UDC: 004.8:004.93 https://doi.org/10.15407/jai2024.03.114

I. Shelehov[1], D. Prylepa[2], Yu. Khibovska[3]
[1,2,3]Sumy State University, Ukraine
116, Kharkivska st., Sumy, 40007
[1]Sumy National Agrarian University, Ukraine
160, Herasyma Kondratieva st., Sumy, 40000
[1]i.shelehov@cs.sumdu.edu.ua
[2]d.prylepa@cs.sumdu.edu.ua
[3]y.khibovska@vbs.sumdu.edu.ua
[1]https://orcid.org/0000-0003-4304-7768
[2]https://orcid.org/0000-0002-4022-5496
[3]https://orcid.org/0000-0001-5832-3134

INFORMATION-EXTREME MACHINE LEARNING OF AN OPHTHALMIC DIAGNOSTIC SYSTEM WITH A HIERARCHICAL CLASS STRUCTURE

Abstract. The paper considers the method of hierarchical information-extreme machine learning for the system of ophthalmic diagnosis of eye pathology. Since the proposed method is developed within the framework of a functional approach to modeling the cognitive processes of natural intelligence, it, unlike neuro-like structures, acquires the properties of flexibility when retraining a diagnostic system and requires an order of magnitude fewer image samples. In addition, the decision rules based on the results of machine learning within the geometric approach in the form of a binary hierarchical structure of recognition classes ensure their practical invariance to the multidimensionality of both the space of diagnostic features and the alphabet of recognition classes. The modified Kullback-Leibler information measure, which is considered as a function of the accuracy of classification solutions, is chosen as a criterion for optimizing the parameters of the machine learning system for diagnosing eye pathologies. A hierarchical information-extreme machine learning algorithm for an ophthalmic diagnostic system for six eye pathologies was developed and programmatically implemented. Based on the results of functional diagnostics, it has been experimentally proved that the constructed decision rules are error-free according to the training matrices of recognition classes of each level of the constructed binary hierarchical structure.

Keywords: information-extreme machine learning, functional categorical model, information criterion, input mathematical description, training matrix, diagnostic system, eye pathology.

Introduction

The diagnosis of ophthalmic diseases plays a pivotal role in preserving vision and providing timely treatment. A multitude of examination methods are employed to ascertain the correct diagnosis, encompassing both non-invasive (slit lamp examination, visual acuity, fundus imaging, ultrasound, and optical coherence tomography) and invasive (e.g., fluorescein angiography) techniques [1–3]. Non-invasive techniques, such as fundus image analysis and optical coherence tomography, are particularly promising due to their safety, ease of use, and the ability to store the results for further analysis and monitoring of the course of disease [4, 5].

The development of computer technology offers the potential for automated analysis of ophthalmic examination results, with the aim of improving diagnostic accuracy and reducing the workload of doctors. However, the implementation of effective systems for computerized diagnostics of eye pathologies is a complex computational task that requires the use of modern methods of machine learning and data mining. Despite the demonstrated efficacy of artificial neural networks, probabilistic models, and clustering methods in retinal image analysis, the quest for optimal approaches to diagnosing ophthalmic diseases persists as a pressing challenge.

The paper proposes a hierarchical information-extreme machine learning method adapted for eye image analysis. The developed approach allows diagnosing a number of typical diseases, such as optic nerve atrophy, glaucoma, retinal detachment, optic neuritis, and retinitis pigmentosa.

Problem statement

Let's consider the problem of information synthesis of a learning system for ophthalmic diagnosis of eye pathologies from images.

Let the alphabet $\{X_m^o \mid m = \overline{1, M}\}$, of recognition classes be given, each of which is characterized by a representative set of images of the object of the corresponding pathology. The input training matrix $\| y_{m,i}^{(j)} \mid i = \overline{1, N}; j = \overline{1, J} \|$, in which the *i*-th column of the matrix is a training set of values for the i-th diagnostic feature and the *j*-th row is a structured vector of features of the recognition class X_m^o, is formed during image processing by reading the values of the brightness gradations of the pixels of the receptor field.

According to the concept of Information Extreme Intelligence (IEI) technology, the input training matrix is transformed into a working binary matrix, which is adapted during the machine learning process to achieve maximum accuracy of diagnostic decisions. Let's consider machine learning with a second level of depth. At the first level, the optimal (in the information sense) geometric parameters of the hyperspherical containers of the recognition classes are determined, and at the second level, the system of control tolerances for diagnostic features is determined. At the same time, the vector of functional parameters, which influence the functional efficiency of a machine learning system for recognition of feature vectors, e.g. a recognition class X_{h,m_h}^o, has the form

$$g_h = < x_{h,m_h}, d_{h,m_h}, \delta_h > , \qquad (1)$$

where h is the level of the hierarchical structure; m_h is the number of classes for which decisive rules are formed at the h hierarchical level in the form of hyperspherical containers (x_{h,m_h} – center and d_{h,m_h} – radius of the corresponding hyperspherical container x_{h,m_h}); δ_h – the system of control tolerances is formed at the h hierarchical level.

Certain restrictions are imposed on the vector of system functioning parameters, which will be further referred to as machine learning parameters:

- the radius d_{h,m_h} of the hyperspherical container for the recognition class X_{h,m_h}^o must be less than the distance between its center and the center of the nearest container of the same hierarchy level;

- the range of acceptable values of the machine learning parameter δ_h is limited by the inequality $\delta_h < \delta_H / 2$, where δ_H is a system of normalized tolerances.

In the process of machine learning an ophthalmic diagnostic system, it is necessary to:

1) optimize the parameters of vector (1) by maximizing the information criterion averaged over all recognition classes $\{X_{h,\,m_h}^o\}$

$$\overline{E}_h = \frac{1}{M} \sum_{h=1}^{S} \sum_{m_h=1}^{M_h} \max_{G_E \cap G_{d_{h,m_h}}} E_{h,m_h}(d_{h,\,m_h}) \; ; \qquad (2)$$

2) using the optimal geometric parameters of containers for recognition classes obtained in the process of machine learning, build decision rules for each level of the hierarchical structure that ensure high overall accuracy of diagnostic decision-making.

3) at the testing stage, it is necessary to decide whether the implementation belongs to one of the recognition classes at the final level of the formed hierarchical structure.

Thus, the task of the information-extreme synthesis of a learning-based eye disease diagnostic system is to optimize its machine learning parameters by maximizing the value of the global information criterion (2) to its theoretical maximum value.

Analysis of recent research and publications

The rapid development of artificial intelligence and machine learning technologies opens new opportunities for automated analysis of ophthalmic examination results and improved accuracy of eye disease diagnosis [1-3]. Ophthalmology is actively researching the use of modern machine learning and pattern recognition methods to diagnose and monitor multimodal eye diseases [4-6]. Analysis shows that the most commonly used methods in ophthalmic research are random forests and support vectors [7-9]. Convolutional neural networks have been widely used in medical image recognition tasks, including eye images [10, 11]. However, there are still unresolved issues to ensure high functional efficiency of known machine learning methods in the context of multidimensional space of diagnostic features and significant overlap in this space of recognition classes characterizing different pathologies. The problematic task is still to provide machine learning methods with flexibility while increasing the power of the recognition class alphabet. In addition, the accuracy of machine learning using known methods depends significantly on the power of the recognition class alphabet, which requires a transition from linear data structures to hierarchical ones that ensure the construction of error-free decision rules according to the training matrix.

A promising way to overcome the above scientific and methodological difficulties is to use the ideas and methods of the so-called information-extreme intelligent technology (IET) of data analysis created by the scientific school of Anatolii Dovbysh, which is based on maximizing the information capacity of the recognition system in the process of machine learning [12-14]. Information-extreme machine learning methods, unlike neuro-like structures, are developed within the framework of a functional approach to modeling the cognitive processes of natural intelligence in the construction and adoption of classification decisions [15].

The purpose of the study

The aim of this work is to improve the accuracy of information-extensive machine learning for ophthalmic diagnostic systems by building a hierarchical binary structure of diagnostic data. Achieving this goal allows us to move from linear multi-class machine learning to two-class machine learning for each level of the binary tree, which is a prerequisite for improving the accuracy of machine learning.

Materials and methods of the study

The information synthesis of a learning ophthalmic diagnostic system is considered within the framework of IEI-technology by optimizing the machine learning parameters given by vector (1) according to the information criterion. The main way to overcome the negative impact of the multidimensionality of the recognition class alphabet on the accuracy of machine learning, which is determined by the full probability of making correct classification decisions, is to switch from a linear data structure to a hierarchical one. As a result, instead of linear multi-class machine learning, two-class machine learning is performed for the recognition classes of each level of the binary hierarchical structure. In this case, at the same level of depth and with the same dimension of the diagnostic feature space, the necessary conditions are created to improve the accuracy of machine learning. At the same time, unlike neural networks, the depth in information-extreme classifiers is determined by the number of machine learning parameters that are optimized by the information criterion.

The functional categorical model of information-extreme machine learning based on a hierarchical data structure is represented as an oriented graph, the attributes of the vertices of which are the sets involved in the machine learning process. At the same time, the edges of the graph are operators that map the corresponding set to another. The input mathematical description of the ophthalmic diagnostic system is considered in the form of a structure

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

where G – is the set of factors that affect the diagnostic system; T – is the set of moments of information reading; Ω – is the space of diagnostic features; Z – is the alphabet of recognition classes; H – is the hierarchical data structure; $Y^{|h|}$ – is the input (Euclidean) training brightness matrix of two recognition classes of the h-th level of the binary tree; $X^{|h|}$ – is the working binary training matrix of two recognition classes of the h-th level of the binary tree given in the Hamming space; g – is the operator of constructing the hierarchical data structure H; f_1 – is the operator of forming the input training matrix; f_2 – is the operator of converting the matrix $Y^{|h|}$ into a binary training matrix $X^{|h|}$.

The functional categorical model of the hierarchical information-extreme machine learning system for ophthalmic diagnosis is shown in Figure 1.

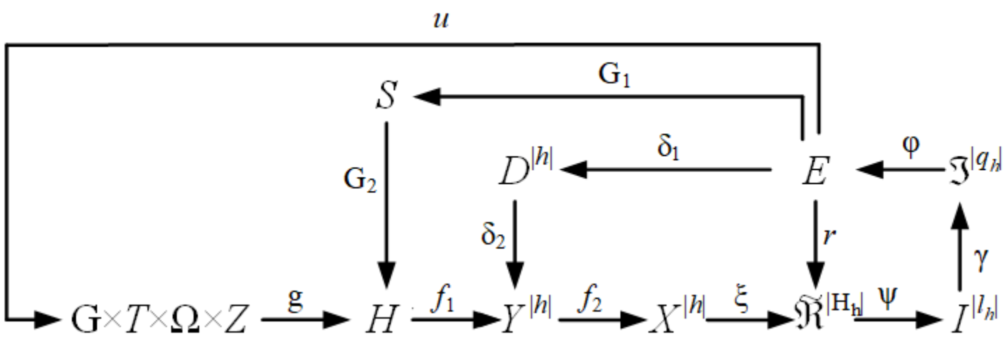


Fig. 1. Functional categorical model of hierarchical machine learning

In Figure 1, the source of information is the Cartesian product of sets $T \times G \times \Omega \times Z$. The term set E contains the values of the information criterion calculated at each step of machine learning. In accordance with the principle of complete composition, this set is common to all cycles of optimization of learning parameters. At each learning step, the operator $r: E \to \tilde{\Re}^{|M_h|}$ restores the hyperspherical containers of the recognition classes in the radial basis of the binary Hamming feature space, forming, in general, their fuzzy partitioning $\tilde{\Re}^{|M_h|}$. The operator ξ maps the binary vectors of the working matrix $X^{|h|}$ to the partition $\tilde{\Re}^{|M_h|}$. Next, the operator $\psi: X_h \to I^{|l_h|}$ tests the main statistical hypothesis $\gamma_1 : x_{h,n}^{(j)} \in X_{h,m}^{o}$. The operator γ calculates the set of precision characteristics $\Im^{|q_h|}$, where $q_h = l_h^2$. At each step of machine learning, the operator φ calculates the information criterion of optimization, which is a function of the precision characteristics. The cycle of optimization of control tolerances is closed through the set $D^{|h|}$, whose elements are the values of control tolerances for diagnostic features. The operator u_H regulates the machine learning process.

Thus, the proposed functional categorical model of information-extreme machine learning optimizes the machine learning parameters of the ophthalmic diagnostic system for each level of the hierarchical structure specified by vector (1). At the same time, the machine learning process automatically generates a binary hierarchical data structure.

The process of information-extreme machine learning based on a hierarchical data structure is performed in the following main stages:

1) based on the input training matrix for a given alphabet of recognition classes, the average structured vectors $\{y_{m_h} \mid m_h = \overline{1, M_h}\}$ of diagnostic features are calculated;

2) for each level of the binary tree, the parameter δ_h of the field of control tolerances for diagnostic features is set. For each i-th feature of the vector y_{m_h}, the lower $A_{HKm_h,i}$ and upper $A_{BKm_h,i}$ control tolerances for diagnostic features are calculated by the formulas

$$A_{HKm_h,i} = y_{m_h,i} - \delta_h;\ A_{BKm_h,i} = y_{m_h,i} + \delta_h;$$

3) a set $\{x_m\}$ of binary averaged vectors of diagnostic features is formed according to the following rule

$$x_{m,i} = \begin{cases} 1,\ if\ \ A_{HKm_h,i} < y_{m_h,i} < A_{BKm_h,i}; \\ 0,\ else; \end{cases}$$

4) at the first (highest) level (h=1) of the hierarchical structure, the training matrices of recognition classes characterized by the highest maximum value of the average

information criterion $\{E_{1,m1}\}$, where m_1 = 1..M, are selected as vertex attributes.

5) for each subsequent (lower) level, the vertex attributes are selected similarly, but from the training matrices of recognition classes that are not used on the higher levels of the binary tree.

6) the process of building the tree continues until all levels of the binary tree are formed.

Thus, the above scheme for each level creates the necessary conditions for the construction of highly accurate decision rules by means of information-extreme two-class machine learning using a linear algorithm.

According to the functional categorical model (Fig. 1), the algorithm for information-extreme machine learning of the system for diagnosing eye diseases based on the hierarchical data structure can be represented for the h-th level of the binary tree as an iterative procedure for finding the global maximum of the criterion (2) averaged by alphabetical $\{X^o_{h,m_h}\}$ recognition classes:

$$\delta^*_{K,h} = \arg\max_{G_{\delta,h}}\{\max_{G_E \cap G_{d_h}} \overline{E}_h(d_h)\}\,. \qquad (3)$$

Unlike a linear algorithm, where the optimal value of the parameter δ_h is determined for the entire set of recognition classes, in information-extreme machine learning based on a hierarchical data structure, the parameter δ_h is determined separately for each level of the hierarchy.

The internal loop of procedure (3) implements a basic algorithm that performs the following functions at each step of machine learning:

- calculates the value of criterion (2);
- finds the global maximum of this criterion;
- determines the optimal geometric parameters of hyperspherical containers for recognition classes.

Input data of the basic algorithm: an array of realizations $\{y^{(j)}_{m_h,i} \mid m_h = \overline{1,M_h}; i = \overline{1,N}, j = \overline{1,n}\}$, a system of control tolerances $\{\delta_{h,i})$ for diagnostic features, selection levels $\{\rho_{m_h,i}\}$ of coordinates of binary averaged feature vectors, which are equal to $\rho_{m_h,i}$ = 0.5 by default.

The geometric parameters of the containers of the recognition classes are optimized using the basic machine learning algorithm according to the procedure [13]:

1) formation of the input structured training matrix of recognition classes of the h-th level of the binary tree;

2) determination of the averaged vector realizations of recognition classes;

3) formation of a binary training matrix according to a given system of control tolerances for diagnostic features;

4) determination of the averaged binary feature vectors by statistical averaging of the corresponding binary training samples;

5) calculation of center-to-center distances in a given recognition class alphabet as code distances between the averaged feature vectors of recognition classes.

6) calculation of the average information criterion for optimizing machine learning parameters at each training iteration;

7) searching for the global maximum of the averaged information criterion in the working area of determining the criterion function;

8) determining the optimal radii of the containers of recognition classes by restoring them in the radial basis of the diagnostic feature space using an iterative procedure at each step of machine learning

$$d^*_{h,m_h} = \arg\max_{G_E \cap G_{d_h}} \overline{E}_h(d_{h,m_h}),\; m = \overline{1,M}_h\,, \qquad (4)$$

9) STOP.

Procedure (3) of forming a binary training matrix according to the system of control tolerances for diagnostic features is implemented in an outer loop with the operator of changing the parameter δ_h of the tolerance field until the maximum value of the information criterion for optimizing machine learning parameters is reached.

The modified Kullback-Leibler measure was used as a criterion for optimizing the machine learning parameters of the ophthalmic diagnostic system for each level of the hierarchical structure of the input data, which for two-probability alternative hypotheses has the form [13, 16, 17]:

$$E_{h,m}^{(k)} = \frac{1}{2}\{2-[\alpha_{h,m}^{(k)}(d)+\beta_{h,m}^{(k)}(d)]\}\times$$
$$\times\log_2\frac{2-[\alpha_{h,m}^{(k)}(d)+\beta_{h,m}^{(k)}(d)]+10^{-p}}{\alpha_{h,m}^{(k)}(d)+\beta_{h,m}^{(k)}(d)]+10^{-p}}. \qquad (5)$$

Due to the limited size of the training sample, estimates of the accuracy characteristics were used in the calculation of the information optimization criterion (5) in the process of implementing the machine learning algorithm.

Based on the optimal geometric parameters of the hyperspherical containers of the recognition classes obtained in the process of machine learning, constructed the decision rules in the form of implicit relations [18, 19]:

$$(\forall X_{m,h}^{o} \in \Re^{|M_h|})(\forall x^{(j)} \in \Re^{|M|})\{if\ [(\mu_{m_h} > 0)\,\&$$
$$\&(\mu_m = \max_{\{m\}}\{\mu_{m_h} \mid m_h = \overline{1,2}\}]\ then\ x^{(j)} \in X_{h,m_h}^{o} \qquad (6)$$
$$else\ x^{(j)} \notin X_{h,m_h}^{o}\}.$$

In expression (6), the function μ_m of vector $x^{(j)}$ membership in the hyperspherical container of the recognition class X_{h,m_h}^{o} is defined by the formula

$$\mu_m = 1 - \frac{d(x^{(j)} \oplus x_{h,m_h}^{*})}{d_{h,m_h}^{*}}. \qquad (7)$$

Thus, the vector of diagnostic features $x^{(j)}$ is classified as belonging to the class from the given alphabet of the corresponding hierarchy level for which the membership function (6) takes on a positive maximum value. In addition, the decisive rules (6), constructed within the framework of the geometric approach, allow for real-time diagnostic decisions, which is important for functional diagnostics.

Experiments

Let's consider information-extreme machine learning for an ophthalmic diagnostic system. As input, we used images of six typical eye pathologies representing the corresponding recognition classes. These classes are ordered according to the proposed methodology for forming a variation series, the visualization of which is shown in Fig. 2. These images served as a test set for validating the developed machine learning approach in the context of the task of computerized diagnosis of eye diseases.

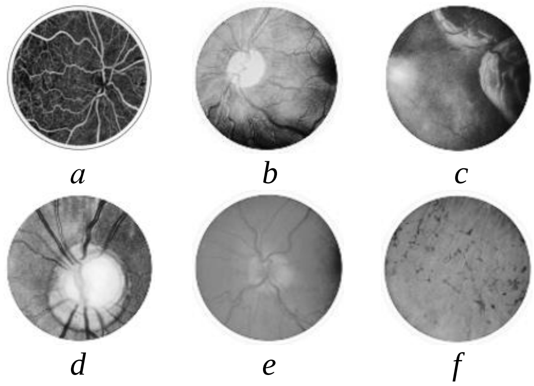


Fig. 2. Images of eye pathologies: *a* – class X_1^o; *b* – class X_2^o; *c* – class X_3^o; *d* – class X_4^o; *e* – class X_5^o ; *f* – class X_6^o

Below are the recognition classes that form the specified recognition class alphabet:

1) normal condition of the eye (recognition class X_1^o);

2) atrophy of the optic nerve (recognition class X_2^o);

3) retinal detachment (recognition class X_3^o);

4) glaucoma (recognition class X_4^o);

5) optic neuritis (recognition class X_5^o);

6) retinitis pigmentosa (recognition class X_6^o);

The input training matrix for the ophthalmic diagnosis system was formed by reading the pixel brightnesses of fragments of eye pathology images with 100×100 pixel size The value of each feature in the input matrix corresponded to the brightness of the corresponding image pixel in the range from 0 to 255. Given the assumption that the images are stationary in brightness, they were scanned in the Cartesian coordinate system. In addition, its transposed matrix was added to the input training matrix, which doubled the dimension of the diagnostic feature space. This approach increased the amount of information and, in accordance with the maximum-distance principle of pattern recognition theory, created the necessary conditions for increasing the average interclass code distance.

Analysis of the results

For the initial alphabet of recognition classes, a linear information-extreme machine learning algorithm with parallel optimization of control tolerances for diagnostic features was implemented. Figure 3 shows a graph of the dependence of the information criterion (5) on the parameter of the control tolerance field obtained in the process of machine learning.

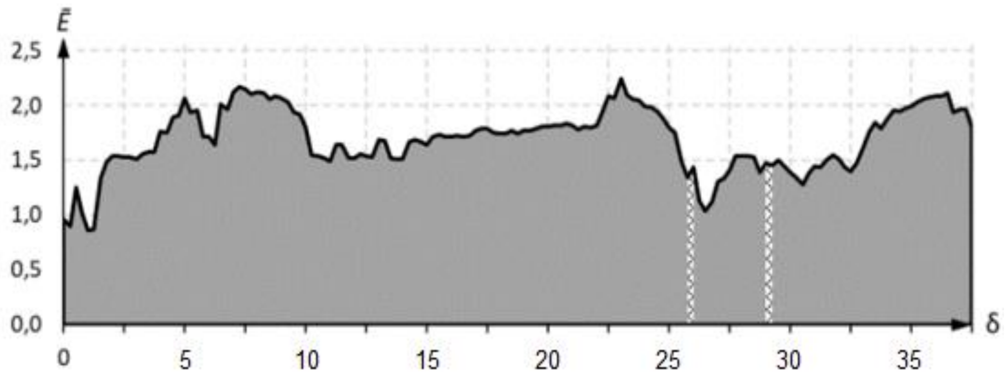
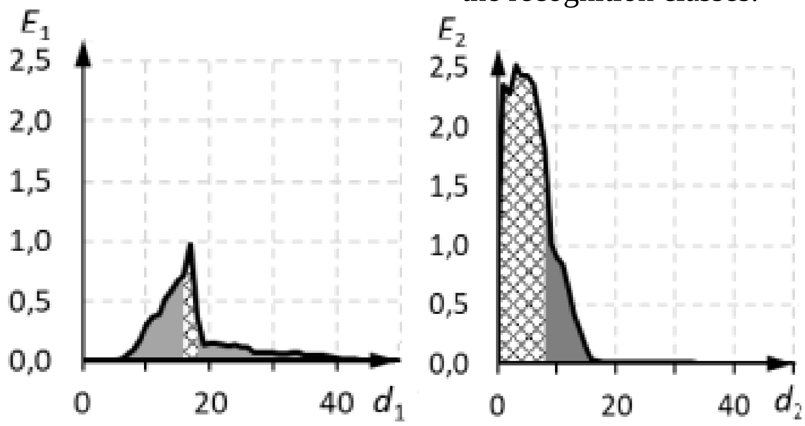


Fig. 3. Graph of dependence of the information criterion on the parameter of the control tolerance field

In Figure 3 and further in the text, double hatching indicates the working areas for determining the criterion function (5), in which the first confidence exceeds 0.5 and the second kind of error is less than 0.5. The analysis of Figure 3 shows that the optimal value of the parameter of the control tolerance field is equal $\delta_{1,1}^* = 28$ (in brightness gradations) to the maximum value of the information criterion $\overline{E}_{1,1}^* = 1,47$.

To build the decision rules (6), it is necessary to determine the optimal geometric parameters of the hyperspherical containers of the recognition classes. Figure 4 shows the graphs of the dependence of the information criterion (5) on the radii of the containers of the recognition classes.



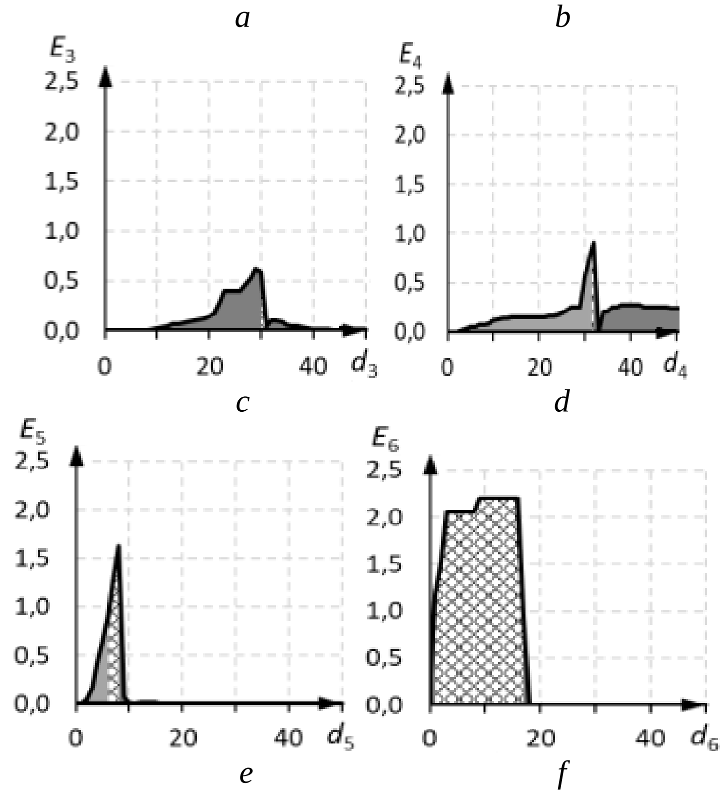


Fig. 4. Graph of dependence of the information criterion on container radii: *a* – recognition class X_1^o; *b* – recognition class X_2^o; *c* – recognition class X_3^o; *d* – recognition class X_4^o; *e* – recognition class X_5^o; *f* – recognition class X_6^o

The results of the functional testing of the ophthalmic diagnosis system based on the developed decision rules showed that the full probability of correct classification is 0.83. To improve the accuracy, an information-extreme machine learning algorithm was implemented for sequential optimization of the control tolerance system, in which the control tolerances obtained from the results of parallel optimization were taken as the starting ones. Figure 5 shows a graph of changes in the average information criterion in the process of parallel-serial optimization of control tolerances.

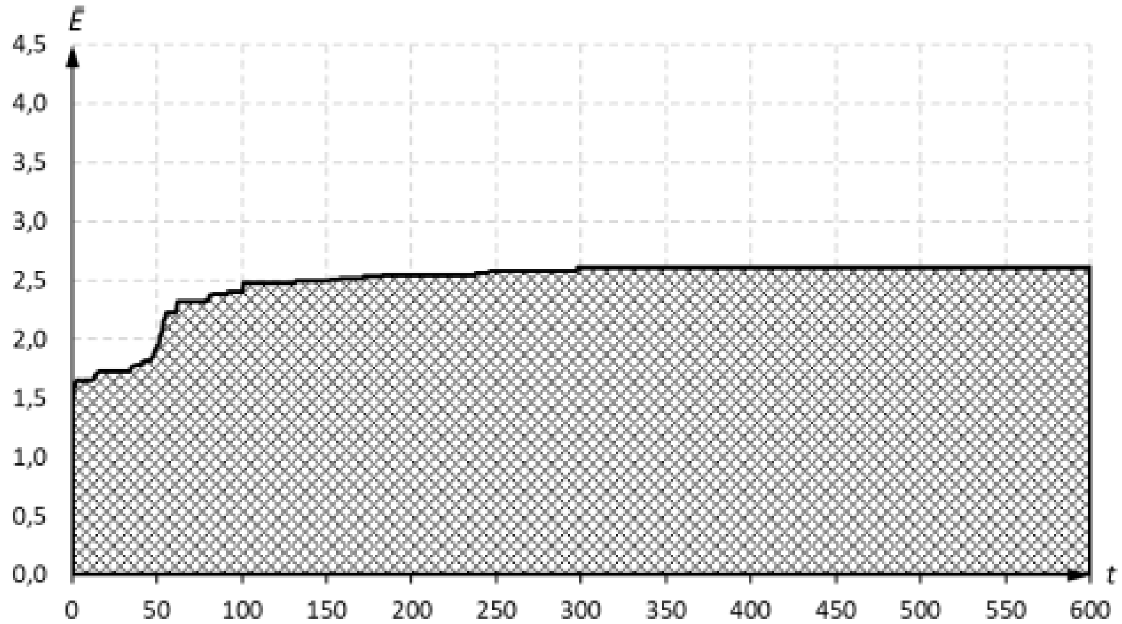


Fig. 5. Graph of changes in the information criterion in the process of sequential optimization of control tolerances

An analysis of Figure 5 shows that the maximum value of the average information criterion was achieved at the 300th iteration and is 2.61, which exceeds its value obtained in parallel optimization. Table 1 shows the results of machine learning.

Table 1. Machine learning results for the first level of the hierarchy

Classes recognition		$E^*_{m_h}$	$d^*_{m_h}$	*Precision characteristics*	
				D^*_1	β^*
X^o_1	$X^o_{1,1}$	4,39	28	1	0
X^o_2	$X^o_{1,2}$	2,61	22	0,86	0
X^o_3	$X^o_{1,3}$	4,39	30	1	0
X^o_4	$X^o_{1,4}$	0,71	10	0,51	0,01
X^o_5	$X^o_{1,5}$	1,57	10	0,72	0,01
X^o_6	$X^o_{1,6}$	3,72	23	0,96	0

An analysis of Table 1 shows that the decision rules provide a full probability of correct classification of 0.92.

In order to build error-free decision rules based on the training matrix, a binary hierarchical structure was constructed. The attributes of the vertices of the first level of the structure are the training matrices of the recognition classes X^o_1 and X^o_3, for which, according to Table 1, error-free decision rules are constructed. The remaining recognition classes were transferred to the second level of the hierarchy. At this level, we implemented linear machine learning with parallel-sequential optimization of the control tolerance system. Figure 6 for this case shows the graph of changes in the average information criterion in the process of parallel-serial optimization of control tolerances.

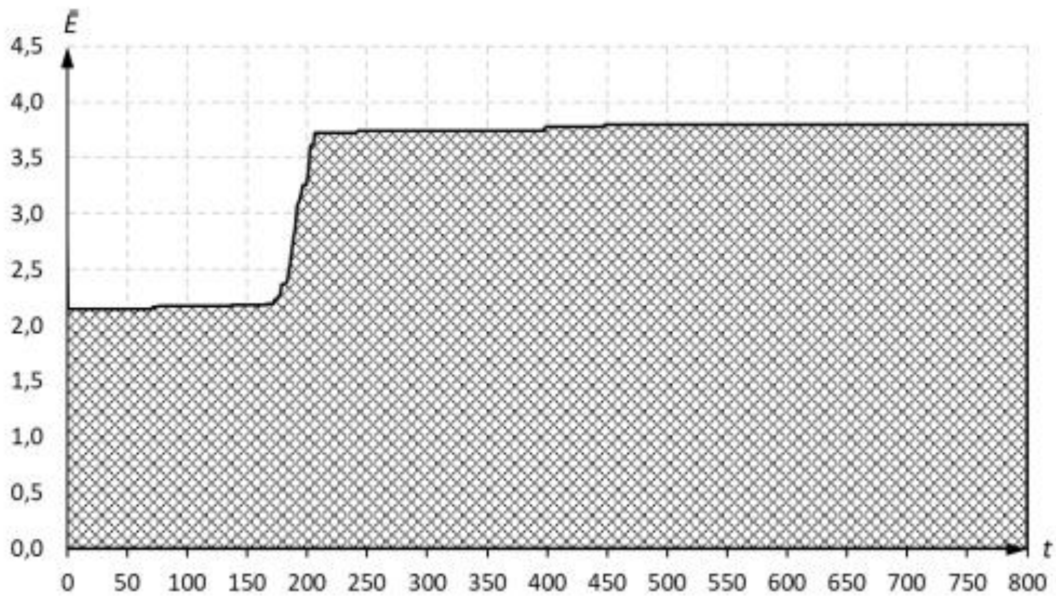


Fig. 6. Graph of changes in the information criterion in the process of sequential optimization of control tolerances of the second level of the hierarchical structure

According to the graph in Figure 6, for the second level of the hierarchical structure, the maximum value of the average information optimization criterion of $\overline{E}^* = 3.80$ was achieved at the 450th iteration. Table 2 shows the results of the parallel-sequential optimization of machine learning parameters for the second level of the hierarchical structure of recognition classes.

Table 2. Machine learning results for the second level of the hierarchy

Class recognition		$E^*_{m_h}$	$d^*_{m_h}$	*Precision characteristics*	
				D_1^*	β^*
X_2^o	$X_{2,1}^o$	4,21	19	0,99	0
X_4^o	$X_{2,2}^o$	4,39	21	1	0
X_5^o	$X_{2,3}^o$	4,39	17	1	0
X_6^o	$X_{2,4}^o$	2,20	19	0,81	0

Thus, at the second level of the hierarchical structure, the recognition classes X_4^o and X_5^o, for which error-free decision rules have been formed, remain. Recognition classes X_2^o and X_6^o are moved to the third level of the hierarchy. Figure 7 shows a graph of the dependence of the average information criterion (5) obtained by parallel optimization of control tolerances.

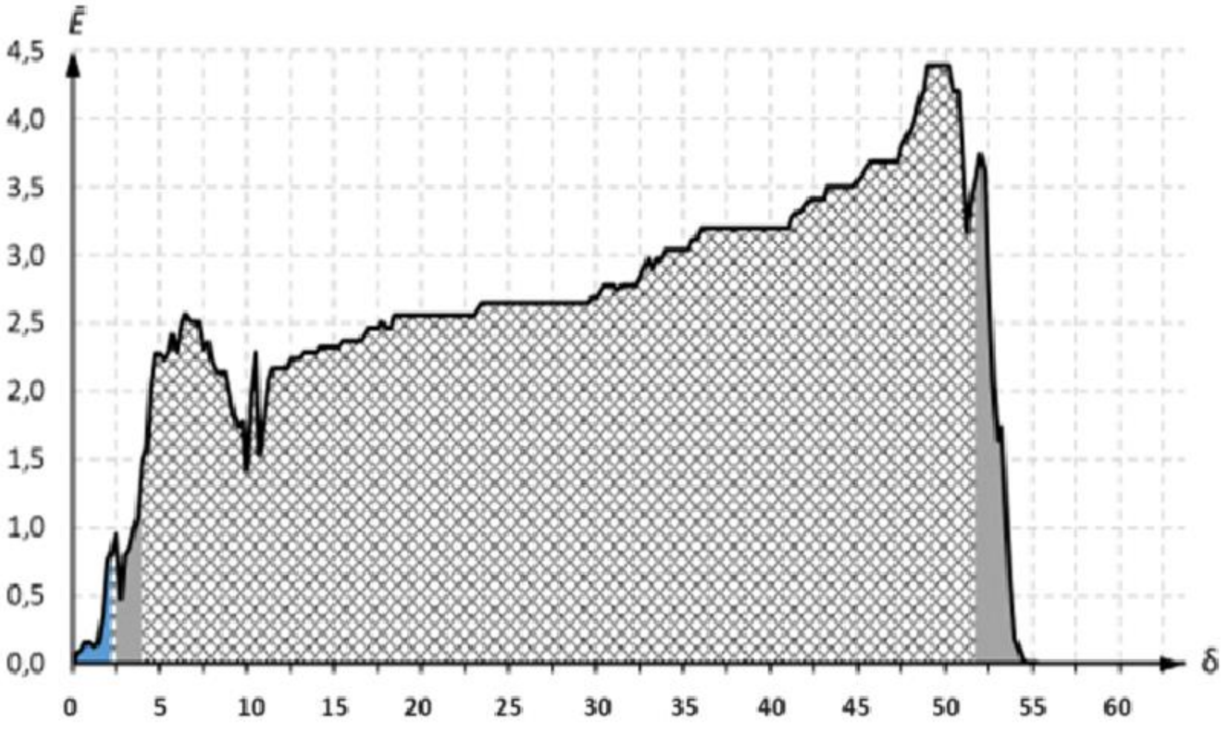


Fig. 7. Graph of dependence of the information criterion on the parameter of the control tolerance field of the third level of the hierarchy

The analysis of Fig. 7 shows that for the third level of the hierarchical structure, the average information criterion reaches a maximum threshold value of $\overline{E}^* = 4.39$ with the optimal parameter $\delta^* = 48$. Table 3 shows the results of parallel-sequential optimization of machine learning parameters for recognition classes at all levels of the hierarchical structure of recognition classes.

Table 3. Machine learning results from all levels of the hierarchy

Class recognition		*Level of hierarchy*	$E^*_{m_h}$	$d^*_{m_h}$	*Precision characteristics*	
					D_1^*	β^*
X_1^o	$X_{1,1}^o$	I	4,39	28	1	0
X_3^o	$X_{1,2}^o$		4,39	30	1	0
X_4^o	$X_{2,1}^o$	II	4,39	21	1	0
X_5^o	$X_{2,2}^o$		4,39	17	1	0
X_2^o	$X_{3,1}^o$	III	4,39	3	1	0
X_6^o	$X_{3,2}^o$		4,39	10	1	0

The analysis of Table 3 shows that the constructed classifiers are error-free according to the training matrices of the levels of the binary hierarchical structure (Fig. 8).

The obtained results of applying information-extreme machine learning to hierarchically organized data open up new opportunities to overcome the problem of the multidimensional alphabet of recognition classes. On the example of developing a system for ophthalmic diagnosis of eye diseases based on pathology images, the possibility of automatic retraining of the system when expanding the alphabet of recognition classes describing the corresponding visual pathologies is demonstrated.

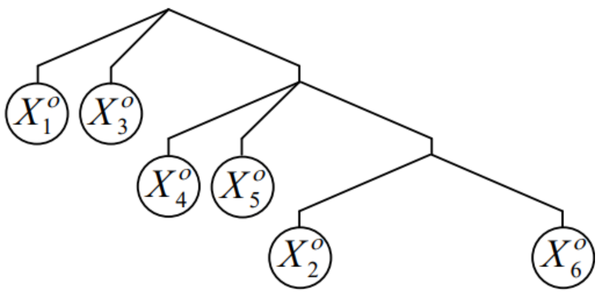


Fig. 8. Three-level hierarchical data structure

It is known that the use of linear machine learning algorithms with a large alphabet capacity leads to a significant decrease in its accuracy due to the high degree of overlap between recognition classes in the space of diagnostic features. The proposed method, unlike linear machine learning, determines the optimal control tolerances for each level of the hierarchical structure of classes separately. At the same time, we demonstrate the feasibility of using parallel-sequential optimization of control tolerances in machine learning. This leads to an increase in both recognition accuracy and machine learning efficiency, since during sequential optimization, the search for the global maximum of the information criterion is carried out in the working areas defined during parallel optimization.

Conclusions

The paper proposes a new approach to pattern recognition with a large number of classes based on a functional-categorical model and information-extreme machine learning. The approach uses a hierarchical data structure, which allows at each level to determine the optimal knowledge of functional parameters in the information sense and to form error-free decision rules for a part of the recognition classes according to the training matrix. Machine learning of the system for recognizing eye pathologies from images is carried out at each level of the hierarchy using a linear algorithm of sufficient depth, which ensures high recognition accuracy and efficiency of the proposed approach.

It is important to note that the obtained rules are error-free according to the training matrices, thus achieving a sufficient depth of machine learning and optimizing other system parameters, including the parameters of forming the input information description, will not lead to an increase in the system's functional efficiency. This approach has significant potential for use in various pattern recognition tasks with a large number of classes.

The research was concluded in the Intellectual Systems Laboratory of Computer Science Department at Sumy State University with the financial support of the Ministry of Education and Science of Ukraine in the framework of state budget scientific and research work of DR No. 0124U000548 “Information Technology for Ensuring the Resilience of the Intelligent Onboard System of Small-Sized Aircraft”.

References

1. Woreta, F.A., Gordon, L.K., & Pérez-González, C.E. (2022). Enhancing Diversity in the Ophthalmology Workforce. Ophthalmology, 129(10), 127–136.
https://doi.org/10.1016%2Fj.ophtha.2022.06.033

2. Lu, W., Tong, Y., Yu, Y., Xing, Y., Chen, C., & Shen, Y. (2018b). Applications of Artificial Intelligence in Ophthalmology: General Overview. Journal of Ophthalmology, 2018, 1–15. https://doi.org/10.1155/2018/5278196

3. Oh, E., Yoo, T. K., & Park, E.-C. (2013). Diabetic retinopathy risk prediction for fundus examination using sparse learning: a cross-sectional study. BMC Medical Informatics and Decision Making, 13(1).
https://doi.org/10.1186/1472-6947-13-106

4. Cho, N. H., Shaw, J. E., Karuranga, S., Huang, Y., da Rocha Fernandes, J. D., Ohlrogge, A. W., & Malanda, B. (2018). IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Research and Clinical Practice, 138, 271–281.
https://doi.org/10.1016/j.diabres.2018.02.023

5. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2019. (2018). Diabetes Care, 42(Supplement 1), S13—S28. https://doi.org/10.2337/dc19-s002

6. Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., ... Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell, 172(5), 1122–1131.e9.
https://doi.org/10.1016/j.cell.2018.02.010

7. Rahimy, E. (2018). Deep learning applications in ophthalmology. Current Opinion in Ophthalmology, 29(3), 254–260.
https://doi.org/10.1097/icu.0000000000000470

8. Catania, L. J., & Nicolitz, E. (2018). Artificial Intelligence and Its Applications in Vision and Eye Care. Advances in Ophthalmology and Optometry, 3(1), 21–38.
https://doi.org/10.1016/j.yaoo.2018.04.001

9. Lu, W., Tong, Y., Yu, Y., Xing, Y., Chen, C., & Shen, Y. (2018). Applications of Artificial Intelligence in Ophthalmology: General Overview. Journal of Ophthalmology, 2018, 1–15.
https://doi.org/10.1155/2018/5278196

10. Bejnordi, B. E., Zuidhof, G., Balkenhol, M., Hermsen, M., Bult, P., van Ginneken, B., Karssemeijer, N., Litjens, G., & van der Laak, J. (2017). Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. Journal of Medical Imaging, 4(04), 1.
https://doi.org/10.1117/1.jmi.4.4.044504

11. Dovbysh, A., Shelehov, I., Romaniuk, A., Moskalenko, R., & Savchenko, T. (2023). Decision-making support system for diagnosis of oncopathologies by histological images. *Journal of Pathology Informatics*, 100193.
https://doi.org/10.1016/j.jpi.2023.100193.

12. Dovbysh, A., & Zimovets, V. (2018). Hierarchical Algorithm of the Machine Learning for the System of Functional Diagnostics of the Electric Drive. In Advanced Information Systems and Technologies: Proceedings of the VI International Scientific Conference (pp. 85-88).

13. Dovbysh A.S. (2009) Osnovy proektuvannia intelektualnykh system: Navchalnyi posibnyk. Sumy: Vydavnytstvo SumDU.

14. Dovbysh, A. S., Budnyk, M. M., Piatachenko, V. Y., & Myronenko, M. I. (2020). Information-Extreme Machine Learning of On-Board Vehicle Recognition System. Cybernetics and Systems Analysis, 56(4), 534–543.
https://doi.org/10.1007/s10559-020-00269-y

15. Shelehov I. V., Barchenko N. L., & Prylepa D. V. (2022). Information-extreme machine training system of functional diagnosis system with hierarchical data structure. *Radio Electronics, Computer Science, Control*, (2). 189–200.
https://doi.org/10.15588/1607-3274-2022-18

16. Papchenko, O., Kuzikov, B., & Shovkoplias, O. (2023). Enhancing functional efficiency in information-extreme machine learning with logistic regression ensembles. Radioelectronic and Computer Systems, (4), 65–74.
https://doi.org/10.32620/reks.2023.4.06.

17. Moskalenko V.V., Dovbysh A.S., Naumenko I. V., Moskalenko A. S., & Korobov A. G. (2018). Improving the effectiveness of training the on-board object detection system for a compact unmanned aerial vehicle. *Eastern-European Journal of Enterprise Technologies*, 4/9(94), 19-26.
doi: 10.15587/1729-4061.2018.139923

18. Dovbysh, A. S., Budnyk, N. N., & Moskalenko, V. V. (2012). Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes. *Journal of Automation and Information Sciences*, 44(10), 35–44.

19. Moskalenko, V. V., Moskalenko, A. S., & Korobov, A. G. (2018). Models and methods of intellectual information technology of autonomous navigation for compact drones. *Radio Electronics, Computer Science, Control*, (3).
https://doi.org/10.15588/1607-3274-2018-3-8.

The article has been sent to the editors 09.05.24.
After processing 24.09.24.
Submitted for printing 30.09.24.