Design and Development of Narrow Band Imaging and Stacked Neural Networks Approach for Cancer Diagnosis

Dr.T. Rammohan[1], Dr.E.Kaliappan[2] ,A.Nagarajan[3]
[1,3] Department of Electrical and Electronics Engineering, Karpagam College of Engineering, Coimbatore, India
[2] Department of Electrical and Electronics Engineering, Easwari Engineering College Chennai, Tamil Nadu, India

ABSTRACT:

In current scenario, narrow band imaging (NBI) is a very powerful image enhancement technology that enhances the visibility of the organ. Laryngeal microvascular networks can be diagnosed and analysed without biopsy and pathological examination. In order to overcome this identified existing problem, a new automation method to identify the lesion and the classify approach is proposed in this article. The developed algorithm is utilized to NBI endoscopic images of laryngeal diseases and the segmentation and the classification accuracies are analysed. The experimental results show the proposed algorithm provides reliable results. This result highly motivates and it proves the feasibility of the new method and supports the investment in further research and development to translate this study into clinical practice. Further-more, to our best knowledge, this is the first time, image processing is used to automatically classify laryngeal tumors in endoscopic videos based on tumor vascularization characteristics. Therefore, the proposed system is represented as an innovation in biomedical and in health informatics.

Keywords: Blood vessel segmentation, computer-aided diagnosis, laryngoscopy, lesion detection, narrow-band imaging (NBI), Markov Random field segmentation, Deep Neural Network.

I. INTRODUCTION

Blood vessels analysis can play an important role in detecting a wide range of diseases. Changes to blood vessel networks are observed as a consequence of many laryngeal diseases. Even the tiniest lesions alter the surrounding vasculature both physiologically and morphologically. Therefore, detecting lesions in the larynx at early stages is one of the most important factors involved in the successful disease treatment. The medical images used as a diagnostic tool which is analyzed and often provide insight and allows enhancing the image features [1]. About 10% of the precancerous lesions are transformed into squamous cells carcinomas, which are one of the most common types of head and neck cancer. Detection and treatment applied at an early disease stage provides highly favorable results, having five year surviving rate between 80% and 90%. However, false cancer detection is also present in modern diagnosis, which helps specialists to avoid overlooking cancer at earlier stages.

To assist the process, a great deal of effort is made by scientists to improve pretreatment evaluation and to provide intra operative data on the pathologies. Automatic detection of larynx cancer employs edge detection using the spatial neighborhood of pixels and incorporates the clustering process to differentiate the types of tissues [2]. In order to obtain the histopathology of an abnormality, the analysis of endoscopic images may be one of the most accessible methods. Such images can potentially be used for lesion detection and classification, as proposed in this paper. Narrow-band imaging (NBI) gives more information about the laryngeal tissue. The NBI is an optical technology that enhances the practitioner's capability to detect and diagnose lesions through endoscopic inspection.

Throat cancer can occur in the esophagus (food pipe), larynx (voice box), thyroid gland or cells lining of the throat (squamous cells). The early detection and characterization of throat cancer helps reduce the need for therapeutic treatment and minimizes pain and suffering [3-5]. It allows better characterization of tissues with the use of a filtered spectrum illumination system that optimizes the scattering and absorbance of the light in tissue. In NBI, wavelengths in the visible spectrum are filtered from the illumination source. These wavelengths coincide with peaks in the absorption spectrum of oxy haemoglobin, so blood vessels are pronounced when viewed in NBI mode. Various disorders of throat are characterized [6] and different characterization methods are evaluated for the automatic detection [6]. Features are modeled by Hidden Markov Model and classified by Support Vector Machine (SVM) to identify the pathology [7].

Therefore, in this article it is proposed to a new processing framework approaches for the automatic detection and classification of laryngeal lesions based on the segmentation and analysis of blood vessel networks in NBI endoscopic video.

II. METHODOLOGY

The methodology of the proposed system consists of (1) preprocessing of the image using median filter. This has the effect of eliminating pixel values which are unrepresentative of their surroundings. (2) Markov Random Field segmentation algorithm is responsible for locating suspect areas within the image. (3) Segmentation method has been proposed separately for blood vessel segmentation. (4) Deep neural network is used to classify the probability of malignancy and statistical analyses of blood vessels.
Consequently, the great merit of this technology over white-light imaging comes from the improved contrast when visualizing microblood vessels and their patterns, as illustrated in Fig. 1.

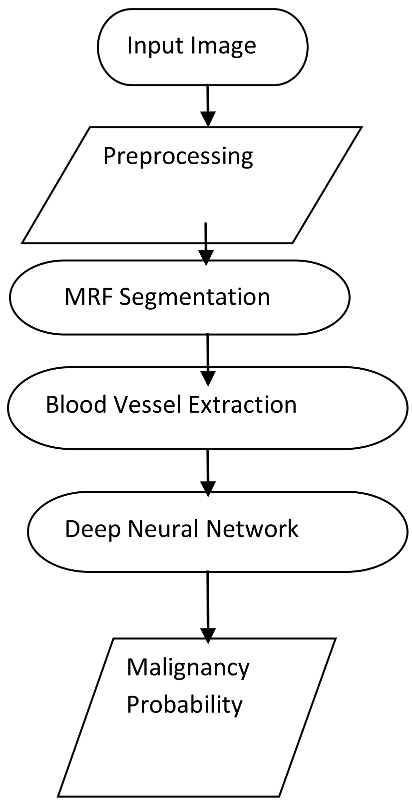

Fig. 1. Block diagram of the proposed system.

A. Preprocessing

Noise present in image is random (not present in the object imaged) variation of brightness or color information in images, and is usually an aspect of electronic noise. It can be produced by digital camera or by the sensor and circuitry of a scanner. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is an unwanted by-product of image capture that adds spurious and extraneous information. The magnitude of image noise can range from almost imperceptible specks on a digital photograph taken in good light. All medical images contain some visual noise. The presence of noise gives an image a mottled, grainy, textured, or snowy appearance. Image noise comes from a variety of sources. Almost all imaging method is free of noise, but noise is much more prevalent in certain types of imaging procedures than in others. Although noise gives an image a generally undesirable appearance, the most important factor is that noise can cover and reduce the visibility of certain features within the image. The loss of visibility is especially significant for low-contrast objects. It is often desirable to be able to perform some kind of noise reduction on an image or signal.

Median filter replaces each pixel value in an image with the neighbor's median value, including itself. This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Median filtering is usually thought of as a convolution filter. Like other convolutions it is based around a kernel, which represents the shape and size of the neighborhood to be sampled when calculating the median. Often a 3×3 square kernel is used. Although larger kernels (*e.g.* 5×5 squares) can be used for more severe smoothing. A small kernel can be applied many times in order to produce a similar but not identical effect as a single pass with a large kernel. A Markov network or MRF is similar to a Bayesian network in its representation of dependencies; the differences being that Bayesian networks are acyclic and directed, whereas Markov networks are undirected and may be cyclic. Thus, a Markov network represents certain dependencies that a Bayesian network cannot (such as cyclic dependencies); on the other hand, it can't represent certain dependencies that a Bayesian network can (such as induced dependencies). The underlying graph of a Markov random field may be finite or infinite.

Gibbs random field is a field where the joint probability density of the random variables is strictly positive. Because, according to the Hammersley – Clifford theorem, it can be represented by a Gibbs measure for an appropriate (locally defined) energy function. The segmentation relies then on a classification step, which can be either non-supervised or supervised [8]. In unsupervised approaches, the pixels of the enhanced image are classified as vessel or non-vessel, by comparing their intensity level to a threshold. The supervised-classification approaches reach better accuracy than the unsupervised one.

Geman and Geman suggested the application of Markov random fields (MRF) for images in early 1984. Their strong mathematical foundation and ability to provide a global optima. Even when defined on local features, this method proved to be the foundation for novel research in the domain of image analysis, de-noising and segmentation. MRFs are completely characterized by their marginal and prior probability distributions, cliques and smoothing

Constraint as well as criterion for updating values. The criterion for image segmentation using MRFs is stated as to find the labeling scheme which has maximum probability for a given set of features. The broad categories of image segmentation using MRFs are supervised and unsupervised segmentation. Segmentation technique [9] requires very accurate edge detection which is not always achievable due to the presence of noise.

B. Supervised image segmentation using MRF and MAP

During image segmentation, the function that MRFs seek to maximize is the probability of identifying a labelling scheme given a particular set of features is detected in the image. This is a restatement of the Maximum a posteriori estimation method.

1. Define the neighborhood of each feature (random variable in MRF terms).Generally this includes 1st order or 2nd order neighbors.
2. Set initial probabilities $P(f_i)$ for each feature as 0 or 1, where $f_i \in \Sigma$ is the set containing features extracted for pixel i and define an initial set of clusters.
3. Using the training data compute the median (μ_i) and variance (σ_i) for each label. This is termed as class statistics.
4. Compute the marginal distribution for the given labeling scheme

$P(f_i/l_i)$

Using Bayestheorem and the class statistics calculated earlier. A Gaussian model is used for the marginal distribution shown in Eq.1.

$$\frac{e^{-(f_i-\mu(l_i))^2/2\sigma(li)}}{\sigma(li)\sqrt{2\pi}} \tag{1}$$

5. Calculate the probability of each class label given the neighborhood defined previously.
Clique potentials are used to model the social impact in labeling.
6. Iterate over new prior probabilities and redefine clusters such that these probabilities are maximized. This is done using a variety of optimization algorithms described below.
7. Stop when probability is maximized and labeling scheme does not change. The calculations can be implemented in log likelihood terms as well.

Optimization algorithms: Models from a variety of fields are adapted and they are set apart by their unique cost functions. These models together constitute optimization. The common trait of cost functions is to make some change in pixel value and difference in pixel label when compared to labels of neighboring pixels.

Iterated conditional modes/gradient descent:

The ICM algorithm tries to reconstruct the ideal labeling scheme by changing the values of each pixel over each iteration and evaluating the energy of the new labeling scheme using the cost function given below,

$$\alpha(1-\delta(li - linitial\ i) + \beta\sum(1-\delta(li, lq(i)) \tag{2}$$

where α is the penalty for change in pixel label and β is the penalty for difference in label between neighboring pixels and chosen pixel in equation (2). Here N(i) is neighborhood of pixel i and δ is the Kronecker delta function.

A major issue with ICM is that, similar to gradient descent, it has a tendency to rest over local maxima and thus not obtain a globally optimal labeling scheme.

C. Blood Vessel Extraction Algorithm Background Exclusion, Thresholding and post filtration

The main purpose of this step is to eliminate background variations in illumination from an image so that the foreground objects may be more easily analyzed. In the proposed algorithm, the background exclusion is obtained by subtracting the original intensity image from the average-filtered image.

Median filter is one of the simplest local operations over an image, which is also called as “neighborhood average method”. The essential idea of a standard moving average filter is to replace the value of the center. The aim of this module is to produce a binary image in which the value of each pixel is either 1 (blood vessel) or 0 (background). Unfortunately, there no thresholding technique is present to determine a unique threshold value which will provide perfect results in all cases. However, in the proposed algorithm, we use Isodata technique. It provides an automatic threshold value for producing a binary image B. This technique divides the histogram into two parts, P_1 andP_2 using an initial threshold value T_0. Subsequently, the median values μ_1 and μ_2 of both the parts are calculated, and a new threshold value is determined which represents the average of μ_1 and μ_2. This process is repeated iteratively until the threshold values T_k and T_{k-1}, converge. The gray-scale/green-channel images $h(x, y)$ are then converted to binary images based on the threshold values T_k. B(x,y)={ 1 if $h(x,y)\geq T_k$ (6)0 otherwise}.Due to thresholding, some unnecessary pixels would appear as noise (false positive) in the resultant binary image, and therefore, some post-processing is adopted to refine the image and retaining the desired objects. For this purpose, a morphological operation (opening) is employed to remove the undesired objects that have fewer than 35 pixels. The bright circle matching to the edge of retina is removed by subtracting a mask image I_{mask} from the binary image, as illustrated in equations (3) to (6).

Let,

$$I_{rgb}=R+G+B \tag{3}$$

$$I_{bw}= inv(I_{rgb}\geq 100) \tag{4}$$

$$I_{mask}=\delta_{se}(I_{bw}) \quad (SE=3 \text{ X } 3 \text{ pixels}) \tag{5}$$

$$BW(x,y)=\{ 1 \quad \text{if } B-I_{mask} > 0 \text{otherwise}\} \tag{6}$$

D. Deep Neural Networks (DNN)

For pattern recognition, artificial neural network usually use the feed-forward technique which includes multi-layer perceptron network [10]. To increase the efficiency this paper indices the Deep Neural Networks (DNN). Deep learning is the name we use for "stacked neural networks"; that is, networks composed of several layers. The layers are made of nodes. Computation happens in a place called node. A node is a loosely patterned on a neuron in the human brain, which fires when it encounters sufficient stimuli. A node combines input from the data with weights that either amplify or dampen that input and with a set of coefficients, or, thereby assigning significance to inputs for the task the algorithm is trying to learn.

Deep-learning networks are different from the single-hidden-layer neural networks by their depth; that is, the number of node layers through which data passes in many steps process of pattern recognition.

Long established machine learning depends on shallow nets, composed of one input and one output layer, and at most one hidden layer in between. More than three layers (including input and output) qualify as "deep" learning. So deep is a strictly defined, technical term that median more than one hidden layer. In deep-learning networks, each layer of nodes trains on a distinct set of features which depends based on the previous layer's output. The further you advance into the neural net, the more complex the features your nodes can recognize, since they aggregate and recombine features from the previous layer.

This is referred as feature hierarchy, and it is also a hierarchy of increasing complexity and abstraction. It makes deep-learning networks capable of handling very large, high-dimensional data sets with billions of parameters that pass through nonlinear functions.

III. EXPERIMENTAL AND SIMULATION RESULTS DISCUSSIONS

The image segmentation technique for detection of throat cancer in endoscopic images has proposed and implanted with the image size of 512x512 pixels containing the cancerous region of the throat is considered which is shown in Fig. 2.

Fig. 2. Input larynx affected image.

The pre-processing stage involves smoothing the color images using the filter which is shown in Fig. 3. The featured image in Fig. 4 shows that the pre-processed image is segmented with Markov Random Field and by the proposed method Deep Neural Network.

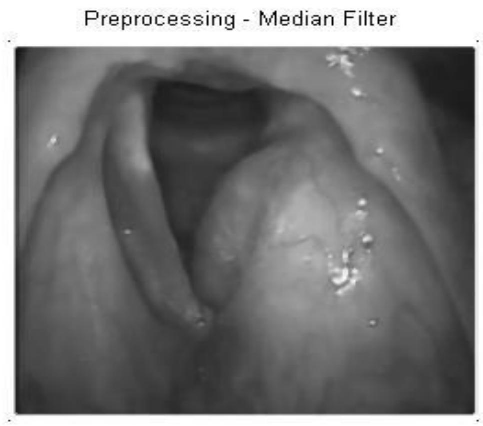

Fig 3. Pre-processed image.

Fig. 4. Feature extracted image.

The feature selection is used to select the features and to reduce the cost computing in the most significant features. DNN is applied after the feature selection to extract the larynx area and the experimental evaluation result shows that our proposed framework works better than the previous methods.

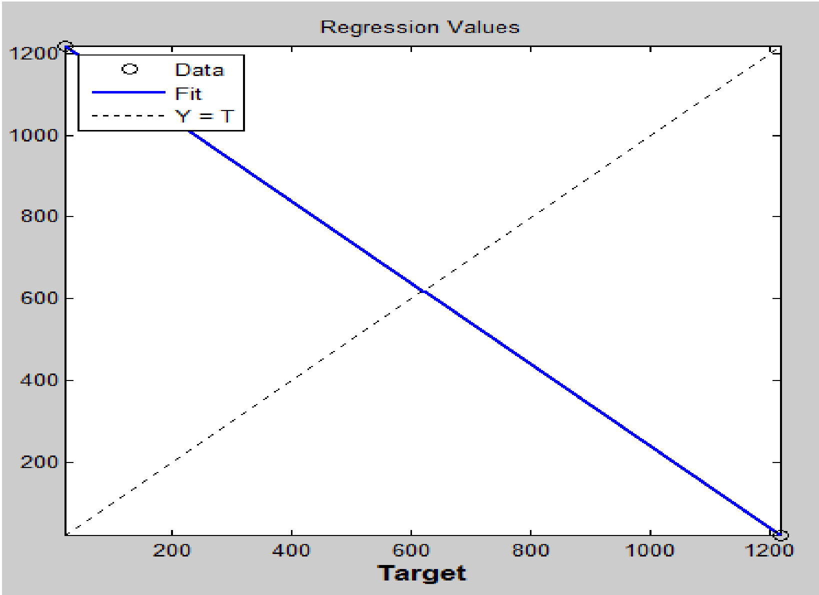

Fig. 5. Regression value of larynx image.

The DNN diagram which is shown in fig 5 is applied to get the regression value which is also known as coefficient determination is shown in fig 6.

Table 1. Execution time of classifier.

Classifier	Training Time	Testing Time
DNN	3.7863	0.0015s

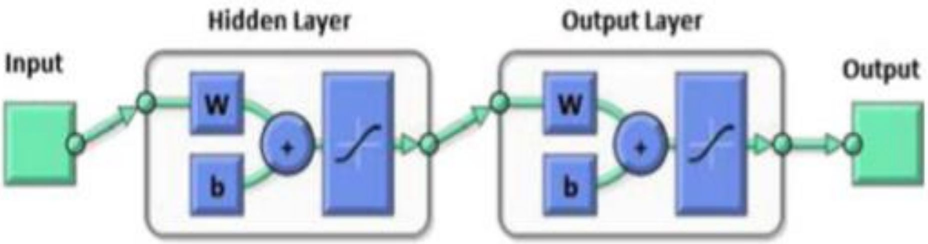

Fig 6. DNN diagram.

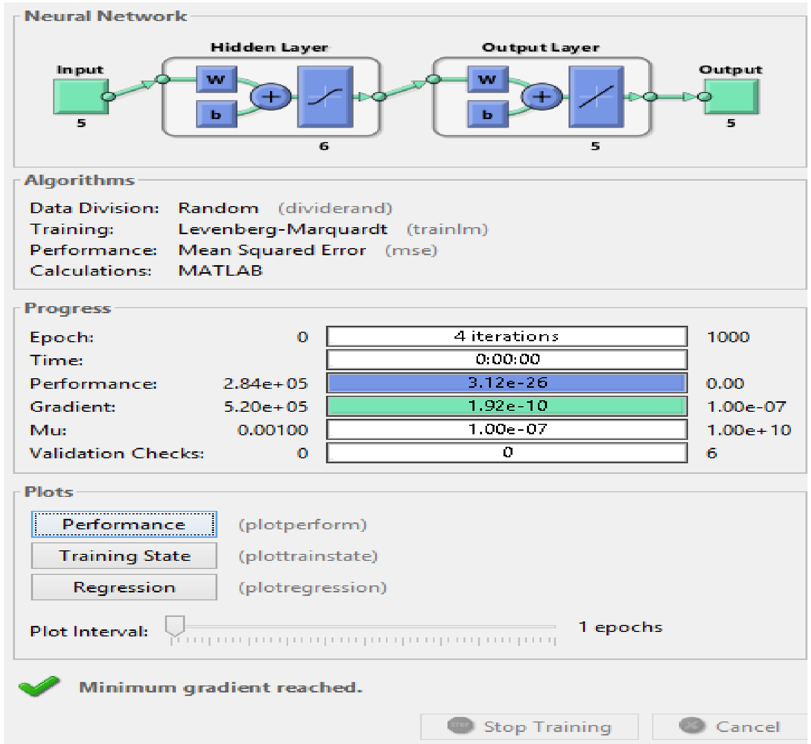

Fig. 7. Simulated framework performance metrics.

Table 1 and 2 describes the time taken for the execution of classifier and the execution time of each and every block.

Table 2. Execution time of each block.

STAGES	EXECUTION TIME (seconds)
Image pre-processing	1.04
MRF Segmentation	1.580
Deep Neural Network	7.387
Malignancy Probability	0.36
Total	29.367

IV. CONCLUSIONS

In this article, a design and development of NBI and stacked neural networks approaches for cancer diagnosis has been successfully simulated through computer simulation with help of the MATLAB/Simulink software platform and experimental analysis. Detection of tumors in Larynx images is a challenging task and the experimental results shown the proficient performance of the designed method. The framework consists of the (1) preprocessing by median filter, (2) Segmentation by Markov Random Field method (3) feature extraction using ORB and GLCM methods and then the (4) DNN methods which makes compromise on computational time for feature selection on accuracy and calculates the image based features reflecting textural and structural information and classifies the malignancy probability. In future, a separate prototype model of designed system is going to be carried out.

References

[1] S. Deepti and B. Dhandra “Abnormality detection in endo-scopic images of throat cancer by morphological operations,” Indian Streams Research Journal, Vol. 1, No. 4, pp. 2011.
[2] Kadria Ezzine, Ahmed Ben Hamida, Zaineb Ben Messaoud and Mondher Frikha “Towards a computer tool for automatic detection of laryngeal cancer,” IEEE 2nd International Conference on Advanced Technologies for Signal and Image Processing, 2016.
[3] http://www.healthline.com/channel/throat-cancer.html
[4] http://www.cancercenter.com/throat-cancer-symptoms.html
[5] http://www.cancer.org/document/PED/content/PED-2-3X-Endoscopy.asp
[6] J.R. Orozco-Arroyave, E.A. Belalcazar-Bolanos “Characterization Methods for the Detection of Multiple Voice Disorders: Neurological, Functional, and Laryngeal Diseases,” *IEEE journal of biomedical and health informatics*,Vol. 19, No. 6, pp. 2168-2194, 2015.
[7] C. M. Travieso-Gonzalez, J. B. Alonso-Hernandez, J. R. Orozco-Arroyave, J. Sole-Casals, and E. Gallego-Jutgla “Automatic detection ´ of laryngeal pathologies in running speech based on the HMM transformation of the nonlinear dynamics,” *Lecture Notes Comput. Sci.*, Vol. 79, No. 11, pp. 136–143, 2013.
[8] F. Rossant, M. Badellino, A. Chavillon, I. Bloch, and M.Paques “A morphological approach for vessel segmentation in eye fundus images with quantitative evaluation,” *J. Med. Imag. Health Informat*., Vol. 1, pp. 42-49, 2011.
[9] S. Hwang and M. celebi “Polyp detection in wireless capsule endoscopy videos based on image segmentation and geometric feature,” *In Proc. IEEE Int. Conf. Acoustics speech signal process.*, pp. 678-681, 2010.
[10] J. K. Basu, D. Bhattacharyya and T. Kim, "Use of Artificial Neural Network in Pattern Recognition", *International Journal of Software Engineering and ts Applications*, Vol. 4, No. 2, pp. 2010.