

Pasichnyk S. M., Shatnyi S. V., Pasichnyk M. S., Gozhenko A. I. Application of neural network information technology for recognition and classification of image presentations of renal cell carcinoma in chronic kidney disease to choose the optimal method of treatment. *Journal of Education, Health and Sport*. 2020;10(10): 279-293. eISSN 2391-8306. DOI <http://dx.doi.org/10.12775/JEHS.2020.10.10.027>
<https://apcz.umk.pl/czasopisma/index.php/JEHS/article/view/JEHS.2020.10.10.027>
<https://zenodo.org/record/4469459>

The journal has had 5 points in Ministry of Science and Higher Education parametric evaluation. § 8. 2) and § 12. 1. 2) 22.02.2019.

© The Authors 2020;

This article is published with open access at Licensee Open Journal Systems of Nicolaus Copernicus University in Torun, Poland
Open Access. This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author (s) and source are credited. This is an open access article licensed under the terms of the Creative Commons Attribution Non commercial license Share alike.
(<http://creativecommons.org/licenses/by-nc-sa/4.0/>) which permits unrestricted, non commercial use, distribution and reproduction in any medium, provided the work is properly cited.

The authors declare that there is no conflict of interests regarding the publication of this paper.

Received: 07.09.2020. Revised: 18.09.2020. Accepted: 09.10.2020.

APPLICATION OF NEURAL NETWORK INFORMATION TECHNOLOGY FOR RECOGNITION AND CLASSIFICATION OF IMAGE PRESENTATIONS OF RENAL CELL CARCINOMA IN CHRONIC KIDNEY DISEASE TO CHOOSE THE OPTIMAL METHOD OF TREATMENT

S. M. Pasichnyk¹, S. V. Shatnyi², M. S. Pasichnyk¹, A. I. Gozhenko³

¹Danylo Halytsky Lviv National Medical University, Lviv 79010, Ukraine

²National University of Water and Environmental Engineering, Rivne 33000, Ukraine

³Ukrainian Research Institute of Transport Medicine, Odessa 65039, Ukraine

Abstract

The information technology of recognition and classification of imaging representations of RCC complicated CKD with use of a neural network is offered. Approaches to architecture design, teaching methods, data preparation for training, training and neural network testing are described. The structural-functional scheme of the neural network is developed, which consists of the input, hidden and output layer, each individual neuron is described by the corresponding activation function with the selected weights. The expediency of using the number of neurons, their type and architecture for the task of recognition and classification of image representations of oncological phenomena of the organism is shown. Data of patients with RCC of complicated CKD, research department of reconstructive and plastic oncurology of NIR, urological department of "Lviv regional hospital", urology department of Lviv urological regional medical - diagnostic center, were used as initial data, on the basis of real observations, a database for

training and education of the neural network was formed. An analysis of the efficiency, speed and accuracy of the neural network, in particular, a computer simulation using modern software and mathematical modeling of computational processes in the middle of the neural network. Software has been developed for preliminary preparation and processing of input data, further training and education of the neural network and directly the process of recognition and classification. According to the obtained results, the developed model and structure of the neural network, its software tools show high efficiency both at the stage of preliminary data processing and in general at the stage of classification and selection of target areas of diseases. The next stage of research is the development and integration of software and hardware system based on parallel and partially parallel computer technology, which will significantly speed up computational operations, achieve the learning and training of the neural network in real time and without loss of accuracy. The presented scientific and practical results have a high potential for integration into existing information and analytical systems, medical analysis the choice of tactics for the treatment of patients with RCC complicated CKD, and health monitoring systems in the preoperative and postoperative periods.

Keywords: renal cell carcinoma; chronic kidney diseases; learning method; testing tool; search process convergence; neuron; activation function; weighting factor.

ЗАСТОСУВАННЯ НЕЙРОМЕРЕЖЕВОЇ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ТА КЛАСИФІКАЦІЇ ОБРАЗНИХ ПРЕДСТАВЛЕНЬ НИРКОВО-КЛІТИННОГО РАКУ ПРИ ХРОНІЧНІЙ ХВОРОБІ НИРОК ДЛЯ ВИБОРУ ОПТИМАЛЬНОГО МЕТОДУ ЛІКУВАННЯ

Пасічник С. М.¹, Шатний С. В.², Пасічник М. С.¹, Гоженко А. І.³

¹Львівський національний медичний університет, ім. Данила Галицького, м. Львів

²Національний університет водного господарства та природокористування, м. Рівне

³ДП Український науково – дослідний інститут медицини транспорту МОЗ України,
м. Одеса

Запропонована інформаційна технологія розпізнавання та класифікації образних представлень нирково - клітинного раку (НКР) ускладненого хронічною хворобою нирок (ХХН) із використанням нейронної мережі. В роботі, описано підходи до проектування архітектури, методів навчання, підготовки даних для проведення навчання, тренування та тестування нейронної мережі. Розроблено структурно-функціональну схему нейронної мережі, яка складається із вхідного, прихованих та вихідного шару, кожен окремий нейрон описаний відповідною активаційною функцією із підібраними ваговими коефіцієнтами хворих НКР ускладненого ХХН. Показано доцільність застосування кількості нейронів, їх тип та архітектуру для проведення задачі розпізнавання та класифікації образних представлень онкологічних явищ організму. В якості вихідних даних використані набори даних хворих НКР ускладненого ХХН, науково – дослідного відділення реконструктивної та пластичної онкоурології НІР, урологічного відділу НКП ЛОР «Львівської обласної лікарні», відділення урології Львівського урологічного регіонального лікувально – діагностичного центр. На основі реальних спостережень, сформована база даних для навчання та тренування нейронної мережі. Проведено аналіз ефективності, швидкодії та точності роботи нейронної мережі, зокрема, проведено комп'ютерну симуляцію із використанням сучасного програмного забезпечення та проведено математичне моделювання обчислювальних процесів в середині структури нейронної мережі. Розроблено програмні засоби для попередньої підготовки та обробки вхідних даних,

подальшого тренування та навчання нейронної мережі та безпосередньо процесу розпізнавання та класифікації. Відповідно до отриманих результатів, розроблена модель та структура нейромережі, її програмні засоби реалізації показують високу ефективність як на етапі попередньої обробки даних, так і в цілому на етапі класифікації та виділення цільових ділянок захворювань. В подальшому наступним етапом досліджень є розробка та інтеграція програмно-апаратної системи на основі розпаралелених та частково-розпаралелених засобів обчислювальної техніки, що дозволить значно пришвидшити обчислювальні операції, досягти виконання процесів навчання та тренування нейронної мережі в режимі реального часу та без втрати точності. Представлені наукові та практичні результати мають високий потенціал для інтеграції в існуючі інформаційно-аналітичні системи, системи медичного аналізу, вибору тактики лікування хворих НКР ускладненого ХХН, моніторингу за станом здоров'я, в доопераційному та післяопераційному періодах.

Ключові слова: нирково – клітинний рак; хронічна хвороба нирок; метод навчання; засіб тестування; збіжність пошукового процесу; нейрон; активаційна функція; ваговий коефіцієнт.

Introduction

Kidney cancer is a malignant tumor that is most often represented by carcinoma and develops either from the epithelium of the proximal tubules and collecting tubules. (renal cell carcinoma (RCC)) Or from the epithelium of the cup-pelvic system (transient - cell cancer) [1, 2]. Renal cell carcinoma is the most common type of malignancy that is localized in the kidney. According to the literature, in adults NKR occurs in 80 - 90% of cases among all malignant kidney diseases [3, 4].

In a large number of clinical cases, several factors operate simultaneously, between which there are quite complex interactions, so an important aspect of the evaluation of certain prognostic and diagnostic data is the creation of a method of information processing. The development of technical means and information technology leads to the emergence of new ways of data processing in order to make effective decisions. In particular, in medicine (bioinformatics, biomedicine, etc.) one of the technologies is accurate forecasting (precision prognosis). This technology allows you to divide a large area into several smaller areas of analysis. These zones

can be divided according to different classification features: size, location, qualitative and quantitative indicators.

Construction of information technology for recognition and classification of imaging representations of renal cell carcinoma using a neural network is an important task. The introduction into medical practice of the latest methods of treatment of patients with oncological pathology makes it possible to choose several alternative treatment regimens. Every doctor asks himself a practical question - which treatment regimen is the most effective in the first stage of therapy, so the problem of optimization and individualization of treatment methods for cancer patients remains relevant. Nowadays, in medical practice, the most common models of correlation-regression prediction analysis have been replaced by more modern and effective neural network prognostic models [5, 6]. A neural network is a branch of artificial intelligence in which phenomena similar to those that occur in the neurons of living beings are used to process signals. One of the most important features of neural networks, which demonstrate its wide capabilities and great potential, is the parallel processing of information.

The composition of artificial neural networks (ANN) includes such elements, the functionality of which is similar to the vast majority of elementary functions of the neuron of a living being. One of the characteristic features of ANN is a large number of properties that are characteristic of the brain of a biological being.

ANNs are able to learn independently on the basis of their own experience, ie generalize previous precedents and interpret for new cases, highlighting the essential features of the information received, which contains and contains redundant information [7].

The principles of ANN are approximation; classification and recognition of certain images; forecasting the course and possible consequences; identification and evaluation, the possibility of modeling complex nonlinear relations of various parameters, fast and simultaneous processing of all data [9, 10]. The elementary structural cell of ANN is a neuron. To solve certain tasks with the help of ANN, first of all, it is necessary to accumulate a sufficient amount of data, and then try to teach the neural network to solve such problems.

In recent decades, many scientific papers have been published on the very effective use of ANN in medical research [11, 12, 13]. The most common work of Tailor A. And co-authors published in 1998. The essence of the work was that on the basis of the created ANN, a prognostic test of oncological pathology in women before surgery was proposed. The parameters

of sensitivity and specificity of this test, unfortunately, were not too high and were - 100 and 98%, respectively [14, 15].

Attention is drawn to the work which demonstrates the created model of classification of cervical cancer risk with the help of ANN (Xiaoping Q, Ning T., 2007). The quality of the ANN model was quite high (sensitivity was 98%, specificity about 97%). The proposed model of risk classification has been successfully used for screening groups of high-risk patients [16].

For example, using ANN in gynecological practice, prognostic models for predicting endometrial hyperplasia have been developed, taking into account the history of patients, clinical data and the results of molecular biological methods. [17].

Quite interestingly, the study was conducted using ANN in predicting the long-term results of treatment of patients with prostate cancer. The method demonstrated high accuracy (prediction error did not exceed 4%) and the ability to optimize the quality of prediction in the subsequent training of the network with increasing number of treated patients. However, these studies do not reflect the solution . Therefore, the recognition and classification of imaging representations of renal cell carcinoma using a neural network is an urgent and important task.

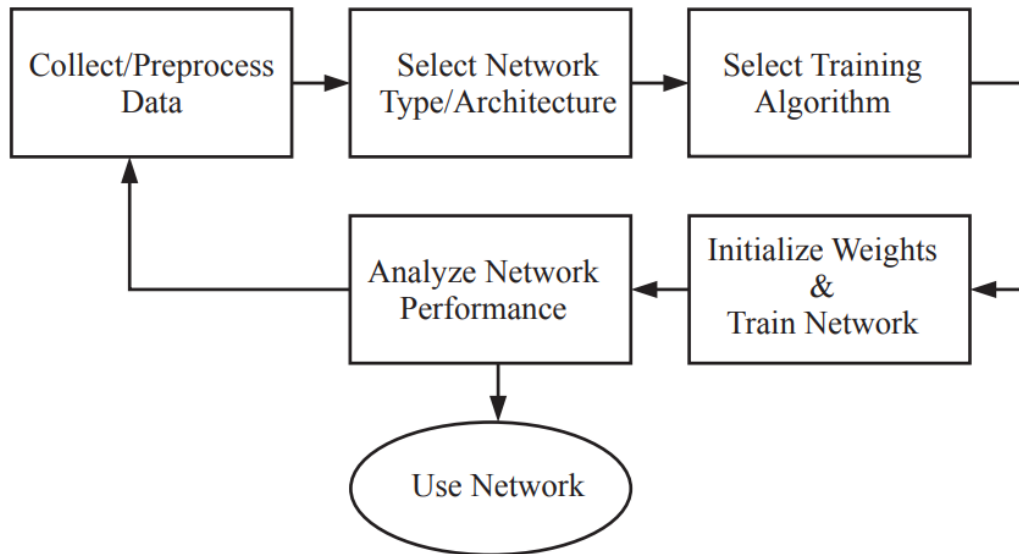
The object of research in this paper are the processes of processing and analysis of image representations for further recognition and classification and determination of the coefficient of belonging to a particular class.

The subject of research is methods, models, structural and functional schemes of information technology of imaging representations of cellular cancer of a kidney with use of a neural network.

The aim of the work is to develop information technology for recognition and classification of imaging representations of renal cell carcinoma complicated by chronic kidney disease using a neural network.

Data pre-processing

This section describes the practical aspects of data preparation and pre-processing, neural network training and testing. Theoretical and practical approaches are not mutually exclusive, so the main approach is to combine the fundamental principles of neural network technologies with practical application. Picture 1 shows the process of neural network operation.



Picture 1 Structural and functional scheme of the neural network functioning process

As can be seen from the picture learning is an iterative process that begins with data accumulation and pre-processing for a better training process. At this stage, the data should be divided into training, reference and test sets. After selecting the appropriate data set, you must select the required network type (multilayer, dynamic, etc.) and architecture (number of layers, number of neurons). Then we choose a training algorithm that is suitable for the neural network and the given problem.

After training the network, it is necessary to analyze the productivity of the network. This analysis can show all existing problems with the data, network architecture or learning algorithm. This process is iterative until the productivity of the developed network reaches the desired and predetermined level.

The process of preliminary training

At this stage, the tasks are grouped into three categories: 1. Data selection; 2. Pre-processing of data; 3. Selecting of network type and architecture. In general, it is very difficult to transfer the accumulated data and rules directly to the neural network, but the quality of the developed neural network will be as high as the quality of data prepared for its training. Training data should cover the full range of input space for which the neural network will be used.

It is impossible to guarantee the performance of the neural network when the data inputs go beyond the training set. Because the proposed neural network for classification is a nonlinear

"black box", it is not capable of qualitative extrapolation of results. It is also not certain that the input data space corresponds to the training set. The proposed design of the neural network to solve a simple problem for which the input vector is small and each element of the input vector can be selected independently of indexing, allows you to operate the input space as a grid.

For many tasks, including the classification of cartographic images, the dimension of the input space is large, which makes the process of abstract data representation impossible and complicated. In this case, all data of the input sets are dependent. The selected area represents the range $[-1..1]$. In this case, the work of the network comes down to finding the function of describing the selected data area. Searching for a function to determine data from outside the selected area is not performing because it is not part of the classification task, it increases the complexity of the network and slows down the overall operation time of the network. There are cases of determining the selected area with insufficient accuracy, the error of reverse spread of the network learning process increases, thereby reducing the accuracy of the classification process and the exit of the search process from the desired area of the data set.

In this case, it is necessary to make pre-processing of the input data set using traditional calculation methods. By analyzing the neural network being studied, it is possible to draw a conclusion about the adequacy of the learning process and, accordingly, the results obtained in the future. Additionally, methods are used to identify the exit of the search process outside the selected area of the data sets on which the learning process took place. This approach does not increase network productivity, but it will make impossible to use the network in tasks for which it is unsuitable. After the process of data collection and preparation, all sets are divided into three groups: 1. Training group; 2. Validation group; 3. Test group. At the time of solving the problem of cartographic images classification, it was determined that the training set occupies 70% of the total volume, the validation set 15%, the test set 15%. It is necessary to take notice that the size of each group may differ depending on the task and data types. The used method of dividing the entire volume of data into groups is to select data randomly from the initial set of prepared data. It has been experimentally confirmed that this method guarantees a satisfactory result in relation to the rapidity, but in future to increase the accuracy it is necessary to insert the classification factors. Also, this approach can be used in the analysis of the learning / training process to identify problems after data division.

Separation of features

The method of features separation is used when the dimension of the input vector row is unnecessary and redundant. The main purpose of this method is to reduce the dimension of the input space by determining the properties of each vector and then use these properties as one of the inputs of the neural network. For the problem of classification of a cartographic area, which is curvilinear, it is not necessary to determine individual spatial hollows, at the stage of preliminary data preparation it is enough to determine their number and give a numerical value to the input of the neural network.

Design of neural network architecture

The next step following the learning process is to develop a network architecture. The basic type of neural network architecture is determined by the type of problem we are solving. Immediately after defining the architecture, the necessary condition is to specify the features of the architecture, in particular the number of neurons and layers, the number of inputs and outputs and the type of neural activation functions.

The posed task of cartographic images classification of soil massifs can be represented as a combination of known tasks, in particular image recognition and clustering. In general, the task is to classify the input data according to the target categories. It is known that optimal for solving this problem are multilayer perceptrons with transfer functions as the basis for source layer of the network. After the analysis and data preparation, it was found to use sigmoid activation functions as the source layer and source neurons. The task of clustering or segmentation leads to group the input data according to their similarity. According to (references) it is advisable to use recurrent and self-organized maps, the main advantage of it is the ability to visualize multidimensional spaces. Figure 2 presents the developed neural network architecture.

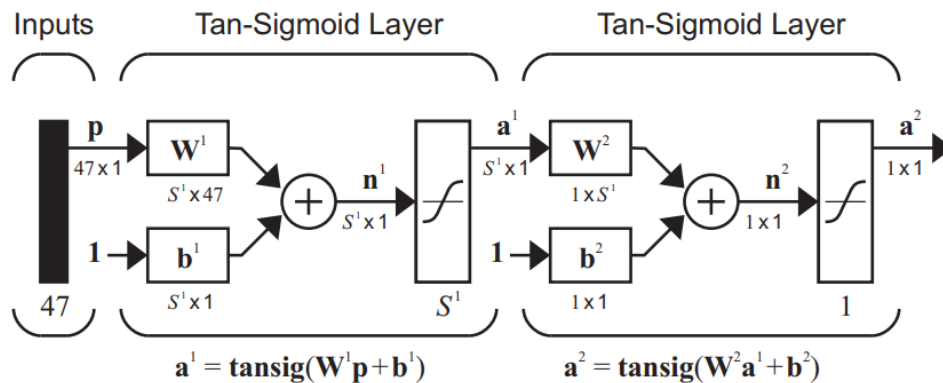


Figure 2. Neural network architecture

At the basis of the network are tangential-sigmoid transfer functions at both levels, which is the standard for a neural network for the problem of classification and pattern recognition. It is possible to use two hidden layers, but it has been experimentally established that one hidden layer is enough to recognize and classify two-dimensional images of soil massifs. The number of neurons in the hidden layer S1 depends on the complexity of the rules of affiliation and is determined at the stage of testing and training of the network.

Computer simulation and analysis of results

Network training was performed using a scalable gradient, which is an effective tool in the task of pattern recognition and classification. Figure 3 shows a graph of change of the average quadratic error in comparison with the number of iterations of the calculating process.

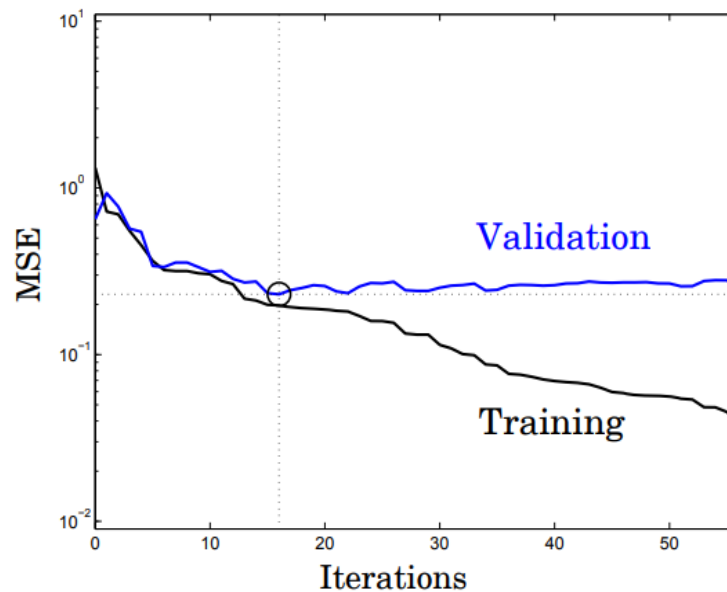


Figure 3. Dependence of the average quadratic error on the number of iterations

It is used the network with 10 neurons in the hidden layer $S1 = 10$. The minimum value of the error was obtained in the sixteenth iterative process, which is shown on Picture 4, and the values of the weight parameters of the neural network were saved at this point. As evident from the presented picture, the learning error did not change during the next forty iterations, so the learning and training process was stopped. The training results are shown on Fig. 4 in the form of coincidences matrix of the trained neural network on the test data.

Confusion Matrix

Output Class	1	13 15.3%	5 5.9%	72.2% 27.8%
	2	1 1.2%	66 77.6%	98.5% 1.5%
		92.9% 7.1%	93.0% 7.0%	92.9% 7.1%
		1	2	Target Class

Figure 4. Matrix of coincidences of the testing process

As seen on a picture the upper left cell shows 13 of the 14 correctly classified soil clusters, cell 2.2 shows 66 of the 71 correctly classified samples. The largest number of misclassified samples is displayed in cell 1.2. Additionally, it was performed the analysis of the recognition quality by the obtained operational characteristic (ROC).

Fig. 5 shows a ROC graph for test data. The ideal trajectory of the ROC path passes through cells 0.0 - 0.1 - 1.1

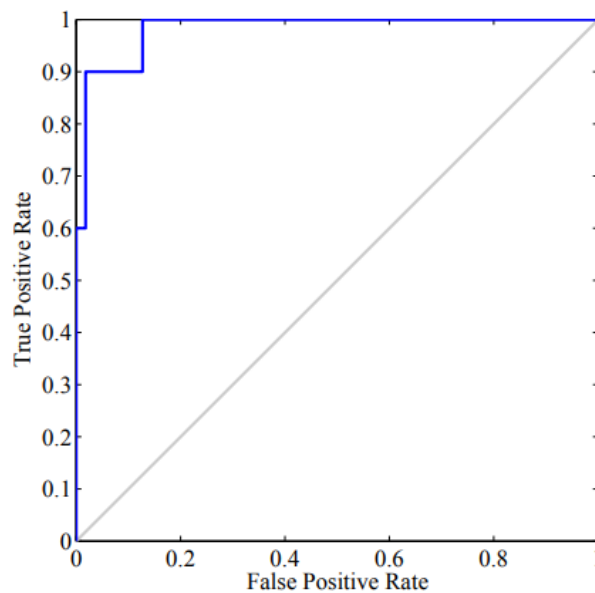


Figure 5. ROC characteristic for test data collection

The results presented in Fig. 5 and Fig. 6 show the distribution of data on training / testing / verification sets. The next stage of the analysis is to determine the sensitivity of the results by conducting a simulation by the Monte Carlo method. The data were divided into 1000 different time intervals, which corresponds to the dynamic change of the data sets of the simulation object. For each data distribution, the neural network was trained with arbitrary weight coefficients. The simulation results were averaged and shown on Fig. 6

Output Class	1	9.11 14.0%	3.69 5.7%	71.2% 28.8%
	2	2.51 3.9%	49.83 76.5%	95.2% 4.8%
		78.4% 21.6%	93.1% 6.9%	90.5% 9.5%
		1	2	
		Target Class		

Figure 6. Averaged values of the Monte Carlo method simulation process

The results reflect the facts of the data sets classification on 12 soil classes for each individual data set. Based on the results, more than 9 classes were determined correctly, with the average quadratic error not exceeding 9.5%.

The averaged results of the Monte Carlo method simulation are similar to the original results, the average value of the error is 9.5%, the average value of the deviation is 3.5%.

Conclusions

The topical issues of development and application of artificial neural networks in medical diagnostic systems, concerning the choice of one or another method of treatment of patients with NKR complicated by CKD are considered in the work. The use of such methods and algorithms allows to increase the efficiency and accuracy of data processing of NKR patients with complicated CKD and their further processing, in order to effectively and efficiently select a

method of treatment of this group of patients. The structural-functional scheme and architecture of the artificial neural network are developed, the optimal methods of training and data preparation for the problem of recognition of the affected areas are selected.

Computer simulations of modes of testing and training of an artificial neural network are carried out, boundary conditions of functioning and the optimum mode of training on speed and duration are defined.

The practical significance of the obtained results is as follows: the use of neural network rationing methods for data pre-processing makes it possible to reduce the total data processing time by more than 20%; the use of the method of classification of cartographic images using a neural network allows to increase the accuracy of determining areas with different classification features up to 73%; the developed information technology makes it possible to increase the speed and level of automation of medical data processing, efficiently and quickly stratify patients and offer one or another type of treatment for a particular patient, compared to other similar systems.

References

1. Clinical oncology / Ed. B.P. Matveeva. - M.: ABV-Press, 2011. -- 934 p. — ISBN 978-5-903018-23-9.
2. Moskvina L. V., Andreeva Yu. Yu., Malkov P. G., Frank G. A., Alekseev B. Ya., Kalpinsky A. S., Pryadilova E. V. Clinically significant morphological parameters of renal cell cancer // *Oncology*. - 2013. - No. 4. - P. 34—39
3. Kuthi L, Jenei A, Hajdu A, Németh I, Varga Z, Bajory Z, et al. Prognostic Factors for Renal Cell Carcinoma Subtypes Diagnosed According to the 2016 WHO Renal Tumor Classification: A Study Involving 928 Patients. *Pathol Oncol Res*. 2017 Jul; 23(3):689–98.
4. Lin SY, Linehan JA, Wilson TG, Hoon DSB. Emerging Utility of Urinary Cell-Free Nucleic Acid Biomarkers for Prostate, Bladder, and Renal Cancers. *Eur Urol Focus*. 2017 Apr;3(2-3):265–72.
5. Haykin S. *Neural networks: full course* / S. Haykin. M.: Publishing house "Williams", 2006. - P. 1104.
6. Comparison of artificial neural networks with logistic regression in prediction of gallbladder disease among obese patients / P. L. Liew, Y. C. Lee, Y. C. Lin et al. // *Dig. Liver Dis*. 2007. – Vol. 39, № 4. – P. 356 – 362.

7. Bozhenko VK Multiparametric analysis of laboratory blood parameters for obtaining diagnostic information in experimental and clinical oncology: author. dis. for a job. scientific. degree doct. honey. Sciences: spec. 14.00.14 "Oncology" / V. K. Bozhenko. - Moscow, 2004. — 32 p.
8. Kryzhanivska AE Locally common cervical cancer: tactics, treatment and prognosis: dis. On the gain. Science. degree of Dr. honey. Science: special. 14.01.07 "Oncology" / AE Kryzhanivska. - Ivano-Frankivsk, 2015. — P. 222-242
9. Osovsky, S. Neural networks for information processing / S. Osovsky. Moscow: Finance and Statistics, 2004. — P. 344.
10. Prediction of survival and complications after gastrostomy in an individual by using clinical factors with an artificial neural network system / T. Takayama, K. Takayama, N. Inoue [et al.] // Eur. J. Gastroenterol. Hepatol. – 2009. –Vol. 21. – P. 1279 – 1285.
11. Radensky P. W. Interactive neural-network-assisted screening. An economic assessment / P. W. Radensky, L. J. Manqo // Acta. Cytol. – 1998. – Vol. 42, № 1. – P. 246 – 252.
12. Usefulness of artificial neural network for differential diagnosis of hepatic masses on CT images / K. Mataka, K. Yoshimitsu, S. Kumazawa [et al.] // Acad Radiol. – 2006. – Vol. 13, № 8. – P. 951 – 962.
13. Radzishevskaya EB Results of statistical analysis of follow-up data of patients with cervical cancer / EB Radzishevskaya, L.Ya. Vasiliev, Ya.E. Vikman // Ukrainian Journal of Radiology. – 2010. – № 18. – P. 65 – 70.
14. Identification of women for referral to colposcopy by neural networks: a preliminary study based on LBC and molecular biomarkers / P. Karakitsos, C. Chrelias, A. Pouliakis [et al.] // J. Biomed. Biotechnol. – 2012. – 2012: 303192. doi: 10.1155 / 2012 / 303192.
15. Zhang Y. X. Artificial neural networks based on principal component analysis input selection for clinical pattern recognition analysis / Y. X. Zhang // Talanta. – 2007. – Vol. 73. – P. 68 – 75.
16. Xiaoping Q. Constructing of the risk classification model of cervical cancer by artificial neural network / Q. Xiaoping, T. Ning, T. Yun // An. International Journal. – 2010. – V. 32. – P. 1094 – 1099.

17. Distinction of cervical cancer biopsies by use of infrared microspectroscopy and probabilistic neural networks / A. Podshyvalov, R. K. Sahu, S. Mark [et al.] // Appl. Opt. – 2005. – Vol. 44, № 18. – P. 3725 – 3734.