

**ПРИЛАДИ І СИСТЕМИ БІОМЕДИЧНИХ ТЕХНОЛОГІЙ**

UDC 616.6:004.67

**DETALIZATION OF RECOGNITION ALGORITHMS IN DIAGNOSING PATIENTS AND EVALUATING THEIR EFFECTIVENESS***Shulyak O. P., Mnevets A. V.**National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute",  
Kyiv, Ukraine**E-mail: [shulyak.alex.47@gmail.com](mailto:shulyak.alex.47@gmail.com), [amnevec-ee22@lil.kpi.ua](mailto:amnevec-ee22@lil.kpi.ua)*

*The issues of detailing recognition algorithms in order to increase the validity of their solutions in diagnosing patients are considered using the example of processing nephrology data. The training of algorithms with a teacher is implied. Procedures for detailing complexes of clinical signs and criteria for comparing such complexes in decision-making are proposed. This means dividing these objects into elements, extracting additional information for them from a priori and current data, and taking them into account in algorithms. Research in the work was focused on the development of software tools for detecting and evaluating additional reserves and opportunities for improving the quality of decisions of recognition procedures by extracting additional useful information from a priori and current data and using them in the process of detailing decision-making procedures. On a specific algorithm, various approaches to such detailing and to the study of its effectiveness were analyzed. Such detailing can be built on the basis of using the experience of clinical practice of observation of patients and their diagnosis in the form of training samples of symptom complexes and (or) observed signals in clinical cases with reliably confirmed diagnoses in the relevant databases.*

*Detailing these algorithmic procedures can lead to the emergence of a multi variance of possible solutions for differently detailed algorithms and require the use of additional procedures for generating a generalizing conclusion based on the results of their mutual consultation. The order and results of detailing are demonstrated in the MatLab environment on two modifications of the proposed algorithm. The introduction reveals the relevance and content of the research. Section 1 reveals the composition of a priori patient data in demo examples and the information that is extracted from them at the training stage. Section 2 proposes two modifications of the algorithm to detalize it. Section 3 proposes software procedures for the statistical evaluation of the performance of the detalization of the algorithms under study. Section 4 describes the refinement of algorithms by introducing weights into the decision criterion, taking into account the spread of values of clinical signs. Section 5 demonstrates the detalization of the algorithms taking into account the information content of the features. The conclusions summarize the results of the work. In general, they are positive.*

**Keywords:** *diagnostics feature-complexes; recognition algorithms decisions reliability reserves-use.*

**Introduction**

The development of information technologies of modern medicine, the technical progress of its tools and systems have led not only to the growth of its achievements but also expanded the horizons of the problem [1–5].

The discovery of more and more profound patterns in the structure and functioning of organs and systems of the human body. Achievements in ensuring the technical accessibility of their observation, research, and management have led to a significant increase in the volume and heterogeneity of patient examination data. Also, increase in the weight of information processing in the processes of diagnosis and treatment, or to the growing need for automation of not only the processing of current data on the state of the human body but also the automation of learning

to work with data [6–10], according to the available a priori and current information [1–5].

The accumulation and formalized use of experience in diagnosing and treating patients, the formation and maintenance of various medically specialized databases of clinical cases with confirmed diagnoses, data on signals, and symptom complexes in various diseases and conditions of the body are becoming increasingly valuable [1–5]. Real opportunities in this direction are noticeably ahead of the pace of their implementation, which indicates a significant unrealized potential for improving the quality of medical care for the population [1–5].

From the available databases, it is possible to extract information about the real diagnostic value of various signals and clinical signs and perform the appropriate refinement of decision-making rules,

correction of recognition procedures and processes, and adjustment of the parameters of recognition structures [1–10].

Data arrays for various clinical cases are widely used in machine learning systems for recognizing the states of the human body [1–10]. Important here is the choice of data composition for training [1–5], whatever the training method itself. The efficiency of recognition procedures depends on this choice [2, 3, 5].

It is important to take a more detailed account of the distributions of the occurrence of different values of signs and signals and their combinations in different cases [1–5], which is especially valuable in early diagnosis for identifying trends in the deterioration of patients' condition and choosing appropriate prevention to maintain health [1–5].

The most important criterion for the effectiveness of the refinement of recognition algorithms, their improvement, and the adjustment of recognition structures are the assessments of the resulting validity, the adequacy of their solutions, which are determined during testing using the same databases [1–5, 10–18].

The particular importance of the study of detailing issues can be traced in connection with the need to open, evaluate and implement the reserves that have not yet been identified. To improve the quality of patient diagnostics due to the insufficiently complete definition and consideration of the characteristics of signals and diagnostic features that can be assessed using the available training samples [1–5, 19–21].

The presence of these reserves is indicated by the fact that in such training there are no strict formalized criteria for completeness, exhaustive nature of extracting information from prior and current data on observed systems. The solution of the questions posed here is investigated by analyzing various options. It needs for detailing the recognition procedures on specific examples. Also with the development of the necessary software tools and its testing with an assessment of the achieved effectiveness within the framework of a specific task. This task based on differential diagnosis of patients using the experience of previous clinical practice in the form of a database that have confirmed diagnoses [2].

The work implies the detailing of such objects (Fig. 1).

Detailing (Fig. 2) in the considered plan means the division of the indicated complex objects into elements, the extraction of additional information for them from the available a priori or current data, and the combination of these elements again into modified complex objects with their addition with new information. Such actions are aimed at increasing the validity of the decisions of recognition procedures.

An example of actions to refine algorithms can be the introduction of additional weights to the elements of composite decision-making criteria for reasons of the spread of feature values.

**OBJECTS OF POSSIBLE DETALISATION  
IN THE DEVELOPMENT AND RESEARCH OF  
RECOGNIZING ALGORITHMS**

- complexes of features that are used to describe the state of observed systems or waveforms;
- criteria for comparing complexes of such features in order to determine their proximity (or difference) to each other;
- composite criteria for making decisions about the type of signals or the type of complexes of clinical features.

Figure. 1. Objects of detail in the development and research of recognition algorithms

**GENERAL CONTENT OF THE DETAILING PROCEDURE  
OF COMPLEX OBJECTS  
IN THE DEVELOPMENT OF RECOGNITION ALGORITHMS**

- the division of such complex objects into elements;
- extracting additional information for them from given to a priori or current data;
- combining the original elements again with their additional information in the modified complex objects.

Figure 2. General content of the procedure for detailing complex objects in the development of recognition algorithms

To illustrate the development of detailing issues for demonstration purposes, a specific recognition algorithm and a specific training set were taken [2]. Developments and research are presented in that order. First, the training sample for research is characterized. Then an algorithm is proposed on which questions of detailing are investigated. The following is the order and an example of evaluating the results of algorithm refinement.

Developments and studies are presented in a number of areas of detail, which are marked with the appropriate section headings.

**Training sample and a priori data extracted from the learning process of recognition algorithms**

In an illustrative example in the study of nephrology data by subjects, the condition of the kidneys is diagnosed [2]. Consideration of random differential diagnosis [1–5]

Symptom complexes are composed of the following results [2]: Age (age of the patient, integer number of years), Length (length of the kidney, mm), Width (width of the kidney, mm), Thickness (thickness of the kidney, mm), Thickness (thickness of the parenchymal layer of the kidney), mm), Velocity (average linear velocity of blood flow through the kidney, cm/s). Index (Purcelo resistance index, relative difference in the velocity of blood flow of the kidney in the phase of systole and diastole of the cardiac cycle), Acceleration (acceleration of arterial blood flow in systole, cm/s<sup>2</sup>). The data of this database corresponds to the format of table 1.

Table 1. Format of complexes of clinical indicators of the patient's kidney

Group	Age	Sex	LR	Length	Width	Thick-ness	Thick-par	Speed	Index	Speed-up
1	74	1	1	115	57	49	16	19,1	0,69	85
2	46	0	1	112	68	88	18	2,3	0,584	459
3	21	1	1	144	##	49	6	16,3	0,707	194

Each set of indicators characterizes a clinical case of medical practice. The data in the table are selected for patients with one of the possible diagnoses: a healthy kidney (1), multiple cysts (2), and hydronephrosis (3). Each diagnosis is shown in the Group column. Sex and LR data are not taken into account. The diagnosis of each patient is confirmed by clinical practice [2], so the data are suitable for training recognition algorithms according to the general methodology of supervised learning [1–5]. There were 22 clinical cases with the first diagnosis, 37 with the second, and 15 with the third, for a total of 122 [2].

The main feature of such symptom-complexes is that they are heterogeneous in physical nature, functional-medical-biological set, diagnostic load, units of measurement, high variability, frequency distribution that reveals different values, measurement magnitude, scatter frequency [1–5, 11, 12, 16, 18]. All this was taken into account when constructing a recognition algorithm for his research and was considered an important condition in performing the refinement of recognition procedures. The presence algorithm was built in such a way as to preserve the ability to work in such conditions.

Scientific research of the algorithm for the choice of students Histograms of the occurrence of various values of signs for various diagnoses is formed. An assessment was made of the mathematical expectation of signs for each diagnosis (Fig. 3).

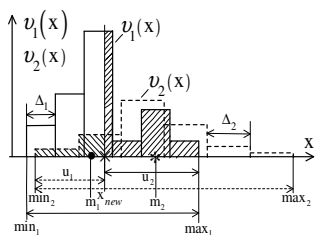


Figure 3. Characteristics of the clinical indicator observed in the evaluation of the recognition algorithm, and evaluation of the assessment of the significance of the sign value for various possible diagnoses.

Here are the characteristics of the same clinical indicator “x” for two different diagnoses, which are conditionally called the first and second. This is reflected in the respective indexes. Each indicator has

its own range of values ( $[min_1, max_1]$ ,  $[min_2, max_2]$ ) for these diagnoses; own histogram of occurrence of values ( $v_1(x), v_2(x)$ ). The second histogram is shown as a dotted line. A dot and an asterisk on the scale of the indicator indicate estimates  $m_1, m_2$  specified mathematical expectations. The number of discretized in histograms on the symptom scale for different diagnoses is the same. Sample sizes  $\Delta_1, \Delta_2$  are different. A separate point shows the value of the indicator for the kidney of a new patient. Her condition needs to be diagnosed using the obtained a priori data.

The results of the analysis of the sample used for the three diagnoses are as follows.

For ranges of indicators, for different diagnoses: Age (21 - 74, 37 - 74, 21 - 68), Length (82 - 135, 98 - 129, 99 - 144), Width (39 - 74, 44 - 89, 41 - 85), Thickness (36 - 72, 37 - 88, 36 - 67), Thickpair (1 - 27, 2 - 11, 6 - 27), Speed (13.3 - 46.6, 2.3 - 43.3, 14.5 - 41.5), Index (0.56 - 0.70, 0.52 - 0.74, 0.69 - 0.80), Speedup (51 - 685, 88 - 656, 98 - 545).

The estimates of mathematical expectations of clinical indicators are as follows: Age (48, 59, 46), Length (112, 111, 118), Width (56, 61, 59), Thickness (48, 50, 52), Thickpair (15, 16, 17), Speed (23, 20, 23), Index (0.64, 0.65, 0.72), Speedup (283, 243, 292).

Histograms of the occurrence of the values of the Speed, Index, Speedup indicators for three diagnoses are shown in fig. 4.

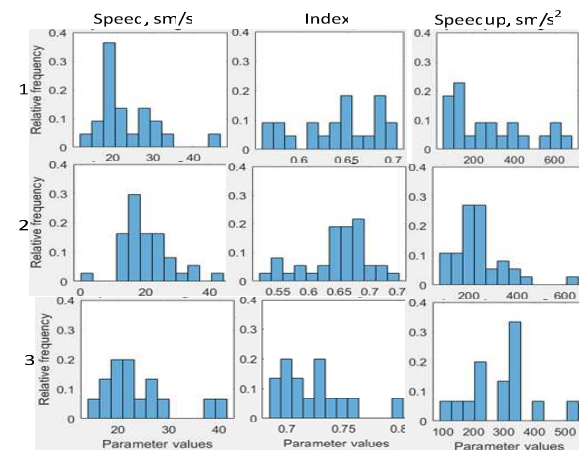


Figure 4. Operations of learning recognition algorithms based on a priori data to ensure their work with symptom complexes in patients.

Histograms of all indicators have 12 intervals on their scales. The results of the sample analysis are used in the work of all algorithms in working out the questions of their detailization.

**Recognizing algorithm for working out on the issues of detailing and its two main modifications**

An algorithm for demonstrating its detail was formed to process clinical signs, taking into account their heterogeneity. The characteristics obtained at the

training stage were used including real histograms of feature distributions.

The decision-making of the algorithm is based on the use of special numerical indicators. An algorithm for demonstrating its detail was formed to process clinical signs, taking into account their heterogeneity. The characteristics obtained at the training stage were used including real histograms of feature distributions.

The decision-making of the algorithm is based on the use of a special numerical indicator  $\mu$ , the evaluation of values or other sign, which is determined from possible diagnoses. The advantage is the diagnosis, for which the value of the severity.

The calculation of this indicator can be seen in Fig. 3. The ideal situation in which the value  $x_{new}$ , the feature witch exactly falls on the estimate of the feature expectation for one of the histograms (on  $m_1$  or  $m_2$ ). Then the algorithm can give preference to the diagnosis for which this histogram was built. In such a situation, the frequencies of the feature values  $x$  falling into the segments to the left and to the right of the mathematical expectation are the same and, according to the normalization condition, are equal to  $1/2$ .

If the sign  $x$  "does not reach" its value to the mathematical expectation, then the sum of the frequencies of the histogram from its edge to the point  $x_{new}$  will be less than  $1/2$ . By its value, it will show the degree of closeness of the current situation to the ideal one, which is implemented in the algorithm.

If the number of histogram samples is fractional, then the frequency for an incomplete sample is taken into account partially in the proportion into which this point divides the sample. The following ratios are used in the algorithm ( $\mu_{max} = \mu(x_{new}=m(x)) = 1/2$ ,  $\mu_r(x_{new}) = \mu(x_{new})/\mu_{max} = 2\mu(x_{new})$ ,  $\mu_r(x_{new}) \in [0;1]$ ), where  $x_{new}$  – received values of the diagnostic sign of the patient.

Each such correspondence is evaluated for each feature, regardless of what the scores for competing diagnoses turn out to be. Separate work with features is replaced by conformity assessment for their complexes. The decision is made in favor of the diagnosis, with maximum compliance. The block diagram of the recognition algorithm is shown in Fig. 5.

The algorithm is built in two modifications. They differ in the order of decision-making (Fig. 6).

The second modification of the algorithm is a refinement of the first modification. It can be expected that a more detailed consideration of the feature values in it will lead to an increase in the validity of the solutions, which was verified statistically on the training sample. The test procedure was as follows.

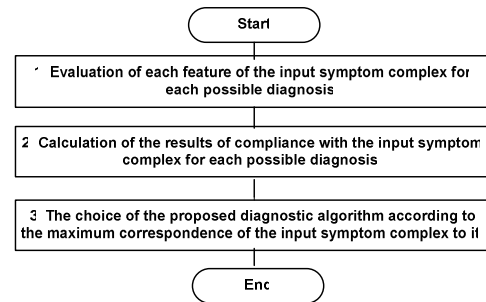


Figure 5. Recognizing algorithm for studying the issues of its detailing.

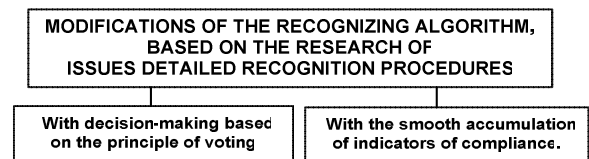


Figure 6. Recognition algorithm modifications.

**The procedure for studying the effectiveness of the performed refinement of algorithms**

The work implements a unified approach to such a study. All algorithms are tested on the same samples that were taken in training. It turns out a comparative assessment of the validity of solutions before and after detailing. We are talking about confirming the trend of its increase due to this detailing. In fact, it introduces additional information into the decision-making process, which is extracted from a priori and (or) current information about the features.

Statistical evaluation of the quality of solutions of algorithms [1 – 5] is done taking into account known diagnoses. When testing algorithms, the calculation of correct and erroneous solutions with errors of various types is carried out [1 – 5]. The results are presented in the form of decision tables [1, 5], which are convenient for calculating the sensitivity (ch), specificity (sp), and overall correctness (validity) val of decisions [1, 5]. The structure of the tables for the three diagnoses is as follows (Table 2).

Table 2. Format of algorithm decision tables

Algorithm solution	case type			sp
	1	2	3	
1	$n_{11}$	$n_{21}$	$n_{31}$	sp1
2	$n_{12}$	$n_{22}$	$n_{32}$	sp2
3	$n_{13}$	$n_{23}$	$n_{33}$	sp3
ch	ch1	ch2	ch3	val

The core of the table is a matrix in which the number of correct (diagonally) and erroneous solutions of the

algorithm of various types (in its remaining cells) was recorded. The first index shows the actual state of the kidney (type of clinical case Group). The second index tells about the type of solutions of the algorithm. The calculation of the sensitivity of the algorithm decisions, the specificity of diagnoses, and the overall validity of the decisions was carried out according to the following formulas

$$ch1 = n_{11}/(n_{11} + n_{12} + n_{13}), ch2 = n_{22}/(n_{21} + n_{22} + n_{23}),$$

$$ch3 = n_{33}/(n_{31} + n_{32} + n_{33}), sp1 = n_{11}/(n_{11} + n_{21} + n_{31}),$$

$$sp2 = n_{22}/(n_{12} + n_{22} + n_{32}), sp3 = n_{33}/(n_{13} + n_{23} + n_{33}),$$

$$val = \frac{n_{11} + n_{22} + n_{33}}{n_{11} + n_{22} + n_{33} + n_{12} + n_{13} + n_{21} + n_{23} + n_{31} + n_{32}}.$$

Sensitivity and specificity are given for specific kidney conditions or types of diagnoses and do not characterize the quality of the algorithm decisions as a whole. Overall validity (correctness) in all cases is important. It is chosen as a criterion for comparing the quality of solutions of algorithms. The calculated validity indicators of the initial modifications of the algorithm are shown in Table 3.

Table 3. Validity of solutions to modifications of the original algorithm

Algorithm solution	Case type			sp
	1	2	3	
<b>Modification with voting (1)</b>				
1	15	14	3	0.47
2	5	14	1	0.70
3	2	9	11	0.50
ch	0.68	0.38	0.73	0.54
<b>Modification with accumulation (2)</b>				
1	16	7	3	0.62
2	3	19	1	0.83
3	3	11	11	0.44
ch	0.73	0.51	0.73	0.62

As expected, the second version of the algorithm gives better solutions, as required. The coincidence of contingents of patients with the same diagnoses was also checked. The coincidence of the solutions is only partial. Council procedures are needed to summarize findings.

**Detailing of algorithms taking into account the spread of values of indicators by introducing weights**

Different indicators have different limits of the scatter of values, its different character and different overlapping of point dislocations for different diagnoses. To form a solution, the algorithm with the accumulation of compliance indicators summarizes the realism indicators of all features:  $\mu_{r\Sigma} = \sum_{j=1}^M \mu_r(x_{newj})$ , where  $\mu_{r\Sigma}$  – the resulted realism of the diagnosis for symptom complex  $M$  of features,  $j$  – number of feature,  $\mu_r(x_{newj})$  – Realism of the diagnosis based on

$j$  feature The diagnosis with the highest value of such a sum wins.

Different signs have different dispersion and it varies depending on the diagnosis. Therefore, it is appropriate to introduce weights into the terms when taking into account the contribution of each feature. The calculation formulas looks like:  $\mu_{r\Sigma} = \sum_{j=1}^M w_j \mu(x_{newj})$  where – specified weights that take into account the dispersion of values. It was requested:  $\sum_{j=1}^M w_j = 1$ . This «1» was distributed in proportion to the accuracy of the spread of signs for each diagnosis.

The accuracy of the dispersion of features was expressed by the relations, where are their root-mean-square deviations from their mathematical expectations in the sample. The sum of the species was reduced to unity. The weights were: Each diagnosis has its own weighting system. A modification of the algorithm with voting is constructed similarly. For each "for" its own weight is used. The decision rule is the same. This is the result of detailing the algorithm. Both versions of the algorithm were studied on a full sample. The feature weights for diagnoses 1, 2, 3 were as follows: Age (0.1975, 0.2344, 0.1684), Length (0.1975, 0.2344, 0.1684), Width (0.1441, 0.1684), 1160, 0.0875), Thickness (0.1154, 0.1031, 0.1075), Thiekpar (0.0735, 0.1036, 0.0489), Speed (0.0643, 0.0598, 0.0574 ), Index(0.3073, 0.2322, 0.4199), Speedup(0.0304, 0.0499, 0.0493). The validity of the solutions is presented in Table. 4.

Table 4. Validity of decisions of algorithms with voting and accumulation with weights

Algorithm modifications	Diagnosis	ch	sp	val
Modification with voting	1	0.682	0.555	0.703
	2	0.757	0.800	
	3	0.600	0.750	
Modification with accumulation	1	0.772	0.607	0.730
	2	0.595	0.880	
	3	1.000	0.714	

In comparison with the initial values, the validity of the solutions increased markedly. The detailing is effective.

**Detailing algorithms taking into account the information content of features**

Such detailing consists in the fact that for each diagnosis a part of the signs is selected, in which the validity of the decisions becomes higher. Feature selection is done during training. Informativity is estimated using the entire sample. The composition of complexes for decision-making is becoming shorter.

The information content of signs is understood as the nature and level of their influence on the validity of decisions, and it can be both positive and negative and different for different diagnoses [11, 16, 17].

Informativity is estimated by discarding signs and registering changes in the validity of decisions in the sample. The changes in the number of errors made in the decisions of the algorithm are counted. The more the validity decreases when a feature is discarded, the more informative it is. The features were ranked by informativity in descending order and selected for use. In the first place, those whose influence was negative were discarded. This was the second, main and more universal way to refine the algorithm in its two versions. An appropriate program procedure has been drawn up.

The first feature selection method was simpler and used the feature of the algorithm, which is that it can make a decision even on one (any) feature. The correct solutions are known. The best among them are those that, when used separately, led to fewer errors. Feature selection was studied for the same modifications of the algorithm on the same statistics.

In the first version of the evaluation of the informativity of features, it is close for two modifications of the algorithm (Fig. 7). The effectiveness of their reduction is shown in Fig. 8. Where method1– algorithm with voting, method2–algorithm with accumulation.

It can be seen that the validity of solutions can become noticeably higher than it was at the beginning, which corresponds to the goal of detalization the algorithm.

Similarly, the possibilities of increasing the validity of solutions in the second method of assessing the information content of features were studied.

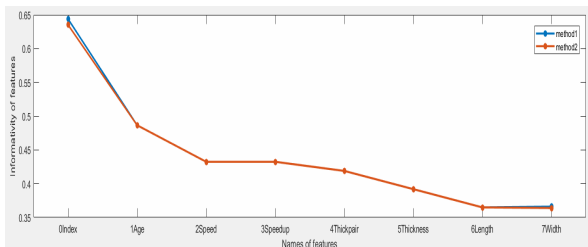


Figure 7. A list of features ranked by information content with its separate assessment

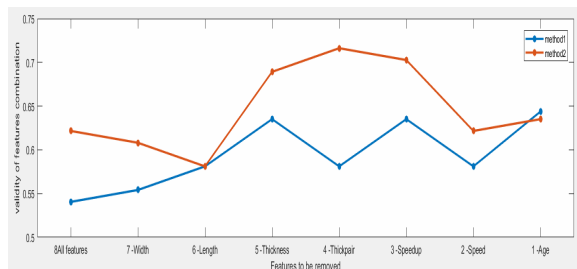


Figure 8. Reducing the composition and changing the validity of the algorithm solutions in the first version of the evaluation of the information content of features

The successive exclusion of features (with a return) to assess their information content led to such results (Fig. 9).

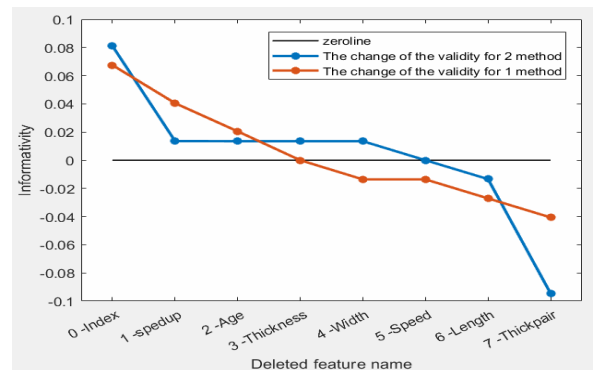


Figure 9. The result of the ranking of features in terms of informativity by their successive exclusion (with a return) for the algorithm with accumulation (2) and voting(1).

Successive exclusion of features from consideration, starting with the worst, led to such results (Fig. 10).

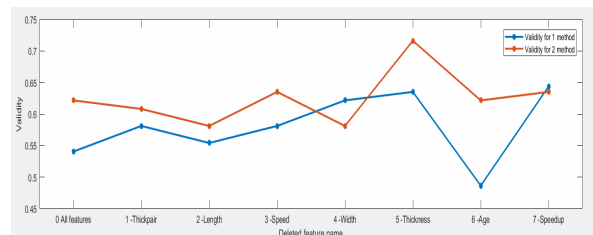


Figure 10. Changing the Validity of Algorithm Decisions in the Second Variant of Feature Selection

Detailing the modifications of the algorithm is effective. The validity of the conclusion was confirmed by the cross-validation method [10]. Four variants of random exclusion of 50% of clinical cases from the sample were taken. The spread and average value of validity for the full and reduced composition of features are in Table. 5.

Table 5. Validity of decisions of algorithms with voting and accumulation with weights

Modification with voting					
Variants of random samples (50% of original samples)				Validity and STD	
1	2	3	4	Mean	CKO.
0,321	0,514	0,633	0,563	0,508	0,133
0,536	0,457	0,516	0,581	0,523	0,05
Modification with accumulation					
Variants of random samples (50% of original samples)				Validity and STD	
1	2	3	4	Mean	CKO
0,321	0,429	0,62	0,618	0,496	0,146
0,536	0,486	0,583	0,655	0,565	0,072

A decrease in the spread of validity estimates confirms the correctness of estimates of its values

[10]. The increase in average validity values confirms the improvement in the quality of solutions, which was required.

### Conclusions

Studies of the detailing of recognition algorithms in different areas, carried out in the work, confirmed on specific examples the manifestation of a tendency to increase the validity of the decisions of recognition procedures.

The proposed approach and software tools provide for the identification, evaluation and implementation of reserves for improving the quality of formed solutions due to a more complete extraction and use of additional information at the training stages from a priori and current data on diagnosed objects.

The considered detailing can provide not only the development of software tools for autopsy and the implementation of reserves to increase the validity of diagnostic decisions about the types of clinical cases in patients. It can also be useful in preparing recognizing structures for their machine learning to determine the rational composition of a priori data that will be used for this purpose. Much depends on the initial composition of features on which such learning will be built.

### References

- [1] A. Genkin, *New informational technology of medical data analysis (OMIS software package)*. Sankt-Peterburg: Politekhnik, 1999.
- [2] V. Djuk, V. Jemanujel', *Information technology in biomedical studies*. Sankt-Peterburg: Politekhnik, 2003.
- [3] M. Antomonov, "Mathematical processing and analysis of biomedical data," 2006.
- [4] V. I. Vasil'ev, *Recognition systems. Reference book*. Kiev: Naukova dumka, 1983.
- [5] O. P. Mincer, Y. V. Voronenko, V. V. Vlasov, *Information technology in health and practical medicine*: Educ. textbook. Kyiv: Vyscha shkola, 2003, p.350.
- [6] J. Anderson, R. S. Michalski, J. G. Carbonell, T. M. Mitchell, and S. Amarel, *Machine learning. an artificial intelligence approach*. Los Altos, CA: Morgan Kaufmann, 1983.
- [7] I. H. Witten and E. Frank, *Data mining. practical machine learning tools and techniques*. 2nd ed. Amsterdam: Elsevier, 2005.
- [8] J. Saez-Rodriguez, M. M. Rinschen, J. Floege, and R. Kramann, "Big science and big data in nephrology," *Kidney International*, vol. 95, no. 6, pp. 1326–1337, 2019.
- [9] O. Niel and P. Bastard, "Artificial Intelligence in nephrology: Core concepts, clinical applications, and Perspectives," *American Journal of Kidney Diseases*, vol. 74, no. 6, pp. 803–810, 2019.
- [10] J. M. Anderson, J. Anderson, E. R. Bareiss, and R. S. Michalski, *Machine learning: An artificial intelligence approach*. Palo Alto, CA: Kaufmann, 1986.
- [11] A. P. Shulyak & A. D. Shachykov, "Criteria and Procedures for Estimating the Informativity and Feature Selection in Biomedical Signals for their Recognition", *RADAP*, no. 66, pp. 79-86, 2016. doi:10.20535/RADAP.2016.66.79-86
- [12] A. P. Shulyak, & A. D. Shachykov, "Analysis of the biomedical signals' structure in the problem of recognition", in *2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO)*, 2014. doi:10.1109/elnano.2014.6873982
- [13] A. P. Shulyak, & A. D. Shachykov, "Decomposition of biomedical signals on mutually orthogonal components in the diagnosis of diseases", in *2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO)*, 2014. doi:10.1109/elnano.2014.6873914
- [14] A. P. Shulyak, & A. D. Shachykov, (2015) "Working off the analysis principles of cyclic biomedical signals' structure for their detection, recognition and classification", *Bull. Kyiv Polytech. Inst. Ser. Instrum. Mak.*, is. 49(1), pp. 169-179, 2015. doi: 10.20535/1970.49(1).2015.47101
- [15] A. P. Shulyak, & A. D. Shachykov, "About the impact of informative features selection in the mutually orthogonal decompositions of biomedical signals for their recognition", in *2016 IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO)*, 2016. doi:10.1109/elnano.2016.7493054
- [16] A. D. Shachykov, V. Shelofast, and A. P. Shulyak, "Modification of selection procedures cyclic characteristics of biomedical signals for their recognition", *Bull. Kyiv Polytech. Inst. Ser. Instrum. Mak.*, no. 53(1), pp. 103–109, 2017. doi:10.20535/1970.53(1).2017.106810
- [17] A. D. Shachykov and A. P. Shulyak, "Features of the use of the form characteristics of biomedical signals during their recognition", *Bull. Kyiv Polytech. Inst. Ser. Instrum. Mak.*, no. 51(1), pp. 131–139, 2016. doi: 10.20535/1970.51(1).2016.78217
- [18] H. Porieva, K. Bahdasariants, & T. Saurova, "Detection of respiratory diseases using BISPECTRAL analysis method", in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, 2020. https://doi.org/10.1109/elnano50318.2020.9088838
- [19] I. D. Mandel, *Cluster analysis*. Moscow: Finance and Statistics, 1988. (in Russian).
- [20] A. Kotov and N. Krasilnikov. (2006) Data clustering.
- [21] B. Duran, P. Odell, *Cluster analysis*. Moscow: Statistics, 1977.

УДК 616.6:004.67

**О. П. Шуляк, А. В. Мневць***Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна***ДЕТАЛІЗАЦІЯ РОЗПІЗНАВАЛЬНИХ АЛГОРИТМІВ У ДІАГНОСТИЦІ ПАЦІЄНТІВ ТА ОЦІНКА ЇЇ РЕЗУЛЬТАТИВНОСТІ**

Розглядаються питання деталізації алгоритмів розпізнавання, з метою підвищення валідності їх рішень у діагностиці пацієнтів на прикладі обробки даних нефрології. Мається на увазі навчання алгоритмів із учителем. Пропонуються процедури деталізації комплексів клінічних ознак та критеріїв порівняння таких комплексів у прийнятті рішень. Мається на увазі поділ цих об'єктів на елементи, вилучення для них додаткових відомостей з апріорних та поточних даних та їх врахування у алгоритмах.

Дослідження в роботі були зосереджені на розробці програмного інструментарію виявлення та оцінки додаткових резервів та можливостей підвищення якості рішень розпізнавальних процедур за рахунок вилучення додаткових корисних відомостей з апріорних та поточних даних щодо їх використання у процесі деталізації процедур прийняття рішень. На конкретному алгоритмі були проаналізовані різні підходи до такої деталізації та дослідження її результативності.

Така деталізація може бути побудована на основі використання досвіду клінічної практики спостереження пацієнтів та їх діагностики у вигляді навчальних вибірок симптомокомплексів та (або) сигналів, що спостерігалися в клінічних випадках з достовірно підтвердженими діагнозами у відповідних базах даних.

Деталізація зазначених алгоритмічних процедур може призвести до виникнення багатоваріантності можливих рішень по-різному деталізованих алгоритмів та вимагати застосування додаткових процедур формування узагальнюючого висновку за результатами їхнього консиліуму.

Порядок і результати деталізації демонструються в середовищі MatLab на двох модифікаціях запропонованого алгоритму.

Розкрито склад апріорних даних пацієнтів у демонстраційних прикладах та відомостей на етапі навчання. Запропоновано дві модифікації алгоритму щодо його деталізації та програмні процедури статистичної оцінки результативності деталізації досліджуваних алгоритмів. Описано деталізацію алгоритмів за допомогою введення вагів в критерій прийняття рішень з урахуванням розкиду значень клінічних ознак.

**Ключові слова:** діагностика пацієнтів; комплекси ознак; розпізнавальні алгоритми; валідність рішень; використання резервів.

*Надійшла до редакції  
06 травня 2022 року*

*Рецензовано  
30 травня 2022 року*

УДК 617.3

**КРИТЕРІЇ ОПОРНИХ ХАРАКТЕРИСТИК СТОПИ ЛЮДИНИ***Цапенко В. В., Терещенко М. Ф.**Національний технічний університет України**«Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна**E-mail: [capenko.valik@ukr.net](mailto:capenko.valik@ukr.net), [agfarkpi@i.ua](mailto:agfarkpi@i.ua)*

*У статті приведено результати дослідження параметрів опорних характеристик стопи людини. Біомеханіка стопи завжди розглядається з точки зору її опорної та ресорної функцій. При перенавантаженнях систем, що підтримують скелетиння - порушуються функції стопи, спотворюється, в цілому, руховий стереотип, відбуваються небажані перерозподіли сил, які передаються на інші відділи опорно-рухового апарату, в результаті чого виникають відхилення від норми, а потім фіксуються патологічні зміни. Провідне місце серед захворювань опорно-рухового апарату займають різні деформації стоп. Діагностика постави людини не була б повною без вимірювання і оцінки стану опорно-ресорних властивостей стопи. Для об'єктивної оцінки розподілу навантаження, викликаного вагою тіла людини та його впливу на локомоції, необхідні клінічний аналіз руху та постурологічне обстеження. При аналізі навантажень на стопу, необхідно враховувати ударні навантажен-*