UDK 004.932.72:617-089

MACHINE LEARNING WITH DIFFERENT DIGITAL IMAGES CLASSIFICATION IN LAPAROSCOPIC SURGERY

M. R. Bayazitov, A. V. Liashenko, D. M. Bayazitov, T. V. Stoeva, T. L. Godlevska

Odesa National Medical University, Odesa, Ukraine

Abstract

The evaluation of the effectiveness of the automatic computer diagnostic (ACD) systems developed based on two classifiers – HAAR features cascade and AdaBoost for the laparoscopic diagnostics of appendicitis and ovarian cysts in women with chronic pelvic pain is presented. The training of HAAR features cascade, and AdaBoost classifiers were performed with images/ frames, which have been extracted from video gained in laparoscopic diagnostics. Both gamma-corrected RGB and RGB converted into HSV frames were used for training. Descriptors were extracted from images with the method of Local Binary Pattern (LBP), which includes both data on color characteristics («modified color LBP» - MCLBP) and textural characteristics, which have been used later on for AdaBoost classifier training. Classification of test video images revealed that the highest recall for appendicitis diagnostics was achieved after training of AdaBoost with MCLBP descriptors extracted from RGB images – 0.708, and in the case of ovarian cysts diagnostics – for MCLBP gained from RGB images – 0.886.

Developed AdaBoost-based ACD system achieved a 73.6% correct classification rate (accuracy) for appendicitis and 85.4% for ovarian cysts. The accuracy of the HAAR features classifier was highest in the case of ovarian cysts identification and achieved 0.653 (RGB) – 0.708 (HSV) values. It was concluded that the HAAR feature-based cascade classifier turned to be less effective when compared with the AdaBoost classifier trained with MCLBP descriptors. Ovarian cysts were better diagnosed when compared with appendicitis with the developed ACD.

Introduction. Automatic computer diagnostic (ACD) / classification of video – images is actual for minimally invasive abdominal surgery and endoscopy [1, 9, 13, 15]. ACD systems developed for tracking laparoscopic instrumentation [4], identifying zones of pathology proved their effectiveness [11, 13, 17].

As far as HAAR features exclude analyzing each pixel of an image, the total time for analysis is shortened [4]. It makes the process of recognition of images congruent with the velocity of video-frames flowing and justifies the exploration of the classifier based on the HAAR-features cascade for video-data analysis. Meanwhile, the main disadvantage of that classifier exploration is confined to a prolonged period of its training, which might be measured in months when tens of thousands of images are used. It is hard to avoid that inconvenience as far as increasing the number of images is proportional to the effectiveness of diagnostics [4].

To strengthen the effectiveness of the HAAR- feature-based classifier, we decided to use both color and texture features for training [3, 11]. Thus texture features such as the heightened mediana of greyscale along with entropy and contrast might be treated as informative differential indices for normal tissue state identification [6, 12, 13]. Besides, as an alternative to HAAR feature-based classifier, we have explored the AdaBoost classifier trained with a minimal number of descriptors gained from the Local Binary Pattern (LBP) method application [11]. The classical method of LBP manipulates with greyscale of color and ignores other colors' information. Instead, the modified LBP method, which includes data on color characteristics (modified color LBP- MCLBP) [10], was used in the present investigation for gaining color and texture descriptors [7]. We used earlier color characteristics as a basis for the distant uroflowgram building up [5] and early diagnostics of diabetes retinopathy [6].

The aim of the work is to work out and compare the effectiveness of ACD based on HAAR- features cascade classifier with AdaBoost- based ACD and trained to distinguish between the normal and pathological state of appendix and ovarium in women with chronic pelvic pain syndrome.

Materials and methods

The next steps performed in the course of collecting data and their analysis:

Calibration of a digital camera, which included white color balance and conversion of color scale into digital code;

The object was located in the frontal position, which was under inspectionю. The deviation from the right angle was up to 15+5 degrees, and the distance to the visualized zone was from 3 to 5 cm [7]. Those images which got in such a fashion used for both ACD training and testing;

Those zones of interest size were 60 x 60 pixels [3, 11]; in the course of the laparoscopic intervention, the speed of video frames was modified via using the low-frequency filter, and the size of the image was artificially modified from 30 x 30 up to 60 x 60 pixels, which was necessary for optimizing classificatory performance.

Gamma-correction of the gained image was performed with the recalculation of gamma-coefficient. The usage of gamma-correction in the course of preprocessing of primarily gathering information permits to identify relations between quantitative pixel characteristics and their actual brightness [7, 17].

Conversion of RGB scale into HSV one; such a conversion justified by the orientation of Haar features on the estimation of the intensity of pixels.

Training to HAAR features classifier, using both RGB and HSV images;

Training AdaBoost classifier with MCLB templates [3, 11]; key features, which used, were confined to mean, entropy, contrast, homogeneity, and excesses.

Results of classification stored at the database, and additional analysis performed later on.

All laparoscopic videos got with a laparoscopic camera with 5 mm aperture diameter Carl Storz Tricam Camera (Carl Storz, Germany) during 2013-2019 years. That camera had the analogous input (PAL 475 horizontal lines), and the incoming signal digitalized with the pixel density of 720 x 576 and capture was made with a video capture card "AVerMedia HD capture Studio 203" (Avermedia, France) and presented at the ACD interface (Fig. 1).
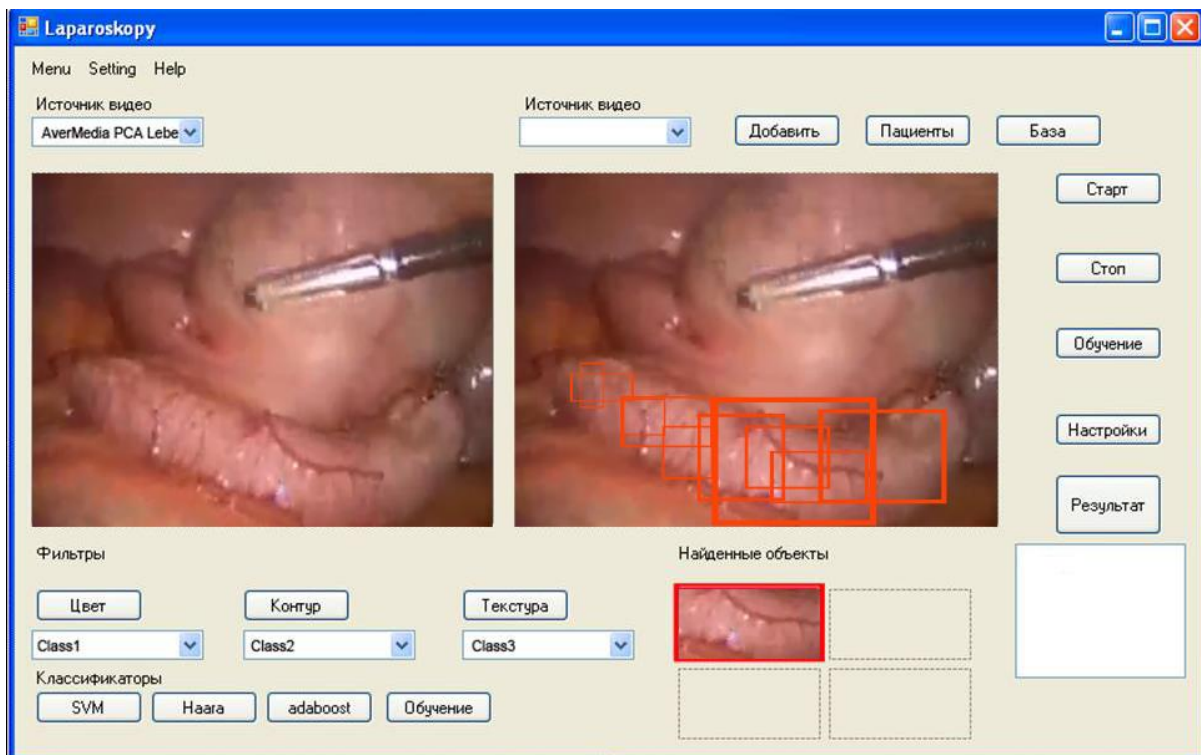
Fig. 1. The interface of the software illustrates zones of appendix inflammation (frames).

Features extraction and classifiers training. Being applied to the RGB scale, MCLBP calculates LBP for R and G channels of normalized RGB color space [10]. It served to get a more stable RGB – MCLBP under different conditions of illumination intensity.

The texture characteristics calculation using HSV - MCLBP was performed via recalculations on the Hue channel, invariant concerning illumination and saturation variability. For LBP calculation, the radius of 1.5 and 12 pixels was applied [10]. The pertinent pattern created for each scale vector, as a result, and the characteristic vector for templates of MCLBP, which included mean, entropy, contrast, homogeneity, and excesses, was determined [7, 10].

For the training of classifiers, 45 laparoscopic video images of patients with appendicitis and 43 with ovarian cysts were used as "positive" ones (Fig. 2). Also, for the classifier training, 40 videos gained from the normal appendix and 35 from the ovarian surface were used as control – "negative" images. Each video contained 2500 – 3000 frames, among which manually those for teaching and testing collections were verified, cropped out, and stored.
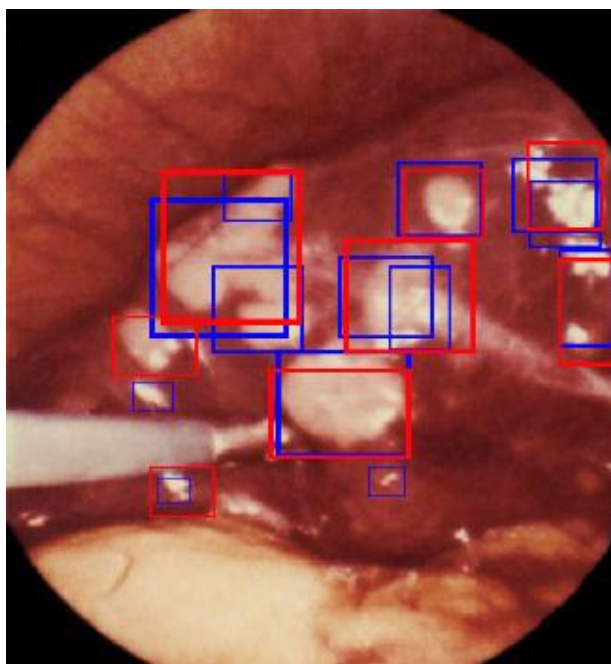
Fig. 2. The general scheme on the ACD work. With the squares, "zones of interest" are automatically defined and tracked.

For classifiers training, the next parameters were explored [11]:

- False-positive rate f = 0,3;

- Windows with the size of the frame as 60x60 pixels;

- Number of positive images - n=1000 for each pathology;

- Number of negative images - n=500;

After cessation of training, the tests were performed to estimate the effectiveness of recognition.

Test session images were different from those which have been used for the training of the classifier. Test control sessions were performed with 243 frames containing appendicitis and 132 frames with ovarian cysts. Three 346 frames with the absence of mentioned pathology were used as a control group – normal appendix (196) and ovarium (150).

Statistical procedures. To assess the performance of our classifiers, we use the measures precision, recall, and F-score [1, 13]. Precision measures the fraction of the detected-positive instances, which are true-positive (TP): Precision = TP/ (TP + FP), where FP – false positive instances.

A recall (TP/ P - number of positive instances) is the fraction of all true-positive instances, which are also detected positively.

F-score (also F-measure or F1-score) is the harmonic mean of precision and recall:

$$F = 2 * Precision * Recall/ Precision + Recall.$$

Accuracy is the proportion of correctly classified items out of all the items classified:

Accuracy = (TP + TN)/ (TP + TN + FP + FN), where TN - the number of true-negative, FN - the number of false-negative instances.

Results

Presented data showed that appendicitis diagnostics in women with chronic pelvic pain was less effective than diagnostics of ovarian cysts (Table 1). Such difference in effectiveness was observed in both pieces of training with RGB and HSV types of images. Meanwhile, the highest value of recall was observed for appendicitis diagnostics after training with RGB images for MCLBP – 0.886 and exceeded such one for training with HSV images for MCLBP, which occupied second place - 0.803. The lowest recall was registered for appendicitis diagnostics with HAAR features- based trained with RGB images (0.440). It should be stressed that in the course of diagnostics of appendicitis as well as diagnostics of cysts MCLBP- based methods revealed better diagnostics results when compared with the based on RGB and HSV images training of the classifiers.

Table 1

Comparative effectiveness of HAAR features-based and Ada-Boost classifiers trained with RGB and HSV images

Classifier	Frames type used for training	True positive	True negative	False positive	False negative	Precision	Recall	F1 Score	Accuracy
Appendicitis									
Haar-features cascade	RGB	107	113	83	136	0.563	0.440	0.494	0.501
	HSV	116	125	71	127	0.601	0.477	0.527	0.549
AdaBoost	MCLBP (RGB)	172	151	45	71	0.793	0.708	0.748	0.736
	MCLBP (HSV)	145	137	59	98	0.711	0.597	0.649	0.642
Ovarian cysts									
Haar-features cascade	RGB	83	105	45	49	0.654	0.629	0.641	0.653
	HSV	94	110	40	38	0.701	0.712	0.706	0.708
AdaBoost	MCLBP (RGB)	117	129	21	15	0.848	0.886	0.867	0.854
	MCLBP (HSV)	106	116	34	26	0.757	0.803	0.779	0.771

It is worth noting that training MCLBP with RGB images raised the number of true positive diagnoses pertained to that gained with HAAR-features cascade classifier in case of appendicitis diagnostics by 1.61 times and in case of ovarian cysts diagnostics – by 1.41 times. The corresponded reduction of false-negative diagnoses was 1.92 and 3.27 times. A less pronounced increase of the number of true positive diagnoses with MCLBP – HSV training was 1.25 times for appendicitis and 1.13 times for ovarian cysts. The reduction of false-negative diagnoses was 1.18 and 1.08 times correspondently.

Discussion

Hence, gained data favor the relatively high effectiveness of laparoscopic diagnostics of the diseases, which cause chronic pelvic pain in women using developed ACD systems. It should stress that such a result is in correspondence with early gained data in the course of liver pathology diagnostics [11]. Altogether our data prove the statement that only the Haar-like feature-based classifier is not enough reliable object classification [4].

One of the reasons for such a difference is that Haar feature-based cascade classifier needs more time for the laparoscopic stream of images recalculation, even though better diagnostic results followed training with HSV images. The vast number of variants of pathological manifestations (shape, color, texture) and the role of different orientations are also crucial for correct Haar feature-based cascade classifier application [2, 16].

Also, our data showed that diagnostics of appendicitis results were weaker than cysts diagnostics independently of images type, which was used for training. Such result points to less potential of HAAR features based classifier for correct video laparoscopic diagnostics when compared with AdaBoost results of diagnostics.

Hence, developed a diagnostic system based on training with modified templates of both RGB and HSV images and minimal MCLBP derived descriptors substantially improved results of classification performed with AdaBoost classifier. Our data showed that MCLBP descriptors got from RGB images drastically reduced false-negative diagnoses - by 1.92 for appendicitis and 3.27 times for ovarian cysts when compared with the corresponded data gained with Haar feature-based classifier exploration. The number of true positive diagnoses raised by 1.61 times and by 1.41 correspondently. Similar tendencies but less pronounced were registered for training with HSV type of images.

It should be noted that results on RGB images usage were better than those based on HSV image exploration. This fact is in favor of a better description of tissue properties in the RGB color scale [8, 17]. Although the net advantage of HAAR features based classification – shortened number of data for the machine learning as well as preventing overfitting of trained

classifier [14, 16], the exploration of MCLBP-based training of AdaBoost proved to be more effective for laparoscopic ACD of diseases which caused chronic pelvic pain in women. It is also possible that a cascade classifier trained with MCLBP descriptors could demonstrate better diagnostics performance as well, and AdaBoost trained with HAAR descriptors.

Altogether, our data points to a positive perspective with the usage of MCLBP descriptors for AdaBoost classifier training to resolve automatic diagnostics problems in laparoscopic surgery.

Conclusions

The ACD of laparoscopic images based on the AdaBoost classifier permitted effectively classify appendicitis, and ovarian cysts in women who suffered from chronic pelvic pain with the highest recall gained with MCLBP from HSV images used for training – up to 0.803, and for MCLBP RGB – up to 0.886 correspondently.

MCLBP used descriptors for training the AdaBoost classifier proved to increase the precision, recall, F1 score, and accuracy of automatic diagnostics of appendicitis and ovarian cysts.

Perspectives of future investigations. It is supposed to continue training both classifiers with increased laparoscopic images and extended forms of laparoscopy–defined pathology. Also, exploring the best version of developed ACD is supposed to perform in the operation room.

Disclosure

None of the authors has any conflict of interest to disclose. We declare that we are familiar with the official position on publication ethics, and this article was prepared in correspondence with actual demands and recommendations.

References

1. Albisser, Z.: Computer-aided screening of capsule endoscopy videos. Master's Thesis, University of Oslo. (2015) .Available at: https://www.duo.uio.no/handle/10852/47642

2. Bay, H, Tuytelaars, T, Van Gool, L.: Surf: Speeded up robust features. Computer vision–ECCV. 3951, 404-417 (2006). doi: 10.1007/11744023_32

3. Bayazitov, D.N., Kresyun, N.V., Buzinovsky, A.B., Bayazitov, N.R., Lyashenko, A.V., Godlevsky, L.S., Prybolovets, T.V., Bidnyuk, K.A: The effectiveness of automatic laparoscopic diagnostics of liver pathology using different methods of digital images classification. Pathologiya [Ukraine]. 14, 182–187 (2017). doi: 10.14739/2310-1237. 2017.2.109219

4. Fatiev, D.: Faculty of Science and Technology Department of Computer Science Object tracking for improved telementoring and telestration. Master's Thesis in Telemedicine and E-health, The Arctic University of Norway. (2015). Available at: https://hdl.handle.net/10037/7764

5. Godlevsky L. S., Kresyun N. V., Martsenyuk V. P., Shakun K. S., Tatarchuk T. V., Prybolovets K. O., Kalinichenko L. F., Karpinski M., Gancarczyk T. Information System for Early Diabetic Retinopathy Diagnostics based on Multiscale Texture Gradient Method. World Academy of Science, Engineering and Technology International Journal of Medical and Health Sciences Vol:14, No:8, 2020, waset.org/Publication/10011395. Available at: https://repo.odmu.edu.ua:443/xmlui/handle/123456789/8296

6. Godlevsky L.S., Shakun K.S., Martsenyuk V.P., Tatarchuk T.V., Stoeva T.V., Godlevska T.L., Shakun I.K.,Klos- Witkowska A. (2019). Dynamic changes of the color intensity of collected urine as a basis for a distant uroflowmetry. In: Proc. The 10 th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. 18-21 September, Metz, France. Pp. 308-312. DOI: 10.1109/IDAACS.2019.8924436

7. Guo, Z., Zhang, L., Zhang, D.: A completed modeling of local binary pattern operator for texture classification. In: IEEE Transactions on Image Processing, pp. 1657-1663, Publisher: Institute of Electrical and Electronics Engineers (2010). doi: 10.1109/TIP.2010.2044957

8. Hashimoto, D.A., Rosman G., Rus, D., Meireles O.R.: Artificial Intelligence in Surgery: Promises and Perils. Ann.Surg. 268, 70-76. (2018). doi: 10.1097/SLA.0000000000002693

9. Lahane, A., Yesha, Y., Grasso, M., Joshi, A., Park, A., Lo, J.: Detection of unsafe action from laparoscopic cholecystectomy video. In: Proc. 2nd ACM SIGHIT International Health Informatics Symposium, pp. 315-322. Association for Computing Machinery Press, New York (2012). https://doi.org/10.1145/2110363.2110400

10. Ledoux, A., Losson, O., Macaire, L.: Color local binary patterns: compact descriptors for texture classification. Journal of Electronic Imaging, 25(6), 061404. (2016). doi: 10.1117/1.JEI.25.6.061404

11. Lyashenko, A.V., Bayazitov, N.R., Godlevsky, L.S., Bayazitov, N.R., Buzinovsky, A.B.: Informational -technical system for the automatized laparoscopic diagnostics. Radio Electronics, Computer Science, Control [Ukraine]. 4, 90-96. (2016). Available at: http://repo.odmu.edu.ua:80/xmlui/handle/123456789/5532

12. Madad Zadeh, S., Francois, T., Calvet, L. Chauvet, P., Canis, M., Bartoli, A., Bourdel, N.: SurgAI: deep learning for computerized laparoscopic image understanding in gynaecology. Surg. Endosc. (2020). https://doi.org/10.1007/s00464-019-07330-8

13. Neofytou, M.S., Tanos, V., Constantinou, I., Kyriacou, E.C., Pattichis M.S., Pattichis, C.S.: Computer-aided diagnosis in hysteroscopic imaging. IEEE J. Biomed. Health. Inform. **19**(3), 1129-1136. (2015). doi: 10.1109/JBHI.2014.2332760

14. Petscharnig, S., Schöffmann, K.: Learning laparoscopic video shot classification for gynecological surgery. Multimed. Tools Appl. **77,** 8061–8079. (2018). doi: 10.1007/s11042-017-4699-5

15. Shu, Y., Bilodeau, G.A., Cheriet F.: Segmentation of laparoscopic images: Integrating graph-based segmentation and multistage region merging. In: Proc. 2nd Canadian Conference on Computer and Robot Vision (CRV'05) (2005). doi: 10.1109/ CRV. 2005.74

16. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In:Proc IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 511-518. (2001). doi**:** 10.1109/CVPR.2001.990517

17. Zhou, M, Bao, G, Geng, Y, Alkandari, B, Li, X.: Polyp detection and radius measurement in small intestine using video capsule endoscopy. In: Biomedical Engineering and Informatics (BMEI), 2014 7th International Conference on, IEEE, pp. 237-241. (2014). doi: 10.1109/BMEI.2014.7002777