

## RESEARCH ARTICLE

# Automatic Colorectal Polyp Detection in Colonoscopy Video Frames

Geetha K<sup>1</sup>, Rajan C<sup>2\*</sup>

### Abstract

Colonoscopy is currently the best technique available for the detection of colon cancer or colorectal polyps or other precursor lesions. Computer aided detection (CAD) is based on very complex pattern recognition. Local binary patterns (LBPs) are strong illumination invariant texture primitives. Histograms of binary patterns computed across regions are used to describe textures. Every pixel is contrasted relative to gray levels of neighbourhood pixels. In this study, colorectal polyp detection was performed with colonoscopy video frames, with classification via J48 and Fuzzy. Features such as color, discrete cosine transform (DCT) and LBP were used in confirming the superiority of the proposed method in colorectal polyp detection. The performance was better than with other current methods.

**Keywords:** Colon cancer- colonoscopy - Local Binary Pattern (LBP) - J48 - Fuzzy - Discrete Cosine Transform (DCT).

*Asian Pac J Cancer Prev*, **17** (11), 4869-4873

### Introduction

Colorectal cancers are the second most common type of cancer occurring in women, and the third most among men. Colorectal polyps are a precursor to colorectal cancers. They usually lead to cancer, if the polyps are not treated. Colon capsule endoscopy (CCE) is novel alternate to traditional techniques like colonoscopy or computed tomography (CT) colonography (Mamonov et al., 2014). Colorectal polyps are anomalous growths in the mucosa of colon or rectum. They have two distinctive shapes: pedunculated or sessile. The typical morphology of such polyps are the cornerstone (Filipe and Condessa, 2011). A polyp is typically a protrusion in folds. Curvature techniques may be employed for singling out the polyp.

Colonoscopy is a useful method for screening for colorectal cancers. Survival is more likely if the cancer is discovered at an earlier stage prior to metastasis to lymph nodes or other organs happens (Muthukudage et al., 2011). Detection of a polyp is more frequent at the time of withdrawing the colonoscope and so the procedure should not be rushed (Riley, 2011).

CAD helps in detecting polyps by identifying the generic shapes that look like bumps. In recent times, various prototype colon CAD frameworks with CRC have been proposed for detecting polyps so as to reduce False Positive (FP) results (Li et al., 1999). Several visual attributes are used, like texture, colour, shape as well as other combinations. Color and texture are generally combined for the analysis of colonoscopy image. Textural attributes like those based on local binary patterns (LBP), texture spectrum (TS) or gray level co-occurrence matrix

(GLCM) are utilized as a resolution in relevant research.

Colour-textural features, textural features, and colour features are computed separately and then used together as the basis of the classifier. LBP operators offer efficient methods to analyse textures. They have a fundamental theoretical model and merge features of structure and statistic texture analysis methods.

Manivannan (2015) suggested colonoscopy as well as histology (cell) images learning extremely discriminative local attributes and image representations for achieving the best classification possible for medical images. Various methods were proposed for this purpose.

Manivannan and Trucco (2015) suggested a new weak-supervised features learning method, learning discriminative local attributes from image-level labeled data for classifying images. Unlike already present features learning methods that presume that a set of extra data in the format of matching/non-matching pairs of local patches are provided for learning the attributes the method utilizes solely the image-level labels that are more easy for obtaining.

Fu et al., (2014) suggested a CAD framework for colonoscopy imaging for classifying colorectal polyps according to their types. The components of the suggested framework contained image enhancement, features extraction, features selection, as well as polyps classification. The textural analyses features, spatial domain features, as well as spectral domain features were taken from the first part of the PCT. Support Vector Machine (SVM) was utilized for classifying the colorectal polyps. Hence, the suggested CAD model improved the quality of the diagnoses of colorectal polyps.

<sup>1</sup>Department of Information Technology, Excel Engineering College, <sup>2</sup>Department of Information Technology, K S Rangasamy College of Technology, India. \*For Correspondence: rajancsg@gmail.com

Tajbakhsh et al., (2016) provided the results of the study into the design of a CAD model to detect polyps in colonoscopy videos. The model has its basis in a hybrid context-shape method that uses context data for removing non-polyp structures as well as shapes data for reliably localizing polyps.

**Material and Methods**

Figure 1 shows the flowchart for the proposed technique. The techniques used are detailed here.

*Color Features*

Features qualifying the colour of LD muscles were extracted from LD scan. The following metrics were computed for R, G as well as B colour functions of every LD scan of muscle colour attributes:  $\mu_R, \mu_G, \mu_B$  that represents which statistic the average colour feature,  $\sigma_R, \sigma_G, \sigma_B$  standard deviation, representing the non-uniformity of colours, and  $M_R, M_G, M_B$  third moments which indicate skewness or imbalance of colours [10].

Colour moments represent colour features for qualifying coloured scans. Features which are able to be included are mean ( $\mu$ ), standard deviation ( $\sigma$ ), skewness ( $\theta$ ), as well as kurtosis ( $\gamma$ ). For RGB color space, 3 attributes are extracted from every plane, R, G as well as B. The formulae for obtaining the moments are the following:

$$\begin{aligned} \mu &= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N P_{ij} \\ \sigma &= \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (P_{ij} - \mu)^2} \\ \theta &= \frac{\sum_{i=1}^M \sum_{j=1}^N (P_{ij} - \mu)^3}{MN\sigma^3} \\ \gamma &= \frac{\sum_{i=1}^M \sum_{j=1}^N (P_{ij} - \mu)^4}{MN\sigma^4} \end{aligned} \tag{1}$$

as well as N indicate the image dimensions.  $P_{ij}$  indicates values of colour at ith column and jth row.

*Local Binary Pattern (LBP) Feature*

The LBP feature is garnering greater interest currently in the domain of textures classification as well as retrieval. For the former, LBP is transformed into a rotational invariant form. Currently, imaging technology permits scholars to gather a huge quantity of digitized images in the IR domain. The development of an automated system for searching, retrievals, or classification of images from the database is significant (Jabid et al., 2010). LBP is the most common textural descriptor in IR analyses. Colon scans are split into many regions from which the LBP features distribution are taken and concatenated into an improved features vector for usage as an image features descriptor.

Every pixel is labelled via thresholding P-neighbor values with centre value and converting result into binary numbers via

$$LBP_{p,r}(x_c, y_c) = \sum_{p=0}^{r-1} s(g_p - g_c) 2^p, s(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Wherein  $g_c$  represents the gray value of the centre pixel  $x_c, y_c$  while  $g_p$  represents the gray values of

equally spaced pixels P on the circumference of circles with radius R. The values of neighbours which are not exactly on the pixel locations are estimated via bi-linear interpolation (Mamonov et al., 2014). Implies that the signs of differences in a neighbourhood are realized as P-bit binary numbers, resulting in  $2^p$  distinctive values for binary patterns. Individual pattern values are capable of describing texture data at center pixel  $g_c$ .

LBP technique uses 28 =256 possible texture units rather than the 38 =6561 units used in the textural spectrum technique, resulting in a more effective abstraction of textures that lead to a comparable textures discrimination performance. LBP of a 3 × 3-pixel neighbourhood is predicted thus:

(i) Original 3×3 neighbourhood (Figure 2a) is thresholded to 2 levels (0 and 1) utilizing centre pixel value.

(ii) Values of the pixels in the thresholded neighbourhood are multiplied by specific weights designated to the related pixels.

(iii) Values of the 8 pixels are summed for obtaining one value for the related pattern.

LBP feature vectors are created through histogram bins of the LBP value distribution in the image area.

*Discrete Cosine Transform (DCT)*

DCT is a type of transform that is like Discrete Fourier Transform (DFT) and associated with Fourier. It possesses a strong energy compression feature and reaches a near-optimal value in compression efficacy. The single dimension DCT (Zhang et al., 2011) transformation may be given thus by (Muthukudage et al., 2011):

$$X(k) = \alpha(k) \sum_{n=0}^{N-1} x(n) \cos \left[ \frac{\pi(2n+1)k}{2N} \right] \quad 0 \leq k \leq N-1 \tag{3}$$

Inverse DCT transformation may be expressed as (4):

$$x(n) = \sum_{k=0}^{N-1} \alpha(k) X(k) \cos \left[ \frac{\pi(2n+1)k}{2N} \right] \quad 0 \leq k \leq N-1 \tag{4}$$

In (4),  $\alpha(0) = \sqrt{1/N}, \alpha(k) = \sqrt{2/N}, 1 \leq n \leq N-1$

*Fused Scores*

Fusion methods are utilized in several applications. Within biometric applications, they merge matching scores from various modalities on the basis of generalized densities predicted from score spaces, although this needs huge amounts of training datasets for approximating density distributions.

Linear fusion, resulting in fused scores s from base classifier scores  $s = \sum_k s_k w_k$ , are a generic method for combining probabilities with non-negative weights, i.e.  $w_k \geq 0$ . This has its basis in intuition as higher probability outputs from base classifiers represent huge probabilities after fusion, and distinct weights describe relative performance of classifiers. The protocol has given powerful performance (Li et al., 2012) particularly with respect to generalizations on unknown testing data sets. For combining visual as well as text searches, several fusion techniques are able to be used. The combinations lead to enhanced results instead of singular modality. Retrieval of texts generally leads to enhanced performance

Table 1. Summary of Results for Features Extracted from Single Frame

	Classification Accuracy	Sensitivity	Specificity	F Measure
J48-Color+LBP	88.9	0.847	0.910	0.836
J48-Color+DCT	89.6	0.855	0.917	0.846
J48-Color+DCT+LBP	92.6	0.898	0.940	0.890
Fuzzy-Color+LBP	92.2	0.889	0.938	0.884
Fuzzy-Color+DCT	93.5	0.919	0.942	0.904
Fuzzy-Color+DCT+LBP	94.3	0.932	0.949	0.916

over visual retrieval, in medical retrieval. Therefore, the best combination strategies are to be chosen for improving performance appropriately. Earlier fusions include unimodal features before making decisions. As decision is based on information source, purely multimodal feature representation is guaranteed. Unimodal feature vectors are merged into one vector via weighting schemes.

#### J48

Classification refers to the procedure of constructing a model of classes from a set of records which comprise class labels. Decision Tree Algorithm refers to the finding of the way in which feature vector functions in several cases. Furthermore, based on the training samples, the classes for the new samples are discovered. The protocol produces the rules for predicting target parameter. With the assistance of tree classification protocol, the critical distribution of data is understandable (Kaur and Chhabra, 2014).

J48 classifier is a simple C4.5 decision tree for classification. It builds a binary tree. The decision tree method is useful in classification problems. Through this method, trees are built for modelling the classification procedure. Once the trees are built, they are applied to all tuples in the dataset and resulting in classifications for the tuples.

INPUT :  $D$

OUTPUT :  $T$

DTBUILD (\* $D$ )

{

$T = \varnothing$ ;

$T =$  Build root node & label with splitting feature;

$T =$  Include arc to root node for every split predicate and label;

For every arc do

$D =$  Database built through application of splitting predicate to  $D$ ;

If terminating point is arrived at for this route, then

$T' =$  construct leaf node and label with correct class;

Else

$T' =$  DTBUILD( $D$ );

$T =$  include  $T'$  to arc;

When constructing trees, J48 ignores missing values, that is, values for those items may be estimated on the basis of what is known regarding the features values for other records. The fundamental notion is the division of data into ranges on the basis of feature values for the item that is discovered in the training samples. J48 permits classification through either decision trees or rules created from them.

#### Fuzzy Classifier

Fuzzy Classifier learns fuzzy rules as well as

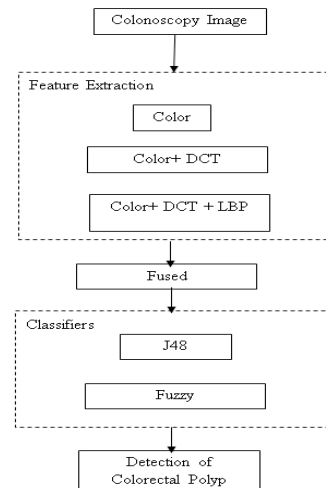


Figure 1. Flow Chart for the Suggested Technique

unordered rule sets. The protocol induces rules for all classes distinctly utilizing the “one class – other classes” division scheme. When classifiers are trained utilizing a class, other classes are not taken into consideration. This assists in the achievement of a state where there is not a single main rule and the sequence of classes in the training procedure is not relevant (Gasparovica and Aleksejeva, 2011).

Denoting fuzzy rules is distinctive where intervals are replaced by fuzzy intervals known as fuzzy sets, with trapeze kind membership functions (Hühn and Hüllermeier, 2009).

$$I^F(v) = \begin{cases} 1 & \phi^{c,L} \leq v \leq \phi^{c,U} \\ \frac{v - \phi^{s,L}}{\phi^{c,L} - \phi^{s,L}} & \phi^{s,L} \leq v \leq \phi^{c,L} \\ \frac{\phi^{s,U} - v}{\phi^{s,U} - \phi^{c,U}} & \phi^{c,U} \leq v \leq \phi^{s,U} \\ 0 & \text{else} \end{cases} \quad (5)$$

Where  $\phi^{c,L}$  as well as  $\phi^{c,U}$  represent the lower as well as upper bounds of the membership of fuzzy sets. FURIA has disadvantages such as when records are enclosed by rules of greater than a single class equally, confidence factor is to be calculated. The primary enhancements to RIPPER are associated with branching. However, the primary benefit of the protocol is a rule stretching technique, that resolves the issue that fresh records that ought to be classified utilizing the induced rules may not be encompassed by the earlier induced rules. The abstraction of the fuzzy rules is also distinct – the intervals are substituted by fuzzy intervals, which are fuzzy sets, with trapeze-type membership functions

Table 2. Summary of Results for Features Extracted from Five Frames

	Classification Accuracy	Sensitivity	Specificity	F Measure
J48-Color+LBP	91.5	0.889	0.927	0.875
J48-Color+DCT	92.2	0.898	0.934	0.885
J48-Color+DCT+LBP	93.6	0.923	0.942	0.906
Fuzzy-Color+LBP	94.0	0.923	0.949	0.912
Fuzzy-Color+DCT	95.7	0.953	0.959	0.937
Fuzzy-Color+DCT+LBP	96.2	0.953	0.966	0.943

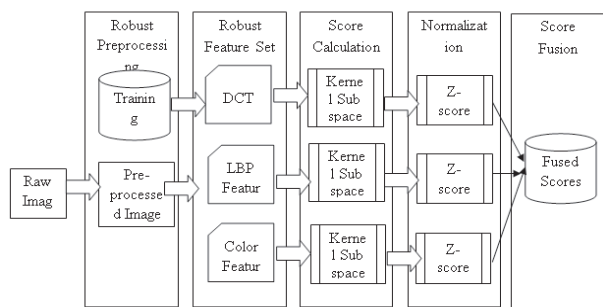


Figure 2. The Overall Architecture of Our Multi-Feature Subspace

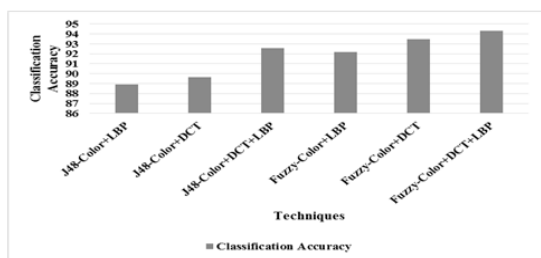


Figure 3. Classification Accuracy

**Results**

For experiments, 235 colon images with polyp and 468 normal images are considered. Table 1 and Figure 3 to 6 shows the classification accuracy, Sensitivity, Specificity and F measure for features extracted from one frame

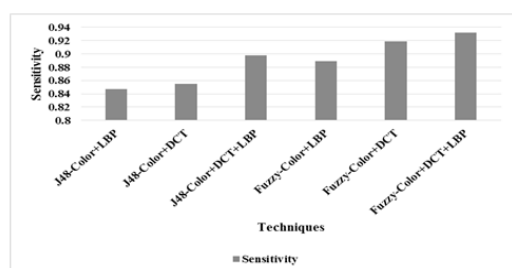


Figure 4. Sensitivity

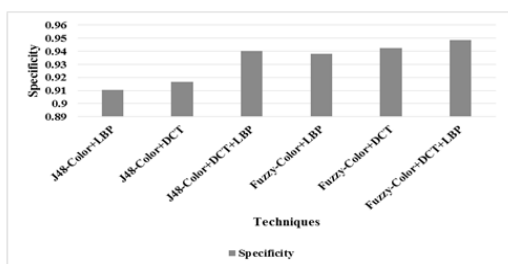


Figure 5. Specificity

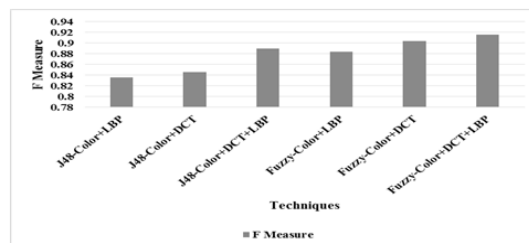


Figure 6. F measure

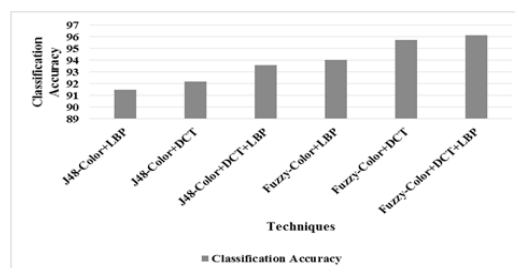


Figure 7. Classification Accuracy

respectively. Table 2 and figure 7 to 10 shows the same for features extracted from five frames.

From Table 1 as well as Figure 3 it is seen that the classification accuracy of Fuzzy-color, DCT and LBP for single frame performs better by 5.91% than J48-color LBP, by 5.1% than J48-color, DCT, by 1.83% than J48-color, DCT and LBP, by 2.3% than Fuzzy-color, LBP and by 0.91% than Fuzzy-color, DCT.

From table 1 as well as figure 4 it is seen that the

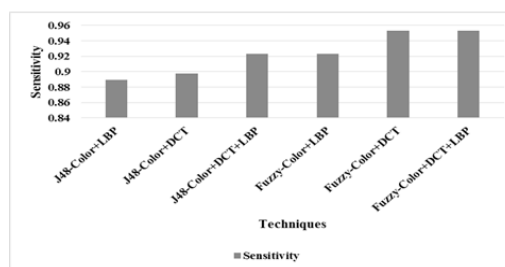


Figure 8. Sensitivity

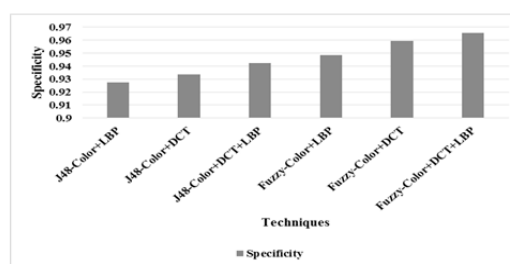


Figure 9. Specificity

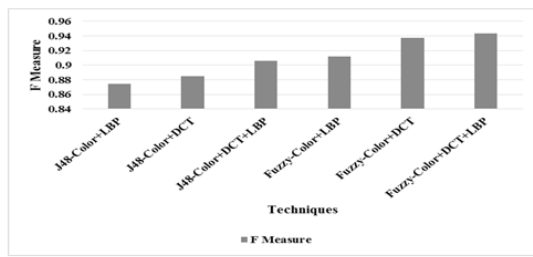


Figure 10. F measure

sensitivity of Fuzzy-color, DCT and LBP for single frame performs better by 9.6% than J48-color LBP, by 8.6% than J48-color, DCT, by 3.71% than J48-color, DCT and LBP, by 4.7% than Fuzzy-color, LBP and by 1.4% than Fuzzy-color, DCT.

From Table 1 as well as figure 5 it is seen that the specificity of Fuzzy-color, DCT and LBP for single frame performs better by 4.13% than J48-color LBP, by 3.42% than J48-color, DCT, by 0.9% than J48-color, DCT and LBP, by 1.13% than Fuzzy-color, LBP and by 0.68% than Fuzzy-color, DCT.

From Table 1 as well as figure 6 it is seen that the F measure of Fuzzy-color, DCT and LBP for single frame performs better by 9.15% than J48-color LBP, by 7.94% than J48-color, DCT, by 2.88% than J48-color, DCT and LBP, by 3.6% than Fuzzy-color, LBP and by 1.38% than Fuzzy-color, DCT than J48-color, DCT and LBP, by 4.7% than Fuzzy-color, LBP and by 1.4% than Fuzzy-color, DCT.

From Table 2 as well as Figure 7 it is seen that the classification accuracy of Fuzzy-color, DCT and LBP for five frame performs better by 4.99% than J48-color LBP, by 4.23% than J48-color, DCT, by 2.69% than J48-color, DCT and LBP, by 2.24% than Fuzzy-color, LBP and by 0.45% than Fuzzy-color, DCT.

From Table 2 and figure 8 it is observed that the sensitivity of Fuzzy-color, DCT and LBP for single frame performs better by 6.93% than J48-color LBP, by 5.97% than J48-color, DCT, by 3.18% than J48-color, DCT and LBP, by 3.18% than Fuzzy-color, LBP and no difference than Fuzzy-color, DCT.

From Table 2 and Figure 9 it is observed that the specificity of Fuzzy-color, DCT and LBP for five frames performs better by 4.1% than J48-color LBP, by 3.37% than J48-color, DCT, by 2.5% than J48-color, DCT and LBP, by 1.79% than Fuzzy-color, LBP and by 0.66% than Fuzzy-color, DCT.

From Table 2 as well as Figure 10 it is seen that the F measure of Fuzzy-color, DCT and LBP for five frame performs better by 7.55% than J48-color LBP, by 6.4% than J48-color, DCT, by 4.01% than J48-color, DCT and LBP, by 3.37% than Fuzzy-color, LBP and by 0.63% than Fuzzy-color, DCT.

## Discussion

Colorectal cancers occur in colon or in the rectum. They typically develop initially as a colorectal polyp, a growth within the colon or the rectum which becomes cancerous later on. The results of the experiment show

that the classification accuracy of Fuzzy-color, DCT and LBP for single frame outperforms by 5.91% than J48-color LBP, by 5.1% than J48-color, DCT, by 1.83% than J48-color, DCT and LBP, by 2.3% than Fuzzy-color, LBP and by 0.91% than Fuzzy-color, DCT. Likewise, the classification accuracy of Fuzzy-color, DCT and LBP for five frame outperforms by 4.99% than J48-color LBP, by 4.23% than J48-color, DCT, by 2.69% than J48-color, DCT and LBP, by 2.24% than Fuzzy-color, LBP and by 0.45% than Fuzzy-color, DCT.

## Reference

- Bhuvana S, Bhuvaneshwari AJ (2015). Computer aided automatic detection of polyp for colon tumor. *Comput Med Imaging Graph*, **4**, 431-38.
- Filipe JC, Condessa AND. (2011). Detection and Classification of Human Colorectal Polyps. *Comput Med Imaging Graph*, **10**, 1-10.
- Fu JJ, Yu YW, Lin HM, Chai JW, Chen CCC (2014). Feature extraction and pattern classification of colorectal polyps in colonoscopic imaging. *Comput Med Imaging Graph*, **38**, 267-75.
- Gasparovica M, Aleksejeva L (2011). Using fuzzy unordered rule induction algorithm for cancer data classification. *Breast Cancer*, **13**, 1-7.
- Hühn J, Hüllermeier E (2009). FURIA: An algorithm for unordered fuzzy rule induction. *Data Min Knowl Discov*, **19**, 293-319.
- Jabid T, Kabir MH, Chae O (2010). Robust facial expression recognition based on local directional pattern. *ETRI J Journal*, **32**, 784-94
- Kaur G, Chhabra A (2014). Improved J48 classification algorithm for the prediction of diabetes. *Int j comput appl technol*, **98**, 7-13.
- Li J, Tan J, Martz FA, Heymann H (1999). Image texture features as indicators of beef tenderness. *Meat Sci*, **53**, 17-22.
- Liu J, McCloskey S, Liu Y (2012). Local expert forest of score fusion for video event classification. In computer vision-ECCV 2012. Springer Berlin Heidelberg, pp. 397-410.
- Mamonov AV, Figueiredo IN, Figueiredo PN, Tsai YHR (2014). Automated polyp detection in colon capsule endoscopy. *IEEE Trans Med Imaging*, **33**, 1488-1502.
- Manivannan S (2015). Visual feature learning with application to medical image classification (Doctoral dissertation, University of Dundee). pp 1-154
- Manivannan S, Trucco E (2015). Learning discriminative local features from image-level labelled data for colonoscopy image classification. In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI) pp 420-23.
- Muthukudage J, Oh J, Tavanapong W, Wong J, De Groen PC (2011). Color based stool region detection in colonoscopy videos for quality measurements. In Pacific-Rim Symposium on Image and Video Technology. Springer Berlin Heidelberg pp 61-72.
- Riley SA (2008). Colonoscopic polypectomy and endoscopic mucosal resection: a practical guide. *Br Soc Gastroenterol*, **8**, 1-22.
- Tajbakhsh N, Gurudu SR, Liang J (2016). Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Trans Med Imaging*, **35**, 630-44.
- Zhang Y, Hao P, Wang M, Guo C (2011). Research on classification and detection of colon cancer's gene expression profiles. *Biomed Res Int*, **6**, 2792-800.