Transforming image descriptions as a set of descriptors to construct classification features

Volodymyr Gorokhovatskyi[1], Iryna Tvoroshenko[1], Olena Yakovleva[1,2,3]
[1]Department of Informatics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[2]Department of Economics and Finance, Bratislava University of Economics and Management, Bratislava, Slovakia
[3]Department of AI R&D, SYTOSS Ltd., Bratislava, Slovakia

Article Info

Article history:

Received Sep 26, 2023
Revised Nov 7, 2023
Accepted Nov 13, 2023

Keywords:

Computer vision
Descriptor
Distance matrix
Image classification
Keypoint

ABSTRACT

The article develops methods to solve a fundamental problem in computer vision: image recognition of visual objects. The results of the research on the construction of modifications for the space of classification features based on the application of the transformation of the structural description through the decomposition in the orthogonal basis and the implementation of the distance matrix model between the components of the description are presented. The application of the system of orthogonal functions as an apparatus for the transformation of the description showed the possibility of a significant gain in the speed of processing while maintaining high indicators of classification accuracy and interference resistance. The synthesized feature systems' effectiveness has been confirmed in terms of a significant increase in the rate of codes and a sufficient level of efficiency. An experimental example showed that the time spent calculating the relevance of descriptions according to their modified presentation is more than ten times shorter than for traditional metric approaches. The developed classification features can be used in applied tasks where the time of visual objects' identification is critical.

Corresponding Author:

Iryna Tvoroshenko
Department of Informatics, Kharkiv National University of Radio Electronics
61166 Kharkiv, Ukraine
Email: iryna.tvoroshenko@nure.ua

1. INTRODUCTION

The formation of a productive feature system to ensure effective image recognition in computer vision is a key task of data science and big data [1]–[5]. The priority direction of the latest research is the study of the nature of the analyzed data to establish existing or hidden regularities in their composition [4], [6], and implement the acquired knowledge into the process of recognition [7]–[11]. The presentation of the description in multi-valued spaces, especially through its transformation [9], [10], [12]–[16], makes it possible to focus on significant classification properties of the data, which will provide a sufficient level of efficiency in the applied aspect. The description is a reflection of objective reality–an image signal reflected in the parameters of the new space.

Traditional methods are being improved. These methods use projection values to look at input data in the space of orthogonal functions and use those values as features for recognition [6], [7]. These methods have a strict formal justification, and they can universally process any data. The effectiveness of these methods depends on the degree of distributability of the selected system functions about the analyzed data [12], [17]–[20]. When processing orthogonal systems about a given base of image descriptions, it is possible to slightly reduce the full set of decomposition elements to create a limited subset that meets the needs of the productive classification. In this case, in proportion to the degree of reduction, there is a possibility of achieving a significant reduction in computing costs [21]. When applying the decomposition apparatus, the basic set for designing descriptions is already set, so it does not need to be formed in the learning process. Here, the learning technology can contribute both to direct adaptation to the etalon base and to the formation of an effective subset, which makes it possible to perform an effective analysis in a limited period [5], [7], [13], [15], [22].

Structured methods of classification use a limited set of vectors to describe the key points in an image to show what the object looks like [23]–[26]. Methods using descriptors of the keypoints are physically based on information and features of the image itself. The use of descriptions in the form of a binary descriptors' set is particularly effective for embedded systems: unmanned aerial vehicles, mobile devices, and robotic and satellite systems [26]–[28]. The value of the relevance measure for a pair of image descriptions is estimated as the proximity or distance between two sets of vectors of approximately the same power [5], [10], [13]. The basis for determining such a measure is the distance matrix, which contains the value of the metric between all pairs of elements of two sets according to the "one-to-one" rule. The relevance of sets based on the principle of voting is also calculated using the distance matrix [2], [14], [18]. The distance matrix refers to the structural metric data models; it is derived from the set of values of the analyzed data, and it contains more meaningful classification information about the differences between data sets than the direct values of the elements of these sets. Statistics of distances, even within one set (description), significantly reflect the properties of the data, and they can independently be the basis for classification.

Computer vision system researchers are putting more and more effort into creating new image classification methods that can be used in real life [1], [9], [29], [30]. We can add recognition by a set of descriptors of keypoints [5], [8], [10], [13] to the list of methods that work well when there are geometric changes and disturbances. The implementation of this device ensures invariance to geometric transformations of images, high speed of data analysis, and decent results. These methods, compared to neural networks, do not require long-term training as they extract feature information directly from the image. The possibility to implement training procedures, including those with a teacher, in such classifiers, where detailed descriptions of classes are considered, also contributes to further improvement of their indicators. An example of an effective solution to the problem of plant disease classification in agriculture is a hybrid approach based on deep learning methods [31], [32].

Because image signals are multidimensional and spread out in space, along with a well-defined set of features, many statistical methods [2], [8], [10], [33] and methods for analyzing spatial data have become popular. These methods make processing much easier [3], [4], [7], [15]. These methods include the decomposition of data in orthogonal space [2], [9], [12], [17], and the use of the distance matrix model [2], [4], [14], [34]. One of the directions to develop structural methods is the transformation of the existing features' system (a set of descriptors) by creating a modified space with the necessary classification features. Such structural description transformations as a hierarchical representation [35], cluster representation [15], division of descriptors into groups by distance to the medoid [13], spatial analysis of the system of non-intersecting fragments [8], [33], and application of distributions for description components are offered and successfully implemented in the classification problem [10]. Such transformations not only speed up the operation of the classifier but also ensure a sufficient level of accuracy. High-speed methods [3], [5], [18] based on the presentation of descriptions by "centers" of data have been developed, but their effectiveness may not be sufficient, as it largely depends on the coordination of the method of selecting centers and the content of the data.

Gorokhovatsky *et al.* [14] proposes the use of a distance matrix on a set of classes as a criterion for the quality of the classification. The result is a model for finding the best threshold to see that the parts of the description are the same, which is a solution to the two-criteria optimization problem. The 0.32 threshold that was found through experiments is the lowest level of aggregated quality for speeded up robust features (SURF) descriptors with 64 dimensions. Metric relations between elements within a structural description can be a source for constructing final features. The matrix of distances between description components is a component of clustering procedures in the space of features, with further construction of a classifier model within the framework of the "bag of words" technology [2], [4], [15]. The use of the matrix of pairwise distances in the tasks of visualization of multidimensional objects based on Kohonen maps is known in information retrieval systems [2], [3] when assessing the classification results [14]. The distance matrix tool can be implemented on a set of image classes that meet the criteria of the inaccuracy matrix [2].

A variant of the effective presentation can be related to the transformation of the image in orthogonal space. Application of the system of Walsh functions and transformations related to them [12], [17] is of practical interest due to its simple software and hardware implementation. The implementation of piecewise constant Walsh basis functions is characterized by insignificant computational costs since it is not related to multiplication operations and it is implemented in the space of integers. The obtained set of integer feature vectors is a solid basis for building a classifier. Description components have characteristics, the most important of which are informativeness and metric relationships within classes. The value of these parameters is evaluated with the help of special criteria both at the training stage and according to the measurement result.

Based on this, the description can be reduced to speed up the classification process [14], [34]. We conclude that the conceptual decomposition of the component descriptors' set and the formalism of the distance matrix for the composition of class descriptions can be the sources for building a new effective system of classification features. It is proposed to study the possibility and effectiveness of their application in structural classification models.

The purpose of the work is to develop structural methods of image classification in the aspect of constructing new spaces for classification features. The transformation of the description into a set of descriptors by implementing the component decomposition apparatus based on the system of orthogonal functions and based on the values of the distance matrix for the elements of the description causes a reduction in computational costs in the analysis of numerous sets of multidimensional components. The tasks of the research consist of implementing orthogonal decomposition of the etalon and input images, using the distance matrix apparatus to obtain integral characteristics of the description, constructing a metric classifier in the transformed space of features, and studying the effectiveness of the developed classifiers' modifications (performance, speed, immunity) using simulation modeling.

2. METHOD

2.1. Scheme of space modification features

In the space of structural features, the image description $Z=\{z_i\}_{i=1}^{s}$ is a finite set of s vectors Z_i, $s=card\ Z, z_i\in B^n, z_i=\{z_{i,j}\}_{j=1}^{n}$, B^n – vector space of n dimension with binary components {0,1} [5], [8]. We consider the descriptor of the image keypoints as a point of n -dimensional discrete signal. We will interpret the transformation of Z description as a mapping $B^n\rightarrow R$, where R is the newly created space of the classification features. The scheme of R implementation is presented in Figure 1.

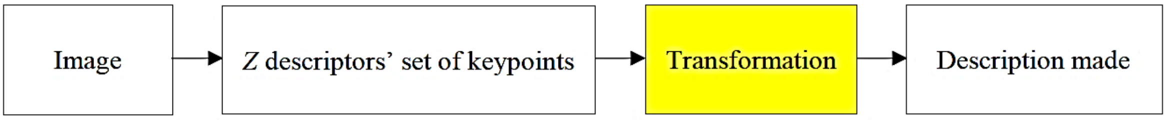


Figure 1. Scheme of description transformation application

The article studies two approaches to the modification of the feature space: orthogonal decomposition and representation in the form of a distance matrix between the components of the description. The formation of a new space is studied in terms of the application of a modified description in the classification problem.

2.2. Presentation of the description in the orthogonal basis and classification

In any finite vector space there is at least one orthogonal system of vectors $W=\{\{w_{j,b}\}_{j=1}^{n}\}_{b=1}^{n}$, so that any vector $z\in B^n, z=\{z_b\}_{b=1}^{n}$ can be represented in it by a tuple $\alpha=\{\alpha_j\}_{j=1}^{n}$, as shown in (1).

$$\alpha_j=(z\cdot w_j)=\sum_{b=1}^{n} z_b w_{j,b}. \quad (1)$$

Coefficient (1) is defined as the scalar product of z vector on w_j vector of the orthogonal system [7], [12], [17], [36]. α_j parameters are the vector's coordinators in W basis. For each $z_i\in Z$ we have the value of the expansion vector α_i. It follows from the introduction of the vector space model to classification that any descriptor $z_i\in Z$ of an image illustrated can be inversely presented in a single way as a linear combination of basis vectors, and the selected basis constitutes exact n vectors, i.e.

$$z_i=\left(\alpha_i\cdot w_j\right)=\sum_{b=1}^{n}\alpha_{i,b}w_{j,b}. \quad (2)$$

Note that in n -dimensional space of vectors theoretically there is a set of different orthogonal bases [36]. Considering the fact that the Walsh functions formally take the value {1,-1} and do not directly enter the space B^n, reconstruction (1) actually reflects the transformation $B^n\rightarrow C^n$ into space C^n vectors with integer components. Just at the same time space $B^n\subset C^n$ is subspace for C^n space, therefore, it can be assumed that the transformations (1), (2) are carried out in C^n space. The normalization coefficient for the Walsh functions to the orthonormal system, which provides the inverse of the transformations (1), (2), is a constant for n, so it can be ignored when calculating the features.

As a result of the usage of component-wise transformation $z_i \rightarrow \alpha_i$ description Z will look as $\alpha(Z) = \{\alpha_i\}_{i=1}^{s}$, i.e. keypoints descriptors' set $\{Z_i\}_{i=1}^{s}$, $z_i \in B^n$ is transformed into a set of $\{\alpha_i\}_{i=1}^{s}$, $\alpha_i \in C^n$, of the same size and number. The problem arises about the effectiveness of such a transformation in the implementation of classification. Transformations (1), (2) can be written in vector form using the Hadamard matrix A [12]. The complete set of Walsh functions of n dimension forms an orthogonal Hadamard matrix A in size of $n \times n$, which consists of vectors of n Walsh functions vectors $W_1, \dots, W_n$. We apply the decomposition of the description in the basis of a family of rectangular basis functions-discrete Walsh functions, which are vectors of integers (value +1,-1) of finite dimension (power of two).

We regard description Z as the rectangular matrix $Z = \{z_{i,j}\}$, $i = \overline{1,s}$, $j = \overline{1,n}$ in size of $s \times n$, whose rows contain s descriptors. Transformation Z can be given as the multiplication of rectangular and square matrices by (3).

$$U = Z \times A. \tag{3}$$

As a result of (3), we get a rectangular matrix $U = \{u_{i,j}\}$, $i = \overline{1,s}$, $j = \overline{1,n}$ in size of $s \times n$, the rows of which contain expansion vectors α for Z description descriptors. Taking into account that in our research the set of Walsh functions is considered and applied in full, we will not analyze the variety of ways to arrange Walsh functions systems (Walsh, Hadamard, Peli, and Trachtman), each of which has a number of its own properties and corresponding areas of application [17]. Traditionally, classification consists of establishing the degree of relevance $\Theta(Z, E_k)$ between the description $\{Z_i\}_{i=1}^{s}$ of the analyzed image and the components $E_k = \{e_i(k)\}_{i=1}^{s}$ of the database $E = \{E_k\}_{k=1}^{N}$ etalon descriptions [5], [9], [13]. The class m of the object is determined according to the extremum of the function $\Theta(Z, E_k)$ as shown in (4).

$$m = arg \underset{k=1,\dots,N}{extr} \Theta(Z, E_k). \tag{4}$$

After the transformation, the classification process will be based on the definition of the modified relevance measure $\Theta^{\alpha}(\alpha(Z), \alpha(E_k))$ for transformed descriptions $\alpha(Z)$ and $\alpha(E_k)$ in a new feature space. Each of the transformed descriptions is a set of vectors with C^n components. Seeing the available possibility of data reduction in the newly created space, we will also have in mind some reduced space C^{α}, obtained from C^n by reducing the number of applied Walsh functions to $q << n$. Paying attention to the limited range of values of the input signal and the controlled transformation in the space of integers, when building a classifier, it is possible to accurately estimate the range of values of functions Θ, Θ^{α} and use this knowledge for the classification.

One of the options either Θ or Θ^{α} is metric for sets of vectors. This can be the Tanimoto (Jacquard) distance for sets A, B, which contains the ratio of the number of elements of a symmetric difference and the union of sets [2] by (5).

$$T(A, B) = \frac{card(A \Delta B)}{card(A \cup B)}. \tag{5}$$

Measure (5) reflects the quantitative characteristics of equivalent and different elements of the compared sets. When applying the metric (5), its limited value is important, for which the elements of the vector space are considered equivalent [2], [14]. This problem is one of the key ones in multidimensional data spaces. The choice of the equivalence threshold significantly affects the classification result. We offer two ways to solve it. One of them determines the limit value as shown in (6).

$$\rho_{\lim}(a,b) = \alpha \rho_{\max}(a,b), \tag{6}$$

as percentage α of the theoretically determined maximum of the metric. For example, for Hamming metric, which for vectors in size of 256 bits varies in the interval [0, 256], with a percentage of $\alpha = 25\%$ it is possible to determine $\rho_{lim}(a,b) = 0.25 \times 256 = 64$. The method of obtaining $\rho_{lim}(a,b)$ is more practical according to the result of the analysis of applied experimental data, where ρ_{max} is calculated for a set of etalons. Calculation time ρ_{max} does not affect the costs for classification and it is carried out at the previous stage of data analysis.

We will evaluate the performance of the classification method by the accuracy indicator pr, which is calculated by the ratio of the number of correctly classified objects r_p to the total number r that was used in the experiment [2].

$$pr = r_p/r. \quad (7)$$

Immunity is also an important indicator of the effectiveness of image recognition methods. It is characterized by the value of the accuracy of classification in the conditions of interference [1], [9], [18]. If the effect of additive interference on the image $B(x,y)$ is described by the model $B_\xi(x,y) = B(x,y) + \xi(x,y)$, and the interference $\xi(x,y)$ is characterized by the root mean square deviation σ with zero mathematical expectation, then the signal-to-noise ratio is described as $\mu = B_m/\sigma$, where B_m – amplitude characteristic (for example, the average brightness value). An indicator of immunity is dependence $pr(\mu)$ for the developed method.

2.3. Reduction in modified space

Considering the fact that the definition of the transformed feature system is based on the integration process of decomposition of the input description data according to the orthogonal spectrum, in space C^n there is a possibility of forming a compact presentation of $\alpha^*(Z)$ and $\alpha^*(E_k)$ with a shortened volume of coefficients. Thus, the transformation to reduced space $C^n \to C^\alpha$ is carried out. Reducing the dimensionality of the data space implements the promising idea of reducing the volume of calculations for Θ^α by reducing the size of the feature vector and reducing the time of classification [6], [9], [12]. One of the traditional ways consists in selecting from among the set $\{\alpha_i\}_{i=1}^{s}$ the most informative subset $\{\alpha_i\}_{i=1}^{q}$, $q << s$, which at the same time provides a sufficiently high-quality classification.

The same processing could be attempted to be done directly for the input set $\{z_i\}_{i=1}^{s}$ features, for example, according to the established informativeness criterion [37]. However, direct reduction can lead to a significant loss of information in proportion to the volume of the modified description. At the same time, the calculated projections $\{\alpha_i\}_{i=1}^{s}$ not only have a multispectral sense of data but also contain integrated information to a greater extent due to the decomposition process (1) according to the system of basis functions; therefore, they can be successfully used as a basis for productive data reduction.

Reduced subset $\{\alpha_i\}_{i=1}^{q}$ formation can be performed unconventionally by selecting a compact part among the functions of the system W, that is, using only a subset for classification W_q from W, $W_q \subset W$. In this case, there is no need to form a complete system of decomposition coefficients for both the input image and etalon set E. Such processing is popular in spectral methods, whereby filtering insignificant components of the spectrum due to some increase in the signal recovery error, the redundancy of the representation is reduced and the vector of the processed data is shortened [6], [17]. At the same time, the performance indicators of the systems decrease slightly.

If the filtering procedure succeeds in selecting a highly informative part of the spectrum that ensures high-quality classification, then it is possible to significantly reduce the computational costs of classification while ensuring the required level of effectiveness. The subset of the most informative part of the spectrum of Walsh functions obtained as a result of sequestration is the result of training on a given set of classes; its formation directly depends on the analyzed data space and is adapted to it. For other variants of etalon data collections, the subset of the spectrum components may have a completely different composition. Set the problem by transforming it into a new space based on mapping T: $\alpha(Z) \to \alpha^*(Z)$, or actually T: $W \to W_q$, due to the application of orthogonal decomposition according to a compact system W_q of Walsh functions to form shortened descriptions $\alpha^*(Z)$ and $\alpha^*(E_k)$, which provide sufficient indicators of the classification effectiveness. Such a reflection $T = T(E,\ W)$ is a function both E and a composition of the orthogonal system of functions W.

Due to the presence of significant correlations between the elements of the image of real nature, which finds its reproduction in the descriptors of the keypoints, the main energy in the discrete spectrum tends to be concentrated in a relatively small number of samples corresponding to slowly oscillating basis functions. Therefore, without significant damage to image restoration or classification, small spectral coefficients can be completely zeroed (or their analysis discarded), and significant spectrum elements can be used in the recognition process [17]. Dispersion analysis of orthogonal transformation coefficients is currently the main tool to evaluate their significance in the signal representation model (1) since the root mean square error of recovery depends on dispersion characteristics [17], [18]. Therefore, dispersion is a natural criterion for choosing a set of significant expansion coefficients. Analyze the curve of normalized variances depending on the number of the Walsh function, placed in descending order, and introduce a subset of the largest coefficients into the classification.

The application of the proposed apparatus in the classification task, where image descriptions are provided by a set of descriptors, has its features and possibilities [19]. The fact is that the coefficients with the largest variances do not always provide the required level of image resolution. The second factor is the possibility of performing dispersion analysis of data both within the full base of etalons and within individual representatives of it, which can affect the efficiency. An additional possibility is a dispersion analysis for the system of fragments of descriptive description, which implements window processing [12].

For sets E descriptions of etalons, we get matrix (3) in size of $Ns \times n$, calculate for each column of the matrix U vectors of mathematical expectation Mu_i and variance σ_i^2.

$$Mu_i = \frac{1}{Ns}\sum_{j=1}^{Ns} u_{i,j}, \sigma_i^2 = \frac{1}{Ns-1}\sum_{j=1}^{Ns}(u_{i,j} - Mu_i)^2, \tag{8}$$

when determining indicators (8) for a separate etalon, we accept $N = 1$. Rank the list of Walsh functions w_i by the value of the variance σ_i^2 for spectrum components: $w_1, w_2, \ldots, w_n$, so that the condition $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_n^2$ to be met. Use the resulting list as the basis of the reduction procedure.

2.4. Formalism of the distance matrix for a set of descriptors

Accept $A = \{a_i\}_{i=1}^{s}$, $B = \{b_j\}_{j=1}^{s}$ are two finite sets n -dimensional vectors with equivalent power s, $A \subset R^n$, $B \subset R^n$, $card\ A = s$, $card\ B = s$, R^n is the space of numerical vectors of dimension n. The choice of equal power somewhat simplifies the analysis of sets and can be achieved by directed adjustment of their power. Sets $A,\ B$ of descriptors of the keypoints for images are in practice classified as multisets, as they often contain close and even equivalent elements [5], [13].

Consider the formalism of the distance matrix M in size of $s \times s$ for the composition of elements of two sets $A,\ B$.

$$M[A, B] = \{\{m_{i,j}\}_{i=1}^{s}\}_{j=1}^{s}, m_{i,j} = \rho(a_i, b_j), \tag{9}$$

where $\rho(a_i, b_j)$ – the distance in space R^n, $a_i \in A$, $b_j \in B$. For binary vectors as ρ the computationally efficient Hamming distance [2] can be applied. Formalism (9) specifies a metric relation Ω for elements of a pair of sets: $\Omega[A, B] \to M$, where each pair of elements $a_i \in A$, $b_j \in B$ is assigned a value of ρ. As we know, the distance between objects is directly related to the probability of their equivalence in the metric-statistical theory of image classification [7], [27].

Each i -row of the matrix M corresponds to i -component of the description and contains a set of distances from i -element of the set A to elements from B. Given the symmetry property for the metric $\rho(a_i, b_j) = \rho(b_j, a_i)$, the square matrix M is symmetric. However, in the classification problem, one of the sets is considered an etalon (for example, A); therefore, it is advisable to analyze rows or columns M [34]. A special case for M is the relationship $\Omega[A, A]$, when the sets match. Then the matrix M will display the range of distance values for the elements of the set A, which can also be a significant feature for classification. An important distinction for classification is that the matrix M preserves invariance to geometric transformations of the image. This property directly follows from the invariance of descriptor values. If the value of the descriptor is preserved during geometric transformations, then the distances between the descriptors also do not change. It is clear that in real conditions both the invariance of the descriptors and the invariance of the distance matrix are fixed and verified approximately.

Based on the distance matrix M with the introduction of limit values for the metric for the equivalence of two vectors, such experimental results of operations on sets as intersection, union, difference, and symmetric difference can be determined [1], [7], [36]. The power of the sets obtained as a result of these actions is also a classification feature. If the elements a_i, b_j in the sample aspect are considered equivalent $(a_i \approx b_j)$ when fulfilling some boundary condition for the distance between them by (10).

$$a_i \approx b_j \mid \rho(a_i, b_j) \leq \delta_\rho, \tag{10}$$

where δ_ρ–given limited value for ρ, then the intersection $A \cap B$ can be defined as a set of elements A, for which in the corresponding row of the matrix M at least one condition (10) is fulfilled. The determination of the intersection, difference, and union of vectors' sets is required, for example, to calculate the value of the Tanimoto distance (5).

One of the direct applications of the distance matrix apparatus consists in the reduction (selection) of the most effective elements of sets for classification based on the introduced criterion [2]. At the same time, the goal of shortening the description while maintaining the necessary effectiveness of the classification is realized. For binary vectors-descriptors of keypoints, according to research results, two vectors can be considered equivalent to each other within the hamming distance of 25% of the maximum distance. For example, for accelerated-KAZE (AKAZE) binary vectors [25] with 488 dimensions, two descriptors are considered equal if the Hamming distance for them does not exceed the value $\delta_\rho = 122$.

Consider some base from N etalon images in the form of set E descriptions: $E = \{E_1, E_2, \ldots, E_N\}$. $E-$ this is an educational sample, which is at the same time the basis for classification by comparison with the etalon [8]. Each etalon description $E_k \subset E$ in the formalism of the classifier represents a separate class. Description of the etalon $E_k = \{e_v(k)\}_{v=1}^{s}$ – a finite set of keypoint descriptors, $e_v(k) \in B^n$, $s = card\ E_k$ – the number of descriptors in a set. Each descriptor $e_v(k)$ of the base E has a parameter k of the class number, and the total number of features-descriptors in the base set E is $card\ E = sN$.

Based on the values of the distance matrices $M[E_k, E_k]$ and $M[E_k, E_j]$, $\forall j \neq k$ it is possible to determine the value of the informativeness criterion for the base elements E and to effectively reduce etalon descriptions to reduce computing costs (classification time) [13], [27]. Descriptors with a low value of the criterion are filtered out as insignificant. Classification is based on a subset of informative description descriptors. Reduction is carried out at the stage of data preprocessing, and its implementation does not directly affect the time spent during classification.

2.5. Classification models based on the distance matrix

The matrix of distances between the components of the description is a component of clustering procedures in the feature space with the further construction of a classifier model within the framework of the “bag-of-words” technology [2], [15], [35]. As a rule, clustering is carried out for set E, each etalon E_k gets a modified description $E_k = (d_1, \ldots, d_u)_k$ in the form of a quantitative representation based on the system of u formed clusters. The classification is carried out by comparing the vector of the cluster representation of the object and etalons or by competitively assigning the elements of the description of the object to the cluster centers of the database E [15].

The application of the distance matrix is the calculation of significant classification characteristics of the description. Such a parameter is the medoid of the set, which is used as a center of description in the classification task. A medoid is an element of the set with the smallest total distance to the rest of the elements [2], [13]. The value of the medoid and ranking elements by aggregated distance is included in the construction of modifications of classifiers based on a shortened description [13].

Medoid $med\ A$ for a set A based on the matrix $M[A, A]$ is defined as the descriptor with the smallest value of the sum in the row (column) of the matrix as shown in (11).

$$z = med\ A | z \in A, z = arg\ \min_{i} \sum_{j=1}^{s} m_{i,j}. \tag{11}$$

The value of matrices $M[A, A]$, $M[A, B]$, $B \neq A$, $B \subset E$ will be used as a basis for evaluating the effectiveness of the classification within the framework of the database E. This can be done using parameters such as classification accuracy concerning component composition. The accuracy of classification (7) is estimated by the ratio of the number of correctly identified components of the set A to their number s. The fact that the classification is correct for $z \in E_k$ is established by the predicate $right(z)$.

$$right(z, M) = 1 | z \in E_k, k = arg\ \min_{i,j} \{m_{i,j}\}_{j=1}^{Ns}, \tag{12}$$

i.e. if the minimum in the row of distance matrices for the etalon base E is achieved for the etalon to which the analyzed descriptor belongs a priori. Based on this preliminary criterion analysis of the set of distance matrices $M[E_k, E_j]$, $j, k = 1, \ldots, s$ formally, it is possible to establish a guaranteed level of classification efficiency for any newly created system of features, without conducting detailed experimental research. Based on the distance matrix, the analysis of the etalon database may be the basis for the need to improve the selected system of features to ensure the desired level of performance.

Consider a set of distance matrices for the etalon base. If images $M[Z, Z]$ of analyzed objects Z or etalon classes E_k are set in the form of matrices $M[E_k, E_k]$, $k = \overline{1, s}$ then the degree of relevance for a pair of descriptions can be represented as establishing a new relation between metric relations displayed by matrices $\{M[E_k, E_k]\}_{k=1}^{N}$ and $M[Z, Z]$. Such a relationship can be created based on statistics of values $M[E_k, E_k]$, $k = \overline{1, s}$. It is known that the use of distributions instead of data values significantly reduces computational costs for the classification [8], [10], [18]. Considering the known range of the metric value, construct a histogram (distribution) for the matrix as the number of elements that take a fixed value x distance.

$$q_k[x] = card\{m_{i,j} | m_{i,j} = x\}. \tag{13}$$

Now the classification can be carried out in the space of vectors with integer components, to which the values of the histograms $q_k[x]$ belong. Specifically, for the Hamming metric applied to binary descriptors in size of 512, we have the range of the histogram argument $x \in [0, \ldots, 512]$. We will carry out the classification based on the metric-statistical approach by calculating and optimizing on a set of etalons for the distance (for example, Manhattan distance) between the distributions of the object and the etalon by (14).

$$\rho(q_Z[x], q_k[x]) = \sum_x | \ q_Z[x] - q_k[x]|, \quad (14)$$

where $q_Z[x]$ – a histogram of distance matrix values for an object.

Taking into account that the order of descriptors in the description changes during geometric transformations, the degree of relevance of two matrices can be established only by integral criteria, including distributions. In addition, the classification criterion can be, for example, a vector of sums $r_j = \sum_{i=1}^{s} m_{i,j}$, $j = \overline{1,s}$ for the values of the columns (rows) of the matrix $\{\{m_{i,j}\}_{i=1}^{s}\}_{j=1}^{s}$, since its components retain their invariance until the order of the elements is changed [34].

For a classifier based on elemental analysis of the description, for an arbitrary descriptor $z \in Z$ for an arbitrary object descriptor, we obtain a set of distance matrices for the complete set of etalons.

$$M(z) = \begin{Bmatrix} m_{1,1} & \dots & m_{1,s} \\ m_{2,1} & \dots & m_{2,s} \\ & \dots & \\ m_{N,1} & \dots & m_{N,s} \end{Bmatrix}, \quad (15)$$

where the first index is the etalon number, and the second is the descriptor number as a part of the etalon. The list (15) can be the basis for building a classifier K, which refers to the parsed descriptor Z of one of the classes $K: z \rightarrow \{1, \dots, N\}$.

In general, a class k for descriptor $z \in Z$, is defined as the argument of the optimum on the set of classes for the function F of matrix argument (15).

$$k = arg \ \underset{i=1,\dots,N}{opt} F[M(z)] \quad (16)$$

the result is determined by the model F. In (16), the determination of the minimum or based on the entire list in size of $N \times s$ can be organized, or separately for each of the etalons (row) with the following definition of the winning class.

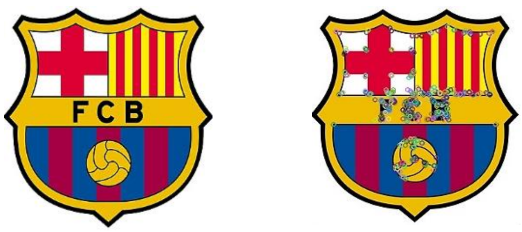
Enter vector $\{h_i\}_{i=1}^{N}$ with integer values to accumulate class votes. Based on implementation K for each descriptor $z \in Z$ according to (16), we determine the class number k, and then increment the battery $h_k = h_k + 1$ to the class number. According to the result of processing the description Z object, we accumulate a vector $\{h_i\}_{i=1}^{N}$. Class Z is defined as the argument of the maximum number of votes by (17).

$$k = arg \ \underset{i=1,\dots,N}{max} h_i || h_k \geq \sigma_h, \quad (17)$$

where $\delta_h -$ threshold for the minimum number of votes, which is set experimentally for a given base. If the inequality in (17) does not hold, the class of the object is not established (refusal of the classification).

3. RESULTS AND DISCUSSION

To study the efficiency of the proposed classification methods (efficiency, speed, immunity), computer simulation was carried out. To research the method of classification based on features of orthogonal decomposition, a base of five images of football club emblems was selected. Images have a size of 325×325 pixels (joint photographic experts group (JPEG) format, Figure 2). The etalon is located on a light background as shown in Figure 2(a). Using the Python programming system and the open source computer vision (OpenCV) library of computer vision algorithms [8], [23], [38], [39], 500 oriented FAST and rotated BRIEF (ORB) descriptors in size of 256-bit were generated for each image as shown in Figure 2(b).



(a) (b)

Figure 2. An example of (a) an etalon and (b) its image with coordinates of keypoints

The range of values of the received transformed data is fixed, which makes it possible to evaluate the classification indicators. The numbers of the first 16 sorted Walsh functions by the value of the square of the variance on the set of descriptors for the etalon base are as follows: 0, 45, 106, 208, 176, 10, 85, 77, 109, 80, 209, 9, 3, 54, 72, 197. The computer time required to calculate the Tanimoto metric between the coefficients of the orthogonal representation was 29 s, and with the representation in the space of 16 selected Walsh functions–1.7 s, respectively. As we can see, the use of a reduced feature space reduces computational costs by 17 times. At the same time, both classifiers (traditional and modification) provide error-free classification for the training sample.

Apply the Tanimoto metric for sets in the classifier for the transformed data space using a threshold ρ_{lim} of 25% of the maximum metric value for etalon data descriptors. At the same time, the maximum value of the metric, which means the threshold ρ_{lim}, depends on the implemented data space. The influence of additive noise on classification indicators was researched. Note that the Tanimoto distance between the etalon and the noisy etalon gradually increases with an increase in the noise level (decrease μ), and this affects the classification accuracy. An experimental study of the interference resistance of the developed modifications of the classifier using the Tanimoto metric (5) and the selected threshold ρ_{lim} under the influence of additive noise showed that indicator $pr(\mu)$ for the method with a full set of 256 Walsh functions, as well as for the method with a reduced representation of 16 Walsh functions, equals 1 (classification without errors) at $\mu \geq 1.2$ and decreases to 0.9 at $\mu = 1$.

For the traditional method (without the application of the Walsh functions apparatus), the interference resistance indicator $pr(\mu)$ is almost within the same limits. For the method with a reduced representation of 8 Walsh functions, the value $pr(\mu)$ decreases to 0.9 at $\mu = 1.2$. We see that the developed method is not inferior to the traditional one in terms of interference resistance. In general, we have quite high noise immunity indicators for applied systems [27], as the proposed method works without errors even with a low signal-to-noise ratio close to 1! At the same time, the time of classification in modified data spaces is reduced in proportion to the number of Walsh functions used. The experiment showed a reduction in time costs in comparison with the full set of 256 Walsh functions for 16 Walsh functions by 19 times, for 8 Walsh functions– by 30 times as shown in Table 1. At the same time, the classification time for the traditional method by calculating the Tanimoto metric for descriptions without transformation is 12 times higher than for the modification of 16 Walsh functions.

Table 1. Experimental estimation of classification time, *s*

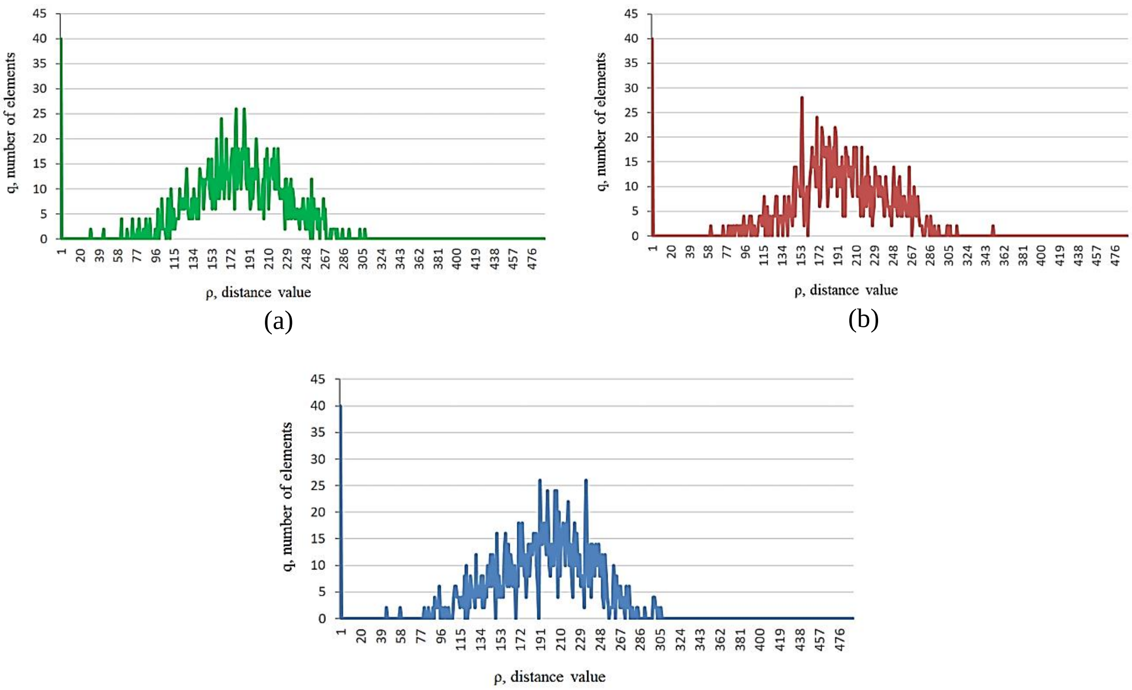
Method	256 Walsh functions	16 Walsh functions	8 Walsh functions	Traditional
Time	228	12	7.5	140

To simulate the classifier using the distance matrix apparatus, the KAZE detector in the improved version of AKAZE, which forms descriptors of 488 bits [25], [34], was selected. The software was created based on the Java programming language and the OpenCV library of computer vision algorithms, 40 AKAZE descriptors of the machine learning database (MLDB) type were generated for each image, which makes it possible to change the size of the descriptor [25], [26]. The experiments were carried out for three classes of images of animals (horses) in size: 300×200 pixels as shown in Figure 3. An example of an image Figure 3(a) and selected coordinates of the keypoints is shown in Figure 3(b).

Because the amount of data in the distance matrix depends on the number of generated keypoints, for modeling, the number of the keypoints $s = 40$ in each of the three etalons was. The choice of the keypoints is a parameter that affects the classification result since usually different numbers of the keypoints are obtained for different etalons. Specifically, in the experiment, 40 points from each etalon were randomly taken. We constructed experimental distributions (13) of distance values separately for each of the etalons, which are shown in different colors in Figure 4 (the result for the first etalon is shown in Figure 4(a), for the second etalon in Figure 4(b), for the third etalon in Figure 4(c). As you can see, visually the distributions are quite different, although almost all their values are in the limited range of 0-320. The Manhattan distance (14) is: etalons (1-2) – 888, (1-3) – 1016, (2-3) – 988. Taking into account the fact that formally the range of distance change (14) is 0-3200 ($2s^2$), it can be argued that the integral features of distributions (13) can be implemented for classification since the difference between them for different etalons reaches 30% of the maximum possible. At the same time, the average distance within each of the descriptions has a value of 182, 190, 194. If we focus on these values, the relative weight of the obtained distances in percentage terms will be even more significant.

(a) (b)

Figure 3. An example of (a) an image and (b) the coordinates of the generated keypoints



(a) (b) (c)

Figure 4. Histograms of values of distance matrices for (a) the first, (b) second, and (c) third etalons

Now analyze the enlarged quantized representation for the etalon matrices in the interval system of the discrete ranges 0-50, 51-100, ..., 401-488, and get an effective reduced dimension of 10 for each of the distributions $q_k[x]$ (Figure 5). In Figure 5, the numbers on the X-axis indicate the rightmost boundary of the discrete range. The histograms in Figure 5 are obtained by counting the number of values within the range in Figure 4. In this quantized space, you can visually see a significant difference between the analyzed distributions within the range 100-300 and approximately the same values for the classes in the remaining ranges. This makes it possible to focus on the implementation of analysis and calculations only in a shortened range of distances of 100-300. We experimentally verified that this fact is also valid for other images, which expands the possibilities of simplification to speed up calculations in the classification process.

In Figure 5 we see confirmation of the well-known fact of significant proximity to each other in the space of multidimensional vectors, which was noticed by researchers earlier [2]–[4]. In addition, in Figure 5, one can visually notice a significant difference in the description of image 3 from the others, while for images 1 and 2, some proximity is observed. In metric values, it was: (1-2) – 216, (1-3) – 620, (2-3) – 408. Naturally, the distances between images in the quantized space decreased somewhat, but the principle ability to distinguish between images of several classes remained.

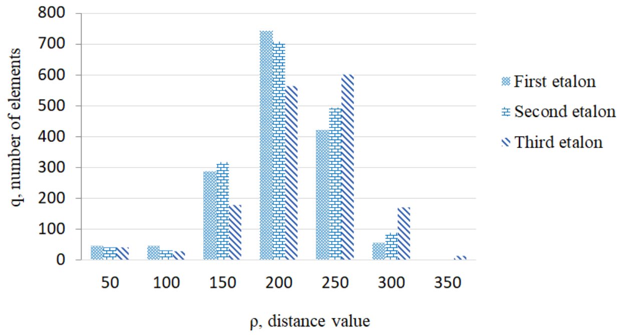


Figure 5. Histograms of distance matrix values in discrete ranges

The experimentally calculated distances between the sum vectors $\{r_j\}_{j=1}^{s}$ for the columns of the etalon distance matrices were: (1-2) – 30524, (1-3) – 44452, (2-3) – 28694. Considering the average experimental values of the distances for the etalons in the range of 182-194, we can consider the obtained vectors of sums to be sufficiently significant classification features. Further integration of the distance values by expanding the ranges showed that the differences between the etalon distributions naturally decrease. For 5 equal ranges, the classification of the analyzed images is complicated.

The experimental calculation of accuracy (7) of the classification showed that for a set of descriptors of the training sample (120 vectors in the researched etalons), all the descriptors are classified correctly. Thus, the accuracy index about the set of etalon components is equal to 1. In the distance matrices, zeros are found only for the etalon itself. This can be explained by the high dimensionality of the data (488), which practically excludes random coincidences of bits in multidimensional vectors. At the same time, it is clear that for real situations under the influence of obstacles or geometric transformations of the image, the accuracy index may naturally be somewhat lower. Taking into account the fact that the experimentally calculated Tanimoto distance for the descriptions used in the experiment was 0.5-0.6, we can see that, despite the significant similarity of the structural descriptions of the analyzed images, the proposed simplified classification models provide a sufficient level of their distinction in practice, and the speed of processing for them ten times higher.

4. CONCLUSION

Utilizing an orthogonal function system for image description transformation shows the potential for a substantial increase in processing speed while maintaining high levels of classification accuracy and resistance to interference. The critical factors for optimizing the effectiveness of this representation involve selecting a metric to map the modified descriptions and determining a threshold to determine the similarity of components in the newly generated data space. The utilization of the Tanimoto measure, along with a threshold equivalent to 25% of the maximum metric value, has demonstrated its effectiveness in evaluating description components. The utilization of the Walsh function apparatus not only resulted in a tenfold reduction in computing expenses but also maintained relatively good indicators of classification efficiency.

By implementing the distance matrix model, it became possible to create efficient integrated features in the form of one-dimensional data distributions and vectors representing the sum of the matrix columns. This resulted in reduced computational expenses while maintaining the classification effectiveness of the original data sample. A software simulation was conducted to experimentally evaluate the effectiveness of image classification in newly created feature spaces and the time required to calculate the relevance of descriptions. This was compared to the traditional approach of voting and calculating metrics on a set of descriptors.

The research introduces a novel approach to image classification by enhancing the structural method. This is achieved through the implementation of description transformation using orthogonal decomposition of data and the construction of feature models based on distance matrices in the descriptor space. Additionally, the research proposes methods to reduce descriptions in newly created spaces, resulting in reduced computational costs for classification. The practical significance of this work lies in the development of classification models in the transformed data space, verification of the effectiveness and resilience of the proposed modifications using image examples, and the creation of software applications to implement the developed classifiers in computer vision systems. Research perspectives may pertain to developing a range of proposed models for the creation and examination of changes in descriptions inside extensive databases.

ACKNOWLEDGEMENTS

The work is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under project No. 09I03-03-V01-00115. The authors acknowledge the support of the Department of Informatics, Kharkiv National University of Radio Electronics, Ukraine, in numerous help and support to complete this article.

REFERENCES

[1] R. Scherer, "Computer vision methods for fast image classification and retrieval," *Studies in Computational Intelligence*, pp. 107–118, 2020, doi: 10.1007/978-3-030-12195-2.

[2] P. Flach, "Model ensembles," in *Machine Learning*, Cambridge University Press, 2012, pp. 330–342.

[3] C. D. Manning, P. Raghavan, and H. Schutze, "Web search basics," *An Introduction to Information Retrieval, Cambridge, England: Cambridge University Press*, pp. 421–441, 2009.

[4] Y. Nong, "Algorithms for mining classification and prediction patterns," *Data Mining: Theories, Algorithms, and Examples, Boca Raton, FL, USA: CRC Press*, pp. 21–136, 2013.

[5] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, and M. Zeghid, "Tools for fast metric data search in structural methods for image classification," *IEEE Access*, vol. 10, pp. 124738–124746, 2022, doi: 10.1109/ACCESS.2022.3225077.

[6] S. M. Alessio, *Digital signal processing and spectral analysis for scientists*. Springer International Publishing, 2016.

[7] W. K. Pratt, "Unitary transforms," *Digital Image Processing*. Wiley, pp. 189–216, Jun. 2007, doi: 10.1002/9780470097434.ch8.

[8] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, S. Gadetska, and M. Al-Dhaifallah, "Methods of classification of images on the basis of the values of statistical distributions for the composition of structural description components," *IEEE Access*, vol. 9, pp. 92964–92973, 2021, doi: 10.1109/ACCESS.2021.3093457.

[9] N. V. Vlasenko and O. V. Sytnik, "Classification of video-objects in attribute space of the walsh functions," *Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika)*, vol. 72, no. 19, pp. 1777–1785, 2013, doi: 10.1615/TelecomRadEng.v72.i19.60.

[10] S. V. Gadetska, V. O. Gorokhovatskyi, N. I. Stiahlyk, and N. V. Vlasenko, "Statistical data analysis tools in image classification methods based on the description as a set of binary descriptors of key points," *Radio Electronics, Computer Science, Control*, no. 4, pp. 58–68, Jan. 2022, doi: 10.15588/1607-3274-2021-4-6.

[11] S. N. Prajwalasimha, S. R. Kavya, and T. Z. Ahmed, "Design and analysis of pseudo hadamard transformation and non-chaotic substitution based image encryption scheme," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 3, pp. 1297–1304, Sep. 2019, doi: 10.11591/ijeecs.v15.i3.pp1297-1304.

[12] N. Ahmed and K. R. Rao, "Walsh-hadamard transform," in *Orthogonal Transforms for Digital Signal Processing*, Springer Berlin Heidelberg, 1975, pp. 99–152.

[13] V. A. Gorokhovatskiy, "Compression of descriptions in the structural image recognition," *Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika)*, vol. 70, no. 15, pp. 1363–1371, 2011, doi: 10.1615/TelecomRadEng.v70.i15.60.

[14] A. V. Gorokhovatsky, V. A. Gorokhovatsky, A. N. Vlasenko, and N. V. Vlasenko, "Quality criteria for multidimensional object recognition based upon distance matrices," *Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika)*, vol. 73, no. 18, pp. 1661–1670, 2014, doi: 10.1615/TelecomRadEng.v73.i18.50.

[15] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, and M. Zeghid, "Cluster representation of the structural description of images for effective classification," *Computers, Materials and Continua*, vol. 73, no. 3, pp. 6069–6084, 2022, doi: 10.32604/cmc.2022.030254.

[16] R. Walsh, I. Osman, and M. S. Shehata, "Masked embedding modeling with rapid domain adjustment for few-shot image classification," *IEEE Transactions on Image Processing*, vol. 32, pp. 4907–4920, 2023, doi: 10.1109/TIP.2023.3306916.

[17] K. Premkumar, H. Twinky, S. Naveen, M. Kathiresan, and R. Aravind, "A secured communication in watermarking using fast walsh hadamard transformation in image processing," in *6th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2022 - Proceedings*, Dec. 2022, pp. 1186–1191, doi: 10.1109/ICECA55336.2022.10009348.

[18] L. Shapiro and G. Stockman, "Pattern recognition concepts," *Computer Vision, Upper Saddle River*, pp. 107–136, 2001.

[19] V. Gorokhovatskyi and I. Tvoroshenko, "Image classification based on the Kohonen network and the data space modification," *CEUR Workshop Proceedings*, vol. 2608, pp. 1013–1026, 2020, doi: 10.32782/cmis/2608-76.

[20] V. D. Babu and K. Malathi, "Large dataset partitioning using ensemble partition-based clustering with majority voting technique," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 2, pp. 838–844, Feb. 2023, doi: 10.11591/ijeecs.v29.i2.pp838-844.

[21] Y. Liu, Y. Huang, S. Wang, W. Lu, and H. Wu, "Modality coupling for privacy image classification," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 4843–4853, 2023, doi: 10.1109/TIFS.2023.3301414.

[22] A. J. M. S. Arockiam and E. S. Irudhayaraj, "Reclust: an efficient clustering algorithm for mixed data based on reclustering and cluster validation," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 1, pp. 545–552, Jan. 2023, doi: 10.11591/ijeecs.v29.i1.pp545-552.

[23] J. Nunez-Iglesias, "ORB feature detector and binary descriptor." https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_orb.html (accessed Sep. 14, 2023).

[24] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proceedings of the IEEE International Conference on Computer Vision*, Nov. 2011, pp. 2564–2571, doi: 10.1109/ICCV.2011.6126544.

[25] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE features," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7577 LNCS, no. PART 6, Springer Berlin Heidelberg, 2012, pp. 214–227.

[26] O. Yakovleva and K. Nikolaieva, "Research of descriptor based image normalization and comparative analysis of surf, sift, brisk, orb, kaze, akaze descriptors," *Advanced Information Systems*, vol. 4, no. 4, pp. 89–101, Dec. 2020, doi: 10.20998/2522-9052.2020.4.13.

[27] R. M. Tymchyshyn, O. Y. Volkov, O. Y. Gospodarchuk, and Y. P. Bogachuk, "Modern approaches to computer vision," *Upravlâûŝie sistemy i mašiny*, no. 6 (278), pp. 46–73, Dec. 2018, doi: 10.15407/usim.2018.06.046.

[28] S. A. K. Tareen and Z. Saleem, "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK," in *2018 International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET 2018 - Proceedings*, Mar. 2018, vol. 2018-Janua, pp. 1–10, doi: 10.1109/ICOMET.2018.8346440.

[29] M. Chaabi, M. Hamlich, and M. Garouani, “Product defect detection based on convolutional autoencoder and one-class classification,” *IAES International Journal of Artificial Intelligence*, vol. 12, no. 2, pp. 912–920, Jun. 2023, doi: 10.11591/ijai.v12.i2.pp912-920.

[30] N. M. Hai, T. T. Van, and T. V. Lang, “A method for semantic-based image retrieval using hierarchical clustering tree and graph,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 5, pp. 1026–1033, Oct. 2022, doi: 10.12928/TELKOMNIKA.v20i5.24086.

[31] R. H. Hridoy, A. D. Arni, and A. Haque, “Improved vision-based diagnosis of multi-plant disease using an ensemble of deep learning methods,” *International Journal of Electrical and Computer Engineering*, vol. 13, no. 5, pp. 5109–5117, Oct. 2023, doi: 10.11591/ijece.v13i5.pp5109-5117.

[32] R. H. Hridoy, T. Yeasmin, and M. Mahfuzullah, “A deep multi-scale feature fusion approach for early recognition of jute diseases and pests,” in *Lecture Notes in Networks and Systems*, vol. 436, Springer Nature Singapore, 2022, pp. 553–567.

[33] V. Gorokhovatsky and S. V. Gadetska, “Determination of relevance of visual object images by application of statistical analysis of regarding fragment representation of their descriptions,” *Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika)*, vol. 78, no. 3, pp. 211–220, 2019, doi: 10.1615/telecomradeng.v78.i3.20.

[34] V. Gorokhovatskyi, O. Peredrii, I. Tvoroshenko, and T. Markov, “Distance matrix for a set of structural description components as a tool for image classifier creating,” *Advanced Information Systems*, vol. 7, no. 1, pp. 5–13, Mar. 2023, doi: 10.20998/2522-9052.2023.1.01.

[35] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, and M. Al-Dhaifallah, “Classification of images based on a system of hierarchical features,” *Computers, Materials and Continua*, vol. 72, no. 1, pp. 1785–1797, 2022, doi: 10.32604/cmc.2022.025499.

[36] E. Zeidler, *Applied functional analysis: main principles and their applications*. Springer Science and Business Media, 2012.

[37] A. Oliinyk, S. Subbotin, V. Lovkin, O. Blagodariov, and T. Zaiko, “The system of criteria for feature informativeness estimation in pattern recognition,” *Radio Electronics, Computer Science, Control*, vol. 0, no. 4, pp. 85–96, Mar. 2018, doi: 10.15588/1607-3274-2017-4-10.

[38] P. F. Alcantarilla, J. Nuevo, and A. Bartoli, “OpenCV Java documentation (4.8.0-pre).” https://docs.opencv.org/4.x/javadoc/index.html (accessed Sep. 18, 2023).

[39] V. Gorokhovatskyi, I. Tvoroshenko, O. Kobylin, and N. Vlasenko, “Search for visual objects by request in the form of a cluster representation for the structural image description,” *Advances in Electrical and Electronic Engineering*, vol. 21, no. 1, pp. 19–27, May 2023, doi: 10.15598/aeee.v21i1.4661.