ORIGINAL ARTICLE

Development and evaluation of a deep learning model to improve the usability of polyp detection systems during interventions

Markus Brand[1] | Joel Troya[1] | Adrian Krenzer[1,2] | Zita Saßmannshausen[1] | Wolfram G. Zoller[3] | Alexander Meining[1] | Thomas J. Lux[1] | Alexander Hann[1]

[1]Interventional and Experimental Endoscopy (InExEn), Internal Medicine II, Gastroenterology, University Hospital Würzburg, Würzburg, Germany

[2]Artificial Intelligence and Knowledge Systems, Institute for Computer Science, Julius-Maximilians-Universität Würzburg, Würzburg, Germany

[3]Department of Internal Medicine and Gastroenterology, Katharinenhospital, Stuttgart, Germany

Correspondence
Joel Troya, Universitätsklinikum Würzburg, Medizinische Klinik und Poliklinik II, Oberdürrbacher Str. 6, 97080 Würzburg, Germany.
Email: TroyaSebas_J@ukw.de

Abstract

Background: The efficiency of artificial intelligence as computer-aided detection (CADe) systems for colorectal polyps has been demonstrated in several randomized trials. However, CADe systems generate many distracting detections, especially during interventions such as polypectomies. Those distracting CADe detections are often induced by the introduction of snares or biopsy forceps as the systems have not been trained for such situations. In addition, there are a significant number of non-false but not relevant detections, since the polyp has already been previously detected. All these detections have the potential to disturb the examiner's work.

Objectives: Development and evaluation of a convolutional neuronal network that recognizes instruments in the endoscopic image, suppresses distracting CADe detections, and reliably detects endoscopic interventions.

Methods: A total of 580 different examination videos from 9 different centers using 4 different processor types were screened for instruments and represented the training dataset (519,856 images in total, 144,217 contained a visible instrument). The test dataset included 10 full-colonoscopy videos that were analyzed for the recognition of visible instruments and detections by a commercially available CADe system (GI Genius, Medtronic).

Results: The test dataset contained 153,623 images, 8.84% of those presented visible instruments (12 interventions, 19 instruments used). The convolutional neuronal network reached an overall accuracy in the detection of visible instruments of 98.59%. Sensitivity and specificity were 98.55% and 98.92%, respectively. A mean of 462.8 frames containing distracting CADe detections per colonoscopy were avoided using the convolutional neuronal network. This accounted for 95.6% of all distracting CADe detections.

Conclusions: Detection of endoscopic instruments in colonoscopy using artificial intelligence technology is reliable and achieves high sensitivity and specificity.

Accordingly, the new convolutional neuronal network could be used to reduce distracting CADe detections during endoscopic procedures. Thus, our study demonstrates the great potential of artificial intelligence technology beyond mucosal assessment.

KEYWORDS
CADe, colonoscopy, deep learning, instrument, intervention

INTRODUCTION

Artificial intelligence (AI) for colonic polyp detection is the most important application of this new technology in gastrointestinal endoscopy to date. Efficiency and functionality of these computer-aided detection (CADe) systems have been demonstrated in several randomized trials.[1–6] However, CADe systems also show many false positive (FP) detections.[7,8] These false markings can affect the examiner's concentration. If a false detection occurs in addition to a relevant finding, the examiner's attention may be distracted, leading to missed findings in the worst case.[9]

Daily use of CADe systems shows that endoscopic interventions (especially biopsies and polypectomies) lead to many false activations of CADe systems. In this case, false positive activations occur due to the inserted instruments (forceps, needle, snare), but also due to intervention on the mucosa itself (injection, resection, clipping). In addition, there are a significant number of non-false but not relevant detections, since the polyp has already been previously detected. To enable the investigators to put their full concentration on the intervention, no distracting AI signals regarding polyp detection should be visible during the procedure.

Therefore, the aim of the current study was to develop and evaluate an AI system that reliably detects introduced instruments in order to disable the CADe system during an intervention and avoid distracting detections.

METHODS

Training dataset

Data from nine different center in Germany, two university hospitals, one community-based hospital and six gastroenterology practices, were collected retrospectively from March 2019 to August 2021. A total of 519,856 images were selected from 580 randomly selected different colonoscopy videos for building the training dataset. Of all images in the training dataset, 144,217 (27.7%) contained a visible instrument (Figure 1). The types of instruments used for training the model included graspers, hot and cold snares, injection needles and clips. No minimum or maximum number of images per instrument in a colonoscopy was predefined for training the model. Images of good and poor quality (e.g., blurry images) were chosen for model training in order to represent a real-life scenario. The colonoscopies were performed using different processors including CV-190 and CV-170

Key summary

Summarize the established knowledge on this subject

- Multiple computer-aided diagnosis (CADe) systems for polyp detection are currently introduced into clinical practice.
- CADe results in multiple distracting detections, especially during therapeutic interventions when instruments are visible.

What are the significant and/or new findings of this study?

- Development and evaluation of a deep learning model to recognize visible instruments that are used for therapeutic intervention in gastrointestinal endoscopy.
- Our model automatically prevents distracting CADe detections during therapeutic interventions.

(Olympus Europa SE & Co. KG, Hamburg, Germany), Image1 S (Karl Storz SE & Co. KG, Tuttlingen, Germany) and EPK-i7000 (Pentax Europe GmbH, Hamburg, Germany) and were recorded using a standard computer with a video grabber (DeckLink Mini Recorder, Blackmagic Design Pty Ltd., Melbourne, Australia) and a custom recording software. Representative images of the four different processor types are displayed in Supplementary Figure 1.

To reduce almost identical images in the dataset, images from the same colonoscopy were filtered to exclude neighboring images. For training the convolutional neuronal network (CNN), the dataset was split into a train (90%) and a validation (10%) dataset. To prevent bias, all images of one colonoscopy were either included in the train or the validation dataset.

Preprocessing and CNN training

Initially, a region of interest for each used processor type was defined and images were cropped accordingly. Afterward, images were zero padded and resized to a dimension of 512×512 pixels to yield uniform images. For train data, the image augmentation pipeline (Supplementary Code Section 1) was applied. All images underwent the standard procedure of image normalization (Supplementary Code Section 2), so that each color and brightness value are standardized. Resulting
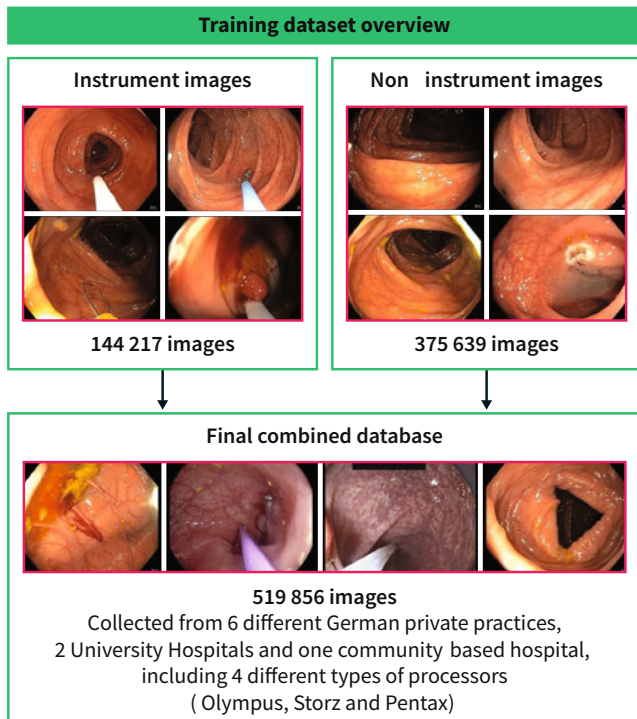
Training dataset overview

Instrument images

144 217 images

Non instrument images

375 639 images

Final combined database

519 856 images
Collected from 6 different German private practices,
2 University Hospitals and one community based hospital,
including 4 different types of processors
(Olympus, Storz and Pentax)

FIGURE 1 Characteristics of the training dataset containing images with and without visible instruments captured using the four different endoscopy processor types

constant values for this were calculated on all train data images and were always used to normalize each input image. CNN training is described in detail in Supplementary Material. The current instrument detection software is freely available to download for research purposes (https://github.com/Maddonix/instrument_detection).

Model testing

The CNN for detecting visible instruments was tested in the withdrawal phase of a set of 10 full-length colonoscopy videos by analyzing its performance in each single image of the videos. To stabilize the prediction results, a running mean function was applied. This was performed to avoid erroneous suppressions of AI detections caused by our CNN. Here, we assigned the current video image the majority label of itself as well as the previous 14 video images, representing a threshold of 467 ms. This same dataset was used to test the change in performance of a CADe system (GI Genius, Medtronic Inc., Ireland, Version March 2020) in the reduction of distracting activations with the developed instrument detection system.

Ethics approval

Patients provided written informed consent prior to video recording. The ethics committee of the University hospital Würzburg approved retrospective analysis of the data used in this study.

Statistical analysis

Two evaluations were statistically analyzed: the capabilities of the instrument detection system and the reduction of CADe distracting activations. Per-frame sensitivity and specificity, accuracy, and Receiver Operating Characteristic (ROC) were calculated for both evaluations. Sensitivity has been defined as the ratio between the number of frames with a visible instrument that were correctly detected (TP) and the total number of frames with a visible instrument (TP+FN). Specificity was defined as the ratio between the number of frames without a visible instrument that were correctly assessed (TN) and the sum of the total number of frames with a false detection and the TN frames (FP+TN). Accuracy was defined as the ratio between the number of correct system assessments (TP+TN) and the total number of frames. Metrics where weighted average to compensate for the imbalance of images with/out a visible instrument. For the calculation of the weighted average metrics the parameter "average" in every used function of the sklearn.metrics module from scikit-learn 1.0.2 package was set to "weighted". All calculations were performed using Python Software (version 3.6).

RESULTS

Characteristics of the patient cohort

The test dataset comprised 10 full-length colonoscopy videos from 10 different patients. Men and women were equally represented, the mean age was 57.1 (interquartile range; 46–65) and the mean Boston Bowel Preparation Scale was 6.9 (range; 6–9) (Supplementary Table 1). The total duration of the withdrawal phase, with the duration of interventions included, was 1 h and 25 min, corresponding to 153,623 single video frames. During this time, instruments were visible for a total of 7 min and 12 s on the screen These 10 videos included a total of 12 different interventions, where 19 different endoscopic through the scope instruments were used: 4 cold snares, 11 graspers, 1 hot snare, 2 needles and 1 clip (Table 1).

Performance of the instrument detection system

The CNN overall accuracy achieved in the detection of visible instruments in the test dataset was 98.59%. Sensitivity and specificity were 98.55% and 98.92%, respectively. The grasper was the instrument that was best detected by the system, with a sensitivity of 99.08% and a specificity of 99.36%, whereas the snare, with a sensitivity of 98.21% and a specificity of 98.51%, was the most difficult instrument to detect, probably because often only the wire was visible. Representative images of a grasper, a snare and a false positive detection of the CNN with the corresponding heat map that depicts the image areas that are recognized as an instrument are

TABLE 1 Characteristics and performance of the instrument detection system in the test dataset

Intervention	Type of instrument	Number of visible instrument frames	Sensitivity (%)	Specificity (%)	Disturbing CADe activations (frames)	Disturbing CADe activations avoided (frames)	False-avoided CADe activations (frames)	Total number of CADe activations
Video 1 Polypectomy	Snare	728	98.60	99.51	377	330	13	3352
Polypectomy	Snare	1262						
Video 2 Polypectomy	Grasper	142	99.77	99.76	6	6	0	396
Video 3 Polypectomy	Grasper	269	99.21	99.68	137	102	5	1252
Video 4 Polypectomy	Grasper	174	98.87	98.43	8	8	2	302
Video 5 Polypectomy	Needle	407	99.22	99.64	1232	1184	54	7834
	Snare	931						
Video 6 Polypectomy	Grasper	1136	99.31	99.67	161	150	6	531
Video 7 Polypectomy	Snare	2760	98.01	98.35	741	736	50	1577
Video 8 Polypectomy	Needle	2493	96.90	96.58	1923	1906	204	7737
	Hot snare	1048						
	Clip	292						
Video 9 Polypectomy	Snare	255	99.33	99.62	101	84	21	1361
Video 10 Random biopsies	5x Grasper	751	98.14	98.98	153	122	2	1099
Polypectomy	Grasper	871						

Abbreviation: CADe, Computer-aided detection system.

presented in Figure 2. The ROC curve illustrating the diagnostic ability of our instrument detection system is depicted in Figure 3. No marked differences in performance were observed relating to BBPS value, that ranged from 6 to 9 (Supplementary Table 1).

Reduction of CADe distracting activations

A total of 25,441 activations were triggered by the CADe system in the test dataset. These activations included detected polyps and false positive detections. 4839 activations (19.02%) occurred when an instrument was visible and were regarded as distracting CADe activations. Especially significant was the amount of distracting activation caused by snares or needles. Our system was able to avoid 4628 of these activations, representing a sensitivity value of 95.64%. Regarding the number of CADe activations that were falsely avoided, the value in frames amounts to 357. Out of those, 292 contained polyps that were previously detected. This implies that the system has a specificity of 98.62% in terms of performance in preventing distracting CADe activations. The overall accuracy was of 99.17%.

The metrics of the developed instrument detection system per intervention and its performance in preventing distracting CADe activations are presented in Table 1 and exemplarily illustrated in Figure 4. In addition, Figure 5 and Video S1 present a graphical example of how the AI system works.

DISCUSSION

Since the introduction of commercially available AI-systems for colorectal polyp detection, the use of these promising systems in daily practice is increasing. The great potential of AI-systems is currently in the field of diagnostics, as CADe systems support the examiner in real time and with high sensitivity.[6,10,11] Since CADe systems have been trained with diagnostic polyp images, they achieve high sensitivity for native polyps in the colon.[12] However, changes to the mucosa in the course of an intervention (e.g., injection) lead to false positive detections, as the systems have not been trained for such situations. The instruments used during the intervention also lead to many false positive detections that may disturb the investigator's concentration. In addition, there are many non-false but irrelevant detections because the polyp causing the intervention has been previously detected. To enable the investigators to put their full concentration on the intervention, no distracting AI signals should be visible during the intervention. This could be achieved by suppressing the CADe signal during the intervention, since polyp detection is not necessary during the intervention.

Currently, the endoscopist can only manually turn off the CADe system and turn it back on after the procedure. Some systems require a button to be pressed on the processor, as not all systems can be controlled via a button on the endoscope. However, it is possible that the endoscopist forgets to turn the system back on after the procedures. Therefore, automatically stopping and starting the

CADe system would increase the comfort for the endoscopist and prevent the CADe system from being accidentally switched off permanently.

Our novel AI system detects inserted instruments with high sensitivity and specificity. Therefore, the system can capture the time frame of an endoscopic intervention with high accuracy. This would

Raw image	Heatmap overlay

FIGURE 2 Grasper (upper row), snare (middle row), and a false positive detection (lower row) of the instrument detecting CNN with the corresponding gradient-weighted class activation mapping (Grad-CAM). Grad-CAM images on the right side visualize areas responsible for the CNN prediction as an instrument

enable the suppression of the CADe signal for the duration of the intervention to focus the investigator's concentration on the intervention. The suppression relates not only to false positive detections but instead to all CADe detections during an intervention that do not add value to the endoscopic image.

The requirements for such a tool detection system are very high, as suppression of the CADe signal outside of an intervention (false positive instrument detection) may increase the risk of missing other visible polyps. Our study shows that our new AI system achieves a very high specificity, which is sufficient for this purpose. To obtain this high specificity, our system was trained with a large number of images from multiple centers using different endoscopy processors. The number of training images we used is comparable to the number used in development of other CADe systems.[13,14] In addition, the optimized algorithm presents only a short delay of 467 ms, that allows for the real-time use in combination with a CADe system.

Since the sensitivity of our AI system is in a high range, the instruments introduced were missed in only a few frames during an intervention. This applies in particular to the insertion and removal of an instrument where only a small portion of it is visible at the edge of the endoscopic view. Once the instrument is in the normal working position, it is quickly and reliably detected by the AI system. Thus, the crucial part of the intervention is captured by our instrument detection system. However, a problem with instrument detection arises when an instrument is pressed so firmly into the mucosa that it is barely visible. In this situation, the instrument recognition works accordingly worse. Nevertheless, our video analysis showed that the new AI system significantly reduced the number of false-positive CADe detections during an endoscopic intervention. While many publications on AI systems only use short, specially selected video sequences in the evaluation phase, our system was tested on full-length colonoscopy, which brings the results much closer to the real examination situation.[15]

Interestingly, the commercially available CADe system seems to generate more detections when a snare is used in comparison to a grasper. There might be different explanations for this phenomenon.
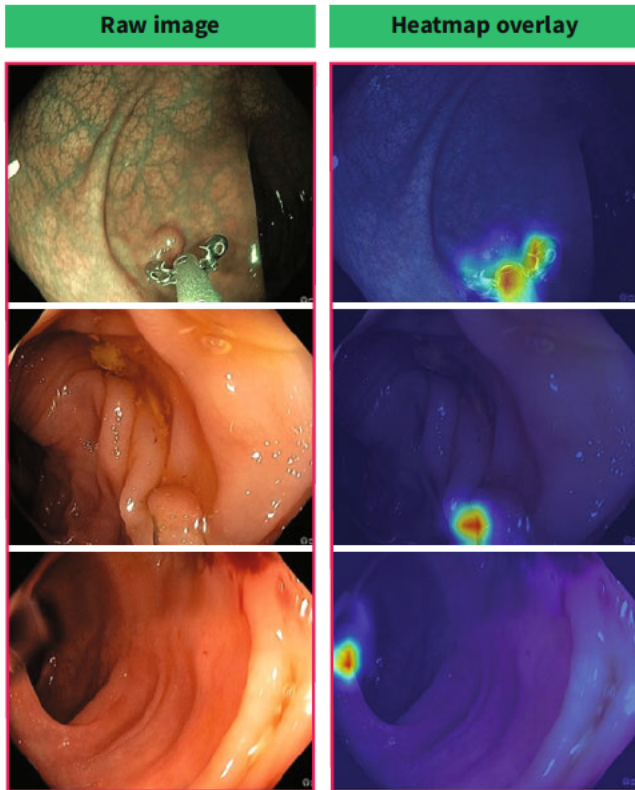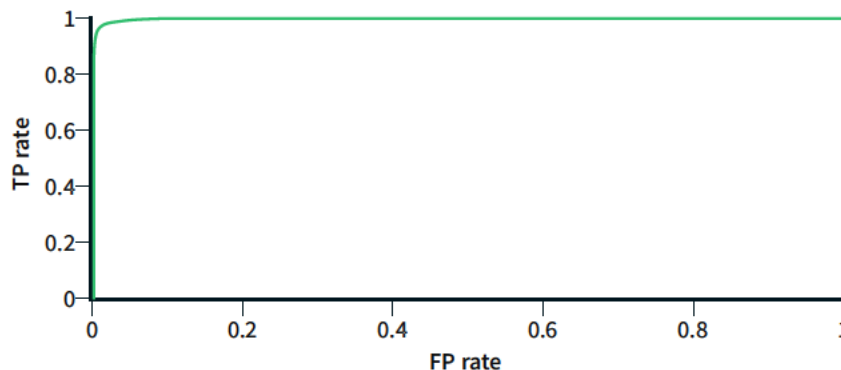

FIGURE 3 Receiver Operating Characteristic Curve of the instrument detection CNN visualizes specificity. While adjusting classification thresholds, the TP rate reaches 96.58% while maintaining a FP rate of 1% resulting in an area under the curve of 0.9971. CNN, convolutional neuronal network; FP, false positive; TP, true positive
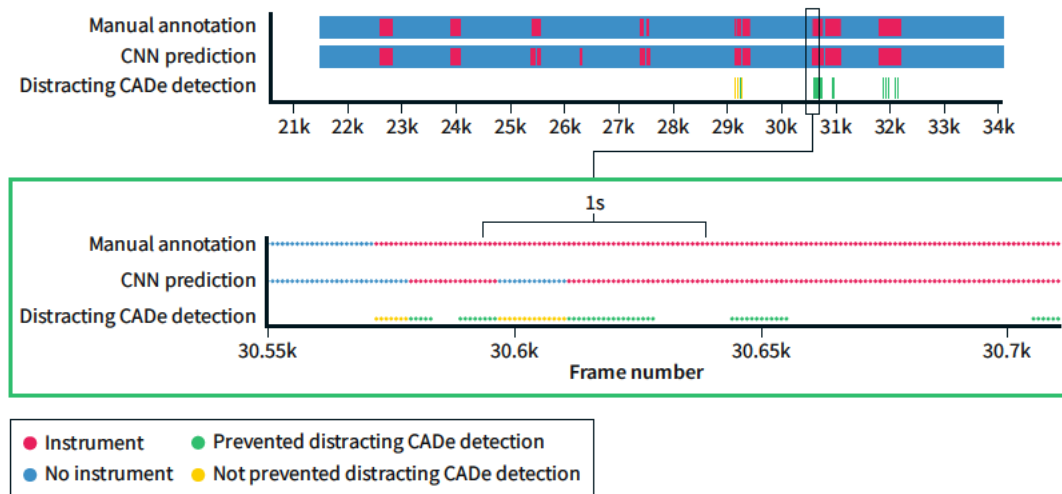
FIGURE 4 Schematic overview of the images with (red) and without (blue) visible instruments in a coloscopy video. The first row represents the manual annotations of whether the corresponding image contains a visible instrument. The second row represents the predictions output by our CNN. The third row represents the distracting CADe activations successfully prevented (green) or unsuccessfully prevented (yellow) by using the developed instrument detection CNN. The inset shows 160 frames (one dot per frame) which correlate to 5.33 s in the video. CADe, Computer-aided detection system; CNN, convolutional neuronal network
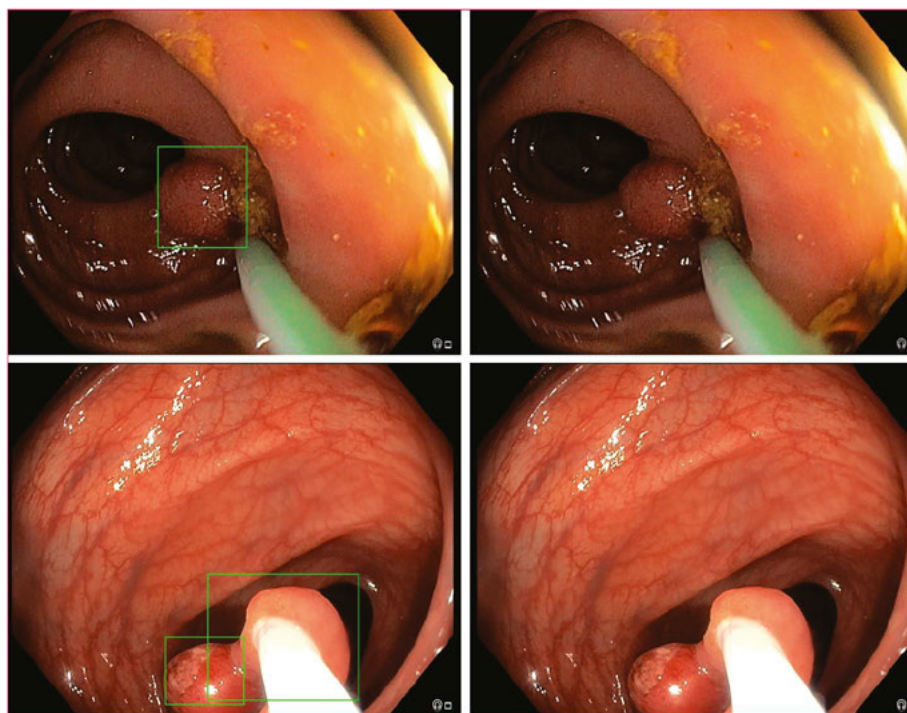

FIGURE 5 Single images of a polypectomy involving a needle for submucosal injection (upper row) and a snare (lower row) using the computer-aided polyp detection system (CADe) (left) and the additional CADe preventing instrument detection system (right). Video S1: Head-to-head comparison of a colonoscopy video sequence with (right) and without (left) the use of the instrument detection convolutional neuronal network

The first one is that the snare produces folds around the polyp by pressing the wire against the mucosa. Those folds are then falsely interpreted as polyps by the commercially available CADe system. Another explanation is that the total time that snares are visible in our dataset is longer than that of graspers. The longer visibility results in more CADe detections. Lastly, the not openly available training dataset of the commercially available CADe system might contain images of graspers and snares in an unbalanced manner.

The implementation of our freely available instrument detection AI system in an existing CADe system could be done by controlling the input signal of the examination monitor. Here, the instrument detection AI would analyze the raw endoscopy processor output

signal in parallel with the CADe system. During withdrawal with no visible instrument, the CADe signal would be displayed to the examiner. Thus, allowing the examiner to fully benefit from the high polyp detection rate of a commercially available CADe system. When an instrument is detected by the new AI, an automatic switch would display the raw processor signal instead. Alternatively, the AI system for instrument recognition can also be integrated as a filter directly into an existing or newly developed CADe system. By integrating our system in a single CADe system, the process would be more comfortable.

To the best of our knowledge, an AI system for instrument detection using deep learning methods has not been developed in gastrointestinal endoscopy. Therefore, we present the first AI system in the field that enables recognition of endoscopic interventions by instrument detection. In addition to the mentioned application, the system could also be useful for automated recording of intervention times or withdrawal time. This could potentially help in obtaining objective data to assess the quality of colonoscopies.[16]

However, our study has several limitations. The new AI system was evaluated using previously recorded videos. Therefore, the mentioned implementation of the AI system in daily practice must be tested in future prospective studies to evaluate clinical benefit. Another possibility would be to investigate (e.g. by eye-tracking) whether the examiner's attention could be better focused on the intervention by reducing distracting CADe signals.[17] Other limitations include that, to facilitate the generation of a large-annotated training dataset in a short period of time, no predefined protocol was used for video selection. A quantitative identification of the causes of false positive detections of our CNN need to be evaluated in future studies.

CADe systems already achieved a remarkable benefit in randomized controlled trials. Future developments of those systems include improving usability by adding customizable features. The commercially available system that was used in our study for example, presents a well-studied CADe function.[5] Other systems incorporate computer-aided diagnosis (CADx) that is only turned on when virtual chromoendoscopy is activated by the examiner.[18] In the case of the ENDO-AID CADe system by Olympus the examiner has the possibility to choose how many CADe detection boxes should be maximally displayed on the screen. The examiner can even choose from two different CADe modes that presumably present different sensitivities.[19] In other words, in general, these devices are incorporating customizable modules that increase the usability and, therefore, the value of the device. Our work aligns in this direction by contributing to the prevention of distracting CADe activations during interventions.

In conclusion, our study shows that instrument detection using AI technology is reliable and achieves high sensitivity and specificity. Therefore, the new AI system could be helpful to reduce distracting CADe detections during endoscopic procedures. Although the clinical benefit of the new AI system needs further evaluation, our study demonstrates the great potential of AI technology beyond mucosal assessment.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Markus Brand https://orcid.org/0000-0002-3495-5206
Joel Troya https://orcid.org/0000-0002-7992-0146
Adrian Krenzer https://orcid.org/0000-0002-1593-3300
Alexander Hann https://orcid.org/0000-0001-8035-3559

REFERENCES

1. Wang P, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. Lancet Gastroenterol & Hepatol. 2020;5(4):343–51. https://doi.org/10.1016/s2468-1253(19)30411-x
2. Wang P, Liu P, Glissen Brown JR, Berzin TM, Zhou G, Lei S, et al. Lower adenoma miss rate of computer-aided detection-assisted colonoscopy vs routine white-light colonoscopy in a prospective tandem study. Gastroenterology. 2020;159(4):1252–61:e1255. https://doi.org/10.1053/j.gastro.2020.06.023
3. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut. 2019;68(10):1813–19. https://doi.org/10.1136/gutjnl-2018-317500
4. Su JR, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). Gastrointest Endosc. 2020;91(2):415–24:e414. https://doi.org/10.1016/j.gie.2019.08.026
5. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology. 2020;159(2):512–20:e517. https://doi.org/10.1053/j.gastro.2020.04.062
6. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. Gastrointest Endosc. 2021;93(1):77–85:e76. https://doi.org/10.1016/j.gie.2020.06.059
7. Hassan C, Badalamenti M, Maselli R, Correale L, Iannone A, Radaelli F, et al. Computer-aided detection-assisted colonoscopy: classification and relevance of false positives. Gastrointest Endosc. 2020; 92(4):900–4:e904. https://doi.org/10.1016/j.gie.2020.06.021

8. Holzwanger EA, Bilal M, Glissen Brown JR, Singh S, Becq A, Ernest-Suarez K, et al. Benchmarking definitions of false-positive alerts during computer-aided polyp detection in colonoscopy. Endoscopy. 2021;53(09):937–40. https://doi.org/10.1055/a-1302-2942

9. Hann A, Troya J, Fitting D. Current status and limitations of artificial intelligence in colonoscopy. United European Gastroenterol J. 2021;9(5):527–33. https://doi.org/10.1002/ueg2.12108

10. Ebigbo A, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA, Jr., et al. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. Gut. 2020;69(4):615–6. https://doi.org/10.1136/gutjnl-2019-319460

11. Ueyama H, Kato Y, Akazawa Y, Yatagai N, Komori H, Takeda T, et al. Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. J Gastroenterol Hepatol. 2021;36(2):482–9. https://doi.org/10.1111/jgh.15190

12. Misawa M, Kudo SE, Mori Y, Hotta K, Ohtsuka K, Matsuda T, et al. Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). Gastrointest Endosc. 2021;93(4):960–7:e963. https://doi.org/10.1016/j.gie.2020.07.060

13. Wang P, Xiao X, Glissen Brown JR, Berzin TM, Tu M, Xiong F, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng. 2018;2(10):741–8. https://doi.org/10.1038/s41551-018-0301-3

14. Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology. 2018;155(4):1069–78:e1068. https://doi.org/10.1053/j.gastro.2018.06.037

15. Lu Z, Xu Y, Yao L, Zhou W, Gong W, Yang G, et al. Real-time automated diagnosis of colorectal cancer invasion depth using a deep learning model with multimodal data (with video). Gastrointest Endosc. 2021. https://doi.org/10.1016/j.gie.2021.11.049

16. Kaminski MF, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. United European Gastroenterol J. 2017;5(3):309–34. https://doi.org/10.1177/2050640617700014

17. Troya J, Fitting D, Brand M, Sudarevic B, Kather JN, Meining A, et al. The influence of computer-aided polyp detection systems on reaction time for polyp detection and eye gaze. Endoscopy. 2022. https://doi.org/10.1055/a-1770-7353

18. Weigt J, Repici A, Antonelli G, Afifi A, Kliegis L, Correale L, et al. Performance of a new integrated computer-assisted system (CADe/CADx) for detection and characterization of colorectal neoplasia. Endoscopy. 2022;54(02):180–4. https://doi.org/10.1055/a-1372-0419

19. Koo CS, Dolgunov D, Koh CJ. Key tips for using computer-aided diagnosis in colonoscopy - observations from two different platforms. Endoscopy. 2021. https://doi.org/10.1055/a-1701-6201

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brand M, Troya J, Krenzer A, Saßmannshausen Z, Zoller WG, Meining A, et al. Development and evaluation of a deep learning model to improve the usability of polyp detection systems during interventions. United European Gastroenterol J. 2022;10(5):477–84. https://doi.org/10.1002/ueg2.12235