



OPEN ACCESS

Automated sizing of colorectal polyps using computer vision

Mohamed Abdelrahim ¹, Hiroyasu Saiga,² Naoto Maeda,² Ejaz Hossain,¹ Hitoshi Ikeda,² Pradeep Bhandari¹

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2021-324510>).

¹Gastroenterology, Queen Alexandra Hospital, Portsmouth, UK

²Medical AI Research Department, NEC Corporation, Tokyo, Japan

Correspondence to

Pradeep Bhandari, Gastroenterology, Queen Alexandra Hospital, Portsmouth, UK; pradeep.bhandari@porthosp.nhs.uk

Received 28 February 2021
Accepted 23 June 2021

MESSAGE

Colorectal polyp size is an important biomarker that influences management decisions, but currently used subjective methods are flawed. We explored two computer vision (CV) techniques for binary classification of polyp size as either ≤ 5 mm or > 5 mm. First, we used premeasured phantom polyps (22 such polyps' videos) fixed on a pig colon model to explore the concept of automated sizing using structure from motion (SfM) approach and compared it with the sizing by 10 independent endoscopists: overall, average diagnostic accuracy of the SfM system (85.2%) was superior to endoscopists judgement (59.5%). Second, we developed a deep learning model based on convolutional neural networks (CNN) and found 80% accuracy in 10 videos of human polyps. Real-time automated polyp sizing when combined with artificial intelligence (AI) assisted polyp characterisation could improve polyp management strategies.

IN MORE DETAILS

CV techniques

CV can be defined as the ability of machines to process and understand visual data, automating the type of tasks the human eye would normally be required to do. In order to perform automated polyp size classification, we employed two types of CV techniques, SfM and deep learning (DL).

SfM is a photogrammetric imaging technique that algorithmically recovers three-dimensional (3D) structure of an object from multiple two-dimensional (2D) images and is commonly used in topographic studies. SfM finds matching points in input images and recovers the 3D structure by solving the epipolar constraint equation derived from these matching points, as briefly illustrated in [figure 1](#). The algorithm calculates a camera's pose as a rotation matrix and a translation vector using matching points. Finally, we apply mathematical formulas to compute the distance between the polyp and endoscope, and that distance is used to compute polyp size in real time. Compared with DL, this SfM technique uses less data making it relatively easier and quicker to convert into a clinical device.

DL, based on neural networks, classifies the input images using a large amount of training data. The quality and accuracy of the training data set are very important in this technique. The biggest challenge here is obtaining accurately sized polyp (in the absence of a validated sizing system) images and videos. We have explored both SfM and DL in this

study and developed two separate models. These models were designed to categorise polyps as either category A (≤ 5 mm) or category B (> 5 mm).

EXPERIMENTS AND RESULTS

Evaluation of SfM technique

The phantom polyps were made from silicone sealant and precisely cut into different sizes ranging from 1 mm to 10 mm. These artificial polyps were designed to mimic real human polyps in shape and colour and included both sessile and flat morphologies. The polyps were accurately measured, calibrated and validated by two researchers independent of each other. The polyps were fixed on a pig colon model without altering their shape or size. We used a Fujifilm colonoscope to examine the pig colon, and video was recorded for this examination. Recorded videos were reviewed and annotated to label different polyps and their predetermined size, and we used these videos to evaluate SfM-based sizing system. [Figure 2](#) illustrates the experimental setting and environment and online supplemental video 1 shows an example of the video recording outputs used for development and testing of the system.

The SfM model was tested on 22 videos of phantom polyps equally divided between the two size categories (≤ 5 mm and > 5 mm). We also asked 10 endoscopists of varying degrees of colonoscopy experience to watch the same 22 videos and categorise polyps as either ≤ 5 mm or > 5 mm. Mean diagnostic accuracy was calculated and compared between the two groups using t-test.

Overall, average diagnostic accuracy of the automated sizing system in the animal model was 85.2%, compared with 59.5% in the endoscopist group ($p < 0.0001$). In category A (≤ 5 mm), the automated sizing system and endoscopists showed diagnostic accuracy of 81.2% and 66%, respectively. In category B (> 5 mm), the automated sizing system accuracy was 87.5%, whereas the endoscopists significantly underestimated the polyp sizes in this category and achieved an accuracy of 42.3%. [Table 1](#) summarises the results of this experiment.

Evaluation of CNN DL technique

Here, we used a DL model based on VGG-16 architecture for binary polyp size classification. We used 219 colonoscopy videos containing 301 polyps for training and validation. 80% of the polyp sequences were used for training and the remaining 20% for validation. These polyps were reviewed and sized by three experts and the mean expert's size estimation



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Abdelrahim M, Saiga H, Maeda N, *et al*. Gut Epub ahead of print: [please include Day Month Year]. doi:10.1136/gutjnl-2021-324510

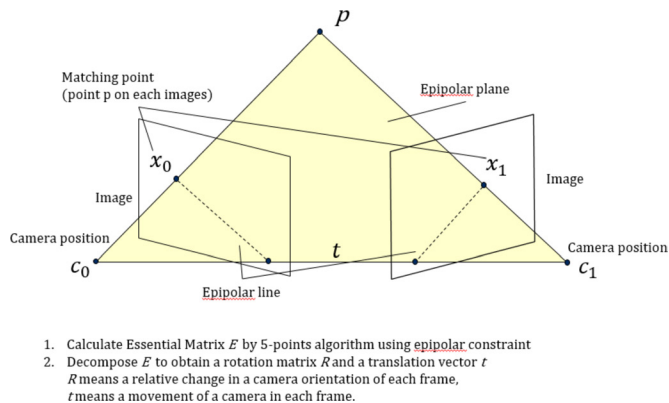


Figure 1 This explains the concept of the structure from motion. We can obtain the relative camera movement using the epipolar constraint equation.

was used for training of the AI model. We employed general data augmentation techniques, such as flip, random cropping and colour conversion. Testing was performed on a completely separate data set containing 10 real-time colonoscopy video recordings. These videos were all recorded with forcep-assisted sizing and were reviewed and sized by three experts, and we used the mean expert’s size as the ground truth. The CNN model achieved an accuracy of 80% in classifying polyps as ≤ 5 mm or >5 mm. **Table 2** summarises the diagnostic accuracy of our system in human polyp video recordings, and online supplemental video 2 shows how our sizing system works on real human polyps.

COMMENTS

Polyp size is related to the risk of cancer and influences surveillance intervals and therapeutic approaches.¹ The 5 mm threshold is particularly important because it influences the implementation of optical diagnosis-based strategies including resect and discard/diagnose and leave approaches and also allows endoscopists to make therapeutic choices of cold versus hot polypectomy.² However, visual size estimation by endoscopists has a significant interobserver variability and error rate,³ and other methods showed variable and sometimes contradicting results.⁴



Figure 2 Images (A) and (B) are examples of the phantom polyps as viewed by the endoscope in the pig colon model. Image (C) shows the pig colon model being scoped. Image (D) shows the real-time endoscopy view during the experiment.

Table 1 Accuracy of automated polyp sizing SfM model and endoscopists in binary classification of colorectal polyps based on their size in an experiment setting (n=22)

	≤ 5 mm (category A)	>5 mm (category B)	Overall (all polyps)
Computer vision accuracy	81.2%	87.5%	85.2%
Endoscopists accuracy	66%	42.3%	59.5%
P value	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$

SfM, structure from motion.

AI is rapidly becoming part of our endoscopy reality, and data on AI-assisted polyp characterisation look very promising, but without accurate sizing, it has limited implication on our practice. However, to date, there is a lack of data on AI-assisted sizing of colorectal polyps, and this report provides one of the very first experiences in this area.

We decided to use artificially created polyps for the early conceptualisation and development of automated polyp sizing systems because human estimation of size, as well as other current methods, are imperfect. This also applies to measuring the size of polyps after removal, especially for small polyps. AI models are only as accurate as the quality and accuracy of their input data, which is why we preferred to come up with this innovative approach to develop a robust AI model.

We developed both SfM-based approach and CNN-based approach separately. We found that SfM-based approach is an appropriate way to algorithmically compute the size of our artificially created polyps in the pig experiment, but the technique proved very challenging when applied to real-world situation (data not shared here). CNN-based approach worked well for the classification of polyp sizes in real time, but it requires a large amount of high-quality and accurately sized data. Therefore, we have been studying both approaches in parallel and have shared the best results in the tables. Overall, SfM worked well for the pig experiment and CNN for the human colon experiment.

Development of this concept is fraught with challenges. We have just highlighted reasons for not using real human polyps in the early developmental phase due to lack of criteria/measures for accurate sizing. Moreover, using SfM technique on artificial polyps in pig colon makes it challenging for the system to find accurate matching points due to the smooth glistening surface and uniform texture of pig colon. This is less of a problem when working on real human polyp videos given the pit and vascular pattern of human colon, but that will create another challenge

Table 2 Accuracy of an automated polyp sizing CNN model in binary classification of colorectal polyps based on their size (n=10)

Polyp number	Size (mm)	Ground truth category	AI category
P1	2	A	A
P2	4	A	A
P3	2	A	A
P4	3	A	A
P5	4	A	A
P6	5	A	A
P7	7	B	A
P8	8	B	B
P9	4	A	B
P10	9	B	B
Overall accuracy of AI model			80%

AI, Artificial Intelligence ; CNN, convolutional neural network.

from other factors like light reflections. Technical solutions, for example, image preprocessing, could allow for mitigation of these issues. On the other hand, DL approach needs a large amount of data to train the network and more stringent measures are needed to ensure the accuracy of polyp sizing and quality of training data sets (ground truth).

We have proven the feasibility of developing and applying two different CV techniques for automated polyp sizing. In the process, we have identified various challenges and strengths of each of these techniques and this will allow us to develop a final product, which can be used for real-time automated sizing of polyps during colonoscopy.

Contributors MA: study concept and design, drafting of the manuscript, analysis and interpretation of data and critical revision of the manuscript. HS: study concept and design, software implementation, analysis and interpretation of data, and revision of the manuscript. NM: study concept and design, software implementation, analysis and interpretation of data, and revision of the manuscript. EH: data acquisition and revision of the manuscript. HI: study concept and design, software implementation, analysis and interpretation of data, and revision of the manuscript. PB: study concept and design, drafting of the manuscript, analysis and interpretation of data, and critical revision of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests MA and EH has no competing interests. PB received a research grant from NEC Japan, Fujifilm, Olympus, Pentax, Boston Scientific,

Medtronic and 3-d Matrix. HS, NM, HI are part of the engineering team employed by NEC Japan who collaborated with the remaining authors.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Mohamed Abdelrahim <http://orcid.org/0000-0001-5167-8776>

REFERENCES

- 1 Vleugels J, Dekker E. Does polyp size matter? *Endosc Int Open* 2017;5:E746–8.
- 2 Ignjatovic A, East JE, Suzuki N, *et al*. Optical diagnosis of small colorectal polyps at routine colonoscopy (detect InSpect characterise Resect and discard; discard trial): a prospective cohort study. *Lancet Oncol* 2009;10:1171–8.
- 3 Chaptini L, Chaaya A, Depalma F, *et al*. Variation in polyp size estimation among endoscopists and impact on surveillance intervals. *Gastrointest Endosc* 2014;80:652–9.
- 4 Rex DK, Rabinovitz R. Variable interpretation of polyp size by using open forceps by experienced colonoscopists. *Gastrointest Endosc* 2014;79:402–7.