Search for Visual Objects by Request in the Form of a Cluster Representation for the Structural Image Description

Volodymyr GOROKHOVATSKYI [1], *Iryna TVOROSHENKO* [1], *Oleg KOBYLIN* [1], *Nataliia VLASENKO* [2]

[1]Department of Informatics, Kharkiv National University of Radio Electronics, Nauky Ave. 14, 61166 Kharkiv, Ukraine
[2]Department of Informatics and Computer Engineering, Simon Kuznets Kharkiv National University of Economics, Nauky Ave. 9-A, 61166 Kharkiv, Ukraine

volodymyr.horokhovatskyi@nure.ua, iryna.tvoroshenko@nure.ua, oleg.kobylin@nure.ua, nataliia.vlasenko@hneu.net

DOI: 10.15598/aeee.v21i1.4661

Article history: Received Aug 09, 2022; Revised Jan 10, 2023; Accepted Feb 12, 2023; Published Mar 31, 2023.

Abstract. *The key task of computer vision is the recognition of visual objects in the analysed image. This paper proposes a method of searching for objects in an image, based on the identification of a cluster representation of the query descriptions and the current image of the window with the calculation of the relevance measure. The implementation of a cluster representation significantly increases the speed of identification or classification of visual objects while maintaining a sufficient level of accuracy. Based on the development of models for the analysis and processing of a set of descriptors of keypoints, we have obtained an effective method for the identification of visual objects. A comparative experiment with the traditional method has been conducted, where a linear search for the nearest descriptor was implemented for identification without using a cluster representation of the description. In the experiment, a speed gain for the developed method has been obtained in comparison with the traditional one by approximately 5.2 times with the same level of accuracy. The method can be used in applied tasks where the time of object identification is critical. The developed method can be applied to search for several objects of different classes. The effectiveness of the method can be increased by varying the values of its parameters and adapting to the characteristics of the data.*

Keywords

Computer vision, detector, Hamming metric, k-means method.

1. Introduction

Detection, identification, and classification of objects are the key tasks of modern computer vision systems that establish instances of visual objects of a certain class (e.g., people, animals, or cars) on the digital images [1], [2], [3] and [4]. Such intellectual tasks are solved for general purposes of research of methods for identification of different types of objects in accordance with the unified framework for imitation of human vision and cognition [5], [6], [7], [8] and [9] and for the application purposes according with specific application scenarios, such as detection of pedestrians, faces, text, movements, etc. In recent years, the rapid development of methods of deep learning has contributed to the new achievements to the subject of detection, which results in the breakthrough and progress, especially for applied implementations [10], [11], [12] and [13]. Detection objects have been used now in many real-world applications, such as autonomous driving, robot vision, video surveillance of moving objects, and more.

Structural methods of image classification have become popular because of their applied efficiency for computer vision tasks, where identification or classification of visual objects is carried out [1], [3] and [14]. Here, traditionally, the set of points of a recognized object is formed by analyzing the part of the image that is highlighted by a scanning frame, namely window that allows to partially exclude background objects during the analysis. For each position of the frame, the decision is made about identification (two classes) or classification (several classes as etalons) based on the relevance value of the query and the image inside the window.

When implementing structural classification methods, the function of image brightness is represented by the set of keypoints, each of which is described by the vector of features - the keypoints descriptor [2].

The formal statement of the classification task based on the description as the set of keypoint descriptors is given in the literature [1].

Identification of visual objects on the scene image can be successfully implemented for method of matching the description of the fragment of the object image and the cluster representation of the etalon as the query for search [2] and [3]. Only the cluster representation due to the significant compression of the description (as a rule, the volume of the analysed description is 500 keypoint descriptors and more) allows to search for the object in real-time. Due to the transition from the set to multidimensional data centres, computational costs, and decision-making time are significantly reduced [3], [4], [5], [7], [15], [16], [17] and [18].

The purpose of the article is to develop for a method for searching visual objects in the image using the cluster representation for the structural description of the query image.

Research tasks are:

- Development of mathematical and software models of data mining when determining the measure of relevance of structural descriptions of the window and query.
- Study of the features of model use to determine relevance with implementing clustering of query data.
- Evaluation of the effectiveness of the developed method in according with the results of the analysis of specific images.

2. Related Works

The identification and classification of objects on the image is the key task of intelligent computer vision systems [1] and [13]. Now researchers mainly focus on methods that are directly aimed at applied implementation. Due to the multi-dimensional and spatial nature of the image signal, statistical approaches have become the most popular for solving this task [1], [3], [10], [11], [12], [13], [14], [19] and [20]. Recently, specialized software tools have been developed based on prior training of the neural network within some fixed image base [2], [4], [10], [11], [19], [20], [21], [22] and [23]. For example, the You Only Look Once (YOLO) network divides images into parts and provides constraints and confidence parameters for each part simultaneously [19] and [23]. The series of improvements based on YOLO has been created, and new versions have been proposed that further improve the parameters of versatility, confidence, and accuracy while maintaining the high identification rate [19], [21], [22], [23], [24], [25] and [26]. However, the limitations of such systems are the need for prior long-term training and the dependence of the application results on the specific base on which the training is carried out.

Despite the existence of effective systems based on machine training, the development and validation of new methods of object search continue [1], [2], [3], [4], [10], [11], [12], [14], [20], [27], [28], [29] and [30]. The new promising direction is the use of descriptions of visual objects as the set of keypoint descriptors. This apparatus provides high-speed data analysis and allows for classification to determine in detail the characteristics of the object detected in the image. It is acceptable to combine different methods to increase efficiency. Additional implementation of training for such systems will further improve their characteristics [2], [4] and [14].

It should be noted that the classification methods based on a set of descriptors by their nature differ from the YOLO apparatus [21], [22], [23] and [24] positively by the simplicity of technical implementation, direct application without prior long-term training, universality concerning the variability of the etalon base.

Note that the cluster representation of the structural description of the image as a set of descriptors [14], [15], [20], [33] and [36] improves the computing performance of classifiers tenfold compared to traditional methods [3], [14] and [15]. It is explained by the implementation of a two-stage search for optimal matching of the components of the object as part of the etalon (through the centres of the clusters) instead of a full-fledged linear search. The method of comparing descriptions in the form of vectors of quantitative cluster composition [1], [3], [15] and [34] also has advantages in the computational sense due to the implementation of a granular presentation of the analysed data in the form of a "bag of words" model [2], [20] and [34]. But the applied application of these methods requires a deeper study since their effectiveness depends significantly on the influence of several factors: the way of separating the background and objects from each other in the image, the composition of the etalon base, the chosen method of clustering, the number of descriptors in the description, the value threshold for the equivalence of descriptors, etc.

The proposed research contains the results of an in-depth study of applied features for the technical implementation of the cluster apparatus for identifying a given object.

3. Mathematical Identification Models Descriptions for Query Image

Let us universally describe the recognizable visual object (request, etalon) as a finite set $Z = \{z_v\}_{v=1}^{s}$, where $z_v \in Z$ are keypoints descriptors, $s = card\, Z$ is its cardinality [2]. For binary descriptors Oriented FAST and Rotated BRIEF (ORB) $Z \subset B^n$, B^n - the space of binary vectors of dimension [11] and [12]. We apply the cluster partition of set Z through reflection $Z \to T$. As a result, the description of the input image of the object will be represented by M disjoint clusters:

$$Z = T\left(Z\right) = \left\{T_k\left(Z\right)\right\}_{k=1}^{M}, T_k\left(Z\right) \cap T_j\left(Z\right) = \phi, \tag{1}$$

where $T_k\left(Z\right)$ is a set of elements of a fixed cluster.

The choice of the number M of clusters is an exclusively applied problem and depends on the content of the analysed data. With an increase in M, the accuracy of the analysis of data groupings increases, but the processing time also increases. In our research, the value $M \in \{3, \ldots, 10\}$ is used for descriptions in the form of a set of descriptors [1], [2], [3], [4] and [11].

Based on the clustering result for each cluster $T_k\left(Z\right)$ from the description of query Z, we will determine the parameters of the centres $T_k\left(Z\right)$ and capacities of $c_k\left(Z\right)$ clusters:

$$c_k\left(Z\right) = card\, T_k\left(Z\right),\ k = \overline{1, M}. \tag{2}$$

Now let us consider windows n fixed by the number $W_1, \ldots, W_u$, $W_i \subset B^n$ hich are separate fragments of the image inside which the desired objects can be located, represented by a set of keypoint descriptors. Such fragments can be synthesized in the established order of the image review, depending on the applied problem [13]. The number of fragments affects the processing time. To ensure the equivalence of the influence of the analysed data on the analysis result, we will consider the parameter value for each description from the set of windows $W_1, \ldots, W_u$ to be the same:

$$card\left(Z\right) = card(W_1) = \cdots = card\left(W_u\right) = s. \tag{3}$$

Condition (Eq. (3)) can always be practically achieved by fixing the value s for query Z and selecting s elements from sets $W_1, \ldots, W_u$ of larger size. Otherwise, additional standardization of data by the number of description elements is required.

We will reduce the identification to the establishment of the relevance degree $\rho\left(W_i, Z\right)$ of the object W_i and query Z presented in the cluster form, followed by a decision based on the value $\rho\left(W_i, Z\right)$. For each descriptor $w \in W_i$, we competitively determine the nearest cluster centre in the set of vectors $\{b_j\left(Z\right)\}$ according to the nearest neighbour procedure:

$$d = \arg \min_{j=1,\ldots,M} \rho\left(w, b_j\left(Z\right)\right),\ d \in \{1, 2, \ldots, M\}, \tag{4}$$

where ρ is the distance between the object descriptor and centre b_j from the cluster system for the query. The processing procedure Eq. (4) is sometimes referred to as designing for multiple cluster centres [3] and [31].

By using binary descriptors and centres in Eq. (4), the Hamming distance can be applied. For the most common clustering procedures, where vector data with non-integer components (k-means, hierarchical classification, etc. [20], [31], [32] and [33]) are used, the Manhattan distance can be applied.

Based on the results of processing Eq. (4) $\forall w_a \in W_i$, the number of $h_1, h_2, \ldots, h_M$ elements of the analysed description, assigned to one of the cluster centres $\{b_j\}_{j=1}^{M}$, is calculated:

$$h_j = \sum_{a=1}^{s} f_a\left[w_a \to \{b_j\}\right], \tag{5}$$

where f_a is a logical function that determines the assignment of the description element to the corresponding centre j of the query cluster according to the concurrency model Eq. (4).

The procedure for implementing function f_a to ensure filtering of interference, which is certainly present in the images, should be based on the value of threshold δ_p for the minimum value in Eq. (4) [1]. Decision $w_b \to b_j$ is made under condition $\rho(w_a, b_j) \leq \delta_\rho$, where δ_ρ is determined experimentally, based on the composition of the etalon images of the analysed base.

Based on the calculation of the components of the vector Eq. (5), we define the relevance measure as the distance γ between the integer vectors $h = \{h_1, h_2, \ldots, h_M\}$ for the request and the

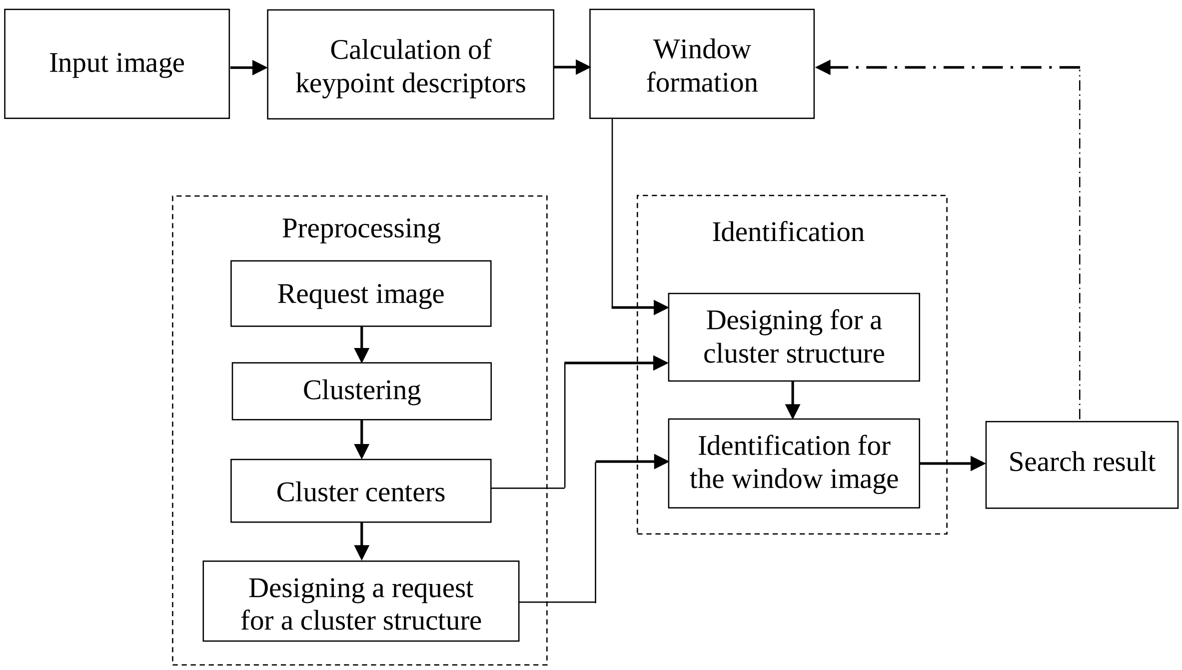


Fig. 1: The object search scheme in the image.

current window $\gamma\left(h\left[Z\right], h\left[W_i\right]\right)$. Depending on the obtained value; we determine the identification decision W_i. At γ, we can take the Manhattan distance in a M -dimensional vector space. Also, in this case, the similarity of vectors, for example, the correlation coefficient, cans a measure of relevance [34].

Note that in models Eq. (4) and Eq. (5) descriptions W_i are processed independently of each other, which makes it possible to make decisions about several search objects in the image [35], [36] and [37].

Consider a step-by-step implementation of the proposed method in the form of pre-processing and identification stages. The pre-processing stage does not affect the time for making an identification decision and contains the following steps:

- Calculate the keypoints descriptors of the etalon request.
- We carry out clustering of the structural description of the etalon.
- Determine the centres and powers of the clusters.

The identification stage can be viewed as a sequence of actions (see Fig. 1):

- Define the set of keypoint descriptors of the recognized image fragment (window).
- We project the considered window description onto the structure of the cluster representation of the request (cluster centres).
- Determine the measure of relevance (distance, similarity) of the fragment views and cluster query centres.
- By the value of the relevance measure, we make a decision regarding the identification of the request and the window image.

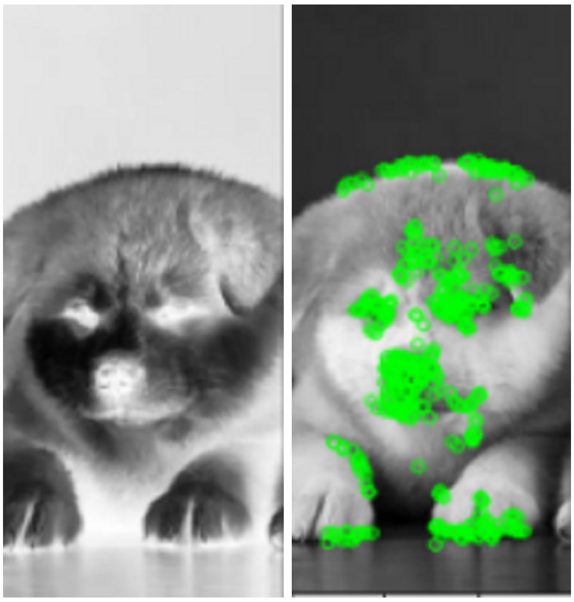
Thus, the essence of identification is to establish significance for the degree of relevance to the query and the composition of the analysed window, projected onto the centres of the clusters for the query.

4. Software Simulation Results

For the research, the Jupyter Notebook software environment was used on the Google Colaboratory service. A program that simulates the search on-demand method written in the Python language using specialized libraries for working with images: Scikit-image and OpenCV [38]. The input image contains 4 objects (puppies), of which object No. 1 (puppy on the left) is used as a request. Fig. 2 shows the input image, and Fig. 3 contains its gray-scale representation, which is processed by the keypoints detector. Some visually noticeable difference between object No. 4 Fig. 2 (puppy on the right) from the rest, it turned out in the experiment too. Human vision easily perceives this object as a puppy since the image created by the human brain is turned on. An artificially intelligent system [27] and [30] decides solely based on a set of informative image points for which keypoint descriptors are calculated [29].

Fig. 2: Analysed image.

Fig. 3: Image of Fig. 2 gray-scale.



(a) (b)

Fig. 4: Request (etalon) and its coordinate's keypoints.

The image size is 600×314 pixels. The search for objects identical to the query was performed by scanning with a frame 1/4 sizes relative to the input image. Fragment No. 1 (the first puppy on the left) with a size of 150×314 was taken as an etalon request (see Fig. 4). The ORB detector is used, which forms a description in the form of a set of about 500 binary descriptors with a dimension of 256 bits [37]. Implementing the method was carried out under conditions of the same number of descriptors in the descriptions of the request and the fragment under consideration.

The discussed method assumes a comparison of an equal number of components. This condition is achieved by changing the used number of keypoints of the considered fragment. Clustering for the description of the request was performed using the k-means method using the Manhattan metric and the number of clusters $k = 3$.

Figure 4 shows the set of coordinates (centres of the green rings) obtained by the ORB detector for the query image. As you can see, the main visual information is quite clearly highlighted by the detector.

First, a pilot experiment was carried out in which the image fragments under consideration (the current scanning window) were independently clustered during the scanning and identification process. Thus, even for object No. 1 (etalon), re-clustering was performed. The centres of the clusters were different for the same image. With such complicated processing, it is difficult to count on success. But even in this case, objects No. 1 and No. 3 of the scene were identified. They showed a fairly small value of the Manhattan distance between the data histograms in the range from 46–115 (the maximum value of the distance is 500).

The next experiment was to implement a method where the cluster representation was performed at a time for a request. The considered structural descriptions of the current image were projected onto the fixed centres of the query clusters by establishing the closest one. The computational costs for this approach are much less. Data clustering occurs only once per query.

The histogram of the cluster representation of the request (object No. 1) for the number of 500 keypoint descriptors is shown in Fig. 5. Columns of the histogram contain the number of image descriptors assigned to the corresponding cluster.

The results of the experiment showed that objects No. 1–No. 3 Fig. 2 are identified accurately (distances are 0, 80, 90), and object No. 4, according to the results of the analysis, showed a significant difference (distance 212 at a maximum of 500). One fragment containing parts of two different objects showed a distance of 162, which is closer to the etalon than object No. 4 (see Fig. 6). Experiments were also carried out with a different number of keypoints, which were

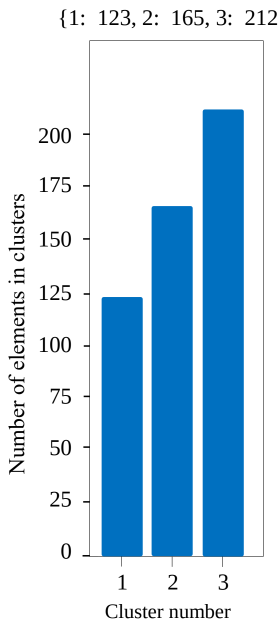


Fig. 5: Histogram of request projections with the number of keypoints 500.

randomly selected from a description of 500 points. With a decrease in the number of keypoints, the computational efficiency improves, but the resolution properties decrease [3].

The best-applied efficiency was shown by the processing option, with the number of keypoint sequel to 200. It is for the request image in Fig. 4, the number of descriptors in the clusters was a vector (60, 46, 94). According to the simulation results, all puppy objects No. 1–No. 4 were identified correctly (distances 0, 26, 30, 42 with a maximum of 200), while the other analysed windows showed significantly larger distances.

Fig. 6: Fragment with a distance of 162.

Figure 6 demonstrates general difficulties for artificial intelligence systems that may arise in the process of identifying objects against a complex background. We solved this problem by an experimental selection of system parameters. In general, to improve performance, the query image can be expanded by additional training of the system. Due to this, information about the characteristic composition of object No. 4 will be included in the image.

For a fixed number of 200 keypoint descriptors, we carried out a comparative experiment, where for identification we implemented the traditional voting method based on a linear search for the nearest descriptor without using the preliminary procedure of cluster presentation of the description. The experiment showed a gain in speed for the developed method in comparison with the traditional 5.2 times. Here, the value of the gain depends on the parameter of the number of clusters and increases with an increase in the number of clusters within 2–8.

As seen from the experiment, the effectiveness of the method can be enhanced by changing the values of its parameters and adapting to the properties of the data.

5. Conclusion

The proposed methods for searching for objects on the image using the clustering apparatus are characterized by the high speed of data processing and sufficient efficiency. The experiment, conducted for the task of identifying several objects in the image with the selection of fixed parameters of the software model for the studied method and the traditional approach, has confirmed the effectiveness and showed a gain in processing speed of more than 5 times. The effectiveness of the developed method can be enhanced by training and choosing such parameters as the size of the description, the compression ratio of the descriptor set, the choice of the informative subset of the description, and the choice of the clustering method. The developed method can be applied to the multiclass situation when instead of identifying “object - background” in each window, classification into several classes is carried out.

The novelty of the research consists of the development and experimental development of the method for searching for objects in the image of the visual scene using clustering implementation for query description data, which contributes to increasing the search performance and provides sufficient efficiency. The practical significance of the work is increasing the depth of analysis of visual data and the speed of classification, confirming the effectiveness of the proposed methods using examples of images, creating applied software tools for studying and implementing classification methods in the latest computer vision systems. Further stages of research can be the construction of hierarchical feature systems according to the features of structural description, as well as considering, when calculating the relevance, the weight characteristics of clusters, reflecting the number of their elements.

Acknowledgment

The authors acknowledge the support the Department of Informatics, Kharkiv National University of Radio Electronics, and Department of Informatics and Computer Engineering, Simon Kuznets Kharkiv National University of Economics, Ukraine, in numerous help and support to complete this article. The authors are grateful to Anton Sanzharovskyi for his participation in the implementation of software modeling.

Author Contributions

V.G. and I.T. conceived of the presented idea and methodology. All authors planned and carried out the simulations. V.G., I.T. and N.V. contributed to the interpretation of the results. V.G. supervised the project. All authors discussed the results and contributed to the final manuscript.

References

[1] DARADKEH, Y. I., 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*. 2021, vol. 9, iss. 1, pp. 13417–13428. ISSN 2169-3536. DOI: 10.1109/ACCESS.2021.3051625.

[2] GOROKHOVATSKY, V. A., A. V. GOROKHOVATSKY and A. Y. BERESTOVSKY. Intellectual Data Processing and Self-Organization of Structural Features at Recognition of Visual Objects. *Telecommunications and Radio Engineering*. 2016, vol. 75, iss. 2, pp. 155–168. ISSN 1943-6009. DOI: 10.1615/TelecomRadEng.v75.i2.50.

[3] GOROKHOVATSKYI, O., V. GOROKHOVATSKYI and O. PEREDRII. Analysis of Application of Cluster Descriptions in Space of Characteristic Image Features. *Data*. 2018, vol. 3, iss. 4, pp. 1–10. ISSN 2306-5729. DOI: 10.3390/data3040052.

[4] TVOROSHENKO, I. S. and V. O. GOROKHOVATSKY. Modification of the branch and bound method to determine the extremes of membership functions in fuzzy intelligent systems. *Telecommunications and Radio Engineering*. 2019, vol. 78, iss. 20, pp. 1857–1868. ISSN 1943-6009. DOI: 10.1615/TelecomRadEng.v78.i20.80.

[5] TUAN, D. L. T., T. SURINWARANGKOON, K. MEETHONGJAN and V. T. HOANG. Ensemble Feature Selection Approach Based on Feature Ranking for Rice Seed Images Classification. *Advances in Electrical and Electronic Engineering*. 2020, vol. 18, iss. 3, pp. 198–206. ISSN 1804-3119. DOI: 10.15598/aeee.v18i3.3726.

[6] DARADKEH, Y. I., V. GOROKHOVATSKYI, I. TVOROSHENKO, S. GADETSKA and M. AL-DHAIFALLAH. Methods of Classification of Images on the Basis of the Values of Statistical Distributions for the Composition of Structural Description Components. *IEEE Access*. 2021, vol. 9, iss. 1, pp. 92964–92973. ISSN 2169-3536. DOI: 10.1109/ACCESS.2021.3093457.

[7] GHAHREMANI, M., Y. LIU and B. TIDDEMAN. FFD: Fast Feature Detector. *IEEE Transactions on Image Processing*. 2021, vol. 30, iss. 1, pp. 1153–1168. ISSN 1057-7149. DOI: 10.1109/TIP.2020.3042057.

[8] TVOROSHENKO, I. S. and V. O. GOROKHOVATSKY. Effective tuning of membership function parameters in fuzzy systems based on multivalued interval logic. *Telecommunications and Radio Engineering*. 2020, vol. 79, iss. 2, pp. 149–163. ISSN 1943-6009. DOI: 10.1615/TelecomRadEng.v79.i2.70.

[9] PEHNELT, T. and P. LAFATA. Optimizing of Passive Optical Network Deployment Using Algorithm with Metrics. *Advances in Electrical and Electronic Engineering*. 2017, vol. 15, iss. 5, pp. 866–876. ISSN 1804-3119. DOI: 10.15598/aeee.v15i5.2285.

[10] DARADKEH, Y. I. and I. TVOROSHENKO. Technologies for Making Reliable Decisions on a Variety of Effective Factors using Fuzzy Logic. *International Journal of Advanced Computer Science and Applications*. 2020, vol. 11, iss. 5, pp. 43–50. ISSN 2158-107X. DOI: 10.14569/IJACSA.2020.0110507.

[11] GOROKHOVATSKY, A. V., V. A. GOROKHOVATSKY, A. N. VLASENKO and N. V. VLASENKO. Quality Criteria for Multidimensional Object Recognition Based Upon Distance Matrices. *Telecommunications and Radio Engineering*. 2014, vol. 73, iss. 18, pp. 1661–1670. ISSN 1943-6009. DOI: 10.1615/TelecomRadEng.v73.i18.50.

[12] KARAMI, E., S. PRASAD and M. SHEHATA. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. In: *Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference*. St. johns: arXiv, 2015, pp. 1–5. ISBN 978-1-5108-3881-9. DOI: 10.48550/arXiv.1710.02726.

[13] SHAPIRO, L. and G. STOCKMAN. *Computer vision*. 1st ed. Upper Saddle River: Prentice Hall PTR, 2001. ISBN 978-0-1303-0796-5.

[14] GADETSKA, S. V. and V. O. GOROKHOVATSKY. Statistical Measures for Computation of the Image Relevance of Visual Objects in the Structural Image Classification Methods. *Telecommunications and Radio Engineering*. 2018, vol. 77, iss. 12, pp. 1041–1053. ISSN 1943-6009. DOI: 10.1615/TelecomRadEng.v77.i12.30.

[15] DARADKEH, Y. I., V. GOROKHOVATSKYI, I. TVOROSHENKO and M. ZEGHID. Cluster representation of the structural description of images for effective classification. *Computers, Materials & Continua*. 2022, vol. 73, iss. 3, pp. 6069–6084, ISSN 1546-2226. DOI: 10.32604/cmc.2022.030254.

[16] DARADKEH, Y. I. and I. TVOROSHENKO. Application of an Improved Formal Model of the Hybrid Development of Ontologies in Complex Information Systems. *Applied Sciences*. 2020, vol. 10, iss. 19, pp. 1–17. ISSN 2076-3417. DOI: 10.3390/app10196777.

[17] SERT, E. and I. T. OKUMUS. Segmentation of Mushroom and Cap Width Measurement Using Modified K-Means Clustering Algorithm. *Advances in Electrical and Electronic Engineering*. 2014, vol. 12, iss. 4, pp. 354–360. ISSN 1804-3119. DOI: 10.15598/aeee.v12i4.1200.

[18] KOBYLIN, O., V. GOROKHOVATSKYI, I. TVOROSHENKO and O. PEREDRII. The Application of Non-Parametric Statistics Methods in Image Classifiers Based on Structural Description Components. *Telecommunications and Radio Engineering*. 2020, vol. 79, iss. 10, pp. 855–863. ISSN 1943-6009. DOI: 10.1615/TelecomRadEng.v79.i10.30.

[19] SAFI, M. E. and E. I. ABBAS. Robust Face Recognition Algorithm with a Minimum Datasets. *Diyala Journal of Engineering Sciences*. 2021, vol. 14, iss. 2, pp. 120–128. ISSN 1999-8716. DOI: 10.24237/djes.2021.14211.

[20] SZELISKI, R. *Computer Vision: Algorithms and Applications*. 1st ed. London: Springer-Verlag, 2011. ISBN 978-1-84882-934-3.

[21] DEWI, C., R.-C. CHEN, Y.-T. LIU, X. JIANG and K. D. HARTOMO. Yolo V4 for Advanced Traffic Sign Recognition With Synthetic Training Data Generated by Various GAN. *IEEE Access*. 2021, vol. 9, iss. 1, pp. 97228–97242. ISSN 2169-3536. DOI: 10.1109/ACCESS.2021.3094201.

[22] HSU, W.-Y. and W.-Y. LIN. Adaptive Fusion of Multi-Scale YOLO for Pedestrian Detection. *IEEE Access*. 2021, vol. 9, iss. 1, pp. 110063–110073. ISSN 2169-3536. DOI: 10.1109/ACCESS.2021.3102600.

[23] LI, H., L. DENG, C. YANG, J. LIU and Z. GU. Enhanced YOLO v3 Tiny Network for Real-Time Ship Detection From Visual Image. *IEEE Access*. 2021, vol. 9, iss. 1, pp. 16692–16706. ISSN 2169-3536. DOI: 10.1109/ACCESS.2021.3053956.

[24] DAI, J., Y. LI, K. HE and J. SUN. R-FCN: Object Detection via Region-based Fully Convolutional Networks. *Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS16)*. Red Hook: Curran Associates Inc., 2016, pp. 1–11. ISBN 978-1-5108-3881-9.

[25] HUANG, J., V. RATHOD, C. SUN, M. ZHU, A. KORATTIKARA, A. FATHI, I. FISCHER, Z. WOJNA, Y. SONG, S. GUADARRAMA and K. MURPHY. Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE COMPUTER SOCIETY, 2017, pp. 3296–3297. ISSN 1063-6919. DOI: 10.1109/CVPR.2017.351.

[26] LI, Z., C. PENG, G. YU, X. ZHANG, Y. DENG and J. SUN. Light-Head R-CNN: In Defense of Two-Stage Object Detector. *arXiv*. 2017, pp. 1–9. ISSN 2331-8422. DOI: 10.48550/arXiv.1711.07264.

[27] FILATOV, V. and A. KOVALENKO. *Advances in Spatio-Temporal Segmentation of Visual Data: Studies in Computational Intelligence*. 1st ed. Cham: Springer, 2019. ISBN 978-3-030-35479-4.

[28] KUBICEK, J., J. TIMKOVIC, M. PENHAKER, D. OCZKA, V. KOVAROVA, A. KRESTANOVA, M. AUGUSTYNEK and M. CERNY. Detection and Segmentation of Retinal Lesions in Retcam 3 Images Based on Active Contours Driven by Statistical Local Features. *Advances in Electrical and Electronic Engineering*. 2019, vol. 17, iss. 2, pp. 194–201. ISSN 1804-3119. DOI: 10.15598/aeee.v17i2.3045.

[29] KUMAR, V., A. NAMBOODIRI and C. V. JAWAHAR. Semi-supervised annotation of faces in image collection. *Signal, Image and Video Processing*. 2018, vol. 12, iss. 1, pp. 141–149. ISSN 1863-1703. DOI: 10.1007/s11760-017-1140-5.

[30] ZAMULA, A. and S. KAVUN. Complex systems modeling with intelligent control elements. *International Journal of Modeling,*

Simulation, and Scientific Computing. 2017, vol. 8, iss. 1, pp. 1–19. ISSN 1793-9623. DOI: 10.1142/S179396231750009X.

[31] KOHONEN, T. *Self-Organizing Maps*. Heidelberg: Springer-Verlag, 2001. ISBN 978-3-540-67921-9.

[32] K-Means Clustering Implementation in Python. In: *Kaggle* [online]. 2023. Available at: https://www.kaggle.com/code/andyxie/k-means-clustering-implementation-in-python/notebook

[33] XIONG, H. and Z. LI. *Data Clustering: Algorithms and Application*. 1st ed. Boca Raton: CRC Press, 2014. ISBN 978-1-4665-5822-9.

[34] AHMAD, M. A., V. GOROKHOVATSKYI, I. TVOROSHENKO, N. VLASENKO and S. K. MUSTAFA. The Research of Image Classification Methods Based on the Introducing Cluster Representation Parameters for the Structural Description. *SSRG International Journal of Engineering Trends and Technology*. 2021, vol. 69, iss. 10, pp. 186–192. ISSN 2231-5381. DOI: 10.14445/22315381/IJETT-V69I10P223.

[35] GOROKHOVATSKYI, V. and I. TVOROSHENKO. Image Classification Based on the Kohonen Network and the Data Space Modification. In: *Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020)*. Zaporizhzhia: CEUR Workshop Proceedings, 2020, pp. 1013–1026. ISSN 1613-0073. DOI: 10.32782/cmis/2608-76.

[36] MIHALIK, J. and I. GLADISOVA. Color Content Descriptors of Images by Vector Quantization. *Advances in Electrical and Electronic Engineering*. 2020, vol. 18, iss. 4, pp. 264–273. ISSN 1804-3119. DOI: 10.15598/aeee.v18i4.3799.

[37] ORB feature detector and binary descriptor. In: *Scikit-image* [online]. 2023. Available at: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_orb.html.

[38] OpenCV (Open Source Computer Vision Library). In: *OpenCV* [online]. 2023. Available at: https://docs.opencv.org/4.x/index.html.

About Authors

Volodymyr GOROKHOVATSKYI born in 1956. Doctor of Engineering, professor, graduated from Kharkiv National University of Radio Electronics with a degree in “Applied Mathematics” in 1978. In 1984 he defended his Candidate’s dissertation on the topic “Development and research of normalization algorithms for image recognition”. In 2010 he defended his Doctoral dissertation on the topic “Structural-hierarchical methods of analysis and image recognition under the conditions of the influence of spatial distortions”. Research interests: visual pattern recognition, artificial intelligence, multidimensional data analysis. Now he is a Professor in the Department of Informatics, Kharkiv National University of Radio Electronics.

Iryna TVOROSHENKO (corresponding author) born in 1980. Ph.D. in Technical Sciences, associate professor Department of Informatics at the Kharkiv National University of Radio Electronics. Graduated from Kharkiv National University of Radio Electronics with a degree in Intelligent Integrated Systems in 2002. She defended his candidate’s dissertation on the topic “Methods and Models for the Operational Estimation of the States of Complex Objects Using Fuzzy Logic” in 2010. Research interests: image and pattern recognition in computer vision systems, structural methods of image classification and recognition, fuzzy methods in artificial intelligence appliances. Currently, she is the Deputy Head Department of Informatics of the Kharkiv National University of Radio Electronics.

Oleg KOBYLIN born in 1973. Ph.D. in Technical Sciences, associate professor, graduated from Kharkiv State Technical University of Radio Electronics, the School of Computer Engineering in 1995. He defended his Candidate’s dissertation on the topic “Methods and Models of Adaptive Normalization Systems of Image Processing” in 2007. Research interests include evolving hybrid systems of computational intelligence: image segmentation, spectral image analysis, identification, forecasting, clustering, and diagnostics. Currently, he is the Head Department of Informatics of the Kharkiv National University of Radio Electronics.

Nataliia VLASENKO born in 1988. Ph.D. in Technical Sciences, in 2010 graduated from Kharkiv National University of Radio Electronics with a degree in “Information technology design”. In 2014 she defended her Candidate’s dissertation on the topic “Feature description models and their transformation in image recognition”. Research interests: visual pattern recognition, artificial intelligence, decision making in information systems. She is currently an associate professor with the Department of Informatics and Computer Engineering, Simon Kuznets Kharkiv National University of Economics.