IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 15, 2021, accepted June 25, 2021, date of publication June 29, 2021, date of current version July 6, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3093457*

# Methods of Classification of Images on the Basis of the Values of Statistical Distributions for the Composition of Structural Description Components

**YOUSEF IBRAHIM DARADKEH[1], VOLODYMYR GOROKHOVATSKYI[2], IRYNA TVOROSHENKO[2], SVITLANA GADETSKA[3], AND MUJAHED AL-DHAIFALLAH[4], (Member, IEEE)**

[1]Department of Computer Engineering and Networks, College of Engineering at Wadi Addawasir, Prince Sattam Bin Abdulaziz University, Al-Kharj 11991, Saudi Arabia
[2]Department of Informatics, Kharkiv National University of Radio Electronics, 61166 Kharkiv, Ukraine
[3]Department of Higher Mathematics, Kharkiv National Automobile and Highway University, 61002 Kharkiv, Ukraine
[4]Systems Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

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

This work did not involve human subjects or animals in its research.

**ABSTRACT** The article considers the problem of image recognition in computer vision systems. The results of the development of the method for image classification, using a structural approach, are presented. The classification method is based on calculating the values of statistical distributions for the set of description descriptors. The distribution vector for a fixed set of classes is based on the calculation of the degree of similarity with the integral characteristics for the descriptions of the etalon base. Two options for constructing the classifier on the principles of object – etalon and object descriptor – etalon, which differ in the degree of integration of the solution, are proposed. The median for the set of vectors describing the etalon is used as the aggregate characteristic of the etalon descriptions. The experimental evaluation of the effectiveness of the developed classifiers in terms of verification of performance and evaluation of the probability of correct classification according to the results of processing of applied images based on three etalons are carried out. The values of precision and completeness indicators for the method object descriptor – etalon, which has demonstrated the significant advantage over the integrated approach, are given. At the same time, both proposed in the experiment methods classify the set of etalons without error. The methods of mathematical statistics, intellectual data analysis, image recognition, the apparatus for calculating the relevance of the system of the features, as well as simulation modelling, are used in this research. Based on the study and the experiment, it was found that the processing time of the images for the developed method is approximately 7 times less than for the traditional method, without reducing the accuracy. The perspective of further research is to study the interference immunity of the developed methods and evaluate their applied effectiveness for three-dimensional image collections.

**INDEX TERMS** Component distribution, computer vision, description relevance, descriptor, image classification, keypoint, median set, ORB detector, precision of classification, structural description.

## I. INTRODUCTION

The implementation of statistical methods as an apparatus of the intellectual data analysis for the construction of classifiers for images of visual objects in computer vision systems is aimed to achieve results during solving the applied tasks. This realization is performed on the basis of studying the content, structure and properties of etalon data, as well as on the basis of implementation of this knowledge to the classification process [1]–[6]. An element of the image space,

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

when using structural recognition methods in the vector data environment, is a finite set of descriptors of keypoints of the image [1], [7].

Recently, statistical distributions have become the primary mean for data analysis in the image recognition systems [1]–[4], [8]. If the description of a recognizable visual object is given by a set of multidimensional vectors, then the statistical apparatus becomes a key workable way to decide on the class of the object [9]–[11]. Probabilistic distributions of structural description components in the system of blocks for keypoint descriptors have shown their high efficiency in terms of quality of classification and processing speed. In [3], [4], [12], [13] the procedure of calculating descriptors, construction of probability distributions on their set and some methods of classification on the basis of distributions are considered in detail. There is the urgent need to implement the apparatus of distributions in the general form for the set of multidimensional descriptors of the description in accordance with the defined data classes, which are determined by the given etalon database [1], [4], [8], [12].

During the classification process, the training the classifier within the fixed database of etalons images is as the way to transfer information from lower (etalon descriptions as a set of descriptors of keypoints) to the upper levels of the data hierarchy (classification), that allow to generalize knowledge of lower levels [3], [5], [6], [14]. This principle of investigation emphasizes the effectiveness of the hierarchical method of data processing in the proposed methods.

The aim of the article is to develop and study the properties of the image classifier on the basis of constructing the ensemble of distributions for the composition of the components of the structural description. The use of integral and independent models for making the classification decisions is proposed. This approach provides the effective classification in the image space as the set of keypoint descriptors.

The objectives of this research are:

- To construct the classification models in the newly created space of statistical characteristics;
- To analyze the parameters that affect the efficiency of the implementation of classification models;
- To evaluate experimentally the efficiency of classifiers by means of software simulation.

The article proposes and demonstrates:

- Formal statement of the classification task (Section III). The main components are given to describe the required elements. The stages of the classifier construction are analyzed in detail. The implementation of the represented ideas provides the universality and significant effectiveness of the classification;
- The process of calculating the values of the distributions for the components (Section IV). The logical processing of input data is discussed in order to eliminate possible interference by means of analyzing the values of distances or the degree of proximity of the descriptor value to each of the classes;
- Classifier models (Section V). The methods for constructing the classifier are considered with the definition of the object class according to the integral distribution of components and the determination of the object class according to the number of component voices. The way to increase the efficiency of ensemble classification methods by adapting to existing data is described. It consists in selecting the subset of elements that are within the given distance from the center, or forming the fixed number of elements closest to the center of the description. The developed method of classification on the basis of statistical distributions of components of descriptions is based on the principle of division of data components in accordance with the classes. Mathematical models for definition of precision and completeness of the given modifications are presented;
- Results of computer simulation (Section VI). Software models have been developed for the implementation of the proposed classification methods in computer vision systems. Testing is performed on the basis of real images. The method of classification on the basis of statistical distributions of the components of the descriptions according to the results of experiments has confirmed its efficiency and effectiveness for the images classification. Using the concentrated part of the description data allows to improve its distinction with other descriptions.

The article presents the method of constructing the classification models in the synthesized data space, which confirmed the quality of the detected modifications of data analysis on the examples of images.

## II. RELATED WORKS

The formal formulation of the classification task using the description of the image as the set of descriptors of keypoints is formulated in [1], [2]. These scientific articles also research the features of the model of image representation as the set of multidimensional vectors and also the advantages of applying transformations of structural description to the feature space in the form of statistical distributions. One of the effective statistical method is cluster representation [15] and granulation using the apparatus of fuzzy sets [3]. However, the effectiveness of these methods depends significantly on the composition of the data; in addition, they require additional computational costs at the classification stage.

The use of more universal methods of statistical classification based on etalon information contributes not only to the generalization of image representation, but also to a more detailed identification of the degree of consistency of the analyzed and etalon images [1], [3]. Statistical approaches allow to solve one of the key problems during the implementation of structural methods – to reduce a sufficiently large amount of computational costs during the processing of bulk vector sets.

The powerful mean of reducing the dimensionality concerning the classification task is the use of method of data aggregation based on the definition of description centers, usually in the form of the average value or median [2], [5], [16], [17]. Despite some conventionality of these characteristics, especially in multidimensional space, the use of data centers significantly reduces the processing time, it has been confirmed by the experimental results [17], [18].

Taking into account the structure of the analyzed data in the form of the vector set, the use of ensemble processing means looks naturally promising [2], [3], [16]. In this case, the resulting classification solution is made on the basis of the set of local solutions for individual descriptors. Articles [3], [5], [6], [19], [20] contain models for constructing ensemble solutions and evaluating the effectiveness of classification systems based on simple classifiers for individual components. Such important advantages of ensemble solutions, as resistance to distortions of individual components and ensuring higher accuracy of data analysis or training, are discussed.

Some works are of particular scientific interest with deep applied content [21]–[23], but they do not contain universal statistical application, in particular, in terms of the use of existing statistical distributions.

On the other hand, the works [24], [25], devoted to the solution of problems of recognition by means of structural analysis, are based on a priori given distributions of data, while the use of nonparametric methods of intellectual analysis of data can provide the more accurate result.

The material of this study is not directly related to the approaches that evaluate textural features (patterns, etc.) [26]–[28], although it uses the apparatus of keypoint descriptors that reflect the local features of the images. Our investigation is concerned with assessing the significance of individual combinations for descriptor sets, which is intended for more universal data analysis and aims to highlight real objects.

## III. FORMAL STATEMENT OF THE CLASSIFICATION TASK

Let's consider the multidimensional space $B^n$ of any binary vectors of dimension $n$, where we will construct images of the object and etalons. Let's present the multiset of vectors $Z \subset B^n$, $Z = \{z_v\}_{v=1}^{s}$ as the description of the image of the visual object in the space of sets of descriptors of keypoints.

Let's consider the database of etalons $E = \{E_1, E_2, \ldots, E_N\}$ of dimension $N$, $E_i = \{e_v(i)\}_{v=1}^{s}$, $s = card\, E_i$ – the number of descriptors in the set $E_i$, $i$ – the number of the etalon in the database $E$ [1]. The features of the object and etalons are vectors $z_k \in B^n$, the finite set of which creates the description.

Let's specify $\forall(z_k, z_\tau)$, $z_k \in B^n$, $z_\tau \in B^n$ the distance $\rho : B^n \times B^n \to [0, \infty]$ in the vector space $B^n$. The example is the Hamming metric. For binary data of the given dimension, the range of values of this metric is fixed – $[0, n]$. Distance is the fundamental criterion of equivalence for the set of descriptors. It reflects the visual similarity of the pixel boundaries of keypoints for the image brightness function, which is showed by the descriptor [18]. Equivalence $z_k \equiv z_\tau$ for two descriptors is determined on the basis of some threshold $\delta_\rho$ for the value of the metric:

$$z_k \equiv z_\tau \ : \rho(z_k, z_\tau) \le \delta_\rho. \tag{1}$$

In the base of etalons $E = \{E_1, E_2, \ldots, E_N\}$, each fixed etalon description $E_i$ represents the separate class for the classifier and takes the form of the finite set of descriptors of keypoints with $B^n$.

The objective of the research is to build the classifier $K : B^n \to [1, 2, \ldots, N]$ based on the construction of the system of statistical features according to the learning outcome based on the material of the etalon set $E = \{E_1, E_2, \ldots, E_N\}$.

The idea of classifier construction is as follows:

- For each descriptor of the recognized object or etalons we specify the degree of belonging to the existing classes in the form of the statistical distribution;
- On the basis of the formed system of component distributions we create the integrated ensemble measure of relevance of the description of the analyzed object and each of the etalons;
- We apply the created measure for the classifier using optimizing the value of relevance in the class system.

On the basis of the available base of descriptions of etalons, the new space of images of component data as a part of values of their probabilistic measure of belonging to classes is created through training. The implementation of this approach, based on the use of the solution of the ensemble of components, provides the universality and significant effectiveness of the classification.

## IV. CALCULATION OF DISTRIBUTION VALUES FOR COMPONENTS

Let's calculate some center of description $\alpha(i) = F(E_i)$ – an aggregate vector that reflects the properties of the set $E_i$ [1], [2], [18] on the basis of the etalon description $E_i = \{e_v(i)\}_{v=1}^{s}$ from the base $E$ in $n$-dimensional vector space

$$\alpha(i) = (\alpha_1(i), \alpha_2(i), \ldots, \alpha_n(i)). \tag{2}$$

The center of the description $\alpha$ can be determined, for example, by calculating the average value or median for the fixed set of vectors [5], [16]. In this case the vectors $\alpha(i)$ are as the statistical characteristics for each etalon, they are the base of classification.

The separate vector $z_v \in Z$, $v = \overline{1, s}$ of etalon and object can be formally considered as the element of the ensemble of $n$-component vectors

$$z_v = (z_{v,1}, z_{v,2}, \ldots, z_{v,n}). \tag{3}$$

For the etalon $E_i$ each component in (3) is associated with the number $i$.

Let us consider the range of analyzed data in terms of assigning the component of the description to the system of etalon classes on the base of the definition of some membership function with values from the range 0…1 [18], [19]:

$$\mu : B^n \to [0, 1], \ \mu(z_v, i) \in [0, 1], \tag{4}$$

the arguments of the function $\mu$ are the descriptor of the description and the class number.

The membership function $\mu$ is determined on the basis of a basic characteristic in Data Science – the ratio of the values of measures that reflect the number of favorable cases (similarity $\eta(z_v, i)$ to the particular class) and the total number of cases $N$ (sum of similarities to all available classes) [1], [19]

$$\mu(z_v, i) = \frac{\eta(z_v, i)}{\sum_{i=1}^{N} \eta(z_v, i)}. \tag{5}$$

The total number of cases is specified by the number $N$ of classes. The degree of similarity $\eta(z_v, i)$ of the element to the class can be specified through the distance $\rho$ in the vector space to the center of the class, for example, through the Manhattan metric

$$\rho(z_v, i) = \sum_{k=1}^{n} \left| z_{v,k} - \alpha_k(i) \right|. \tag{6}$$

For the case where $\alpha(i) \in B^n$, $z_v \in B^n$ instead of (6), we can apply the Hamming distance $\chi(z_v, i)$ (number of mismatched bits) in space $B^n$, then the similarity $\eta(e_v, i)$ can be defined as

$$\eta(z_v, i) = n - \chi(z_v, i). \tag{7}$$

For each element $z_v$ using expression (5) let's calculate the value of the vector $d$ of its statistical distribution according to the set $N$ of classes

$$\begin{aligned} d &= (d(1), d(2), \ldots, d(N)), \\ d(i) &= \mu(z_v, i), \ \sum_{i=1}^{N} d(i) = 1. \end{aligned} \tag{8}$$

From the statistical point of view, the vector $d$, defined for the random descriptor of the etalon or object, reflects the degree of proximity of the value of the descriptor to each of the classes.

Let's consider the matrix $D = \{\{d_k(i)\}_{k=1}^{s}\}_{i=1}^{N}$ that, similarly to the fuzzy representation, specifies the value of the measure of belonging (5) for all components of the analyzed description [3], [18]. The value of $s$ means the number of descriptors in the etalon set, and the value of s is the same for all etalons and the input image. The row of the matrix $D$ is the distribution by etalon classes for descriptor of the input image. The sum of the values for each row is equal to 1. We believe that this calculation of distributions is close in meaning to the values of the membership function on the set of etalons in the image space.

In fact, the matrix $D$ determines the distribution of data sets according to the priori defined classes.

The values of the matrix $D$ allow the case to introduce logical processing of input data to remove possible interference (i.e. false descriptors) by analyzing the values of distances (6) or values (8) using the threshold. At the same time, this analysis can be directly implemented in the classification process.

## V. CLASSIFIER MODELS

Based on the matrix $D$, let's construct the classifier $K$ that implements the reflection for the structural description of the random object

$$K : D \rightarrow [1, 2, \ldots, N] \tag{9}$$

from the set of distributions of the ensemble of data components to the set of classes.

At the initial stage, let's construct and apply the distribution (8) according to the data classes directly for the set $E_i$ of database etalons to verify the operability. It is clear that for each representative $E_i \in E$ the classifier $K$ should receive number of the corresponding etalon, as the solution, which description transmits to the input of the classifier. This is the primary principle of the adequacy of the functioning of the classifier, which must confidently distinguish descriptions from the set of etalons (training sample). For example, in the data distribution for the 1st etalon (1st column of the matrix $D$), the first component must dominate over the others. Similarly, for the 2nd etalon, the dominant element of distribution should be the 2nd element, etc.

Let's consider some ways to construct the classifier (Figure 1).

**1. Determining the class of the object according to the integral distribution of components**.

Let's define the column of the matrix $D$ with the maximum sum of elements

$$K : j = \arg\max_{i} \sum_{v=1}^{s} d_v(i), \tag{10}$$

which integrally specifies the class $j$ of the object through aggregation of the distributions of each of the classes (the separate column of the matrix $D$) for the whole set of components of the description. Classification (10) corresponds to the most plausible solution, as it is based on adding the values of the same type of distributions [1]–[5]. In general, the classifier (10) implements the principle of analysis object – etalon based on the aggregation of data of the entire description.

**2. Determining the class of the object according to the number of voices of the components.**

Let's calculate the maximum value for each row of the matrix $D$

$$c_v = \arg \max_{i=1,\ldots,N} \{d_v(i)\}, \tag{11}$$

that is, we define separately for each descriptor of the description of the most important class according to the distribution vector that corresponds to the mode parameter [4], [8], [16].

For each descriptor $z_v$ of the description $Z$ we increment the number $h_i$ of voices of the $i$-th class elements which is corresponding to the value $c_v$ finding by the expression (11)

$$h_i = \begin{cases} h_i + 1, & c_v = i, \\ h_i, & c_v \neq i. \end{cases} \tag{12}$$

According to the result of definition (12) for the whole set of descriptors of the description $Z$ we obtain the vector

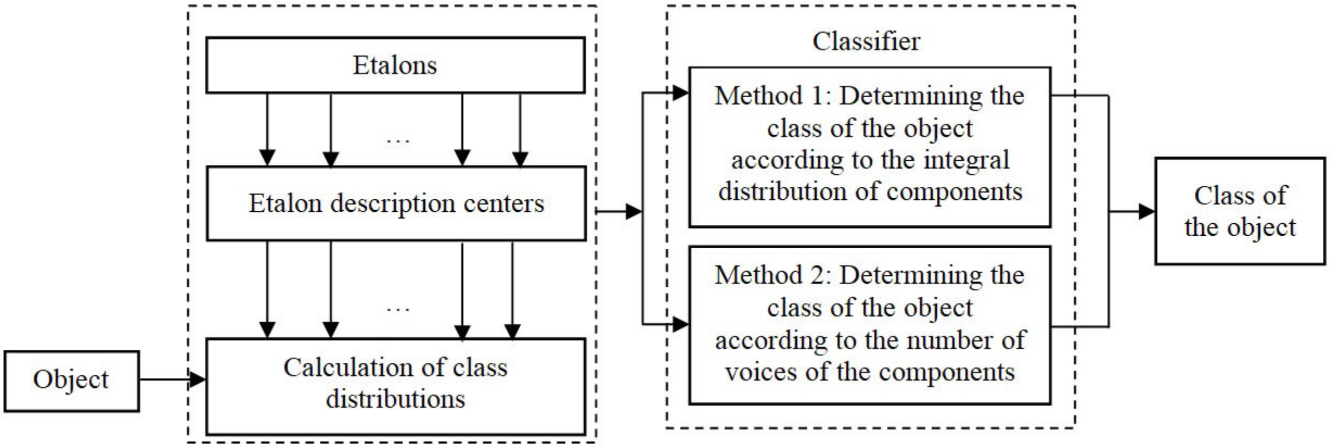



**FIGURE 1.** **Classification scheme in two ways.**

of the voices

$$h = (h_1, h_2, \ldots, h_N), \tag{13}$$

on the basis of which we determine the class number

$$j = \arg\max_b h_b, \tag{14}$$

which has gained the maximum value among the whole set of voices of object descriptors. This is the method of voting on the set of descriptors, where the class is determined on the basis of the distribution mode [5], [20].

The classifier (11)–(14) implements the principle of analysis object descriptor – etalon, where the class is predefined for each component of the description.

The considered variants (10), (11)–(14) of the construction of the classifier can naturally be interpreted within the framework of the theory of ensemble models [4], [8], [9], [20]. Due to the creation and aggregation of feedbacks of component classifiers (local solutions), the strong classifier with guaranteed higher efficiency of decision-making is synthesized. The most considered approaches correspond to the boosting model [9], [20].

According to research, the ensemble of classifiers in most cases provides better accuracy of data analysis or training. However, it requires solving a number of problems, such as increasing time and computational costs, the complexity of the results interpretation, justification and choice of ways to combine local solutions [2], [20]. In our case, the set of local classifiers consists of basic models of the same type, i.e. is homogeneous. The ensemble classification solution $\Theta$ can be represented as the combination of a finite set of local solutions $\theta_v$

$$K : j = \arg \mathop{opt}_i \Theta[\{\theta_v(d_v(i))\}_{v=1}^{s}]. \tag{15}$$

Note, that in ensemble models of image analysis in order to take into account only significant local solutions we often use a system of threshold parameters [1]–[4]. It provides separation of interference and is generally aimed at increasing the reliability of operation. For example, in the modification of the classifier (11)–(14) the class $d_m = \max_{i=1,\ldots,N}\{d_v(i)\}$ for the local solution is determined only if the condition is met, which compares the value of the optimum $d_m$ with the threshold $\delta_d$ or with the nearest local optimum $d_{m-1}$:

$$d_m > \delta_d \text{ or } d_m > \lambda d_{m-1}, \tag{16}$$

where $\lambda$ is the numerical coefficient.

In the proposed approach, the probabilities of belonging to all existing classes are calculated explicitly. It allows performing the further logical processing of their values to identify and take into account significant differences between classes for elements of the description.

Note, that the threshold $\delta_d$ can be specified on the base of the results of training with a teacher on the set of descriptors for etalons. The obtained value of mode must be not less than its value for their etalon. This analysis at the stage of training the classifier must be done for all etalons. As the threshold $\delta_d$ we have to choose the largest mode among all the obtained modes [2], [17].

One of the means to increase the effectiveness of ensemble classification methods by adapting to existing data is, on the base of introduction of logical procedures, to select some concentrated (or most informative) subset of elements, which will be the basis for the classification decision [2], [17], [29], [30].

Regarding the description as the set of keypoint descriptors, such a procedure may consist in selecting the subset of elements within a given distance from the center, or forming the fixed number of elements closest to the center of the description.

If $Z \subset B^n$ is the description, $s = card\, Z$ – its power, then let's introduce procedure $L$ of compression

$$L(Z) \to Z^*, Z^* \subset B^n, card\, Z^* = s^*, s^* \ll s. \tag{17}$$

The use of the filtering method (17) along with the significant reduction in computational time often contributes to the improvement of classification indicators [17], [31]. Compression $L$ should be directly aimed at selecting the informative part of the description. It is implemented at the stage of preprocessing, so it does not affect the time of classification.

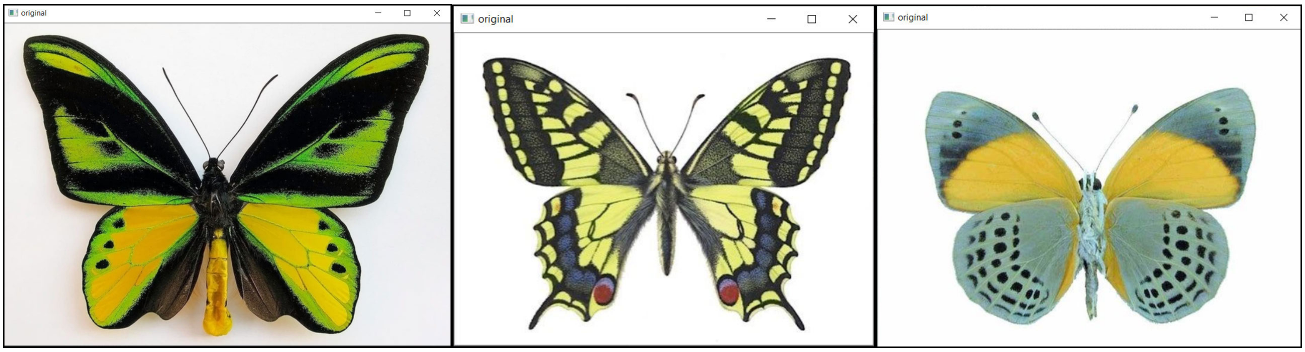



**FIGURE 2.** Etalon images of butterflies Chimera, Machaon, and Davis.

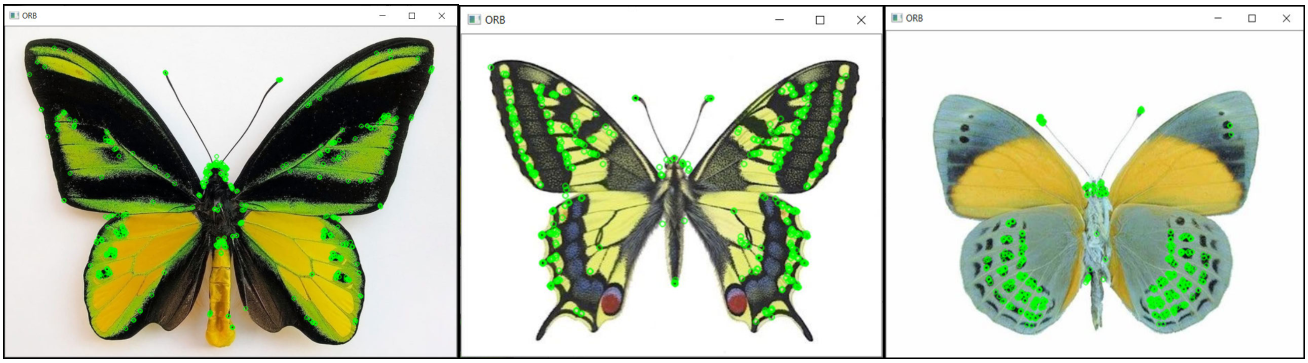



**FIGURE 3.** Examples of selected of keypoints coordinates by the ORB descriptor (Chimera, Machaon, Davis).

According to the results of pre-processing of the whole database $E$ of etalon data (the set of all etalon descriptors) on the basis of the distribution matrix, it is possible to evaluate the effectiveness of the proposed classification methods on the set of etalon descriptions.

For example, the precision Prec of classification can be estimated by the ratio of the total number TP of descriptors of the etalon base, for which the class is correctly defined, to their total number in accordance with the model [5], [20]

$$\mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}. \tag{18}$$

The index of completeness Compl (frequency of truly positive results) can be estimated as the number of descriptors correctly referred to the class (etalon)

$$\mathrm{Compl} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}. \tag{19}$$

Indicators (18), (19) are often used in machine learning systems [5], [32], [33].

## VI. RESULTS OF COMPUTER SIMULATION

The simulation has been performed in PyCharm 2020 using the OpenCV library and the Python 3.6 programming language [1], [34]–[36]. An ORB keypoint detector [35], [37], [38] of $n = 256$ dimension was used to determine the descriptors of keypoints.

The developed classifier models have been implemented on the example of images of butterflies from the database Leeds Butterfly [34], [39]. The size of the images is $600 \times 600$ pixels. The classes of etalon images (Figure 2) and the coordinates of the formed of keypoints (Figure 3) are illustrated. The number of calculated descriptors in the description of each of the etalon is $s = 500$.

By counting the number of 1-bits in each digit of the description of each of the etalons, vectors $\alpha(i)$, $i = 1, 2, 3$ have been obtained, which were logically reduced to the binary form for further implementation of the Hamming metric. The centers $\alpha(i)$ were actually determined by averaging the values of the description descriptors. The calculated values of the distances between the centers of the etalons were 56, 72, 72. This result indicates the significant their proximity in this aspect (28% of the maximum distance). After all, the values of the Hamming distance between the two vectors of the analyzed data are in the range $0, \ldots, 256$.

In accordance with the classification model (10), the values of the sum vectors in expression (10) were obtained as follows:

- For the 1st etalon (Chimera) – (173.01; 165.44; 161.51);
- For the 2nd etalon (Machaon) – (165.45; 170.94; 163.65);
- For the 3rd etalon (Davis) – (164.97; 163.69; 171.18).

According to the method (10) all etalons are classified correctly (the maximum is observed in the corresponding component), but the difference between the maximum of the integrated distribution and the nearest value is rather insignificant (within 3.5%–4.5%). This indicates the significant similarity of images in the formed space of features using the model (10).

According to the classifier model (11)–(14), three accumulated maximums of descriptor voices ($h_1, h_2, h_3$) were obtained for each of the etalons: (281, 111, 108), (144, 249, 107), (140, 86, 274).

From the obtained numerical values, we see that this classifier more confidently performs image recognition. The difference between the maximum of voices and its closest competitor exceeds 42%, and for some etalons (Chimera, Machaon) reaches 60%. This indicates the stronger perspective on the application of object descriptor – etalon classification models compared to the integrated analysis model (10).

Our experiments have confirmed that the proposed procedure for data analysis largely depends on the method of center formation $\alpha(i)$. The separate research has been carried out to determine the centers according to the median value for the set of detectors of the description [1], [17]. The distance between the medians of different etalons varies in the wider range than the distance between the centers-average values. In our experiment, there was a deviation from the maximum distance for the median of etalons in the range of 16%…46%. Thus, it turned out that the median for the set of descriptors is more sensitive to the content of the description than the average value.

In the process of research for the method of integral analysis (10) the experiments have been carried out to select the informative subset (fixed number) of description descriptors closest by Hamming's distance to the formed center in the form of the median using the model of logical processing (17) [17]. This action must be performed for the etalon descriptions and for the recognized image. With such a transformation, the values of the median and distributions for the components do not change, which means that additional calculations consist only in the formation of the compressed informative description (17).

The mentioned above compression significantly reduces the processing time in proportion to the reduction in the number of descriptors. Experimentally modified descriptions of etalons obtained in this way by selecting 50, 100 and 200 descriptors from a total number of 500.

The research has showed that using the proposed modification of compression, all etalon descriptions are classified correctly, and the calculated weight values $\{d(i)\}_{i=1}^{N}$ of the classes for the images of the etalons to some extent are improved with the decrease in the number of components.

With the decrease in the number of descriptors in the description, there is a corresponding concentration of values of the distributions around the centers, which contributes to the increase in indicators that correspond to the correct class. This, in turn, improves the effectiveness of the classification, as the class of the object is determined more confidently.

The difference between the maximum value of the distribution and the nearest value increases for different etalons from 3% for the complete description (500 descriptors) to 15% (200 descriptors), 21% (100 descriptors) and 25% (50 descriptors) for the modified description. Further compression of the analyzed descriptions according to the model (17) leads to the loss of informativeness of the data composition.

Note, that for general reasons, the effectiveness of the classification decision depends entirely on the available base of etalons and the set of input images, i.e. the content of the applied task. It is possible that for some applications, where there are significant differences between the images of the etalons, the most effective method may be the integrated method (10), based solely on the median values.

For the purpose of comparative analysis, experiments have been performed where the compressed subset of the description of 200–300 descriptors was selected by random selection and using specialized procedures for the analysis of the set of vectors [32]. It is confirmed, that such methods of data filtering do not have the perspective of application in the research approaches, as the effectiveness of method (10) significantly deteriorates, the indicator of correct classification, even for input etalon images, is reduced.

Let's now evaluate the effectiveness of the classifier object descriptor – etalon (11)–(14) according to the precision and completeness indicators (18), (19). Table 1 contains quantitative characteristics of the number of descriptors separately for each of the three etalons referred by the developed classifier to the corresponding class. In this case, all three etalons (the number of descriptors is equal to 500 for each etalon image) are classified correctly.

Table 2 contains the calculated values of Prec and Compl according to the Table 1.

Tables 1 and 2 are obtained directly for method 2 (classification by the number of votes of the components) as more promising in practice. The etalon images as input images were submitted to the classifier (the test set was taken from the Leeds Butterfly database). Table 1 contains the number of correctly and incorrectly classified descriptors as a result of testing the software model. Table 2 provides traditional indicators of precision (18) and completeness (19) for a set of descriptors. Based on the learning outcomes, the centers (medians) of the etalon images were calculated. The set of etalon image descriptors was used as a test set.

The average value of both indicators among the three analyzed etalons of the base was approximately 0.536. As you can see from Table 2, for the considered images the values of performance indicators are not high enough, as they are not close to 1, as it was obtained in our research [2] from the calculation example. This can be explained by the significant similarity of the researched objects visually (Figure 2) and in the constructed space of features. In addition, it is clear that, the accuracy naturally decreases with increasing number of classes.

According to the calculation according to Table 1 for truly negative results, which correspond to the fact of correct recognition of descriptors of other classes, the indicators Prec and Compl are much more significant: their average value is 0.768, and the minimum value is 0.716. This confirms sufficiently high effectiveness of the developed approach. Indeed, even in the presence of the high level of similarity of the

**TABLE 1.** Classification results.

| | | Number of descriptors | Number of descriptors for the rest etalons | Total |
|---|---|---|---|---|
| **Etalon 1** (Chimera) | Referred to 1 | 281 | 144 + 140 = 284 | 565 |
| | Referred to 2 and 3 | 219 | 356 + 360 = 716 | 935 |
| | Total | 500 | 1000 | 1500 |
| **Etalon 2** (Machaon) | Referred to 2 | 249 | 111 + 86= 197 | 446 |
| | Referred to 1 and 3 | 251 | 389 + 414 = 803 | 1054 |
| | Total | 500 | 1000 | 1500 |
| **Etalon 3** (Davis) | Referred to 3 | 274 | 107 + 108 = 215 | 489 |
| | Referred to 1 and 2 | 226 | 393 + 392 = 785 | 1011 |
| | Total | 500 | 1000 | 1500 |

**TABLE 2.** Values of Prec and Compl indicators for truly positive results.

| | **Etalon1** (Chimera) | **Etalon 2** (Machaon) | **Etalon 3** (Davis) | **Minimum** |
|---|---|---|---|---|
| **Prec** | 0.497 | 0.55 | 0.56 | 0.497 |
| **Compl** | 0.562 | 0.498 | 0.548 | 0.498 |

researched objects, the number of identified different descriptors for different descriptions is still actually much higher, which has been confirmed by experimental calculations.

Given the undoubted influence of the multidimensionality factor of the data used (descriptor size is 256 components, number of description descriptors is 500), it can be considered necessary for providing reliable recognition of images of similar content in databases to apply the reduction of dimensionality (compression or concentration) of descriptions by introducing clustering or hashing procedures [1], [15], [32], which can contribute to obtain more significant indicators of the effectiveness of the classification.

The essence of the study is to develop methods of structural classification based on descriptions from a set of descriptors of keypoints. It is clear that the effectiveness of any classification method directly depends on the base of etalon images. The key idea of the paper is to study and check the efficiency of classification using the apparatus of distribution of descriptions for etalon classes.

The figures in Table 2 are not impressive, but they only show the evaluation results for the specific descriptors of the description.

At the same time, in the experiment all etalon images are classified correctly by both proposed methods, that is the actual precision and completeness of the classification in terms of the description of the input image is 1. This classification model corresponds to boosting [5], [12], [20].

The key result of the analysis is that in this case, despite the not very high performance of each of the components of the description, their common classification solution provides a guaranteed high performance of the classifier (14), (15).

Testing showed that on many etalons the proposed method worked without errors, therefore, the accuracy of classification is equal to 1.

The speed of classification compared to the traditional approach of descriptor voting, where linear search is used, improves proportionally number of etalons. For the test example, the processing time was approximately 5 seconds (developed method) and 34 seconds (traditional), which corresponds to a gain of approximately 7 times. The main reason for the gain is that in the developed method the input descriptors are metrically compared only with the centers of etalon descriptions.

The article does not set the task of obtaining efficiency higher than all known methods; we are just beginning to study a promising method.

It is clear that with increasing number of etalons in the applied database Leeds Butterfly the accuracy of classification for this method may decrease. This can be explained by the significant similarity of the shape of objects (butterflies) in the database. At the same time, for other image databases where there is insignificant similarity, the accuracy may be maintained at a high level. Construction of the classification accuracy curve depending on the number of classified etalons or other undesirable effects (noise, interference, background images) may be a prospect for further research in the implementation of the developed method.

## VII. CONCLUSION

The developed method of classification on the basis of statistical distributions of components of descriptions is constructed on the principle of separation of data components according to the classes. According to the results of experiments this method confirms its efficiency and effectiveness for the image classification. Its effectiveness for three-dimensional image databases can be enhanced by the adapted choice of metrics or similarity measures, by the choice of way to form centers for etalon descriptions, by the introduction of logical processing or compression of the structural description. More confident experimental results of classification have been shown by the model which has an optimum in the vector of distributions for some descriptors (distribution mode). The use of the concentrated part of the description data allows to improve its distinction with other descriptions.

For the case of multiclass classification, the uncertainty in decision-making increases with increasing number of classes. As a result, the precision decreases, as it is difficult to ensure that the new etalons do not resemble existing ones. The solution is to make a decision about each class separately, or ranking according to the confidence indicator.

The scientific novelty of the research is the development of the effective method of image classification based on the introduction of the apparatus of statistical distributions for the components of the description. It implements in-depth analysis in the data space and ensures the effectiveness of the classification even for similar images. The classifier is implemented in the variants of integration of component distributions according to the classes and on the basis of mode analysis for distributions of individual components. The use of the median as the center of description has the advantage over the average value.

The practical significance of the research is:

- Construction of classification models in the synthesized data space;
- Confirmation of the efficiency of the proposed modifications of data analysis by the examples of images;
- Development of software models for the implementation of the proposed classification methods for the computer vision systems.

The perspective of the development may be related to the research of the interference immunity of the developed methods and the evaluation of their applied performance in relation to three-dimensional image collections, where the distribution coefficients can take small absolute values.

## REFERENCES

[1] V. O. Gorokhovatskyi and S. V. Gadetska, "Structural methods of image classification according to the description of the set descriptors of key-points," in *Statistical Process and Data Mining in Structural Image Classification Methods (Monograph)*. Kharkiv, Ukraine: FLP Panov A. N., 2020, pp. 10–28.

[2] V. O. Gorokhovatskyi, S. V. Gadetska, N. I. Styahlyk, and N. V. Vlasenko, "Classification of images based on an ensemble of statistical distributions by etalon classes for structural description components," *Radio Electron. Comput. Sci. Control*, vol. 4, pp. 85–94, Dec. 2020, doi: 10.15588/1607-3274-2020-4-9.

[3] Y. I. Daradkeh, I. Tvoroshenko, V. Gorokhovatskyi, L. A. Latiff, and N. Ahmad, "Development of effective methods for structural image recognition using the principles of data granulation and apparatus of fuzzy logic," *IEEE Access*, vol. 9, pp. 13417–13428, Jan. 2021, doi: 10.1109/ACCESS.2021.3051625.

[4] R. O. Duda, P. E. Hart, and D. G. Stork, "Unsupervised learning and clustering," in *Pattern Classification*. Hoboken, NJ, USA: Wiley, 2000, pp. 517–580.

[5] P. Flach, "Features," in *Machine Learning. The Art and Science of Algorithms That Make Sense of Data*. New York, NY, USA: Cambridge Univ. Press, 2012, pp. 298–329.

[6] C. Celik and H. S. Bilge, "Content based image retrieval with sparse representations and local feature descriptors: A comparative study," *Pattern Recognit.*, vol. 68, pp. 1–13, Aug. 2017, doi: 10.1016/j.patcog.2017.03.006.

[7] X. Zhang, F. X. Yu, S. Karaman, and S.-F. Chang, "Learning discriminative and transformation covariant local feature detectors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4923–4931, doi: 10.1109/CVPR.2017.523.

[8] F. C. Porter, "Testing consistency of two histograms," 2008, *arXiv:0804.0380*. [Online]. Available: https://arxiv.org/abs/0804.0380

[9] Y. I. Daradkeh and I. Tvoroshenko, "Technologies for making reliable decisions on a variety of effective factors using fuzzy logic," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 5, pp. 43–50, 2020, doi: 10.14569/IJACSA.2020.0110507.

[10] G. Sharma and B. Schiele, "Scalable nonlinear embeddings for semantic category-based image retrieval," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1296–1304, doi: 10.1109/ICCV.2015.153.

[11] J. Wu, W. Lin, G. Shi, Y. Zhang, W. Dong, and Z. Chen, "Visual orientation selectivity based structure description," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4602–4613, Nov. 2015, doi: 10.1109/TIP.2015.2460467.

[12] R. Szeliski, "Recognition," in *Computer Vision: Algorithms and Applications*. London, U.K.: Springer-Verlag, 2010, pp. 655–718.

[13] J. Yang, W. Zhang, X. Li, T. Zhou, and B. Ou, "Full reference image quality assessment by considering intra-block structure and inter-block texture," *IEEE Access*, vol. 8, pp. 179702–179715, Oct. 2020, doi: 10.1109/ACCESS.2020.3028282.

[14] Y. I. Daradkeh and I. Tvoroshenko, "Application of an improved formal model of the hybrid development of ontologies in complex information systems," *Appl. Sci.*, vol. 10, no. 19, p. 6777, Sep. 2020, doi: 10.3390/app10196777.

[15] O. Gorokhovatskyi, V. Gorokhovatskyi, and O. Peredrii, "Analysis of application of cluster descriptions in space of characteristic image features," *Data*, vol. 3, no. 4, p. 52, Nov. 2018, doi: 10.3390/data3040052.

[16] V. Gorokhovatskyi and I. Tvoroshenko, "Image classification based on the Kohonen network and the data space modification," in *Proc. CEUR Workshop Comput. Modeling Intell. Syst. (CMIS)*, vol. 2608, Apr. 2020, pp. 1013–1026. [Online]. Available: http://ceur-ws.org/Vol-2608/

[17] V. O. Gorokhovatskyi, S. V. Gadetska, and R. P. Ponomarenko, "Logical analysis and processing of data for the classification of images on the basis of formation of statistical center description," *Control, Navigat. Commun. Syst.*, vol. 4, no. 56, pp. 43–48, Apr. 2019, doi: 10.26906/SUNZ.2019.4.043.

[18] I. S. Tvoroshenko and V. O. Gorokhovatsky, "Intelligent classification of biophysical system states using fuzzy interval logic," *Telecommun. Radio Eng.*, vol. 78, no. 14, pp. 1303–1315, 2019, doi: 10.1615/TelecomRadEng.v78.i14.80.

[19] S. Kim and I. S. Kweon, "Biologically motivated perceptual feature: Generalized robust invariant feature," in *Proc. Asian Conf. Comput. Vis. (ACCV)*, vol. 3852, Jan. 2006, pp. 305–314, doi: 10.1007/11612704_31.

[20] N. B. Paklin and V. I. Oreshkov, "Data Mining: Clustering," in *Business Analytics: From Data to Knowledge: Textbook*. Saint Petersburg, Russia: Peter, 2013, pp. 308–341.

[21] K. M. Hamdia, M. Arafa, and M. Alqedra, "Structural damage assessment criteria for reinforced concrete buildings by using a fuzzy analytic hierarchy process," *Underground Space*, vol. 3, no. 3, pp. 243–249, Sep. 2018, doi: 10.1016/j.undsp.2018.04.002.

[22] Y. Yang, X. Wang, B. Sun, and Q. Zhao, "Channel expansion convolutional network for image classification," *IEEE Access*, vol. 8, pp. 178414–178424, Sep. 2020, doi: 10.1109/ACCESS.2020.3027879.

[23] X. Zhang, "Content-based E-commerce image classification research," *IEEE Access*, vol. 8, pp. 160213–160220, Aug. 2020, doi: 10.1109/ACCESS.2020.3018877.

[24] H. M. R. Afzal, S. Luo, M. K. Afzal, G. Chaudhary, M. Khari, and S. A. P. Kumar, "3D face reconstruction from single 2D image using distinctive features," *IEEE Access*, vol. 8, pp. 180681–180689, Oct. 2020, doi: 10.1109/ACCESS.2020.3028106.

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

[26] L. Armi and S. Fekri-Ershad, "Texture image classification based on improved local quinary patterns," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 18995–19018, Jul. 2019, doi: 10.1007/s11042-019-7207-2.

[27] G. Deep, L. Kaur, and S. Gupta, "Local quantized extrema quinary pattern: A new descriptor for biomedical image indexing and retrieval," *Comput. Methods Biomech. Biomed. Eng. Imag. Vis.*, vol. 6, no. 6, pp. 687–703, Jul. 2017, doi: 10.1080/21681163.2017.1344933.

[28] S. Fekri-Ershad and F. Tajeripour, "Color texture classification based on proposed impulse-noise resistant color local binary patterns and significant points selection algorithm," Jun. 2019, *arXiv:1906.11010*. [Online]. Available: https://arxiv.org/abs/1906.11010

[29] T. Mensink, J. Verbeek, F. Perronnin, and G. Csurka, "Distance-based image classification: Generalizing to new classes at near-zero cost," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2624–2637, Nov. 2013.

[30] A. Iscen, G. Tolias, P.-H. Gosselin, and H. Jegou, "A comparison of dense region detectors for image search and fine-grained classification," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2369–2381, Aug. 2015, doi: 10.1109/TIP.2015.2423557.

[31] Y. Liu, F. Wei, J. Shao, L. Sheng, J. Yan, and X. Wang, "Exploring disentangled feature representation beyond face identification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Dec. 2018, pp. 2080–2089, doi: 10.1109/CVPR.2018.00222.

[32] J. Leskovec, A. Rajaraman, and D. J. Ullman, "Clustering," in *Mining of Massive Datasets*. New York, NY, USA: Cambridge Univ. Press, 2014, pp. 241–276.

[33] 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, Oct. 2020, doi: 10.1109/ACCESS.2020.3028740.

[34] *Learning Models for Object Recognition From Natural Language Descriptions*. Accessed: Feb. 20, 2021. [Online]. Available: https://www.josiahwang.com/projects/leedsbutterfly

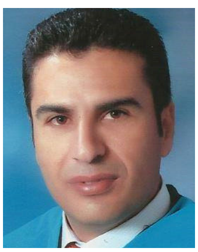
[35] *ORB Feature Detector and Binary Descriptor*. Accessed: Feb. 22, 2021. [Online]. Available: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_orb.html

[36] *OpenCV Open Source Computer Vision*. Accessed: Feb. 24, 2021. [Online]. Available: https://docs.opencv.org/master/index.html

[37] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571, doi: 10.1109/ICCV.2011.6126544.

[38] E. Karami, S. Prasad, and M. Shehata, "Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images," 2017, *arXiv:1710.02726*. [Online]. Available: https://arxiv.org/abs/1710.02726

[39] J. Wang, K. Markert, and M. Everingham, "Learning models for object recognition from natural language descriptions," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Sep. 2009, pp. 1–11, doi: 10.5244/C.23.2.



**YOUSEF IBRAHIM DARADKEH** received the Doctor of Engineering Sciences (Ph.D. and P.Eng.) degrees in computer engineering and information technology (computer systems engineering and computer software engineering).

He was a dynamic academician having more than 15 years of experience in teaching and scientific research development and administration experience. He was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, University of Calgary, Canada. He has taught wide spectrum of computer science, computer engineering and networks, and computer software engineering courses (undergraduate and graduate degrees). He has an excellent experience in designing courses that bridge the gap between academia and industry and follow the accreditation requirements. He is currently an Associate Professor with the Department of Computer Engineering and Networks, College of Engineering at Wadi Addawasir, Prince Sattam Bin Abdulaziz University, Saudi Arabia. He is a Senior Scientific Researcher and the Assistant Dean for the Administrative Affairs. He is also a well-known and respected scientist internationally. He has published over 90 high-quality refereed research articles in the international journals and conferences. He has also published two books, one chapter, and edited book in the most prestigious publications. His international recognition of scientific achievements is demonstrated by numerous invitations to participate with the program committees of the international conferences and foreign journals and lecturing with renowned scientific centers around the world. He has a membership of the International Academy of Science and Engineering for Development (IASED).

**VOLODYMYR GOROKHOVATSKYI** received the M.Sc. degree in applied mathematics and engineering from the 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 tool of artificial intelligence, in 2010.

He received an Internship from Dresden Technical University. He is currently working as a Professor with the Department of Informatics, KNURE. He has more than 250 scientific articles and six monographs. His research interests include the image and pattern recognition in computer vision systems, structural methods of image classification and recognition, and artificial intelligence.

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

She is currently an Associate Professor with the Department of Informatics, Kharkiv National University of Radio Electronics. She has published 146 scientific articles and educational and methodical articles include four study guides, six monographs, 33 articles, 79 abstracts of reports, 11 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. She was awarded the Diploma of the Scholar in the field of science (in informatics and computer science) named after V. V. Sviridov for significant achievements.

**SVITLANA GADETSKA** received the Ph.D. degree in physics and mathematics, in 1993.

She is currently working as an Associate Professor with the Department of Higher Mathematics, Kharkiv National Automobile and Highway University. She has published more than 90 scientific and methodical articles. Her research interests include computer vision systems and mathematical modeling.

**MUJAHED AL-DHAIFALLAH** (Member, IEEE) received the B.Sc. and M.Sc. degrees in systems engineering from the King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, and the Ph.D. degree in electrical and computer engineering from the University of Calgary, Calgary, AB, Canada.

Since 2020, he has been an Associate Professor of systems engineering with the King Fahd University of Petroleum and Minerals. His current research interests include nonlinear systems identification, control systems, optimization, artificial intelligence, and renewable energy.