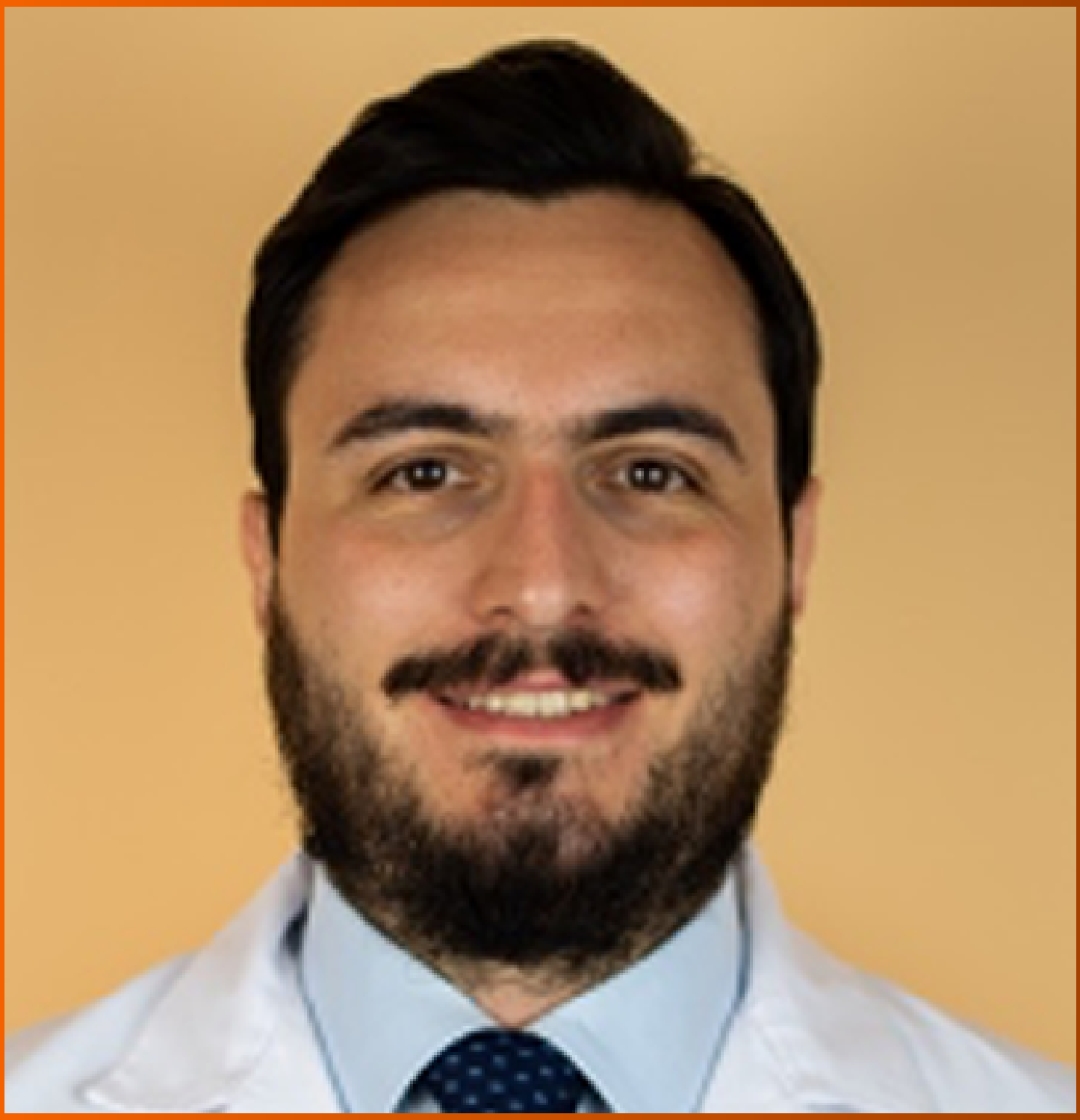
ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2020 October 28; 1(2): 28-43



Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 1 Number 2 October 28, 2020

EDITORIAL

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 1 Number 2 October 28, 2020

Contents

ABOUT COVER

Editorial board member of *Artificial Intelligence in Gastrointestinal Endoscopy*, Dr. Marcello Maida received his medical degree in 2009, summa cum laude, at the University of Palermo (Italy), where he went on to perform his clinical training in gastroenterology. He is currently a consultant gastroenterologist at S. Elia - Raimondi Hospital in Caltanissetta (Italy), a member of the governing council of the Young Italian Gastroenterologists Association, and a manager of Gastrolearning, the Multimedia Educational Platform of Italian Medical Schools of Gastroenterology. Dr. Maida's research interests span several areas of gastroenterology, including endoscopy and liver diseases. He is author of several peer-reviewed publications indexed on PubMed, including some of the best referral journals in gastroenterology. (L-Editor: Filipodia)

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE*, *Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
June 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Krish Ragunath, Sahin Coban, Fatih Altintoprak

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
October 28, 2020

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

E-mail: bpgoffice@wjgnet.com https://www.wjgnet.com

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2020 October 28; 1(2): 28-32

DOI: 10.37126/aige.v1.i2.28

ISSN 2689-7164 (online)

EDITORIAL

Artificial intelligence in Barrett's esophagus: A renaissance but not a reformation

Karen Chang, Christian S Jackson, Kenneth J Vega

ORCID number: Karen Chang 0000-0002-1523-1587; Christian S Jackson 0000-0003-4229-4206; Kenneth J Vega 0000-0002-2432-6123.

Author contributions: All authors planned the editorial and reviewed all data included; all authors wrote the editorial and revised it for intellectual content; and approved the final submitted version; Vega KJ is the editorial guarantor.

Conflict-of-interest statement: All authors have no conflict of interests to declare.

Manuscript source: Invited manuscript

Received: October 10, 2020
Peer-review started: October 10,

Karen Chang, Department of Internal Medicine, University of California, Riverside School of Medicine, Riverside, CA 92521, United States

Christian S Jackson, Gastroenterology Section, VA Loma Linda Healthcare Syst, Loma Linda, CA 92357, United States

Kenneth J Vega, Division of Gastroenterology and Hepatology, Department of Medicine, Augusta University-Medical College of Georgia, Augusta, GA 30912, United States

Corresponding author: Kenneth J Vega, MD, Professor, Division of Gastroenterology and Hepatology, Department of Medicine, Augusta University-Medical College of Georgia, 1120 15th Street, AD 2226, Augusta, GA 30912, United States. kvega@augusta.edu

Abstract

Esophageal cancer remains as one of the top ten causes of cancer-related death in the United States. The primary risk factor for esophageal adenocarcinoma is the presence of Barrett's esophagus (BE). Currently, identification of early dysplasia in BE patients requires an experienced endoscopist performing a diagnostic endoscopy with random 4-quadrant biopsies taken every 1-2 cm using appropriate surveillance intervals. Currently, there is significant difficulty for endoscopists to distinguish different forms of dysplastic BE as well as early adenocarcinoma due to subtleties in mucosal texture and color. This obstacle makes taking multiple random biopsies necessary for appropriate surveillance and diagnosis. Recent advances in artificial intelligence (AI) can assist gastroenterologists in identifying areas of likely dysplasia within identified BE and perform targeted biopsies, thus decreasing procedure time, sedation time, and risk to the patient along with maximizing potential biopsy yield. Though using AI represents an exciting frontier in endoscopic medicine, recent studies are limited by selection bias, generalizability, and lack of robustness for universal use. Before AI can be reliably employed for BE in the future, these issues need to be fully addressed and tested in prospective, randomized trials. Only after that is achieved, will the benefit of AI in those with BE be fully realized.

Key Words: Barrett's esophagus; Artificial intelligence; Machine learning; Cognitive neural networks; Computer aided diagnosis; Endoscopy

2020
First decision: October 22, 2020
Revised: October 26, 2020
Accepted: October 27, 2020
Article in press: October 27, 2020
Published online: October 28, 2020

P-Reviewer: Yu CH
S-Editor: Wang JL
L-Editor: A
P-Editor: Ma YJ

Core Tip: Screening and surveillance in patients with Barrett's esophagus (BE) remain problematic in regards to accuracy and adherence. This occurs in spite of recommendations and advances in endoscopic imaging. Artificial intelligence (AI) algorithms assist in endoscopic evaluation of BE by identifying potential targets for biopsy. This may occur by increasing endoscopic efficiency and diagnosing accuracy by decreasing procedure time. AI in BE has been developed by expert endoscopists and appear to perform similarly among them. At this point, the benefit of AI in BE may be for use by non-expert endoscopists and trainees to maximize BE endoscopic evaluation.

Citation: Chang K, Jackson CS, Vega KJ. Artificial intelligence in Barrett's esophagus: A renaissance but not a reformation. *Artif Intell Gastrointest Endosc* 2020; 1(2): 28-32
URL: https://www.wjgnet.com/2689-7164/full/v1/i2/28.htm
DOI: https://dx.doi.org/10.37126/aige.v1.i2.28

INTRODUCTION

In 2020, the United States is estimated to record over 18000 new esophageal cancer cases and over 16000 deaths[1]. Furthermore, esophageal cancer remains in the top ten of cancers diagnosed and cause of cancer related death nationally. One common risk factor for esophageal adenocarcinoma (EAC) is the presence of Barrett's esophagus (BE). Currently, identification of early dysplasia requires an experienced endoscopist performing a diagnostic endoscopy consisting of random 4-quadrant biopsies to be taken every 1-2 cm within appropriate surveillance intervals based on absence or presence of dysplasia seen in the random biopsies[2-5]. Unfortunately, adherence to this recommendation remains inconsistent, particularly with low-grade dysplasia. Its subtle appearance and discontinuous nature can make it difficult to accurately biopsy areas for tissue pathology to confirm or rule out the diagnosis. In addition, there is significant difficulty for endoscopists to distinguish BE with low-grade dysplasiafrom high-grade dysplasia (HGD) or early adenocarcinoma. To combat this, high-definition white light, narrow band imaging (NBI), probe-based confocal endomicroscopy (pCLE), volumetric laser endomicroscopy (VLE) and optical computed tomography among others have all been tested and employed an in attempt to increase biopsy yield for accurate diagnosis[6-9]. However, early EAC is often flat and difficult to distinguish from the surrounding non-dysplastic Barrett's mucosa, even with these endoscopic advances. The rate-limiting step among of these technologies is that they are operator dependent, requiring hand-eye coordination to distinguish and biopsy suspicious areas, often-taking years to acquire the necessary skill set. Theoretically, artificial intelligence (AI) can assist in this by using methods of deep learning to identify and process - in real-time - endoscopic data that may not consciously appreciated by humans such as subtle changes in color and texture to aid in taking targeted biopsies rather than random biopsies.

There have been recent advances in the development and testing of AI and various machine learning (ML) algorithms to improve the ability to identify dysplastic and malignant mucosa. Previously, computer algorithms were trained to classify a patient's likelihood for EAC based on symptoms or compare patient biopsy cDNA microarrays to known EAC samples. These methods drew us closer to accurately diagnosing dysplasia and malignant mucosa, but their sensitivities/specificities could not match the parameters outlined in American Society for Gastrointestinal Endoscopy's Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) criteria for new technologies. PIVI criteria recommends that the sensitivity should be at least 0.90, specificity should be at least 0.80 and a negative predictive value of at least 0.98 for detecting HGD or BE[10]. AI makes use of several methods of ML. One commonly used method is the cognitive neural network (CNN). In CNN, each node (or "neuron") is connected to other nodes in a way that mimics real human neural networking. Several layers of neurons can exist to make a single decision to call a grouping of pixels on an image either normal tissue or dysplasia. Multiple recent studies have already experimented with the capabilities of such computer-aided diagnosis (CAD) (Table 1). The advantages that AI appears to confer per-endoscopy is a removal of the inter-observer or intra-observer variability in identification of non-

Table 1 Computer-aided diagnosis of Barrett's esophagus

Ref.	Year	Study design	Lesions	Imaging modality	Image qualification	Teaching dataset	Validation method	Outcomes	Compared to expert/current standard
van der Sommen *et al*[11]	2016	Retrospective	HGD, early EAC	WLI	High quality, clear visible/absence of lesions	100 images	LOO	Per-image SPEC/SENS: 83%/83%; Per-patient SPEC/SENS: 86%/87%	Inferior
de Groof *et al*[12]	2019	Retrospective	Non-dysplastic and dysplastic BE	WLI	1280 × 1024 pixels – HD	60 images	LOO	Accuracy: 0.92; SENS: 0.95; SPEC: 0.85	NA
Swager *et al*[13]	2017	Retrospective	HGD, early EAC	VLE	High quality image database	60 images	LOO	AUC: 0.95, 0.89, 0.91	Superior
Ebigbo *et al*[15]	2020	Prospective	Early EAC	WLI	1350 × 1080 pixels and 1600 × 1200 pixels – HD	129 images	LOO	Accuracy: 0.899; SENS: 0.837; SPEC: 1.00	NA

AUC: Area under the curve; BE: Barrett's esophagus; EAC: Esophageal adenocarcinoma; HD: High definition; HGD: High-grade dysplasia; LOO: LEAVE-one-out; NA: Not available; SENS: Sensitivity; SPEC: Specificity; VLE: Volumetric laser endomicroscopy; WLI: White light imaging.

normal lesions, combined with rapid, objective analysis of all visual inputs in such a way that is consistent and not subject to fatigue. CAD can allow endoscopists to take targeted, high-yield biopsies in real-time. Compared to taking random biopsies per the Seattle protocol or using enhanced imaging, CAD may increase efficiency and accuracy for making a diagnosis by limiting the chance of missing neoplastic mucosa. Moreover, CAD may decrease risk by decreasing sedation time secondary to decreased procedure length.

Recent studies would indicate that CAD can be successful in the detection of neoplastic lesions in BE. Von Der Sommen *et al*[11] developed a ML algorithm that used CAD to analyze texture and color in static images to detect early neoplastic lesions in BE. The sensitivity and specificity were between 0.90 to 1.00 and 0.65 to 0.91 respectively. In a study by Groof *et al*[12], six experts identified likely neoplastic tissue in the same image and used these expert-delineated images to train the computer algorithm to identify neoplastic BE and non-dysplastic BE in test cases. The resulting sensitivity and specificity of the computer algorithm was 0.95 and 0.85 respectively. Swager *et al*[13] used CAD on *ex vivo* VLE images to retrospectively detect non-dysplastic BE and HGD or early adenocarcinoma. They were able to achieve a sensitivity of 0.90 and specificity of 0.93 while using VLE as the reference images rather than high-definition white light endoscopy.

Though the data is promising, nearly all research has focused on training an algorithm on a set of retrospectively gathered images. Because of this, these studies are unfortunately subject to selection bias since the images are often curated for high definition and typically from a single endoscopy center. Therefore, the algorithms are usually overtrained on a relatively small sample set and not generalizable to other images of poorer quality or a population with different incidence and/or prevalence of BE. A sparing number of prospective or real-time studies currently exist and these are performed on a rather small number of samples. Furthermore, standardization of AI systems is proving difficult, given that the details of the algorithm are in a "black box" and inaccessible to critique and direct modifications. The struggles that have been encountered in using AI for identification of Barrett's mucosa have been encountered in identifying early esophageal cancers. Though promising, the thresholds to detect early esophageal cancer are below PIVI criteria which may be secondary to limited images and lack of ability to identify images in real time. Hashimoto *et al*[14] may have found a way to overcome previous difficulties by being able to create a faster algorithm which allowed for a real time video overlay using a large database of images. Using this technique, Hashimoto *et al*[14] were able to identify early esophageal neoplasms with high accuracy.

The process of standardization of ML algorithms poses a difficult challenge. The algorithm may be different for white light endoscopy compared to NBI, VLE or pCLE. It is possible that subtle differences such as the brand of endoscope, wavelength of

light or white balance could impact specificity or sensitivity of a tested algorithm. There is no guarantee that a single algorithm would work both in populations of high prevalence of BE and populations of low prevalence. Ideally, several algorithms should be tested prospectively and compared to the current gold standard of random biopsy in large, multicenter randomized clinical trials. Some of these studies are currently ongoing. User databases such as ImageNet or GastroNet contain samples of labeled images for use for training and testing of algorithms, but there is need for databases of patients with varying prevalence of risk factors for BE to determine if a single algorithm is robust enough to accurately diagnose BE nationwide.

To date, the ML platforms used have been developed by expert endoscopists. A recent study published by Ebigbo *et al*[15] used real-time AI to identify cancer in BE and found that the AI system performed in a similar fashion to the expert endoscopist. Such programs can also help train non-experts and gastroenterology fellows alike by giving real-time feedback, thus propagating more expert endoscopists in a shortened timeframe. Of course, endoscopists who are not BE experts can also benefit as well.

CONCLUSION

AI represents a renaissance in endoscopy, but not a reformation. The benefit may lie in the improvement in recognition of dysplastic and malignant tissue among non-expert endoscopists or gastroenterology fellows, since expert endoscopists have similar performance to AI. Generalizability, robustness of a single or few algorithms that can apply to either different imaging modalities or diverse populations, and the ability to easily modify an algorithm are current obstacles that need to be addressed before we can reliably use AI in endoscopic management of BE.

REFERENCES

1 **American Cancer Society**. About Esophagus Cancer. American Cancer Society website. Updated March 20, 2020. Available from: URL: https://www.cancer.org/content/dam/CRC/PDF/Public/8614.00.pdf

2 **Vennalaganti PR**, Kaul V, Wang KK, Falk GW, Shaheen NJ, Infantolino A, Johnson DA, Eisen G, Gerson LB, Smith MS, Iyer PG, Lightdale CJ, Schnoll-Sussman F, Gupta N, Gross SA, Abrams J, Haber GB, Chuttani R, Pleskow DK, Kothari S, Goldblum JR, Zhang Y, Sharma P. Increased detection of Barrett's esophagus-associated neoplasia using wide-area trans-epithelial sampling: a multicenter, prospective, randomized trial. *Gastrointest Endosc* 2018; **87**: 348-355 [PMID: 28757316 DOI: 10.1016/j.gie.2017.07.039]

3 **Dalal KS**, Coffing J, Imperiale TF. Adherence to Surveillance Guidelines in Nondysplastic Barrett's Esophagus. *J Clin Gastroenterol* 2018; **52**: 217-222 [PMID: 27984402 DOI: 10.1097/MCG.0000000000000777]

4 **Holmberg D**, Ness-Jensen E, Mattsson F, Lagergren J. Adherence to clinical guidelines for Barrett's esophagus. *Scand J Gastroenterol* 2019; **54**: 945-952 [PMID: 31314608 DOI: 10.1080/00365521.2019.1641740]

5 **Roumans CAM**, van der Bogt RD, Steyerberg EW, Rizopoulos D, Lansdorp-Vogelaar I, Sharma P, Spaander MCW, Bruno MJ. Adherence to recommendations of Barrett's esophagus surveillance guidelines: a systematic review and meta-analysis. *Endoscopy* 2020; **52**: 17-28 [PMID: 31529444 DOI: 10.1055/a-0995-0134]

6 **Johnson EA**, De Lee R, Agni R, Pfau P, Reichelderfer M, Gopal DV. Probe-Based Confocal Laser Endomicroscopy to Guide Real-Time Endoscopic Therapy in Barrett's Esophagus with Dysplasia. *Case Rep Gastroenterol* 2012; **6**: 285-292 [PMID: 22754488 DOI: 10.1159/000338835]

7 **Everson MA**, Lovat LB, Graham DG, Bassett P, Magee C, Alzoubaidi D, Fernández-Sordo JO, Sweis R, Banks MR, Wani S, Esteban JM, Ragunath K, Bisschops R, Haidry RJ. Virtual chromoendoscopy by using optical enhancement improves the detection of Barrett's esophagus-associated neoplasia. *Gastrointest Endosc* 2019; **89**: 247-256. e4 [PMID: 30291849 DOI: 10.1016/j.gie.2018.09.032]

8 **Belderbos TDG**, van Oijen MGH, Moons LMG, Siersema PD. Implementation of real-time probe-based confocal laser endomicroscopy (pCLE) for differentiation of colorectal polyps during routine colonoscopy. *Endosc Int Open* 2017; **5**: E1104-E1110 [PMID: 29104910 DOI: 10.1055/s-0043-117948]

9 **Trindade AJ**, Inamdar S, Smith MS, Chang KJ, Leggett CL, Lightdale CJ, Pleskow DK, Sejpal DV, Tearney GJ, Thomas RM, Wallace MB. Volumetric laser endomicroscopy in Barrett's esophagus: interobserver agreement for interpretation of Barrett's esophagus and associated neoplasia among high-frequency users. *Gastrointest Endosc* 2017; **86**: 133-139 [PMID: 27899321 DOI: 10.1016/j.gie.2016.11.026]

10 **Sharma P**, Savides TJ, Canto MI, Corley DA, Falk GW, Goldblum JR, Wang KK, Wallace MB, Wolfsen HC; ASGE Technology and Standards of Practice Committee. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on imaging in Barrett's Esophagus. *Gastrointest Endosc* 2012; **76**: 252-254 [PMID: 22817781 DOI: 10.1016/j.gie.2012.05.007]

11 **van der Sommen F**, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, Bergman JJ, de With PH, Schoon EJ. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. *Endoscopy* 2016; **48**: 617-624 [PMID: 27100718 DOI: 10.1055/s-0042-105284]

12 **de Groof J**, van der Sommen F, van der Putten J, Struyvenberg MR, Zinger S, Curvers WL, Pech O, Meining A, Neuhaus H, Bisschops R, Schoon EJ, de With PH, Bergman JJ. The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019; **7**: 538-547 [PMID: 31065371 DOI: 10.1177/2050640619837443]

13 **Swager AF**, van der Sommen F, Klomp SR, Zinger S, Meijer SL, Schoon EJ, Bergman JJGHM, de With PH, Curvers WL. Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy. *Gastrointest Endosc* 2017; **86**: 839-846 [PMID: 28322771 DOI: 10.1016/j.gie.2017.03.011]

14 **Hashimoto R**, Requa J, Dao T, Ninh A, Tran E, Mai D, Lugo M, El-Hage Chehade N, Chang KJ, Karnes WE, Samarasena JB. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). *Gastrointest Endosc* 2020; **91**: 1264-1271. e1 [PMID: 31930967 DOI: 10.1016/j.gie.2019.12.049]

15 **Ebigbo A**, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA Jr, Papa J, Palm C, Messmann H. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. *Gut* 2020; **69**: 615-616 [PMID: 31541004 DOI: 10.1136/gutjnl-2019-319460]

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2020 October 28; 1(2): 33-43

DOI: 10.37126/aige.v1.i2.33

ISSN 2689-7164 (online)

MINIREVIEWS

Understanding deep learning in capsule endoscopy: Can artificial intelligence enhance clinical practice?

Amporn Atsawarungruangkit, Yousef Elfanagely, Akwi W Asombang, Abbas Rupawala, Harlan G Rich

ORCID number: Amporn Atsawarungruangkit 0000-0003-0622-6839; Yousef Elfanagely 0000-0002-3056-3811; Akwi W Asombang 0000-0002-8658-6772; Abbas Rupawala 0000-0001-5669-0686; Harlan G Rich 0000-0003-3538-7631.

Author contributions: Atsawarungruangkit A and Elfanagely Y contributed equally to this work including literature review, study selection, and manuscript writing; Asombang AW, Rupawala A, and Rich HG critically revised the manuscript and provided supervision.

Conflict-of-interest statement: The authors declare no conflict of interests.

Amporn Atsawarungruangkit, Akwi W Asombang, Abbas Rupawala, Harlan G Rich, Division of Gastroenterology, Warren Alpert School of Medicine, Brown University, Providence, RI 02903, United States

Yousef Elfanagely, Department of Internal Medicine, Brown University, Providence, RI 02903, United States

Corresponding author: Amporn Atsawarungruangkit, MD, Academic Fellow, Division of Gastroenterology, Warren Alpert School of Medicine, Brown University, 593 Eddy Street, POB 240, Providence, RI 02903, United States. amporn_atsawarungruangkit@brown.edu

Abstract

Wireless capsule endoscopy (WCE) enables physicians to examine the gastrointestinal tract by transmitting images wirelessly from a disposable capsule to a data recorder. Although WCE is the least invasive endoscopy technique for diagnosing gastrointestinal disorders, interpreting a WCE study requires significant time effort and training. Analysis of images by artificial intelligence, through advances such as machine or deep learning, has been increasingly applied to medical imaging. There has been substantial interest in using deep learning to detect various gastrointestinal disorders based on WCE images. This article discusses basic knowledge of deep learning, applications of deep learning in WCE, and the implementation of deep learning model in a clinical setting. We anticipate continued research investigating the use of deep learning in interpreting WCE studies to generate predictive algorithms and aid in the diagnosis of gastrointestinal disorders.

Key Words: Capsule endoscopy; Deep learning; Machine learning; Wireless capsule endoscopy; Small bowel capsule; Video capsule endoscopy

Core Tip: Wireless capsule endoscopy is the least invasive endoscopy technique for investigating the gastrointestinal tract. However, it takes a significant amount of time for interpreting the results. Deep learning has been increasingly applied to interpret capsule endoscopy images. We have summarized deep learning's framework, various

Manuscript source: Unsolicited manuscript

Received: September 21, 2020
Peer-review started: September 21, 2020
First decision: September 25, 2020
Revised: October 1, 2020
Accepted: October 13, 2020
Article in press: October 13, 2020
Published online: October 28, 2020

P-Reviewer: Kawabata H
S-Editor: Wang JL
L-Editor: A
P-Editor: Wang LL

characteristics in published literature, and application in the clinical setting.

Citation: Atsawarungruangkit A, Elfanagely Y, Asombang AW, Rupawala A, Rich HG. Understanding deep learning in capsule endoscopy: Can artificial intelligence enhance clinical practice? *Artif Intell Gastrointest Endosc* 2020; 1(2): 33-43
URL: https://www.wjgnet.com/2689-7164/full/v1/i2/33.htm
DOI: https://dx.doi.org/10.37126/aige.v1.i2.33

INTRODUCTION

Since 1868, endoscopy has been constantly evolving and improving to assess the lumen and mucosa of the gastrointestinal tract, including the esophagus, stomach, colon, and parts of the small bowel[1]. Despite its utility, endoscopic examination of the small intestine is limited by its length and distance from accessible orifices[2-4]. This limitation is a factor that contributed to the development of wireless capsule endoscopy (WCE).

Developed in the mid-1990s, WCE utilizes an ingestible miniature camera that can directly view the esophagus, stomach, entire small intestine, and colon without pain, sedation, or air insufflation[2,5-7]. An important clinical application of WCE is the evaluation of gastrointestinal bleeding after a high quality bidirectional conventional endoscopy and colonoscopy does not identify a source of bleeding[5]. A typical WCE study lasts 8 to 12 h and generates 50000-100000 images. Reviewing that quantity of images requires significant time effort and training. Additionally, abnormalities in the gastrointestinal tract may be present in only one or two frames of the video which may be missed due to oversight[2]. An automatic computer-aided diagnosis system may aid and support physicians in their analysis of images captured by WCE.

Artificial intelligence (AI), an aspect of computer-aided design, has been rapidly expanding and permeating in academia and industry[8]. AI involves computer programs that perform functions associated with human intelligence[9,10]. Specific features of AI include computer learning and problem solving. AI was first described as the development of computer systems to perform tasks that require human intelligence, which can include decision making and speech recognition[11]. Many techniques of AI have been proposed to facilitate the recognition and prediction of patterns[12].

Machine learning (ML) is an application of AI that provides systems with the ability to automatically learn and improve from experience without explicit programming[13]. ML can recognize patterns from datasets to create algorithms and make predictions[10,12]. A tremendous breakthrough in ML has been the development of deep neural networks (also known as deep learning)[13]. Deep learning consists of massive multilayer networks of artificial neurons that can automatically discover useful features. To put it simply, deep learning can extract more patterns from high dimensional data[5,12]. Several deep learning models have been reported in the literature and are differentiated by their application[12]. Convolutional neural network (CNN), a type of deep learning, is highly effective at performing image analysis[8,13,14]. Given CNN's utility in image analysis, applications for CNN have extended into the medical field, including gastroenterology[8,14]. The main drawback of deep learning is a long training time. Advances in graphic processing units, however, have drastically reduced the training time of deep learning from days or weeks to hours or days[15].

ML and CNN have been increasingly explored and applied to diagnostic images found in radiology, pathology, and dermatology[15-18]. Likewise, ML and CNN have utility in endoscopy and WCE through image-based interpretation without alteration of the existing procedures[8,11]. Current applications of ML and CNN in gastroenterology include polyp detection, esophageal cancer diagnosis, and ulcer detection through image-based interpretation from WCE. WCE is among the top interests of AI researchers in gastroenterology.

LITERATURE REVIEW

We conducted a literature review on December 15, 2019 and updated it on March 31, 2020 on PubMed/MEDLINE database and IEEE Xplore digital library. The search phrase used for query data in PubMed/MEDLINE database was ("Capsule Endoscopy") AND ("Deep Learning" OR "Neural Network" OR "Neural Networks"). Similarly, the search phrase used for query data in IEEE Xplore digital library was ("All Metadata":"Capsule Endoscopy") AND (("All Metadata":"Deep Learning") OR ("All Metadata":"Neural Network") OR ("All Metadata":"Neural Networks")). As presented in Figure 1, we found 50 records in PubMed/MEDLINE database and 71 records in IEEE Xplore digital library. After removing 14 duplicate records, the total number of distinct records were 107.

Only articles written in English language or available in English translation were considered. Conference abstracts, review articles, magazine articles, and unpublished studies were excluded to ensure quality. At this stage, two authors (AA and YE) independently reviewed whether the studies met the above inclusion criteria based on the title and abstract. Then, the articles that passed the initial screening were independently reviewed again based on the full-text articles to locate all included studies within a predefined scope of this article.

USE OF DEEP LEARNING FOR CLASSIFYING GASTROINTESTINAL DISORDERS

The most common indication for using WCE is the evaluation of small intestinal bleeding. WCE has also be used to diagnose other small intestinal disorders, such as celiac disease, Crohn's disease, polyps, and tumors, for the evaluation of esophageal pathology in non-cardiac chest pain, and for colon cancer screening. As shown in Table 1, previous studies have focused on the use of deep learning for classifying gastrointestinal diseases and lesions identified on WCE images. Unsurprisingly, a frequently investigated outcome in published literature is bleeding. Deep learning models have enhanced WCE's ability to detect bleeding lesions (including suspected blood content and angioectasia) with relatively high sensitivity and specificity[19-27]. In addition to bleeding, researchers have also used deep learning models in WCE to classify other gastrointestinal lesions such as ulcers[19-21,28-32], Crohn's disease[33], polyps[7,19-21,34], celiac disease[6], and hookworm[35].

Deep network architectures

The deep network architecture is the full arrangement of neural networks in deep learning models covering input layer, hidden layers, and output layer. Although there were some variations with the deep network architecture, 16 out of 17 studies in Table 1 used CNN-based architectures in their deep learning models. The choice of deep network architectures depends on the classification objectives and individual research group. Nevertheless, many research groups prefer to use the well-known CNN-based architectures when classifying WCE images or benchmarking the performance of their custom deep learning architectures. These prebuilt CNN-based architects include LeNet[25], AlexNet[25,27,31,32], GoogLeNet[6,25,31], VGG-Net[25], ResNet[20,22,30], RetinaNet[29], Single Shot MultiBox Detector[23,28,34], and Xception[33].

WCE devices

In addition to variations in the deep learning architect, researchers had some variation in WCE device. There were three brands of WCE devices mentioned in these deep learning studies: PillCam (Medtronics), NaviCam (Ankon Technologies), and MiroCam (IntroMedic). Deep learning models can be incorporated with each device. However, different devices have different sizes and qualities of raw images, brightness, and camera angles. Since these devices are not standardized, the application of a specific deep learning model may not perform at the same prediction accuracy when applied universally to the other WCE devices.

Image resolution

Although the size and quality of the original WCE images is dependent on the device, image resolution is dependent on training time, deep network architecture, and lesion types. Intuitively, physicians prefer a higher image resolution when making an image-based diagnosis. However, higher image resolutions can lead to an increase in

Table 1 Deep learning applications in wireless capsule endoscopy for classifying gastrointestinal disorders

Ref.	Class/outcome variable	Deep network architecture	Device/image resolution	Training and internal validation dataset	Testing/external validation dataset	Accuracy (%)/AUC	Sensitivity (%)/specificity (%)
Majid *et al*[19], 2020, NA	Multiple lesions (bleeding, esophagitis, ulcer, polyp)	CNN with classical features fusion and selection	NA/224 × 224 pixels	70% of 12889 images from multiple databases	30% of 12889 images from multiple databases	96.5/NA	96.5/NA
Ding *et al*[20], 2019, China	Multiple SB lesions[1]	CNN (ResNet 152)	SB-CE by Ankon Technologies/480 × 480 pixels	158235 images from 1970 patients	113268334 images from 5000 patients	NA/NA	99.88/100 (per patient); 99.90/100 (per lesion)
Iakovidis *et al*[21], 2018, NA	Multiple SB lesions[2]	CNN and iterative cluster unification	(1) NA/489 × 409 pixels; and (2) MiroCam CE/320 × 320 pixels	(1) 465 images from 1063 volunteers; and (2) 852 images	(1) 233 images from 1063 volunteers; and (2) 344 images	(1) 89.9/0.963; and (2) 77.5/0.814	(1) 90.7/88.2; and (2) 36.2/91.3
Aoki *et al*[22], 2020, Japan	Bleeding (blood content)	CNN (ResNet50)	Pillcam SB2 or SB3 CE / 224 × 224 pixels	27847 images from 41 patients	10208 images from 25 patients	99.89/0.9998	96.63/99.96
Tsuboi *et al*[23], 2019, Japan	Bleeding (SB angioectasia)	CNN (SSD)	Pillcam SB2 or SB3 CE/300 × 300 pixels	2237 images from 141 patients	10488 images from 28 patients	NA/0.998	98.8/98.4
Leenhardt *et al*[24], 2019, France	Bleeding (SB angioectasia)	CNN-based semantic segmentation	Pillcam SB3 CE / NA	600 images	600 images	NA/NA	96/100
Li *et al*[25], 2017, China	Bleeding (intestinal hemorrhage)	CNNs: (1) LeNet; (2) AlexNet; (3) GoogLeNet; and (4) VGG-Net	NA/NA	9672 images	2418 images	NA/NA	(1) 99.91/96.2; (2) 99.96/98.72; (3) 100/98.73; and (4) 99.96/98.72
Jia *et al*[26], 2017, Hong Kong, China	Bleeding (both active and inactive)	CNN	NA/240 × 240 pixels	1000 images	500 images	NA/NA	91.0/NA
Jia *et al*[27], 2016, Hong Kong, China	Bleeding (both active and inactive)	CNN (Inspired by AlexNet)	NA/240 × 240 pixels	8200 images	1800 images	NA/NA	99.2/NA
Aoki *et al*[28], 2019, Japan	Ulcer (erosion or ulceration)	CNN (SSD)	Pillcam SB2 or SB3 CE/300 × 300 pixels	5360 images from 115 patients	10440 images from 65 patients	90.8/0.958	88.2/90.9
Wang *et al*[29], 2019, China	Ulcer	CNN (RetinaNet)	Magnetic-guided CE by Ankon Technologies/480 × 480 pixels	37278 images from 1204 patient cases	9924 images from 300 patient cases	90.10/0.9469	89.71/90.48
Wang *et al*[30], 2019, China	Ulcer	CNN (based on ResNet 34)	Magnetic-guided CE by Ankon Technologies/480 × 480 pixels	80% of dataset from 1416 patients	20% of dataset from 1416 patients	92.05/0.9726	91.64/92.42
Alaskar *et al*[31], 2019, NA	Ulcer	CNN: (1) GoogLeNet; and (2) AlexNet	NA /(1) 224 × 224 pixels; and (2) 227 × 227 pixels	336 images	105 images	(1) 100/1; and (2) 100/1	(1) 100/100; and (2) 100/100
Fan *et al*[32],	(1) Ulcer; and (2) Erosion	CNN (AlexNet)	NA/511 × 511 pixels	(1) 5500 images; and (2) 7410	(1) 2750 images; and (2) 5500 images	(1) 95.16/0.9891; and (2)	(1) 96.80/94.79; and (2)

2018, China				images		95.34/0.9863	93.67/95.98
Zhou *et al*[6], 2017, USA	Celiac disease	CNN (GoogLeNet)	Pillcam SB2 CE/512 × 512 pixels	8800 images from 11 patients	8000 images from 10 patients	NA/NA	100/100
Klang *et al*[33], 2020, Israel	Crohn's disease	CNN (Xception)	Pillcam SB2 CE/299 × 299 pixels	Experiment 1: 80% of 17640 images from 49 patients; Experiment 2: Images from 48 patients	Experiment 1: 20% of 17,640 images from 49 patients; Experiment 2: Images from 1 individual patient	Experiment 1: 95.4-96.7/0.989-0.994; Experiment 2: 73.7–98.2/0.940-0.999	Experiment 1: 92.5-97.1/96.0-98.1; Experiment 2: 69.5-100/56.8-100
Saito *et al*[34], 2020, Japan	Polyp (protruding lesion)	CNN (SSD)	Pillcam SB2 or SB3 CE/300 × 300 pixels	30584 images from 292 patients	17507 images from 93 patients	84.5/0.911	90.7/79.8
Yuan *et al*[7], 2017, Hong Kong, China	Polyp	Deep neural network	Pillcam SB CE/64 × 64 pixels	Unknown proportion of 4000 images from 35 patients	Unknown proportion of 4000 images from 35 patients	98/NA	98/99
He *et al*[35], 2018, Israel	Hookworm	CNN	Pillcam SB CE/227 × 227 pixels	10 out of 11 patients (436796 images from 11 patients)	1 individual patient (11-fold cross-validation)	88.5/NA	84.6/88.6

[1]Abnormal classes include (1) inflammation; (2) ulcer; (3) polyps; (4) lymphangiectasia; (5) bleeding; (6) vascular disease; (7) protruding lesion; (8) lymphatic follicular hyperplasia; (9) diverticulum; and (10) parasite.
[2]Various lesions include gastritis, cancer, bleeding, ulcer, vascular anomalies, polypoid anomalies, and inflammation anomalies. AUC: Area under the receiver operating characteristic curve; CE: Capsule endoscopy; CNN: Convolutional neural networks; NA: Not available; SB: Small bowel; SSD: SingleShot Multi Box Detector.

trainable parameters, floating-point operations, memory requirements, and training time. To counteract this, original images are often modified (either cropped or resized) to lower image resolution. As illustrated in Table 1, image resolution can range from 64 × 64 pixels to 512 × 512 pixels. The typical range of resolution is 240 × 240 pixels to 320 × 320 pixels. It is worth noting that all studies using the images captured by NaviCam (Ankon Technologies) selected the original image resolution of 480 × 480 pixels[20,29,30].

Data partitioning

A collection of WCE images labeled by physicians is the main data source, which is commonly referred to as a dataset. As a part of data pre-processing, the dataset is typically divided into two groups. This creates two different datasets from the labeled WCE images. The first dataset is for training and internally validating the deep learning models. Once the final model is selected, the second dataset is used for testing the performance of the model with the data the model has not seen. Hence, data partitioning is one of the factors that could impact the predive performance of deep learning models[36].

There were two common approaches for dividing the initial dataset identified during the literature review. The first was to partition the data based on the aggregated images. The second was to partition the data per patient or video. The ratio of the two datasets varied dependent on the study, but common ratios included 50:50, 70:30 and 80:20[19,24,30,33]. The second approach to partition was often used when

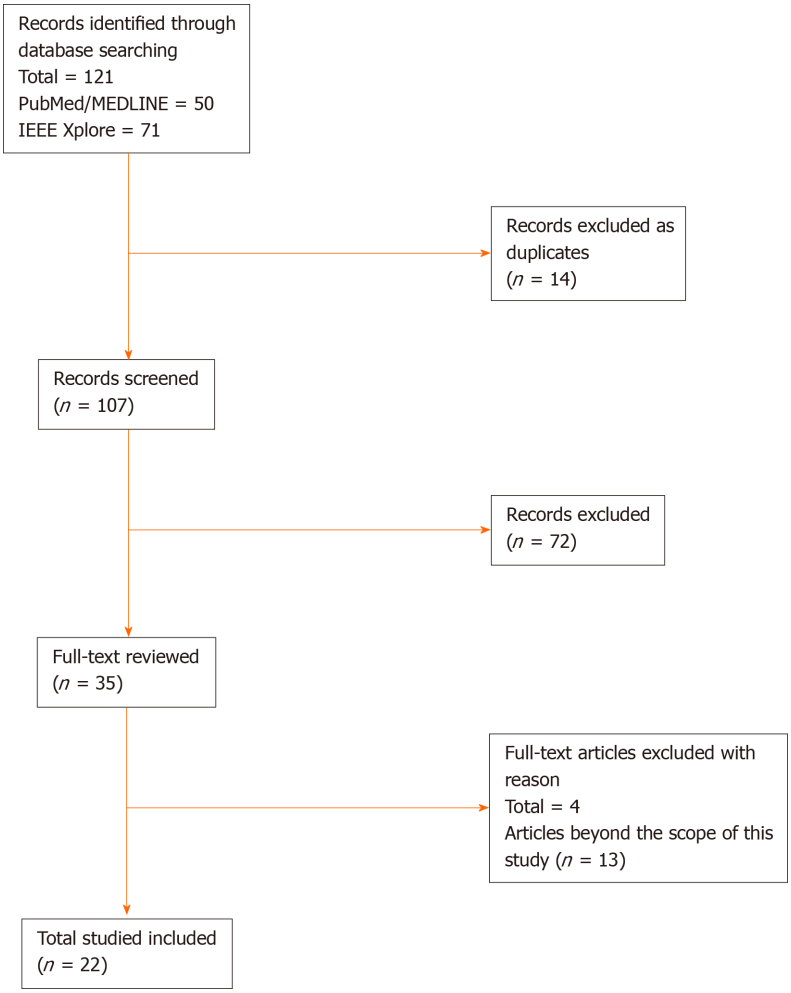


Figure 1 Study selection.

evaluating the predictive performance per patient[6,20,32,33]. Therefore, we can notice that the data partitioning approach in WCE images highly depends on the study design.

Performance metrics

In medical literature, the most popular performance metrics are accuracy, sensitivity, specificity, and area under the curve (AUC). In the case of WCE images, where few WCE images are true lesions, accuracy and specificity can be skewed by deep learning models correctly identifying normal mucosa. For this reason, in data science, the focus on performance evaluation is on true positive classification[37]. In other words, data scientists prefer their models correctly classify the small number of positive images (*e.g.*, angioectasia, tumor, or ulcer) rather than correctly classifying the normal mucosa images. Instead of accuracy and sensitivity, precision [true positive/(true positive + false positive)], recall [true positive/(true positive + false negative)], and F1 score (a harmonic mean of precision and recall) are the common performance metrics used by data scientists. It is worth noting that precision and recall are also known as positive predictive value and sensitivity respectively. Unfortunately, only a limited number of studies fully reported these set of performance metrics, especially F1 score[19,25-27]. In short, it is important to consider the performance metrics when determining or comparing the performance of deep learning models.

USE OF DEEP LEARNING FOR CLASSIFYING NON-DISEASE OBJECTS

The main goal when analyzing WCE images is to detect abnormalities in the gastrointestinal tract. However, it is also helpful to detect normal mucosa and anatomical landmarks. As shown in Table 2, only two studies were designed to classify non-disease objects. The first study used deep learning to classify the

Table 2 Deep learning applications in wireless capsule endoscopy for classifying non-disease objects

Ref.	Class/outcome variable	Deep network architecture	Device/image resolution	Training and internal validation dataset	Testing/external validation dataset	Accuracy (%)/AUC	Sensitivity (%)/specificity (%)
Seguí *et al*[38], 2016, Spain	Scenes (turbid, bubbles, clear blob, wrinkles, wall)	CNN	Pillcam SB2 CE/100 × 100 pixels	100000 images from 50 videos	20000 images from 50 videos	96/NA	NA/NA
Zou *et al*[5], 2015, NA	Organ locations (stomach, small intestine, and colon)	CNN (AlexNet)	NA/480 × 480 pixels	60000 images	15000 images	95.52/NA	NA/NA

AUC: Area under the receiver operating characteristic curve; CE: Capsule endoscopy; CNN: Convolutional neural networks; NA: Not available; SB: Small bowel.

complexities within the endoluminal scene, including turbid, bubbles, clear blob, wrinkles, and wall[38]. Although these images may not contribute to a final diagnosis, they can be used to characterize small intestine motility and to help rule out negative images. The second study created a predictive model for identifying organ locations such as the stomach, intestine, and colon[5]. Organ classification can be used to calculate the passage time of WCE in each organ and to determine if there are any motility disorders in the gastrointestinal tract. An important aspect of physician review of a WCE study is the identification of anatomical landmarks such as first images of the stomach, duodenum, and cecum which ultimately helps calculate capsule transit time through the small bowel. This transit time is vital to determining the location of lesion in the small bowel that may help guide treatment with deep enteroscopy techniques.

USEFULNESS OF DEEP LEARNING MODELS IN CLINICAL PRACTICE

An ideal goal for WCE would be the creation of a fully automated system for interpreting WCE images and generating accurate reports at least equivalent to conventional reading by physicians. Two retrospective studies compared the performance of conventional reading to the deep learning assisted reading (Table 3)[20,39]. The average reading times of deep learning assisted reading in both studies was less than 6 min. The average conventional reading time varied from 12 to 97 min depending on the expertise of the reader and the scope of WCE reading. In terms of overall lesion detection rate, there was a 3%-8% improvement of deep learning assisted reading over conventional reading. Interestingly, the accuracy of the deep learning model (as calculated during development) was higher than the actual detection rate. These findings may reflect the real-world challenges impacting human and deep learning model collaborations. An additional limitation was that there was no clear definition on how reading time was determined (*e.g.*, from data preprocessing to final report generation).

CHALLENGES

The goal when creating a deep learning model is to best fit your target function. Overfitting is a classic problem that can occur after creating the initial deep learning model. Overfitting occurs when a model learns the detail and noise of the training data too well to the extent that it negatively impacts the performance of the model on new data. Despite the standard methods for dividing datasets during training and testing, the detection rate in deep learning assisted trials are not as good when compared to the rates during the initial training and testing process[20,39]. The decreased performance could indicate that the model fits the training dataset too closely and does not perform well with an unseen dataset. Another explanation could be imperfect human and machine collaboration. Since the human physician is the one who makes the final diagnosis based on the information provided by the deep learning model, the misdetection could be derived from how human physicians use or trust the judgment

Table 3 Deep learning applications in wireless capsule endoscopy for improving the reading efficiency of wireless capsule endoscopy

Ref.	Experiment type	Scope of WCE reading /device	Conventional reading	Deep learning assisted reading	*P* value
Aoki *et al*[39], 2019, Japan	Retrospective study using anonymized data	SB section only/Pillcam SB3	mean reading time (min): Trainee: 20.7; Expert: 12.2	mean reading time (min): Trainee: 5.2; Expert: 3.1	< 0.001
			Overall lesion detection rate: Trainee: 47%; Expert: 84%	Overall lesion detection rate: Trainee: 55%; Expert: 87%	NS
Ding *et al*[20], 2019, China	Retrospective study by randomly selected videos	Small bowel abnormalities/SB-CE by Ankon Technologies	mean reading time ± standard deviation (min): 96.6 ± 22.53	mean reading time ± standard deviation (min): 5.9 ± 2.23	< 0.001
			Overall lesion detection rate: 41.43%	Overall lesion detection rate: 47.00%	NA[1]

[1]In per-patient analysis, deep learning assisted physician significantly outperformed conventional reading in detecting lymphangiectasia, lymphatic follicular, hyperplasia, inflammation, protruding lesion, and polyps. CE: Capsule endoscopy; NA: Not available; NS: Not significant; SB: Small bowel.

from deep learning models.

Traditionally, the risk stratification scores developed by one research team can be validated by another research team. Unfortunately, we have not seen the same level of transferability in deep learning research for WCE yet. As a result, the trials are very limited to their own research group and can be very difficult to have third party validation.

Each deep learning model is designed for a specific task that is based on the availability of positive lesions in their own dataset. Given this, it is questionable if it is even possible or effective to integrate these models. Integration can be even more complicated by the fact that each research group may use different devices, image resolutions, network architectures, and labeling practices.

One common barrier in medical device-related research is the use of proprietary file format. For example, the video file from PillCam device is stored in *.gvi and *.gvf file[40]. Thus, it may be difficult to extract data that is stored in the proprietary file format without help from the manufacturer. Such constraints may impact model integration and deployment. For example, it may take a longer time to prepare the files from deep learning models to use in a clinical setting. Also, there is no guarantee that the image resolution would be equivalent to the one seen in the proprietary reading software after extraction. For this reason, researchers should explore the pros and cons of each device available in their market to compare features and select the one that best aligns with their research goals.

Data preprocessing is the most time-consuming task in AI research. It is necessary to transform raw data into a ready-to-use and efficient format. Having a high-quality dataset is one of the key factors for creating a predictive model. By spending a lot of time extracting the data and labeling it, the dataset is a valuable asset to the research group. Ideally, high-quality datasets should be publicly available for researchers to use. However, there are a limited number of such datasets.

CONCLUSION

Since 2006, CNN-based architecture has proven to be an effective method for analyzing image data in various fields. Researchers have increasingly adopted CNN-based architecture for solving image classification problems. In our literature review, seventeen papers were identified that applied deep learning in WCE to classify gastrointestinal disorders. Our literature review demonstrated that the majority of CNN-based deep learning models were nearly perfect with regard to accuracy, sensitivity, specificity, and AUC[9].

There were only a few studies applying deep learning models to address non-disease objects, such as organ location and scenes in normal mucosa images (*e.g.*, turbid, bubbles, clear blob, wrinkles, and wall). These non-disease objects are important building blocks toward a fully automated system and can aid in the identification of "landmarks" such as the first images of each bowel segment.

Although there seems to be an increasing amount of deep learning research on classifying WCE images, we are still in the early stages of investigating the utility of

deep learning in enhancing clinical practice. The studies we identified often reflected the more standard view of WCE, as a means to view areas of the small bowel not accessible by upper and lower endoscopy. As the scope of WCE grows beyond the small bowel, we expect to see deep learning research on WCE expand accordingly. In addition, deep learning could enhance WCE capability to become highly effective in clinical practice and patient care by improving the speed and accuracy of WCE reading as well as predicting the location of abnormalities. Regardless of existing limitations and constraints, we expect the research and development in this area will continue to grow rapidly in the next decade.

The studies gathered in this literature review were indexed by PubMed. We also investigated publications concerning the utility of deep learning in computer science, medical image processing, mathematical modeling, and electrical engineering. Unfortunately, we cannot ensure that we identified every publication outside of PubMed.

In addition, it is difficult to compare one deep learning model to another based on their performance metrics alone. Most researchers have focused more on reporting traditional performance metrics without F1 score. The best practice for comparing these models would be to benchmark their performances on the same dataset that the models have never been trained on. To do so, researchers would need to make their trained models publicly available (*e.g.*, uploading them to GitHub). This would allow clinical trials on deep learning models to expand outside their research group.

The idea of using computational algorithms for analyzing WCE images is not entirely new. The earliest study identified was published in 2006[41]. Universal to all these studies was a central hypothesis investigating the ability of computational algorithms to improve the efficiency of reading WCE studies, specifically in terms of time and accuracy. The prospect of a fully automated system for interpreting WCE images would benefit patient care because of fast and accurate diagnoses of gastrointestinal medical conditions such as bleeding, polyps, Crohn's disease, and cancer.

REFERENCES

1 **Moglia A**, Menciassi A, Schurr MO, Dario P. Wireless capsule endoscopy: from diagnostic devices to multipurpose robotic systems. *Biomed Microdevices* 2007; **9**: 235-243 [PMID: 17160703 DOI: 10.1007/s10544-006-9025-3]

2 **Li B**, Meng MQ, Xu L. A comparative study of shape features for polyp detection in wireless capsule endoscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2009; **2009**: 3731-3734 [PMID: 19965014 DOI: 10.1109/IEMBS.2009.5334875]

3 **Mishkin DS**, Chuttani R, Croffie J, Disario J, Liu J, Shah R, Somogyi L, Tierney W, Song LM, Petersen BT; Technology Assessment Committee; American Society for Gastrointestinal Endoscopy. ASGE Technology Status Evaluation Report: wireless capsule endoscopy. *Gastrointest Endosc* 2006; **63**: 539-545 [PMID: 16564850 DOI: 10.1016/j.gie.2006.01.014]

4 **Hirasawa T**, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J, Tada T. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018; **21**: 653-660 [PMID: 29335825 DOI: 10.1007/s10120-018-0793-2]

5 **Zou Y**, Li L, Wang Y, Yu J, Li Y, Deng W. Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network. Proceedings of the 2015 IEEE International Conference on Digital Signal Processing (DSP); 2015 July 21-24; Singapore, Singapore. IEEE, 2015: 1274-1278 [DOI: 10.1109/ICDSP.2015.7252086]

6 **Zhou T**, Han G, Li BN, Lin Z, Ciaccio EJ, Green PH, Qin J. Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method. *Comput Biol Med* 2017; **85**: 1-6 [PMID: 28412572 DOI: 10.1016/j.compbiomed.2017.03.031]

7 **Yuan Y**, Meng MQ. Deep learning for polyp recognition in wireless capsule endoscopy images. *Med Phys* 2017; **44**: 1379-1389 [PMID: 28160514 DOI: 10.1002/mp.12147]

8 **Sahiner B**, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, Summers RM, Giger ML. Deep learning in medical imaging and radiation therapy. *Med Phys* 2019; **46**: e1-e36 [PMID: 30367497 DOI: 10.1002/mp.13264]

9 **Le Berre C**, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, Danese S, Peyrin-Biroulet L. Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology* 2020; **158**: 76-94. e2 [PMID: 31593701 DOI: 10.1053/j.gastro.2019.08.058]

10 **Ruffle JK**, Farmer AD, Aziz Q. Artificial Intelligence-Assisted Gastroenterology- Promises and Pitfalls. *Am J Gastroenterol* 2019; **114**: 422-428 [PMID: 30315284 DOI: 10.1038/s41395-018-0268-4]

11 **Thakkar SJ**, Kochhar GS. Artificial intelligence for real-time detection of early esophageal cancer: another set of eyes to better visualize. *Gastrointest Endosc* 2020; **91**: 52-54 [PMID: 31865996 DOI: 10.1016/j.gie.2019.09.036]

12 **Mahmud M**, Kaiser MS, Hussain A, Vassanelli S. Applications of Deep Learning and Reinforcement Learning to Biological Data. *IEEE Trans Neural Netw Learn Syst* 2018; **29**: 2063-2079 [PMID: 29771663

DOI: 10.1109/TNNLS.2018.2790388]

13 **Spann A**, Yasodhara A, Kang J, Watt K, Wang B, Goldenberg A, Bhat M. Applying Machine Learning in Liver Disease and Transplantation: A Comprehensive Review. *Hepatology* 2020; **71**: 1093-1105 [PMID: 31907954 DOI: 10.1002/hep.31103]

14 **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078. e8 [PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037]

15 **Kang R**, Park B, Eady M, Ouyang Q, Chen K. Single-cell classification of foodborne pathogens using hyperspectral microscope imaging coupled with deep learning frameworks. *Sens Actuators B Chem* 2020; **309**: 127789 [DOI: 10.1016/j.snb.2020.127789]

16 **Yasaka K**, Akai H, Kunimatsu A, Kiryu S, Abe O. Deep learning with convolutional neural network in radiology. *Jpn J Radiol* 2018; **36**: 257-272 [PMID: 29498017 DOI: 10.1007/s11604-018-0726-3]

17 **Ehteshami Bejnordi B**, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak JAWM; the CAMELYON16 Consortium, Hermsen M, Manson QF, Balkenhol M, Geessink O, Stathonikos N, van Dijk MC, Bult P, Beca F, Beck AH, Wang D, Khosla A, Gargeya R, Irshad H, Zhong A, Dou Q, Li Q, Chen H, Lin HJ, Heng PA, Haß C, Bruni E, Wong Q, Halici U, Öner MÜ, Cetin-Atalay R, Berseth M, Khvatkov V, Vylegzhanin A, Kraus O, Shaban M, Rajpoot N, Awan R, Sirinukunwattana K, Qaiser T, Tsang YW, Tellez D, Annuscheit J, Hufnagl P, Valkonen M, Kartasalo K, Latonen L, Ruusuvuori P, Liimatainen K, Albarqouni S, Mungal B, George A, Demirci S, Navab N, Watanabe S, Seno S, Takenaka Y, Matsuda H, Ahmady Phoulady H, Kovalev V, Kalinovsky A, Liauchuk V, Bueno G, Fernandez-Carrobles MM, Serrano I, Deniz O, Racoceanu D, Venâncio R. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* 2017; **318**: 2199-2210 [PMID: 29234806 DOI: 10.1001/jama.2017.14585]

18 **Esteva A**, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115-118 [PMID: 28117445 DOI: 10.1038/nature21056]

19 **Majid A**, Khan MA, Yasmin M, Rehman A, Yousafzai A, Tariq U. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. *Microsc Res Tech* 2020; **83**: 562-576 [PMID: 31984630 DOI: 10.1002/jemt.23447]

20 **Ding Z**, Shi H, Zhang H, Meng L, Fan M, Han C, Zhang K, Ming F, Xie X, Liu H, Liu J, Lin R, Hou X. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* 2019; **157**: 1044-1054. e5 [PMID: 31251929 DOI: 10.1053/j.gastro.2019.06.025]

21 **Iakovidis DK**, Georgakopoulos SV, Vasilakakis M, Koulaouzidis A, Plagianakos VP. Detecting and Locating Gastrointestinal Anomalies Using Deep Learning and Iterative Cluster Unification. *IEEE Trans Med Imaging* 2018; **37**: 2196-2210 [PMID: 29994763 DOI: 10.1109/TMI.2018.2837002]

22 **Aoki T**, Yamada A, Kato Y, Saito H, Tsuboi A, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Nakahori M, Tanaka S, Koike K, Tada T. Automatic detection of blood content in capsule endoscopy images based on a deep convolutional neural network. *J Gastroenterol Hepatol* 2020; **35**: 1196-1200 [PMID: 31758717 DOI: 10.1111/jgh.14941]

23 **Tsuboi A**, Oka S, Aoyama K, Saito H, Aoki T, Yamada A, Matsuda T, Fujishiro M, Ishihara S, Nakahori M, Koike K, Tanaka S, Tada T. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig Endosc* 2020; **32**: 382-390 [PMID: 31392767 DOI: 10.1111/den.13507]

24 **Leenhardt R**, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, Becq A, Marteau P, Histace A, Dray X; CAD-CAP Database Working Group. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc* 2019; **89**: 189-194 [PMID: 30017868 DOI: 10.1016/j.gie.2018.06.036]

25 **Li P**, Li Z, Gao F, Wan L, Yu J. Convolutional neural networks for intestinal hemorrhage detection in wireless capsule endoscopy images. Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME); 2017 July 10-14; Hong Kong, China. IEEE, 2017: 1518-1523 [DOI: 10.1109/ICME.2017.8019415]

26 **Jia X**, Meng MQ. Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features. *Annu Int Conf IEEE Eng Med Biol Soc* 2017; **2017**: 3154-3157 [PMID: 29060567 DOI: 10.1109/EMBC.2017.8037526]

27 **Jia X**, Meng MQ. A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2016; **2016**: 639-642 [PMID: 28268409 DOI: 10.1109/EMBC.2016.7590783]

28 **Aoki T**, Yamada A, Aoyama K, Saito H, Tsuboi A, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Tanaka S, Koike K, Tada T. Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest Endosc* 2019; **89**: 357-363. e2 [PMID: 30670179 DOI: 10.1016/j.gie.2018.10.027]

29 **Wang S**, Xing Y, Zhang L, Gao H, Zhang H. A systematic evaluation and optimization of automatic detection of ulcers in wireless capsule endoscopy on a large dataset using deep convolutional neural networks. *Phys Med Biol* 2019; **64**: 235014 [PMID: 31645019 DOI: 10.1088/1361-6560/ab5086]

30 **Wang S**, Xing Y, Zhang L, Gao H, Zhang H. Deep Convolutional Neural Network for Ulcer Recognition in Wireless Capsule Endoscopy: Experimental Feasibility and Optimization. *Comput Math Methods Med* 2019; **2019**: 7546215 [PMID: 31641370 DOI: 10.1155/2019/7546215]

31 **Alaskar H**, Hussain A, Al-Aseem N, Liatsis P, Al-Jumeily D. Application of Convolutional Neural Networks for Automated Ulcer Detection in Wireless Capsule Endoscopy Images. *Sensors (Basel)* 2019; **19** [PMID: 30871162 DOI: 10.3390/s19061265]

32 **Fan S**, Xu L, Fan Y, Wei K, Li L. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. *Phys Med Biol* 2018; **63**: 165001 [PMID: 30033931 DOI: 10.1088/1361-6560/aad51c]

33 **Klang E**, Barash Y, Margalit RY, Soffer S, Shimon O, Albshesh A, Ben-Horin S, Amitai MM, Eliakim R, Kopylov U. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc* 2020; **91**: 606-613. e2 [PMID: 31743689 DOI: 10.1016/j.gie.2019.11.012]

34 **Saito H**, Aoki T, Aoyama K, Kato Y, Tsuboi A, Yamada A, Fujishiro M, Oka S, Ishihara S, Matsuda T, Nakahori M, Tanaka S, Koike K, Tada T. Automatic detection and classification of protruding lesions in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest Endosc* 2020; **92**: 144-151. e1 [PMID: 32084410 DOI: 10.1016/j.gie.2020.01.054]

35 **He JY**, Wu X, Jiang YG, Peng Q, Jain R. Hookworm Detection in Wireless Capsule Endoscopy Images With Deep Learning. *IEEE Trans Image Process* 2018; **27**: 2379-2392 [PMID: 29470172 DOI: 10.1109/TIP.2018.2801119]

36 **Korjus K**, Hebart MN, Vicente R. An Efficient Data Partitioning to Improve Classification Performance While Keeping Parameters Interpretable. *PLoS One* 2016; **11**: e0161788 [PMID: 27564393 DOI: 10.1371/journal.pone.0161788]

37 **Sokolova M**, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag* 2009; **45**: 427-437 [DOI: 10.1016/j.ipm.2009.03.002]

38 **Seguí S**, Drozdzal M, Pascual G, Radeva P, Malagelada C, Azpiroz F, Vitrià J. Generic feature learning for wireless capsule endoscopy analysis. *Comput Biol Med* 2016; **79**: 163-172 [PMID: 27810622 DOI: 10.1016/j.compbiomed.2016.10.011]

39 **Aoki T**, Yamada A, Aoyama K, Saito H, Fujisawa G, Odawara N, Kondo R, Tsuboi A, Ishibashi R, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Nakahori M, Tanaka S, Koike K, Tada T. Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading. *Dig Endosc* 2020; **32**: 585-591 [PMID: 31441972 DOI: 10.1111/den.13517]

40 **PillCam® Capsule Endoscopy**. User Manual RAPID® v9.0, DOC-2928-02, Novermber 2016, p.123. Available from: https://www.pillcamcrohnscapsule.eu/assets/pdf/DOC-2928-02-PillCam-Desktop-SWv9UMEN.pdf

41 **Spyridonos P**, Vilariño F, Vitrià J, Azpiroz F, Radeva P. Anisotropic feature extraction from endoluminal images for detection of intestinal contractions. *Med Image Comput Comput Assist Interv* 2006; **9**: 161-168 [PMID: 17354768 DOI: 10.1007/11866763_20]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com