)

where \( \alpha_p \) is given by (5):

\[
\alpha_p = \frac{\beta_p}{\sum_{p=1}^{P} \beta_p}
\]  

(5)

If the input vector \( x \) is close to some training vector \( x_p \), the corresponding coefficient \( \alpha_p \) will also be large, hence the output \( y \) will be close to training output \( y_p \) [5].

PROBABILISTIC NEURAL NETWORKS

Probabilistic neural networks are also known as Bayes-Parzen estimators. Let \( x \) be an \( m \)-dimensional vector that belongs either to class \( K_1 \) or class \( K_2 \). Let the corresponding probability density functions of \( K_1 \) and \( K_2 \) be \( F_1(x) \) and \( F_2(x) \). The Bayes decision rule tells us that the \( x \) belongs to \( K_1 \) if equation (6) is true, and it belongs to \( K_2 \) otherwise.

\[
\frac{F_1(x)}{F_2(x)} > \frac{L_1P_2}{L_2P_1}
\]  

(6)

where \( P_1 \) and \( P_2 \) are the prior probabilities of the occurrence of classes \( K_1 \) and \( K_2 \), respectively. \( L_1 \) is the cost function associated with the misclassification of the vector \( x \) as belonging to class \( K_1 \) while it belongs to class \( K_2 \), and \( L_2 \) is the cost function associated with misclassification of the vector \( x \) as belonging to class \( K_2 \) while it belongs to class \( K_1 \). From equation (6) it can be seen that \( F_1(x), F_2(x), L_1, \) and \( L_2 \) are needed to find the most probable class of the input vector \( x \). The density functions of classes are obtained by Parzen windows.

\[
F(x) = \frac{1}{(2\pi)^{m/2} \sigma^m n} \sum_{i=1}^{n} \exp\left(-\frac{(x - x_i)^T (x - x_i)}{2\sigma^2}\right)
\]  

(7)

where \( n \) is the number of training vectors, \( m \) is the dimension of the input space, \( i \) is the pattern number, and \( \sigma \) is the adjustable smoothing parameter [6].

PRINCIPAL COMPONENT ANALYSIS

Component analysis is a way to find the "right" features from the data in a least-squares sense. Essentially, a set of correlated variables is transformed into a set of uncorrelated variables, which are ordered by decreasing variability. Given a data set of \( n \) \( d \)-dimensional observations \( x_1, \ldots, x_n \), the mean vector \( \mu \) and \( d \times d \) covariance matrix \( \Sigma \) is computed. The mean vector is given by

\[
\mu = \frac{1}{n} \sum_{k=1}^{n} x_k
\]  

(8)

and the covariance matrix is given by

\[
\Sigma = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \mu)(x_k - \mu)^T
\]  

(9)
Next the eigenvectors and eigenvalues are computed. Eigenvectors represent basis of the new vector space, and their corresponding eigenvalues give the "importance" of each eigenvector. We sort the eigenvalues in decreasing order, and call the largest eigenvalue as $\lambda_1$ and corresponding eigenvector as $e_1$, the second eigenvalue $\lambda_2$ with eigenvector $e_2$, and so on. We chose the top $k$ eigenvectors and assume that the rest of the dimensions generally contain noise. We form a $dxk$ projection matrix $A$, whose columns consist of the selected eigenvectors. The original data is projected onto new principal components space by

$$x' = A'(x - \mu),$$

where $x'$ is the projection of observation $x$ and $A'$ is the transpose of $A$.

**RECEIVER OPERATION CHARACTERISTIC ANALYSIS**

ROC analysis is part of the “Signal Detection Theory” developed for using in fields like military, where making a wrong decision carries big risks, and in medical decision making, where a wrong decision can result in death. In such fields it is acceptable to not decide at all for some instances, but it is very important to make right decisions.

The notions of sensitivity and specificity are made use of in order to detect the accuracy of the diagnosis. Let’s define TP (true positive) as the case of diagnosing an ill patient as ill, TN (true negative) as the case of diagnosing an healthy patient as healthy, FP (false negative) as the case of diagnosing a healthy patient as ill, and FN (false positive) as the case of diagnosing an ill patient as healthy. Then, the sensitivity and specificity can be given as below [8].

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

$$\text{Specificity} = \frac{TN}{TN + FP}$$

Sensitivity can be interpreted as the proportion of patients with disease whose tests are positive, whereas specificity can be interpreted as the proportion of patients without disease whose tests are negative.

The increase in one of the sensitivity or specificity usually results in a decrease in the other.

A ROC curve is a plot of sensitivity vs. specificity. The ROC analysis estimates a curve, which describes the inherent tradeoff between sensitivity and specificity of a diagnostic test.

**SIMULATION RESULTS**

The database describes the diagnosis of SPECT images. Each of the patients is classified into either of two categories: normal and abnormal. The database of 267 SPECT images (of patients) was processed to extract features that summarize the original SPECT images. There are 22 different sensors which get information from body in two phases: stress and rest times. As a result, 44 continuous feature patterns were created for each patient. There are no missing feature values for samples.

In SPECTF database all of the features are continuous integer values. Predicted class output for overall diagnosis is binary valued (0 for normal and 1 for abnormal class). In SPECTF training dataset, there are equally distributed 80 instances in two classes with 44 features. In SPECTF heart test dataset there are 267 samples of the database including training samples which were used for training. Class distribution of test dataset is 55 samples for class 0 (normal perfusion) and 212 samples for class 1 (abnormal perfusion).

Principal component analysis has been used to decrease the number of the features, thus amount of data. The main goal of using PCA is to reduce the dimensionality of a dataset while retaining as much information as is possible. Several reduced datasets have been tested (9, 7, 5, 4, 3 features), but at the end it has been decided to use a dataset with four features in this paper because of the highest accuracy in results.

In previous works, with the same dataset, the CLIP3 machine learning algorithm achieved 77.0% accuracy [3]. Multilayer Layer Perceptron (best structure result) achieved %77 accuracy, Support Vector Machines with sequential minimal optimization algorithm achieved %73 accuracy and k-nearest neighbor algorithm (with the best k result) achieved %78 accuracy.

The simulations were realized by MATLAB 7.0 Neural Network Toolbox. First, non-reduced database has been simulated, then reduced database with four features has been simulated. It has been observed that the reduced database results are better than original database. The classification results of the database with four attributes in each training and test set for RBF, GRNN, and PNN are given in the Tables 1-3.

![Figure2](image)

**Table 1** Training and test results and percentages for RBF
The sensitivity and specificity values are computed as in equation (11,12) at optimum spread value. Sensitivity shows test strength in finding true positive instances whereas specificity shows test strength in avoiding false positives. According to these definitions RBF is better than other network structures, as can be seen from Table 4.

**Table 4 Training and test sets sensitivity and specificity values for all net structure at optimum spread value.**

**ROC CURVES**

![ROC CURVES](image)

CONCLUSION AND DISCUSSION

From the medical point of view our goal is to semi-automate cardiac SPECT diagnosis process in order to assist a cardiologist in diagnosing cardiac SPECT images, to make this procedure easier, more consistent, and efficient.

The achieved results are encouraging because of the high correctness of diagnoses. According to simulation results, successful classification ratios have been achieved for all types of statistical neural networks (especially 95% in RBF). In previous studies which are mentioned in the simulation results section, this ratio has not been reached. RBF is the most successful one, but the other statistical neural networks (GRNN and PNN) still can be used for pattern recognition problems. PNN and GRNN are often used for function approximation. It has been shown that ([5],[6]), when given a sufficient number of hidden neurons, both of these structures can approximate a continuous function (81% accuracy for both PNN and GRNN structures is a better result compared to previous works mentioned in simulation result section.

The biggest disadvantages of statistical neural networks are memory and time consumption when dimensions of input data and volume of training sets increase.

On the other hand, Principal Component Analysis was very successful in this study: reduced (from 44 to 4 features) dataset gave better results than the original dataset in terms of both efficiency and classification success.

According to Figure 3, RBF is the method with the largest area under the curve and PNN follows it. This observation is compatible with sensitivity and specificity values in Table 4, which is complementary to Figure 3.

Results show that the proper neural-based statistical pattern recognition models will play a fundamental role in medical signal processing and image analysis.

REFERENCES