HAL open science

Estimating Polyp Size From a Single Colonoscopy Image Using a Shape-From-Shading Model

Josué Ruano, Diego Bravo, Diana Giraldo, Martín Gómez, Fabio González, Antoine Manzanera, Eduardo Romero

To cite this version:

Josué Ruano, Diego Bravo, Diana Giraldo, Martín Gómez, Fabio González, et al.. Estimating Polyp Size From a Single Colonoscopy Image Using a Shape-From-Shading Model. 2024 IEEE International Symposium on Biomedical Imaging (ISBI), May 2024, Athens, France. pp.1-5, ⟨10.1109/ISBI56570.2024.10635358⟩. ⟨hal-04879338⟩

HAL Id: hal-04879338

https://hal.science/hal-04879338v1

Submitted on 10 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Autorisation HAL

HAL Authorization

ESTIMATING POLYP SIZE FROM A SINGLE COLONOSCOPY IMAGE USING A SHAPE-FROM-SHADING MODEL

Josué Ruano[a], Diego Bravo[a], Diana Giraldo[a,b] Martín Gómez[c], Fabio A. González[d], Antoine Manzanera[e] and Eduardo Romero[a]

[a] Computer Imaging and Medical Applications Laboratory (CIM@LAB)
[b] imec-Vision Lab, Department of Physics, University of Antwerp, Antwerp, Belgium
[c] Hospital Universitario Nacional de Colombia, Unidad de Gastroenterología, Bogotá, Colombia
[d] Machine Learning, Perception and Discovery Lab (MindLab)
[e] Unité d'Informatique et d'Ingénierie des Systémes, ENSTA-Institut Polytechnique de Paris, France
[a,c,d] Universidad Nacional de Colombia, Bogotá, Colombia

ABSTRACT

Colonoscopy (CO) is the most useful procedure to estimate the polyp size as part of surveillance and therapeutic management to prevent Colorectal cancer. Studies have reported a high rate of misestimated lesions by experts, reaching a relative accuracy of 67%. This work presents a method to estimate polyp size from a raw CO frame. A shape-from-shading model trained with synthetic images estimates a depth map to reconstruct the three-dimensional ($3d$) colon structure. Polyp is segmented in RGB image with a U-Net, to compute diameter in pixels. This diameter is projected onto the $3d$ surface to compute polyp size in millimeters. The method obtained a mean absolute error of $2.16\ mm$ and relative accuracy of 88.17% in 2 802 synthetic images, and $2.06\ mm$ in 100 real images. In a binary classification task ($\leq 10\ mm$ or $\geq 10\ mm$), the method achieved macro F1 scores of 89% and 72% in the synthetic and real databases respectively.

***Index Terms*—** Colonoscopy, Synthetic database, Depth estimation, polyp size estimation

1. INTRODUCTION

Colorectal cancer (CRC) represents a significant global health issue being the second most mortal cancer worldwide in 2020 [1]. The CRC is usually manifested as polypoid (elevated) or non-polypoid (flat) masses growing from the intestinal mucosa, which will eventually evolve to adenocarcinoma or hyperplasia[2, 3]. Colonoscopy (CO) is the most useful procedure for diagnosing CRC, whereby a gastroenterologist visualizes the colon wall searching polyps [4]. The size of any detected polyp plays a major role in deciding the surveillance intervals, i.e. patients with polyps larger than $10\ mm$ must have CO surveillance within 3 years [5]. Also, particular therapeutic strategy is defined based on polyp size [6] as follows: diminutive and small ($\leq 10\ mm$) lesions may be immediately and completely removed (polypectomy) using cold biopsy forceps or cold snare: large polyps ($> 10\ mm$) are treated with hot snare, whereas huge polyps ($> 20\ mm$) or large number of polyps require a colectomy or proctocolectomy [7, 8]. Usually, experts estimate polyp size by visual examination and less frequently, they place an instrument with known sized (e.g. biopsy forceps or linear probes) near the lesion and compare their diameters [9]. However, different studies have reported a high rate of misestimated lesions reaching a relative accuracy of 67%, overestimates in most cases, with considerable inter and intra-observer variability [6, 10]. Nevertheless, the real size of polyps in-vivo is not possible to obtain during a CO since polyps lose mass immediately after resection, being smaller in ex-vivo estimations [11]. Hence, most clinical trials set as ground truth the size estimation of the most experienced specialist using a linear probe [10]. Recently, a laser probe called virtual scale endoscopy has reported fewer misestimated lesions, but this tool has not been incorporated into clinical practice yet [6, 9]. Nevertheless, estimating polyp size during a routine CO needs to be improved or supported with second readers, and other workflows must be proposed to validate the performance of current and new strategies.

Some computational strategies have been proposed to estimate polyp size or classify the lesions in different size ranges. For instance, Kwak et al. [12] approximate the diameter of a polyp propagating a reference distance computed between the main vessels of the mucosa. First, a W-Net, an ensemble of two consecutive U-Nets trained with four retinal image datasets that include reference distances, is used to segment the colonic vessels. Then, the reference distance of main vessels in a CO image is determined by searching a similar configuration of vessels in the retinal image dataset. In a set of 30 polyps, this work obtained a Concordance Correlation Coefficient of 0.96 between the estimated sizes and the lesion measurements after resection. On the other hand, the work presented by Abdelrahim et al. [13] evaluates two approaches to classify polyp size as $\leq 5\ mm$ or $\geq 5\ mm$: automated sizing using Structure from Motion (SfM) in 22 premeasured polyps fixed in ex-vivo pig colon and a convolutional neural network (CNN), i.e. VGG-16, trained with 301 human polyps and tested with 10. The SfM system showed a superior polyp size classification accuracy of 85.2% compared to 59.5% obtained by experts in ex-vivo models, while the CNN achieved an accuracy of 80% using the average estimation of three expert gastroenterologists as ground truth for training and testing the network. Another approach presented by Itoh et al. [14] constructed a simplified virtual colon model from nine CT scans, i.e., without tissue appearance nor specular reflections, to emulate virtual CO explorations and obtain synthetic RGB images with depth maps. Then, a Cycle-Generative Adversarial Network (Cycle-GAN) learns the colon shape from synthetic images to estimate a depth map from real RGB images. Simultaneously, a You-Only-Look-Once network

Further author information: (Send correspondence to Dr. Eduardo Romero). Eduardo Romero: E-mail: edromero@unal.edu.co, Telephone: +57 (1) 3 16 54 91.

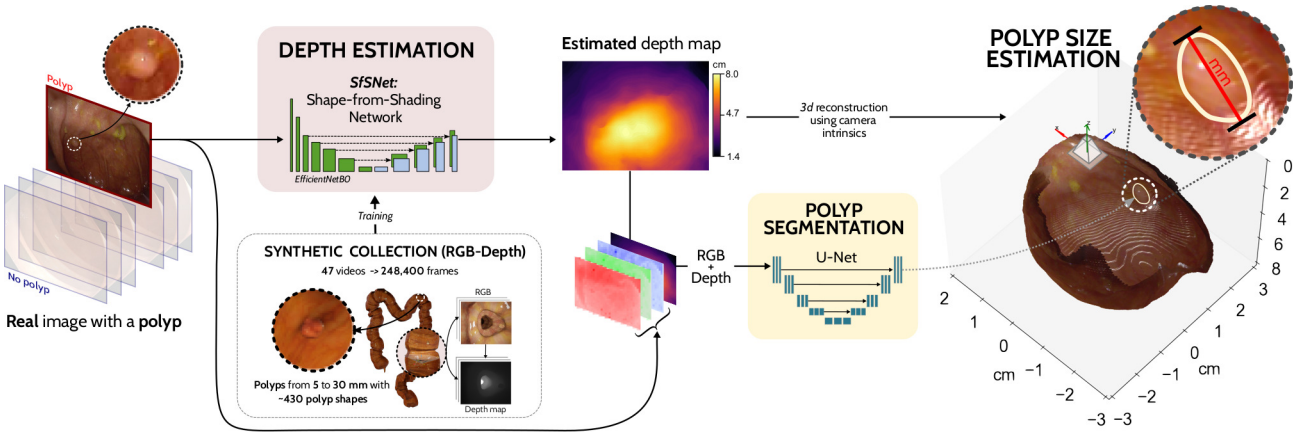

Fig. 1. Pipeline of the proposed approach is composed of 5 steps. First, a frame with the presence of a polyp is selected by an expert during a colonoscopy procedure. Second, a depth map is estimated using a shape-from-shading model which was trained with a synthetic collection (see Section 2.1). Third, segmentation of the polyp is achieved using a U-Net trained with RGB+Depth images. The diameter is computed in two reference points from the polyp segmentation. Fourth, the estimated depth map with intrinsic camera parameters serves to reconstruct the $3d$ shape of the colon wall. Fifth, the Euclidean distance between the $3d$ position of the two reference points establishes the polyp size.

(YOLOv3) localizes a lesion in a bounding box and then uses it to extract depth information. The extracted depth region is used to feed a CNN which estimates the polyp size using 30 subjects for training. Performance was assessed in nine CO images from different subjects obtaining a mean absolute error of $1.36\,mm$ w.r.t. an expert visual estimation. However, these strategies are limited in a) validating their performance since variability and number of lesion sizes were scarce [12, 13, 14], b) they required different image modalities to train their methods [14, 12], c) a rigid colon wall was assumed to estimate size from motion [13], d) the methods ought specific structures present in the images with a particular position of CO device [12]. In addition, these works did not release the source code of their methods and used private databases, so it is not possible to compare with them.

A main contribution of this work is the automatic estimation of polyp size using a single image based on a shape-from-shading model learned in synthetic CO images. The size estimated for this approach is used to evaluate a binary classification task, i.e. whether a polyp is either $\leq 10mm$ or $> 10mm$. In clinical practice, this task is very useful for determining whether a patient requires a CO in the upcoming three years or guiding whether a polyp should be removed with cold biopsy forceps or hot snare. Additionally, we provided a synthetic colonoscopy collection with depth maps, segmentation of lesions, and size annotations.

2. METHODOLOGY

The polyp size is defined as its diameter D, which is the largest Euclidean distance between the lesion's contour points in the three-dimensional ($3d$) reconstruction of the colon wall, estimated in millimeters. First, $3d$ reconstruction of the colon wall is obtained by estimating a depth map from a single image using a shape-from-shading neural model trained with a collection of synthetic CO images. Second, the polyp is segmented using a U-Net, which is fed with RGB images enriched with the estimated depth map as an additional channel. The polyp segmentation is used to compute the two furthest pixels from the contour, which are back-projected onto the $3d$ reconstruction where the polyp size D is computed. Finally, the estimated size is used to assess a task of clinical relevance: a binary task between $\leq$ 10 mm or $>$ 10 mm to support endoscopic surveillance and therapeutic management. Performance of these steps is two-fold validated: with synthetic and real CO collections. A pipeline of the method is presented in Fig 1 and described hereafter.

2.1. Estimating colon depth based on Shape-from-Shading.

The depth map of a CO frame associates to each visible colon wall point the distance to the focal plane, therefore this map captures the $3d$ colon shape. Colon depth may be estimated by learning the variations of shading with respect to both the distance and orientation of the camera, a technique known as shape-from-shading. As this estimation is an ill-posed problem by the multiple approximations that may co-exist of a $3d$-scene from a $2d$-image, different constraints regularize the solution, such as: a) the camera obeys a pinhole model with constant focal length, b) the light source's position is close to the camera's optical center, and c) the mucosa reflection is mostly Lambertian. These constraints facilitate capturing the colon shape as a function of a light source which dims out as the distance increases by a simple intensity attenuation model, an intensity inverse square law. In these conditions, a deep network can learn such model, i.e., a depth map may be reliably obtained given a single RGB image by a pixel-level regression task. Such regression is performed by supervised learning of the Shape-from-shading network ($SfSNet$) [15] that consists of an encoder-decoder layout, i.e., the encoder is the EfficientNetB0 architecture, the decoder includes four up-sampling blocks, and between the encoder-decoder, long skip connections are incorporated [15]. A customized loss function $\mathbf{L}$ balances a smooth reconstruction of the depth map with the L_z term, while it also preserves the high-frequency details or borders with L_e and L_c, respectively related to the edge (gradient ∇) and curvature (Hessian matrix H) of the depth maps, defined as:

$$\mathbf{L}(d, \hat{d}) = w_1 L_z(d, \hat{d}) + w_2 L_e(\nabla(d), \nabla(\hat{d})) + w_3 L_c(H(d), H(\hat{d}))$$

where $\hat{d}$ is the estimated depth, d is the ground truth. $w_1 = 0.1$, $w_2 = 0.3$ and $w_3 = 0.6$ were set through ablation study. All terms are defined in the $\mathcal{L}_1$ norm.

2.2. Three-dimensional reconstruction of colon wall

$3d$ reconstruction of the colon wall is used to obtain the $3d$ positions of each pixel in the RGB image, allowing estimation of any distance or size, such as polyp diameter. This reconstruction is achieved by using the estimated depth map of an RGB image to back-project from the $2d$ image coordinate system to the $3d$ world coordinate system. Therefore, a pixel at $2d$ coordinates (u, v) with depth d is mapped to the (x, y, z) $3d$ world coordinates in millimeters using the intrinsic parameters of the camera, as follows $\begin{pmatrix} x & y & z \end{pmatrix}^T = dK^{-1} \begin{pmatrix} u & v & 1 \end{pmatrix}^T$ with:

$$K^{-1} = \begin{pmatrix} 1/f & 0 & -c_x/f \\ 0 & 1/f & -c_y/f \\ 0 & 0 & s \end{pmatrix}$$

Where f, the focal distance, s a scale factor, and (c_x, c_y) the focal center. The result of this transformation is a point cloud forming a partial $3d$ reconstruction of the colon wall.

2.3. Polyp segmentation strategy

The polyp contour is essential to find the two reference points in the RGB image that form the diameter or size of a polyp. A fully automatic approach is proposed using a classic but still relevant segmentation method, a U-Net, which serves to segment the polyp contour. The first layer of this network is adapted to support RGB-Depth images, i.e. the estimated depth map is concatenated as an additional channel to the RGB image, with the aim of improving the segmentation task using $3d$ shape information. Lastly, a specialist confirms the network's segmentation and makes any necessary corrections since a misestimation may occur, e.g. in tiny sessile polyps with poorly defined edges.

2.4. Polyp size estimation

The two reference points calculated in the polyp segmentation step are projected to the $3d$ reconstruction of the colon wall. After, the Euclidean distance is computed between the points, representing the largest diameter $\hat{d}$ or size of a polyp. This estimation is obtained in millimeters.

2.5. Datasets

- ***SUN Database:*** This study [16] comprised 100 subjects who underwent CO procedures and were diagnosed with polyps at Showa University Northern Yokohama Hospital. The detected polyps have an average size of $6.2\,mm \pm 3.72\,mm$ and most of them are low-grade adenomas according to pathological analysis. In terms of shape, 59 are polypoid lesions (Ip and Is) and the rest are non-polypoid of type IIa according to Paris classification [2]. This collection provides 49 136 raw images with different perspectives of the polyps, and every frame was annotated by an expert gastroenterologist with a bounding box. Also, polyp size was estimated by the expert, but this estimation does not correspond to all frames. For this reason, an external expert selects a frame that better correlates with the provided size and manually delineates the lesion. 100 raw CO images with different polyps were taken from this database and these are used only for evaluating the complete approach.
- ***Synthetic collection:*** This database [15] provides 47 synthetic CO videos (248 400 frames) generated at a spatial resolution of $1\,280 \times 1\,080$ pixels and 15 frames per second. The videos were rendered in a virtual environment in which $3d$ colon models were constructed using averaged colon measurements from 500 adults. These models contain approximately 430 synthetic polyps with sizes from 3 to 30 millimeters. In this work, the depth estimation network is trained and tested with this collection since depth maps are provided. In addition, a segmented ground truth mask and polyp size are provided for 2 802 frames in the testing set, frames used to challenge the complete approach for polyp size estimation.
- Two collections with pixel-wise image segmentation of polyps are used to train a U-net. First, the ***Kvasir-SEG*** dataset which has 1000 polyp images with a resolution of 332×487 to $1\,920 \times 1\,072$ pixels. Each image contains at least one polyp with a ground truth polyp mask [17]. Second, the ***CVC-ClinicDB*** has 612 images of 384×288 resolution. The frames are extracted from 25 different studies with 29 sequences. Each image contains one polyp with its ground truth mask [18].

3. EVALUATION AND RESULTS

Performance of the proposed approach was two-fold assessed:

1. Polyp size estimation: This task is evaluated with the Relative Accuracy (relACC), Mean Absolute Error (MAE), and Mean Bias Error (MBE).

2. Binary classification between $polyp \leq 10\ mm$ **and** $polyp > 10\ mm$ **classes:** F1 score (F1) and MAE are computed for each class and macro F1 score (macro-F1) for the complete task.

Both tasks are challenged in synthetic and real CO databases. For synthetic collection, the method is assessed by comparing the estimated size $\widehat{D}$ with the ground truth size D for each frame, being D a precise measurement obtained from ground-truth depth maps. For the real collection, SUN Database, the estimated size $\widehat{D}$ is compared with the estimation performed by an expert gastroenterologist $\widehat{D}_{expert}$ to assess the similarity between both estimations.

3.1. Experimental setup

The three intermediate tasks were evaluated and tested as follows:

- ***Depth estimation:*** the SfSNet was trained and tested using the synthetic collection split into 65% for training, 15% for validating, and 20% for testing. During the training, five common hyper-parameters of the network were optimized in 75 trials. This network reached an accuracy of 95.65% and a root-mean-square error of $4.51\ mm$.
- ***Polyp segmentation:*** U-net was trained with CVC-ClinicDB and Kvasir-SEG collections, with a total of 1 449 images with the presence of polyps (90% for training and 10% for validation). Hyper-parameter optimization was performed during training using a grid-search strategy for initial learning rate and batch size. Performance of this network in the synthetic database was an average DICE-score of 0.89 ± 0.09, and in the SUNDatabase was 0.71 ± 0.27. An expert verified whether the segmentation provided by the method was acceptable to estimate polyp size. The inaccurate segmentations were replaced using manual segmentation by the expert. For the synthetic database, 217 of 2 800 polyps segmentations were corrected, and for the SUN Database, were 16 of 100.
- ***Intrinsic camera parameters:*** the focal length f, focal axis points (c_x, c_y) and the scale s are used to reconstruct the $3d$ structure of colon. The virtual camera used to generate the synthetic collection was set to a $f = 1.755\ mm$ (448.13 pixels) corresponding to a wide-angle camera of $110°$, $(c_x, c_y) = (0, 0)$ and $s = 1$. Since the SUN database did not provide these intrinsics of the CO device, the f and s employed in the virtual

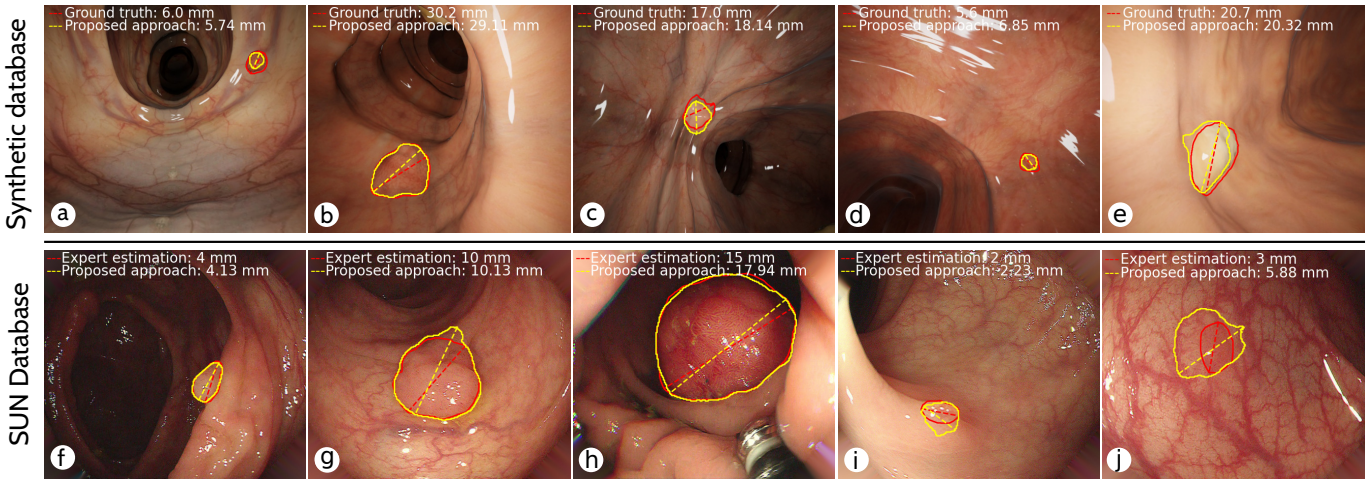

Fig. 2. Size estimations of the proposed approach on 10 images from SUN Database.

camera had to be adjusted for the SUN database in an optimization task. First, this database was split randomly into 90% for testing the complete workflow and 10% for calibrating f and s. With this subset, the polyp size estimation error was minimized while varying the f and s during 100 trials. The result of this process was $f = 3.159\ mm$ and $s = 4.1$.

3.2. Polyp size estimation

The proposed approach achieved a reliable polyp size estimation in the synthetic database with MAE of $2.16\,mm$ and relAcc of 88.17% (see Table 1). For SUN Database, the estimated size in these real images was very similar w.r.t. estimation of the expert obtaining an MAE of $2.06\ mm$. MBE of $-0.46\ mm$ indicated that the method tended to slightly underestimate size in the synthetic set, and for SUN Database, to overestimate minimally with MBE of $0.03\ mm$.

Database	MAE (mm)	MBE (mm)	relAcc (%)
SyntheticDb	2.16	-0.46	88.17
SUNDb	2.06	0.03	63.07

Table 1. Evaluation metrics computed in the Synthetic and SUN Database for polyp size estimation task.

Figure 2 illustrates the size estimation and segmentation results in the Synthetic and SUN databases. The red contour stands for the ground truth delineation and the yellow contour is the estimated segmentation, showing a reliable overlap as shown from *a* to *i*. The segmentation method fails in lesions with total or partially fuzzy borders (*j*). Considering that the variability of lesions in the SUN Database is high in terms of size, appearance, and the depth where these are located, the estimated sizes by the method were outstanding. Also, the estimations are very similar compared to the specialist with several years of expertise.

3.3. Binary classification between $polyp \leq 10\ mm$ and $polyp > 10\ mm$ classes

The estimated size served to classify a polyp in two classes as follows: polyps smaller than or equal to $10\ mm$ in size (class 1) or polyps larger than $10\ mm$ in size (class 2). Replicating this procedure with ground truth size in the synthetic database, a confusion matrix was computed comparing labels established by the method and ground truth labels, in which an outstanding performance (see Table 2) was reached with a macro F1 score of 89%. MAE for both classes was less than $2.3\,mm$. For the SUN Database, confusion matrix was built with labels established by the method and those computed using the estimated size by the expert. Classification made by the expert and the method also was similar, producing a macro F1 score of 72%. MAE in class $polyp > 10\ mm$ was high because the estimated depth in CO frames with large polyps is less precise.

Database	Class	N. frames	MAE(mm)	F1(%)
SyntheticDb	$\leq 10\ mm$	408	1.26	81
	$> 10\ mm$	2 394	2.31	94
SUNDb	$\leq 10\ mm$	76	1.78	92
	$> 10\ mm$	14	4.28	52

Table 2. Evaluation metrics computed in the Synthetic and SUN Database for the binary classifications task.

4. CONCLUSIONS AND DISCUSSION

This work developed a method for estimating polyp size in millimeters from a single CO image. The proposed technique uses shading in colon wall to estimate $3d$ colon structure in a world coordinate system. The results reveal that the estimations by the proposed approach were equivalent to those obtained by expert gastroenterologists. With a relative accuracy above 80% obtained by the proposed method, the estimated size was useful to classify these lesions as smaller or larger than $10\ mm$. This categorization is extremely important in clinical practice since it is utilized to decide CO surveillance as well as the surgical method to remove the lesion. In comparison to state-of-the-art approaches, this approach is closer to clinical application since it does not require elaborate protocols and is capable of handling artifacts without requiring preprocessing. Source code and subset of the synthetic database which contains the polyp segmentations with their sizes are publicly available in `https://github.com/jaruanob/PolypSizeEstimation_SfSNet.git` so that new works can be compared with this approach. Future work includes increasing the amount of validation data as well as the number of experts with whom the method is compared.

5. COMPLIANCE WITH ETHICAL STANDARDS

This work was conducted retrospectively using human subject data made available in open access. Ethical approval was not required as confirmed by the license attached with the open-access data.

6. ACKNOWLEDGMENT

- This work was supported in part by the project with code 110192092354 and entitled "Program for the Early Detection of Premalignant Lesions and Gastric Cancer in urban, rural and dispersed areas in the Department of Nariño" of call No. 920 of 2022 of MinCiencias.
- This work was partially supported by project BPIN 2019000100060 "Implementation of a Network for Research, Technological Development and Innovation in Digital Pathology (RedPat) supported by Industry 4.0 technologies" from FCTeI of SGR resources, which was approved by OCAD of FCTeI and MinCiencias.

7. REFERENCES

[1] Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] H Inoue, H Kashida, S Kudo, M Sasako, T Shimoda, H Watanabe, S Yoshida, M Guelrud, CJ Lightdale, K Wang, et al., "The paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach and colon.," *Gastrointest. Endoscopy*, vol. 58, pp. S3–S43, 2003.

[3] Noam Shussman and Steven D. Wexner, "Colorectal polyps and polyposis syndromes," *Gastroenterology Report*, vol. 2, no. 1, pp. 1–15, 01 2014.

[4] Colin J Rees, Roisin Bevan, Katharina Zimmermann-Fraedrich, Matthew D Rutter, Douglas Rex, Evelien Dekker, Thierry Ponchon, Michael Bretthauer, Jaroslaw Regula, Brian Saunders, Cesare Hassan, Michael J Bourke, and Thomas Rösch, "Expert opinions and scientific evidence for colonoscopy key performance indicators," *Gut*, vol. 65, no. 12, pp. 2045–2060, 2016.

[5] Cesare Hassan, Giulio Antonelli, Jean-Marc Dumonceau, Jaroslaw Regula, Michael Bretthauer, Stanislas Chaussade, Evelien Dekker, Monika Ferlitsch, Antonio Gimeno-Garcia, Rodrigo Jover, et al., "Post-polypectomy colonoscopy surveillance: European society of gastrointestinal endoscopy (esge) guideline - update 2020," *Endoscopy*, vol. 52, pp. 687–700, 8 2020.

[6] Claire Haumesser, Melissa Zarandi-Nowroozi, Mahsa Taghiakbari, Roupen Djinbachian, Maria Abou Khalil, Sacha Sidani, Jeremy Liu Chen Kiow, Benoit Panzini, Ioana Popescu Crainic, and Daniel von Renteln, "Comparing size measurements of simulated colorectal polyp size and morphology groups when using a virtual scale endoscope or visual size estimation: Blinded randomized controlled trial," *Digestive Endoscopy*, vol. 35, pp. 638–644, 7 2023.

[7] Alan Moss and Kumanan Nalankilli, "Standardisation of polypectomy technique," *Best Practice & Research Clinical Gastroenterology*, vol. 31, pp. 447–453, 8 2017.

[8] Mike T. Wei, Christine Y. Louie, Yu Chen, Jennifer Y. Pan, Susan Y. Quan, Robert Wong, Ryanne Brown, Melissa Clark, Kristin Jensen, Hubert Lau, and Shai Friedland, "Randomized controlled trial investigating cold snare and forceps polypectomy among small polyps in rates of complete resection: The tinypolyp trial," *The American journal of gastroenterology*, vol. 117, pp. 1305–1310, 8 2022.

[9] Ryo Shimoda, Takashi Akutagawa, Michito Tomonaga, Tatsuro Murano, Kensuke Shinmura, Masato Yoshioka, Yuichi Teramura, Fumiaki Kiyomi, and Hiroaki Ikematsu, "Estimating colorectal polyp size with a virtual scale endoscope and visual estimation during colonoscopy: Prospective, preliminary comparison of accuracy," *Digestive Endoscopy*, vol. 34, pp. 1471–1477, 11 2022.

[10] Takahiro Utsumi, Takahiro Horimatsu, Yoshitaka Nishikawa, Akira Teramoto, Daizen Hirata, Mineo Iwatate, Shinwa Tanaka, Nobuaki Ikezawa, Masaya Esaki, Shozo Osera, et al., "Factors associated with inaccurate size estimation of colorectal polyps: A multicenter cross-sectional study," *Journal of Gastroenterology and Hepatology*, vol. 36, pp. 2224–2229, 8 2021.

[11] M. Izzy, M. Virk, A. Saund, J. Tejada, F. Kargoli, and S. Anand, "Accuracy of endoscopists' estimate of polyp size: A continuous dilemma," *World Journal of Gastrointestinal Endoscopy*, vol. 7, no. 8, pp. 824, 2015.

[12] Min Seob Kwak, Jae Myung Cha, Jung Won Jeon, Jin Young Yoon, and Jong Wook Park, "Artificial intelligence-based measurement outperforms current methods for colorectal polyp size measurement," *Digestive Endoscopy*, vol. 34, pp. 1188–1195, 9 2022.

[13] Mohamed Abdelrahim, Hiroyasu Saiga, Naoto Maeda, Ejaz Hossain, Hitoshi Ikeda, and Pradeep Bhandari, "Automated sizing of colorectal polyps using computer vision," *Gut*, vol. 71, pp. 7–9, 1 2022.

[14] Hayato Itoh, Masahiro Oda, Kai Jiang, Yuichi Mori, Masashi Misawa, Shin Ei Kudo, Kenichiro Imai, Sayo Ito, Kinichi Hotta, and Kensaku Mori, "Uncertainty meets 3d-spatial feature in colonoscopic polyp-size determination," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, vol. 10, pp. 289–298, 2022.

[15] Josué Ruano, Martín Gómez, Eduardo Romero, and Antoine Manzanera, "Leveraging a realistic synthetic database to learn shape-from-shading for estimating the colon depth in colonoscopy images," *arXiv preprint arXiv:2311.05021*, 2023.

[16] Masashi Misawa, Shin ei Kudo, Yuichi Mori, Kinichi Hotta, Kazuo Ohtsuka, Takahisa Matsuda, Shoichi Saito, Toyoki Kudo, Toshiyuki Baba, Fumio Ishida, Hayato Itoh, Masahiro Oda, and Kensaku Mori, "Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video)," *Gastrointestinal Endoscopy*, vol. 93, pp. 960–967.e3, 4 2021.

[17] Debesh Jha, Pia H. Smedsrud, Michael A. Riegler, Pål Halvorsen, Thomas de Lange, Dag Johansen, and Håvard D. Johansen, "Kvasir-seg: A segmented polyp dataset," *Lecture Notes in Computer Science*, vol. 11962 LNCS, pp. 451–462, 2020.

[18] J. Bernal, J. Sánchez, and F Vilariño, "Towards automatic polyp detection with a polyp appearance model," vol. 45, pp. 3166–3182, 2012.