Received 13 January 2026, accepted 27 January 2026. Date of publication 00 xxxx 0000, date of current version 00 xxxx 0000.

Digital Object Identifier 10.1109/ACCESS.2026.3659757

Feature Space Quantization and Application of Metrics in Structural Methods of Image Recognition

VOLODYMYR GOROKHOVATSKYI[1], IRYNA TVOROSHENKO[1], OLENA YAKOVLEVA[1,2,3], AND MONIKA HUDÁKOVÁ[2]

[1]Department of Informatics, Kharkiv National University of Radio Electronics, 61166 Kharkiv, Ukraine
[2]Department of Economics and Finance, Economics and Management Institute, Bratislava University of Economics and Management, 851 04 Bratislava, Slovakia
[3]Department of AI Research and Development, SYTOSS Ltd., 831 04 Bratislava, Slovakia

Corresponding author: Iryna Tvoroshenko (iryna.tvoroshenko@nure.ua)

ABSTRACT This research aims to simplify the decision-making rules and reduce computational load when using structural methods of image classification that employ a description in the form of a set of keypoint descriptors. The main attention is paid to the implementation of the classifier and studying the properties of quantization tools for the basic set of descriptors of etalon descriptions. As a result of quantization, the etalons and the analyzed image are represented by a vector of numbers through projection onto a compact set of defined centers. This approach enables the use the metrics apparatus and provides a high-speed classification compared to the traditional voting method. Improvements to the method by introducing a threshold for assigning descriptors to the system's centers are discussed. The gain in classification speed for the considered data processing schemes increases in proportion to the number of etalons. The paper presents the results of software modeling of the proposed approaches for the experimental base, which includes images of historical monuments. The test sample is formed as a set of images from the etalon database and images outside the database with a set of geometric transformations of shift, scale, and rotation in the field of view applied to them. Practical issues of choosing threshold parameters to establish the equivalence of descriptors and to minimize the number of class votes to ensure the required levels of classification accuracy are researched. Testing has confirmed a significant acceleration of the classification process and a sufficiently high level of classification accuracy using quantization and the Tanimoto metric. In particular, a tenfold increase was achieved in the modeling performed. The peculiarities of forming an etalon base and choosing a range of parameters of geometric transformations for the proposed method were experimentally studied. The considered methods should be implemented in applied computer vision tasks with high requirements for the speed of classification of visual objects.

INDEX TERMS Computer vision, feature voting, image recognition, keypoint descriptors, quantization in multidimensional space, speed of classification, structural methods, Tanimoto metric.

I. INTRODUCTION. LITERATURE REVIEW

Computer vision systems are now an essential tool for achieving sustainable human development. They enable environmental protection through monitoring and control, increase the economic efficiency of processes through automation, improve the quality and safety of life, expand access to services, and reduce the need for human resources.

Simplification of decision-making procedures remains a pressing issue for modern intelligent systems. In this regard,

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Sharif.

the introduction of effective intellectual tools significantly expands the functional abilities of modern automated computer vision systems [1], [2].

The possibility of using the traditional mathematical apparatus provides a clearer formalization and justification of the analysis processes when creating effective classifiers in applied applications [3], [4].

It's important to note that the modern theory of pattern recognition is based on four basic principles of data analysis and processing that contributed to the creation and implementation of the following categories of developed methods [4]:

1) Methods based on decision theory.
2) Methods based on orthogonal transformations.
3) Structural methods.
4) Neural networks.

At the same time, the latest applied approaches often combine these categories and use an expanded arsenal of these principles in a combined form, relying on their positive properties for synthesized feature systems. For example, the structural method developed in the works [5], [6], [7] uses both the comparison with an etalon (decision theory) and the formation of an effective feature system (orthogonal transformations, neural networks).

Researchers consider the following shortcomings of modern neural networks to be the most important for the practice of visual object recognition [4], [8], [9]:

1) The need to develop effective learning algorithms.
2) The large dimensionality of networks and the need for significant amounts of memory.
3) Ensuring invariance to affine transformations and other deformations of visual images.
4) Lack of versatility (limited range of tasks for a particular network, dependence of the network architecture on the nature of the data and the type of task).

Therefore, designing an effective neural network to solve a specific practical problem is now considered to be primarily an art. Nevertheless, the neural networking apparatus is successfully developing. New templates are being introduced to transform the features of input images [10], which can be successfully used to train a neural network.

At the same time, the use of structural methods, where the description of objects is represented by a set of keypoint descriptors, avoids the above limitations and requires only a fixed etalon base and an effective classification method in a certain feature space [6], [7], [11].

In this case, the key task in classification structural methods based on the model of comparison with the etalon is to determine the similarity or difference for two sets of multidimensional vectors describing visual objects. In theoretical terms, this task is solved by determining the intersection power of the sets of vectors describing the recognized object and the etalon. However, the peculiarities of applied computer vision problems are that not only data vectors (keypoint descriptors) with the complete coincidence of components should be considered equivalent when making a decision, but also vectors with the distance between them within a given threshold [5].

The value of such a threshold is a parameter in establishing equivalence that directly affects the classification performance.

Thus, each data description vector should be considered together with a specific neighborhood that characterizes the degree of equivalence for a subset of vectors in a multidimensional space. Such neighborhoods can naturally overlap for different pairs, groups, or classes of vectors, which generally causes ambiguity and vagueness of decisions when implementing traditional methods of determining the intersection.

In addition, researches have shown that the use of a model with the formation of statistical centers for structural descriptions does not provide decent classification results [6], [11], since these descriptions for images often consist of descriptors that are close to each other.

This problematic situation can be solved by projecting the existing description in the form of a set of vectors into a new quantized data space, which can be obtained by clustering or another method of data quantization, for example, by the Kohonen network, hashing, etc. [6], [12], [13].

The data within a quantum are now considered equivalent to the center and to each other at the same time, which formally eliminates the ambiguity in determining equivalence for any pair of vectors [14]. In fact, this is the preliminary stage of data classification in quantized space. Data quanta here form separate classes for descriptors. This stage, in turn, makes it possible to apply a mathematically based metric apparatus to assess the relevance of descriptors.

The proposed modification of the image description can be considered as one of the ways to construct a transformed feature space, in which a relatively simple and effective separation of the analyzed images is possible in practice. This approach conforms to an important statement of pattern recognition theory that in order to successfully solve applied problems, it is necessary to strive for an effective reduction in the dimensionality of the space and simplification of the decisive rules [4], [6]. This idea echoes Leonardo da Vinci's famous statement "Simplicity is the highest refinement".

In the quantized space of multidimensional data, the mathematical apparatus of fuzzy sets or multisets can be used to evaluate the intersection of two descriptions in the form of sets [15], [16], [17]. In this case, the obtained data quantization centers formally create the base set for the multiset representation (or the domain of the fuzzy set definition).

The given etalon descriptions can now also be preliminarily represented in the form of multisets, i.e., integer vectors containing the number of elements of a fixed etalon in each cluster. And the classification process can be performed by comparing the description of the input object with the generated projections of the etalons or with the centers of the clusters to make a decision [6]. An additional important positive impact is that due to the quantization of the feature space, a significant acceleration of the classification

process is achieved since, in this way, it is possible to avoid computationally expensive procedures of linear search using the nearest neighbor model on the full set of etalon descriptors [12], [18], [19].

Data quantization has also found application in the problem of image classification by neural network methods with the use of a deep learning machine [18]. However, the effectiveness of such neural network approaches depends on the synthesized network structure, which is associated with the implementation conditions. When implementing deep learning, the problem of forming a compact set of features is also relevant.

As a result of the introduction of quantization, time-consuming procedures for voting for a multivalued set of descriptors of two sets are avoided, which is a variant of the similarity measure without preserving the symmetry property [4], [6]. In addition, along with the chance of a thorough application of metrics, it becomes possible to use a variety of other developed models of the metric space apparatus, for example, correlation coefficient, statistical criteria, etc.

An equally important factor is to ensure a decent level of performance of the recognition system under the influence of external factors: interference, background, geometric transformations [6], [11]. With the introduction of quantized data representation, it becomes possible to control the classification process by introducing a threshold for assigning to cluster centers as a parameter for filtering out false data and interference. The idea behind this control is to ensure better consistency with the a priori etalon data. Such a threshold should ensure positive perception of data similar to the etalon data and eliminate false image components.

Means of quantizing a set of features are now successfully used in applied computer vision tasks to reduce the amount of computation, including the use of neural networks [20], [21], [22].

Quantization strategies are being developed to improve the ability to express diversity in feature parameters [20]. Image search networks are being improved to reduce memory usage and decrease the density of image features in the database to optimize resources in practical applications [21].

Means of reducing computational complexity in deep convolutional neural networks (CNN) are being developed for their deployment on peripheral devices with limited resources [22].

It is worth mentioning that the development of high-dimensional data hashing methods in image recognition systems makes it possible to ensure computational efficiency with some loss of data semantics [23], [24], [25], [26], [27], [28], [29], [30]. The introduction of specific hash codes makes it possible to compare image and text descriptions while ensuring the necessary degree of hash representation differentiation in Hamming space [23], and the use of genetic algorithms helps to improve the representation of local features in image recognition [24]. Methods of controlled hashing are being developed using hash function training tools with subsequent quantization of the generated hash codes [25]. It should be pointed out that the method of direct training of hash codes slightly reduces the impact of quantization losses but leads to some increase in computational complexity [26]. A model of a quantized artificial neural network (QANN) has been developed that combines quantum feature extraction with classical neural network topologies to improve the compression of recognized images [27]. Attention is also paid to the problem of minimizing quantization errors based on the principle of relative data similarity, but a practical solution is still ahead [28].

When using satellite remote sensing for computer vision tasks related to environmental monitoring and digital environmental security, a dynamic binary visual transformer was developed, which proved to be effective in capturing complex patterns and subtle features in multispectral images [29].

According to researchers [30], the implementation of vector quantization methods for image recognition generally guarantees insignificant quantization errors. It also ensures high accuracy of sample analysis in precision farming and sustainable smart city development tasks.

The material of this paper develops the research of authors' previous works [5], [6], [7], [12], [18], [31], which are aimed at improving structural methods of image classification in terms of increasing their speed while ensuring a decent level of classification accuracy. The proposed improvement practically leads to a simplification of the applied classification rules. The main attention is paid to such aspects as the introduction of metrics in the classification process, the justification of introducing a threshold for filtering descriptors when assigning to the center system, evaluating the characteristics of the etalon database to guarantee a decent performance, comparing the accuracy and speed of classification for the proposed approach and the traditional voting method by means of software modeling.

The analysis reveals that balancing such key criteria as the complexity of implementation, accuracy and speed of implementation, sensitivity to data variations, and specificity or universality of use remains a challenging task for various applied classification models. There is hope that this research will make a contribution to solving this problem. The results of this study will be valuable for researchers in the fields of image recognition, multivariate data analysis, machine learning, and decision-making methods.

II. MATERIALS AND METHODS

A. FORMAL DEFINITION OF THE CLASSIFICATION TASK

The classification as part of a database of N etalons (representatives or prototypes of classes) will be performed, which is specified in the form of a finite set E of descriptions of etalon images: $E = \{E_1, E_2, \ldots, E_N\}$.

The set E is a training set, which is simultaneously the basis for the construction and implementation of the classifier, including comparing the descriptions of the analyzed objects and the etalons [6], [31].

Each etalon description E_k in the classifier formalism represents a separate class. The class k with a fixed etalon description E_k is formally understood as a certain infinite set of images obtained from the etalon image (prototype class with the number k) by applying to the etalon a multi-parameter group of geometric transformations, which most often in applied applications includes displacement, rotation, scaling, the effect of which does not remove the object of interest from the field of view [5].

The next step is to define the space B^n of all binary vectors of dimension n. It is known that $card\ B^n = 2^n$. B^n will be identified with the space of descriptors of keypoints of the image obtained by some keypoint detector [11], [32]. The description of the analyzed object Z as well as the description of a separate etalon $E_k = \{e_v(k)\}_{v=1}^{s}$, $E_k \subseteq B^n$ will be considered as a finite set of binary vectors of power s in the space B^n, $e_v(k) \in B^n$, $s = card\ E_k$ – the number of descriptors in the set. Each descriptor $e_v(k)$ in the database E is characterized by the parameter k of the class number. In general, the number of features – descriptors in the base set E is $card\ E = sN$. For simplicity, the number of s descriptors is assumed to be the same for all E_k.

The process of R structural classification for the analyzed object $Z = \{z_v\}_{v=1}^{s}$, $z_v \in B^n$ using the traditional method based on the voting procedure [31] will be carried out in two stages as $R = R_2R_1$. The R_1 stage implements a local decision to determine the class for a separate descriptor z_v of the object Z, and the R_2 stage determines the value of the class parameter for the analyzed object based on the decisions (votes) of the components of the entire description Z. In fact, the stage R_1 in such a model implements the nearest neighbor method on the training set E, where the class number for the object descriptor is defined as the class of the nearest element from E. And the R_2 stage implements the procedure for assessing the relevance of descriptions using a homogeneous voting model with a decision based on the highest number of votes on a set of classes.

The R_1 classifier, using the linear search method, performs an element-by-element analysis of the input description $Z = \{z_v\}_{v=1}^{s}$ of the object and assigns each $z_v \in Z$ descriptor to one of the classes according to the rule

$$R_1 : z_v \to 1, \ldots, N. \tag{1}$$

The class k of the analyzed descriptor z_v in the model (1) is defined as an argument of the minimum of a certain distance on the set E

$$R_1 : k = arg \min_{i=1,\ldots,N;d=1,\ldots,s} \rho\left(z_v, e_d(i)\right). \tag{2}$$

Here $\rho : B^n \times B^n \to [0, n]$ is the distance in the vector space B^n. The most effective in the space B^n of binary vectors is the use of in (2) of the Hamming metric [6].

In fact, the classification model (2) not only determines the class k for the descriptor z_v of an object but also establishes a correspondence for a pair of elements $z_v \to e_d(k)$ of the sets Z and E_k. This correspondence can be described as a binary equivalence relation between the elements z_v, $e_d(k)$ of the two sets of descriptors [19]. At the same time, it is assumed that some elements of Z in the implementation of (2) may not be assigned to any of the etalon classes at all if the obtained correspondence is insignificant.

The significance of the correspondence for the descriptors is determined by checking the logical rule [5]

$$\rho_m = \min_{i=1,\ldots,N;d=1,\ldots,s} \rho\left(z_v, e_d(i)\right), \rho_m \le \delta_\rho, \tag{3}$$

where δ_ρ is some threshold for the minimum ρ_m of the metric ρ. Condition (3) establishes equivalence for a pair of descriptors, taking into account the allowable limit δ_ρ. In fact, according to rule (3), the equivalence of any pair of descriptors is determined with a certain tolerance, and each descriptor z_v in the space B^n is considered as a certain multidimensional ball containing equivalent to z_v descriptors z_x under the condition $\rho(z_v, z_x) \le \delta_\rho$.

By performing R_1 for each $z_v \in Z$ in accordance with (2) and (3), the class number k is determined, and then increment the vector – the vote accumulator $h_k = h_k + 1$ for the resulting class number according to the model (2). In fact, the vector $\{h_i\}_{i=1}^{N}$ calculated by aggregation over a set of object description elements is a distribution (histogram) of the number of votes of object descriptors in the existing class system [14], [15], [16].

Now, let's build the rule R_2 based on the vector $\{h_i\}_{i=1}^{N}$ – an accumulator of class scores with integer values for the accumulated number of votes obtained by applying model (2) to the entire set of object descriptors Z. The rule R_2 defines the class of the object is by as the argument of the maximum

$$R_2 : Z \to E_k \mid \left(k = arg \max_{i=1,\ldots,N} h_i\right) \& \left(h_k \ge \delta_h\right), \tag{4}$$

where δ_h is a threshold for the minimum acceptable value of the maximum among the number of votes. If the $h_k \ge \delta_h$ condition is not met, the class of the object is considered unspecified (refusal to classify; there is no reason to assign the object to any of the available classes).

The value of δ_h is determined experimentally for a given database of E etalons. As a rule, it is chosen as the minimum number of votes (with a tolerance) required to confidently classify the test sample based on the transformations of the etalons. Another option is a statistically significant number of votes (half of the possible number). From a general theoretical point of view δ_h is a compromise in the task of distinguishing a finite set of etalons from the virtually unlimited number of other images that can be input to a classification system [5], [31].

The sequence of classification rules R_1, R_2 implements a classifier based on an ensemble of solutions of a set of classifiers obtained for a set of constituent components of an object. It implements the principles of structural analysis and alignment with the etalon database and provides resistance to spatial distortions of individual components (descriptors) due to the possible influence of interference and background in the process of image analysis.

The classification performance will be traditionally evaluated by the accuracy criterion pr, which reflects the ratio of the number of q test descriptions with correct class evaluation to the total number of Q experiments performed

$$pr = q/Q. \tag{5}$$

B. TANIMOTO METRIC FOR DETERMINING RELEVANCE

The Tanimoto metric $\mu(A, B)$ is one of the most widely used metrics for evaluating the similarity of two sets A, B, since its value integrally reflects the relevance of a pair of sets to the entire set of their components, unlike several other metrics [17], [18]. The Tanimoto metric can also take into account the repetition of elements, i.e., the variant of multisets. This metric is defined as the ratio of the powers of symmetric difference and union for a pair of compared sets A, B and takes values within the interval $[0 \ldots 1]$

$$\mu(A, B) = card(A \Delta B) / card(A \cup B). \tag{6}$$

At the same time, expression (6) can be presented in a more practical form

$$\mu(A, B) = \frac{card(A) + card(B) - 2card(A \cap B)}{card(A) + card(B) - card(A \cap B)}. \tag{7}$$

From (7), it is clear that the main task for implementing the calculation of the Tanimoto metric is to determine the power for the intersection of two finite sets since their individual powers are usually known (this is the number of descriptors in the description).

C. CALCULATING A METRIC BETWEEN DATA SETS IN A QUANTIZED VECTOR SPACE

Let's perform quantization (clustering) to jointly describe the complete database E from N etalons E_k to M quanta. Quantization is implemented by a self-learning procedure and is intended to represent the existing set of etalon data in the form of a fixed set of clusters containing similar (equivalent, close, equal) elements. This technique is a kind of segmentation (partitioning) of the set of descriptors of the existing etalon database on M into groups.

The result is a system of centers $W = \{w_j\}_{j=1}^{M}$ of quanta and a representation of the database E in the form of a partition $E = \cup_{j=1}^{M} T_j$, where the constituent data segments – clusters T_j are obtained in such a way that they do not overlap. At the same time, the centers w_j, depending on the self-learning procedure may not belong to the space B^n at all, but they can be transformed into B^n, for example, by rounding the components w_j. From the point of view of the multiset theory, the obtained centers create a certain base set, on which it is now possible to universally project the description of an image as an arbitrary set of binary vectors from the space B^n [6].

First, let's design according to the scheme [33] for the composition of existing etalon descriptions. In accordance with the design scheme, for an arbitrary descriptor $z \in E_k$ of an etalon or object, the number j of the cluster (quantum) is set according to the competitive rule

$$z \in T_j \mid j = arg \min_{i=1,\ldots,M} \rho(w_i, z). \tag{8}$$

Rule (8) implements one of the variants of the "bag of words" classification model, where the base set is the set of cluster centers. In expression (8), the Hamming metric can be used at $w_j \in B^n$.

As a result of applying (8) to the entire set of description components, a representation of each description E_k is numerically obtained in the form of a vector of integers $\varphi[E_k] = \{\varphi_{k,i}\}_{i=1}^{M}$ with the number of components corresponding to the number M of clusters. The value of $\varphi_{k,i}$ is the number of elements of the etalon E_k assigned to the cluster with the number j as a result of the analysis (8). The obtained data of the cluster representation for the E database can be considered as a matrix $\{\{\varphi_{k,i}\}_{i=1}^{M}\}_{k=1}^{N}$, the rows of which contain a quantitative decomposition of the composition of the etalon by the cluster system, and the columns contain a numerical representation of the composition of each cluster by the class system. The scheme for projecting the description onto the set of quantum centers is shown in Figure 1.

In the process of classification, the design is performed by applying rule (8) to the set of elements $z_v \in Z$ of the analyzed object. A numerical representation as a vector $\varphi[Z] = \{\varphi_i(Z)\}_{i=1}^{M}$ is obtained. This vector, like the $\varphi[E_k]$ representation of the etalon, can be interpreted as a function of the object elements belonging to the set of cluster centers.

The metric models are applied to calculate the power (9), intersection (10), union (11), and symmetric difference (12) for a pair of multisets A, B (etalon – object) [17], [34]

$$card(Z) = \sum_{j=1}^{M} \varphi_j, \tag{9}$$

$$card(A \cap B) = \sum_{j=1}^{M} min\left[\varphi_j(A), \varphi_j(B)\right], \tag{10}$$

$$card(A \cup B) = \sum_{j=1}^{M} max\left[\varphi_j(A), \varphi_j(B)\right], \tag{11}$$

$$card(A \Delta B) = \sum_{j=1}^{M} \left|\varphi_j(A) - \varphi_j(B)\right|. \tag{12}$$

These metrics are the set-theoretic equivalent for the intersection and union of two sets. Next, let's determine the value $\mu(A, B)$ of the Tanimoto metric by expressions (6) or (7). The parameter M here acts as the dimension of the compared multisets. Relations (9), (10), (11), and (12) are used in the application of theories of fuzzy sets and triangular norms [17], [35], [36]. Also, to obtain the membership function of a set – the intersection of two sets with normalized membership functions – models such as the algebraic product $\varphi_j(A) \times \varphi_j(B)$ (elementwise multiplication of membership function values) and the bounded sum $max[\varphi_j(A) + \varphi_j(B) - 1, 0]$ are used. All these actions on sets are considered in more detail in applications of the theory of triangular norms. A triangular norm is a two-place function $f(x, y), f : [0, 1]^2 \to [0, 1]$ with a domain of definition in the form of a closed unit square and a set of values in the form of a closed unit interval [17], [34].

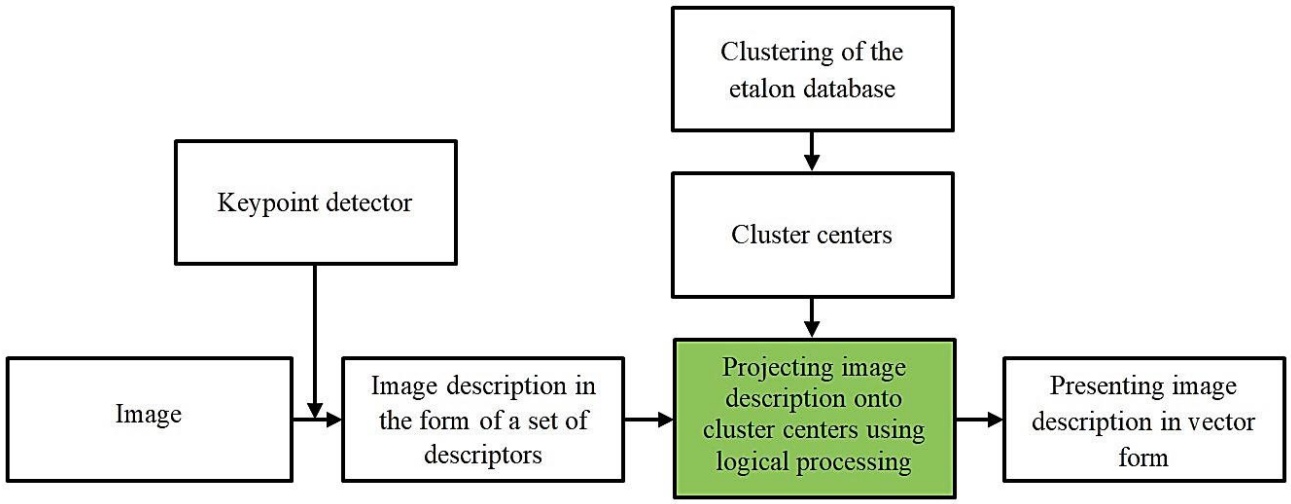


FIGURE 1. Projecting an image description onto a set of centers.

The resulting classification decision on the class of the object with the description Z is taken as an argument for the minimum distance on the set $\{\varphi[E_i]\}$ for the projections of the etalons

$$R_2 : Z \to E_k \mid k = arg \min_{i=1,\ldots,N} \mu\left(\varphi[Z], \varphi\left[E_i\right]\right). \quad (13)$$

The traditional method of classification using voting implements a sequence of models (2), (3), and (4).

The classifier using quantization based on the center system W first obtains the vector representation $\varphi(Z) = [\varphi_1, \varphi_2, \ldots, \varphi_M]$ for the recognized object according to model (8), and then, using the metric μ to calculate the relevance of integer vectors, determines the object class as an argument to the minimum of the metric μ on the set of cluster representations for the database etalons E.

A variant of the metric μ in (13) can be the Manhattan distance or any other distance for the space of integer vectors. Using the rule (13) in clustering methods instead of (2) and (4) for the traditional method gives a gain in computational complexity roughly proportional to the value of $\gamma = s/M$. Note that in order to ensure the rejection of extraneous images, the calculated optimal value of the metric μ_v in (13) must satisfy the condition of significance $\mu_v \leq \delta_\mu$, where δ_μ is the minimum metric value between clustered representations of descriptions.

In the practical application of the metric apparatus for classification, special attention should be paid to determining the center of the cluster according to rule (8) when forming $\varphi[Z]$. The minimum of the metric ρ_j obtained in (8) must necessarily satisfy condition (3) of the significance of $\rho_j \leq \delta_\rho$. This requirement is necessary in practice to filter out false descriptors that appear under the influence of the background [5].

Another practical limitation is the requirement for approximately the same number of descriptors in the descriptions of the object and the etalons. If the powers of the sets differ significantly, it is necessary to additionally normalize the data by the number of descriptors to make a balanced classification decision. At the same time, the use of the Tanimoto metric makes it possible to classify without additional data normalization.

In the quantized vector space, other similarity measures can be applied with low computational requirements, such as correlation coefficient, statistical criteria, etc.

D. JUSTIFICATION OF THE CHOICE OF THRESHOLD PARAMETERS. FILTERING BY DISTANCE TO CENTERS

It's important to note that the classification process discussed here, using metric correlations between descriptors and cluster centers, as well as between cluster views, relies on a number of parameters, the most important of which are thresholds for making a decision:

1) δ_ρ – to establish the equivalence of a pair of descriptors.
2) δ_h – to be crucial for maximizing the number of class votes.
3) δ_μ – to minimize the metric value between cluster views of descriptions.

In most applications, due to the natural proximity of data in a multidimensional space, the threshold δ_ρ is traditionally chosen as 25% (or less) of the defined maximum of the metric value ρ for a pair of descriptors [18].

The thresholds δ_h, δ_μ determine the required degree of significance of the measure accumulated by the description composition when deciding on the proximity of descriptors and the class of the object. They are data-dependent, so they are estimated based on the results of analyzing a priori etalon information in the classification database [2], [7], [19], [37], [38].

Taking into account the specifics of the classification method based on the use of the cluster representation, a new

threshold $\delta_{\rho 1}$ for the distance will be introduced between the descriptor and the cluster center, which will be implemented in the procedure (8) when determining the optimal cluster center by the minimum distance for the descriptors of the analyzed description.

If the minimum distance in the model (8) does not exceed the value $\delta_{\rho 1}$, then the descriptor will not be assigned to any of the centers of the cluster system.

This condition can be formalized by checking the truth of some predicate $Pr(z, \{w_i\}, \delta_{\rho 1})$:

$$Pr\left(z, \{w_i\}, \delta_{\rho 1}\right) : \min_{i=1,\ldots,M} \rho\left(w_i, z\right) \leq \delta_{\rho 1}. \qquad (14)$$

This predicate takes the value of one (true) when the inequality in (14) is fulfilled and zero otherwise. Thus, with the predicate Pr, a procedure for filtering false descriptors is actually introduced based on the criterion of distance from the system of centers on which the classification is based.

In the proposed processing variant (14), the threshold $\delta_{\rho 1}$ will have a rather significant impact on the classification result, as its implementation enables, to some extent, the filtering out of false descriptors that are far from the established system of centers. The parameter $\delta_{\rho 1}$ regulates the degree of such rejection. The value of the $\delta_{\rho 1}$ parameter can vary within the permissible range of values of the ρ metric for descriptors.

Note that the innovation (14) can be preliminarily applied to the etalon descriptions as well. As a result of such filtering, the cluster representations of the etalons may change in the direction of reduction, and perhaps their power may even become unequal for different etalons. This situation is automatically taken into account by using the Tanimoto metric, whose model includes the volumes of the compared descriptions. Another approach is to introduce normalization for the filtered etalon view by dividing it by the total number of description descriptors that satisfy the set condition with a threshold of $\delta_{\rho 1}$.

A similar filtering procedure can also be applied to the composition of clusters. A way to filter data is to use the ''champion list'' model for the data within each of the created clusters. To do this, let's define a certain procedure $D(T_j, w_j, d)$ for data processing that selects a fixed number d of elements from each cluster T_j with the closest distance to its center w_j. As a result, the number of elements in the quantized etalon descriptions will be reduced. The condition $card\left(T_j\right) = d$ will be fulfilled, and the total amount of etalon data will be equal to Md as the product of the number M of clusters by the fixed number d of elements in each cluster. This modification reduces the size of the cluster representation to a fixed size but, in turn, also requires the use of the Tanimoto metric or normalization of the resulting representation for the etalons.

By changing the number d of data points in the cluster using the D procedure or varying the threshold $\delta_{\rho 1}$, the formation of data volume in the cluster representation can be controlled. With the transformed compressed data volume, the cluster centers can be updated so that they more accurately approximate the class distribution, which is used in the classification process. In fact, with this innovation, D is produced as a procedure for filtering the descriptor space according to the criterion of their consistency with the system of cluster centers. The introduction of D leads to a compression of the cluster representation and some redistribution of the number of elements of different classes within the cluster. This approach allows the classifier to better adapt to the available data to improve efficiency. Note that the D procedure can be implemented at the stage of data preparation for classification, and its implementation does not directly affect the computational cost of classification.

It is obvious that different schemes for implementing the proposed innovation (14) can be considered, including the way it is used exclusively for filtering the analyzed data of the input description, leaving the clustering result for the descriptions of the etalon database unchanged.

The thresholds δ_{ρ} and $\delta_{\rho 1}$ do not directly depend on the database, but the threshold δ_h depends on these thresholds. The thresholds δ_h and δ_{μ} explicitly depend on the database. They are set based on a cross-calculation of the number of votes or the Tanimoto distance between the etalons. The basis is the percentage for the largest interclass number of votes and the smallest Tanimoto distance. The threshold $\delta_{\rho 1}$ depends on the influence of external factors and sets the required level of filtering false data during classification. At the same time, it should ''pass'' data as close to the centers as possible to ensure effective classification.

In general, the thresholds for the metric, number of votes, and filtering are determined experimentally, and their choice depends on the image database and recognition conditions.

III. EXPERIMENTS AND DISCUSSIONS

The software modeling of the proposed methods was carried out in the Visual Studio Code environment, a modern cross-platform code editor that supports the Python language, including its modern libraries [39].

For each image, the Binary Robust Invariant Scalable Keypoints (BRISK) detector obtained $s = 500$ the most significant parameters in terms of the value of the *response* binary detectors. The response parameter in the BRISK detector reflects the degree of prominence of keypoints among the surrounding pixels. The BRISK descriptor has the dimension $n = 512$. The classification results are generated in an Excel spreadsheet.

Ten images of historical monuments (palaces, castles, churches) in different architectural styles were selected for the research [40]. Of these, five images are included in the etalon database, and the remaining five form a sample that is not included in the database. The etalon images are shown in Figure 2, and the images outside the database are shown in Figure 3.

Figure 4 shows an example of an etalon image with the coordinates of the keypoints generated and selected by the *response* criterion. The *response* parameter obtained by the BRISK detector is a floating point number that varies

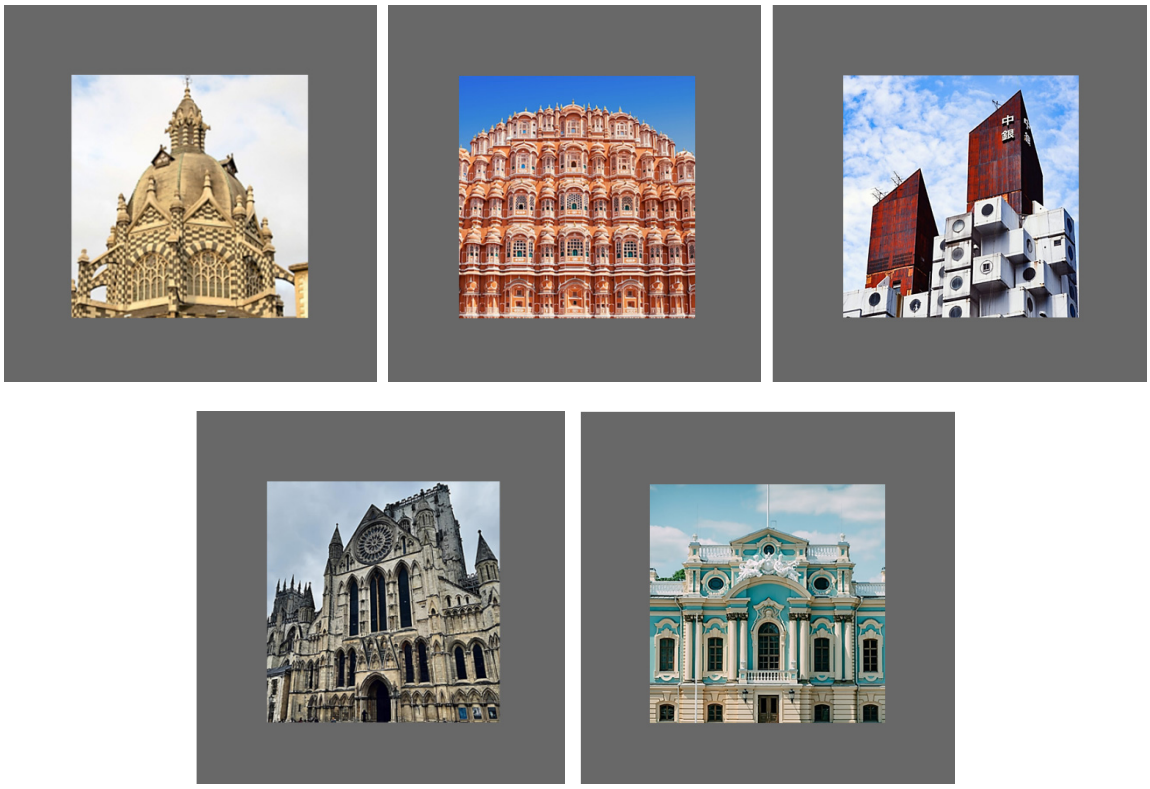


FIGURE 2. Etalon images in the experiment.

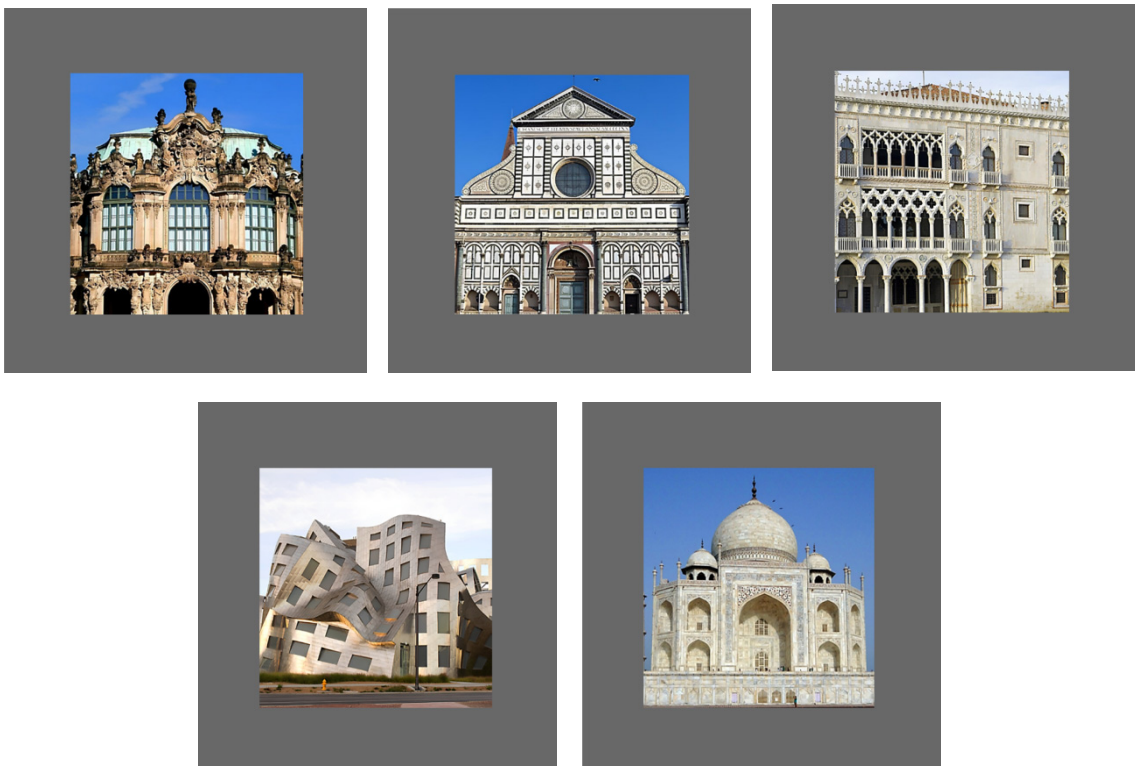


FIGURE 3. Image outside the base.

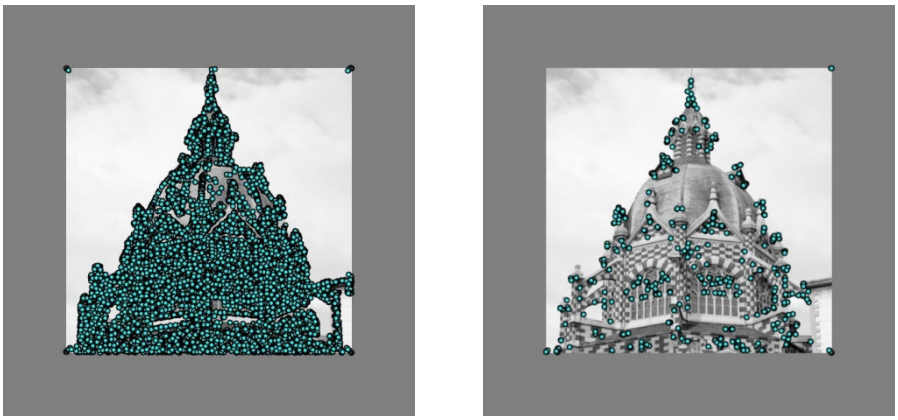


FIGURE 4. Examples of image with coordinates of all keypoints and selected 500 keypoints by *response* parameter.

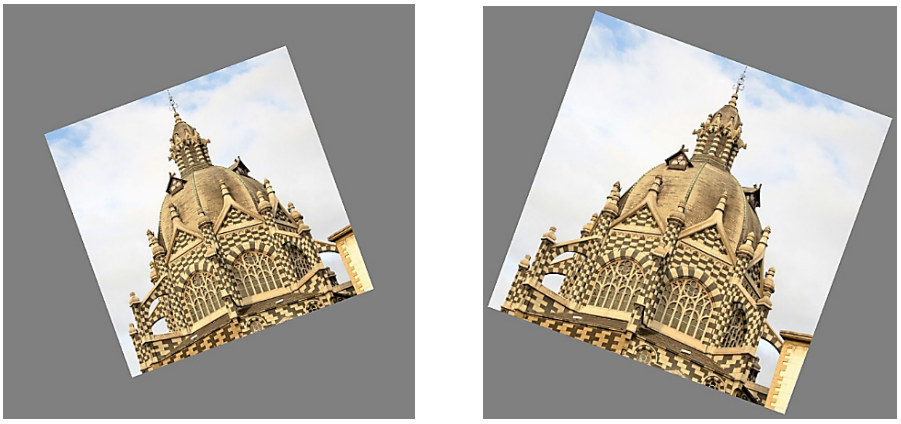


FIGURE 5. Examples of combining rotate and scale transformations.

TABLE 1. Cluster view for etalon images.

Etalon number	Number of etalon elements in clusters									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	101	54	14	45	36	29	47	62	41	71
2	54	26	63	26	23	41	76	74	66	51
3	90	96	21	51	35	76	24	31	36	40
4	69	57	42	58	28	41	31	48	68	58
5	76	52	21	38	52	71	44	40	36	70

from 10 to 245 when calculated for the images in the experiment. The initial number of keypoints obtained is 21000, and then 500 keypoints with the highest *response* values are selected.

At the preparation stage, all images were converted to halftone, cropped to 700×700 pixels, and centered on a gray background of 1100×1100 pixels. This way, displacement of objects is avoided relative to the image frame during transformations and ensure equal conditions for detecting keypoints.

The test sample included etalon images, the images outside the database, and these 10 images transformed by geometric rotation (+20, -20 degrees) and scale (0.9, 1.1) transformations and their combinations with offsets. In total, the performance and speed of processing for a sample of 170 images is evaluated. Examples of images with transformations are shown in Figure 5.

The modeling and evaluation of processing time was performed on an HP EliteBook 850 G5 computer, Intel Core i5-8250U processor (1.70 GHz), 16 GB DDR4 RAM, Windows 11 Pro OS, version 10.0.26100.

The key moment of the experiment was a comparative evaluation of the speed and classification accuracy criteria for the traditional voting method and the proposed method based on cluster representation using the Tanimoto metric.

The second important point was to study the impact of introducing the $\delta_{\rho 1}$ filtering threshold on the classification performance of the cluster-based method.

Clustering for the set of descriptors from the descriptions of the etalon database was performed by the batch self-learning k-means method with a fixed number of clusters $M = 10$. In order to use only the Hamming metric in the formation of clusters, the updated centroids were transformed into a binary form.

Experimental Table 1 contains the composition of the vector representation for the etalon database (Figure 2).

The data in Table 1 contain fundamental information about the obtained cluster representation of the etalons, which largely determines the effectiveness of the classification process in general.

As can be seen from Table 1, the composition of the etalons with respect to the formed centers is relatively homogeneous, which indicates their high joint similarity.

When functioning correctly, the classifier should successfully recognize the etalon samples and assign them to its class, and images outside the database should not be classified into any of the classes.

For the voting method, the equivalence threshold for the pair of descriptors $\delta_\rho = 128$ (25% of the maximum possible value of the 512 Hamming distance for BRISK descriptors), as well as the threshold for the decisive value of the number of class votes – $\delta_h = 250$ (half of the maximum number) were chosen.

Based on the analysis of the value of the Tanimoto metric for the set of etalons in the modified feature space, the threshold $\delta_\mu = 0.23$ was selected as the minimum of this metric. The maximum value of the Tanimoto metric was 0.53.

As this study experimental analysis has shown, the effectiveness of the classifier with cluster representation significantly depends on this actual value (0.23), as it characterizes the degree of discrepancy within the etalon data. In it revealed during the study, from a statistical point of view, to ensure reliable classification under geometric transformations and possible interference, this value should be close to 0.30. The required initial level of distinction for a set of selected etalons can be obtained in such ways as direct selection of etalons, varying the number of clusters, choosing a clustering method, center formation model, etc.

In the course of a detailed experiment, several variants of the $\delta_{\rho 1}$ filtering threshold were researched:

a) Without filtering of the analyzed image and etalons (traditional).
b) Filtering of the analyzed image and etalons.
c) Filtering of the analyzed image only.
d) Filtering of etalons only.

Among these options for applying the $\delta_{\rho 1}$ threshold to filter data on cluster centers, the best method according to the experiment, was chosen in c), when only input descriptions are filtered, and etalon descriptions are used in full (without preliminary filtering).

This experimental fact can naturally be attributed to the properties of the specific data used in the experiment. At the same time, even the reduction of etalon descriptions under the influence of the $\delta_{\rho 1}$ threshold without the influence of geometric transformations showed their confident recognition of the images used in the experiment. At the same time, using filtering with the threshold of $\delta_{\rho 1} = 192$, the value of the Tanimoto metric for the etalons increased slightly and ranged from 0.1 to 0.53, and for images outside the database – from 0.45 to 0.49.

An experimental evaluation for the value of the threshold $\delta_{\rho 1}$ was also carried out.

As a result, the threshold $\delta_{\rho 1} = 192$ was determined (as 37.5% of the maximum value of the Hamming distance). Increasing the $\delta_{\rho 1}$ threshold has almost no effect on the performance, while reducing $\delta_{\rho 1}$ to the level of 128 leads to a significant reduction in the input description and makes classification virtually impossible.

Without the influence of geometric transformations, the classifier modified by the threshold parameter $\delta_{\rho 1} = 192$ with the Tanimoto metric showed an error-free result. At the same time, the volume of filtered descriptions of the input images (database and outside the database) decreased and ranged from 330 to 475, and the value of the Tanimoto metric for the etalons of ''their'' classes increased slightly to the range from 0.05 to 0.20.

As a result of the experiment on a test set of 170 images with geometric transformations, it was found that both approaches – voting and quantization using the Tanimoto metric – demonstrated successful classification with the highest possible accuracy $pr = 1.0$ for the used set of 170 test input images.

Comparative analysis of the computer time to classify one image showed that the method using cluster representation and the Tanimoto metric shows significantly better speed – only 0.1 ms compared to 9.38 ms for voting, i.e., 94 times less! At the same time, the gain in processing time for voting increases proportionally to the number of etalons in the database.

However, it's worth mentioning that the performance of the method with the Tanimoto metric is more dependent on the characteristics of the used image set. In the experiment, it was found that swapping the roles between the etalons and outside the database images (i.e., swapping their places) leads to some decrease in accuracy for this method on the same test set to the level of $pr = 0.93$. The scale transformation has a greater impact on the accuracy decrease. The magnitude of the transformation also matters. This parameter indicates a greater sensitivity of the proposed approach to the formation of the etalon base and the need for careful selection of representative data and the acceptable level of transformation. Suppose a sufficiently significant difference for the etalons is provided (by selecting them or by selecting clustering), for example, at the level of 0.3 according to the Tanimoto metric. In that case, the proposed method provides no less accuracy than voting.

The results obtained in the experiment are shown in Table 2.

Instead, the traditional method under the changed experimental conditions demonstrated stable operation and showed high accuracy. This method is consistent with its status as a proven optimal approach that has already proven itself in a number of similar problems [4], [5], [6], [7].

TABLE 2. Comparative indicators of the proposed and traditional methods.

Method	Classification time, ms	Accuracy	Accuracy for base variation
Traditional	9.38	1.0	1.0
With quantization	0.1	1.0	0.93

Accordingly, in the applied implementation of the method with clustering and the Tanimoto metric, in order to ensure the required level of accuracy, it is worth conducting an additional analysis with a detailed study of the influence of the composition of descriptors, specifics of normalization and filtering of features in the classification process.

The classification speed for the method with quantization of the set of descriptive descriptors, compared to the traditional linear search method, expands proportionally to the increase in the number of etalons and the ratio s/M of the description volume to the number of clusters.

For the values $N = 5$, $s = 500$, and $M = 10$ used in the experiment, we have a performance gain of approximately 94 times. With an increase in the number M of cluster centers, the classification accuracy naturally grows too, but the performance gain decreases. However, the choice of the most effective M value, unfortunately, which is typical for all data analysis tasks, also depends on the composition of the selected database images. After the thorough analysis of the outcomes of this research and the works [6], [18], we can summarize that the selection of a cluster structure with the parameter M that will ensure the required level of accuracy and speed of classification in an applied task with specific descriptions of etalons should exclusively rely on the results of a specific experiment.

The voting method provides high accuracy and stability of classification, but requires significant processing time. The high computational time requirements are explained by the need to thoroughly search through the descriptors of all the etalons for each input image (linear search).

According to the research results, the proposed method accuracy is not inferior to voting in situations with small values of geometric transformation parameters, but it requires a more thorough study of the etalon base and classification conditions. At the same time, the proposed method has significant advantages in processing speed and can be successfully applied to tasks where speed is a critical factor. Examples include real-time systems and devices with computing resource constraints.

IV. CONCLUSION

Quantization in the space of multidimensional data describing visual objects extensively simplifies classification decisions, significantly speeds up computations, and more accurately formalizes and justifies the classification process in applied computer vision tasks.

The introduction of quantization makes it possible to take into account the intrinsic nature of the processed multidimensional data to a greater extent, with the possibility of using the apparatus of multisets or fuzzy sets. Quantization in interaction with the metrics apparatus not only makes it possible to formalize the classification process and avoid ambiguity in the interpretation of data equivalence but also provides a decent level of accuracy with a significant gain in processing speed. The classification time for the researched method practically does not increase with the number of etalons, while the computational cost of the traditional method is directly proportional to the number of etalons.

Classification methods based on quantization of etalon information, in addition to those discussed in this paper, have several further development prospects.

First, it involves improving or analyzing of clustering procedures to select the most effective one in distinguishing between descriptions for the etalons of different classes in the cluster representation. For example, separate clustering on a set of etalons enhances the effectiveness of classification methods using a cluster representation [6]. Another significant improvement in the performance of the classifier based on available a priori information may be the use of class number values for etalon descriptors during clustering.

Secondly, it is the analysis and use of the weighting coefficients of the cluster representation to identify and concentrate the most significant ones.

Thirdly, it controls the scale of the cluster view (number of clusters) using the filtering model (14).

All of these improvements can be made at the preprocessing stage to ensure consistent classification performance across the board for the applied image set.

ACKNOWLEDGMENT

This work was supported by the European Union NextGenerationEU through the Recovery and Resilience Plan for Slovakia under Project 09I03-03-V01-00115. The results of this research were obtained under the international research project EU Horizon Europe ''INITIATE: Supporting European research and innovation through stakeholder collaboration and institutional reform'' (grant agreement No. 101136775). The authors would like to thank Shcherbak Darii for participating in the experiments.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to report regarding the present research.

REFERENCES

[1] H. Li, F. Xie, J. Zhou, and J. Liu, ''Object detection, segmentation and categorization in artificial intelligence,'' *Electronics*, vol. 13, no. 13, p. 2650, Jul. 2024, doi: 10.3390/electronics13132650.

[2] Q. Bai, S. Li, J. Yang, Q. Song, Z. Li, and X. Zhang, ''Object detection recognition and robot grasping based on machine learning: A survey,'' *IEEE Access*, vol. 8, pp. 181855–181879, 2020, doi: 10.1109/ACCESS.2020.3028740.

[3] W. Zhang, G. Chen, P. Zhuang, W. Zhao, and L. Zhou, ''CATNet: Cascaded attention transformer network for marine species image classification,'' *Expert Syst. Appl.*, vol. 256, Dec. 2024, Art. no. 124932, doi: 10.1016/j.eswa.2024.124932.

[4] R. M. Tymchyshyn, O. Y. Volkov, O. Y. Gospodarchuk, and Y. P. Bogachuk, ''Modern approaches to computer vision,'' *Control Syst. Comput.*, vol. 278, no. 6, pp. 46–73, Dec. 2018, doi: 10.15407/usim.2018.06.046.

[5] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, and M. Zeghid, "Improving the effectiveness of image classification structural methods by compressing the description according to the information content criterion," *Comput., Mater. Continua*, vol. 80, no. 2, pp. 3085–3106, Aug. 2024, doi: 10.32604/cmc.2024.051709.

[6] V. Gorokhovatskyi, I. Tvoroshenko, O. Yakovleva, M. Hudáková, and O. Gorokhovatskyi, "Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set," *IEEE Access*, vol. 12, pp. 73376–73385, 2024, doi: 10.1109/ACCESS.2024.3404371.

[7] V. Gorokhovatskyi, I. Tvoroshenko, O. Yakovleva, and M. Hudáková, "Image description compression in classification structural methods," *IEEE Access*, vol. 13, pp. 43631–43641, 2025, doi: 10.1109/ACCESS.2025.3548910.

[8] Y. Parzhin, "Principles of modal and vector theory of formal intelligence systems," 2013, *arXiv:1302.1334*.

[9] C. C. Aggarwal, "Advanced topics in neural networks," in *Neural Networks and Deep Learning*. Cham, Switzerland: Springer, 2023, doi: 10.1007/978-3-031-29642-0.

[10] M. R. Mishra and P. K. Meher, "Efficient detection and classification of brain tumors using a novel local binary pattern," *Academia Eng.*, vol. 2, no. 1, pp. 1–16, Mar. 2025, doi: 10.20935/acadeng 7568.

[11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: 10.1023/B:VISI.0000029664.99615.94.

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

[13] T. N. Mau, Y. Inoguchi, and V.-N. Huynh, "A novel cluster prediction approach based on locality-sensitive hashing for fuzzy clustering of categorical data," *IEEE Access*, vol. 10, pp. 34196–34206, 2022, doi: 10.1109/ACCESS.2022.3162690.

[14] R. Khan, C. Barat, D. Muselet, and C. Ducottet, "Spatial histograms of soft pairwise similar patches to improve the bag-of-visual-words model," *Comput. Vis. Image Understand.*, vol. 132, pp. 102–112, Mar. 2015, doi: 10.1016/j.cviu.2014.09.005.

[15] T. Yu, W. Chen, G. Junfeng, and H. Poxi, "Intelligent detection method of forgings defects detection based on improved EfficientNet and memetic algorithm," *IEEE Access*, vol. 10, pp. 79553–79563, 2022, doi: 10.1109/ACCESS.2022.3193676.

[16] M. Ghahremani, Y. Liu, and B. Tiddeman, "FFD: Fast feature detector," *IEEE Trans. Image Process.*, vol. 30, pp. 1153–1168, 2021, doi: 10.1109/TIP.2020.3042057.

[17] O. Berehulyak and R. Vorobel, "The algebraic model with an asymmetric characteristic of logarithmic transformation," in *Proc. IEEE 15th Int. Conf. Comput. Sci. Inf. Technol. (CSIT)*, vol. 2, Sep. 2020, pp. 119–122, doi: 10.1109/CSIT49958.2020.9321906.

[18] M. Alghieth, "DeepECG-Net: A hybrid transformer-based deep learning model for real-time ECG anomaly detection," *Sci. Rep.*, vol. 15, no. 1, p. 20714, Jul. 2025, doi: 10.1038/s41598-025-07781-1.

[19] M. Isik, "Comprehensive empirical evaluation of feature extractors in computer vision," *PeerJ Comput. Sci.*, vol. 10, p. e2415, Nov. 2024, doi: 10.7717/peerj-cs.2415.

[20] Y. Xie, X. Hou, Y. Guo, X. Wang, and J. Zheng, "Joint-guided distillation binary neural network via dynamic channel-wise diversity enhancement for object detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 1, pp. 448–460, Jan. 2024, doi: 10.1109/TCSVT.2023.3286072.

[21] S. Lee, H. Seong, S. Lee, and E. Kim, "Correlation verification for image retrieval and its memory footprint optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 3, pp. 1514–1529, Mar. 2025, doi: 10.1109/TPAMI.2024.3504274.

[22] Y.-S. Tai, C.-Y. Chang, C.-F. Teng, Y.-T. Chen, and A.-Y. Wu, "Joint optimization of dimension reduction and mixed-precision quantization for activation compression of neural networks," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 42, no. 11, pp. 4025–4037, Nov. 2023, doi: 10.1109/TCAD.2023.3248503.

[23] Y. Ma, M. Wang, G. Lu, and Y. Sun, "Deep hashing similarity learning for cross-modal retrieval," *IEEE Access*, vol. 12, pp. 8609–8618, 2024, doi: 10.1109/ACCESS.2024.3352434.

[24] J. Jiang, D. Kim, B. Kim, and Y.-G. Shin, "GA-VAE: Enhancing local feature representation in VQ-VAE through genetic algorithm-based token optimization," *IEEE Access*, vol. 13, pp. 34286–34295, 2025, doi: 10.1109/ACCESS.2025.3542384.

[25] X. Li, "Non-relaxation deep hashing method for fast image retrieval," *IEEE Access*, vol. 11, pp. 17684–17692, 2023, doi: 10.1109/ACCESS.2023.3244813.

[26] Z. Han, A. B. Azman, F. B. Khalid, and M. R. B. Mustaffa, "Multi-granularity semantic information integration graph for cross-modal hash retrieval," *IEEE Access*, vol. 12, pp. 44682–44694, 2024, doi: 10.1109/ACCESS.2024.3380019.

[27] B. Subbiyan, R. Prabhavathi Neelakandan, K. Leelasankar, R. Rajavel, M. Malarvel, and A. Shankar, "A quantum-enhanced artificial neural network model for efficient medical image compression," *IEEE Access*, vol. 13, pp. 31809–31828, 2025, doi: 10.1109/ACCESS.2025.3542807.

[28] P. V. T. Minh, N. D. D. Viet, N. T. Son, B. N. Anh, and J. Jaafar, "RelaHash: Deep hashing with relative position," *IEEE Access*, vol. 11, pp. 30094–30108, 2023, doi: 10.1109/ACCESS.2023.3259104.

[29] Y. Xie, X. Hou, J. Ren, X. Zhang, C. Ma, and J. Zheng, "Binary quantization vision transformer for effective segmentation of red tide in multispectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 4202814, doi: 10.1109/TGRS.2025.3540784.

[30] Z. Chen, Y.-G. Wang, and L.-C. Li, "Debiased cross contrastive quantization for unsupervised image retrieval," *IEEE Trans. Big Data*, vol. 11, no. 3, pp. 1298–1308, Jun. 2025, doi: 10.1109/TBDATA.2024.3453751.

[31] 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.

[32] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2548–2555, doi: 10.1109/ICCV.2011.6126542.

[33] T. Kohonen, "Learning vector quantization," in *Self-Organizing Maps*. Berlin, Germany: Springer, 2001, pp. 245–261.

[34] C. Alsina, M. J. Frank, and B. Schweizer, "Functional equations involving t-norms," in *Associative Functions: Triangular Norms and Copulas*. Singapore: World Scientific, 2006, pp. 99–151.

[35] Y. Fang and L. Liu, "Angular quantization online hashing for image retrieval," *IEEE Access*, vol. 10, pp. 72577–72589, 2022, doi: 10.1109/ACCESS.2021.3095367.

[36] A. Alazzawi, Q. M. Yas, and B. Albayati, "A group decision-making for selecting multi-deep face recognition models," *Iraqi J. Comput. Sci. Math.*, vol. 6, no. 2, p. 21, May 2025, doi: 10.52866/2788-7421.1262.

[37] M. Hassaballah, H. A. Alshazly, and A. A. Ali, "Analysis and evaluation of keypoint descriptors for image matching," in *Recent Advances in Computer Vision*. Cham, Switzerland: Springer, 2019, pp. 113–140.

[38] V. O. Gorokhovatskyi, I. S. Tvoroshenko, and N. V. Vlasenko, "Using fuzzy clustering in structural methods of image classification," *Telecommun. Radio Eng.*, vol. 79, no. 9, pp. 781–791, Sep. 2020, doi: 10.1615/telecomradeng.v79.i9.50.

[39] *Python AI: Why Is Python So Good for Machine Learning?* Accessed: Jun. 10, 2025. [Online]. Available: https://www.netguru.com/blog/python-machine-learning

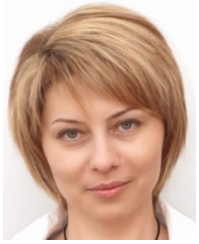
[40] *Tripadvisor*. Accessed: Jun. 12, 2025. [Online]. Available: https://www.tripadvisor.com/

VOLODYMYR GOROKHOVATSKYI received the M.Sc. degree in applied mathematics and engineering from Kharkiv National University of Radio Electronics (KNURE), the Ph.D. degree in management (technical systems), in 1984, and the Dr.Sc. degree in systems and means of artificial intelligence, in 2010.

He received an internship from Dresden Technical University. He is currently a Professor with the Department of Informatics, KNURE. He has more than 200 scientific articles and seven monographs. His research interests include the image and pattern recognition in computer vision systems, structural methods of image classification and recognition, multidimensional data analysis, and artificial intelligence.

IRYNA TVOROSHENKO received the Ph.D. degree in systems and means of artificial intelligence, in 2010.

She is currently an Associate Professor with the Department of Informatics, Kharkiv National University of Radio Electronics. She has published 188 scientific articles and educational and methodical articles include four study guides, seven monographs, 50 articles, 100 abstracts of reports, 14 lecture notes, and 13 methodological instructive regulations. She is fluent in modern programming languages and technologies, computer-aided mathematical modeling, and constantly expanding her range of scientific interests. Her research interests include image and pattern recognition in computer vision systems, structural methods of image classification and recognition, and fuzzy methods in artificial intelligence appliance.



OLENA YAKOVLEVA received the Ph.D. degree in systems and means of artificial intelligence, in 2004. In 2008, she received the academic title of an Associate Professor with the Department of Informatics.

Since 2004, she has been teaching courses on databases, big data, data visualization, artificial intelligence, computer vision, and information systems development with Kharkiv National University of Radio Electronics. Since 2017, she has been a Scientific Consultant in computer vision, machine learning, and intelligent systems with SYTOSS Ltd., Slovakia. Since 2023, she has also been a Researcher with Bratislava University of Economics and Management. She is the author of several publications on computer vision, natural language processing, and intelligent systems. Her research interests include computer vision, machine learning, and intelligent systems and related areas.

MONIKA HUDÁKOVÁ received the Ph.D. degree in branch and cross-sectional economics from the Faculty of Operational Economics, Slovak University of Agriculture in Nitra, Slovakia, in 2000, and the Professor degree in economics and business management from Slovak University of Agriculture, in 2022.

Since 2022, she has been a University Professor with Bratislava University of Economics and Management. In 2022, she held the position of the Vice-Rector of Science and Research. Since 2023, she has been the Vice-Rector of Teaching and Lifelong Learning. She is currently a full-time Professor with the Department of Economics and Finance, where she constantly raises the visibility of issues in the field of economics, management, crisis management, and international management. She is also a Full Member of Czech Academy of Agricultural Sciences and Slovak Academy of Agricultural Sciences. She has received several awards for her long-standing international cooperation and development of science and research. Her scientific research activity is documented by a rich, internationally accepted publication activity and participation in several research projects.