

Design of a Hybrid Intelligent System for Transitional Bladder Cell Carcinoma Diagnosis

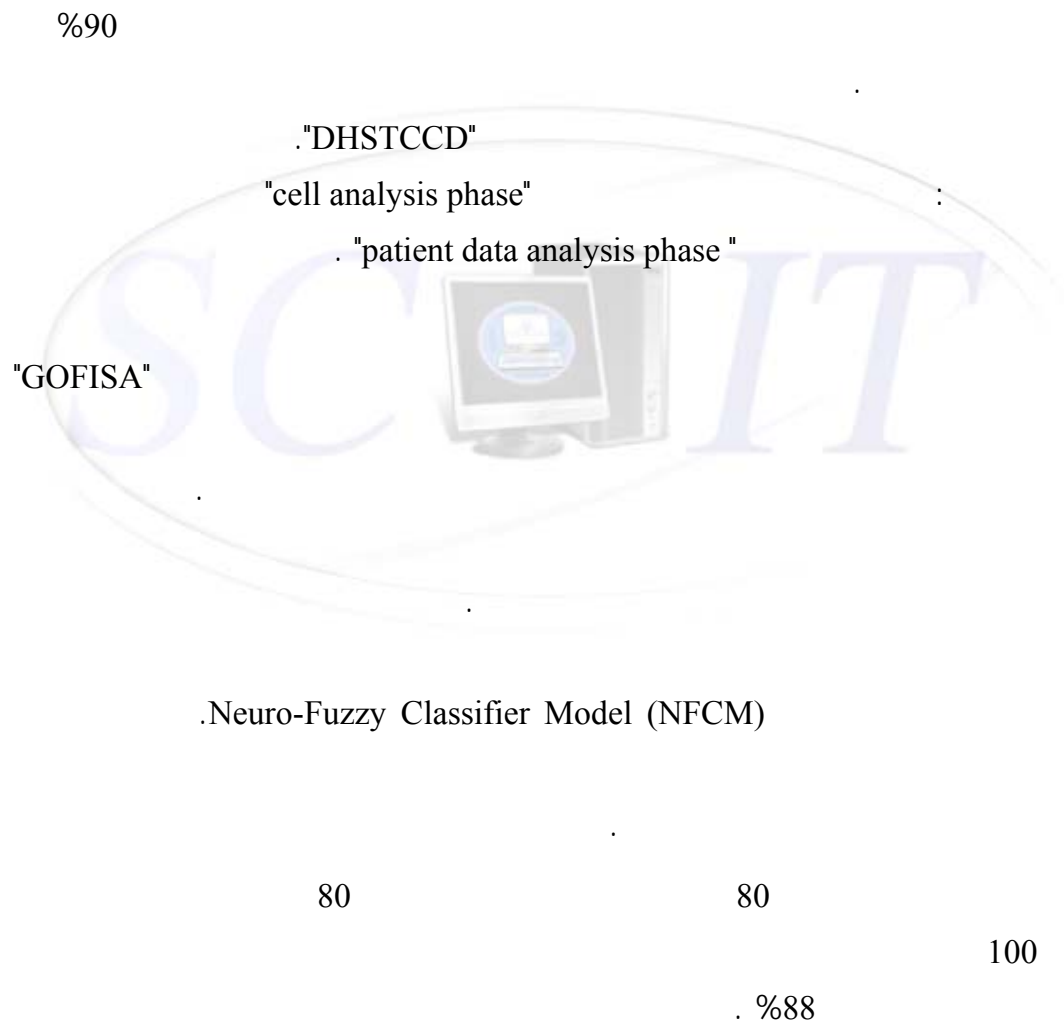
Dr. Nada N. Saleem

Dr. Khalid N. Saleem

ABSTRACT

Transitional Bladder cell carcinoma is one of the most common cancers of the urinary tract, accounting for approximately 90% of Bladder cancers. In this research a new computer-based system "Design of a Hybrid Intelligent System for Transitional Bladder Cell Carcinoma Diagnosis" (DHSTCCD) has been proposed and implemented. The proposed system is composed of two main phases, the first phase is the "cell analysis phase" which consists of three main stages including segmentation stage using "Genetic Optimization Based Fuzzy Image Segmentation Algorithm" (GOFISA), morphometric and photometric feature extraction stage and "Neuro-Fuzzy Classifier Model" (NFCM) stage that has been developed and implemented to classify a set of normal and abnormal cells using hybrid intelligent technique that combines the artificial neural network and fuzzy logic.

The second phase is the "patient data analysis phase", in this phase a rule-based Fuzzy Expert System has been proposed and implemented that uses the laboratory and clinical data and simulates an expert doctor's behavior. The final diagnosis of the patients is determined from the results of the fuzzy expert system and the NFCM.



1. Introduction

The use of computer in medical applications has increased dramatically over the last three decades. Computerized image processing techniques have been used to improve the picture quality, images can be analyzed to highlight areas of interest or to extract meaningful diagnostic features that can provide objective evidence to aid the human decision making process.

Machine learning involves getting computer to induce rules, or statistical models, from a set of "training" examples. These examples can either have the correct class labels associated with them, often called supervised learning, or sub-classes of examples can automatically be found from within the data, often called unsupervised learning [1]. Pattern recognition is a scientific discipline, the goal of which is the classification of objects into a number of categories. In recent years neural computation and hybrid architectures for intelligent systems have appeared as practical technologies with successful applications in this field [2,3]. Although every artificial intelligent technique (i.e., fuzzy logic, neural networks, genetic algorithms and expert systems) has particular computational properties that make it suited for a particular type of problems, there are great advantages in their synergistic utilization [4,5].

Today there is a synergy beginning to form among neural networks, fuzzy logic and genetic algorithms. This synergy has been variously called *Soft Computing* [6,7]. Soft Computing is an area of computing allowing imprecision, uncertainty and partial truth to process and therefore achieves robustness and low solution cost. Hybrid Soft Computing approaches incorporates all the features from individual fields and, moreover, has the ability to overcome difficulties and limitations that characterize each field. The use of intelligent hybrid systems is growing rapidly with successful applications in many areas including process control, robotics, manufacturing, medical diagnosis, etc. [8].

In the present research, a computer-based system (DHSTCCD) model is proposed, the (DHSTCCD) system incorporates Fuzzy logic, Neural Network and Genetic Algorithm which allow it to recognize complex patterns in data and are more flexible than other techniques.

2. Bladder Cancer disease

Bladder cancer is a disease in which the cells lining the urinary bladder lose the ability to regulate their growth and start dividing uncontrollably. This abnormal growth results in a mass of cells that forms a tumor [9].

There are several types of bladder cancer. The most common type is Transitional Cell Carcinoma (*TCC*), the one we deal with. *TCC* is by far the most common histologic variant of bladder cancers, accounting for approximately 90% of urinary bladder cancers. It represents the second most common malignancy of the urinary tract and the fourth most common cancer in men and the eighth most common cancer in women [10].

TCC occurs more often in men than in women and usually affects older adults, the average age at diagnosis is 68 years old [10]. The Bladder walls as shown in Figure (1.1) consist of muscle and the urinary epithelium which contains specialized cells that line the walls of bladder (Transitional cells), from these cells *TCC* of bladder begins.

The epithelial lining of the urinary bladder as shown in Figure (1.2) consists of several cell layers, each cell contains a nucleus and cytoplasm, both enclosed by a thin membrane. The cytoplasm completely surrounds the nucleus. Healthy cells grow and divide in an orderly way. This process is controlled by DNA-the genetic material that contains the instructions for every chemical process in the body. When DNA is mutated, changes occur in these instructions. One result is that cells may begin to grow out of control and eventually form a tumor, a mass of malignant cells [11].

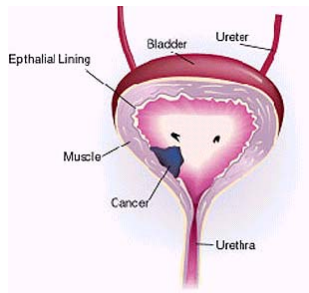
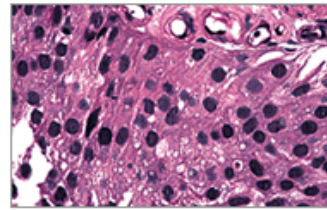


Figure (1.1): Bladder cancer on the inner lining of bladder wall



Figure(1.2): Epithelial lining of the urinary bladder

Risk factor, is something that increases your chance of getting a disease or condition. In general, the following risk factors have been linked to bladder cancer

- **AGE**. The chance of getting bladder cancer increases, as you grow older.
- **RACE**. Caucasians are twice as likely to develop bladder cancer as blacks and Hispanics. Lowest rates of the disease.
- **SEX**. Men are two to three times more likely to get bladder cancer than women.
- **SMOKING**. Smoking is the single greatest known cause of bladder cancer. The risk increases depending on how many cigarettes smoked per day and the number of years you have smoked.
- **INDUSTRIAL CHEMICALS**. Repeated exposure to chemicals used in the manufacture of dyes textiles and paint products may increase the risk of developing bladder cancer years later.
- **CERTAIN DRUGES**. Treatment with the drug cytoxan increases the risk of bladder cancer.
- **CHRONIC BLADDER INFLAMMATION**. Chronic or repeated urinary infection or inflammations predispose to a certain form of bladder cancer.
- **FAMILY HISTORY**. You're at higher risk of bladder cancer if you have family members with TCC.
- **PERSONAL HISTORY**. Having bladder cancer once makes it more likely you will get it again.

Symptoms, the most common symptoms of bladder cancer include:

- blood in the urine (making the urine slightly rusty to deep red)
- Pain during urination.
- Frequent urination or feeling the need to urinate without results.

These symptoms are not sure signs of bladder cancer. Infection, benign tumors, bladder stones or other problems also can cause these symptoms. Anyone with these symptoms should see a doctor so that the doctor can diagnose and treat any problem as early as possible [12].

3. The Proposed DHSTCCD Mode

The proposed system is composed of two main parts: The cell analysis phase which consists of three main stages (segmentation, feature extraction and cell classification) and the patient data analysis phase which is concerned with fuzzy expert system. The basic outlines of the proposed (DHSTCCD) system are illustrated in the Figure (2). The following steps are suggested in a cell analysis phase:

- 1- Preparation of cell specimens.
- 2- Scanning and focusing of the specimen to make it accessible to a computer.
- 3- Searching for cells to analyze.
- 4- Segmentation of cells into nucleus and background.
- 5- Feature extraction: the process of measuring relevant properties from a segmented cell.
- 6- Classification of individual cells.

3.1 GOFISA Image Segmentation Method

The delineation of the cell image into a cell nucleus and background is an essential component of an automatic computerized cell image analysis system in the proposed method. After an image has been captured, the problem of automatic recognition and classification of the cell has two main components. The First one is to be able to isolate the nucleus from the rest of the image. The Second one is to extract useful information from the segmented nucleus and use this information to separate them in various classes.

In this work we have focused on the analysis of cell nuclei because most of the useful diagnostic information is associated with the cell nucleus.

A method named "Genetic Optimization Based Fuzzy Image Segmentation Algorithm" (GOFISA) has been proposed. This method has been used for the segmentation of the nuclei of the selected cells images using hybrid genetic algorithm and fuzzy system. A brief outline of the proposed *GOFISA* algorithm is given in the flowchart in Figure (3).

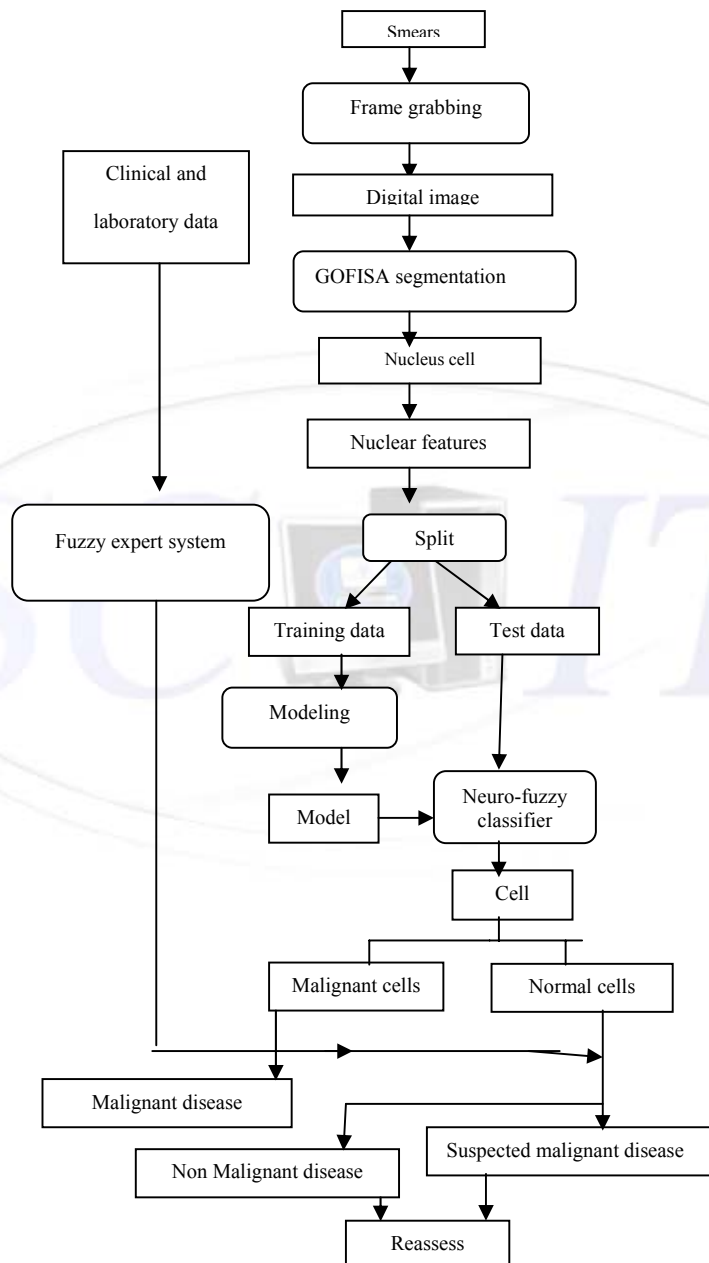


Figure (2): The flowchart of the DHSTCCD algorithm

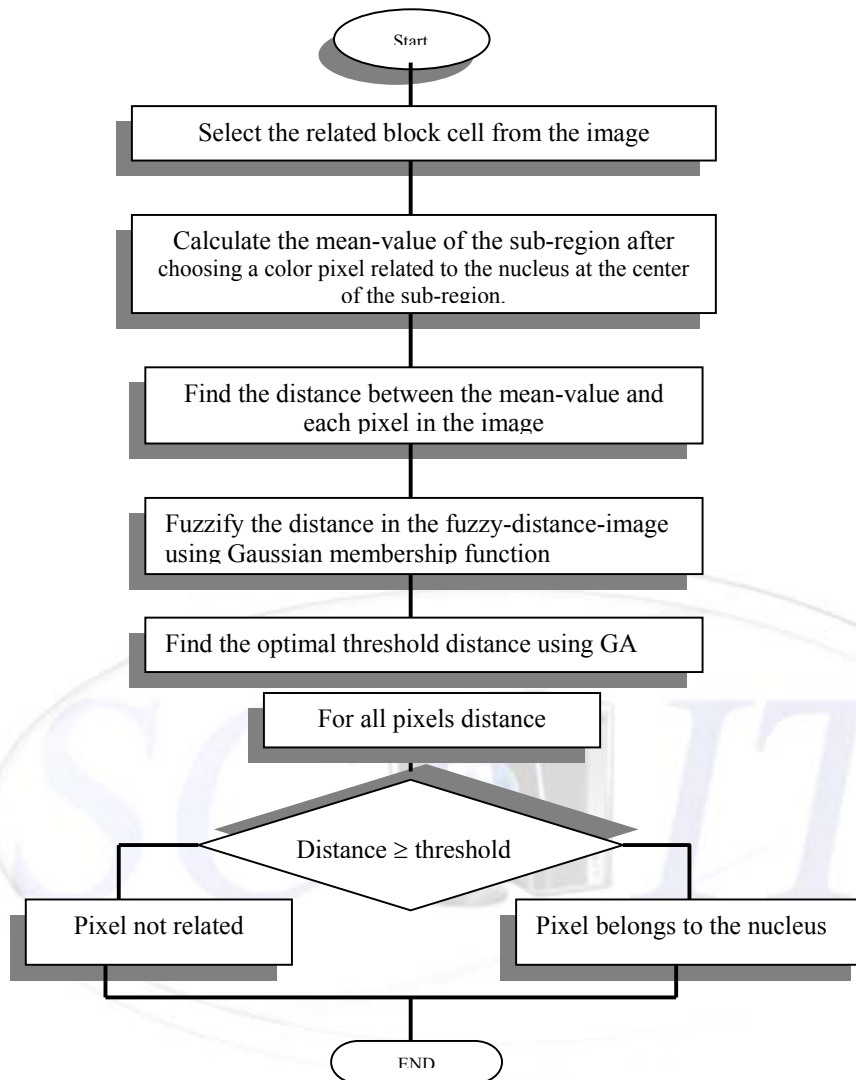


Figure (3): The flowchart of the *GOFISA* algorithm

3.2. Features Extraction

The next step in the computer-assist microscope analysis process of the cell image is to extract descriptive feature measures from the segmented bladder TCC cells nuclei.

The choice of features is closely related to the project that we have worked on. In this work we have focused on the analysis of cell nuclei characteristics to perform classification. There are certain criteria used to identify the cancer cells by visual examination in a microscope. These are quantified in visual examination

and in computerized analysis. Table (1) shows the criteria used for visual examination and the corresponding terms for computerized image analysis.

Several features are extracted from individual segmented cell nucleus. These are morphometric and photometric features.

Table (1): the criteria used for visual examination and the corresponding terms for computerized image analysis.

Visual examination	Computerized analysis
Enlargement of nucleus	Area of nucleus, measured as the number of pixels.
Hyperchromatic nucleus	Integrated optical density of the nucleus, all pixel values of the nucleus are used.
Irregularity of the nuclear border	Shape of the nucleus

- ◀ **Morphometric (shape) features:** Features that are based solely on the spatial arrangements of pixels (Area, Perimeter, Thinness and compactness ratio, Longest and shortest diameters)
- ◀ **Photometric features:** Features that only describe the intensity (gray-level or color) values (Integrated intensity, Mean object intensity).

3.3. Neuro-Fuzzy Classifier Model (NFCM)

Modeling is the subject of constructing models. In fuzzy, neural or hybrid modeling, the model building is generally based on a set of input-output data or sequence of input output data, and the objective is obtaining a system that minimizes the distance between its output and the corresponding data output. Fuzzy-Neural systems are neural networks that topologically or conceptually are structured as a rule based system with **IF-THEN** clauses.

We present a hybrid neuro-fuzzy technique for the adaptive learning of Takaqi-Sugeno type fuzzy inference systems. An optimization technique to determine the number of fuzzy rules and to find the antecedent and consequent parameters is proposed for an adaptive network.

3.3.1. Architecture of the Neuro-Fuzzy Classifier (NFCM)

In this work an adaptive network (NFCM) is a three-layer feed forward network in which can performs Fuzzy IF-THEN rules for pattern classification.

The format of fuzzy **IF-THEN** rules used in the developed classifier is a zero-order Takagi-Sugeno fuzzy model, the rule can be expressed as

$$R_{(k)}: \text{IF } (x_1 \text{ is } A_1^{(k)}) \text{ and } \dots \text{ and } (x_n \text{ is } A_n^{(k)}) \text{ THEN } (y_1 \text{ is } C_{(k)1}) \text{ and } \dots \text{ and } (y_m \text{ is } C_{(k)m})$$

We consider a Multi-Input Multi-Output (MIMO) fuzzy model. The architecture of this Neuro-Fuzzy inference network (NFCM) is shown in Figure (4), a square node (Adaptive node) has parameters while a circle node has none.

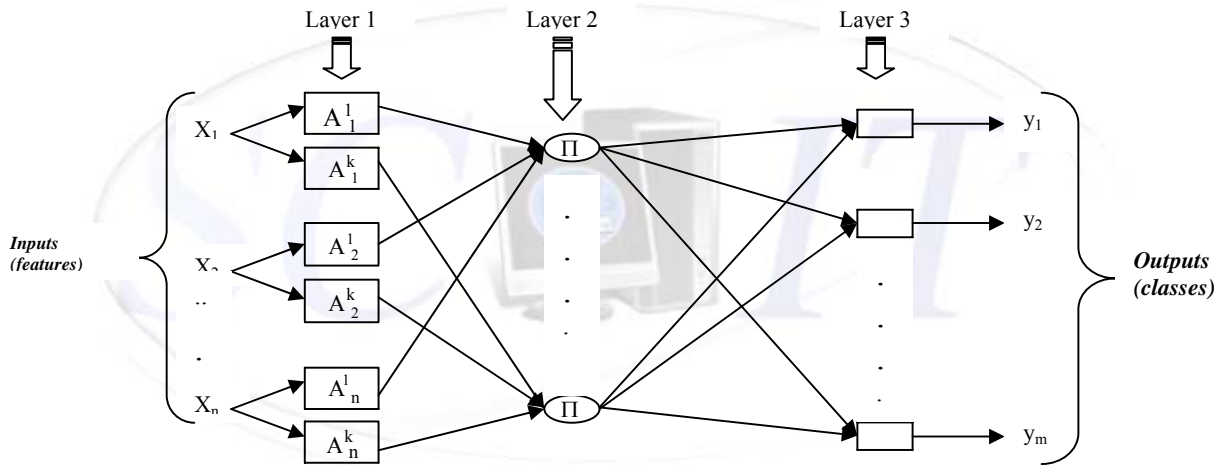


Figure (4): NFCM architecture

3.3.2. Hybrid Learning Strategy

Two phases of learning, structure and parameter learning are used for the neuro-fuzzy (NFCM) network. In the first phase, the network is self-organizing to determine the number of rules and initial parameters of each rule. In the second phase all parameters are adjusted using a supervised learning scheme.

4. A Fuzzy Expert Decision System Design

Expert systems for medical diagnosis support are highly dependent on the quality of data. The data is collected during anamnesis and generally contains expressions that might be considered fuzzy, making it difficult to model them with

conventional computational methods. In this context, fuzzy set theory is an interesting tool to deal with the representation of inaccurate medical entities.

4.1. Purpose: In this work, it is not quite possible to carry out a diagnosis of Transitional Bladder cancer fully based on only image processing. We have developed a rule-based fuzzy expert system that uses the laboratory and other data and simulates an expert-doctor's behavior. Using this data and help from an expert-doctor, the fuzzy rules to determine the patient possibility factor of having Bladder cancer was developed.

4.2. Material and Methods: The clinics and laboratory data for the developed system were taken from collected data. For the design process Haematuria (Haem), Age (Age) and smoking (Smk) are used as input parameters and Bladder cancer risk (Rsk) is used as output. For the inference mechanism the Mamdani max-min inference was used.

4.3. Fuzzy Expert System Structure: Parts of the developed fuzzy rules are shown in Table (2). Total of 54 rules are formed.

Table (2): Fuzzy rules

Rule No.	Age	Smk.	Haem.	Rsk.
Rule 1	Young	Light	Normal	Vlow
...				
Rule 25	Old	Light	2-plus	Middle
...				
Rule 40	Middle	Light	4-plus	High
...				

For example, Rule 40 can be interpreted as follows:

Rule 40: if Age = Middle and Smk = Light and Haem = 4-plus then Rsk = High, i.e. if the patient has middle age and patient's Smk is light and patient's Haem is 4-plus, then patient's Rsk. is High.

The developed fuzzy expert system has a structure as shown in the Figure (5).

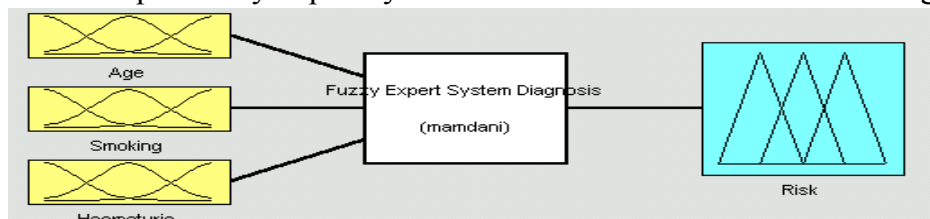


Figure (5): The structure of the fuzzy expert system

4.4. Structure of the fuzzy parameters: The membership of the used parameters are shown in the Figures (6:(1-4)).

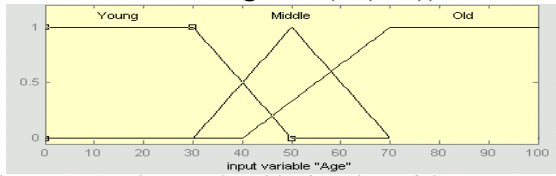


Figure (6.1): The membership function of the Age

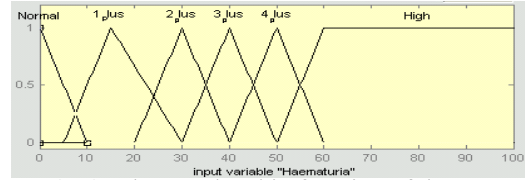


Figure (6.2): The membership function of the Haem.

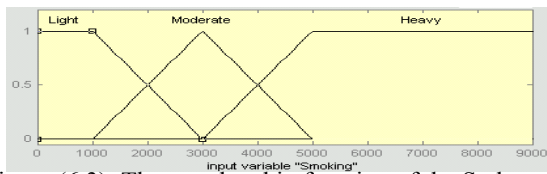


Figure (6.3): The membership function of the Smk.

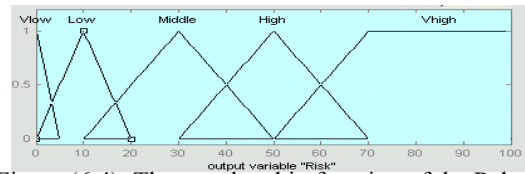


Figure (6.4): The membership function of the Rsk.

4.5. Defuzzification: Truth degrees of the rules are determined for the each rule by aid of the *min* and then by taking *max* between working rules, then we can calculate the crisp output value of the Rsk by center of gravity defuzzifier method.

5. System Implementation and Results

5.1. Implementation of NFCM for Bladder Cells Classification

NFCM is used to classify cells from urinary bladder. The classification is based on cell characteristics derived from images of cells from urine specimens of patients.

The data consists of a set of 100 cells captured from 80 urine examination slides. These cells were captured from normal and abnormal slides see figure (7.(1-2)) using a Twingle digital Web camera connected to the microscope and PC. The high resolution of images is done by using 1000x magnification on the microscope and a web camera resolution of 640×480 pixels in RGB 24-bit colors.

The data is divided into a training set and a test set of equal sizes. The training data is used for building a model for the classification of cells. The test data is used for measuring the performance of the model. The cells can be

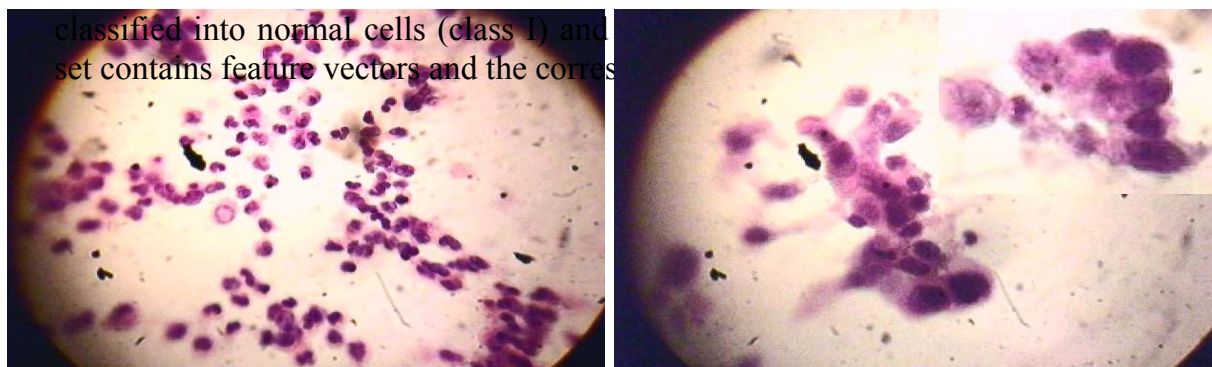


Figure (7.1): Normal slide contains normal cells Figure (7.2): Abnormal slide contains malignant cells

A neuro-fuzzy network with three inputs and two outputs, corresponding to the two classes was considered. The first layer contains six neurons representing membership degrees of each neuron of entry and they act as fuzzy sets representing the terms of the corresponding input variable. In this problem each input variable is connected with two neurons of the first layer, and the six fuzzy sets representing the terms of input variables are shown in Figure (8).

The second layer (layer of rules) is represented by a collection of rules evaluated by a function of aggregation, which is an estimation of truth rate of each rule. The initial number of eight rules was considered for the self-organizing phase. The output layer contains two neurons with values that represent the degree to which the input feature vector belongs to class i .

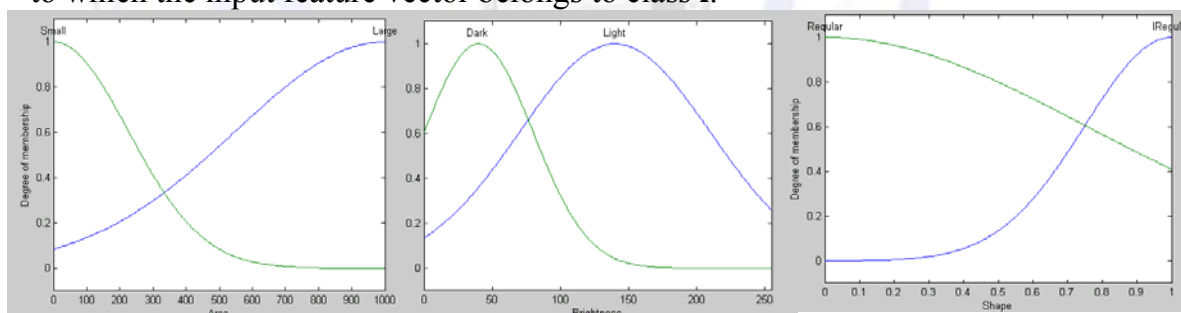


Figure (8): Fuzzy sets of input variables

After the self-organizing learning phase, the network topology was established with a number of four nodes in the second layer, corresponding to the following four fuzzy rules.

1. *If (area is small) and (intensity is light) and (shape is regular) then (class is normal)*
2. *If (area is large) and (intensity is dark) and (shape is irregular) then (class is malignant)*
3. *If (area is large) and (intensity is dark) and (shape is regular) then (class is malignant)*
4. *If (area is small) and (intensity is light) and (shape is irregular) then (class is normal)*

The initial parameters of the network were determined. These parameters were then tuned by the supervised learning phase, which was stopped when 100%

classification rate was achieved or a maximum number of 1000 epochs. The network training error after 100 epochs is shown in Figure (9).

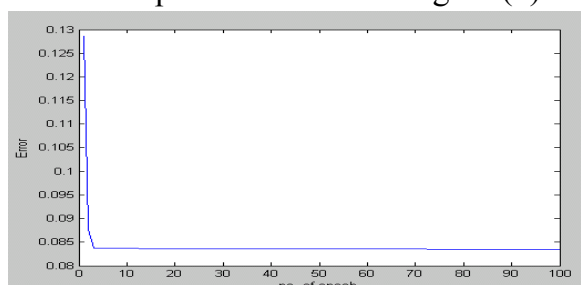


Figure (9): The NFCM Training error

5.2. User Interface

The user interface is an important part of DHSTCCD. It is the task of the interface to handle all the communications between the user and the system. The user interface is organized as in Figure (10).

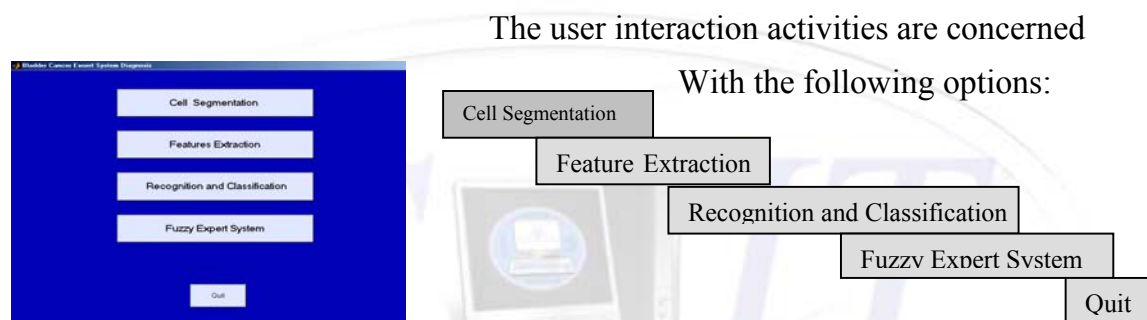


Figure (10): DHSTCCD user interface system

If the user selects segmentation option, DHSTCCD will perform the operation as shown in Figure (11).

If features extraction option is selected, the DHSTCCD will display for each segmented cell the measurement of the following features as shown in Figure (12).

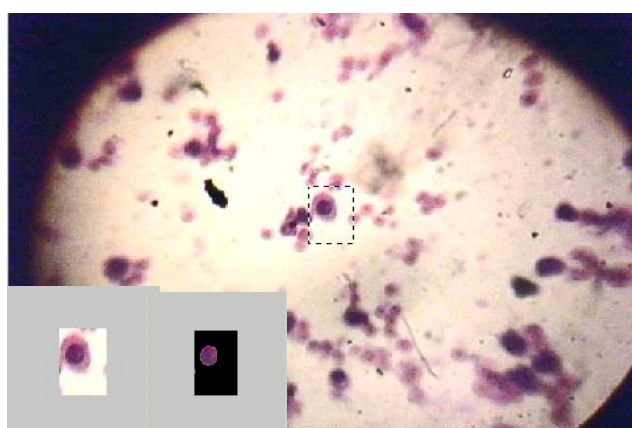


Figure (11): Segmentation process result

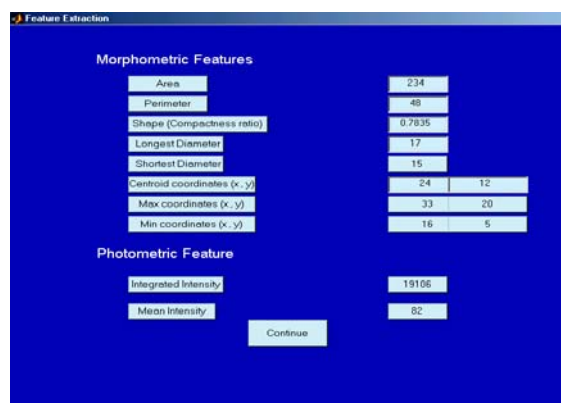


Figure (12): Cell measurements

If the classification and recognition option is selected, the type of segmented cell (normal or abnormal/ malignancy) is determined using NFCM. The final diagnosis is achieved after the selection of the fuzzy expert system option. For example, for Hem = 57, Age = 75, Smk = 2800 the rules 44 and 53 will be fired and we will obtain:

$$\begin{aligned} \text{Rul44} &= \min(\text{old Age, moderate Smk, 4_plus Hem}) = \min(1, 0.8, 0.4) = 0.4 \\ \text{Rul53} &= \min(\text{old Age, moderate Smk, high Hem}) = \min(1, 0.8, 0.6) = 0.6 \end{aligned}$$

From Mamdani *max-min* inference we will obtain the membership degree of our system as $\max(\text{Rul44, Rul53}) = 0.6$, that means high Rsk. Then we can calculate the crisp output of the Rsk center of gravity defuzzifier method.

As seen from Figure (13), the value of Rsk = 76.2. This means that the patient has a probability of having Bladder cancer of 76.2 %, because this is a quite high percentage, doctor has to decide more sophisticated investigations.

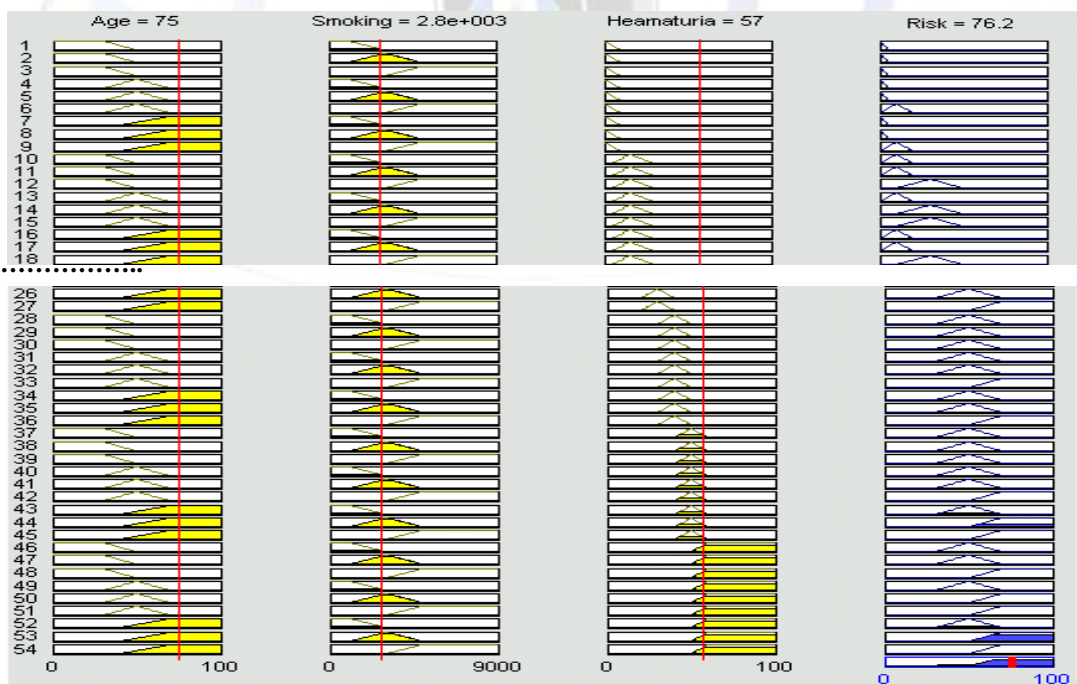


Figure (13): Calculation of the value Rsk

The patient's final diagnosis is obtained by determining the cell type using the Neuro-Fuzzy classifier model (NFCM) and the fuzzy expert system result. If the cell type is malignant, then the patient's diagnosis will be malignant disease even if the result of the fuzzy expert system shows low risk of bladder cancer.

The suspicion of malignant disease would be thought, if the cell type were normal and the fuzzy expert system shows high probability of having bladder cancer, reassessment of the patient is advised. The patient's diagnosis would be non-malignant disease whenever the cell type is normal and the fuzzy expert system shows low risk of having bladder cancer, other pathologies should be sought for.

6. Measuring Performance

To test the performance of a cell classification, six values will be used: False Negative rate (FN%), False Positive rate (FP%), sensitivity, specificity, Positive predictive value (PP%) and Negative predictive value (NP%). The performance is tested on a database of 100 cells. The database contains features of cells. Cells are classified as normal/ negative (N) or malignant/ positive (P). Table (3) shows the system performance of the NFCM.

Table (3): Performance of the NFCM

Sensitivity	Specificity	FN%	FP%	PP%	NP%
88%	90%	12%	10%	89%	88%

7. Conclusions

In this research we have proposed a system for computerized diagnosis of Transitional bladder cell carcinoma, based on fuzzy-neural networks and fuzzy expert system. The proposed system was implemented using Matlab V6.5.

Several variables observable in microscopic images are amenable to morphometric features analysis. The most widely applied are size measurement of individual cells, shape, etc. Computerized image analysis allows accurate and objective evaluation of nuclear morphology showing, for example, that the increase in nuclear size is detected more frequently in carcinomas than in non-

malignant cells. Shape (Compactness ratio) measurement in this study is based on morphometric analysis using size and perimeter data. Increasing abnormalities and irregularities of nuclear features has been reported in Transitional Bladder cell carcinoma.

In this work an NFCM with a two-phase hybrid learning scheme was developed to improve the automatic classification of TCC and could be considered an efficient and robust classification system able to generalize in making decisions about complex input data, improving significantly the diagnostic accuracy. The main features and advantages of the developed network are:

- It is a general framework that combines two technologies, namely neural networks and fuzzy systems.
- The inside of the neural network can be explained in concept of a fuzzy model and hence it can be easily understood.
- The network encodes in its structure the essential design parameters to assemble a fuzzy model, and processes data according to a fuzzy reasoning mechanism.
- The network determines both structure and parameters of the fuzzy rule base via learning from data.

The NFCM exhibited a high performance in correctly classifying Transitional bladder cell under examination into two categories utilizing all the available diagnostic information carried by nuclear descriptors.

- The results of the analysis carried out on 80 smears showed a sensitivity exceeding 88% in the malignancy group and a specificity of 90% in the normal group, and positive predictive value of 89% .
- From the collected data the following medical statics of the DHSTCCD are reached as shown in tables (8.1) and (8.2).

Table (8.1): Patients Symptoms

Risk factors		
Age	50-59	30%
	60-69	27.5%
	70-80	42.5%
Average Age	65	
Gender	Male	75%
	Female	25%
Smoking	75%	
Chronic inflammation	50%	
Vesicle stones	18.75%	
radiotherapy	12.5%	

Table (8.2): Patients Risk factors

Symptoms	
Haematuria	81.25%
Frequency	62.5%
Dysuria	58.75%
Retention of urine	37.5%
Pain	25%

- It seems that the method could be a helpful diagnostic tool for the diagnosis of Transitional bladder cell carcinoma considering the fact that it would be a quick, less invasive and easy test, assisting physician diagnosis.

The developed methodology gives an advantage to be part of an automated computer-aided diagnosis system.

References

- [1] Andrew Peter Bradley, (1996), "Machine Learning for Medical Diagnostics: Techniques for Feature Extraction, Classification, and Evaluation", Ph.D. Thesis, University of Queensland.
- [2] Oscar Fontenla-Romero, Bertha Guijarro-Berdinas, Amparo Alonso-Betanzos, (2000), "Symbolic, Neural and Neuro-Fuzzy Approaches for Pattern Recognition in Cardiotocograms", ESTT 2000.
- [3] Rector Magnificus, (2001), "Adaptive Methods of Image Processing", Ph.D. dissertation, aan de Technische Universiteit Delft, ISBN 90-75691-8.
- [4] Edward Wilson, (2002), "An Overview of Intelligent Systems Technologies for Process Optimization", Myrtle Beach, SC. Symposium sponsored by the Iron and Steel Society's Computer Applications Committee.
- [5] Bob Waterbury, (2002), "Artificial Intelligence Expands Frontiers in Asset Management", Technical report: ISU CS-TR 95-01.
- [6] Computer Design's: Electronic Systems Technology & Design, (1995), "Fuzzy, neural and genetic methods train to overcome complexity", Vol. 34 Issue 5, pp. 59-66.
- [7] James J. Buckley, Esfandiar Eslami, (2003), "Advances in Soft computing", A Springer-Verlag Company, Physica-Verlag Heidelberg.
- [8] Spyros G. Tzafestas, Konstantinos D. Blekas, (1999), "Hybrid Soft Computing Systems: A Critical Survey with Engineering Applications", in S.G. Tzafestas, editor, Soft Computing in System and Control Technology, Singapore: Word Scientific, pp. 119-168.
- [9] Michael J. Droller, Steven E. Canfield, (2002), "Where we stand with tumor markers for urothelial malignancy", J. Urology, Vol. 4, pp. 63-82.
- [10] Olivia Aranha, David P. Wood, Fazul H. Sarkar, (2000), "Ciprofloxacin Mediated Cell Growth Inhibition, S/G2_M Cell Cycle Arrest, and Apoptosis in a Human Transitional Cell Carcinoma of the Bladder Cell Line", Clinical Cancer Research, Vol. 6, pp. 891-900.
- [11] Michael S. Cookson, Jeffrey M. Holzbeierlein, (2002), "Superficial Bladder Cancer: An update for practicing urologists", Contemporary Urology, Vol. 3, pp. 62-78.
- [12] Martin Burchardt, Tatjana Burchardt, Ahmad Shabsigh, (2000), "Current Concepts in Biomarker Technology for Bladder Cancers", Clinical Chemistry, pp. 595-605.