

AI everywhere in endoscopy, not only for detection and characterization



Authors

Cesare Hassan¹, Yuichi Mori^{2,4}, Giulio Antonelli^{3,4}

Institutions

- 1 Gastroenterology Unit, Nuovo Regina Margherita Hospital, Rome, Italy
- 2 Department of Translational and Precision Medicine, "Sapienza" University of Rome, Italy
- 3 Clinical Effectiveness Research Group, University of Oslo, Oslo, Norway
- 4 Digestive Disease Center, Showa University Northern Yokohama Hospital, Yokohama, Japan

Bibliography

Endosc Int Open 2021; 09: E627–E628

DOI 10.1055/a-1373-4799

ISSN 2364-3722

© 2021. The Author(s).

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

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

Corresponding author

Cesare Hassan, MD, PhD, Gastroenterology Unit, Nuovo Regina Margherita Hospital, Rome, Italy

Fax: +390658446533

cesareh@hotmail.com

Time is life! While saving many patients from cancer-related mortality, gastrointestinal endoscopy still pays an unacceptable price in terms of missed diagnosis and post-endoscopy cancer [1–4]. Isn't it true that nearly 80% of early Barrett's-related neoplasia are missed by community endoscopists [5]? Similar estimates are likely for early gastric cancer missed in non-expert centers. Time is also money! How much are we wasting in the duplication between endoscopic prediction and post-endoscopic confirmation? Suboptimal competence in differentiating between adenomatous and hyperplastic polyps or predicting precancerous gastric lesions leads to burdensome costs for pathology.

If time is life, artificial intelligence (AI) is the answer! Isn't it true that AI is fast, smart, and brilliant? It can compute millions of mathematical operations every second, it analyzes a higher number of frames than the human eye, and it comes with accuracy values equivalent or superior to our best experts. The data for AI associated with colonoscopy show at least a one-third increase in the adenoma detection rate, a 50% increase in adenomas per colonoscopy, and an equivalent drop in the polyp miss rate [6]. AI also can save money! At least in the endocytoscopy setting, very high accuracy values for polyp characterization have been shown with use of the technology, irrespective of the endoscopist [7,8]. This results in raising the bar on standards and saving money that otherwise would be wasted on pathology. No doubt, colonoscopy teaches that if time is the problem, AI is the fastest answer!

Houston, we have a problem! If AI is the ideal solution for speed and accuracy, why do we still suffer from so many pitfalls in everyday endoscopy? Why don't we have the same AI support that we enjoy in colonoscopy for upper gastrointestinal endoscopy? Why don't we have additional modules for colonoscopy, such as superficial T1 invasion prediction or quantification of bowel preparation score or inflammatory disease activity? How long should we wait before benefiting from AI-based quantification of mucosal exploration, blind spot alarm, or inspection timing monitoring? Similarly, why aren't the available modules being continuously updated? Why doesn't AI learn from its own mistakes in real time as does speech recognition software on our mobile phones? Ultimately, why is AI development in gastrointestinal endoscopy so painfully slow?

There is an elephant in the room! What if behind AI a hidden "huge" time-consuming factor exists? Let's take a step back. How do we produce or train an AI algorithm? Unlike the classic machine-learning methods, the latest AI technologies, such as deep learning, allow automatic extraction of the object features by the machine itself. The practicality of this, thus, is that the accuracy of an AI system primarily depends on the accuracy of the training data. This is why we need to know everything about AI training. How many endoscopists collected the data? What was their level of expertise? How many patients were included? How many lesions and what type?

The fact that only rigorous and robust training ensures high-level AI performance has its own drawback. Because such train-

ing must be supervised by humans, it requires an extremely laborious phase, which is manual annotation of endoscopic images. Behind the scenes, AI is supported by a new but little-known form of human slavery. To provide just one example, in order for a machine to learn a single case of a colorectal polyp, approximately 3,000 frames must be manually annotated, which sometimes takes an hour or more. The hyperbolic multiplication for thousands and thousands of cases explains why human-supervised training is the time-consuming step in AI implementation. First, the scarcity of manpower for manual annotation forces most researchers to focus only on one AI module at a time. Second, for each software, the available capacity in terms of administered cases is still suboptimal; most of the available software has learned no more than a few hundred cases. Let's be honest! Most AI engines in colonoscopy have never been presented with enough cases of rarer but highly clinically relevant lesions, such as non-granular laterally spreading tumors or depressed invasive cancers [9].

Annotation-related burden is primary responsible for the persistent delay in AI implementation throughout the field of gastrointestinal endoscopy. Thus, any methods of bypassing it – such as the one presented by Hansen U et al. in this issue of the journal – should be prioritized. By incorporating a deep learning module in the annotation software, the authors increased the number of annotated images per minute by 5- to 6-fold. This had two clinically relevant implications. First, if to develop an AI model took an average of 1 year, the new annotation system compressed that time to 2 to 3 months. Second, the necessary manpower was minimized because only approximately 3% of the frame sample needed to be hand-annotated while the remaining had the benefit of being propagated automatically.

Addressing the time-consuming phase of deep learning training is expected to speed up the development of AI modules for the entire spectrum of gastrointestinal diseases and to also reduce the substantial cost of development of AI software. This may also lead to faster updating of the available modules, including adding rare but clinically relevant lesions to machine learning. Ideally, endoscopists should gain the ability to annotate real-time relevant images on the endoscopy screen, which would give the system immediate feedback, similar to what happens with speech recognition software. Of note, advances in technology and consideration of ethics and data-transfer issues should be part of this process.

Additional evidence is, however, needed before deep learning-based propagation is incorporated into the annotation process of AI applied to gastrointestinal endoscopy. What was effective for polyp annotation may not be generalizable to more subtle flat neoplastic lesions, such as upper gastrointestinal neoplasia. Similarly, the same propagation software should be validated on more complex visual patterns, such as inflammatory bowel disease or bowel preparation scores. Selection or

spectrum disease bias cannot be excluded when considering the limited number of cases included in this study.

Conclusion

In conclusion, the novelty of AI-based polyp detection and characterization has generated great expectations for a more universal AI approach to gastrointestinal endoscopy in both the endoscopy community and the patient population, including upper gastrointestinal endoscopy, and a more complete quality assurance assessment. To meet such complex needs, an acceleration in supervised training is mandatory, and automatization of the annotation process is likely to play a pivotal role. What may appear as a mere technicality may have universal life-saving benefits for patients with neoplasia.

Competing interests

The authors declare that they have no conflict of interest.

References

- [1] Zhao S, Wang S, Pan P et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. *Gastroenterology* 2019; 156: 1661–1674.e11
- [2] Hernanz N, Rodríguez de Santiago E, Marcos Prieto HM et al. Characteristics and consequences of missed gastric cancer: A multicentric cohort study. *Dig Liver Dis* 2019; 51: 894–900
- [3] Pimenta-Melo AR, Monteiro-Soares M, Libânio D et al. Missing rate for gastric cancer during upper gastrointestinal endoscopy: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol* 2016; 28: 1041–1049
- [4] Rodríguez de Santiago E, Hernanz N, Marcos-Prieto HM et al. Rate of missed esophageal cancer at routine endoscopy and survival outcomes: A multicentric cohort study. *United Eur Gastroenterol J* 2019; 7: 189–198
- [5] Bergman JJGHM, de Groof AJ, Pech O et al. An interactive web-based educational tool improves detection and delineation of Barrett's esophagus-related neoplasia. *Gastroenterology* 2019; 156: 1299–1308.e3
- [6] Hassan C, Spadaccini M, Iannone A et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2020; doi:10.1016/j.gie.2020.06.059
- [7] Mori Y, Kudo S-E, East JE et al. Cost savings in colonoscopy with artificial intelligence-aided polyp diagnosis: an add-on analysis of a clinical trial (with video). *Gastrointest Endosc* 2020; 92: 905–911.e1
- [8] Mori Y, Kudo S-E, Misawa M et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Ann Intern Med* 2018; 169: 357–366
- [9] Hassan C, Bhandari P, Antonelli G et al. Artificial intelligence for non-polypoid colorectal neoplasms. *Dig Endosc* 2020; doi:10.1111/den.13807