Validation Techniques for Relating Ensemble Methods in Polyp Detection

Nagesh B S, N P Kavya

Abstract: Endoscopy is one of the most efficient colon screening technique through which the polyps are identified and treated. This manual process of identifying the polyp has chance of missing some polyps while diagnosis. To overcome this an efficient computer aided detection technique need to designed, there are several computing techniques and algorithms available, this research article tries compare different techniques used in polyp detection and also proposes several performance evaluation metrics which can be used to find the proficiency of methods in identifying polyps in endoscopy videos. The article presents various identification methods, evaluation methods include various parameters like performance metrics. The proposed ensemble approaches are discussed in polyp detection mechanism.

Keywords: Ensemble, Endoscopy, polyp detection. Video mining

I. INTRODUCTION

The evaluation study tries to compare the performance of several polyp identification techniques with respect to clinical as well as technical context. The work tries to covers efficiency of convolutional neural network (CNN) technique and ensemble approaches in identifying polyps.

Clinical background:

Colorectal and gastrointestinal cancer is one of the dangerous disease spreading across and causing fatalities across the world.
The disease has a platform derived from gastrointestinal polyps, which are abnormal gastrointestinal tissue growths that develop on mucous membrane of gastro intestinal tract. The polyps are initially not malignant but, they may become malignant over a period of time and range to neighbouring tissues and organs like organs like liver, lungs etc.

Technical Background:

The continuous improvement of the skills of physicians is through well versed training and regular practice, the outcome of endoscopy has been advanced due latest technological advances.
The technological shifts are categorized into two groups
a) Device improvement
b) Computer support systems development

The system improvements are desired as mentioned below Enhancing image resolution, wide angled cameras covering colon wall surface, Zooming and magnifying techniques and autofluorescence imaging.

II. REVIEW ON COMPUTER AIDED POLYP DETECTION METHODS

Several historical advances in the field of polyp detection are available for researches to proceed further. The detection techniques are be classified into three types. **Custom-made techniques, End to End Methods** and **Ensemble techniques.**

Custom techniques

The popular techniques to identify polyps are based on primitive image processing techniques which are used to find borders of polyps. Only few are discussed in Table 1.

Table 1: custom methods

Authors	Methods
Iwahori .[2]	Hessian filters used to enhance features of image and Primitive image processing techniques
Mojtaba Akbari[1]	Convolutional neural network and Otsu thresholding.
Silva . [3]	Hough transform technique to identify boundaries of polyps
Kang . [5]	analyses curvatures of boundaries detected using any of the above methods
Kang and Doraiswami [6]	Searching ellipsoidal shapes typically associated with polyps.
Hwang et al. [7]	Approach includes curvature analysis and shape fitting techniques

End to End Methods

Table 2 tries to explore few end to end methods proposed by several authors, where the works utilize texture and colour information in polyp detection

Validation Techniques for Relating Ensemble Methods in Polyp Detection

Table 2: End to end methods

Authors	Methods
Karkanis et al. [8]	Usage of colour wavelets
Ameling et al. [9]	usage of co-occurrence matrices
Gross et al. [10]	Usage of local binary patterns.
Angermann et al. [11]	Active learning methodologies are used to strengthen the trade-off among performance and computation.
Park and Sargent [12]	deep learning techniques in polyp detection

Ensemble Approaches

Table 3 shows few ensemble approaches which associates numerous procedures for detecting polyps.

Table 3: Ensemble approaches

Authors	Methods
Tajbakhsh . [13]	Associates feature extraction long with edge detection to increase accurateness in polyp detection.
Yoon and Bae.[14]	Work is grounded on extreme knowledge along with the feature learning
Silva. [4],	Utilizes handmade features to filter image regions which are not informative.

III. ENSEMBLE METHODS FOR POLYP IDENTIFICATION

Method 1: The approach follows full localization pattern, which involves two portions, detection and localization. The detection considers two groups of images, first group comprising polyps and the second group which does not have polyps. The two group of images are considered for training. The Global image features like texture, shape etc. are considered and the similarity values among input image frame and the training images along with their results are analysed, the input frame is of polyp class or non-polyp class is concluded by the detection module.

The proposed localization scheme follows implementation as pre-processing sequence by converting frames from RGB to YCbCr colour space, sub-images and borders removal, flare hiding and smoothing is done by low pass filtering and polyps shape is used to know the position, i.e. Predicting polyps by oval shaped regions offering local features that distinguish polyps from neighbouring tissues. The energy map having maximum values which are calculated by elliptical form of polyps' common presence decides polyp's position.

Method 2: The proposed approach follows region of interest, where Initially, Regions of Interests are analysed so that it may comprise a polyp. This is done using size and shape features [5]. The scanning of images are done first and presence of circular or elliptical forms in polyps promotes to use Hough transform as first filtering phase. After region of interest are identified, the second phase considers texture to remove the region of interests where actual polyps are not present. To accomplish this, the boosting based learning procedure by means of texture features are calculated after co-occurrence matrices.

Method 3: This method involves of two phases. The first phase, geometric features are used to generate a set polyp candidates. i.e. From the endoscopic frame, a basic edge pixels set is obtained. Classification and feature extraction methods are used to refine the derived edge map. The edge classification has an intention of the module is to eliminate pixels of non polyps residing in boundaries by means of early edge map. The voting system uses the retained edges geometry which focuses polyp contenders, i.e. Objects with curve boundary in the retained edge maps. A bounding case for every produced contender grounded on produced map of voting , which is estimated by voting method. Followed by a second phase that uses an ensemble approach, where convolution neural networks, where individual kind of features are considered by each of them. That are considered to every single candidate bounding box. As last phase, outcomes of the CNNs are used to calculate average for producing confidence value of a taken candidate of polyp.

IV. DEFINITIONS AND PERFORMANCE METRICS

a. **Definitions**: The definition of few abbreviations used in performance metrics are defined below

Polyp detection: is defined as the skilled procedure to decide the existence of polyp endoscopy frame.

Polyp localization: is the ability to give the position of a polyp in an image. An efficient polyp detection technique must pick frames comprising polyps and the technique should ignore remaining. The technique should specify locations of every polyps existing in the frame.

True Positive (TP): indicates the detections output is inside the polyp.

False Positive (FP): indicates the detections that shows outside polyp.

False Negative (FN): indicates the absence of indication in pictures with a polyp.

True negative (TN): is indicated when the technique does not offer some output for image without polyps.

False Positive (FP): indicates detections provided for image frames without a polyp.

b. **Performance Metrics**

The table 4 below shows various performance metrics that are used for performance evaluation.

The performance evaluation decides the techniques, which must considered while detecting the polyps. There are several metrics proposed and available to evaluate the techniques. Only few of the metrics are shown below.

Table 4: Metrics

$Precision = TP/(TP-FP)$
$Recall = TP/(TP-FN)$
$Specificity = TN/(TN+FP)$
$F1\ Score = 2*(Recall * Precision) / (Recall + Precision)$
$F2\ Score = 5*(Recall * Precision) / 4(Recall + Precision)$

V. RESULTS AND DISCUSSIONS

Gastrointestinal dataset comprises ten video sequences with total 7894 images. The video sequences are collected by gastroenterologists in gastro intestinal examinations using an Olympus NBI endoscope.

Images with at least one polyps are constructed to cover diverse polyp presences as likely. Images are marked by skilled endoscopy technicians.

The experts are questioned to sketch margins of polyps existing present in frame. The binary mask is generated using the boundaries which represent real polyp area in the image, which is used for authentication.

The below Figure 1 shows outline and Figure 2 shows masking.
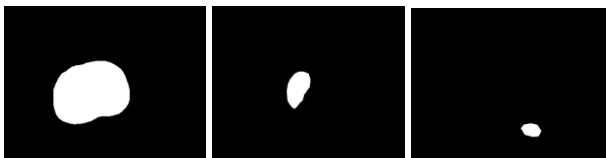

Figure 1: polyp outline shown.

Figure 2: Masking polyps region.

F1-score is used to compare various approaches, because F1 score offers a stability among missed polyps and wrong highlights. The statistical approach of F1 score is done on only videos having polyp and possibly non-polyp image frames, because in still image analysis the number of samples are not large enough to give statistically applicable inferences and the investigation in frames that does not have polyps may lead F1 score as zero.

VI. VALIDATION TECHNIQUES

The analysis aims to discover proficiencies of a polyp discovery technique. The validation techniques are applied on images that contain minimum on polyp and some images are taken where polyps are not clear or other features are abstracting the polyps like vessels and other folds.
A well sophisticated graphical user interface can be used by experts to tag independently each frame of the considered database. After deriving annotations, performance analysis is categorised into two major groups: 1) Images with polyps

2) Without polyps. The study tries examine differences among performance for images with and without considering F1-score, Recall and Precision.

VII. CONCLUSION

The paper proposes a technique to conduct validation study for comparing the performance of various polyp discovery techniques. The common and proved performance metric can be considered for validation and the open source database available for research purpose can be used. The work displays Clinical and technical tests. The proposed methods try to combine different techniques of identifying polyps and comparing the efficiency of methods. The various ensemble approaches used for detecting polyps are discussed and their performance comparison technique by considering various performance metrics like Recall, specificity, precision, F1 Score and F2 score are discussed in the paper. The article also proposes a technique of building a test data by a clinical experts utilising the available image processing a graphical user interface applications. The research article aims to contribute towards building the computer aided software which assist the practitioners to detect the polyps and also identify the malignant polyps at early stage.

REFERENCES

1. Mojtaba Akbari "Polyp Segmentation in Colonoscopy Images Using Fully Convolutional Network " arXiv:1802.00368
2. I. Yuji et al., "Automatic polyp detection in endoscope images using a Hessian filter," in Proc. MVA, 2013, pp. 21–24, 2013.
3. J. Silva, A. Histace, O. Romain, X. Dray, and B. Granado, "Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer," Int. J. Comput. Assist. Radiol. Surgery, vol. 9, no. 2, pp. 283–293, 2014.
4. H. Zhu, Y. Fan, and Z. Liang, "Improved curvature estimation for shape analysis in computer-aided detection of colonic polyps," in Proc. Int. Workshop Comput. Challenges Clin. Opportunities Virtual Colonoscopy Abdominal Imag., 2010, pp. 9–14.
5. J. Kang and R. Doraiswami, "Real-time image processing system for endoscopic applications," in Proc. IEEE Can. Conf. Electr. Comput. Eng. (CCECE), vol. 3. May 2003, pp. 1469–1472.
6. S. Hwang, J. Oh, W. Tavanapong, J. Wong, and P. C. de Groen, "Polyp detection in colonoscopy video using elliptical shape feature," in Proc. IEEE Int. Conf. Image Process., vol. 2. Oct. 2007, pp. II-465–II-468.
7. S.A.Karkanis, D.K.Iakovidis, D.E.Maroulis, D.A.Karras, and M. Tzivras, "Computer-aided tumor detection in endoscopic video using color wavelet features," IEEE Trans. Inf. Technol. Biomed., vol. 7, no. 3, pp. 141–152, Sep. 2003.
8. S Ameling, "Texture based polyp detection in colonoscopy," in Bildverarbeitung für die Medizin. Berlin, Germany: Springer, 2009, pp. 346–350.
9. S. Gross et al., "A comparison of blood vessel features and local binary patterns for colorectal polyp classification," Proc. SPIE, vol. 6918, Feb. 2009.
10. Q. Angermann, A. Histace, and O. Romain, "Active learning for real time detection of polyps in videocolonoscopy," Procedia Comput. Sci., vol. 90, pp. 182–187, Jul. 2016.
11. S. Y. Park and D. Sargent, "Colonoscopic polyp detection using convolutional neural networks," Proc. SPIE, vol. 9785, Mar. 2016.
12. N. Tajbakhsh, S. R. Gurudu, and J. Liang, "Automated polyp detection in colonoscopy videos using shape and context information," IEEE Trans. Med. Imag., vol. 35, no. 2, pp. 630–644, Feb. 2016.
13. S.-H. Bae and K.-J. Yoon, "Polyp detection via imbalanced learning and discriminative feature learning," IEEE Trans. Med. Imag., vol. 34, no. 11, pp. 2379–2393, Nov. 2015.

AUTHORS PROFILE

Nagesh B S holds a Master degree in computer applications from Kuvempu University and pursuing research under VTU, Belagavi, at R & D Centre, Department of CSE, RNSIT, Bengaluru has published 4 research papers in reputed international journals. The main areas of research are Data Sciences, Video Mining, Machine Learning, and Artificial Intelligence, etc.

Dr. N P Kavya holds Bachelor of Engineering in Computer Science and Engg. along with MS in software systems and Ph.D. in computer science form VTU Belagavi . She has a vast experience of 24 years in the field education and research. She is currently a Professor in Computer science and Engg., Department , RNSIT, Bengaluru. She has published around 90 research papers in reputed international journals including IEEE, Elsevier, Springer (SCI and Web of Science). Has 94+ citations in Google scholar as on Oct 2019. Her main areas of expertise are expertise are Machine Learning, Artificial Intelligence, Big Data analytics etc.