UDC 004.93

Lukac M.[1], Abdiyeva K.[2], Kameyama M.[3]

[1]Dr., Assistant Professor of Department Computer Science, Nazarbayev University, Astana, Kazakhstan
[2]Post-graduate Student, ROSE Laboratory, Nanyang Technological University, Singapore
[3]Dr., Professor of the Graduate School of Information Sciences, Tohoku University, Sendai, Japan

EVALUATION OF COMPONENT ALGORITHMS IN AN ALGORITHM SELECTION APPROACH FOR SEMANTIC SEGMENTATION BASED ON HIGH-LEVEL INFORMATION FEEDBACK

In this paper we discuss certain theoretical properties of the algorithm selection approach to the problem of semantic segmentation in computer vision. High quality algorithm selection is possible only if each algorithm's suitability is well known because only then the algorithm selection result can improve the best possible result given by a single algorithm. We show that an algorithm's evaluation score depends on final task; i.e. to properly evaluate an algorithm and to determine its suitability, only well formulated tasks must be used. When algorithm suitability is well known, the algorithm can be efficiently used for a task by applying it in the most favorable environmental conditions determined during the evaluation. The task dependent evaluation is demonstrated on segmentation and object recognition. Additionally, we also discuss the importance of high level symbolic knowledge in the selection process. The importance of this symbolic hypothesis is demonstrated on a set of learning experiments with a Bayesian Network, a SVM and with statistics obtained during algorithm selector training. We show that task dependent evaluation is required to allow efficient algorithm selection. We show that using symbolic preferences of algorithms, the accuracy of algorithm selection can be improved by 10 to 15% and the symbolic segmentation quality can be improved by up to 5% when compared with the best available algorithm.

Keywords: algorithm selection, algorithm suitability, computer vision.

NOMENCLATURE

ALE is Automated Labeling Environment;
BN is Bayesian Network;
CPMC is Constrained Parametric Min-Cuts for Automatic Object Segmentation;
FFT is Fast Fourier Transform;
HOG is Histogram of Oriented Gradients;
IA is Iterative Analysis;
MSER is Maximally Stable Extremal Regions;
SIFT is Scale Invariant Feature Transform;
SDS is Simultaneous Detection and Segmentation;
SVM is Support Vector Machine.

INTRODUCTION

The algorithm selection problem has been introduced by Rice [1] and since has been used in various applications. Recently it has been applied to machine vision and image processing [2, 3]. While in general the algorithm selection process works well [4–7], for more complex problem spaces, problems that are related to feature selection, evaluation and algorithm suitability have been recorded and reported [3, 8].

Algorithm selection is in general seen as a secondary solution to a problem because to select best algorithm from a set of available algorithms several preconditions must be satisfied: knowledge about the problem, knowledge about the algorithms, algorithm suitability and distinctive features for each algorithm must be known. Consequently, algorithm selection is neither easy to apply nor the least expensive solution. However, for complex problems that have large feature spaces including problems that deal with real-world situations and environment, algorithm selection is a viable alternative. The concept behind algorithm selection is in the algorithm separation: an algorithm that would deal with the problem successfully for all combinations of environmental conditions will be too complex but a set of more specific algorithms for subsets of conditions will provide better and cheaper solutions when applied on a case by case basis.

To obtain result improvement from a case by case selected set of algorithms, high quality selection mechanism with a minimal precision of selection is required: for a set of inputs, the selected algorithms must be such that the cumulative result is better than the best of the available algorithms. This implies that the algorithm selection mechanism must be able to select the best algorithm as often as possible.

A reliable algorithm selection implies that the set of available algorithms have been evaluated in a very strict setting and in a task dependent manner. As will be shown, task specific evaluation provides data that can be used for algorithm selection because only such evaluation results can be used to predict reliably algorithm results on new untested input data. This means that evaluation of a single algorithm cannot be seen as a holistic process but rather as a precise and specific process that is not generalizable.

Finally, the algorithm selection presented in this paper is situated within the framework for high level image understanding. We show that unlike standard feature-only based algorithm selection approaches, the high level symbolic description greatly improves the accuracy of algorithm selection as well as the final result of high level understanding.

1 PROBLEM STATEMENT

In this paper we analyze several problems of the algorithm selection:
– the impact of the high level symbolic understanding of image content on the accuracy of algorithm selection;
– the impact of algorithm evaluation on the algorithm selection process;
– the impact of feature for object recognition evaluation on the algorithm selection process.

2 REVIEW OF LITERATURE

Algorithm selection was introduced by Rice [1] in the context of selection of scheduling algorithm in computer operating system. Since then it has been applied to various problems and fields of research in different ways and granularity. In image processing and computer vision the algorithm selection has been used to determine the best algorithm for segmentation of artificially generated images of noisy geometrical shapes [4]. In [7] algorithm selection was used for determining the best algorithm for the segmentation of biological cell images and [5] used algorithm selection in a performance predicting framework.

For segmentation of more complex natural images [2] proposed an algorithm selection approach using machine learning and composition: final segmentation was created from partial segmentation from best algorithms for different regions of the image. The method showed that despite results with high accuracy of selection the final result was only as good as the best available algorithm. Finally, a more specific approach was used to select parameters in single algorithm for segmentation in [9].

In [10] uses depth information to estimate whole image properties such as occlusions, background and foreground isolation and point of view estimation to determine type of objects in the image. All the modules of this approach are processed in parallel and integrated in a final single step. An airport apron analysis is performed in [11] where the authors use motion tracking and understanding inspired by cognitive vision techniques. Finally, the image understanding can also be approached from a more holistic approach such as for instance in [12] where the intent is only to estimate the nature of the image and distinguish between mostly natural or artificial content.

Currently there is a large amount of work combining segmentation and recognition and some of them are [13, 14]. In [15] uses an interleaved object recognition and segmentation in such manner that the recognition is used to seed the segmentation and obtain more precise detected objects contours. In [16] objects are detected by combining part detection and segmentation in order to obtain better shapes of objects. More general approaches such as [17] build a list of available objects and categories by learning them from data samples and reducing them to relevant information using some dictionary tool. However this approach does not scale to arbitrary size because the labels are not structured and ultimately require complete knowledge of the whole world.

3 MATERIALS AND METHODS

In [8] an alternative approach to image understanding was proposed: an algorithm selection platform with verification of the high-level symbolic interpretation of the image content was proposed. This platform is used as basis of research in this paper and is shown in Fig. 1.

The platform works in two distinct modes and integrates both the algorithm selection from features and algorithm selection from high-level feedback. Initially, the input image is processed (box 1) by algorithm selected using the algorithm selection mechanism (box 3) that uses only the image features (Loop 1). The resulting high-level description of the image (obtained from the object recognition), is verified for logical contradictions (box 4) both on the context, on the part-level, on the location and on the relative size level. If the verification does not detect any high-level symbolic contradiction the processing stops and outputs the current high level description. If however a logical contradiction is detected, a hypothesis that solves the contradiction is generated (box 5). The image region that corresponds to the contradiction and to the hypothesis is used to extract local features, to determine local context information and to estimate attributes of the possible objects located in the selected image region. These three sources of information are used in the meta level to estimate what other algorithm should be used to correct the contradiction (Loop 2). This second loop is iterated over all contradictions until all contradictions are resolved. For the rest of this paper the presented system will be referred to as Iterative Analysis (IA).

Notice that the proposed system uses a twofold processing convergence. First convergence of the approach is to obtain a non-contradictory high-level description (contradiction resolution). The second convergence is the match between a description without contradiction and a set of algorithms (algorithm matching). The proposed approach thus combines processing quality with the meta-processing algorithm matching. This approach thus enables to exploit each algorithm's strongest points on an application, image features and image content basis.
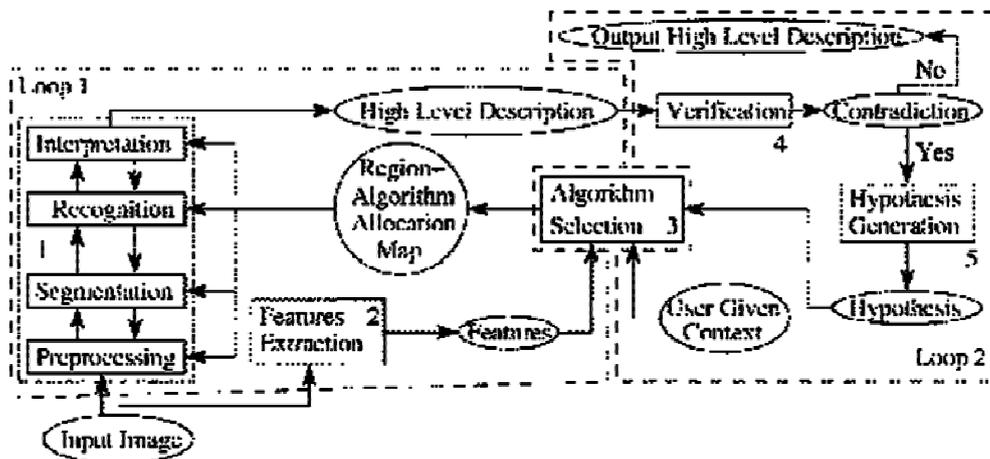

Figure 1 – Algorithm selection platform with verification of the high-level symbolic interpretation of the image content

The concept behind the processing in box 1, Figure 1 is that each algorithm used is a network of various component algorithms. Box 1 shows the general classical robotic sequential processing that uses four components processing levels: the preprocessing, segmentation, recognition and interpretation. However as in this paper the algorithms used are performing the semantic segmentation the interpretation is obtained by a single common algorithm. Also the selection is not limited to these four processing blocks but rather is intended to accommodate various algorithm networks.

As a final note some specific information about the selection process is required. In the initial loop of the IA processing, features extracted from the input images are FFT coefficients, Gabor features, wavelets, gist, color average, intensity average, edges, covariant features, SIFT, HOG, MSER and textures. All these features are transformed into a histogram and are concatenated into a single vector per image (or per region) of 5000 values.

For all loops after the initial one, the hypothesis is represented as a set of attributes. These attributes are obtained using the regprops function in Matlab. These attributes have been discretized in order to simplify the representation but to allow discrete representation of each of the available hypotheses.

In this paper the platform uses algorithms performing semantic segmentation: first segment an image and recognize regions as objects. The result of such processing is fed to the interpretation and verification according to the above platform description.

4 EXPERIMENTS

In order to assess an algorithm processing quality, it is necessary to evaluate its performance with respect to some training data set and ground truth. Each evaluation experiment was designed using real algorithm selection data. The algorithms used in our classification task are the ALE [11, 18], CPMC with recognition [14] and the SDS [19]. The three algorithms have similar performance results shown in Table 1. Here the numbers given in the original papers may vary due to different set up, initialization and training conditions of the original and this experiments.

Consequently most of the algorithms that perform the semantic segmentation task first segments an image using some well-known segmentation algorithm and then apply the object recognition (there are other algorithms that are not using this order such as [15]).

Let us assume that an algorithm is evaluated for the quality of segmentation i.e. it evaluates whole image segmentation by comparing the result of processing to a human provided ground truth. Figure 2a shows an example of input image, Fig. 2b – human generated ground truth and Fig. 2c–Fig. 2d – the result of a segmentation algorithm. Fig. 2b – Fig. 2c have also their f-value shown in the parentheses. F-value is one of the standard measures used to determine the accuracy of computer generated segmentation [20]. According to the f-measure in this case of evaluation the algorithm generating the result shown in Fig. 2c is superior (closer when pixel-to-pixel comparison is done with human segmentation in Fig. 2b) to the algorithm which result is shown in Fig. 2d.

Now let's look at the same algorithms in the task of semantic segmentation. In semantic segmentation and input

Table 1 – Results of semantic segmentation algorithms on the VOC2012 validation dataset

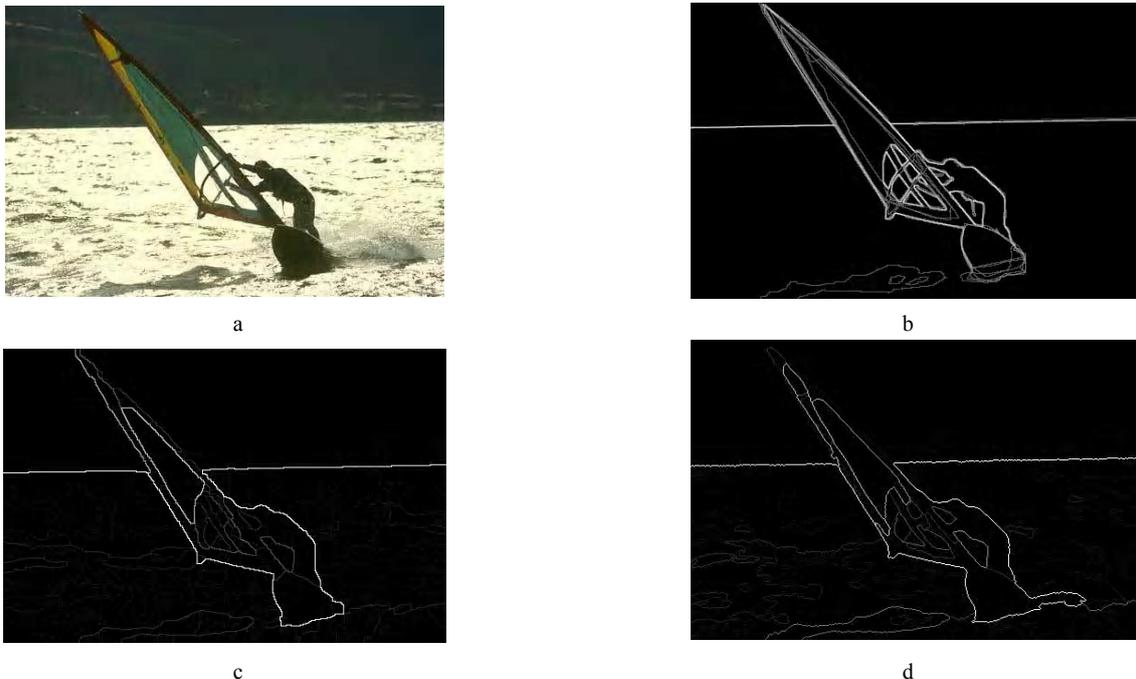| Name	Result
ALE	47.8%
CPMC	48.3%
SDS	49.9%

a

b

c

d

Figure 2 – Example of algorithmic Segmentation: a – input image, b – human ground truth (0.9), c – result of algorithm from [21] (0.92), d – result of algorithm from [22] (0.87)

image is segmented and then each region is labeled form a set of available object label set. In the task of semantic segmentation, the two best algorithms for image segmentation shown in Fig. 2c, Fig. 2d will not have the same f-values. In fact, algorithms with much lower f-value f= 0.77 (in the task of segmentation and with result shown in Fig. 3b) will have much higher resulting score because the regions obtained from the detected regions are more precise for object detection and labeling.

The reason for such change of the score is possible because in image segmentation the algorithm's result is evaluated by comparing the obtained boundaries with the ground truth generated by human. However, in the case of semantic segmentation the evaluation is made first by determining the boundaries of the target object and then the detection of the correct object is tested. This means that segmenting the whole image and comparing it to a set of human generated ground truth will result in more variation because even humans will not generally agree on how to segment a whole scene. This is because the evaluation is done with respect to a human segmentation that depends on feeling and intuition. However, when segmenting an image to determine object boundary the disparity between humans is much smaller. The semantic segmentation can be automatically judged on whether or not the correct object was correctly detected. Consequently despite the fact that some segmentation might be close enough to a human like segmentation it might not be well suited for the segmentation of a particular object.

Thus for two different tasks, the score of the final evaluation of a same algorithm might not be the same and the algorithm that had a good result in one tasks will have much lower result score for another task. But a statistical evaluation of algorithms might not be sufficient to determine advantages and disadvantages precisely enough. Figure 4 shows the standard model of robotics where multiple processes are formed into a set of consecutive algorithmic steps. The combination of algorithms can result in non-linear result that would not be observed otherwise. Thus it is necessary to evaluate the component algorithms as well

so that individual suitabilities can be determined and impact on the result of the entire computation.

Similarly to the segmentation study a change of result can be obtained in recognition. Various features have different accuracy and ability to detect and recognize an object.

Using various features for detection (using the bag of words recognition model) it can be shown that depending on the region used to extract the feature descriptors and on the features extracted, the recognition accuracy will change. For instance assume that a segmentation algorithm such as [21, 25] is used for segmentation. The results of the segmentation are boundaries that indicate main regions of the image where the features for recognition should be extracted and the recognition model should be applied. Depending on what features are extracted the accuracy will change depending on the image. In some cases there will be no detection and in some other cases the detection will be a success.

Figure 5 shows the results of calculating the bounding box after two different features (SIFT and HOG) have been used for object recognition. In this case we extract features form the whole images. The features and the descriptors extracted are used to recognize a motor-bike and then the same feature descriptors are used to generate a bounding box. The idea behind this experiment is to assess the importance of a region in recognition of a motor-bike given that segmentation occurred prior to recognition.

The bounding box method determination is shown in Fig. 6 and Fig. 7.

Following the standard bag of words object recognition method a model of the object being detected is available as a set of histograms of feature clustered centers. Input image is first used as input for feature extraction, the feature descriptors are then clustered into k centers and a new histogram is constructed with bins corresponding to the k centers. Once the histogram is obtained, it is compared to all histograms in the model database and four closest matches are saved. Finally features corresponding to four best matching bins (each from one of the model histograms) are used to determine which descriptors and consequently which key points are used to determine the bounding box (Figure 7).

a   b

Figure 3 – Example of two algorithms for segmentation with lower f-value:
a – result of segmentation of input image from Figure 2a using the algorithm from [23], b – result of segmentation of input image from Figure 2a using the algorithm from [24]

Preprocessing → Segmentation → Recognition → Understanding

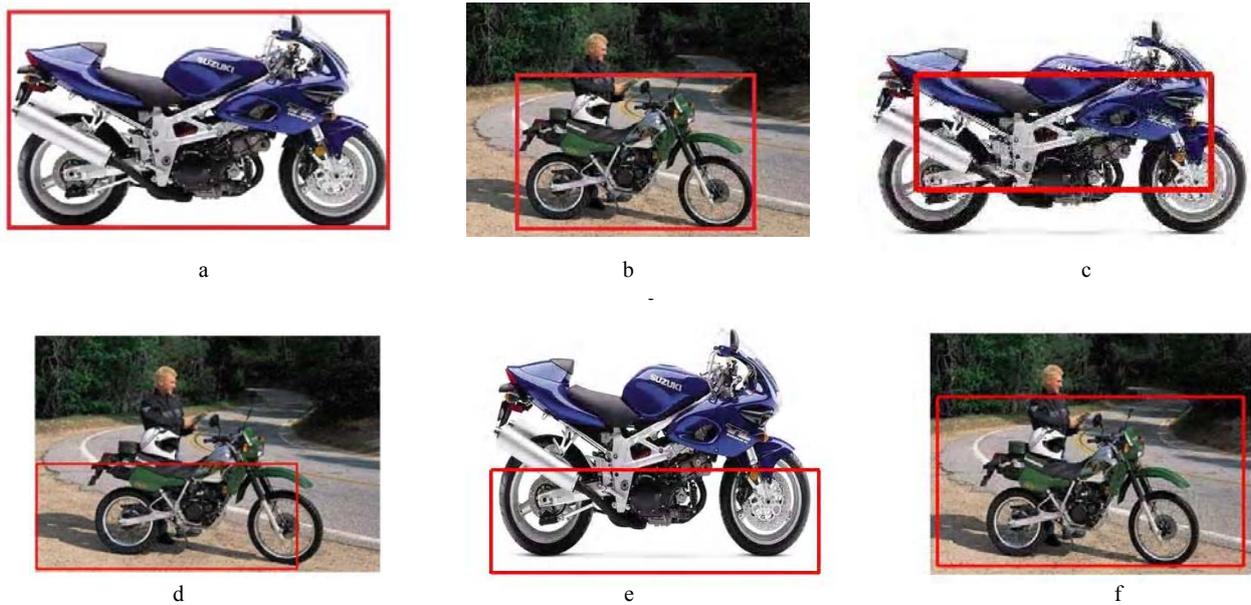Figure 4 – Typical example of processing required for semantic segmentation

Figure 5 – Comparison of Bounding boxes obtained from SIFT and HOG features: a – Input Image 1 and Bounding Box by Human, b – Input Image 2 and Bounding Box by Human, c – Bounding box from SIFT, d – Bounding box from SIFT, e – Bounding Box from HOG, f – Bounding Box from HOG
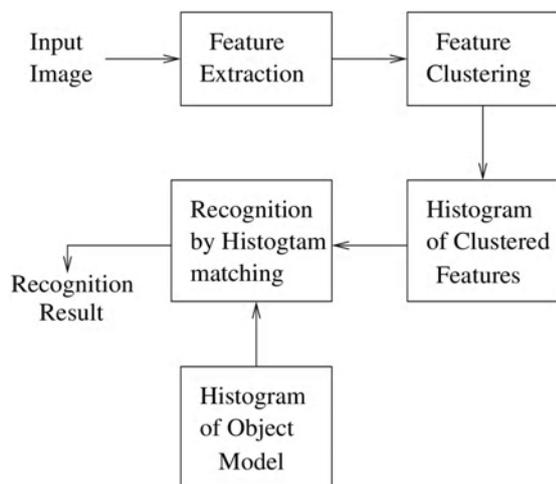

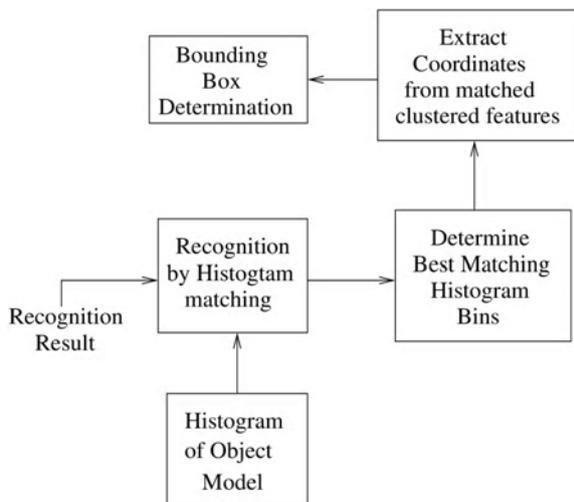Figure 6 – Schema of Bag-of-words recognition algorithm

Figure 7 – Schema of Bounding Box extraction from a successful object recognition using Bag-of-words recognition algorithm

Using the method for determining the bounding box as an evaluation of the feature descriptor relevance to the motor-bicycle model, it can be seen that having different regions used to extract the feature descriptors would have a significant impact on object detection. For instance if only the upper left region (Region 1 in Fig. 8) of the bicycle would be contained in a single region no detection would happen if the HOG features would be used. Also if the SIFT features would be used it is possible that positive detection might not occur as not enough of the significant SIFT descriptors for successful detection are contained in the Region 1 only. On the other hand if the bottom of the motorcycle would be contained in a single region (Region 3 in Fig. 8) the HOG features would not be able to detect the motorcycle.

Consequently, using various features for only recognition or for semantic segmentation can have considerably different results as both the segmentation and the recognition are sensitive and difficult operations. Their evaluation is thus highly task dependent.

In the software platform previously introduced the algorithm selection is iterated through several iterations. The stopping condition for the processing of the image is either no more improvement is possible due to having tried all available algorithms or no more improvement is possible as the new hypothesis generated is the same as the previous one.

Initially, the algorithms are selected using only the image features but after the first processing loop the hypothesis generated is used for algorithm selection. The features have been successfully used for algorithm selection in various approaches, however in general such algorithm selector is limited due to the fact that many algorithms are designed for particular symbolic and semantic context.

The semantic segmentation results in a set of symbolically labeled regions and thus analyzing the obtained regions by various algorithms it is possible to conclude that various algorithms have affinities for different objects.
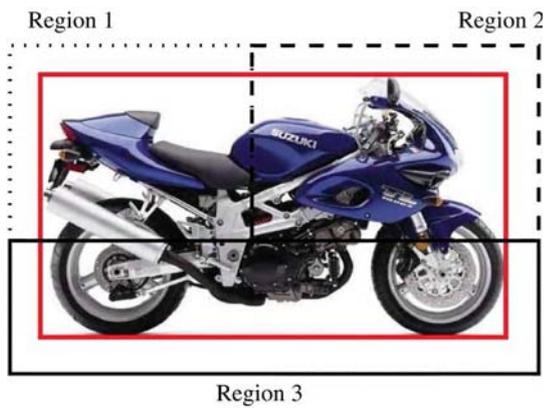
Figure 8 – Example of three different regions obtained as a result of a possible segmentation

Such affinity for particular objects can be due to the following reasons:

– The environment in which the particular object is captured has particular interaction with the object that is favorable to be detected by a particular algorithm.

– The object itself has a set of features that a particular algorithm is better suited for detection and segmentation.

Consequently we asked: what is the impact of symbolic information (content related) on the accuracy of algorithm selection?

To answer the above question we conducted a set of experiments. The experiments were carried using the VOC2012 [26] database. The dataset used is not the standard VOC2012 validation set but a reduced one in order to allow applying our platform. This means that only images where multiple objects to be segmented are present. The dataset is thus reduced and contains only ~300 images out of ~1500 images contained in the VOC2012 dataset.

5 RESULTS

The platform was initially designed to use Bayesian Network (BN) because the probabilistic inference is well suited to deal with missing variables. Consequently and because the two different modes of algorithm selection (features only and features with high level description), a single trained BN can be used. However selection of algorithms using Bayesian Network is still problematic and thus two alternative algorithm selectors were used for comparison. These two selection mechanisms are SVM and Statistics from training.

In a first experiment we compared the BN and the SVM because both of these algorithm selectors work on similar principles. Both BN and SVM are used in the initial and all further iterations of the IA platform. In the first iterations only features are used to select algorithm while in all further loops the features from the contradiction region as well as the hypothesis is used. The main difference between using the BN and SVM is that SVM requires incomplete input information imputation [28] while the BN is well suited to handle missing input information by design. This means that for the first iteration, the SVM is provided with average values of the hypothesis in order for the input vector has the desired and fixed length.

The results of comparison of the precision of the BN and of the SVM are shown in Table 2.

The problem of using the Bayesian Network is that it requires discrete input information. However most of the features extracted from input image are continuous and unbounded. Consequently it is required to cluster the input information and only then use it as input to the BN. This however has in most of the cases a dramatic influence on the performance of the probabilistic algorithm selection.

As can be seen the impact of hypothesis attributes is significant in the case of SVM, however in the case of BN it is difficult to evaluate as the overall precision is too low. The general increase of algorithm selection using the features and attributes compared to the selection using only features is up to 10% of accuracy.

The final evaluation of the high level information (feedback) in our system is the usage of statistical accuracy of each algorithm. The accuracy represents the percentile average of the f-measure of each semantic segmentation algorithm. Table 3 shows the accuracy of semantic segmentation for each of the three used algorithms: ALE [18], CPMC [14] and SDS [19]. Each columns shows average accuracy for each of the categories of objects that are to be recognized and segmented and overall average accuracy in the bottom row. The statistical information provided was obtained by evaluating the VOC2012 validation data set that contains approximately 1500 images.

The last column in Table 3 shows for each class of objects the best algorithm based on the statistical accuracy of each algorithm. This means that in the platform and during iterations all but the first one, for each hypothesis algorithm will be selected only using the best algorithm listed in the rightmost column. Using this approach we evaluated the proposed Iterative Analysis method described in this paper. The result comparison is shown in Table 4. It shows average precision for each algorithm for the test dataset.

Table 2 – Comparison of BN and SVM in selection accuracy using only features vs. features and attributes

Task	Algorithm	Features	Features+Attributes
2-class	SVM	58%	65%
2-class	BN	46%	55%
3-class	SVM	43%	46%
3-class	BN	39%	43%

Table 3 – Average accuracy of different algorithms on the training set

| Accuracy for each class (intersection/union measure) | Algorithm Names | | | Best Algorithm |
	ALE	SDS	CPMC	
Background	71.711	84.937	83.098	SDS
Aeroplane	52.096	60.927	64.404	CPMC
Bicycle	27.558	26.823	17.965	ALE
Bird	36.684	56.239	50.783	SDS
Boat	38.656	47.003	45.036	SDS
Bottle	43.643	48.465	41.605	SDS
Bus	65.787	70.559	69.104	SDS
Car	58.338	60.723	60.733	CPMC
Cat	63.789	59.847	56.524	ALE
Chair	24.001	20.815	11.663	ALE
Cow	64.853	42.112	52.842	ALE
Dining Table	41.339	38.694	19.406	ALE
Dog	55.190	51.535	48.995	ALE
Horse	58.998	43.653	43.899	ALE
Motorbike	56.909	52.300	52.858	ALE
Person	49.107	61.649	46.707	SDS
Potted Plant	31.408	37.360	40.563	CPMC
Sheep	53.563	51.829	49.285	ALE
Sofa	38.598	22.375	49.208	ALE
Train	53.910	56.288	58.319	CPMC
Average Accuracy	48.473	50.089	47.048	

Table 4 – Results of semantic segmentation accuracy on the test data set

Accuracy for each class (intersection/union measure)	Algorithm's Name				Best Algorithm
	ALE	SDS	CPMC	IA	
Background	54.878	80.061	77.478	62.157	SDS
Aeroplane	0.000	0.000	0.000	0.000	--
Bicycle	26.799	31.913	13.515	27.624	SDS
Bird	22.070	37.042	59.947	21.932	CPMC
Boat	0.000	0.000	0.000	0.000	--
Bottle	37.445	50.990	39.280	50.226	SDS
Bus	44.212	12.034	71.156	45.412	CPMC
Car	52.788	34.924	31.873	56.241	IA
Cat	63.939	65.552	62.707	63.802	SDS
Chair	19.113	22.355	7.800	19.014	SDS
Cow	33.093	0.000	0.000	30.991	ALE
Dining Table	39.155	50.907	23.997	40.169	SDS
Dog	60.085	49.253	49.827	59.148	ALE
Horse	46.406	27.761	27.155	47.128	IA
Motorbike	61.154	28.477	33.949	61.697	IA
Person	46.362	63.940	46.068	57.947	SDS
Potted Plant	25.762	36.391	25.045	23.245	SDS
Sheep	69.008	66.129	27.191	69.008	IA
Sofa	29.672	17.062	11.806	29.702	IA
Train	43.602	0.000	28.651	51.174	IA
TVmonitor	31.320	62.904	53.201	37.091	SDS
Average Accuracy	38.422	35.128	32.935	40.653	IA

6 DISCUSSION

Notice that not all algorithms tested have an average score in all categories: this is due to the fact for the images that contained the object cow was not detected not even once by neither SDS nor CPMC. Moreover observe that our approach IA is best only in few categories but in most of the categories is relatively close to the best one. As a result of using the statistical information for algorithm selection the IA approach results in the best semantic segmentation.

As a final comment on the importance of high level image description and content understanding, Figure 9 shows the results of three different semantic segmentation algorithms (Fig. 9c–Fig. 9e) and the result obtained by IA platform (Fig. 9f) that uses features and features and hypothesis attributes for algorithm selection. In the experiment illustrated in Fig. 9 the input image is shown in Fig. 9a. The first algorithm selected generated the result shown in Fig. 9c. The obtained semantic segmentation was analyzed for shape, proximity, position and relative size contradiction [27] and a hypothesis solving the contradiction is generated.

Using this hypothesis a new algorithm (Fig. 9e) was selected and the two results of semantic segmentation are merged. The result is shown in Fig. 9f. Notice the replacement of the chair (red region) from the initial result without removing any part of the sofa (green region).

CONCLUSION

In this paper we described some theoretical properties of algorithm selection. In particular we discussed the importance of the proper evaluation and the importance of hypothesis in the algorithm selection. The results show that for algorithms that are context sensitive – and most of algorithms used in real world application are context sensitive – the iterative approach proposed in this paper improves the overall computer vision and image understanding. The high level information was demonstrated to be very important – using only the statistics on the class level segmentation accuracy the algorithm selection approach provides best results and outperforms all the used algorithms.

Several extensions to this work are planned. The statistical information obtained during testing is not precise enough and thus will be explored in combination of features from the contradiction regions for increased accuracy of algorithm selection. The features used so far in the algorithm selection also require accuracy improvement by finding richer features. Such features have been recently obtained by the use of convolutional neural networks and we plan to integrate them into the IA platform. Finally, the hypothesis used is a simple object label obtained from measured properties such as objects
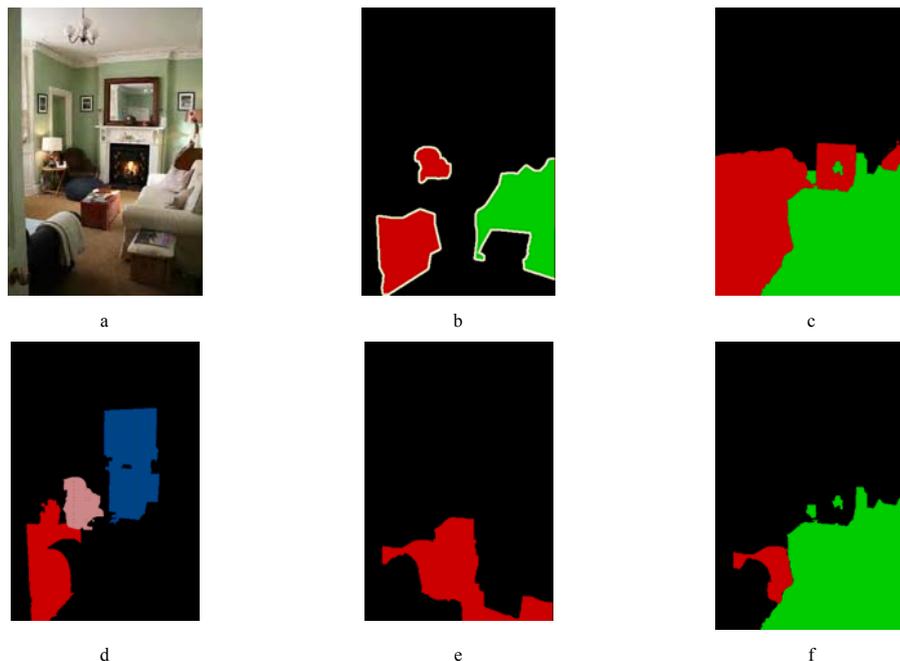

Figure 9 – An example of different stages of processing an input image using the algorithm selection platform: a – Input Image, b – Ground Truth, c – ALE Result, d – SDS Result, e – CPMC Result, f – IA Result

proximity, relative size and so on: to increase the accuracy of hypothesis generation a deeper semantic model connecting more object attributes and the objects with the environment of the world is to be build and used in close future.

REFERENCES

1. Rice J. The algorithm selection problem / J. Rice // Advances in Computers. – 1976. – Vol. 15. – P. 65–118.
2. Lukac M. Machine learning based adaptive contour detection using algorithm selection and image splitting / M. Lukac, R. Tanizawa, M. Kameyama // Interdisciplinary Information Sciences. – 2012. – Vol. 18, № 2. – P. 123–134.
3. Lukac M. Natural image understanding using algorithm selection and high level feedback / M. Lukac, M. Kameyama, K. Hiura // SPIE Intelligent Robots and Computer Vision XXX: algorithms and Techniques. – 2013. DOI: 10.1117/12.2008593
4. Zhang Y. Optimal selection of segmentation algorithms based on performance evaluation / Y. Zhang and H. Luo // Optical Engineering. – 2000. – Vol. 39, № 6. – P. 1450–1456.
5. Yong X. Optimal selection of image segmentation algorithms based on performance prediction / X. Yong, D. Feng, Z. Rongchun // Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing (VIP2003). – 2003. – P. 105–108.
6. Yu L. Feature selection for high-dimensional data: A fast correlation-based filter solution / L. Yu, H. Liu // Proceedings of the 20th International Conference on Machine Learning. – 2004. – P. 856–863.
7. Takemoto S. Algorithm selection for intracellular image segmentation based on region similarity / S. Takemoto, H. Yokota // Ninth International Conference on Intelligent Systems Design and Applications. –2009. – P. 1413–1418. DOI: 10.1109/ISDA.2009.205
8. Lukac M. Bayesian-network-based algorithm selection with high level representation feedback for real-world intelligent systems / M. Lukac, and M. Kameyama // Information Technology in Industry. – 2015. – Vol. 3, № 1. – P. 10–15.
9. Peng B. Parameter selection for graph cut based image segmentation / B. Peng, V. Veksler // In Proceedings of the British Conference on Computer Vision. – 2008. – P. 16.1–16.10. DOI: 10.5244/C.22.16
10. Hoiem D. Closing the loop on scene interpretation / D. Hoiem, A. A. Efros, M. Hebert // Proc. Computer Vision and Pattern Recognition (CVPR). – 2008. – P. 1–8. DOI: 10.1109/CVPR.2008.4587587
11. Ferryman J. Automated scene understanding for airport aprons / [J. Ferryman, M. Borg, D. Thirde and other] // Proceedings of 18th Australian Joint Conference on Artificial Intelligence. – 2005. – P. 593–603. DOI: 10.1007/11589990_62
12. Oliva A. Modeling the shape of the scene: a holistic representation of the spatial envelope / A. Oliva, A. Torralba // International Journal of Computer Vision. – 2001. – Vol. 42, № 3. – P. 145–175.
13. Ladicky L. Graph cut based inference with co-occurrence statistics / L. Ladicky, C. Russell, P. Kohli, and P. Torr // In Proceedings of the 11th European conference on Computer vision. –2010. – P. 239–253. DOI: 10.1007/978-3-642-15555-0_18
14. Carreira J. Object recognition by sequential figure-ground ranking / J. Carreira, F. Li, C. Sminchisescu // International Journal of Computer Vision. –2012. – Vol. 98, № 3. –P. 243–262.
15. Leibe B. Robust object detection with interleaved categorization and segmentation / B. Leibe, A. Leonardis, B. Schiele // International Journal of Computer Vision. – 2008. – Vol. 77. – P. 259–289.
16. Finding animals: Semantic segmentation using regions and parts / [Arbelaez P., Hariharan B., Gu C. and other] // International Conference on Computer Vision and Pattern Recognition. – 2012. – P. 3378–3385. DOI: 10.1109/CVPR.2012.6248077
17. Li L.-J. Towards total scene understanding: classification, annotation and segmentation in an automatic framework / L.-J. Li, R. Socher, L. Fei-Fei // Computer Vision and Pattern Recognition (CVPR). –2009. – P. 2036–2043. DOI:10.1109/CVPR.2009.5206718
18. Ladicky L. Inference methods for crfs with co-occurrence statistics / L. Ladicky, C. Russell, P. Kohli, P. Torr // International Journal of Computer Vision. – 2013. – Vol. 103, № 2. – P. 213–225.
19. Hariharan B. Simultaneous detection and segmentation / B. Hariharan, P. Arbelaez, R. Girshick, J. Malik // European Conference on Computer Vision (ECCV). – 2014. – P. 297–312. DOI: 10.1007/978-3-319-10584-0_20
20. Martin M. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics / M. Martin, C. Fowlkes, D. Tal, J. Malik // Proceedings of 8th International Conference on Computer Vision. – 2001. – P. 416–423. DOI: 10.1109/ICCV.2001.937655
21. Arbelaez P. Contour detection and hierarchical image segmentation / P. Arbelaez, M. Maire, C. Fowlkes, J. Malik // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2011. – Vol. 33, № 5. – P. 898–916.
22. Arbelaez P. Boundary extraction in natural images using ultrametric contour maps / P. Arbelaez // Computer Vision and Pattern Recognition Workshop. – 2006. – P. 182–190. DOI: 10.1109/CVPRW.2006.5.
23. Ren X. Multi-scale improves boundary detection in natural images / X. Ren // Proceedings of the 10th European Conference on Computer Vision. –2008. – P. 533–545. DOI: 10.1007/978-3-540-88690-7_40
24. Dollar P. Supervised learning of edges and object boundaries / P. Dollar, Z. Tu, S. Belongie // IEEE Computer Vision and Pattern Recognition (CVPR). – 2006. – P. 1964–1971. DOI: 10.1109/CVPR.2006.298
25. Using contours to detect and localize junctions in natural images / [M. Maire, P. Arbelaez, C. Fowlkes, J. Malik] // Conference on Vision and Pattern Recognition. – 2008. – P. 1–8. DOI: 10.1109/CVPR.2008.4587339
26. The pascal visual object classes (VOC) challenge / [M. Everingham, L. Van Gool, C. K. I. Williams and other] // International Journal of Computer Vision. – 2010. – Vol. 88, № 2. – P. 303–338.
27. Lukac M. Bayesian-network-based algorithm selection with high level representation feedback for real-world information processing / M. Lukac M. Kameyama // IT in Industry. – 2015. – Vol. 3, № 1. – P. 10–15.
28. Handling missing values in support vector machine classifier / [K. Pelckmans, J. De Brabanter, J.A.K. Suykens and other] // Neural Networks. – 2005. – Vol. 18, № 5–6. – P. 684–692.

Лукач М.[1], Абдиева К.[2], Камеяма М.[3]

[1]Д-р философии, ассистент кафедры компьютерных наук, Университет им. Назарбаева, Астана, Казахстан

[2]Аспирант, лаборатория ROSE, Наньянский технологический Университет, Сингапур

[3]Д-р наук, профессор, профессор школы информатики, Университе Тохоку, Сендай, Япония

ОЦЕНИВАНИЕ КОМПОНЕНТНЫХ АЛГОРИТМОВ ДЛЯ ВЫБОРА АЛГОРИТМА СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ НА ОСНОВЕ ОБРАТНОЙ СВЯЗИ С ВЫСОКИМ УРОВНЕМ ИНФОРМАЦИИ

Обсуждаются некоторые теоретические свойства подхода по выбору алгоритма для решения проблемы семантической сегментации в компьютерном зрении. Высококачественный выбор алгоритма возможен, только если пригодность каждого алгоритма хорошо известна, потому что только тогда результат выбора алгоритма может улучшить наилучший возможный результат, полученный одним алгоритмом. Показано, что оценка алгоритма зависит от конечной задачи; т.е. для того чтобы правильно оценивать алгоритм и определить его пригодность, необходимо использовать только хорошо сформулированные задачи. Когда пригодность алгоритма известна алгоритм может быть эффективно использован для задачи, применяясь в наиболее благоприятных условиях, определяемых в ходе оценивания. Оценивание, зависящее от задачи, продемонстрировано на сегментации и распознавании объектов. Кроме того, обсуждается важность символического знания высокого уровня в процессе отбора. Важность этой символической гипотезы продемонстрировано на наборе экспериментов по по обучению байесовской сети и SVM, а также с помощью статистических данных, полученные во время обучения селектора алгоритма. Показано, что для выбора эффективного алгоритма требуется оценивание, зависящее от

задачи. Показано, что используя символические предпочтения алгоритмов, точность выбора алгоритма может быть улучшена на 10–15%, а качество символической сегментации может быть улучшено до 5% по сравнению с наилучшим доступным алгоритмом.

Ключевые слова: выбор алгоритма, пригодность алгоритма, компьютерное зрение.

Лукач М.[1], Абдієв К.[2], Камеяма М.[3]
[1]Д-р філософії, асистент кафедри комп'ютерних наук, Університет ім. Назарбаєва, Астана, Казахстан
[2]Аспірант, лабораторія ROSE, Наньянський технологічний Університет, Сінгапур
[3]Д-р наук, професор, професор школи інформатики Універсам Тохоку, Сендай, Японія

ОЦІНЮВАННЯ КОМПОНЕНТНИЙ АЛГОРИТМІВ ДЛЯ ВИБОРУ АЛГОРИТМА СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ НА ОСНОВІ ЗВОРОТНОГО ЗВ'ЯЗКУ З ВИСОКИМ РІВНЕМ ІНФОРМАЦІЇ

Показано, що оцінка алгоритму залежить від кінцевого завдання; тобто для того щоб правильно оцінювати алгоритм і визначити його придатність, необхідно використовувати тільки добре сформульовані завдання. Коли придатність алгоритму відома, алгоритм може бути ефективно використаний для завдання, застосовуючись у найбільш сприятливих умовах, обумовлених у ході оцінювання. Оцінювання, залежне від завдання, продемонстровано на сегментації і розпізнаванні об'єктів. Крім того, обговорюється важливість символічного знання високого рівня у процесі відбору. Важливість цієї символічної гіпотези продемонстровано на наборі експериментів з навчання байєсівської мережі та SVM, а також за допомогою статистичних даних, отриманих під час навчання селектора алгоритму. Показано, що для вибору ефективного алгоритму потрібно оцінювання, залежне від завдання. Показано, що використовуючи символічні переваги алгоритмів, точність вибору алгоритму може бути поліпшена на 10–15%, а якість символічної сегментації може бути покращена до 5% у порівнянні з найкращим доступним алгоритмом.

Ключові слова: вибір алгоритму, придатність алгоритму, комп'ютерний зір.

REFERENCES

1. Rice J. The algorithm selection problem, *Advances in Computers*, 1976, Vol. 15, pp. 65–118.
2. Lukac M., Tanizawa R., Kameyama M. Machine learning based adaptive contour detection using algorithm selection and image splitting, *Interdisciplinary Information Sciences*, 2012, Vol. 18, No. 2, pp. 123–134.
3. Lukac M., Kameyama M., Hiura K. Natural image understanding using algorithm selection and high level feedback, *SPIE Intelligent Robots and Computer Vision XXX: algorithms and Techniques*, 2013. DOI: 10.1117/12.2008593
4. Zhang Y., Luo H. Optimal selection of segmentation algorithms based on performance evaluation, *Optical Engineering*, 2000, Vol. 39, No. 6, pp. 1450–1456.
5. Yong X., Feng D., Rongchun Z. Optimal selection of image segmentation algorithms based on performance prediction, *Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing (VIP2003)*, 2003, pp. 105–108.
6. Yu L., Liu H. Feature selection for high-dimensional data: A fast correlation-based filter solution, *Proceedings of the 20th International Conference on Machine Learning*, 2004, pp. 856–863.
7. Takemoto S., Yokota H. Algorithm selection for intracellular image segmentation based on region similarity, *Ninth International Conference on Intelligent Systems Design and Applications*. 2009, pp. 1413–1418. DOI: 10.1109/ISDA.2009.205
8. Lukac M., Kameyama M. Bayesian-network-based algorithm selection with high level representation feedback for real-world intelligent systems, *Information Technology in Industry,* 2015, Vol. 3, No. 1, pp. 10–15.
9. Peng B., Veksler V. Parameter selection for graph cut based image segmentation, *In Proceedings of the British Conference on Computer Vision*, 2008, pp. 16.1–16.10. DOI: 10.5244/C.22.16
10. Hoiem D., Efros A. A., Hebert M. Closing the loop on scene interpretation / D. Hoiem, // Proc. Computer Vision and Pattern Recognition (CVPR), 2008, P. 1–8. DOI: 10.1109/CVPR.2008.4587587
11. Ferryman J. Borg M., Thirde D., Fusier F., Valentin V., Bremond F., Thonnat M., Aguilera J., Kampel M. Automated scene understanding for airport aprons, Proceedings of 18th Australian Joint Conference on Artificial Intelligence, 2005, pp. 593–603. DOI: 10.1007/11589990_62
12. Oliva A., Torralba A. Modeling the shape of the scene: a holistic representation of the spatial envelope, *International Journal of Computer Vision*, 2001, Vol. 42, No. 3, pp. 145–175.
13. Ladicky L., Russell C., Kohli P., and Torr P. Graph cut based inference with co-occurrence statistics, *In Proceedings of the 11th European conference on Computer vision*, 2010, pp. 239–253. DOI: 10.1007/978-3-642-15555-0_18
14. Carreira J. Li F., Sminchisescu C. Object recognition by sequential figure-ground ranking, *International Journal of Computer Vision*, 2012, Vol. 98, No. 3, pp. 243–262.
15. Leibe B. Leonardis A., Schiele B. Robust object detection with interleaved categorization and segmentation, *International Journal of Computer Vision,* 2008, Vol. 77, pp. 259–289.
16. Arbelaez P. Hariharan B., Gu C., Gupta S., Bourdev L., Malik J. Finding animals: Semantic segmentation using regions and parts, *International Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3378–3385. DOI: 10.1109/CVPR.2012.6248077
17. Li L.-J., Socher R., Fei-Fei L. Towards total scene understanding: classification, annotation and segmentation in an automatic framework, *Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 2036–2043. DOI:10.1109/CVPR.2009.5206718
18. Ladicky L., Russell C., Kohli P., Torr P. Inference methods for crfs with co-occurrence statistics, *International Journal of Computer Vision*, 2013, Vol. 103, No. 2, pp. 213–225.
19. Hariharan B., Arbelaez P., Girshick R., Malik J. Simultaneous detection and segmentation, *European Conference on Computer Vision (ECCV)*, 2014, pp. 297–312. DOI: 10.1007/978-3-319-10584-0_20
20. Martin M., Fowlkes C., Tal D., Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, *Proceedings of 8th International Conference on Computer Vision*, 2001, P. 416–423. DOI: 10.1109/ICCV.2001.937655
21. Arbelaez P., Maire M., Fowlkes C., Malik J. Contour detection and hierarchical image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011,Vol. 33, No. 5, pp. 898–916.
22. Arbelaez P. Boundary extraction in natural images using ultrametric contour maps, *Computer Vision and Pattern Recognition Workshop*, 2006, pp. 182–190. DOI: 10.1109/CVPRW.2006.5.
23. Ren X. Multi-scale improves boundary detection in natural images, *Proceedings of the 10th European Conference on Computer Vision*, 2008, pp. 533–545. DOI: 10.1007/978-3-540-88690-7_40
24. Dollar P., Z. Tu, S. Belongie Supervised learning of edges and object boundaries, IEEE Computer Vision and Pattern Recognition (CVPR), 2006, pp. 1964–1971. DOI: 10.1109/CVPR.2006.298
25. Maire M. Arbelaez P., Fowlkes C., Malik J. Using contours to detect and localize junctions in natural images, *Conference on Vision and Pattern Recognition*, 2008, P. 1–8. DOI: 10.1109/CVPR.2008.4587339
26. Everingham M., Van Gool L., Williams C. K. I., Winn J., Zisserman A. The pascal visual object classes (VOC) challenge, *International Journal of Computer Vision*, 2010, Vol. 88, No. 2, pp. 303–338.
27. Lukac M., Kameyama M. Bayesian-network-based algorithm selection with high level representation feedback for real-world information processing, *IT in Industry*, 2015, Vol. 3, No. 1, pp. 10–15.
28. Pelckmans K., De Brabanter J., Suykens J.A.K., De Moor B. Handling missing values in support vector machine classifier, *Neural Networks*, 2005, Vol. 18, No. 5–6, pp. 684–692.