Journal of Pathology Informatics 14 (2023) 100193

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

# Decision-making support system for diagnosis of oncopathologies by histological images

Anatoliy Dovbysh [a,*], Ihor Shelehov [a], Anatolii Romaniuk [b], Roman Moskalenko [b], Taras Savchenko [a]

[a] Department of Computer Science, Sumy State University, 2 Rymskogo-Korsakova Street, Sumy, Sumy Region, Ukraine
[b] Department of Pathology, Sumy State University, 31 Pryvokzalna Street, Sumy, Sumy Region, Ukraine

ARTICLE INFO

*Keywords:*:
Machine learning
Information criterion
Histological image
Breast cancer
Computer-aided detection
Hierarchical information-extreme machine learning

ABSTRACT

The aim of the study is to increase the functional efficiency of machine learning decision support system (DSS) for the diagnosis of oncopathology on the basis of tissue morphology. The method of hierarchical information-extreme machine learning of diagnostic DSS is offered. The method is developed within the framework of the functional approach to modeling of natural intelligence cognitive processes at formation and acceptance of classification decisions. This approach, in contrast to neuronal structures, allows diagnostic DSS to adapt to arbitrary conditions of histological imaging and flexibility in retraining the system by expanding the recognition classes alphabet that characterize different structures of tissue morphology. In addition, the decisive rules built within the geometric approach are practically invariant to the multidimensionality of the diagnostic features space. The developed method allows to create information, algorithmic, and software of the automated workplace of the histologist for diagnosing oncopathologies of different genesis. The machine learning method is implemented on the example of diagnosing breast cancer.

## Introduction

The use of full-slide optical microscopy in histological studies still remains the main highly reliable method for diagnosing oncopathologies of various origins. The main advantage of this method is the ability to obtain a large amount of visual diagnostic information, but the reliability of the diagnosis depends on the training level and the histologist experience. Therefore, the creation of DSS for the diagnosis of oncopathology on histological images is an urgent task, as it allows you to create an automated workplace of a histologist. The main way to solve this problem is to use for the information synthesis of diagnostic DSS ideas and machine learning methods. The article considers the machine learning method, which in contrast to existing Data Mining methods, allows to give diagnostic DSS adaptability properties of decisive rules to arbitrary conditions of histological imaging and flexibility in retraining the system by expanding the recognition classes alphabet that characterize different tissue structures.

## Background

Modeling cognitive processes inherent in a person in the formation and making of classification solutions is one of the most difficult tasks today. To solve this problem, machine learning methods and pattern recognition are widely used.[1,2] The work[3] shows an example of a computerized system for diagnosing breast cancer. In this system, for 2 classes of recognition (benign and malignant tumor), the input data for machine learning and testing were the breast tissue cytological examination results. The diagnostic signs structured vector consisted of only 9 characteristics of breast tissue cells. Since machine learning DSS was carried out by different methods, the greatest reliability for the alphabet of the 2 classes was obtained by implementing the method of reference vectors.[4] But it should be noted that in the case of increasing the power of the recognition classes alphabet with a constant space of diagnostic features, the reference vectors method convergence is not guaranteed. In Beckwith,[5] the input data for machine learning and multilayer convolutional neural network (CNN) testing was a set of full-slide images, which consisted of 400 microscopic images of lymph node tissues. As a result, the best version of machine learning CNN correctly recognized 92.5% of tumors on the test set of slides, and the pathologist only 73.3%. But the main disadvantage of CNN is the sensitivity to the multidimensionality of the diagnostic features space and the recognition classes alphabet. In Moskalenko and Korobov,6 to reduce the impact of multidimensionality, it is proposed to use input extractors built on CNN, but this approach inevitably leads to information loss.

Thus, the use of traditional methods of intellectual data analysis of Data Mining,[7] including multi-convoluble CNN,[8,9] for information synthesis DSS for diagnosing oncopathology by histological images, does not always provide high functional efficiency of machine learning due to a number of scientific and methodological limitations:

* Corresponding author.
*E-mail addresses:* a.dovbysh@cs.sumdu.edu.ua (A. Dovbysh), i.shelehov@cs.sumdu.edu.ua (I. Shelehov), moskalenko@med.sumdu.edu.ua (R. Moskalenko).

http://dx.doi.org/10.1016/j.jpi.2023.100193
Received 20 June 2022; Received in revised form 17 January 2023; Accepted 17 January 2023
Available online 26 January 2023

1. arbitrary conditions for the formation of histological images;
2. significant intersection of recognition classes in the diagnostic features space;
3. the space multidimensionality of diagnostic signs full of slide histological images;
4. influence of uncontrolled perturbing factors.

One of the promising approaches to the information synthesis of diagnostic DSS is the use ideas and methods of so-called information-extreme intelligent technology (IEIT) data analysis, which is based on maximizing the information capacity of the system in machine learning.[10–12] The idea of IEIT methods, as in CNN, is to adapt the input mathematical description in the machine learning to the maximum possible probability of making the correct diagnostic decisions. But the main advantage of information-extreme machine learning methods is that, unlike neuro-like structures, they are developed as part of a functional approach to modeling cognitive processes inherent in man in the formation and adoption of classification decisions. This approach, in contrast to structural methods, allows the information-extreme machine learning methods to provide flexibility in retraining the system by expanding the recognition classes alphabet. In addition, built on the geometric approach of the decisive rules practically solve the problem of many dimensions of the dictionary of recognition features. It should be emphasized that the information-extreme machine learning methods are not interactive training matrix is an order of magnitude smaller than the image samples. In Naumenko,[12] an example of information synthesis within the IEIT diagnostic DSS for 4 recognition classes is given. But since the linear information-extreme machine learning algorithm was implemented, it was not possible to build highly reliable decisive rules.

The purpose of the study is to increase the functional efficiency of DSS for the diagnosis of oncopathology on histological images by information-extreme machine learning on a hierarchical data structure.

## Problem statement

Consider the formalized formulation of the problem of hierarchical information-extreme machine learning DSS for the oncopathology diagnosis by histological images. We will consider the hierarchical structure of data in the form of the so-called decursive binary tree. The data structure in the form of a binary tree will be called decursive, in which the attribute from the top of the upper tier is transferred to the its stratum top of the lower tier. In our case, training matrices of the corresponding recognition classes are considered as vertex attributes. Final executions from which attributes are not transferred will be called final. Thus, as the power of the recognition class alphabet increases, the decursive hierarchical structure is divided into strata, each of which consists of the 2 closest in binary space Hamming features of the recognition classes. This allows for their classification to use a linear algorithm of information-extreme machine learning of the required depth. In this case, in contrast to neuro-like structures, the depth of information-extreme machine learning is determined not by the hidden layers number, but by the machine learning parameters number that are optimized by the information criterion.

Suppose that the recognition classes alphabet is given in the form of a decursive hierarchical structure $\left\{X^{o}_{h,s,m} | h=\overline{1,H}; s=\overline{1,S}; m=\overline{1,2}\right\}$, where H is the hierarchical structure tiers number; S – he number of executions on the h- th tier. The 2 recognition classes of each stratum characterize histological tissue samples with different morphological structures of the breast. For each recognition class, the input training matrix of brightness $\| y^{(j)}_{h,s,m,i} | i=\overline{1,N}, j=\overline{1,n} \|$ is formed by processing histological images, where N is the diagnostic features number in the structured vector of the recognition class $X^{o}_{h,s,m}$; n – the structured vectors number of the recognition class diagnostic features $X^{o}_{h,s,m}$, which form a training matrix.

It is known that the concept of IEIT is to transform the input learning matrix of the type "object-property" into a given in the Hamming binary space working learning matrix, which in machine learning the process through acceptable transformations adapts to the maximum full probability of making the right classification decisions. Suppose that a set of $\{g_m\}$ structured vectors of machine learning parameters that affect the functional efficiency of diagnostic DSS is given in the Hamming space. We will limit ourselves to the depth of machine learning of DSS of the second level, which allows to optimize the geometric parameters of hyperspherical containers and recognition classes and the control tolerances system for diagnostic features. In this case, the machine learning parameters vector for the decisive rules construction for the recognition of diagnostic features vector, for example, class $X^{o}_{h,s,m}$ will be presented in the form of a structure

$$g_{h,s} = <x_{h,s,m}, d_{h,s,m}, \delta_{h,s}> \quad (1)$$

where $x_{h,s,m}$ is a binary averaged vector of diagnostic features, the vertex of which determines the geometric center of the recognition class container $X^{o}_{h,s,m}$; $d_{h,s,m}$ – code distance, which determines the radius of the hyperspherical container of the recognition class $X^{o}_{h,s,m}$; $\delta_{h,s}$ – the decursive tree h-th tier s-th stratum parameter, which is equal to half of the symmetric field of control tolerances for the diagnostic feature.

At the same time, restrictions are set on the machine learning parameters:

1. the range of values of the radius of the container recognition class $X^{o}_{h,s,m}$:

$$d_{h,s,m} \in [0; d(x_{h,s,m} \oplus x_{h,s,c}) - 1]$$

where $d(x_{h,s,m} \oplus x_{h,s,c})$ is the Hamming code distance between the vector of diagnostic features $x_{h,s,m}$ of the recognition class $X^{o}_{h,s,m}$ and a similar vector $x_{h,s,c}$ of the nearest neighboring recognition class $X^{o}_{h,s,m}$;

1. for bilateral symmetric fields of tolerances of a diagnostic sign there is:

$$\delta_{h,s} \in [0; \delta_H/2]$$

where $\delta_H$ is the field of normalized tolerances for diagnostic features, which determines the values range of the corresponding control tolerances.

At the stage of machine learning it is necessary:

1. to optimize the parameters of vector (1) according to the alphabetical average of recognition criteria information criterion, which is calculated at the k-th step of machine learning for the decursive tree h-th tier s-th stratum:

$$\overline{E}^{(k)}_{h,s} = \frac{1}{2}\Sigma^{2}_{m=1} \max_{G_E \cap \{k\}} E^{(k)}_{h,s,m} \quad (2)$$

where $E^{(k)}_{h,s,m}$ is the criterion of information capability of DSS calculated on the k-th step of machine learning to recognize structured vectors of diagnostic signs of class $X^{o}_{h,s,m}$; $G_E$ is the criterion of information capability of DSS calculated on the k-th step of machine learning to recognize structured vectors of diagnostic signs of class; (k) – an ordered set of machine learning DSS steps.

1. According to the optimal geometric parameters of the recognition classes containers obtained in the machine learning process, construct decisive rules for each stratum of the hierarchical structure, which guarantee a high full probability of making the correct diagnostic decisions;
2. at the examination of the stage, it is necessary to make a classification decision on the belonging of the features vector to one of the recognition classes of the corresponding final stratum.

## Materials and methods

Within the IEIT input description of diagnostic DSS consists of:

1. The recognition features dictionary, the power of which is determined by the size of the frame full slide histological image;
2. The recognition classes alphabet, which characterize different morphological structures of breast tissue;


**Fig. 1.** Frame image of interest areas: a – class $X_1^o$; b – class $X_2^o$; c – class $X_3^o$; d – class $X_4^o$.

3. the input training matrix of the frame pixels brightness of the histological image;
4. working binary training matrix, which in the machine learning process adapts to the maximum full probability of making the correct diagnostic decisions.

Histological imaging of the tissues morphology with malignant neoplasms of the breast was obtained by optical microscopy at the Sumy Regional Oncology Center (Sumy, Ukraine). The image was 1728x923 pixels in size and was divided into 1400 frames of 54x54 pixels, from which an alphabet of 4 recognition classes was formed by expert evaluation: class $X_1^o$ – histological imaging of the tissues morphology with malignant neoplasms of the breast was obtained $X_2^o$ – invasive cell growth; class $X_3^o$ – glass nuclei and class $X_4^o$ – connective tissue. Representatives of these recognition classes are shown in Fig. 1.

The input information description of diagnostic DSS will be presented in the structure form:

$$I_B = \ <G, T, \Omega, Z, H, Y^{|S|}, X^{|S|}; g, f_1, f_2>$$

where G is the set of factors that affect the diagnosis; T - the time moments set of reading information from the histological image; Ω - space of diagnostic signs; Z - alphabet of recognition classes; H - hierarchical data structure in the form of a binary hierarchical tree; $Y^{|S|}$ - set of input (Euclidean) pixel brightness training matrices of the histological image frame for all strata of the decursive tree; $X^{|S|}$ - working binary training matrix, transformed into a space of Hamming diagnostic signs; g - a decursive tree construction operator; $f_1$ - the set formation operator of $Y^{|S|}$ the decursive tree strata training matrices; $f_2$ - operator transformation of matrices $Y^{|S|}$ into working binary matrices $X^{|S|}$. Cartesian product of sets $G \times T \times \Omega \times Z$ specifies the information source.

The categorical functional model for the second level of depth of information-extreme machine learning of diagnostic DSS according to the decursive hierarchical data structure is shown in Fig. 2.

The operator g shown in Fig. 2 forms a decursive binary tree H from the information source, and the operator $f_1$ generates input (Euclidean) fuzzy learning matrices $Y^{|S|}$ for all strata of the decursive tree. Operator $f_2$ by comparing the recognition features with the given control tolerances forms for all strata, respectively, a set of $X^{|S|}$ binary working matrices, which in the machine learning process are adapted to the maximum probability of making correct classification decisions. The term set $E_S$, the elements of which are calculated at each step of machine learning values of the information criterion for each stratum, is common to all contours of learning parameters optimization. Operator $r : E \rightarrow \tilde{\Re}_S^{|2|}$ at each step of machine learning restores in the radial basis of the Hemming space of diagnostic features containers of recognition classes, which form for each stratum of the partition $\tilde{\Re}_S^{|2|}$. The operator ξ reflects the partition $\tilde{\Re}_S^{|2|}$ on the distribution of binary vectors of the features of the working matrix $X^{\|S\|}$. Next, the operator $\psi : X^{\|S\|} \rightarrow I^{|2|}$ forms a set $I^{|2|}$ of 2 alternative statistical hypotheses by testing the basic statistical hypothesis $\gamma_1 : x_{h,s,n}^{(j)} \in X_{h,s,m}^o$. The operator γ determines the set $\Im^{\|Q\|}$ of the exact characteristics of diagnostic solutions, where $Q = C^2$, and the operator φ calculates the set E the set E of the information optimization criterion values, which is a functional of the exact characteristics.

The optimization contour of control tolerances is closed through the term set D, the elements of which are the values of control tolerances for diagnostic features. Operator $u_H$ regulates machine learning the process.

Thus, information-extreme machine learning of diagnostic DSS is a purposeful search for the global maximum of the information criterion (2) optimization of geometric parameters of recognition classes containers, which are restored in the radial basis of the Hamming space of diagnostic features.

As the power of the recognition class alphabet increases, the decursive hierarchical structure is divided into final strata, each of which consists of the 2 closest features of the recognition classes in the binary space. This allows for their classification to apply a linear algorithm of information-extreme machine learning of the required depth to build highly reliable decisive rules. According to the categorical model (Fig. 3), information-extreme algorithm of machine learning DSS with optimization of machine learning parameters (1), we present in the form of an iterative procedure for finding the global maximum of the alphabetically averaged recognition of information criterion (2) for the s-th stratum of the $h$-th tier of the decursive tree:

$$\delta_{h,s}^* = \arg\left\{ \max_{G_\delta} \left\{ \max_{G_E \cap (k)} E_{h,s}^{(k)} \right\} \right\} \tag{3}$$

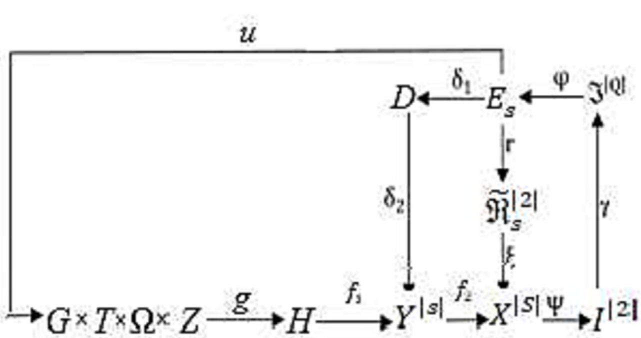


**Fig. 2.** Categorical model of machine learning diagnostic DSS.

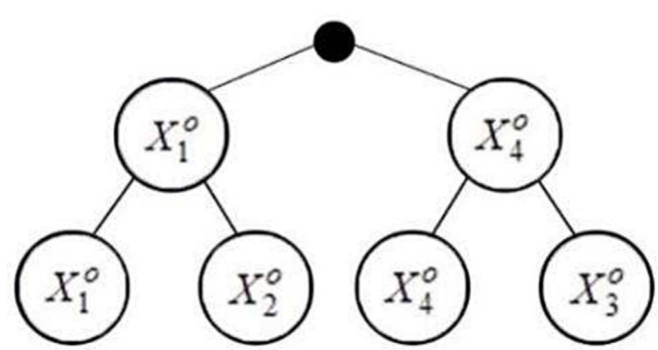


**Fig. 3.** Decursive data structure.

where $\delta^*_{h,s}$ - is the optimal (hereinafter in the information sense) control tolerances field parameter for the recognition classes of the s-th stratum of the h-th tier of the decursive tree; $G_\delta$ – parameter values valid range $\delta_{h,s}$.

The input information for the information-extreme machine learning algorithm is array $\{y^{(j)}_{h,s,m,i}\}$ and a system of fields of normalized tolerances $\{\delta_{H,\ i}\}$ and a system of fields of normalized tolerances sets the values range of the corresponding control tolerances. Consider scheme of the information-extreme machine learning algorithm with optimization of control tolerances for diagnostic features according to the procedure (3):

1. resetting the counter of recognition classes: $m \coloneqq 0$

2.$m : m + 1$

3. resetting the counter for changing parameter $\delta_{h,s} : \delta_{h,s} \coloneqq 0$;

4.$\delta_{h,s} \coloneqq \delta_{h,s} + 1$;

5. the lower $A_{H,i}$ and upper $A_{B,i}$ control tolerances on diagnostic signs according to rules are calculated

$$A_{H,i} = y_{h,s,m,i} - \delta_{h,s}; A_{B,i} = y_{h,s,m,i} + \delta_{h,s}$$

where $y_{h,s,m,i}$ is the average value of the i-th diagnostic feature of the features structured vector $x_{h,s,m}$;

6. resetting of the radius change steps counter of the hyperspherical container: $k \coloneqq 0$;

7.$k \coloneqq k + 1$;

8. a binary training matrix $\{X_{h,s,m,i}\}$, is formed, the elements of which are calculated according to the rule

$$x^{(j)}_{h,s,m,i}[k] = \begin{cases} 1, if\ A_{Hk,i}[k] < y^{(j)}_{h,s,m,i} < A_{Bk,i}[k]; \\ 0, if\ \ else; \end{cases}$$

9. an array formation of averaged binary vectors-realizations $\{x_{h,s,m}\}$, the elements of which are determined by the rule

$$x_{h,s,m,i} = \begin{cases} 1, if \dfrac{1}{n} \sum\limits_{j=1}^{n} x^{(j)}_{h,s,m,i} > \rho_m; \\ 0, if\ \ else; \end{cases}$$

where $\rho_m$ is the selection level of the coordinates of the binary vector $x_{h,\ s,\ m}$, which by default is 0,5.

10. the optimization information criterion of machine learning parameters is calculated;

11. if $k \leq N$, then point 7 is fulfilled, otherwise – point 12;

12. if $\delta < \delta_H$, then point 4, otherwise – point 13;

13. the maximum value of the information criterion in the working area of determining its function is determined;

14. if $m < M - 1$, then point 2 is implemented, otherwise – point 15;

15. the global maximum of the average information criterion $\overline{E}^*_{h,s}$ in the working area of definition of its function is defined;

16. the optimal values of parameter $\delta^*_{h,s}$ and respectively, the lower $A^*_{H,i}$ and upper $A^*_{B,i}$ control tolerances for all diagnostic features are determined.

As a criterion for optimizing the machine learning parameters DSS used a modified information measure Kullback in the form

$$E^{(k)}_{h,s,m} = \frac{1}{n_{min}} log_2 \left\{ \frac{2n_{min} + 10^{-r} - \left[K^{(k)}_{1,h,s,m} + K^{(k)}_{2,h,s,m}\right]}{\left[K^{(k)}_{1,h,s,m} + K^{(k)}_{2,h,s,m}\right] + 10^{-r}} \right\} \quad (4)$$

$$\left[n_{min} - \left(K^{(k)}_{1,h,s,m} + K^{(k)}_{2,h,s,m}\right)\right]$$

where $K^{(k)}_{1,h,s,m}$ is the events number in which the feature vectors of the recognition class $X^o_{h,s,m}$ do not erroneously belong to it; $K^{(k)}_{2,h,s,m}$ is the events number in which erroneously belong to the recognition class $X^o_{h,s,m}$vectors of features of the neighboring stratum recognition class; nmin is the minimum size of a representative training sample; $10^{-r}$ is a small enough number to avoid zero division.

To increase the functional efficiency of diagnosing at the second level of the depth of information-extreme machine learning DSS, an algorithm was implemented to consistently optimize control tolerances. In this case, the control tolerances obtained at the stage of parallel optimization were accepted as the starting point for sequential optimization, which was carried out according to the procedure.

$$\delta^*_{h,s} = arg \bigoplus_{l=1}^{L} \left\{ \max_{G_{\delta_i}} \left[ \frac{1}{2} \sum_{m=1}^{2} \max_{G_{E_m} \cap \{k\}} E^{(k)}_{h,s,m}(l) \right] \right\}, i = \overline{1, N} \quad (5)$$

where $\oplus$ is the symbol of the repeat operation; L - the runs number of the procedure of sequential optimization of control tolerances due to suboptimal starting values of control tolerances for all characteristics; $E^{(k)}_{h,\ s,\ m}(l)$ - calculated on the k-th step of the l-th run of machine learning information criterion.

According to the optimal geometric parameters of the recognition classes containers obtained in the machine learning process, decisive rules were constructed

$$\left(\forall X^0_{h,s,m} \epsilon \Re^{|2|}_{h,s}\right) \left(\forall x^{(j)} \in \Re^{|2|}_{h,s}\right) \left[ if \left( \mu_{h,s,m} = \max_{\{m\}} \{\mu_{h,s,m}\} \right) \right]$$

$$then\ x^j \epsilon X^0_{h,s,m}\ else\ x^j \notin X^0_{h,s,m} \quad (6)$$

where $x^{(j)}$ is a recognizable vector; $\mu_{h,s,m}$ – a membership function of vector $x^{(j)}$ of the recognition class container $X^o_{h,s,m}$.

In expression (6), the membership function for a hyperspherical container of recognition class $X^o_{h,s,m}$ is determined by the formula:

$$\mu_{h,s,m} = 1 - \frac{d\left(x^*_{h,s,m} \oplus x^{(j)}\right)}{d^*_{h,s,m}}$$

where $x^*_{h,s,m}$, $d^*_{h,s,m}$ are the optimal geometric parameters of machine learning: the average binary implementation and the radius of the hyperspherical container, respectively.

To automatically construct a hierarchical structure in the form of a binary decursive tree for the above recognition classes, a variation series was formed by increasing the average brightness of the training matrices of the tissue morphology images shown in Fig. 1. As the strata recognition classes of the first tier, the most remote classes $X^o_1$ and $X^o_4$ were chosen. The lower tier strata were formed on the principle of the nearest neighbor.

Fig. 3 shows the hierarchical data structure for the 4 recognition classes in the form of a binary decursive tree.

To achieve a high full probability of making the right diagnostic decisions, it is necessary to give the diagnostic DSS the property of personalization. That is, DSS should not be diagnosed on the basis of averaged data for a group of different patients, but take into account the individual performance of each patient. For this purpose, the average brightness values of each frame were calculated, which formed a variation series in their growth. A fragment of the variation series from each of the 4 recognition classes is shown in Fig. 4.

Initially, the attributes of recognition classes $X^o_1$ and $X^o_4$, which are included in the first stratum of the hierarchical structure, were determined (Fig. 4). To find them, the variation series, in which the smallest average value of brightness was 60.5 and the largest 248, was divided into 26

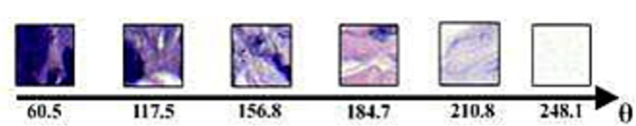


**Fig. 4.** Ordered sequence of frames.

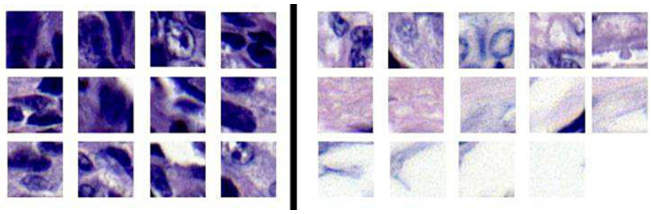

**Fig. 5.** Sampling of tissue and cell morphology.

intervals, each of which had a size of 7 units of brightness. The next step was to select frames whose brightness was approximately average for the corresponding interval. As a result, a sample of the frame brightness average values was formed. Then, the sample was divided into 2 parts in the place of the largest difference between the frames clarity average values (Fig. 5).

The last stage is the search for all possible pairs from different parts of the sample (Fig. 5). The information-extreme machine learning process of DSS with parallel optimization of the control tolerances system for diagnostic signs according to the procedure (3) is performed. The pair of frames, which has the maximum value of the optimization information criterion (5), have the largest interclass distance is placed in the decursive data structure (Fig. 3), respectively, as recognition classes attributes $X_1^o$and $X_4^o$ (Fig. 6).

Next, consider the definition in the hierarchical structure (Fig. 3) of the attributes of the recognition classes vertices $X_2^o$ and $X_3^o$. Since in the decursive tree the attribute of the top of the upper tier is transferred to the corresponding stratum of the lower tier, it is necessary to choose a representative attribute no longer from 1400 frames, but only from frames belonging to these classes. Since the attributes of the strata of the lower tiers are chosen on the principle of the nearest neighbor, the recognition class attribute $X_2^o$ is chosen training frame matrix for the minimum-maximum value of the information criterion optimization, which indicates the closest similarity of this frame with the representative frame vertices of the recognition class $X_1^o$. For the second final stratum, the the top attribute of the recognition class $X_3^o$. In the case of the recognition classes alphabet expanding by a similar procedure, the vertices attributes of all strata of the lower tiers are determined.

Fig. 7 shows the final view of the hierarchical decursive structure (Fig. 3) with the attributes of all recognition classes.

Thus, the formation of a hierarchical data structure in the form of the decursive binary tree allows to divide the recognition classes set into pairs of the nearest neighboring classes, for which optimization is carried out according to the above linear algorithm of information-extreme machine learning. When retraining diagnostic DSS due to the emergence of a new recognition class, a new variational series of the average values brightness of the frames of the extended alphabet is formed and a hierarchical data structure is built according to the considered scheme.

The functional effectiveness verification of information-extreme machine learning was carried out during the operation of DSS in the examination mode according to the geometric decisive rules (6). The algorithm of the exam is similar to the algorithm of DSS operation directly in the diagnostic mode.

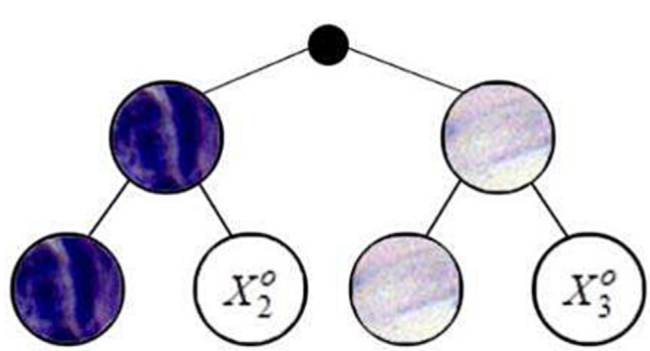


**Fig. 6.** Partially filled hierarchical data structure.

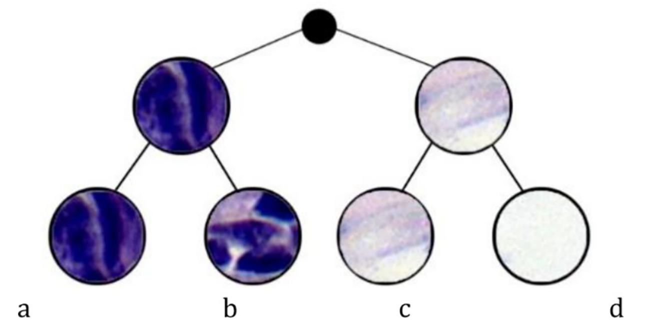


**Fig. 7.** Hierarchical data structure: a – class $X_1^o$; b – class $X_2^o$; c – class $X_4^o$; d – class $X_3^o$.

## Results

The information synthesis of diagnostic DSS was carried out by information-extreme machine learning with optimization of control tolerances for diagnostic features. The input training matrix was formed by screening in the Cartesian coordinate system of histological images (Fig. 2). The brightness values of the RGB components, which were read in each pixel of the frame the receptor field, were considered as signs of recognition. The training matrix for each recognition class consisted of 54 time-structured vectors, each of which had 162 diagnostic features. Information-extreme machine learning was carried out according to the data hierarchical structure in the form of a decursive tree (Fig. 7). For each final execution, the optimal parameters of machine learning were determined (6).

Fig. 8 shows the dependence graphs of the average normalized criterion (4) on the parameter of the control tolerances field obtained for each final stratum of the lower tier when implementing the information-extreme machine learning algorithm by procedure (3) parallel optimization of control tolerances of diagnostic features.

Fig. 8 dark areas indicate the working (permissible) areas for determining the function of the information criterion (4), in which the errors of the first and second kind are less than the first and second values, respectively. Analysis of Fig. 9a shows that in the information-extreme machine learning process on the training matrices of the recognition classes of the first final stratum, the optimal value of the control tolerances field: $\delta_1^* = 27$ (hereinafter in brightness gradations) with the alphabetically criterion $\overline{E}_1 = 0,26$. In Fig. 8b, the maximum values of the information criterion belong to the plateau-type area. In this case, the optimal value is taken as the value of parameter $\delta_2$ with the minimum value of the coefficient $\eta_\delta$, which characterizes the intersection degree of the 2 recognition classes:

$$\eta_s = \frac{\overline{d}_s(\delta)}{d[x_{h,s,1}(\delta) \oplus x_{h,s,2}(\delta)]} \to \max_{\delta} \qquad (7)$$

where $\overline{d}_s(\delta)$ is the average radius of the recognition classes containers $X_{h,s,m}^o$; $d[x_{h,s,1}(\delta) \oplus x_{h,s,2}(\delta)]$ – code center distance of recognition classes $X_{h,s,1}^o$ and $X_{h,s,2}^o$.

Taking into account condition (7) for the second stratum, the optimal value of the parameter of the control tolerances field is equal to $\delta^* = 18$ at the maximum limit value of the normalized information criterion $\overline{E}_2 = 1$.

Since for the first final execution (Fig. 8a), a low maximum value of the information criterion was achieved, in addition, a consistent optimization of the system of control tolerances was performed according to procedure (5). Thus, the control tolerances received at a stage of parallel optimization were accepted as starting at consecutive optimization. Fig. 9 shows the changes graph in the average normalized information criterion (5) in the

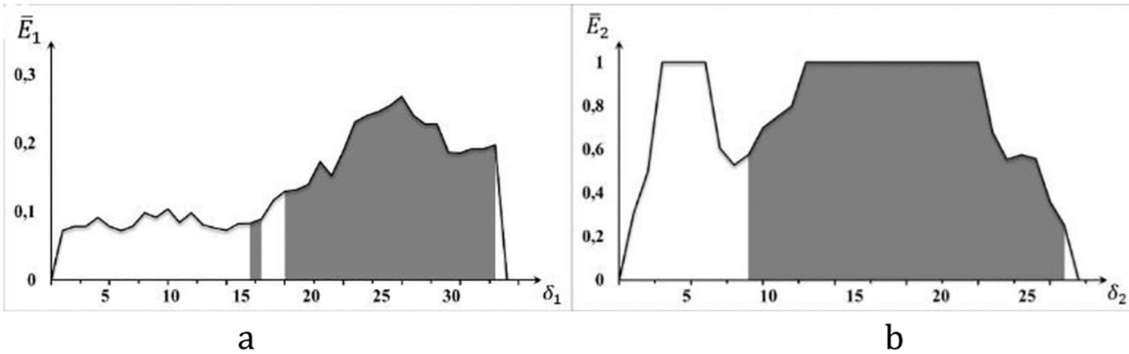


**Fig. 8.** The dependence graphs of the information criterion on the parameter of the control tolerances field: a – the first stratum; b – the second stratum.

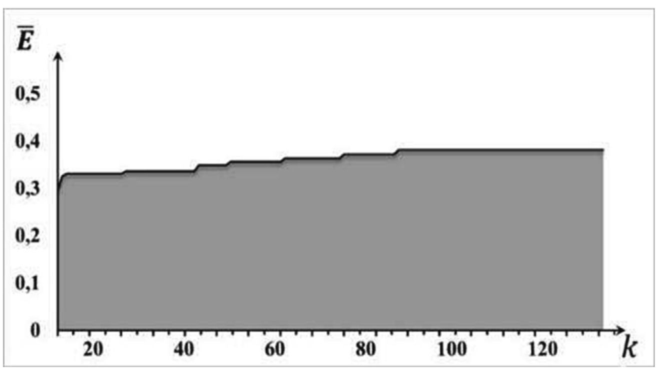


**Fig. 9.** The changes graph of the information criterion in the successive optimization process of control tolerances for the first final execution.

information-extreme machine learning process diagnostic DSS with consistent optimization of the control tolerances system for diagnostic features.

The obtained result (Fig. 9) shows that the information criterion of optimization reached the maximum value of $\overline{E}^* = 0.38$ in the second run of the procedure (5). The sequence number of the run is determined by the ratio of the iterations number $k$ (machine learning steps) to the features number $N$ in the structured vector of recognition features. Thus, in the machine learning process with consistent optimization of control tolerances, the value of the information criterion still remains low, but it has increased compared to procedure (4) of parallel optimization of control tolerances.

To build the decisive rules (6) it is necessary to know the geometric parameters of the recognition classes the containers. Fig. 10 shows the dependence graphs of the information criterion (5) on the radii of the recognition classes the containers of the final strata of the decursive tree.

Analysis of Fig. 10 shows that the optimal radii of the recognition classes containers are equal to: $d_1^* = 31$ (hereinafter in the Hamming code units) for class $X_1^o$; $d_2^* = 33$ for class $X_2^o$; $d_3^* = 40$ for class $X_3^o$ and $d_4^* = 8$ for class $X_4^o$. The average value of the normalized information criterion is $\overline{E} = 0.7$.

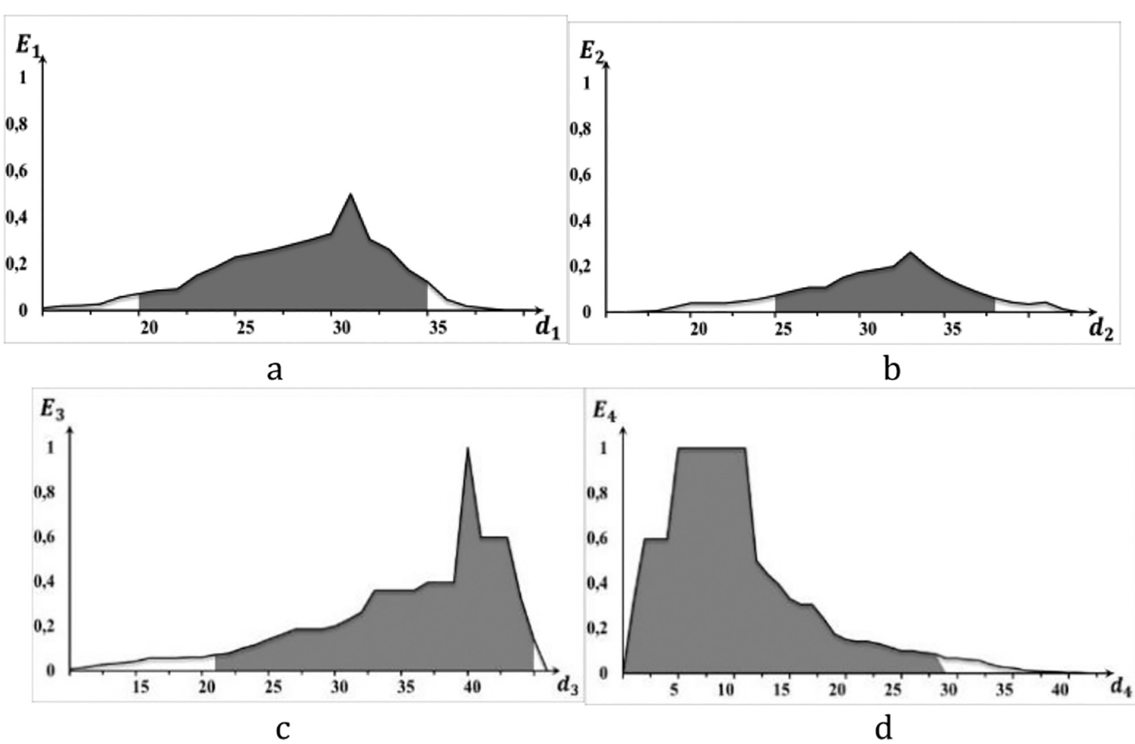


**Fig. 10.** The dependence graphs of information criterion of optimization on radii of recognition classes containers: a – class $X_1^o$; c – class $X_2^o$; c – class $X_3^o$; d – class $X_4^o$.

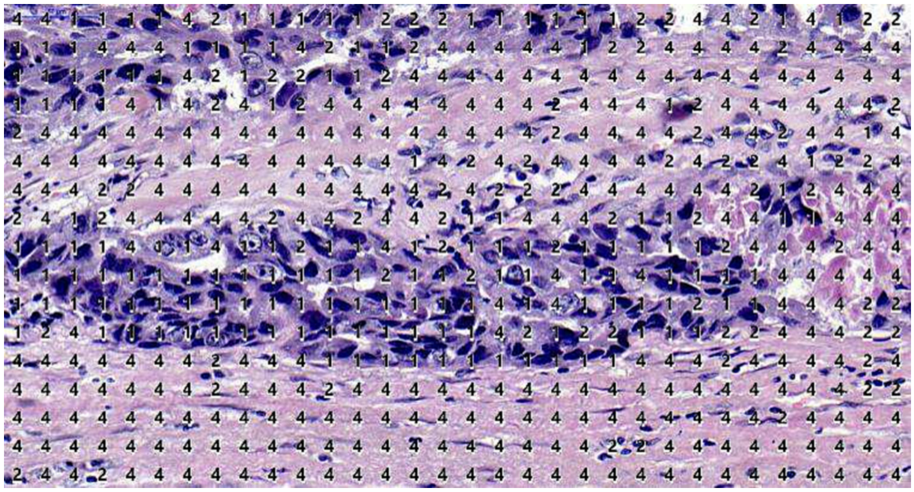

**Fig. 11.** Digitized histological representation of the results of frame identification.

Testing the functional effectiveness of machine learning DSS was carried out in the examination mode. Fig. 11 shows a digitized histological image obtained according to the decisive rules (6). In the figure, the numbers correspond to the ordinal numbers of the recognition classes.

Visual analysis of Figs 2 and 11 confirms a fairly high recognition reliability of histological images.

## Discussion

The lack of highly reliable DSS for the oncopathology diagnosis by histological images is due to the shortcomings of machine learning modern methods, including neuro-like structures, which is especially evident in the large amounts data and alphabets analysis of high-power recognition classes. A promising direction to increase the functional efficiency of DSS for the oncopathology diagnosis by histological images is the application of a functional approach to modeling the cognitive processes of natural intelligence in the formation and adoption of classification decisions. As part of this approach, the above method of information-extreme machine learning DSS for the oncopathology diagnosis on the basis of tissue morphology, which is an alternative to neuro-like structures and eliminates their main shortcomings. Based on the hierarchical information-extreme machine learning results, the decisive rules are characterized by a sufficiently high reliability for the 4 recognition classes. But the analysis of Fig. 10 shows that the decisive rules are not infallible according to the training matrix, because it was not possible to reach the maximum value of the information criterion for optimizing the machine learning parameters. This fact is confirmed by visual analysis of the segmented image (Fig. 11) of the morphological structure of the tissue. Therefore, the subject of further research is to increase the functional efficiency of diagnostic DSS by increasing the depth of information-extreme machine learning. At the same time, among the additional optimization parameters, the parameters of forming the input information description of the diagnostic DSS will play an important role. Such machine learning parameters can be, for example, the the image frame size, the RGB components weights of the histological image[12] and so on. In addition, a promising way to increase the reliability of histological image segmentation is to assess the informativeness of diagnostic features by the method of information-extreme machine learning, proposed in the authors work.[13]

## Conclusion

Within the framework of a functional approach to modeling cognitive processes of natural intelligence, for the first time an information-extreme machine learning method of diagnostic DSS based on a hierarchical data structure in the form of a decursive binary tree was developed. The proposed method allows you to automatically generate an input learning matrix, take into account the individual characteristics of the patient and retrain the system when expanding the recognition classes alphabet. As a result, deep information-extreme machine learning of diagnostic DSS is carried out for each pair of the nearest neighboring classes by a linear algorithm.

The obtained results allow to create an automated workplace of a histologist for diagnosing oncopathologies of different pathogenesis. The constructed decisive rules are not infallible according to the educational matrix. Therefore, to increase the functional efficiency of information-extreme machine learning diagnostic DSS, it is necessary to increase the depth of machine learning by optimizing additional parameters, including the parameters of the input training matrix.

## Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Anatoliy Dovbysh reports financial support, administrative support, and statistical analysis were provided by Sumy State University. Anatoliy Dovbysh reports a relationship with Sumy State University that includes: employment. Anatoliy Dovbysh has patent issued to Licensee. Anatolii Romaniuk has patent issued to Licensee. Roman Moskalenko has patent issued to Licensee.

## Acknowledgments

This research has been performed with the financial support of grants of the Ministry of Education and Science of Ukraine No. 0122U000773 “Application of artificial intelligence to provide automation and standardization of the Gleason system in the diagnosis of prostate cancer”

## References

1. Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intel 2017;39(12): 2481–2495.
2. Avramenko VV, Demianenko VM. Operative recognition of standard signal types. Радіоелектроніка, інформатика, управління 2020;2. C:75–83. https://doi.org/10.15588/1607-3274-2020-2-8.
3. Luo W, Hassell LA. Use cases for digital pathology. Digital Pathology: Historical Perspectives, Current Concepts and Future Applications. Springer Intern. Publ; 2016. p. 5-15.
4. Gerrit van den Burg PJ, Groenen.. GenSVM: a generalized multiclass support vector machine. J Mach Learn Res 2016;17(224):1-42.
5. Beckwith BA. *Standards for digital pathology and whole slide imaging. Digital Pathology: Historical Perspectives, Current Concepts and Future Applications.* Springer Intern. Publ. 2016:87–97.

6. Moskalenko VV, Korobov AG. Extreme algorithm of the system for recognition of objects on the terrain with optimization parameter feature extraction. Radio Electronics, Computer Science, Control 2017;2:38–45.
7. Xu G, Zong Y, Yang YZ. *Applied Data Mining.* CRC Press. 2013:284.
8. Rozhkova NI, Roitberg PG, Varfolomeeva AA, et al. Neural network-based segmentation model for breast cancer X-ray screening. Sechenov Med J 2020;11(3):4-14. https://doi.org/10.47093/2218-7332.2020.11.3.4-14.
9. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Cham: Springer; 2015. p. 234–241.
10. Dovbysh AS, Rudenko MS. Information-extreme learning algorithm for a system of recognition of morphological images in diagnosing oncological pathologies. Cybernet Syst Anal 2014;50(1):157–163. https://doi.org/10.15587/1729-4061.2016.71930.
11. Dovbysh AS, Budnyk MM, Piatachenko VYu, Myronenko MI. Information-extreme machine learning of on-board vehicle recognition system. Cybernet Syst Anal 2020;56(4): 534–543. https://doi.org/10.1007/s10559-020-00269-y.
12. Naumenko I, Myronenko M, Savchenko T. Information-extreme machine training of on-board recognition system with optimization of RGB-component digital images. Radioelect Comput Syst 2021;4:59–70. https://doi.org/10.32620/reks.2021.4.05.
13. Dovbysh I, Shelehov S, Pylypenko OB. Estimation of informativeness of recognition signs at extreme information machine learning of knowledge control system. CEUR Workshop Proc 2019;2362:143–152.