Method of Endoscopic Images Analysis for Automatic Bleeding Detection and Segmentation

Nataliia Obukhova, Alexandr Motyko, Alexandr Pozdeev
Saint Petersburg Electrotechnical University "LETI"
Saint Petersburg, Russia
{natalia172419, motyko.alexandr}@yandex.ru, puches4@gmail.com

Boris Timofeev
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
timofeev-boris36@mail.ru

***Abstract*—Automatic detection and segmentation of bleeding on endoscopic images is an important task for use in clinical decision support systems, as well as in wireless capsule endoscopy. This can help to draw the physician's attention to a dangerous situation in time or reduce the length of the video sequence analyzed by the physician. The proposed method is based on block-based segmentation using local features, most of which are determined by color characteristics. Proposed method was tested on open database of endoscopic images KVASIR. Experiment shows that method can effectively perform block segmentation with an acceptable level of random and systematic error. The main advantages of the method are relatively low requirements for the size of a training database and high-speed performance.**

I. INTRODUCTION

The evolution from video providing systems to intelligent CDSS systems determined significant role of automatic image analysis. In such task, as cancer diagnostic, the detection of early cancer markers is the most logical and promising way of cancer prevention nowadays.

There are number of such markers: changing of vascular shape, specific mucosa degradation, the appearance and growing of polyps, etc. The one of the markers that is valuable in any diagnostic case is bleeding. The detection of bleeding in endoscopic examination is a popular task for endoscopic CDSS as well as for wireless capsule endoscopy. It helps to draw the physician attention on some abnormal situation in observed organ. First of all, it is very useful to detect bleeding automatically in long-lasting examination like capsule endoscopy of the gastrointestinal tract. It supports the physician during long examination of the received video data and highlights the fragments worthy of attention.

The existing approaches for automatic bleeding detection realize different methods and algorithms of digital image processing and machine learning. They could be classified by the image features they rely on for the analysis and by the decision strategy they use for detection. Suspected blood indicator (SBI) is a technique to detect bleeding from endoscopic images. However, in accordance with research and meta reports, its sensitivity and specificity are not sufficiently well [1], [2], [3]. Many approaches use color or texture features as the base for analysis [4], [5], [6], [7]. The others are based on contour analysis or various unsupervised segmentation techniques [8], [9]. As for the decision-making mechanism it is sometimes simple rule-based algorithm, or some kind of Naive Bayesian approach, but often it is based on powerful techniques as Support Vector Machines [10] or Artificial Neural Networks [11]. Therefore, recent advancement in the area of endoscopic bleeding detection algorithms resulted in establishing a great number of image processing techniques. However, it is obvious that there is still a strong need for a detailed study of the particular visual features, which are considered by the physicians during the examination of the images, and which were not sufficiently analyzed and described yet [12]. Moreover, every solution should be adapted for specific video system, sensor, data and environment.

Thus, the task of automatic bleeding detection cannot be considered as solved today. It requires more researches, new algorithms and solving strategies in the field of artificial intelligence. The modern artificial intelligence is based on technologies of machine learning and data mining. The fact is – elaborating a system with automatic analysis function is a very complex and long-lasting task. This task requires a lot of training data and many learning strategies and data mining algorithms should be tested or developed during the work.

The developing process of the algorithm for automatic bleeding detection for endoscopic video system consists of the set of stages which are described in the following sections. Specification of the database used in the investigation, data acquisition and brief description of the proposed method are given in section II. In section III the set of original local features and the procedure for extracting them from the image are described. Features analysis and filtering are given in section IV. Section V is devoted to choosing the best suitable machine learning algorithms for current task and testing them for finding the best solution. The post-processing procedure and experimental results are presented in section VI.

II. PROPOSED METHOD

In our investigation we used free-access The Kvasir Dataset [13]. The expanded Kvasir dataset contains 8000 endoscopic images of the gastrointestinal tract - 8 different classes (about 1000 images with polyps), checked and marked by experienced endoscopists. The resolution of images varies from 720x576 up to 1920x1072 pixels. The images were obtained by different devices in various conditions and examinations.

The dataset contains the extracted visual feature descriptors for all the image from the Kvasir Dataset. The extracted visual features are stored in the separate folders and files named accordingly to the name and the path of the corresponding image files. The extracted visual features are the global image features, namely: JCD, Tamura, ColorLayout, EdgeHistogram, AutoColorCorrelogram and PHOG.

However, the dataset does not contain any ground truth information about the presence of bleeding on the images. Moreover, the available features are global, related to the whole picture and do not provide the possibility to use them for segmentation task.

Thus, in our research we had to prepare the dataset by ourselves. We searched through all the images and manually collect about 90 images with bleeding from the entire dataset. The examples of those pictures are represented in Fig. 1.

Next the ground truth information had to be obtained. By using image editor, we created manually the masks for selected images. The examples are shown in Fig. 2 (the corresponded masks for the pictures from Fig. 1)

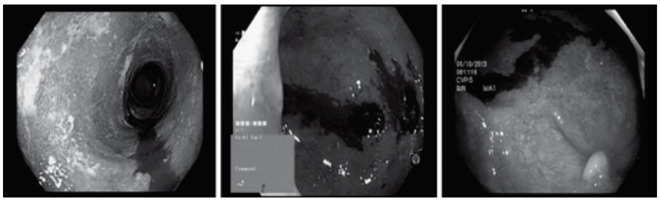


Fig. 1. Pictures with bleeding from Kvasir dataset

Fig. 2. Ground truth information for images from Fig. 1

The texture and geometry of bleeding varies within very wide range, which makes it difficult to use these features for the segmentation task. Color is the most appropriate feature for determining bleeding. The features provided by Kvasir are global and due to this reason, it would be possible to use them for images classification task. The result obtained by this strategy would be image classification if there is presence of bleeding or not. The disadvantage of this approach is the absence of bleeding localization on image. Such information of bleeding area and square is also unavailable.

Moreover, in our case we had only about 90 images in positive class (with bleeding). A small amount of data would make learning impossible.

For that reason, we proposed the block-based approach for features extraction. Block size may vary depending on the application and characteristics of a particular sensor. Block diagram of proposed algorithm is shown in Fig. 3. The main advantages of this approach are following:

- Small number of images with bleeding does not allow to train global classifier, that process the entirely image as one sample. But the extraction the small samples with and without bleeding significantly enlarge the dataset and enhance the diversity of samples, represented in each class.
- The block-based method allows not only classify the image on correspondence to some class, but provides the more valuable result of segmentation.

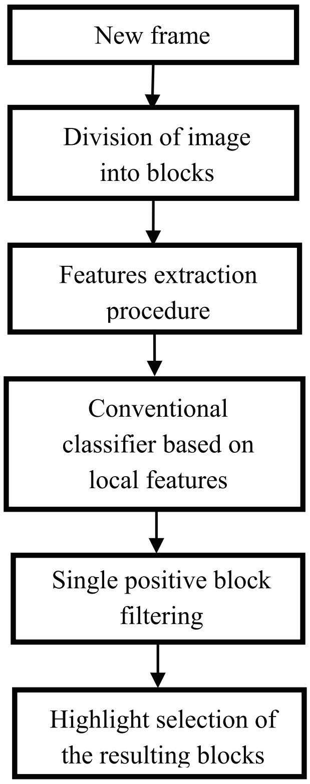


Fig. 3. Block-scheme of proposed algorithm

In the next sections the solution to the task of the current research – to form set of features for image analysis and bleeding detection and to design the method and the algorithms for their automatic extraction is described.

III. Features extraction

For features automatic extraction a specific software was designed. It gives the possibility to divide the images on blocks (16x16 pixels is default size, that can be tuned) and for each block it is possible to calculate different characteristics (brightness or color information, the amount of details etc.). According to the created masks of bleeding the software is able to separate the blocks by two classes – with bleeding and without it.

The process of data acquisition is very flexible. It can be tuned from which images the negative samples (without bleeding) should be extracted (from the images with the presence of bleeding samples, from the images without bleeding at all etc.), the software provides different data augmentation procedures (random crop, scaling).

The screenshot of feature extraction procedure is shown in Fig. 4. The image is divided into blocks. If more than half of the block area falls into the bleeding area on the ground image, this block is attributed as positive class.

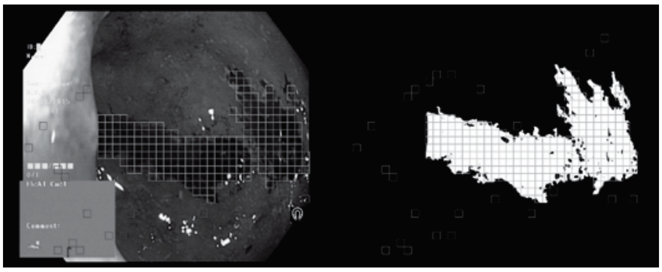


Fig. 4. The screenshot of the software for feature extraction

Thus, the algorithm of features mining consists of the following steps:

- To collect the pictures with bleeding;
- To make the ground truth masks;
- To divide the images to blocks;
- To attribute all the blocks with bleeding from corresponded images as positive samples, the blocks without bleeding from the images randomly selected from the whole DB as negative samples;
- As the amount of pictures with bleeding is small the data augmentation procedure for positive samples has to be processed to balance the number of samples in two classes.

It should be mentioned that the final datasets contain about 80000 negative samples and about 40000 positive samples.

The following features (calculated as the mean values for extracted blocks) were extracted and investigated. Color and brightness:

- R, G, B, r, g, b (normalized), B/R, G/R,
- X, Y, Z,
- L, a, b,
- Y, Cr, Cb,
- H, S, V

and texture:

- Ros, HOG.

The features from different color spaces allows to highlight various characteristic of the samples from both classes. The R, G, B values their normalized values and ratios and the X, Y, Z values represents the presence and weight of different channels in samples.

The L, a, b and Y, Cr, Cb color spaces allows to analyze the brightness and color information separately. For example, the conversion from R, G, B to L, a, b can be provided by the following nonlinear transformation (first the R, G, B is transformed to X, Y, Z and normalized on reference white values to become x_r, y_r, z_r):

$$\begin{gathered} L = 116 f_y - 16 \\ a = 500(f_x - f_y) \\ b = 200(f_y - f_z) \end{gathered} \tag{1}$$

where

$$\begin{gathered} f_x = \begin{cases} \sqrt[3]{x_r}, if\ x_r > e \\ \dfrac{k\, x_r + 16}{116}, otherwise \end{cases}, \\ f_y = \begin{cases} \sqrt[3]{y_r}, if\ y_r > e \\ \frac{k\, y_r + 16}{116}, otherwise \end{cases}, \\ f_z = \begin{cases} \sqrt[3]{z_r}, if\ z_r > e \\ \dfrac{k\, z_r + 16}{116}, otherwise \end{cases}, \end{gathered} \tag{2}$$

where $e = 0.008856$, $k = 903.3$.

The Hue-Saturate-Value (H, S, V) model is processed by following equations:

$$\begin{gathered} MAX = maximum(R, G, B), \\ MIN = minimum(R, G, B), \\ H = \begin{cases} 0, if\ MAX = MIN \\ 60\frac{G-B}{MAX-MIN} + 0, if\ MAX = R\ \&\ G \geq B \\ 60\frac{G-B}{MAX-MIN} + 360, if\ MAX = R\ \&\ G < B \\ 60\frac{B-R}{MAX-MIN} + 120, if\ MAX = G \\ 60\frac{R-G}{MAX-MIN} + 240, if\ MAX = B \end{cases} \\ S = \begin{cases} 0, if\ MAX = 0 \\ 1 - \dfrac{MIN}{MAX}, otherwise \end{cases}, \\ V = MAX. \end{gathered} \tag{3}$$

The Ros (Rosenfeld-Troy) metric corresponds to the amount of details on the image block (normalized by noise level).

The HOG features [14] is a popular technique to discover shapes within an image. The methods describe any shape of structure in the region by capturing information about gradients. To this purpose, the image is divided into small (usually 8x8 pixels) cells and blocks of 2x2 cells (in our case). Each cell has a fixed number of gradient orientation bins. Each pixel in the cell votes for a gradient orientation bin with a vote proportional to the gradient magnitude at that pixel. Thus each shape will be descripted by a set of HOG features and similar shapes will have close HOG values.

IV. Features Analysis and Filtering

On the next stage the obtained features were investigated. The main goals here:

- to eliminate the features with high mutual linear dependence – as they only enlarge the features space dimensionality but don't provide any additional information useful for classification;
- to filter the features which do not differ from different classes – as they don't provide separability.

The covariation matrix was computed and the features with high covariance were rejected as they may worsen the learning process. The rest of features were investigated for their separability potential. The final features set (after features filtering) included: H, S, a, b and Ros. The covariation matrix of the final features set is displayed in Table I.

TABLE I. THE COVARIATION MATRIX OF THE FINAL FEATURES SET

	H	S	a	b	Ros
H	1	-0.045	-0.750	-0.425	-0.026
S	-0.045	1	0.199	0.303	-0.145
a	-0.750	0.199	1	0.750	0.089
b	-0.425	0.303	0.750	1	0.007
Ros	-0.026	-0.145	0.089	0.007	1

It is important to mention, that the final features set seems very reasonable. There are not features connected with brightness, because brightness isn't the real feature of bleeding. The strongest features are connected with hue, saturation, color distribution and the amount of details. That corresponds to prior expectations.

V. INVESTIGATION AND CHOOSING OF CLASSIFICATION STRATEGY

The following ML methods were under investigation:

1) *Linear Discriminant Analysis* [15] – as the base level evaluation of the performance.
2) *Support Vector Machines* with radial-based function kernels [16] – as the most suitable algorithm in case of relatively small datasets.
3) *Random Decision Forests* [17].
4) *ADA Boost* [18].

The last two methods are very good also in case of nonlinear separability of classes. That is why it was the reasonable strategy to include them in investigation.

For classifier effectiveness estimation, we calculate the following metrics for the so-called confusion matrix: True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN). There are three commonly used performance measurements including accuracy, sensitivity and specificity based on confusion matrix values.

The accuracy of classifiers is the percentage of correctness outcomes among the test sets exploited in the study:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}. \quad (4)$$

The sensitivity is referred as the true positive rate and the specificity as the true negative rate:

$$sensitivity = \frac{TP}{TP + FP}, \quad (5)$$

$$specificity = \frac{TN}{TN + FN}. \quad (6)$$

The numeric results were obtained by 10-fold cross validation method [19]. In 10-fold cross-validation, the original sample is randomly partitioned into 10 equal sized subsamples. Of the 10 subsamples, a single subsample is retained as the validation data for testing the model, and the remaining 9 subsamples are used as training data. The cross-validation process was then repeated 10 times, with each of the 10 subsamples used exactly once as the validation data. The results were then averaged to produce a single estimation.

The obtained results are represented in the Table II.

TABLE II. THE 10-FOLD CROSS VALIDATION RESULT

Classifier	Accuracy	Sensitivity	Specificity
Linear	0.928	0.742	0.972
SVM	0.943	0.855	0.964
RDF	0.944	0.847	0.968
AdaBoost	0.929	0.788	0.964

The achieved metrics can be considered a successfully result for the classifier building. Moreover, the additional inference test was provided.

The dataset with 90 images containing bleeding was divided on the training set (70 images) and the test set of 20 images. Additional 20 images without bleeding were added to the test set for correct measurement of the false positive triggering.

After building the classifier on the base of training set (70 images with bleeding) the following metrics were achieved for the test set:

- accuracy = 0.95;
- sensitivity = 0.85;
- specificity = 0.97.

The results for reserved test set confirms the success of classifier building and of solving bleeding detection task.

Thus, the second task of current research – to investigate different classification strategy for automatic bleeding detection and estimate their effectiveness was successfully solved.

VI. EXPERIMENTAL RESULTS

The following inference results (Fig. 5 - 7) represent the analysis of reserved test set. The main things to mention here:

- the bleeding is successfully detected on the images, obtained by different devices, in various conditions, shapes, light and contrast;
- block segmentation causes some selection error that requires additional numerical evaluation;
- the images from negative class (without bleeding) did not triggered much false positives.

To improve the result, we used single positive block filtering as postprocessing procedure. The idea is based on assumption that the bleedings area in overwhelming most cases is manifested as several spatially connected blocks. Thus, the single positive blocks are related to false positives.

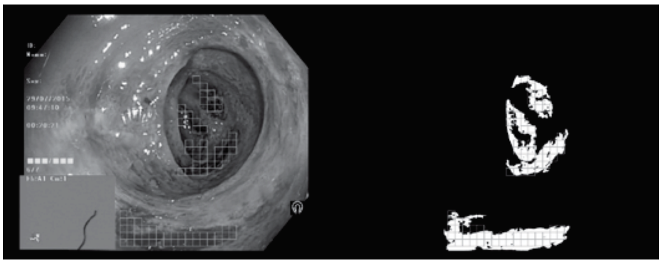


Fig. 5. The inference result of bleeding detection

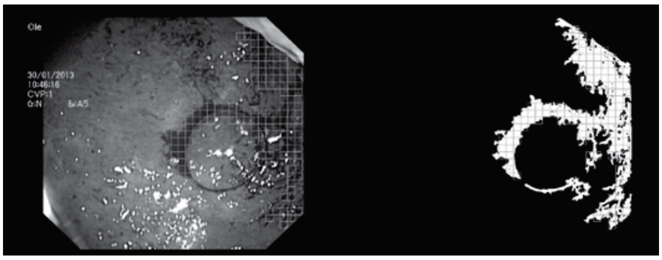


Fig. 6. The inference result of bleeding detection

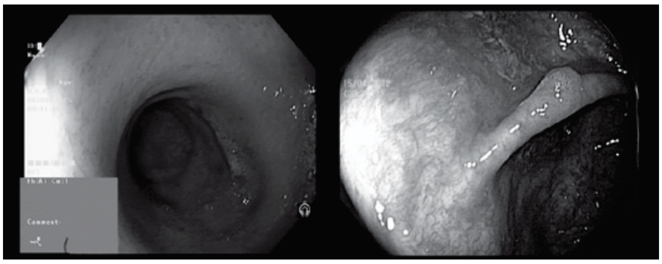


Fig. 7. The inference result of bleeding detection (there is no bleeding on the images)

The magnitude of the segmentation error in allocating an object of interest is an important characteristic of the segmentation method using blocks. The segmentation error in this case is an incomplete or redundant allocation of the object of interest. In the first case, only the part of the corresponding image is assigned to the object: the selected part is significantly smaller than the area of the object of interest in the image plane. Redundant segmentation is the assignment of image parts to an object that are not relevant to it: the selected part is much larger than the real area of the object in the image plane.

The image of the object of interest resulting from the segmentation is a set of blocks. Even if the bleeding segmentation was successful, the segmented image will not exactly match the real one. The latter describes an object with pixel accuracy, segmented with block accuracy. There is a systematic error caused by non-compliance of the discretization levels:

$$E_{syst} = \left| S_e - N_p^0 * s_b \right|, \tag{7}$$

where N_p^0 - the number of blocks assigned to the object as a result of segmentation, S_e - area of the object of interest in the image (the number of pixels related to the object), s_b - block area in pixels.

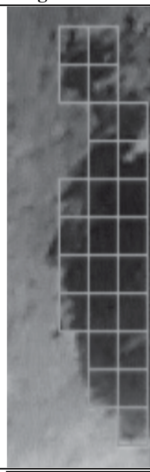
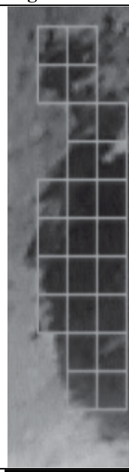
The systematic error is the difference between the real area of the object in the image and the area of its block approximation, which most closely corresponds to the real image of the object.

Random error is the accuracy of determining the number of blocks related to the object of interest N_p. Blocks that do not belong to it can be assigned to an object of interest or, conversely, the corresponding blocks are not included. Random error in the *i*-th image is the difference between the number of blocks assigned to the object N_p^i from N_p^0:

$$E_r^i = \left| N_p^0 - N_p^i \right|. \tag{8}$$

Table III shows an example of an image fragment with an ideal block segmentation and segmentation using a trained classifier. The number of blocks for the ideal case is 25, while the classifier has found 24 blocks. Thus, the random error for a current image fragment is 4%.

TABLE III. EXAMPLE OF IDEAL BLOCK SEGMENTATION AND SEGMENTATION USING A TRAINED CLASSIFIER

	Image fragment	**Ideal segmentation**	**Real segmentation**
image			
ground truth			
number of blocks	**-**	**25**	**24**

Estimates of random and systematic errors across the entire database with bleeding for a block size of 16 are following:

$$m\left(E_r^i\right) = 11.2\%,$$

$$m\left(E_{syst}^i\right) = 12.2\%.$$

The random and systematic errors are permissible, which makes it possible to use the block approach in the problem of bleeding segmentation.

VII. CONCLUSION

Method of endoscopic images analysis for automatic bleeding detection and segmentation was proposed. During the research the following tasks were solved:

- the set of features for image analysis and bleeding detection was formed;
- the method and the algorithm for features automatic extraction were designed and implemented in software.
- different classification strategy for automatic bleeding detection were investigated and their effectiveness for clinical decision support systems in endoscopy were estimated;
- the random and systematic errors are permissible.

This approach was investigated in condition of the relatively small total amount of positive samples in Dataset, which leads to using of conventional ML methods (SVM, RDF, ADA Boost). The general advantage here: fast processing and training combined with precision results. The block-based image analysis gives us the possibility to solve at the same time 2 tasks: sample classification and image segmentation.

The results show high quality of segmentation for images, obtained in different condition, with various sensors, different light sources and different diagnosis. Better results covering more complex and rare occurrences can be achieved with the corresponded enlargement of the training database.

ACKNOWLEDGMENT

The work was supported by the Russian Foundation for Basic Research, grant No. 17-07-00045.

REFERENCES

[1] S. Han, J. Fahed, and D.R. Cave, “Suspected Blood Indicator to Identify Active Gastrointestinal Bleeding: A Prospective Validation”, *Gastroenterology Res*. vol.11(2), Apr. 2018, pp. 106–111.

[2] Z. Liao, R. Gao, C. Xu and Z.S. Li, “Indications and detection, completion, and retention rates of small-bowel capsule endoscopy: a systematic review”, *Gastrointest Endosc*. vol.71(2), 2010, pp. 280–286.

[3] J. M. Buscaglia et al., “Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study”, *Clinical gastroenterology and hepatology*, vol.6(3), Mar.2008, pp. 298–301.

[4] Y. S. Jung, Y. H. Kim, D. H. Lee and J. H. Kim, “Active blood detection in a highresolution capsule endoscopy using color spectrum transformation”, *in Proc. BioMedical Engineering andInformatics, BMEI 2008, International Conference on*, vol.1, 2008, pp.859-862.

[5] P. Guobing, Y. Guozheng, Q. Xiangling and C. Jiehao, “Bleeding detection in wireless capsuleendoscopy based on probabilistic neural network”, *Journal of Medical Systems*, 2011, pp.1477-1484.

[6] S. Sainju, F. M. Bui and K. Wahid, “Bleeding detection in wireless capsule endoscopy based on color features from histogram probability”, *in Proc. 26th Annual IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, May 2013,

[7] P.M. Szczypiński et al, “Texture and color based image segmentation and pathology detection in capsule endoscopy videos”, *Computer Methods and Programs in Biomedicine*, vol.113(1), 2014, pp. 396-411.

[8] Y. Fu, M. Mandal and G. Guo, “Bleeding region detection in WCE images based on colorfeatures and neural network”, *in Proc. Circuits and Systems (MWSCAS), IEEE 54th InternationalMidwest Symposium on*, 2011, pp.1-4.

[9] G. Pan, F. Xu and J. Chen, “A novel algorithm for color similarity measurement andthe application for bleeding detection in WCE”, *International Journal of Image, Graphics and Signal Processing*, vol. 3(5), 2011, pp.1-7.

[10] Y. Xiong et al, “Bleeding detection in wireless capsule endoscopy based on MST clustering and SVM”, *in Proc. IEEE Workshop on Signal Processing Systems (SiPS)*, 2015.

[11] J. Xiao and M.Q. Meng, “A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images”, *in Proc. IEEE Eng Med Biol Soc.*, 2016, pp.639-642.

[12] Brzeski, A. Blokus and J. Cychnerski, “An Overview of Image Analysis Techniques in Endoscopic Bleeding Detection”, *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 1(6), 2013, pp.1350-1357.

[13] K. Pogorelov et. al, “Kvasir: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection”, *in Proc. of the 8th ACM on Multimedia Systems Conference (MMSYS)*, 2016, pp. 164-169.

[14] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection”, *in Proc. of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2005, pp.1-8.

[15] S. Mika et al. “Fisher Discriminant Analysis with Kernels”, *in Proc. of the IEEE Conference on Neural Networks for Signal Processing IX*, 1999, pp. 41–48.

[16] Y.-W. Chang et al. “Training and testing low-degree polynomial data mappings via linear SVM”, *Journal Machine Learning Research*, vol.11, 2010, pp. 1471–1490.

[17] T. Hastie, R. Tibshirani and J. Friedman, “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”. Heidelberg: 2nd ed. Springer-Verlag, 2009.

[18] Y. Freund, R. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting”, *Journal of Computer and System Sciences*, vol. 55, 1999, pp 119-139.

[19] R. Christensen “Thoughts on prediction and cross-validation”, Department of Mathematics and Statistics University of New Mexico, 2015, Web: http://www.math.unm.edu/~fletcher/Prediction.pdf.