Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol 3, No.2, 2012
www.iiste.org

Artificial Neural Network based Cancer Cell Classification (ANN – C3)

Guruprasad Bhat[1*] Vidyadevi G Biradar[2] H Sarojadevi[2] Nalini N[2]

1. Cisco Systems, SEZ Unit, Cessna business park, Marathahalli-Sarjapur outer ring road, Bangalore, Karnataka, India – 560 103
2. Nitte Meenakshi Insitute of Technology, Bangalore , Karnataka, India – 560 064

*guruprasadbharatibhat@gmail.com,
vgb2011@gmail.com,hsarojadevi@gmail.com,nalinaniranjan@hotmail.com

Abstract

This paper addresses the system which achieves auto-segmentation and cell characterization for prediction of percentage of carcinoma (cancerous) cells in the given image with high accuracy. The system has been designed and developed for analysis of medical pathological images based on hybridization of syntactic and statistical approaches, using Artificial Neural Network as a classifier tool (ANN) [2]. This system performs segmentation and classification as is done in human vision system [1] [9] [10] [12], which recognize objects; perceives depth; identifies different textures, curved surfaces, or a surface inclination by texture information and brightness.

In this paper, an attempt has been made to present an approach for soft tissue characterization utilizing texture-primitive features and segmentation with Artificial Neural Network (ANN) classifier tool. The present approach directly combines second, third, and fourth steps into one algorithm. This is a semi-supervised approach in which supervision is involved only at the level of defining structure of Artificial Neural Network; afterwards, algorithm itself scans the whole image and performs the segmentation and classification in unsupervised mode. Finally, algorithm was applied to selected pathological images for segmentation and classification. Results were in agreement with those with manual segmentation and were clinically correlated [18] [21].

Keywords: Grey scale images, Histogram equalization, Gausian filtering, Haris corner detector, Threshold, Seed point, Region growing segmentation, Tamura texture feature extraction, Artificial Neural Network(ANN), Artificial Neuron, Synapses, Weights, Activation function, Learning function, Classification matrix.

1. Introduction

In the modern age of computerized fully automated trend of living, the field of automated diagnostic systems plays an important and vital role. Automated diagnostic system designs in Medical Image processing are one such field where numerous systems are proposed and still many more under conceptual design due explosive growth of the technology today. From the past decades, we have witnessed an explosive growth of Digital image processing for analysis of the data that can be captured by digital images and artificial neural networks are used to aggregate the analyzed data from these images to produce a diagnosis prediction with high accuracy instantaneously where digital images serve as tool for input data [20] [21]. Hence in the process of surgery these automated systems help the surgeon to identify the infected parts or tumors in case of cancerous growth of cells to be removed with high accuracy hence by increasing the probability of survival of a patient. In this proposal one of such an automated system for cancer cell classification which helps as a tool assisting surgeon to differentiate cancerous cells from those normal cells i.e. percentage of carcinoma cells, instantaneously during the surgery. Here the pathological images serve as input data. The analysis of these pathological images is directly based on four steps: 1) image filtering or enhancement, 2) segmentation, 3) feature extraction, and 4) analysis of extracted features by pattern recognition system or classifier [21]. Since neural network ensembles are used as decision makers even though network takes more time to adapt behavior, once it is trained it classifies almost instantaneously due to electrical signal communication of nodes in the network.

2. System architecture

The ANN – C3 architecture is shown in figure 1. It comprises of five distinct components, as show below. Each component is described briefly in subsequent sections.

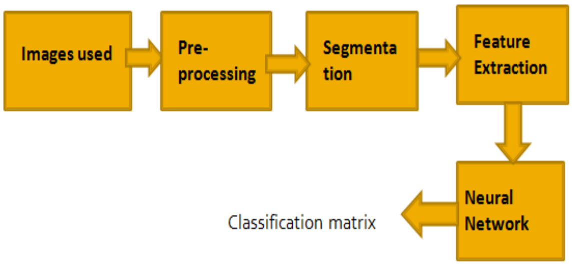

Figure 1: ANN - C3 system architecture

2.1 Images used

This system is designed and verified to take grey scale pathological images as input. Grey scale pathological images help to identify affected cells makes these images for analysis of cancerous growth of cells.

2.2 Pre-processing

Grey scale pathological imaging process may be dirtied by various noises. Perform an image pre processing task to remove noise in a pathological image first. To remove the noise the Histogram equalization or Gaussian filter based median filtering is done [5] [6] [8] [19].

2.3 Segmentation

Segmentation includes two phases. First phase deals with threshold detection and the later one with similar region identification. For threshold detection various methods like GUI selection, graphical method or corner detectors can be used. GUI selection reduces automation and graphical method fails when multiple objects are present in an input data. Since this design mainly deals with multiple objects (cells) in an input image, Haris corner detectors are used to find threshold. In second phase, threshold points detected by corners serve as seed point for segmentation. Four neighborhood based region growing segmentation is used increase the speed compare to eight neighborhood and increase the accuracy compared to region split and merge i.e. trade off between accuracy and speed. A brief discussion of Haris corner detector and 4-neighborhood region growing Segmentation is done in section III [11] [13].

2.4 Feature Extraction

Neural network classifiers are those differ from traditional classifiers like Bayesian and k – nearest neighborhood classifiers in various aspects from type of input data to output representation. Since the neural networks are used as classifiers in this design which takes only numerical data as input rather than any kind of data as input by Bayesian and k – nearest neighbor classifiers, the input image data has to be converted to numerical form. This conversion is done by extracting tamura texture features. A brief discussion of tamura texture feature is done in section IV.

Tamura texture features:

The human vision system (HVS) permits scene interpretation ‘at a glance’ i.e. the human eye ‘sees’ not scenes but sets of objects in various relations to each other, in spite of the fact that the ambient illumination is likely to vary from one object to another—and over the various surfaces of each object—and in spite of the fact that there will be secondary illumination from one object to another. These variations in the captured images are referred as tamura texture features, even the same texture features are observed by surgeon to differentiate carcinoma cells and non-carcinoma cells.

2.5 Neural Network

Supervised feed-forward back-propagation neural network ensemble used as a classifier tool. As discussed previously, neural network differs in various ways from traditional classifiers like Bayesian and k – nearest neighbor classifiers. One of the main differences is linearity of data. Traditional classifiers like Bayesian and k – nearest neighbor requires linear data to work correctly. But neural network works as well for non-linear data because it is simulated on the observation of biological neurons and network of neurons. Wide range of input data for training makes neural network to work with higher accuracy, in other words a small set of data or large set of similar data makes system to be biased [22]. Thus neural network classifier requires a large set of data for training and also long time to train to reach the stable state. But once the network is trained it works as fast as biological neural networks by propagating signals as fast as electrical signals.

3. Haris corner detector and 4-neighborhood region growing segmentation

3.1 Haris corner detector

A corner can be defined as the intersection of two edges. A corner can also be defined as points for which there are two dominant and different edge directions in a local neighborhood of the point. An interest point is a point in an image which has a well-defined position and can be robustly detected. This means that an interest point can be a corner but it can also be, for example, an isolated point of local intensity maximum or minimum, line endings, or a point on a curve where the curvature is locally maximal.

In practice, most so-called corner detection methods detect interest points in general, rather than corners in particular. As a consequence, if only corners are to be detected it is necessary to do a local analysis of detected interest points to determine which of these real corners are.

A simple approach to corner detection in images is using correlation, but this gets very computationally expensive and suboptimal. Haris corner detector is one such corner detector, which uses differential of the corner score with respect to direction directly, instead of using shifted patches. This corner score is often referred to as autocorrelation.

The algorithm of haris corner detector as follows:

Without loss of generality, we will assume a grayscale 2-dimensional image is used. Let this image be given by I. Consider taking an image patch over the area (u,v) and shifting it by (x,y). The weighted sum of squared differences (SSD) between these two patches, denoted S, is given by:

$$S(x,y) = \sum_u \sum_v w(u,v)\,(I(u+x,v+y) - I(u,v))^2 \quad (1)$$

$I(u+x,v+y)$ can be approximated by a Taylor expansion . Let I_x and I_y be the partial derivatives of I, such that

$$I(u+x,v+y) \approx I(u,v) - I_x(u,v)x + I_y(u,v)y \quad (2)$$

This produces the approximation

$$S(x,y) \approx \sum_u \sum_v w(u,v)(I_x(u,v)x + I_y(u,v)y)^2 \quad (3)$$

This can be written in matrix form:

$$S(x,y) \approx (x\ y)A\begin{pmatrix}x\\y\end{pmatrix} \quad (4)$$

Where A is the structure tensor,

$$A = \sum_u \sum_v w(u,v)\begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2\rangle & \langle I_xI_y\rangle \\ \langle I_xI_y\rangle & \langle I_y^2\rangle \end{bmatrix}$$

(5)

This matrix (5) is a Harris matrix, and angle brackets denote averaging (i.e. summation over (*u*,*v*)). If a circular window is used, then the response will be isotropic [16].

A corner (or in general an interest point) is characterized by a large variation of *S* in all directions of the vector (x,y). By analyzing the eigenvalues of *A*, this characterization can be expressed in the following way: *A* should have two "large" eigenvalues for an interest point. Based on the magnitudes of the eigenvalues, the following inferences can be made based on this argument:

If $\lambda_1 \approx 0$ and $\lambda_2 \approx 0$ then this pixel (x , y) has no features of interest.

If $\lambda_1 \approx 0$ and λ_2 has some large positive value, then an edge is found.

If λ_1 and λ_2 have large positive values, then a corner is found.

Haris and Stephens noted that exact computation of the eigenvalues is computationally expensive, since it requires the computation of a Square root, and instead suggest the following function M_c, where κ is a tunable sensitivity parameter:

$$M_c = \lambda_1\lambda_2 - k(\lambda_1+\lambda_2)^2 = \det(A) - k\,\mathrm{trace}^2(A) \qquad (6)$$

Therefore, the algorithm does not have to actually compute the Eigen value decomposition of the matrix *A* and instead it is sufficient to evaluate the determinant and trace of *A* to find corners, or rather interest points in general.

The value of κ has to be determined empirically, and in the literature values in the range 0.04 - 0.15 have been reported as feasible.

The covariance matrix for the corner position is A^{-1}, i.e.

$$\frac{1}{\langle I_x^2\rangle\langle I_y^2\rangle - \langle I_x I_y\rangle^2}\begin{bmatrix} \langle I_y^2\rangle & -\langle I_x I_y\rangle \\ -\langle I_x I_y\rangle & \langle I_x^2\rangle \end{bmatrix} \qquad (7)$$

Compute x and y derivatives of image

$$I_x = G_\sigma^x * I \qquad (8)$$

$$I_y = G_\sigma^y * I \qquad (9)$$

Compute product of derivatives of each image

$$I_{x2} = I_x * I_x \qquad (10)$$

$$I_{y2} = I_y * I_y \qquad (11)$$

3. Compute the sums of products of derivatives at each pixel

$$S_{x2} = G_{\sigma 1} * I_{x2} \qquad (12)$$

$$S_{y2} = G_{\sigma 1} * I_{y2} \qquad (13)$$

$$S_{xy} = G_{\sigma 1} * I_{xy} \qquad (14)$$

Define at each pixel (x, y) the matrix

$$H(x, y) = \begin{bmatrix} S_{x2}(x,y) & S_{xy}(x,y) \\ S_{xy}(x,y) & S_{y2}(x,y) \end{bmatrix} \qquad (15)$$

Compute the response of the detector at each pixel

$$R = \mathrm{Det}(H) - k(\mathrm{Trace}(H))^2 \qquad (16)$$

Threshold on value R. Compute nonmax suppression.

3.2 4-neigbourhood region growing Segmentation

Segmentation is the process of identifying the region of interest from the input image. Considering an input image I being read and converted to the greyscale image .let's assume the seed point to be (x , y). If the seed point is provided by the GUI then a function getpts() will make sure the x and y axes values have been fetched. To create a mask we'll convert all the pixels in the image I' to 0 and call the image J.

In order to discover the neighbors we will use four pixel connectivity [14]. Starting with the seed point the algorithm looks for the 4 pixels surrounding the pixel in consideration. Every time a surrounding pixel is considered, the region mean is calculated and checked with that of the pixel in consideration and added to the region. Similarly as the pixel is added to the region corresponding pixel in the image J is highlighted to 1 which would result in the highest intensity hence illuminating the pixel. As the segmentation continues the region into consideration is intensified in the image J resulting in the segmentation of the affected area, which later can be combined with the original image and displayed to the user.

5. Tamura Textutre feature extraction

Tamura texture feature concepts proposed by Tamura et al in 1978. These tamura texture features corresponding to human perception and these features examined by 6 different constituent features. Six features are: [15]

Coarseness – Coarseness is the numerical value describing whether texture is coarse or fine.

Contrast – Contrast defines whether texture contrast is high or low.

Directionality – Directionality defines whether texture pallets are oriented in single direction or not i.e. directional or non-directional.

Line-likeness – Line-likeness correspond to pattern elements i.e. whether texture formed by lines i.e. line-like or blob-like.

Regularity – Regularity defines the interval in which patterns repeated. If patterns are repeated in regular interval then the texture is regular else it is said to be Irregular.

Roughness – Roughness defines the whether the surface is rough or smooth.

In these six features, Coarseness, Contrast and Directionality correspond to strong human perception and these features are calculated pixel-wise by creating 3-D histogram of these three features. Estimation of these three features are described in subsequent sections.

Coarseness relates to distances of notable spatial variations of grey levels, that is, implicitly, to the size of the primitive elements (texels) forming the texture. The proposed computational procedure accounts for differences between the average signals for the non-overlapping windows of different size:

At each pixel (x,y), compute six averages for the windows of size $2^k \times 2^k$, k=0,1,...,5, around the pixel.

At each pixel, compute absolute differences$E_k(x,y)$ between the pairs of nonoverlapping averages in the horizontal and vertical directions.

At each pixel, find the value of k that maximises the difference $E_k(x,y)$ in either direction and set the best size $S_{best}(x,y)=2^k$.

Compute the coarseness feature F_{crs} by averaging $S_{best}(x,y)$ over the entire image. Instead of the average of $S_{best}(x,y,$ an improved coarseness feature to deal with textures having multiple coarseness properties is a histogram characterising the whole distribution of the best sizes over the image.

Contrast measures how grey levels q; q = 0, 1, ..., q_{max}, vary in the image **g** and to what extent their distribution is biased to black or white. The second-order and normalised fourth-order central moments of the grey level histogram (empirical probability distribution), that is, the variance, σ^2, and kurtosis, α_4, are used to define the contrast:

$$F_{con} = \sigma / \alpha_4^n \qquad (17)$$

Where,

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \qquad (18)$$

$$\sigma^2 = \sum_{q=0}^{q_{mn}} (q-m)^2 \Pr(q|g) \quad (19)$$

$$\mu_4 = \sum_{q=0}^{q_{mn}} (q-m)^4 \Pr(q\ \lg) \quad (20)$$

and *m* is the mean grey level, i.e. the first order moment of the grey level probability distribution. The value n=0.25 is recommended as the best for discriminating the textures.

Degree of *directionality* is measured using the frequency distribution of oriented local edges against their directional angles. The edge strength $e(x,y)$ and the directional angle $a(x,y)$ are computed using the Sobel edge detector approximating the pixel-wise *x*- and *y*-derivatives of the image:

$$e(x,y)=0.5(|\Delta x(x,y)|+|\Delta y(x,y)|) \quad (21)$$

$$a(x,y)=\tan^{-1}(\Delta x(x,y)/\Delta y(x,y)) \quad (22)$$

where $\Delta_x(x,y)$ and $\Delta_y(x,y)$ are the horizontal and vertical grey level differences between the neighbouring pixels, respectively. The differences are measured using the following 3 × 3 moving window operators:

$$\begin{matrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{matrix} \qquad \begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{matrix}$$

A histogram $H_{dir}(a)$ of quantised direction values *a* is constructed by counting numbers of the edge pixels with the corresponding directional angles and the edge strength greater than a predefined threshold. The histogram is relatively uniform for images without strong orientation and exhibits peaks for highly directional images. The degree of directionality relates to the sharpness of the peaks:

$$F_{dir} = 1 - r n_{peaks} \sum_{p=1}^{n_{peaks}} \sum_{a \in w_p} (a - a_p)^2 H_{dir}(a) \quad (23)$$

where n_p is the number of peaks, a_p is the position of the *p*th peak, w_p is the range of the angles attributed to the *p*th peak (that is, the range between valleys around the peak), *r* denotes a normalising factor related to quantising levels of the angles *a*, and *a* is the quantised directional angle (cyclically in modulo 180^o). Three other features are highly correlated with the above three features and do not add much to the effectiveness of the texture description.

The *linelikeness* feature F_{lin} is defined as an average coincidence of the edge directions (more precisely, coded directional angles) that co-occurred in the pairs of pixels separated by a distance *d* along the edge direction in every pixel. The edge strength is expected to be greater than a given threshold eliminating trivial "weak" edges. The coincidence is measured by the cosine of difference between the angles, so that the co-occurrences in the same direction are measured by +1 and those in the perpendicular directions by -1. The *regularity* feature is defined as $F_{reg}=1-r(s_{crs}+s_{con}+s_{dir}+s_{lin})$ where *r* is a normalising factor and each $s_{...}$ means the standard deviation of the corresponding feature $F_{...}$ in each subimage the texture is partitioned into. The *roughness* feature is given by simply summing the coarseness and contrast measures: $F_{rgh}=F_{crs}+F_{con}$. These features capture the high-level perceptual attributes of a texture well and are useful for image browsing. However, they are not very effective for finer texture discrimination.

6. Artificial Neural Network

A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects [3] [4] [7]:

1. Knowledge is acquired by the network through a learning process.

2. Interneuron connection strengths known as synaptic weights are used to store the knowledge.

- ❖*Benefits of neural network*
- ❖*Nonlinearity.*
- ❖*Input-output mapping.*

- *Adaptivity.*
- *Contextual information.*
- *Fault tolerance.*
- *VLSI implementability.*
- *Uniformity of analysis and design.*
- *Neurobiological analogy.*
- *Model of a neuron*

A *neuron* is an information-processing unit that is fundamental to the operation of a neural network. We may identify three basic elements of the neuron model: [17] [18]

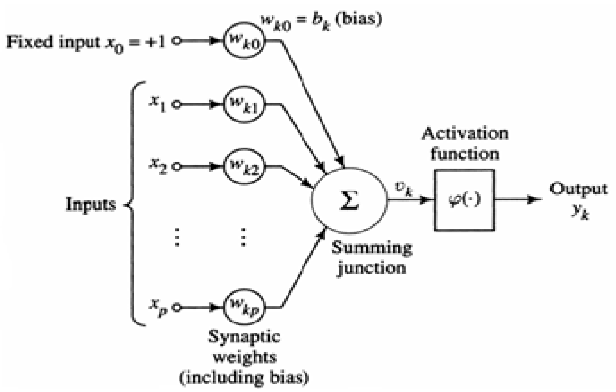

Figure 2:Non-linear model of a neuron.

A set of *synapses*, each of which is characterized by a *weight* or *strength* of its own. Specifically, a signal *xj* at the input of synapse *j* connected to neuron *k* is multiplied by the synaptic weight *wkj*. It is important to make a note of the manner in which the subscripts of the synaptic weight *wkj* are written. The first subscript refers to the neuron in question and the second subscript refers to the input end of the synapse to which the weight refers. The weight *wkj* is positive if the associated synapse is excitatory; it is negative if the synapse is inhibitory.

An *adder* for summing the input signals, weighted by the respective synapses of the neuron.

An *activation function* for limiting the amplitude of the output of a neuron. The activation function is also referred to in the literature as a *squashing function* in that it squashes (limits) the permissible amplitude range of the output signal to some finite value.

Typically, the normalized amplitude range of the output of a neuron is written as the closed unit interval [0, 1] or alternatively [-1, 1].

The model of a neuron also includes an externally applied bias (threshold) *wk0* = *bk* that has the effect of lowering or increasing the net input of the activation function.

Since after feature tamura feature extraction data is in the form of numerical values, Artificial Neural Network classifier suits well for classification. Also non – linearity of the data makes other traditional classifiers like Bayesian and kth – nearest neighbor classifier inefficient compared to ANN classifier. Thus in this system ANN classifier is used as classification tool.

7. Experimental results

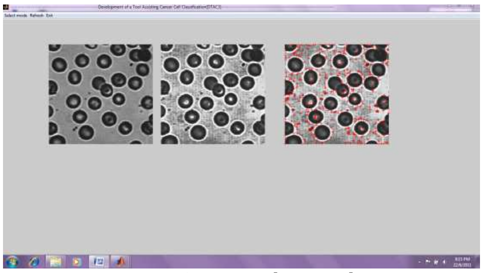

Figure 3: Intermediate result

Figure 3 shows the intermediate result after corner detection. First image in figure 3 is the input image, 2nd image displayed is histogram equalized image. From this histogram equalized image threshold points detected and marked with red + marks as shown in third image of figure3.

From each seed point region is extracted and from extracted region tamura features are calculated. Each feature vector consists of 4 features and n number of such feature vectors can be obtained from single image which helps to prevent the system to be biased.

Extracted feature vectors are sent to neural network. The performance measurement with variable number of hidden layer neurons with single layered feed forward back- propagation network is tabulated in table1:

Index	Number of neurons	Percentage of correct classification
1.	20	96.4286%
2.	21	85.7143%
3.	22	92.8571%
4.	23	92.8571%
5.	24	96.4286%

Table 1- variable number of hidden layer neurons

The performance measurement with variable number of hidden layers with fixed number of neurons of 20 neurons in each layer, feed forward back- propagation network is tabulated in following table:

Index	Number of hidden layers	Percentage of correct classification
1.	1	96.4286%
2.	2	89.2857%
3.	3	85.7143%

Table 2- variable number of hidden layers

8. Conclusion

Even though there is no successful generalized neural network configuration, for a particular application a neural network with acceptable level of accuracy can be designed by selecting suitable number of hidden layers, number of neurons per hidden layer and transfer and learning functions. The performance also depends on the training function parameters like whether it is a batch training or one input at a time. Also we have witnessed the advantages of neural network classifiers over other traditional classifiers like Bayesian and k – nearest neighbor classifiers.

This design can be extended to estimate the number of carcinoma cells per unit area. This estimation helps in automated diagnosis systems like blood purifier in case of blood cancer. Also this can extended to take color image as input with more feature added to feature vector to increase the accuracy of the output.

References

[1]. "An Approach for Discretization and Feature Selection Of Continuous-Valued Attributes in Medical Images for Classification Learning".
[2]. Basavaraj .S. Anami1 and Vishwanath.C.Burkpalli 2 1. Principal, K.L.E.Institute of Technology, Hubli-580030, India 2. Research Scholar, Basaveshwar Engineering College, Bagalkot – 587102, India.
[3]. "Texture based Identification and Classification of Bulk Sugary Food Objects", ICGST-GVIP Journal, ISSN: 1687-398X, Volume 9, Issue 4, August 2009.
[4]. Bing Gong, School of Computer Science and Technology Heilongjiang University Harbin, China, "A Novel Learning Algorithm of Back-propagation Neural Network" ,2009 IITA International Conference on Control, Automation and Systems Engineering.
[5]. Weilin Li, Pan Fu and Weiqing Cao, "Tool Wear States Recognition Based on Genetic Algorithm and Back Propagation Neural Network Model", 2010 International Conference on Computer Application and System Modeling (1CCASM 2010)
[6]. Acharya and Ray, "Image Processing: Principles and Applications", Wiley-Interscience 2005 ISBN 0-471-71998-6
[7]. Russ, "The Image Processing Handbook", Fourth Edition, CRC 2002 ISBN 0-8493-2532-3
[8]. SIMON HAYKIN, Book on "Neural Networks", 2nd edition, A comprehensive edition.
[9]. "Digital Image Processing" by Gonzalez & Woods 2^{nd} edition
[10]. H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," IEEE Trans. On Systems, Man, and Cybernetics, vol. Smc-8, No. 6, June 1978.
[11]. J. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image database".IEEE Intl. Conf. on Image Proc., 1994.
[12]. Castleman K R. "Digital image processing". NJ: Prentice Hall, 1996.
[13]. Manjunath, B., Ma, W.: "Texture features for browsing and retrieval of image data". IEEE Trans on Pattern Analysis and Machine Intelligence 18 (1996) 837842
[14]. Alexander Suhre, A. Enis Cetin, Tulin Ersahin, Rengul Cetin-Atalay, "Classification of cell images using a generalized harris Corner Detector"
[15]. R. Adams and L. Bischof, "Seeded Region Growing", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, pp. 641-647, 1994.
[16]. R. Haralick, "Statistical and structural approaches to texture", Proceedings of the IEEE, vol. 67, pp. 786–804, 1979.
[17]. C. Harris and M.J. Stephens. "A combined corner and edge detector". In Alvey Vision Conference, pages 147–152, 1988. D.
[18]. Ballard and C. Brown, "Computer Vision", Prentice-Hall, Inc., 1982, Chap. 6.
[19]. Davies, "Machine Vision: Theory", Algorithms and Practicalities, Academic Press, 1990, Chap. 18.
[20]. A K Jain, "Fundamentals of Digital Image Processing", Prentice-Hall, 1986, Chap. 9.
[21]. D. Vernon Zhi-Hua Zhou, Yuan Jiang, Yu-Bin Yang, Shi-Fu Chen , "Lung Cancer Cell Identification Based on Artificial Neural Network Ensembles", National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, P.R.China. Artificial Ingelligence in Medicine, 2002, vol.24, no.1, pp.25-36. @Elsevier.
[22]. Leonard Fass , "Imaging and cancer: A review", GE Healthcare, 352 Buckingham Avenue, Slough, SL1 4ER, UK Imperial College Department of Bioengineering, London, UK.
[23]. Jinggangshan, P. R. China , "Application of Neural Networks in Medical Image Processing" ISBN 978-952-5726-09-1, Proceedings of the Second International Symposium on Networking and Network Security (ISNNS '10), 2-4, April. 2010, pp. 02.

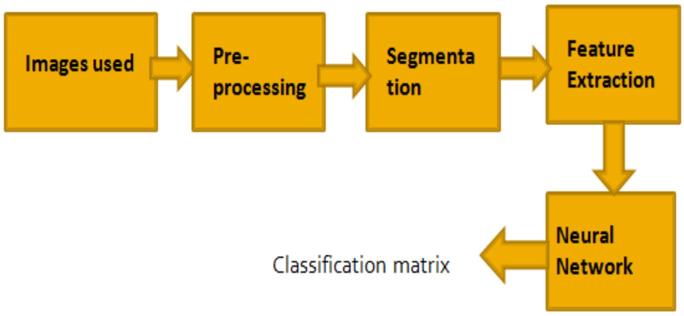

Figure 1: ANN - C3 system architecture

Index	Number of neurons	Percentage of correct classification
1.	20	96.4286%
2.	21	85.7143%
3.	22	92.8571%
4.	23	92.8571%
5.	24	96.4286%

Table 1: Variable number of hidden layer neurons

Index	Number of hidden layers	Percentage of correct classification
1.	1	96.4286%
2.	2	89.2857%
3.	3	85.7143%

Table 2: Variable number of hidden layers