Impact of Artificial Intelligence in Colorectal Polyp Detection and Characterization

Shivaraj Afzalpurkar¹ Mahesh K. Goenka² Rakesh Kochhar³

¹Department of Gastroenterology, Nanjappa Multi-speciality Hospitals, Davangere, Karnataka, India
²Director and Head, Institute of Gastrosciences and Liver, Apollo Multi-speciality Hospitals, Kolkata, West Bengal, India
³Department of Gastroenterology, NIMS University, Jaipur, India

Abstract
Colorectal cancer (CRC) is the third most common cancer in the world. Colonoscopy has contributed significantly to reduction of incidence and mortality of CRC. Integration of artificial intelligence (AI) into colonoscopy practice has addressed the various shortcomings of screening colonoscopies. AI-assisted colonoscopy will help in real-time recognition of type of polyp with probable histology. This will not only save time but will also help to mitigate human errors. Computer-aided detection and computer-aided characterization are two applications of AI, which are being studied extensively with a goal of improvement of polyp and adenoma detection rates. Several studies are being conducted across the globe, which either involve simple decision-making algorithms or complex patterns through neural networks, which imitate the human brain. Most data are collected retrospectively and the research is limited to single-center studies, which might have bias. Therefore, the future research on AI in colonoscopy should aim to develop more sophisticated convolutional neural network and deep learning models that will help to standardize the practice and ensure the same degree of accuracy with all the colonoscopies, irrespective of experience of performing endoscopists. In this review, we will take a closer look at the current state of AI and its integration into the field of colonoscopy.

Keywords
► artificial intelligence
► colonoscopy
► computer-aided detection
► computer-aided diagnosis

Introduction
According to the Global Cancer Statistics 2023, colorectal cancer (CRC) is the third most common cancer with annual incidence of 153,020 new cases and second leading cause of cancer-related mortality in both the sexes in the United States.¹ Widespread utilization of colonoscopy for polyp detection and removal has not only helped in reducing the incidence of CRC in last three decades but also contributed significantly in reducing CRC-related mortality.² Approximately one in four people with CRC have had a prior polyp.³ The key determining factor in detection of adenoma is the quality of the colonoscopy, which is operator-dependent.⁸,⁹ Several measures like optimal bowel preparation, colonoscopy withdrawal time of at least 6 minutes, improvement in cecal intubation rate have been employed in last few years to improve the adenoma detection rate (ADR).¹⁰,¹¹ Couple of studies have shown that up to 27% of adenomas are missed even with meticulous colonoscopy, especially those with size less than 10 mm.¹²,¹³ Every 1% improvement in ADR results in 3% reduction in CRC after colonoscopy.¹⁴

Artificial Intelligence (AI) is being commonly utilized in various aspects of health care system, which supports the medical professionals in diagnosis and treatment of various...
several studies have been published (developed software utilizing CADe and CADx systems and evaluated to detect the polyps and improve ADR. Human errors on visual diagnosis, CADe technology is being easily missed during routine colonoscopies. To mitigate the than 5 mm) and require very careful examination and may be diminutive (less than 5 mm) and require very careful examination and may be usually missed during routine colonoscopies. To mitigate the human errors on visual diagnosis, CADe technology is being evaluated to detect the polyps and improve ADR.

Brain. To ensure accurate results, the problem must be represented by training data. The particular data like images or videos of the polyps are fed into the AI system to predict the same pattern or relevant images, which will help in prediction, detection, or classification of the images. In these conventional ML methods, a researcher will have to manually add the images to train the AI system for identification of similar images or tasks. However, in the advanced ML method, which is called as deep learning, features of the polyps are automatically extracted by the algorithm without any manual input by the researcher. Deep learning is composed of multiple layered neural network and act as a subset of ML. Here multiple datasets are processed and evaluated several times in different layers (which are hidden) till the final output is reached. The evaluation of each layer depends on the output of previous layer. For example, let us assume that we give an input of colonoscopy image with an aim to look for characteristics of polyp to an AI model. The process begins with multiplication of image followed by scanning using different filters. Scoring system is used to score each layer of filters for image color, border, edges, vascular pattern, pit pattern, etc. The output score of each layer will be the input for next layer and final output depicting the type of polyp (benign/malignant or hyperplastic/adenomatous, etc.) is achieved. The term deep learning is used since the workflow involves multiple layers.

During the development of CADe system, numerous images of polyps and real-time videos are collected and used to train this AI system and can then be used in real-time during the colonoscopy procedure for detection of polyps. The outline of development of CADe system is presented by training data. The particular data like images or videos of the polyps are fed into the AI system to predict the same pattern or relevant images, which will help in prediction, detection, or classification of the images. In these conventional ML methods, a researcher will have to manually add the images to train the AI system for identification of similar images or tasks. However, in the advanced ML method, which is called as deep learning, features of the polyps are automatically extracted by the algorithm without any manual input by the researcher. Deep learning is composed of multiple layered neural network and act as a subset of ML. Here multiple datasets are processed and evaluated several times in different layers (which are hidden) till the final output is reached. The evaluation of each layer depends on the output of previous layer. For example, let us assume that we give an input of colonoscopy image with an aim to look for characteristics of polyp to an AI model. The process begins with multiplication of image followed by scanning using different filters. Scoring system is used to score each layer of filters for image color, border, edges, vascular pattern, pit pattern, etc. The output score of each layer will be the input for next layer and final output depicting the type of polyp (benign/malignant or hyperplastic/adenomatous, etc.) is achieved. The term deep learning is used since the workflow involves multiple layers.

During the development of CADe system, numerous images of polyps and real-time videos are collected and used to train this AI system and can then be used in real-time during the colonoscopy procedure for detection of polyps. – Fig. 1 shows the outline of development of CADe system.

### Computer-Aided Colonic Polyp Detection System

In one of the first studies on CADe-based system, Karkanis et al. proposed an approach of polyp detection based on a color feature extraction scheme built on the wavelet decomposition. The whole methodology was applied on the real datasets of colonoscopic videos and the performance of adenomatous polyp detection was estimated and was found to have a sensitivity and specificity of 93.6 and 99.3%, respectively. Most of the polyps in large intestine are diminutive (less than 5 mm) and require very careful examination and may be easily missed during routine colonoscopies. To mitigate the human errors on visual diagnosis, CADe technology is being evaluated to detect the polyps and improve ADR.

### Development of Computer-Aided Detection System

Machine learning (ML) is a technique of deciphering a decision in uncertain conditions using mathematical algorithms, which are automatically built from given data. In simpler words, it is an ability of a computer to learn from experience. This process either involves simple decision-making algorithms or complex patterns through neural networks, which imitate the human brain. To ensure accurate results, the problem must be represented by training data. The particular data like images or videos of the polyps are fed into the AI system to predict the same pattern or relevant images, which will help in prediction, detection, or classification of the images. In these conventional ML methods, a researcher will have to manually add the images to train the AI system for identification of similar images or tasks. However, in the advanced ML method, which is called as deep learning, features of the polyps are automatically extracted by the algorithm without any manual input by the researcher. Deep learning is composed of multiple layered neural network and act as a subset of ML. Here multiple datasets are processed and evaluated several times in different layers (which are hidden) till the final output is reached. The evaluation of each layer depends on the output of previous layer. For example, let us assume that we give an input of colonoscopy image with an aim to look for characteristics of polyp to an AI model. The process begins with multiplication of image followed by scanning using different filters. Scoring system is used to score each layer of filters for image color, border, edges, vascular pattern, pit pattern, etc. The output score of each layer will be the input for next layer and final output depicting the type of polyp (benign/malignant or hyperplastic/adenomatous, etc.) is achieved. The term deep learning is used since the workflow involves multiple layers.

During the development of CADe system, numerous images of polyps and real-time videos are collected and used to train this AI system and can then be used in real-time during the colonoscopy procedure for detection of polyps. – Fig. 1 shows the outline of development of CADe system.

### Key Studies on Computer-Aided Detection System

Many hand-crafted AI models were developed using predefined feature of polyps like color, texture, size, etc., that had a sensitivity of up to 90%. Years later, the real-time CADe

### Table 1 Published trials on computer-aided detection and computer-aided characterization

<table>
<thead>
<tr>
<th>System type</th>
<th>Company</th>
<th>Publication year</th>
<th>Author</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>CADe</td>
<td>CAD-EYE (Fujifilm)</td>
<td>2023</td>
<td>Rondonotti et al.</td>
<td>Italy</td>
</tr>
<tr>
<td>CADe</td>
<td>GI genius (Medtronic)</td>
<td>2022</td>
<td>Mangas-Sanjuan et al.</td>
<td>Spain</td>
</tr>
<tr>
<td>CADe</td>
<td>GI genius (Medtronic)</td>
<td>2022</td>
<td>Hassan et al.</td>
<td>Italy</td>
</tr>
<tr>
<td>CADe</td>
<td>SKOUT</td>
<td>2022</td>
<td>Shaukat et al.</td>
<td>United States</td>
</tr>
<tr>
<td>CADe</td>
<td>CAD-EYE (Fujifilm)</td>
<td>2022</td>
<td>Rondonotti et al.</td>
<td>Italy</td>
</tr>
<tr>
<td>CADx</td>
<td>Endo Brain</td>
<td>2021</td>
<td>Barua et al.</td>
<td>Norway, UK, Japan</td>
</tr>
<tr>
<td>CADe</td>
<td>EndoScreener</td>
<td>2021</td>
<td>Wang et al.</td>
<td>China</td>
</tr>
<tr>
<td>CADe</td>
<td>Yolo V3</td>
<td>2021</td>
<td>Kamba et al.</td>
<td>Japan</td>
</tr>
<tr>
<td>CADe</td>
<td>Endoangel</td>
<td>2021</td>
<td>Yao et al.</td>
<td>China</td>
</tr>
<tr>
<td>CADe</td>
<td>Self-developed</td>
<td>2021</td>
<td>Xu et al.</td>
<td>China</td>
</tr>
<tr>
<td>CADe</td>
<td>Self-developed</td>
<td>2020</td>
<td>Su et al.</td>
<td>China</td>
</tr>
<tr>
<td>CADe</td>
<td>Henan Tongyu</td>
<td>2020</td>
<td>Liu et al.</td>
<td>China</td>
</tr>
</tbody>
</table>

Abbreviations: CADe, computer-aided detection; CADx, computer-aided characterization.

*One latest key trial on CADe or CADx system by each company has been shown in the table.*
system using convolutional neural network (CNN) for polyp detection was used by Urban et al. In CNN consists of convolutional and pooling layers and fully connected layers, making it a specific class of artificial neural networks, which helps in making overall classification. The CNN system developed by Urban et al used representative set of 8,641 images from 2,000 patients and on testing models involving 20 colonoscopy videos, it identified the polyps with an accuracy of 96.4% and receiver operating characteristic curve of 0.991. In the recent cohort study conducted in Singapore by Koh et al, performance of individual endoscopists with AI was compared with their baseline performance without AI. It was found that there was 72.2% improvement in ADR with real-time AI-assisted colonoscopy even for experienced endoscopists. The incorporation of AI-assisted colonoscopy resulted in median 8.5% increase in ADR of individual endoscopists. However, short duration and limited sample size were the major limitations of this study. In one of the recently conducted randomized controlled trials (RCT) involving 660 patients by Repici et al, it was found that the ADR was higher in CADe group in comparison to control group (53.3 vs. 44.5%; relative risk: 1.22; 95% confidence interval: 1.04–1.40; \( p < 0.01 \)). Data from this RCT were compared with that of the previously conducted similar RCT involving six experienced endoscopists. It was found that experience played a minor role as determining factor for ADR. In a similar RCT by Wang et al, it was found that colonoscopy with CADe had better ADR especially for detecting diminutive adenoma, in comparison to the control group (29 vs. 20%). However, there was no significant difference for detecting polyps larger than 5 mm and also there were 39 false positive alarms for the polyps.

In a meta-analysis conducted in 2021, it was found that ADR for small and nonadvanced adenomas improved with AI-based colonoscopy in comparison to non-AI-based colonoscopy (29.6 vs. 19.3%), but there was no significant difference for advanced adenoma. Another recently conducted meta-analysis of RCTs concluded that adenomas and sessile serrated lesions detected by colonoscopy were significantly higher in AI-assisted group in comparison to routine colonoscopy.

**Limitations of Computer-Aided Detection System**

Having false positive alarm during colonoscopy is an important limitation of CADe technology. Previous studies have reported up to an average of 27.3 false positive alarms per colonoscopy. Another limitations of the studies based on CADe system is that majority of the control groups had ADR below 23%, which is lower than the recommended ADR of ≥25%. 

**Computer-Assisted Diagnosis System**

Polyp characterization or classification is equally important once the polyp detection is made. This helps in avoiding unnecessary polypectomies of small or low-risk polyps and also reduces the financial burden on patients by reducing the unnecessary histopathological examinations. The research interest has increased significantly on developing and validating CADx system. This will help us to adopt “diagnose and leave” or “resect and discard” strategy during colonoscopy wherever appropriate. Initially, retrospective studies on application of CADx system used narrow band imaging (NBI), endocytoscopy, autofluorescence imaging (AFI) endoscopy, confocal endomicroscopy, magnifying chromoendoscopy or white light (WL) imaging endoscopy. Tischendorf et al in 2009 and Gross et al in 2011 used magnifying NBI for developing CADx system and reported a diagnostic accuracy of approximately 85 and 93%, respectively. There is limited number of prospective studies on CADx system with utilization of NBI or AFI endoscopy.

A novel intelligent device for optical characterization of colonic polyps in the real time was proposed recently by Biffi et al. This device works on conventional WL endoscopy and without utilization of blue light (chromoendoscopy). Endoscopists were divided into experts (\( n = 10 \)) and nonexperts (\( n = 11 \)) and CADx-classified polyps (513) into “adenoma” or “nonadenoma.” Using histopathology as reference standard, it was found that the accuracy of CADx in WL was comparable to the accuracy of expert endoscopists and superior to that of nonexpert endoscopists. Therefore, this device has potential to support the nonexpert endoscopists to reach the performance of expert endoscopists in characterization of colorectal polyps.

In another prospective pilot study, CADx model was built by utilizing three polyp features like vessel length, vessel circumference, and brightness in the blood vessels. The data were extracted from colonoscopies with NBI. It was found that the sensitivity and specificity of the model for accurate polyp classification was 96.9 and 71.4%, respectively. However, the sensitivity and specificity for accurate prediction of polyp histology by endoscopists were 93.8 and 85.7%, respectively. The major limitation of this study was that the AI algorithm was not fully automated.
Real-time decision outputs from the CADx system is an important advancement in clinical research on screening colonoscopy. Horiuchi et al.38 included 95 patients with 258 rectosigmoid diminutive polyps and reported an accuracy of 91.5% for predicting neoplastic polyps. AFI is a technique in which light elicits the endogenous fluorophores from the mucosa leading to production of green or red image, which is analyzed by a computer. Aihara et al.39 included 32 patients with 102 colonic lesions and reported a sensitivity and specificity of 94.2 and 88.9% for discriminating neoplastic from benign lesions. Both these studies used AFI-based CADx system.38,39 There are a few prospective studies that have shown poor and unfavorable diagnostic performance during real-time application of CADx models. Kuiper et al.40 showed that their AI model did not have sufficient power to differentiate between adenomatous and nonadenomatous polyps. Another model by Rath et al.41 had a sensitivity and specificity of only 81.8 and 85.2%, respectively. EndoBRAIN is the novel AI technology that helps in real-time differentiation of neoplastic and non-neoplastic polyps by analyzing vessel pattern, cell nuclei, and crypt structure.42 Katrevula45 in their recent observational pilot study have demonstrated better diagnostic performance of EndoBRAIN in comparison to endoscopists using endoscopy for predicting neoplastic lesions. – Table 2 summarizes the key published studies on CADx system.

**Simultaneous Use of Computer-Aided Detection and Computer-Aided Characterization System**

Detection and classification of polyps simultaneously by AI system is the emerging field of research interest. There are limited studies on AI models that help in simultaneous polyp detection and characterization. Ozawa et al.47 developed a CNN-based model, which can detect the polyps on WL or NBI with a sensitivity of 92% and then characterize the polyp with an accuracy of 83%. Weigt et al.48 in their recent study developed a new integrated CADe/CADx system using multcenter library of about two million images from 1,572 polyps and found its sensitivity, specificity, and accuracy to be 92.9, 90.6, and 91.7%, respectively. There are many pitfalls in application of AI system during colonoscopy as it may not be cost-effective in majority of the centres, likelihood of false positive detections, and undue increase in the duration of procedure. A recent RCT has shown that CADe did not improve the colonoscopic detection of advanced colorectal lesions.49 Too much relying on advanced technology may lead in deskilling of endoscopists as they decrease he use of their clinical skills and experience.50 Extensive research is needed to implement the models utilizing both CADe and CADx in a single setting in real-world practice.

**Conclusion**

AI is the future of modern medicine. The integration of AI into the field of gastroenterology especially into the screening colonoscopy in particular for polyp detection and characterization has shown promising results. Most data are collected retrospectively and the research is limited to single-center studies. This model will not only lead to a development of an algorithm, which cannot be generalized to other datasets but may lead to bias. Therefore, the future research of AI in colonoscopy should include the development of more sophisticated CNN and deep learning models and studies on real-time prospective data through multicenter trials. Conducting RCTs in multicenter setting will not only give an accurate algorithm but may also help in overcoming various pitfalls. This will standardize the practice and ensure the same accuracy with all the colonoscopies, irrespective of experience of performing endoscopists.

**Conflict of Interest**

None declared.

**References**

Artificial Intelligence in Colorectal Polyp Detection

Azalpurkar et al.


Katrevula A. Real-world experience of ai-assisted endocytoscopy using EndoBRAIN—an observational study from a tertiary care center. J Dig Endosc 2023;14:3–7


