International Science Group

ISG-KONF.COM

INTEGRATION OF SCIENTIFIC BASES INTO PRACTICE

IV

SCIENTIFIC AND PRACTICAL CONFERENCE

12-16 October

Stockholm, Sweden

DOI 10.46299/ISG.2020.IV

ISBN 978-1-64945-864-3

INTEGRATION OF SCIENTIFIC BASES INTO PRACTICE

Abstracts of IV International Scientific and Practical Conference

Stockholm, Sweden
October 12-16, 2020

Library of Congress Cataloging-in-Publication Data

UDC 01.1
The IV th International scientific and practical conference «Integration of scientific bases into practice» (October 12-16, 2020). Stockholm, Sweden 2020. 523 p.

ISBN - 978-1-64945-864-3
DOI - 10.46299/ISG.2020.IV

EDITORIAL BOARD

MECHANISMS OF IMAGE CLASSIFICATION BASED ON DESCRIPTORS OF LOCAL FEATURES

Tvoroshenko Iryna,
Ph.D., Associate Professor
Kharkiv National University of Radio Electronics

Tkachenko Dmytro
Master in informatics
Kharkiv National University of Radio Electronics

Computer vision is a branch of information technology that explores the ability of machines to extract information from images obtained from various sensors and visualize human perception [1].

Among the scientific areas considered by computer vision, we can conditionally distinguish the following [2]: pattern recognition, scene reconstruction, video analysis.

Some problems of image recognition in computer vision can be considered solved (object detection [3], instance recognition [4]), and the problem of image classification (category recognition) remains unsolved [5-9] for the following reasons:

– Visual variability of objects belonging to a certain category;

– Permissible structural differences between items belonging to one category;

– Impossibility of extensive expansion of the training sample, because the combinatorial explosion leads to retraining;

– The relationship between object categorization and external context;

– Misunderstanding of how these mechanisms work in the human mind.

The difficulty may be that there may be situations where none of the classes taught by the recognition system is represented in the image [10]. In this case, the software application must provide information about the absence of existing classes. This fact leads to the use of thresholds for the decision-making mechanism [11, 12].

Most of the existing image classification mechanisms are based on the visual features of images [13-16]:

– Compiling a training sample of images divided by classes;

– Identification of visual features in the training sample;

– Conversion of the received data for the further work;

– Use of processed data to analyze the next input image and make decisions about whether it belongs to a certain class.

Problems of the existing [17-20] mechanism of training and classification are defined:

– Libraries of “etalon” images representing individual classes are formed manually, which reduces flexibility and imposes the need to keep the library up to date;

– The choice of the classifier can largely depend on the characteristics of the processed images, and, in the case of inappropriate images, the reliability of the classification drops sharply;

– With a significant number of classes, the accuracy of classification is reduced, as it becomes difficult to make an unambiguous model of each individual class;

– Image classes are not related to the semantic load, which does not allow further interpretation of the results.

There are many methods of finding points of interest in the image and methods of presenting information about their environment in the form of descriptors [21, 22].

Using descriptor values (multi-dimensional vectors) to find the closest elements between images differs for different applications. In classification problems, the use of the stage of quantization of values by visual words is widespread, which allows us to dramatically reduce the computational complexity of the sample of the nearest elements and simplify the classification model [23]. Descriptors of local features can have different invariance to distortions [24].

Detectors show themselves differently depending on the filling of the scene and the quantitative parameters of the image, while the descriptors have excellent parameters of invariance concerning image distortion [24].

Classification systems are taught on a specially prepared sample of images, divided into classes. To date, many such libraries allow you to successfully evaluate the performance of certain classifiers on the same sets of images. As mentioned earlier, such libraries are compiled manually and the classes represented in them are unrelated. This fact determines the complexity of their maintenance and expansion, as well as the complexity of processing images containing instances of different classes.

There are different variants of the classification problem and the accompanying requirements for the training sample. There are two types of such tasks: categorization of the image as a whole and categorization through localization.

The problem of categorization of the image as a whole is formulated as finding a class in which the function of the probability of belonging of the image to the class reaches the maximum value [25]:

$$\overset{*}{Class} = \arg\max_{Class} f(Image, Class), \tag{1}$$

where $\overset{*}{Class}$ – the most probable class for this image; $\arg\max_{Class}$ – the value of the argument at which the expression reaches a maximum; f – function of probability of belonging of the image of a class; $Image$ is the analyzed image; $Class$ class from the available set.

For training of the corresponding classifiers only distribution of the educational sample on various classes is required.

The problem of categorization by localization [26] involves not only to assign the image to a particular class but also to localize an instance of this class in the image:

$$\overset{*}{Class} = \arg\max_{Class} f(Image \cap area, Class), \tag{2}$$

$$\overset{*}{Class} = \arg\max_{Class} f(\overline{Image \cap area}, \overline{Class}), \quad (3)$$

where $area$ – is the area in the image $Image$.

To be able to work with the following conditions of the task, additional requirements are introduced to the training sample: in addition to the class label itself, the area of presence of this object is selected for each image in the images.

In the vast majority of cases, it is represented by a rectangle, as the use of other geometric shapes is associated with further complexity of processing. For categorization through localization, a separate training of different classifiers is carried out: for selected objects and the surrounding background. In the future, the analysis of input images uses a function that takes into account the values of both classifiers and reaches a maximum in the area of the expected location of the instance of a particular class. In the simplest case, this function is the sum of the values of the classifiers. Tasks are formulated as "finding an instance of the class or their absence".

The paper analyzes the approaches to the use of local visual features for image classification; they can be divided into some groups [24]:

– Bag of Words (BoW);
– Naive Bayesian Nearest Neighbor (NNBN);
– Part-Based;
– Based on segmentation.

In the BoW method, all visual features (which are multidimensional vectors) from the entire training sample are combined into a common array, which is then divided into a given number of visual words.

The resulting sets of visual words are used to decide whether the analyzed image belongs to a particular class. Generalizing or discriminatory methods are used to make a decision. An example of a generalizing method is the Naïve Bayesian classifier, and a discriminating method is a classifier based on the reference vector method. When using a discriminating classifier, the next step is to calculate the representation of the image in the form of a histogram of the distribution of visual words.

Among the advantages of this model in terms of the proposed method is the stability of its results with minor changes in the training sample. The essence of this method fits well into the approach with the automatically obtained training sample. Among the disadvantages is the problem of situations where the already mentioned image shows more than one copy of the object from the training sample or this object is represented in part. High computational complexity at the learning stage, especially the process of clustering the original set of visual features, and the presence of the quantization stage, which reduces the discriminatory possibility of descriptors.

The NBNN method is based on the idea that the process of quantizing the original visual features used in the BoW method reduces their discriminatory potential. An assumption is made: the probability of finding a single descriptor in the class does not depend on the probability of finding others.

In this approach, all the visual features of the educational sample for each class are combined into a common array. When analyzing the input image for each of its

visual features, the closest feature among each of the classes is sought. The closest class is the one for which the sum of the distances will be minimal.

The advantages of this model are the lack of a learning stage and the ability to work with objects presented in part.

Disadvantages include the instability of the results of work on changes in the educational sample: even minor changes in the sample for one class can worsen the overall performance of the system. Increased requirements for the uniformity of the educational sample between different classes: classes, which are characterized by a larger number of extracted visual features, will have greater classifier values due to statistical laws.

The Part-Based approach takes into account the mutual arrangement of visual features and approaches based on segmentation. It is specialized for certain tasks and therefore cannot be used in a general classification problem. To work with a training sample obtained automatically, the classifier must have the stability of the result of the work with minor changes in the training sample.

Based on the variability of the educational sample and its heterogeneity, as well as based on the practical results of experiments, it was decided to use the representation of images in the form of visual words for classification.

In the framework of this work, two classifiers using image representations in the form of visual words were tested: a naive Bayesian classifier and a classifier based on the method of reference vectors.

Based on the work of the above-described classifiers on real data obtained automatically, it was decided to use a classifier based on the method of reference vectors, which showed the best practical results. The experiments performed automatically, without an etalon set of images, showed the ability to filter only synthetic images and images that are the result of any error.

To filter such emissions use:

– Representation of images in the BoW model;

– The arithmetic mean value of histograms is calculated;

– Measure of scattering due to the standard deviation;

– The ratio of the distance of each image to the arithmetic mean to the extent of scatter.

At the same time, visual words are extracted for different classes independently; this makes it possible to achieve a more uniform distribution of them within the class.

A significant problem and prospect for further research is the lack of links between the input data and between the obtained classifiers. There is no mechanism for detecting cases of potentially erroneous classification or the fact that the image does not represent one of the existing classes.

References:

1. Szeliski R. (2010) *Computer Vision: Algorithms and Applications*, London, Great Britain: Springer-Verlag, 957 p.

2. Gorokhovatskyi V., Gadetska S., and Stiahlyk N. (2020) Image structural classification technologies based on statistical analysis of descriptions in the form of

bit descriptor set, *In CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2020)*, 2608, pp. 1027-1039.

3. Sonka M., Hlavac V., and Boyle R. (2014) *Image Processing, Analysis, and Machine Vision*, Atlanta, USA: Thomson-Engineering, 920 p.

4. Gorokhovatskyi V.A. (2018) Image classification methods in the space of descriptions in the form of a set of the key point descriptors, *Telecommunications and Radio Engineering*, 77(9), pp. 787-797. DOI: 10.1615/TelecomRadEng.v77.i9.40

5. Gorokhovatskyi V.O., Gadetska S.V., and Stiahlyk N.I. (2019) Study of statistical properties of the block representation model for a set of key image descriptors, *Radio Electronics Computer Science Control*, 2, pp. 100-107.

6. Duda R.O., Hart P.E., and Stork D.G. (2000) *Pattern classification*, Hoboken, USA: John Wiley & Sons, 738 p.

7. Nong Ye. (2013) *Data Mining: Theories, Algorithms, and Examples*, Florida, USA: CRC Press, 349 p.

8. Tvoroshenko Irina, Ahmad M. Ayaz, Mustafa Syed Khalid, Lyashenko Vyacheslav, and Alharbi Adel R. (2020) Modification of Models Intensive Development Ontologies by Fuzzy Logic, *International Journal of Emerging Trends in Engineering Research*, 8(3), pp. 939-944. DOI: 10.30534/ijeter/2020/50832020

9. Yousef Ibrahim Daradkeh, and Iryna Tvoroshenko (2020) Application of an Improved Formal Model of the Hybrid Development of Ontologies in Complex Information Systems, *Applied Sciences*, 10(19). p. 6777. DOI: 10.3390/app10196777

10. Peters J.F. (2017) *Foundations of computer vision: Computational Geometry, Visual Image Structures and Object Shape Detection*, Cham, Switzerland: Springer International Publisher, 417 p.

11. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Intelligent classification of biophysical system states using fuzzy interval logic, *Telecommunications and Radio Engineering*, 78(14), pp. 1303-1315. DOI: 10.1615/TelecomRadEng.v78.i14.80.

12. Ahmad M. Ayaz, Tvoroshenko Irina, Baker Jalal Hasan, and Lyashenko Vyacheslav (2019) Computational Complexity of the Accessory Function Setting Mechanism in Fuzzy Intellectual Systems, *International Journal of Advanced Trends in Computer Science and Engineering*, 8(5), pp. 2370-2377. DOI: 10.30534/ijatcse/2019/77852019

13. Gorokhovatskyi V., and Tvoroshenko I. (2020) Image Classification Based on the Kohonen Network and the Data Space Modification, *In CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2020)*, 2608, pp. 1013-1026.

14. Gorokhovatskyi V.O., Tvoroshenko I.S., and Peredrii O.O. (2020) Image classification method modification based on model of logic processing of bit description weights vector, *Telecommunications and Radio Engineering*, 79(1), pp. 59-69.

15. Sharma G., and Schiele B. (2015) Scalable Nonlinear Embeddings for Semantic Category-based Image Retrieval, *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 7-13.

16. Gorokhovatsky V.A. (2016) Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions, *Telecommunications and Radio Engineering*, 75(14), pp. 1271-1283.

17. Matarneh Rami, Tvoroshenko Irina, and Lyashenko Vyacheslav (2019) Improving Fuzzy Network Models For the Analysis of Dynamic Interacting Processes in the State Space, *International Journal of Recent Technology and Engineering*, 8(4), pp. 1687-1693. DOI: 10.35940/ijrte.C5582.118419

18. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Modification of the branch and bound method to determine the extremes of membership functions in fuzzy intelligent systems, *Telecommunications and Radio Engineering*, 78(20), pp. 1857-1868. DOI:10.1615/TelecomRadEng.v78.i20.80.

19. M. Ayaz Ahmad, Irina Tvoroshenko, Jalal Hasan Baker, and Vyacheslav Lyashenko (2019) Modeling the Structure of Intellectual Means of Decision-Making Using a System-Oriented NFO Approach, *International Journal of Emerging Trends in Engineering Research*, 7(11), pp. 460-465. DOI: 10.30534/ijeter/2019/107112019

20. Asaad Ma. Babker, Abd Elgadir A. Altoum, Irina Tvoroshenko, and Vyacheslav Lyashenko (2019) Information Technologies of the Processing of the Spaces of the States of a Complex Biophysical Object in the Intellectual Medical System HEALTH, *International Journal of Advanced Trends in Computer Science and Engineering*, 8(6), pp. 3221-3227. DOI: 10.30534/ijatcse/2019/89862019

21. Leutenegger S., Chli M., and Siegwart R. (2011) BRISK: Binary Robust Invariant Scalable Keypoints, *Proceedings of 2011 IEEE International Conference on Computer Vision (ICCV)*, pp. 2548-2555.

22. Rublee E., Rabaud V., Konolige K., and Bradski G. (2011) ORB: an efficient alternative to SIFT or SURF, *Proceedings of 2011 IEEE International Conference on Computer Vision (ICCV)*, pp. 2564-2571.

23. Tvoroshenko I.S., and Gorokhovatsky V.O. (2020) Effective tuning of membership function parameters in fuzzy systems based on multi-valued interval logic, *Telecommunications and Radio Engineering*, 79(2), pp. 149-163. DOI: 10.1615/TelecomRadEng.v79.i2.70.

24. Flah P. (2015) *Machine learning. The science and art of building algorithms that extract knowledge from data*, Moscow, Russia: DMK Press, 400 p., (in Russian).

25. Lyashenko V., Mustafa S.K., Tvoroshenko I., and Ahmad M.A. (2020) Methods of Using Fuzzy Interval Logic During Processing of Space States of Complex Biophysical Objects, *International Journal of Emerging Trends in Engineering Research*, 8(2), pp. 372-377. DOI: 10.30534/ijeter/2020/22822020

26. Daradkeh Y.I., and Tvoroshenko I. (2020) Technologies for Making Reliable Decisions on a Variety of Effective Factors using Fuzzy Logic, *International Journal of Advanced Computer Science and Applications*, 11(5), pp. 43-50. DOI: 10.14569/IJACSA.2020.0110507