ISSN: 0067-2904
GIF: 0.851

Diagnoses System of Varicose Disease

Mays M. Hoobi*
Department of Computer Science, College of science, University of Baghdad, Baghdad, Iraq

Abstract

The diagnoses system of varicose disease has a good level of performance due to the complexity and uniqueness in patterns of vein of the leg. In addition, the patterns of vein are internal of the body, and its features are hard to duplicate, this reason make this method not easy to fake, and thus make it contains of a good features for varicose disease diagnoses. The proposed system used more than one type of algorithms to produce diagnoses system of varicose disease with high accuracy, in addition, this multi-algorithm technique based on veins as a factor to recognize varicose infection. The obtained results indicate that the design of varicose diagnoses system by applying multi- algorithms (Naïve Bayes and Back-Propagation) produced new system with high accuracy and low (FAR & FRR) as soon as possible.

Keywords: Back-Propagation Neural Network, Naïve Bayes Classifier, Varicose.

نظام تشخيص مرض الدوالي

ميس محمد هوبي*
قسم علوم الحاسبات، كلية العلوم، جامعة بغداد، بغداد، العراق

الخلاصة

نظام تشخيص مرض الدوالي يمتاز بمستوى جيد من الاداء حيث يعود السبب في ذلك الى انه أنماط الوريد في الساق تكون وحيده ومعقده . وبالإضافة إلى ذلك ، أنماط الوريد تكون داخل الجسم، و خصائصها من الصعب تكرارها ، وهذا السبب جعل هذه الطريقة صعبة التزوير, وبالتالي جعلها تحتوي على ميزات جيدة لتشخيص مرض الدوالي . للحصول على دقة عالية من النظام المقترح ، يقترح هذا البحث تقنية متعددة الخوارزميات على أساس الأوردة كعامل لتميز الاصابة بمرض الدوالي . تشير النتائج إلى أن تصميم نظام الكشف عن الدوالي من خلال تطبيق خوارزميات متعددة (Naïve Bayes and Back–Propagation) الى إنتاج نظام جديد بدقة عالية و(FRR & FAR) منخفضة قدر الممكن .

1-Introduction:

Biometrics is automated methods of distinguish a person based on a physiological or behavioral characteristics. Veins are the soft, thin-walled tubes that return blood from the arms and legs to the heart, and it relies on biological information on the interior of the body [1]. The varicose vein is a little purple vein that suddenly seems to appear on legs. The diagnosis of varicose veins is made primarily by physical examination with the help of a hand-held Doppler [2]. The aim of this paper is examines the applicability of a new system to diagnose the varicose infection and the degree associated with this infection. The structure of this paper as follows: In section 2, the notion of applying different types of classification algorithms in proposed system. Subsequent to these, section 3 visualizes the results of proposed system are presented with two experiments. Finally, section 4 clarifies the conclusion of proposed system.

*Email:Mais.shms@yahoo.com

2- Proposed System –Varicose Diagnose System (VDS)

Classification is to obtain the class that is most matches to the classified sample. The proposed system VDS focuses on vein varicose disease infection. The objective is to determine whether the proposed VDS could recognize the infected image and determine the degree of this infection of varicose disease. The classification via (VDS) is based on the idea that each patient posse's unique vien features. Due to the increase the accuracy. This paper produces a multimodal system to obtain diagnose system of varicose disease with high accuracy, this system exploits Naïve Bayes Classifier (NBC) and Back- Propagation Neural Network (BPNN) algorithms. The proposed system was mainly consists of several phases as illustrated in Figure-1.

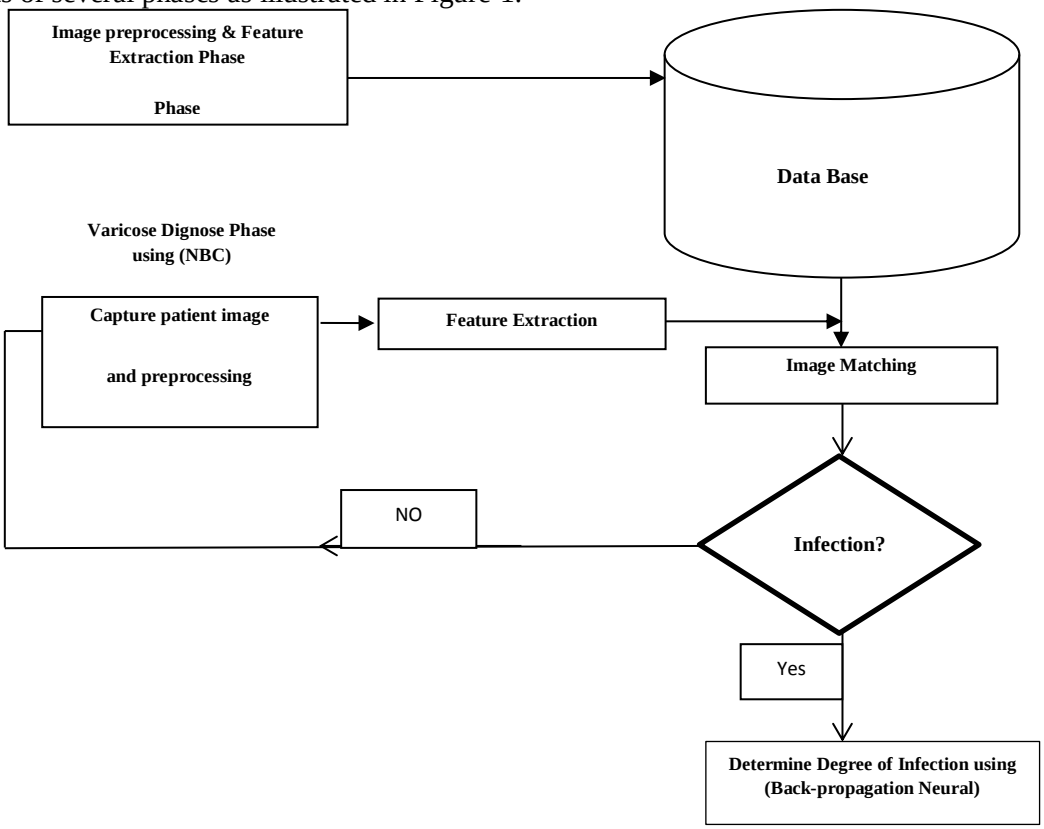

Figure 1- General Structure of VDS

The first phase is image preprocessing and feature extraction phase where the image vector was obtained. Second, a NBC was performed by using image vector to distinguish either there is infection or not based on patient image vector. Finally if there is varicose infection, the VDS uses BPNN algorithm to diagnose the degree of infection among 5 degrees.

2-1Varicose Image Preprocessing and Feature Extraction Phase

At the beginning, the images of legs were obtained , the quality of leg images were bad and required several preprocessing techniques must be used to enhance the image quality. The purpose of this stage is to improve the image quality so that vein patterns can be more easily distinguished. On the other hand, the original image is in RGB format it must be converted into gray scale image, which allows faster processing, as compared to coloured images. A simple type of contrast stretching is applied to enhance the details of image that increase the range of an image to cover whole brightness range. Then, the mean filter is used to reduce the noise and convert the produced gray image to binarized image (black and white). Figure-2 illustrates an example of image preprocessing stages.

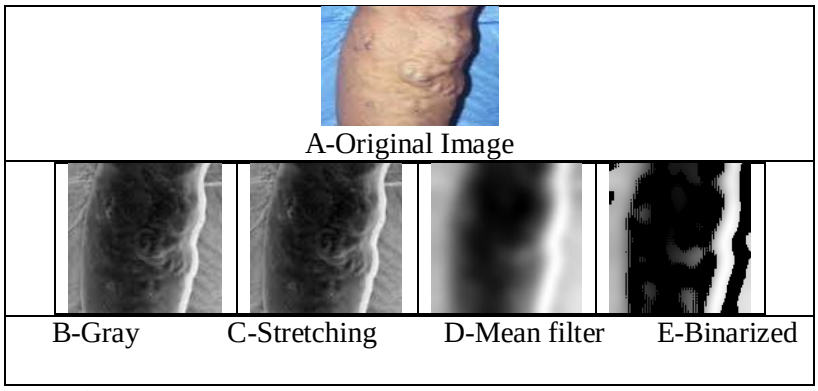

Figure 2 - Preprocessing Stages

The next step is feature extraction. There are several features can be used, however, not all kinds of features are useful and widely used [3]. In this paper the following statistical features are used.
1-Mean [4]

$$Mean = \frac{Sum_i}{Np} \quad (1)$$

Where

$$Sum_i = \sum_{i=1}^{Np} Pi \quad (2)$$

Pi : Is ith picture
2-Variance [4].

$$\sigma^2 = \sum (Pix_i - Mean)^2 / Nsamp \quad (3)$$

Pix: is the value of the ith pixel
N: is the number of samples

2-2 Diagnose Varicose Phase using Naïve Bayes Classifier (NBC)

The Bayesian classification represents a supervised probabilistic learning method as well as a statistical method for classification [5]. A Naïve Bayes Classifier (NBC) is a simple probabilistic statistical classifier based on applying Bayes probability theorem [6].
In general NBC algorithm can be described as show in algorithm (1): [7]

Algorithm (1):

Step1- Establish a training set { x_j , c_j }, j =1, 2….. N for each class , where x_j number of training samples and c_j number of classes .

Step2- Compute a priori information such as probabilities for each template vector and probability density function $p(x \mid c_i)$ as show in equation (4).

$$p(x \mid c_i) = \frac{P(x \mid c_i)P(c_i)}{P(x)} \quad \text{.....(4)}$$

Step3- Given a new un classified measurement Y, use Bayes theorem to obtain the measurement conditioned probability as show in equation (5).

$$P(c_i \mid Y) = \frac{P(Y \mid c_i)P(c_i)}{P(Y)} \quad \text{.....(5)}$$

Step4- Choose ci such that $P(c_i \mid Y) > P(c_j \mid Y)$ for all i ≠j.

NBC consists of two stages, training and classification. Training is the process of learning a model. After the extraction of image vectors, the training process is performed to calculate the conditional probability $P(c_i \mid x)$ which requires mean (μi), variance (σ), as illustrated in equation(6) [7].

$$p(c_i \mid x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\{\frac{-1}{2}(\frac{x-\mu_i}{\sigma})^2\} \text{.....(6)}$$

With the problem of varicose vein infection classification, number of training samples together with the corresponding correct class for each sample are known. NBC classifies the image vector of patient by comparing probability of input vector $p(x)$ with associated class probability $P(c_i)$. In proposed system there are two classes, then the Bayes theorem shows that if $P(c_1 \mid x) > P(c_2 \mid x)$ implies, $\frac{P(x \mid c_1)P(c_1)}{P(x)} \rangle \frac{P(x \mid c_2)P(c_2)}{P(x)}$ .....(7)

So using the fact that $P(c_1) = P(c_2)$ yields, $P(c_1 \mid x) \rangle P(c_2 \mid x) \Rightarrow P(x \mid c_1) \rangle P(x \mid c_2)$.....(8)

And this allows a decision rule, choose c1 if $P(x \mid c_1) > P(x \mid c_2)$ otherwise choose c2.

2-3 Determination Varicose Degree using Back-Propagation Neural Network (BPNN)

BPNN algorithm is used to determine the degree of varicose infection; it is a form of supervised learning for Multi-Layer Perceptron (MLP). The MLP network consists of several "layers" of neurons; typically an input layer, hidden layers, and an output layer. Input layers take the input and distribute it to the layers hidden. These hidden layers do all the necessary computation and output the results to the output layer, which forwards the data to the user as shown in Figure-3 [8].

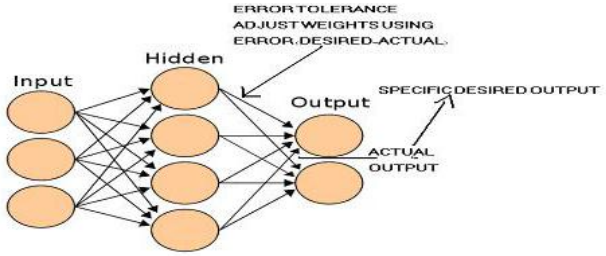

Figure 3- Structure of Back-propagation Neural Network (BPNN)

Error data at the output layer is back-propagated to earlier ones, allowing incoming weights to these layers to be updated. The back propagation algorithm has been widely used as a learning algorithm in feed forward multilayer neural networks [9]. To train the BPNN, a set of image vectors from each infection degree of patient class was required. These image vectors are collected for each patient and stored in databases. Taking these image vectors as an input, comparing the current output with the target output, and adjusting the weight values according to the back-propagation training algorithm. When the error of the training vector set is reduced to a pre-defined threshold which is the total summed squared error less than or equal to threshold, training is stopped, algorithm (2) depicts BPNN algorithm [8].

Algorithm (2).

Step-1(Initialization)

BPNN to determine the infection degree with the parameters adjusted as follows:

a. Number of input layer nodes (IN) depends on the image vector length plus 1 node as bias.

b. Number of hidden layer nodes (HN) is set to 2*IN+1.

c. Number of output layer nodes (ON) is set to 3 since the number 5 represents by 3 bits with binary system to determine which infection degree among 5 degrees

d. Learning rate η is set to (0.1)

e. Momentum α is set to (0.5)

Step- 2 (Compute activation function(sigmoid) for all hidden nodes)

$$h_i = \frac{1}{1 + e^{-\sum_{i=0}^{I} w1_{ij} * x_i}}, \; for\; j = 1..H \qquad \textbf{(9)}$$

Step- 3 (Compute activation function for output nodes)

$$o_i = \frac{1}{1 + e^{-\sum_{i=0}^{H} w2_{ij} * h_i}}, \; for\; j = 1..O \qquad \textbf{(10)}$$

Step- 4 (Compute errors in the output nodes)

$$\delta 2_j = o_j(1-o_j)(y_j - o_j),\ for\ j = 1..O \quad (11)$$

Step- 5 (Compute errors in the hidden nodes)

$$\delta 1_j = h_j(1-h_j)\sum_{i=1}^{c} \delta 2_i \times w2_{ij},\ for\ j = 1..H \quad (12)$$

To train the BPNN, a set of samples from each infection degree was required. These samples are collected for each patient and stored in databases. Training consists of taking an image vector as an input, comparing the current output with the target output, and adjusting the weight values according to the BPNN training algorithm. Figure-4 depicts the proposed VDS with BPNN algorithm.

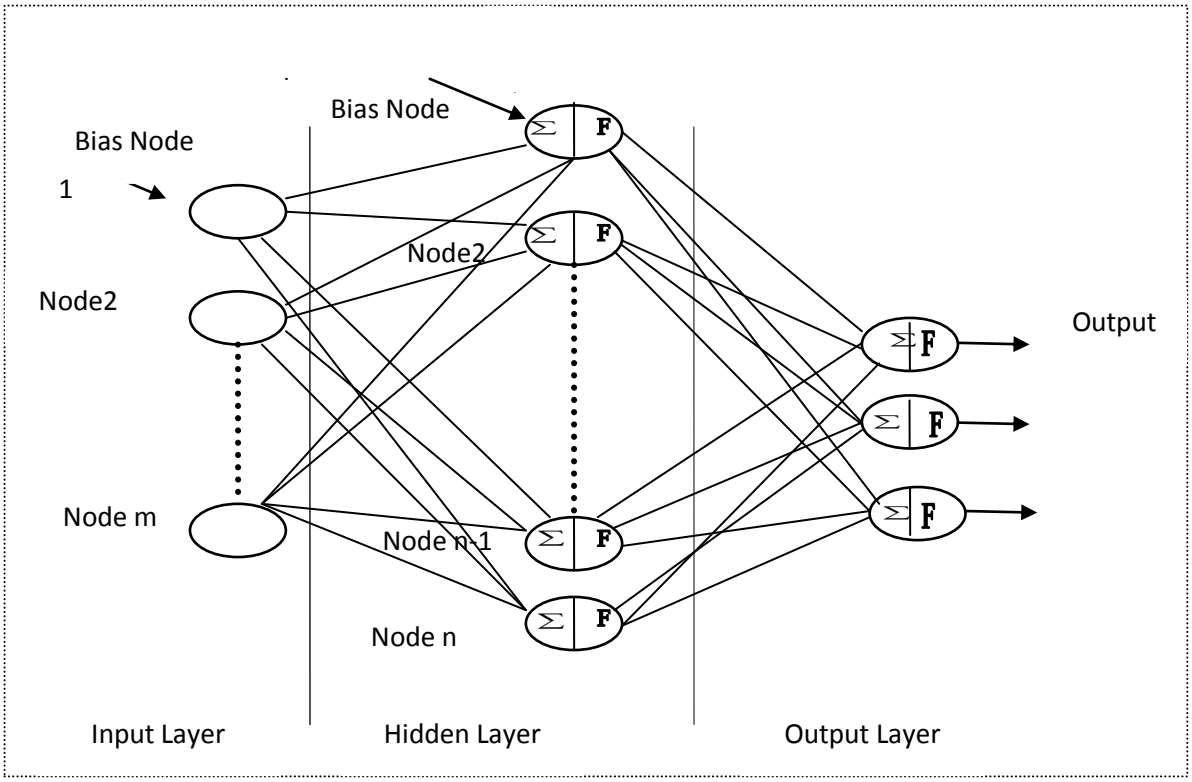

Figure 4- BPNN Structure of VDS

3-Experimental Result Analysis

This section investigates the performance of the proposed VDS. In testing phase the image of the patient's leg is obtained and preprocessed, then the feature vector of the adopted approach is calculated, and the resulting feature vector is compared with those stored in database and the patient's leg is recognized was infect or not by applying NBC, then determine the degree of infection by applying BPNN. The NBC is applied in two stages: training and testing, in training stage the algorithm was trained on 85 samples divided into 50 infected and 35 not infected. While in testing phase two experiments are performed, in the first experiment the NBC was tested on the same samples that it was trained on to ensure the good level of training, While in the second experiment 45 new samples was tested as 25 infected and 20 as not infected and also the produced results indicate to good performance of proposed diagnose system. A total of 112 infected samples are participated in the BPNNexperiments. The participated images are divided into five groups according to infection degree and these 112 samples include the 50 infected samples that used with NBC in previous phase in addition to 62 new infected samples. For each one of these infected samples was classified according to corresponding infection degree as illustrated in Table-1.

Table 1- Samples Distribution According to Infection Degree

Infection Degree	No. of Samples
1	22
2	20
3	18
4	27
5	25

Two experiments are performed independently on different infection samples using BPNN to determine the infection degree. The first experiment applied on the same 112 samples in training phase, while the second experiment applied on the same 50 sample using in training phase in addition to 45 new infected samples. The produced results indicate to high accuracy of VDS when using BPNN algorithm to determine the degree of infection of varicose disease. To evaluate the performance of the proposed VDS, three metrics are used including the False Rejection Rate (FRR) (i.e. the rate at which incorrectly rejects), False Acceptance Rate (FAR) (i.e. the rate at which incorrectly matching the input with a template) and accuracy (Acc) (i.e. the proportion of true results in the population). Tables-2,3 illustrate the results of the proposed VDS.

Table 2- The results of NBC Experiments

Metrics	EXP1	EXP2
FAR%	0	1
FRR%	0	5
Acc%	100	98

Table 3- The results of BPNN Experiments

Metrics	EXP1	EXP2
FAR%	0	5
FRR%	0	3
Acc%	100	97.5

This section demonstrates the good performance of proposed VDS, as illustrated in Tables-2,3. The low rate of two metrics (FAR, FRR) and high (Acc) in second experiment it comes as reflection of a good level of training in both algorithms NBC and BPNN with perfect parameters. In addition, the proposed system VDS is a good system with new design and it can be used in medicine to facilitate the diagnoses operation of varicose disease with low cost.

4-Conclusion

In this paper a classification method with (infected /not infected) samples was presented by applying NBC and BPNN sequentially. The primary goal of the proposed VDS is to determine the varicose infection and its degree with low FAR, low FRR and high Acc. The BPNN was applied on different degrees of infected samples to determine the degree of infection, the tested results demonstrate that both of NBC and BPNN was able to distinguish image is infected or not and determine the degree of infection with low FAR, low FRR and high Acc.

References

1. Masaki W., Toshio, E. and Morito S. and Shigeru S.**2005**.Palm vein authentication technology and its applications, The Biometric Consortium Conference, September 19-21, USA.
2. Alexander R. Schlant R. and Fuster V.**1998** . *The Heart*. Ninth Edition. New York: McGraw-Hill.
3. Fazal M. and Baharum B. **2012**. *The Statistical Quantized Histogram Texture Features Analysis for Image Retrieval Based on Median and Laplacian Filters in the DCT Domain*, IAJIT First Online Publication.
4. Gelman, Andrew. **2008**.*Variance, analysis of*. The new Palgrave dictionary of economics. Second Edition. Basingstoke, Hampshire New York: Palgrave Macmillan.
5. Balagani, K., Phoha, S.V., Ray, V. A. And Phoha, S. **2011**. On the Discriminability of Keystroke Feature Vectors Used in Fixed Text Keystroke Authentication, *Pattern Recognition Letters*, 32(7), pp:1070-1080.

6. Leung, K. C. and Leung, C. H.**2011**.Improvement of Fingerprint Retrieval by a Statistical Classifier, IEEE *Transactions on Information Forensics and Security*, 6(1), pp:59 -69.
7. Ludmila Kuncheva, I. **2003**.*Combining Pattern Classifiers Methods and Algorithms*, Bangor, Gwynedd, United Kingdom.
8. Ham F.M. and Kostanic I. **2001**. *Principles of Neurocomuting for Science and Engineering*. McGraw Hill.
9. Jain, AL. and Sharath Pankanti, H.**2003**. Biometric identification systems, signal processing, *ACM*, 83 (12), pp: 2539–2557.