# A Neural Network Learned Information Measures for Heart Motion Abnormality Detection

M.S Nambakhsh[1,2], Kumaradevan Punithakumar[3,5], Ismail Ben Ayed[5], Aashish Goela[4], Ali Islam[4,5], Terry Peters[1,2,4], and Shuo Li[2,3]

[1] Imaging Research Laboratories, Robarts Research Institute, London, ON, Canada,
[2] Biomedical Engineering Program, University of Western Ontario, London, ON, Canada,
[3] GE Healthcare, London, ON, Canada,
[4] Department of Medical Imaging, University of Western Ontario,
[5] St. Joseph Health Care, London, ON, Canada.

**Abstract.** In this study, we propose an information theoretic neural network for normal/abnormal left ventricular motion classification which outperforms significantly other recent methods in the literature. The proposed framework consists of a supervised 3-layer artificial neural network (ANN) which uses hyperbolic tangent sigmoid and linear transfer functions for hidden and output layers, respectively. The ANN is fed by information theoretic measures of left ventricular wall motion such as Shannon's differential entropy (SDE), Rényi entropy and Fisher information, which measure global information of subjects distribution. Using $395 \times 20$ segmented LV cavities of short-axis magnetic resonance images (MRI) acquired from 48 subjects, the experimental results show that the proposed method outperforms Support Vector Machine (SVM) and thresholding based information theoretic classifiers. It yields a specificity equal to 90%, a sensitivity of 91%, and a remarkable Area Under Curve (AUC) for Receiver Operating Characteristic (ROC), equal to 93.2%.

**Keywords:** Artificial neural network, computer aided diagnosis, information theoretic measures, magnetic resonance imaging.

## 1 Introduction

Detection of left ventricular (LV) disfunctions such as wall motion abnormality is essential to effectively detect coronary heart diseases in their early stage. Conventional visual diagnosis is labour-intensive, tidious and restrictive. Alternatively, in order to analyze such a massive data, computer-aided diagnostic systems have attracted a large number of research studies currently [1–5]. LVleft ventricular motion can lead to promising classification results. However, it is a simple thresholding of each individual measure [4] so that it is not able to diagnose LV diseases optimally. This disability stems from the significant overlap between the distribution of information measures computed for normal and abnormal subjects (cf. plots in Fig. 3).

In this paper, to classify the aforementioned distributions a nonlinear relation between the inputs and clusters of the output is solved. This nonlinear mapping distinctively separates overlapped classes of the inputs. The classifier can be expressed as a

multilayer combination of certain numbers of weighted hyperbolic tangent functions ($tanh$) of the inputs [6].

In order to find the best aforementioned nonlinear relation, we investigate an artificial neural network based on a combination of information theoretic measures. Adaptability and nonlinearity of ANNs has made them well-known, easy to implement, useful tools in pattern recognition [7, 8], as well as in medical image processing [9–11].

Using 395 image sequences, each consisting of 20 segmented LV cavities of short-axis magnetic resonance (MR) functional images, obtained from 48 subjects, the experimental analysis demonstrates that the proposed ANN can lead to a significant improvement over other classifiers that are prevalent in the literature [2]. Moreover, a Bhattacharyya distance analysis [12] (cf. plots in Fig. 3), which is measure of separability of classes, show that the ANN yields better classification ability amidst the stochastic nature of the cardiac motion and image segmentation inaccuracies.

## 2   Preprocessing of data

### 2.1   Tracking the Cavity Boundary and Dynamic Model for Temporal Periodicity

Following the work in [13, 14], an automatic detection of the heart cavity boundary (the endocardium) for each MR cardiac frame $k \in [2, \ldots, K]$ was formulated as the evolution of a closed planar parametric curve towards the endocardium.[6]

Let (x, y) be a Cartesian point on the endocardial boundary and state vector $\xi = [\bar{x}_k^i \; \mathsf{x}_k^i \; \dot{\mathsf{x}}_k^i]^T$ describe the dynamics of the point in x-coordinate direction, where $\dot{\mathsf{x}}_k^i$ and $\bar{x}_k^i$ denote, respectively, velocity and the mean position over a cardiac cycle. Now, let $\{\mathbf{s}_k^i = [\bar{\mathsf{x}}_k^i \; \mathsf{x}_k^i \; \dot{\mathsf{x}}_k^i \; \bar{y}_k^i \; y_k^i \; \dot{y}_k^i]^T : i = 1, \ldots, N\}$ be a sample point on the LV boundary in frame $k$.

Cyclic motion of the points are formulated by a continuous state-space model, from which a Kalman filter is used to estimate the states. In order to analyze wall motion regionally, a symmetric nearest neighbor correspondences [23] is applied by sampling a set of equally-spaced points along the LV boundary [4]. Consequently, the sequence of corresponding points over time is found. Using Spline interpolation, $N_s$ points were sampled along the LV cavity in each frame, and $N$ points were chosen as measurements to the filter. Finally, a Gaussian kernel estimation was applied to the points to obtain the probability density. The information theoretic measures are listed in the [4] (Eq. 3-8).

### 2.2   Artificial Feed-forward Backpropagation Neural Network

To classify the aforementioned distributions, a nonlinear relation between the inputs and clusters of the output must be solved, to distinctively separate overlapped classes of inputs. The nonlinear mapping can be expressed as a multilayer combination of certain numbers of weighted hyperbolic tangent ($tanh$) of the inputs, and is expressed as

$$\mathbf{Y} = F\left[\sum W^{h_{(n-1)}} \ldots \tanh\left(\sum \mathbf{W}^{h_1} . \tanh(\mathbf{W}^{h_0}.\mathbf{X})\right)\right] \tag{1}$$

---

[6] $K$ typically is equal to 20 or 25.

where $\mathbf{X}$ is a matrix of inputs, $F$ a thresholding or step function, $\mathbf{W}^{h_i}$ matrix of weights, $i = 0 \ldots n-1$, in consecutive layers $i$ and $i+1$, and $\mathbf{Y}$ the output classes. We need to find the best $\mathbf{W}$ corresponding to the least error in estimating $\mathbf{Y}$.

**Training:** In this study, a neural network was applied to implement $\mathbf{Y}$ in (1). Before the ANN is prepared for use, it should be trained. The input vector sequentially is introduced to ANN $\mathbf{X} = (X_1, X_2, .. X_n)^T = (S_f, I_f, R_f^\alpha, \mu_r, \mu_v)^T$, where $\mu_r$ and $\mu_v$ are mean radial distance and mean radial velocity respectively, and its corrosponding output vector $\mathbf{Y} \in \{0, 1\}$ are split up into a training set $\{(\mathbf{X}^j, \mathbf{Y}^j); j \in I_r$ and a testing set $\{(\mathbf{X}^j, \mathbf{Y}^j); j \in I_t$. The network in its simplest form comprises a single neuron as shown in Fig. 1a.
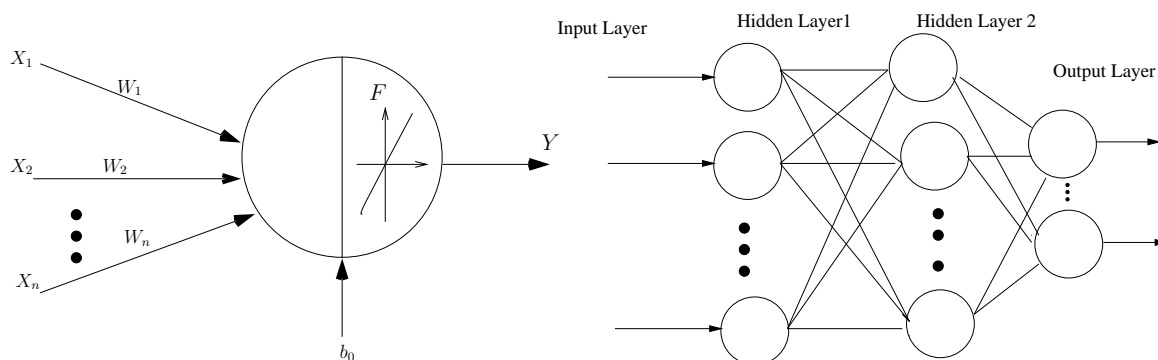


**Fig. 1.** a) A node of ANN model b) General structure of the ANN.

**Testing:** After the training procedure, the weight vector, $\mathbf{W}$, is configured and remains constant. The applied ANN uses a hyperbolic tangent function,

$$F(x) = \tanh(x) \tag{2}$$

as an activation function in hidden layers and step function in the output layer. Thus, having trained the simple network at Fig. 1, the output of this neuron during feeding each testing set $\mathbf{X}^t$ is expressed as

$$\hat{\mathbf{Y}}_k^t = \tanh\left(\mathbf{W}.\mathbf{X}_k^t\right) = \tanh\left(W_1.S_f(k) + W_2.I_f(k) + W_3.R_f^\alpha(k) + b_0\right) \tag{3}$$

where $k$ is the $kth$ sample of the testing set and $b_0$ is bias to the neuron. Note that the right-hand side term in (3) is based on information theoretic measures as inputs to the ANN. Now assuming that the ANN contains a single hidden layer, and that there are $n$ neurons in this layer, the ultimate output is expressed as

$$\mathbf{Y}_{out} = F_{step}(\mathbf{W}^h.\mathbf{Z} + b_{out}) \tag{4}$$

$$\mathbf{Z}(i) = \hat{\mathbf{Y}}^t(i) \tag{5}$$

where $F_{step}$ is a step function, $\hat{\mathbf{Y}}^t(i)$ is the output of $ith$ neuron in the hidden layer (3), $\mathbf{W}^h$ is the weight vector connecting the neurons in the hidden layer to the neuron in the output layer and and $b_{out}$ is bias to the output neuron. Note that the ANN in this study has a single neuron in the output layer.

## 3  Experiments

The data-set contained 395 short-axis image sequences, each consisting of 20 functional 2D-images acquired from 30 normal and 18 abnormal subjects as diagnosed by radiologists based on left ventricular wall motion. The data were acquired on 1.5T MRI scanners with fast imaging with a steady state acquisition (FIESTA) image sequence. The experiments compare the proposed ANN classifier with other classifiers such as Support Vector Machine (SVM), the mean radial displacement, mean systolic radial velocity, SDE, Rényi entropy ($\alpha = 2$) and Fisher information. $\alpha$ was chosen out of a wide variety of experimented $\alpha s$. Also, the same kernel density with bandwidth 0.5 was used for Ab/normal subjects. The results were compared with ground truth classification of the MRI sequences by radiologists. If any of the cardiac segments in an image sequence is abnormal, it is diagnosed as an abnormal.The *leave-one-subject-out* method, the receiver operating characteristic (ROC) curves, and the corresponding area under the curves(AUCs) are used to evaluate the performance of the classifier. Moreover, the ability of classifiers to separate classes was assessed by the Bhattacharyya measure. Table 1 summarizes the results.
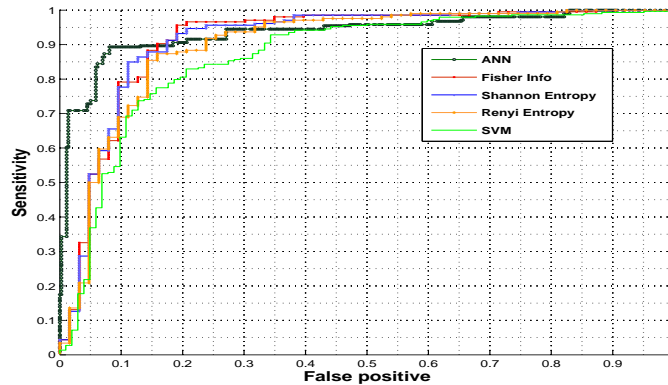


**Fig. 2.** Receiver operating characteristics of the ANN, Information Theoretic, SVM classifier elements.

The ROC curves for classifier elements are shown in Fig.2. The plot shows that the proposed ANN has better classification accuracy than the literature. The corresponding AUCs are listed in Table 1.
The closer the curve to the upper left corner, the better the classification performance.

The ROC curves show that the proposed ANN clearly outperforms other classifier elements.

The AUC represents the average of the classifier sensitivity over specificity resulting from thresholding the outputs. The ANN yielded the highest AUC and, therefore, has the best overall performance. The ANN curve also is closest to the ideal step function, with 91% sensitivity and 90% specificity.

We applied the Bhattacharyya distance metric ($\mathcal{B}$) [12] to evaluate the discriminative power of the classifiers over normal and abnormal hearts.

The higher $\mathcal{B}$, the lesser the overlap (Refer to Fig. 3 for an illustration) and, therefore, the better the discriminative ability of the classifier. The ANN yielded the highest $\mathcal{B}$ as shown in Fig. 3 and reported in Table 1.
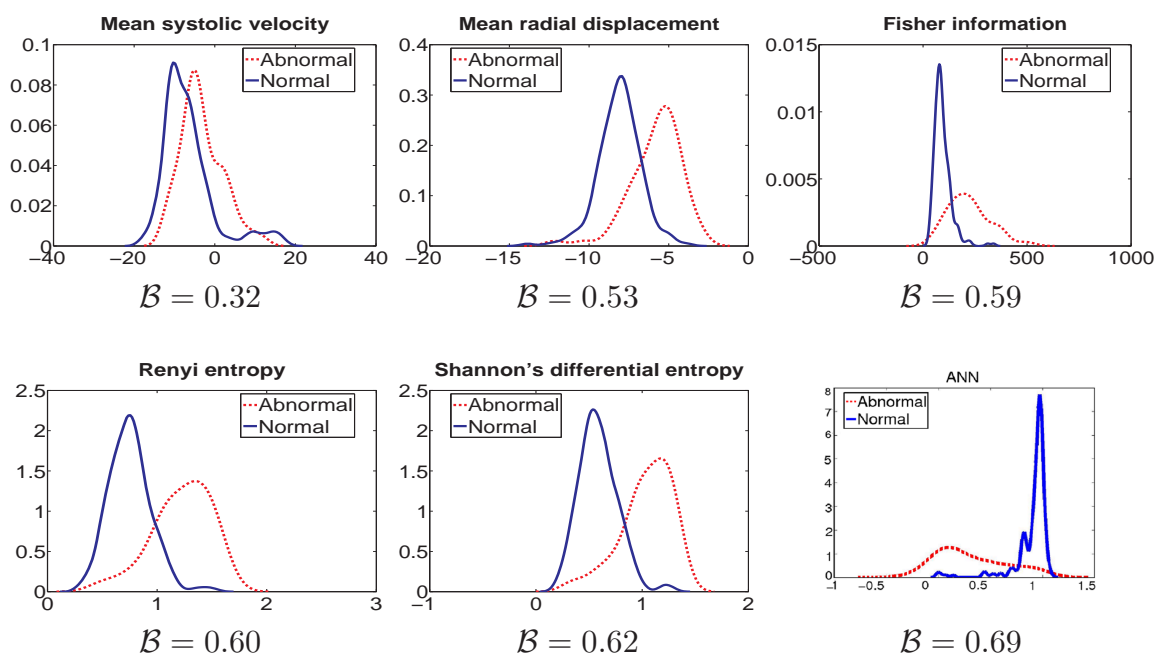


**Fig. 3.** Distribution of normal and abnormal hearts categorized using classifier.

## 4    Conclusions

We derived an ANN to classify the left ventricular based on information theoretic measures. The ANN was fed by global measures of distribution of left ventricular wall motion such as SDE, and Fisher information along with local measures like mean. The structure of ANN was examined by different combinations of these features to determine the best and optimized of them. Moreover, the ANN structure was subject to modifications of number of layers, nodes and activation function to give the best performance. The proposed method is based on analyzing wall motion quantitatively

| Classifier element | AUC (%) | $\mathcal{B}$ | Classification accuracy Abnormal (%) | Normal (%) |
|---|---|---|---|---|
| ANN | 93.2 | 0.69 | 91.0 | 90.0 |
| SVM | 86.50 | 0.61 | 85.0 | 75.0 |
| Mean systolic velocity | 70.8 | 0.32 | 79.4 | 54.9 |
| Fisher information | 89.3 | 0.59 | 84.1 | 85.0 |
| Rényi entropy | 90.8 | 0.60 | 87.3 | 84.5 |
| Shannon's differential entropy | 90.9 | 0.62 | 90.5 | 78.6 |

**Table 1.** The AUC, $\mathcal{B}$ and percentage of classification accuracy corresponding to Fig. 2 and Fig. 3 respectively.

by constructing distributions of the radial distance and a nonlinear classifier in order to separate the overlapped classes. The results, based on ROC, area under the curves, Bhattacharyya distance metrics, and *leave-one-out* cross validation, showed that the proposed ANN can lead to a significant improvement over other classifiers in the literature.

## References

1. Qian, Z., Liu, Q., Metaxas, D.N., Axel, L.: Identifying regional cardiac abnormalities from myocardial strains using spatio-temporal tensor analysis. In Metaxas, D., et. al, eds.: MICCAI 2008. (Volume 5241 of LNCS.) 789–797
2. Qazi, M., Fung, G., Krishnan, S., Bi, J., Bharat Rao, R., Katz, A.: Automated heart abnormality detection using sparse linear classifiers. IEEE Engineering in Medicine and Biology Magazine **26** (2007) 56–63
3. Suinesiaputra, A., Frangi, A., Kaandorp, T., Lamb, H., Bax, J., Reiber, J., Lelieveldt, B.: Automated detection of regional wall motion abnormalities based on a statistical model applied to multislice short-axis cardiac MR images. IEEE Transactions on Medical Imaging **28** (2009) 595–607
4. Punithakumar, K., Li, S., Ben Ayed, I., Ross, I., Islam, A., Chong, J.: Heart motion abnormality detection via an information measure and Bayesian filtering. In Yang, G.Z., et. al, eds.: MICCAI 2009. (Volume 5762 of Lecture Notes in Computer Science.) 373–380
5. Leung, K.Y.E., Bosch, J.G.: Localized shape variations for classifying wall motion in echocardiograms. (In: MICCAI (1) 2007) 52–59
6. Chowdhury, M., Alouani, A., Hossain, F.: Comparison of ordinary kriging and artificial neural network for spatial mapping of arsenic contamination of groundwater. Stochastic Environmental Research and Risk Assessment **24** (January,2010) 1–7
7. Haykin., S.: Neural Network: A Comprehensive Foundation. 2ed, Englewood Cliffs, NJ: Prentice-Hall (1998)
8. Zurada, J.: Introduction to artificial neural systems. West Publishing Co., St. Paul, MN, USA (1992)
9. Veredas, Francisco, M.H., Morente, L.: Binary tissue classification on wound images with neural networks and bayesian classifiers. IEEE Transactions on Medical Imaging **29** (2010) 410–27
10. Huang, S., Chang, R., Moon, W., Lee, Y., Chen, D., Suri, J.: Analysis of tumor vascularity using three-dimensional power doppler ultrasound images. IEEE Transactions on Medical Imaging **27** (2008) 320–330

11. Suzuki, K., Abe, H., MacMahon, H., Doi, K.: Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (Mtann). IEEE Transactions on Medical Imaging **25** (2006) 406–416
12. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (2003) 564–577
13. Ayed, I.B., Li, S., Ross, I.: Embedding overlap priors in variational left ventricle tracking. IEEE Transactions on Medical Imaging **28** (2009) 1902–1913
14. Ben Ayed, I., Lu, Y., Li, S., Ross, I.: Left ventricle tracking using overlap priors. In Metaxas, D., et. al, eds.: MICCAI 2008. (Volume 5241 of LNCS.) 1025–1033