

# Algorithm combining virtual chromoendoscopy features for colorectal polyp classification

**Citation for published version (APA):**

Schreuder, R-M., van der Zander, Q. E. W., Fonollà, R., Gilissen, L. P. L., Stronkhorst, A., Klerkx, B., de With, P. H. N., Masclee, A. M., van der Sommen, F., & Schoon, E. J. (2021). Algorithm combining virtual chromoendoscopy features for colorectal polyp classification. *Endoscopy International Open*, 9(10), E1497-E1503. <https://doi.org/10.1055/a-1512-5175>

**Document license:**

CC BY-NC-ND

**DOI:**

[10.1055/a-1512-5175](https://doi.org/10.1055/a-1512-5175)

**Document status and date:**

Published: 01/10/2021

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Algorithm combining virtual chromoendoscopy features for colorectal polyp classification



## Authors

Ramon-Michel Schreuder<sup>1</sup>, Qurine E.W. van der Zander<sup>2,3</sup>, Roger Fonollà<sup>4</sup>, Lennard P.L. Gilissen<sup>1</sup>, Arnold Stronkhorst<sup>1</sup>, Birgitt Klerkx<sup>1</sup>, Peter H.N. de With<sup>4</sup>, Ad M. Masclee<sup>2</sup>, Fons van der Sommen<sup>4</sup>, Erik J. Schoon<sup>1</sup>

## Institutions

- 1 Department of Gastroenterology and Hepatology, Catharina Cancer Institute, Catharina Hospital Eindhoven, The Netherlands
- 2 Department of Gastroenterology and Hepatology, Maastricht University Medical Center, Maastricht, The Netherlands
- 3 GROW School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands
- 4 Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands

submitted 26.12.2020

accepted after revision 11.5.2021

## Bibliography

Endosc Int Open 2021; 09: E1497–E1503

DOI 10.1055/a-1512-5175

ISSN 2364-3722

© 2021. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,  
70469 Stuttgart, Germany

## Corresponding author

Ramon-Michel Schreuder, Catharina Cancer Institute,  
Catharina Hospital, Michelangelolaan 2, 5623 EJ, Eindhoven,  
The Netherlands  
Fax: +310402399751  
[ramonmichel.schreuder@catharinaziekenhuis.nl](mailto:ramonmichel.schreuder@catharinaziekenhuis.nl)

## ABSTRACT

**Background and study aims** Colonoscopy is considered the gold standard for decreasing colorectal cancer incidence and mortality. Optical diagnosis of colorectal polyps (CRPs) is an ongoing challenge in clinical colonoscopy and its accuracy among endoscopists varies widely. Computer-aided diagnosis (CAD) for CRP characterization may help to improve this accuracy. In this study, we investigated the diagnostic accuracy of a novel algorithm for polyp malignancy classification by exploiting the complementary information revealed by three specific modalities.

**Methods** We developed a CAD algorithm for CRP characterization based on high-definition, non-magnified white light (HDWL), Blue light imaging (BLI) and linked color imaging (LCI) still images from routine exams. All CRPs were collected prospectively and classified into benign or premalignant using histopathology as gold standard. Images and data were used to train the CAD algorithm using triplet network architecture. Our training dataset was validated using a threefold cross validation.

**Results** In total 609 colonoscopy images of 203 CRPs of 154 consecutive patients were collected. A total of 174 CRPs were found to be premalignant and 29 were benign. Combining the triplet network features with all three image enhancement modalities resulted in an accuracy of 90.6%, 89.7% sensitivity, 96.6% specificity, a positive predictive value of 99.4%, and a negative predictive value of 60.9% for CRP malignancy classification. The classification time for our CAD algorithm was approximately 90 ms per image.

**Conclusions** Our novel approach and algorithm for CRP classification differentiates accurately between benign and premalignant polyps in non-magnified endoscopic images. This is the first algorithm combining three optical modalities (HDWL/BLI/LCI) exploiting the triplet network approach.

## Introduction

Colorectal cancer (CRC) remains a global health burden and is the third most commonly diagnosed malignancy and the fourth leading cause of death from cancer worldwide [1].

Colonoscopy is considered the gold standard for detection and treatment of colorectal polyps (CRPs) and (early) CRC and therefore decreasing colorectal cancer incidence and mortality. Optical diagnosis of CRPs, the in-vivo characterization of histology of polyps by endoscopists, is of increasing interest for clin-

ical endoscopy practice. Visual differentiation between benign and premalignant CRPs is an ongoing challenge in clinical endoscopy with accuracies of 71% to 90% in the Dutch bowel cancer screening program, exposing patients to risks of incorrect optical diagnosis [2]. Recent studies conducted in community-based hospitals, have been disappointing [3,4] with accuracies not reaching the 90% negative predictive value (NPV) threshold for preservation and incorporation of valuable endoscopic interventions (PIVI) [5]. PIVI thresholds for CRP classification are only met in highly selective expert endoscopists. Therefore, current guidelines recommend to remove all CRPs except diminutive ( $\leq 5$  mm), non-neoplastic rectosigmoid polyps [5]. To allow implementation of “diagnose-and-leave” and “resect-and-discard” strategies, optical diagnosis has to improve. In case of correct optical diagnosis of CRPs, histological examination could be omitted which could improve the time- and cost-effectiveness of colonoscopy and reduce polypectomy-related complications. High-quality endoscopic examinations are necessary to detect benign and premalignant CRPs. It is known that diagnostic accuracy can be improved by the use of optical enhancement techniques such as narrow-band-imaging (NBI), blue light imaging (BLI), linked color imaging (LCI), by magnifying techniques and by chromoendoscopy [6]. BLI optimizes the visualization of the superficial vascular and mucosal patterns by emitting light with a short wavelength (410 nm) [7]. LCI can enhance color separation for red colors [8].

Computer-aided diagnosis (CAD) of endoscopic images using artificial intelligence (AI) has attracted attention because of its potential for better accuracy and lower interobserver variability; therefore, it can improve non-histological polyp evaluation and may facilitate endoscopists in CRP differentiation [9]. Furthermore, nonexpert endoscopists may achieve accuracy levels more easily, hopefully meeting the PIVI threshold [10].

This study aimed to investigate the diagnostic accuracy of a novel approach and algorithm for CRP malignancy classification using a triplet network architecture, which helps the network to better learn similarities between polyps of the same type, while simultaneously boosting the contrast between polyps of different types. To our knowledge this is the first study exploiting three optical modalities (high-definition white light [HDWL], BLI, and LCI) for improving the automatic characterization of CRPs using AI.

## Methods

### Study sample

Endoscopy images of CRPs were prospectively collected from consecutive patients who underwent colonoscopy because of bowel cancer screening programs between October 2017 and June 2018 in the Catharina Hospital Eindhoven, the Netherlands. Consecutive patients were included and for each patient, each colorectal polyp (CRP) was represented by three serial images from almost exactly the same position and distance to the polyp: one HDWL image, and two optically enhanced images (BLI and LCI) without magnification. Per CRP only three pic-

tures were taken. All consecutive images were collected and were retrieved in lossless TIF format and de-identified. To achieve optimal classification, a region of interest (ROI) was drawn for each polyp. The cropped region ensured the coverage of the polyp area, as well as its surrounding texture. The input images were normalized using statistics derived from ImageNet (mean, standard deviation) and rescaled to 224 by 224 pixels to match the input size of the network. Data augmentation was applied for regularization of the models, using a combination of flipping, shifting and  $\pm 90^\circ$  rotation.

CRPs were divided into two categories: benign (hyperplastic polyps [HPs]) and premalignant (adenomas [ADs], sessile serrated lesions [SSAs] and T1 colorectal carcinomas [T1-CRC]) using histopathology as gold standard. The distribution of histological subtypes of CRPs corresponded to the natural occurrence in the screening population.

### Ethics

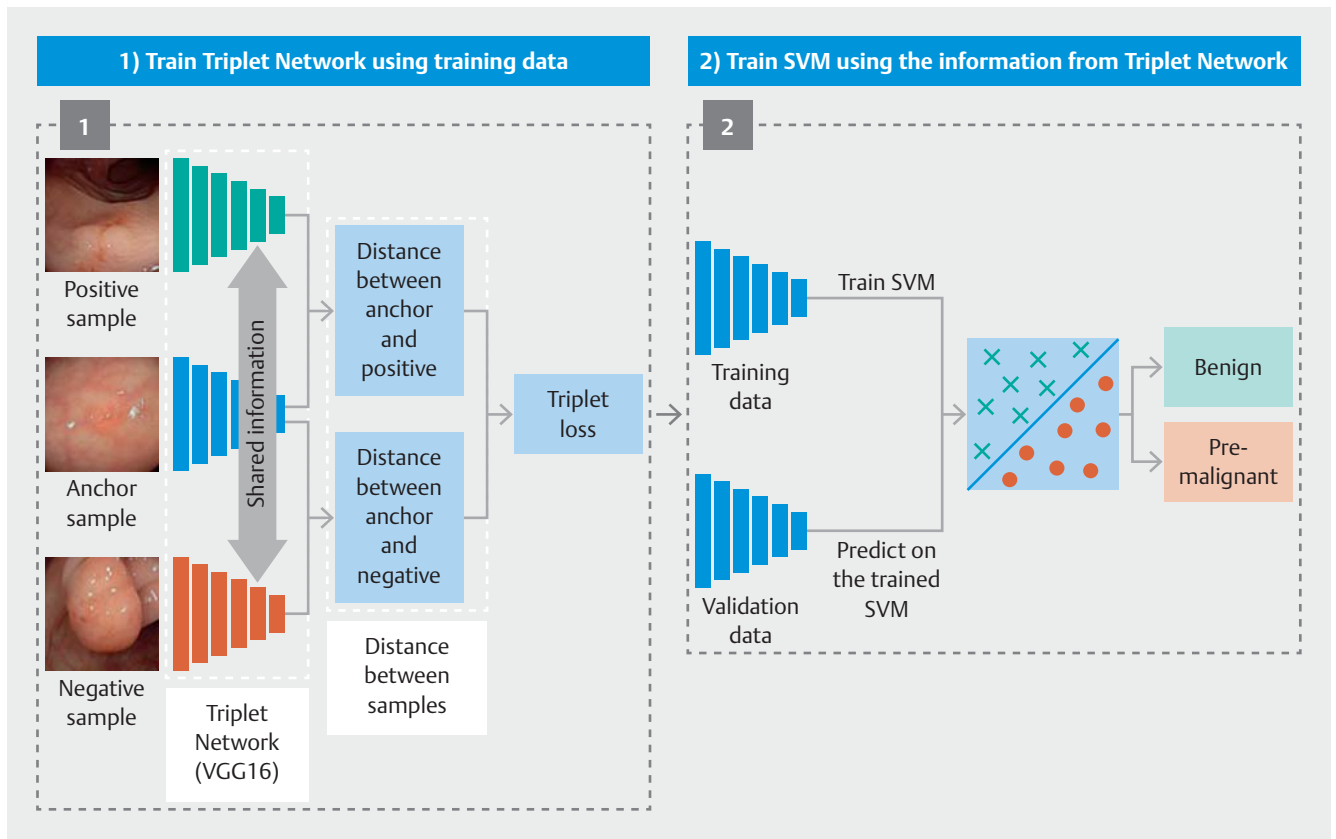
The study was approved by the Local Institutional Review Board of the Catharina Hospital, Eindhoven, the Netherlands (W20.118).

### Endoscopic procedure

Colonoscopies were performed as part of screening colonoscopies. Patients underwent routine bowel preparation, which consisted of 10 mg bisacodyl and 2 L macrogol/electrolytes (Moviprep, Norgine BV, Amsterdam, the Netherlands) given in split doses. All procedures were performed by experienced endoscopists (> 5000 colonoscopies). Patients underwent colonoscopy under conscious sedation (fentanyl and midazolam intravenously) in doses at the discretion of the endoscopist. Endoscopic examination was performed using high-definition colonoscopes (Fujifilm Eluxeo 700 series, Tokyo, Japan) and high-definition monitors. All CRPs detected during examination were resected using either cold snare or hot snare resection. CRPs that could not be resected during initial colonoscopy, were resected during a second colonoscopy using piecemeal endoscopic mucosal resection, endoscopic submucosal dissection, or endoscopic full-thickness resection. Histological diagnosis of CRPs was made and size was estimated by two expert gastrointestinal pathologists. The final classification of benign (HPs) or premalignant (ADs, SSAs, and T1 CRC) was made based on histological findings according to the revised Vienna classification.

### Triplet network

Our CAD algorithm was developed to differentiate between benign and premalignant CRPs, using a state-of-the-art deep learning architecture. Deep learning is a specific category of machine learning that allows learning of a hierarchical representation of data via artificial neural networks multiple layers. The CAD algorithm is capable of automatically learning unique visual properties of polyps in order to classify its malignant potential given a colonoscopy image. During the training process, each CRP-image was compared with histopathological findings. The learned information of the network was updated based on the error between histopathology and prediction. This process



► **Fig. 1** Workflow of the proposed algorithm in which a triplet network was trained using the triplet loss. From the learned features several SVMs were trained for each modality (HDWL, BLI, LCI). The combination of all SVMs was used to increase the classification of the model.

was repeated for every image in the training set until the error was minimal in a process known as backpropagation. Hence, the lower the error, the closer the network is to predict the correct histology.

The triplet network architecture consisted of three identical sub-networks with shared parameters. A triplet contained a query image (anchor), a positive image (positive), and a negative image (negative) where the positive image is more similar to the query image than the negative image. The combination of anchor, positive, and negative sample is called a triplet. The network output consisted of the Euclidean (L2) distance between the anchor and the positive sample, and the anchor and the negative sample. The triplet loss function was associated with the triplets and their individual distances (► **Fig. 1**).

### Training dataset

To preserve balanced training for both classes, the same amount of triplets were selected for both the positive and the negative class. We assigned the anchor and positive samples as benign polyps, hence the negative class was referred to as pre malignant polyps. To achieve balanced training, data augmentation (rotations, horizontal flip, and vertical flip) was applied during training to successfully obtain an equal amount of positive and negative samples in each network, only leaving the anchor sample unaffected. We trained our models using the

stochastic gradient descent (SGD) algorithm with learning rate 0.001, momentum 0.9, and decay 0.001.

### Testing dataset

For internal validation of our training dataset, we used a three-fold cross validation method. In total three networks were trained with the exact same hyperparameters. We randomly shuffled the dataset into three sets ( $d_0$ ,  $d_1$ , and  $d_2$ ), so that all three sets were equal size, while respecting the image-patient association, such that images of the same patient were never represented in different sets. We combined all modalities (HDWL, BLI, and LCI) in our dataset. All three modalities of a unique CRP were represented in the same set. For final classification of the images, a Support Vector Machine (SVM) was employed, which was trained on features provided by the triplet network using a grid-search for hyperparameter optimization (► **Fig. 1**).

### Outcomes and statistical analysis

The primary outcome was the diagnostic accuracy of optical diagnosis of colorectal polyps using a computer-aided-diagnosis (CAD) algorithm. The main outcome measures were accuracy, sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), and area under the receiver operating curve (AUC). Diagnostic time was also evaluated. All statistical

► **Table 1** Polyp characteristics.

Polyp characteristics	n=203
Size, n (%)	
▪ ≤5 mm	85 (41.9)
▪ 6–9 mm	38 (18.7)
▪ ≥10 mm	80 (39.4)
Location, n (%)	
▪ Cecum	8 (3.9)
▪ Ascending colon	35 (17.2)
▪ Transverse colon	31 (15.3)
▪ Descending colon	27 (13.3)
▪ Sigmoid	61 (30.0)
▪ Rectum	41 (20.2)
Histology, n (%)	
▪ Hyperplastic polyp	29 (14.3)
▪ Adenoma	156 (76.8)
▪ Sessile serrated adenomas	12 (5.9)
▪ T1 colorectal carcinoma	6 (3.0)

analyses were performed with IBM SPSS Statistics for Windows version 25 (Armonk, New York, United States).

## Results

We developed a deep-learning algorithm using 609 colonoscopy images of 203 CRPs of 154 prospectively enrolled patients who underwent a colonoscopy examination. Ninety-four patients were male (61%) and the mean age was 65.8 years ( $\pm 7.95$ ).

### Colorectal polyps

In total 203 CRPs were included. ► **Table 1** shows the characteristics of these CRPs. 74 CRPs (36.5%) were located in the proximal colon (cecum, ascending colon, and transverse colon) and 129 CRPs (63.5%) were located in the distal colon (descending colon, sigmoid colon, and rectum). Based on size, 85 CRPs (41.9%) were diminutive ( $\leq 5$  mm), 38 (18.7%) were small (6–9 mm), and 80 (39.4%) were large ( $\geq 10$  mm). Regarding histological evaluation, 29 CRPs (14.3%) were found to be benign (hyperplastic) and 174 CRPs (85.7%) were premalignant. Of these premalignant CRPs, 156 were adenomas (76.8%), 12 SSAs (5.9%), and six T1-CRC (3.0%). ► **Table 2** shows the prevalence of premalignant CRPs according to size.

### Algorithm

The combined CAD algorithm correctly characterized 182 of 203 CRPs: The CAD algorithm showed a diagnostic accuracy of 83.3% for HDWL images. This diagnostic accuracy increased to 84.2% for LCI images and to 89.2% for BLI images. Combining

► **Table 2** Prevalence of premalignant CRPs, according to CRP size.

Size (mm)	Prevalence of premalignant CRPs, n/N (%)
$\leq 5$	61/85 (71.8)
6–9	34/38 (89.5)
$\geq 10$	79/80 (98.8)

CRP, colorectal polyp.

triplet network features with all three-image enhancement modalities resulted in an accuracy of 90.6% for CRP malignancy classification, by pooling the normalized decision values of the three modalities. Of the 174 premalignant CRPs, 156 (sensitivity of 89.7%) and of the 29 benign CRPs, 28 (specificity of 96.6%) were correctly classified. The overall PPV and NPV of the CAD algorithm were 99.4% and 60.9%, respectively. The ROC curve for the combined algorithm performance in differentiating benign from premalignant polyps was calculated and the AUC was 0.932 (► **Table 3**).

► **Table 4** shows the diagnostic performance of the CAD algorithm in the diagnosis of diminutive, small, and large CRPs using a combination of HDWL, BLI, and LCI. The accuracy of our CAD algorithm was 84.7%, 100%, and 92.5% with a NPV of 65.7%, 100%, and 14.3%, for diminutive, small, and large CRPs, respectively. Of the 85 included diminutive CRPs 12 premalignant CRPs were misclassified. Of these 12 CRPs, two images were blurred and two out of focus (► **Fig. 2**). The other eight images were good quality.

We included 80 large CRPs in the dataset, 79 premalignant and one benign. Six premalignant CRPs were classified as benign. Five images of those misclassified CRPs contained SSAs. The other misclassified CRP concerned a blurred image. When we focus on small CRPs, our algorithm correctly classified all 34 premalignant and four benign CRPs.

The classification time of our algorithm was approximately 90 ms per image computed on a single GPU TitanXp.

## Discussion

In our study, we prospectively collected 609 images of 203 CRPs to evaluate the diagnostic ability of our CAD algorithm. To prevent selection bias, all consecutive CRP images were included in the database, regardless of out-of-focus or blurred images.

Our results show that combining a triplet network with additional endoscopic enhancement modalities (HDWL, BLI, and LCI) improves the overall diagnostic accuracy from 83.3% (HDWL), 89.2% (BLI), and 84.7% (LCI) to 90.6% for CRP malignancy classification.

Previous studies mainly focused on NBI magnification images or endocytoscopy, systems that are not readily available in current clinical practice [6, 11–22]. Little research has been done on the performance of CAD systems using HDWL, LCI or BLI techniques without magnification. Comparative studies presenting different CAD systems with non-magnification ima-

► **Table 3** Diagnostic accuracy of CADx based on HDWL, BLI, and LCI.

	Diagnostic performance, % (95% CI)					AUC (95% CI)
	Accuracy	Sensitivity	Specificity	PPV	NPV	
HDWL	83.3 (78.1–88.4)	81.6 (76.3–86.9)	93.1 (89.6–96.6)	98.6 (97.0–100)	45.8 (36.0–55.5)	0.897 (0.849–0.938)
BLI	89.2 (84.9–93.4)	89. (84.8–93.4)	89.7 (85.5–93.8)	98.1 (96.3–99.9)	57.8 (49.3–66.2)	0.927 (0.887–0.960)
LCI	84.2 (79.2–89.2)	82.8 (77.6–88.0)	93.1 (89.6–96.6)	98.6 (97.1–100)	47.4 (37.7–57.0)	0.895 (0.839–0.942)
HDWL + BLI + LCI	90.6 (86.6–94.6)	89.7 (85.5–93.8)	96.6 (94.0–99.1)	99.4 (98.3–100)	60.9 (52.4–69.3)	0.932 (0.894–0.963)

CADx, computer-aided diagnosis; HDWL, high-definition, non-magnified white light; BLI, blue light imaging; LCI, linked color imaging; CI confidence interval; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

► **Table 4** Diagnostic performance of the CADx algorithm in diagnosis of small and large polyps using a combination of HDWL, BLI, and LCI.

	Diagnostic performance, % (95% CI)					AUC (95% CI)
	Accuracy	Sensitivity	Specificity	PPV	NPV	
Diminutive polyps (≤5 mm)	84.7 (77.1–92.4)	80.3 (71.9–88.8)	95.8 (91.6–100)	98.0 (95.1–100)	65.7 (53.5–77.9)	0.874 (0.778–0.942)
Small polyps (6–9 mm)	100	100	100	100	100	1.000
Large polyps (≥10 mm)	92.5 (86.7–98.3)	92.4 (86.6–98.2)	100	100	14.3 (0–34.6)	0.962 (0.911–1.000)

CADx, computer-aided diagnosis; HDWL, high-definition, non-magnified white light; BLI, blue light imaging; LCI, linked color imaging; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.



► **Fig. 2** An example of a blurred image in HDWL, BLI, and LCI.

ges, showed diagnostic accuracies between 75.1% (HDWL) [23] and 78.4% (LCI) [8]. Our diagnostic accuracy was higher compared to results of those previous studies, even with out-of-focus and blurry pictures which were included in our analysis. Prior studies, with accuracies between 82.5% and 97.8% which used BLI or NBI in their algorithm were based on magnified images or videos [11, 12, 14, 16, 21, 22, 24, 25].

The PIVI initiatives state that a “diagnose-and-leave” approach is acceptable for a diminutive hyperplastic polyp when high-confidence endoscopic judgment can achieve ≥90% NPV

for adenomas [5]. The NPV of our CAD algorithm regarding optical diagnosis of small CRPs is 100%, in contrast to the NPV of diminutive or large CRPs being 65.7% and 14.3%, respectively. The difference in NPV between large and small CRPs might be explained by the presence of SSAs. Eighty-three percent (5 of 6) of the misclassified large (≥10 mm) premalignant CRPs were SSAs, which are polyps with malignant potential. It is difficult to diagnose SSAs with a CAD algorithm, with a very high risk to misdiagnose them as non-neoplastic, since they generally have no nuclear abnormalities and mimic hyperplastic

polyps. Besides, since we've included all consecutive CRP's, SSAs occurred limited in our dataset. This might be the reason that some SSAs were misclassified. In the future, larger datasets and preferably videos, with higher numbers of SSAs are needed to train the dataset to achieve a higher accuracy and NPV for the detection of those lesions. However, SSAs are managed clinically in a similar fashion as adenomatous polyps. Therefore, the clinical significance is limited because current guidelines advice to resect all CRPs proximal to the sigmoid colon [26].

A reason for the difference in NPV between diminutive and small CRPs might be the significant percentage (30.8%) of blurred or out of focus images in misclassified images, thus generating false negative analysis. Use of non-magnification images and the size of the CRP relative to the image could be other explanations. Besides, for a CAD system a bigger CRP generate more data and is therefore easier to classify. These findings suggests that experience in making high-quality endoscopic images with appropriate distance, cleaning of lesions and clear focus are particularly important for computer analysis.

Unfortunately, our algorithm was not able to reach the NPV threshold of  $\geq 90\%$  for diminutive CRPs, in contrast to a system developed by Chen et al. (2018) and Byrne et al. (2017) [11, 24]. There are several possible explanations for the differences in NPV between these studies. In contrast to our study, Chen et al. included only high-quality images of diminutive polyps, so selection bias might have occurred. Secondly, maximum magnification power was used for all diagnoses in their study. However, magnifying endoscopes are not regularly used for colonoscopies and are not commercially available in Western, Asian or American countries. Byrne et al. used videos of diminutive CRPs instead of still images. With video it is easier for a computer to generate a diagnosis, because the algorithm receives more information from a video (52 frames/sec) compared to a single still image. In addition, they excluded all the videos (15%) that did not develop at least 50% confidence in the diagnosis by the AI model which also leads to selection bias. In this study, all 203 CRP images were included, regardless of the algorithm confidence in diagnosis. Both abovementioned studies also excluded SSAs in their analysis. Besides, our CAD algorithm used only color as the single parameter included for analysis. In previous magnified NBI-based CAD analysis with accuracies  $\geq 90\%$ , the accuracy of the CAD algorithm is not based on color as a single parameter but also on capillary patterns, microstructure of the mucosa and magnification of images. Analyses based on the NBI system which comprises capillary patterns and microstructure of the mucosa as well, are more likely to have higher accuracy than analyses that only consider color for diagnosis [12, 14, 21, 22].

This study has several strengths. First, to our knowledge this is the first study combining HDWL, BLI, and LCI as imaging enhancement modalities in the development of a CAD algorithm. Second, we used non-magnification colonoscopy for the development of this CAD algorithm, which has practical consequences as mentioned earlier. Third, we included SSAs in the study representing real clinical practice, while prior studies have been criticized for not incorporating SSAs in their analysis. Fourth, state-of-the-art machine learning techniques and deep

learning architectures were used to develop the CAD algorithm. Fifth, because we did not exclude any suboptimal images (e. g. out-of-focus or blurred images), there is no risk of selection bias; therefore, the dataset is comparable to colonoscopy findings in daily practice. The increase in diagnostic accuracy by using the sum of the three modalities is another scientific proof of the fact that all these modalities contain separate information, not always visible to the human eye or appreciated by the human brain.

This study has also certain limitations. First, we did not reach the NPV threshold of  $\geq 90\%$  for diminutive CRPs. This might be explained by the presence of blurred or out of focus images, the size of the CRP relative to the image, the use of non-magnified images and the small size of (diminutive) CRPs in the training data. Second, the amount of SSAs was included according to natural occurrence, but therefore occurred limited in the dataset. This might be the reason that some SSAs were misclassified. In the future, larger datasets and preferably videos with higher numbers of abovementioned CRPs are needed to train the algorithm to achieve a higher accuracy and NPV for the detection of those lesions. Third, because we included all consecutive images in the dataset, some images were blurred or out-of-focus. Therefore, the confidence in the prediction of our CAD algorithm could be underestimated. When the model is used in a true live patient scenario, endoscopists would be able to move the colonoscope and change the image in an attempt to allow the model to build up its confidence. Quality control of images should be implemented in the system when it is going to be used in clinical practice. Because of inclusion of consecutive images during bowel cancer screening colonoscopies, also large ( $> 1$  cm) CRPs were included. These large CRPs would be resected in general. However, we've included all these CRPs, because our goal was to generate input for and train our algorithm, irrespective of size or histology.

Fourth, in the development of the CAD algorithm it was only possible to analyze benign (HP) versus premalignant (AD, SSA, and CRC) because of the relatively small number of SSAs and CRCs in the database. More data are needed to differentiate accurately between CRPs with high-grade dysplasia, endoscopically resectable early cancers versus deep invasive ones. We are currently enrolling patients in a prospective trial for this purpose.

Fifth, we did not use an external test database for validation of our CAD algorithm, but we used threefold cross validation with our internal dataset, this limited the validity and generalizability of this study. As mentioned above, we are including patients for a prospective trial to validate our algorithm with a test database nowadays. More data would improve our algorithm. Sixth, the algorithm's performance in specific disease states of the colon such as inflammatory bowel disease or intestinal bleeding has not been evaluated in the current study as we did not have sufficient data from this subset of cases. Seventh, for this analysis, images had to be made in all three modalities (HDWL, BLI and LCI), which could be more time consuming.

## Conclusions

In conclusion, our novel approach and algorithm for automatic polyp malignancy classification differentiates accurately between benign and premalignant CRPs in endoscopic still images. With an overall accuracy of 90.6%, our CAD algorithm has promising diagnostic potential for predicting histopathology of colorectal polyps based on color analysis. We achieved high accuracy levels compared with similar studies, especially without magnification. To our knowledge, this is the first study combining three imaging modalities (HDWL, BLI, and LCI) using a single AI model based on a triplet network. To meet the PIVI criteria for the “diagnose-and-leave” strategy in the future, the algorithm should be further improved by increasing the amount of high-quality images and videos. To increase generalizability, validation of our CAD algorithm should be performed in a prospective, real-time, multicenter clinical trial. Validation studies in real patients are ongoing.

## Competing interests

This study was financially supported by the Catharina Research Foundation.

## References

- [1] Torre LA, Bray F, Siegel RL et al. Global cancer statistics, 2012. *CA Cancer J Clin* 2015; 65: 87–108
- [2] van de Wetering AJ, Meulen LW, Bogie RM et al. Optical diagnosis of diminutive polyps in the Dutch Bowel Cancer Screening Program: Are we ready to start? *Endosc Int Open* 2020; 8: E257–E265
- [3] Ladabaum U, Fioritto A, Mitani AK et al. Real-time optical biopsy of colon polyps with narrow band imaging in community practice does not yet meet key thresholds for clinical decisions. *Gastroenterology* 2013; 144: 81–91
- [4] Rees CJ, Rajasekhar PT, Wilson A et al. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. *Gut* 2017; 66: 887–895
- [5] Rex DK, Kahi C, O'Brien M et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc* 2011; 73: 419–422
- [6] Muguruma N, Miyamoto H, Okahisa T et al. Endoscopic Molecular imaging: status and future perspective. *Clin Endosc* 2013; 46: 603–610
- [7] Yoshida N, Yagi N, Inada Y et al. Ability of a novel blue laser imaging system for the diagnosis of colorectal polyps. *Dig Endosc* 2014; 26: 250–258
- [8] Min M, Su S, He W et al. Computer-aided diagnosis of colorectal polyps using linked color imaging colonoscopy to predict histology. *Sci Rep* 2019; 9: 2881
- [9] Gross S, Trautwein C, Behrens A et al. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc* 2011; 74: 1354–1359
- [10] Kominami Y, Yoshida S, Tanaka S et al. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; 83: 643–649
- [11] Chen PJ, Lin MC, Lai MJ et al. Accurate classification of diminutive colorectal polyps using computer-aided analysis. *Gastroenterology* 2018; 154: 568–575
- [12] Gross S, Trautwein C, Behrens A et al. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc* 2011; 74: 1354–1359
- [13] Kiesslich R, Burg J, Vieth M et al. Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo. *Gastroenterology* 2004; 127: 706–713
- [14] Kominami Y, Yoshida S, Tanaka SB et al. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; 83: 643–649
- [15] Kudo SE, Mori Y, Wakamura K et al. Endocytoscopy can provide additional diagnostic ability to magnifying chromoendoscopy for colorectal neoplasms. *J Gastroenterol Hepatol* 2014; 29: 83–90
- [16] Mesejo P, Pizarro D, Abergel A et al. Computer-aided classification of gastrointestinal lesions in regular colonoscopy. *IEEE Trans Med Imaging* 2016; 35: 2051–2063
- [17] Misawa M, Kudo SE, Mori Y et al. Characterization of Colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy. *Gastroenterology* 2016; 150: 1531–1532
- [18] Mori Y, Kudo SE, Wakamura K et al. Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos). *Gastrointest Endosc* 2015; 81: 621–629
- [19] Mori Y, Kudo SE, Chiu PW et al. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. *Endoscopy* 2016; 48: 1110–1118
- [20] Takeda K, Kudo SE, Mori Y et al. Accuracy of diagnosing invasive colorectal cancer using computer-aided endocytoscopy. *Endoscopy* 2017; 49: 798–802
- [21] Takemura Y, Yoshida S, Tanaka S et al. Computer-aided system for predicting the histology of colorectal tumors by using narrow-band imaging magnifying colonoscopy (with video). *Gastrointest Endosc* 2012; 75: 179–185
- [22] Tischendorf JJ, Gross S, Winograd R et al. Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. *Endoscopy* 2010; 42: 203–207
- [23] Komeda Y, Handa H, Watanabe TT et al. Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience. *Oncology* 2017; 93: 30–34
- [24] Byrne MF, Chapados N, Soudan F et al. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2019; 68: 94–100
- [25] Cheng Tao Pu LZ, Maicas G et al. Computer-aided diagnosis for characterisation of colorectal lesions: a comprehensive software including serrated lesions. *Gastrointest Endosc* 2020; 92: 891–899
- [26] Ferlitsch M, Moss A, Hassan C et al. Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy* 2017; 49: 270–297