

# Deep Learning and Minimally Invasive Endoscopy: Automatic Classification of Pleomorphic Gastric Lesions in Capsule Endoscopy

Miguel Mascarenhas, MD, MSc<sup>1,2,3,\*</sup>, Francisco Mendes, MD, MSc<sup>1,2,\*</sup>, Tiago Ribeiro, MD, MSc<sup>1,2</sup>, João Afonso, MD, MSc<sup>1,2</sup>, Pedro Cardoso, MD, MSc<sup>1,2</sup>, Miguel Martins, MD, MSc<sup>1,2</sup>, Hélder Cardoso, MD, MSc<sup>1,2,3</sup>, Patrícia Andrade, MD, MSc<sup>1,2,3</sup>, João Ferreira, PhD<sup>4,5</sup>, Miguel Mascarenhas Saraiva, MD, PhD<sup>6</sup> and Guilherme Macedo, MD, PhD<sup>1,2,3</sup>

**INTRODUCTION:** Capsule endoscopy (CE) is a minimally invasive examination for evaluating the gastrointestinal tract. However, its diagnostic yield for detecting gastric lesions is suboptimal. Convolutional neural networks (CNNs) are artificial intelligence models with great performance for image analysis. Nonetheless, their role in gastric evaluation by wireless CE (WCE) has not been explored.

**METHODS:** Our group developed a CNN-based algorithm for the automatic classification of pleomorphic gastric lesions, including vascular lesions (angiectasia, varices, and red spots), protruding lesions, ulcers, and erosions. A total of 12,918 gastric images from 3 different CE devices (PillCam Crohn's; PillCam SB3; OMOM HD CE system) were used from the construction of the CNN: 1,407 from protruding lesions; 994 from ulcers and erosions; 822 from vascular lesions; and 2,851 from hematic residues and the remaining images from normal mucosa. The images were divided into a training (split for three-fold cross-validation) and validation data set. The model's output was compared with a consensus classification by 2 WCE-experienced gastroenterologists. The network's performance was evaluated by its sensitivity, specificity, accuracy, positive predictive value and negative predictive value, and area under the precision-recall curve.

**RESULTS:** The trained CNN had a 97.4% sensitivity; 95.9% specificity; and positive predictive value and negative predictive value of 95.0% and 97.8%, respectively, for gastric lesions, with 96.6% overall accuracy. The CNN had an image processing time of 115 images per second.

**DISCUSSION:** Our group developed, for the first time, a CNN capable of automatically detecting pleomorphic gastric lesions in both small bowel and colon CE devices.

**KEYWORDS:** artificial intelligence; capsule endoscopy; deep learning

*Clinical and Translational Gastroenterology* 2023;14:e00609. <https://doi.org/10.14309/ctg.000000000000609>

## INTRODUCTION

The detection and treatment of multiple gastric lesions is of utmost importance. Conventional esophagogastroduodenoscopy (EGD) is the current standard of care for gastric evaluation, either in a screening setting or in patients with upper gastrointestinal (GI) symptoms, given its ease in identifying and treating gastric lesions. However, upper endoscopy is an invasive examination, with a non-neglectable risk of perforation, bleeding,

infection, or even cardiopulmonary adverse events (1). Furthermore, EGD can be uncomfortable, and the use of sedation techniques during the examination can increase costs (2) related to the procedure itself and result in loss of working days by patients.

Capsule endoscopy (CE) is a minimally invasive examination that allows the entire visualization of the GI tract (3), with each capsule type having its own characteristics (Figure 1). Recently, it has emerged as an alternative for conventional EGD in the

<sup>1</sup>Precision Medicine Unit, Department of Gastroenterology, São João University Hospital, Alameda Professor Hernâni Monteiro, Porto, Portugal; <sup>2</sup>WGO Gastroenterology and Hepatology Training Center, Porto, Portugal; <sup>3</sup>Faculty of Medicine of the University of Porto, Alameda Professor Hernâni Monteiro, Porto, Portugal; <sup>4</sup>Department of Mechanical Engineering, Faculty of Engineering of the University of Porto, Porto, Portugal; <sup>5</sup>Digestive Artificial Intelligence Development, Porto, Portugal; <sup>6</sup>ManopH Gastroenterology Clinic, Porto, Portugal. **Correspondence:** Miguel Mascarenhas Saraiva, MD, PhD. E-mail: miguelmascarenhasaraiva@gmail.com.

\*Miguel Mascarenhas and Francisco Mendes contributed equally to this work.

Received February 13, 2023; accepted June 2, 2023; published online July 3, 2023

© 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology

evaluation of the upper GI tract, especially the small bowel (4,5). With the development of colon CE (CCE), CE-based panendoscopy is an important discussion matter (6), with minimally invasive evaluation of the entire GI tract. However, CE diagnostic performance for gastric lesions is suboptimal, limiting its clinical use (7). The stomach's anatomy is very different from that of the esophagus or duodenum, and its collapsed structure in the absence of insufflation makes the observation of all its surfaces difficult, especially the more proximal regions (8). CE dependence of the peristaltic movements can be a challenge when implementing this technology in the clinical setting. Besides, CE is a time-consuming examination, with reading times for a single examination ranging from 30 to 120 minutes (9).

The large number of image frames presented by CE devices favors the development of artificial intelligence (AI) tools for image analysis. Convolutional neural networks (CNNs) are a multilayer AI architecture with high performance levels for image analysis, inspired in the neural architecture of the human visual cortex, making it suitable for detection of imaging patterns (10). Recently, CNN models have revealed promising results in several fields of medicine (10–13). CE is one of the most studied fields for the development of CNN-based technologies for automatic detection of lesions and normal mucosa (14,15). Moreover, future application of AI tools might increase CE diagnostic yield while, at the same time, shorten its reading time and increase its cost-effectiveness. Nonetheless, the role of this technology in the identification of gastric lesions by using wireless CE (WCE) is yet to be explored. In this project, our group aimed to create a CNN-based system for the automatic classification of multiple gastric lesions using 3 different CE devices, focusing on vascular lesions (angiectasia, varices, and red spots), protruding lesions, ulcers, and erosions.

## METHODS

### Study design

This multicentric multidevice study was based on gastric images obtained from 3 different types of CE devices (PillCam SB3; PillCam Crohn's; and OMOM HD CE system) in 2 different centers (Centro Hospitalar Universitário São João and ManopH), comprising 5,846 CE examinations in 4,372 patients.

This project was developed in a noninterventional fashion (without intervention in the clinical management of each patient involved). This study was performed following the Declaration of Helsinki and with approval from the ethics committee of São João University Hospital/Faculty of Medicine of the University of Porto (No. CE 407/2020). Information potentially associated with the identification of the patients was omitted, and effective data anonymization for researchers involved in CNN development was assured by random number assignment for each patient. A legal team with Data Protection Officer Certification (Maastricht University) ensured the nontraceability of data and conformity with the general data protection regulation.

### CE protocol

CE procedures were conducted using 3 different CE devices: the PillCam SB3 system (Medtronic, Minneapolis, MN), the PillCam Crohn's (Medtronic), and the OMOM HD (Jinshan Science & Technology, Chongqing, Yubei, China). Images from PillCam SB3 and PillCam Crohn's CE were reviewed using PillCam software version 9 (Medtronic), whereas images from the OMOM

HD device were reviewed using the Vue Smart software (Jinshan Science & Technology). Possible patient-identifying information (name, operating number, and date of procedure) was removed by image processing. After that, each extracted frame was stored and labeled with a consecutive number.

Each patient was asked to undergo bowel preparation in line with previous recommendations by the European Society of Gastrointestinal Endoscopy (16). In summary, patients were asked to follow a clear liquid diet on the day before capsule ingestion, with fasting the night before the examination. Before capsule ingestion, patients drank a bowel preparation consisting of 2 L of polyethylene glycol solution. Simethicone was used as an antifoaming agent. Domperidone 10 mg was used as a prokinetic if the capsule remained in the stomach 1 hour after ingestion (which implied image review on the data recorder worn by the patient).

### Classification of lesions

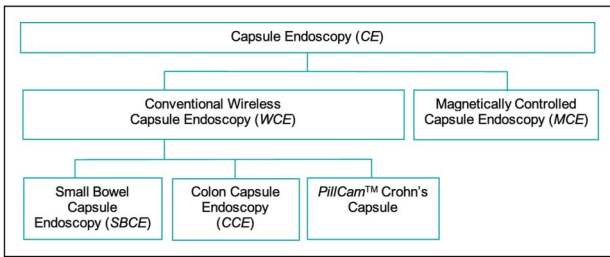
The gastric segment video was reviewed for identification of multiple gastric lesions. This included the first gastric image distal to the esophagogastric junction until the last image before the appearance of the duodenal mucosa. The whole group of gastric lesions included vascular lesions (angiectasia, varices, and red spots), protruding lesions, ulcers, and erosions. The definitions of the different lesions were adapted from classification scores used in small bowel CE (SBCE) (17). Regarding vascular lesions, red spots were defined as a punctuate (<1 mm) flat lesion with a bright red area, within the mucosal layer, without vessel appearance (18). Angiectasia was defined as a distinct reddish lesion constituting tortuous and clustered capillary dilations within the mucosal layer. Varices were defined as raised venous dilatation with a serpiginous appearance. Protruding lesions included polyps, epithelial tumors, subepithelial lesions, nodules, and venous structures (19). Mucosal erosions were defined as minimal loss of epithelial layering surrounded by normal mucosa. Ulcers were defined as depressed loss of epithelial covering, with a whitish base and surrounding swollen mucosa, with an estimated diameter of >5 mm.

Classification of the extracted images was performed by 3 gastroenterologists with CE expertise (M.M.S., H.C., and P.A.), each having read over 1,000 CE examinations before this study. The inclusion of a specific image implied the concordance of classification between at least 2 experts.

### CNN development

After evaluating all the examinations, 12,918 selected gastric images from 107 CE examinations were inserted into a CNN model with transfer learning. The full image data set had 1,407 images of protruding lesions; 994 from ulcers and erosions; 822 from vascular lesions; and 2,851 from hematic residues, with the remaining images being from normal mucosa.

The selected images were divided into 2 different data sets, one training data set (around 90% of the full image data set = 11,289), which was divided into 3 independent subsets, and an independent validation data set (around 10% of the full image data set = 1,629). On division, all images from a given patient were allocated to the same data set (patient-split design). The validation data set was used to evaluate the performance of the CNN model. The study design is summarized with a flowchart presented in Figure 2.



**Figure 1.** Different types of capsule endoscopies that are discussed in this article.

When creating the CNN, the Xception model with its weights trained on ImageNet was used. To transfer this learning to our data, the convolutional layers of the model were kept. We removed the last fully connected layers and attached fully connected layers based on the number of classes used to classify our endoscopic images.

Our group used 2 blocks, each having a fully connected layer, followed by a dropout layer of 0.25 drop rate. After these 2 blocks, a dense layer with a size defined as the number of categories to classify was added. A learning rate of 0.0001, batch size of 128, and 20 epochs were set by trial and error. We used TensorFlow 2.3 and Keras libraries to prepare the data and run the model. The analyses were performed with a computer equipped with an Intel Xeon Gold 6130 processor (Intel, Santa Clara, CA) and a NVIDIA Quadro RTX 4000 graphic processing unit (NVIDIA, Santa Clara, CA).

**Performance measures and statistical analysis**

For each image, the CNN model calculated the probability for each category (pleomorphic gastric lesions vs normal), with a given probability (Figure 3). A higher probability translated into a greater CNN prediction confidence. The software-generated heatmaps localized features that predicted a lesion probability (Figure 4). The CNN’s output was compared with a consensus classification provided by 2 WCE-experienced gastroenterologists.

At the first experiment, a 3-fold cross-validation was performed, with the division of the development data set into 3 even-sized image subsets. The primary performance measures included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. These measures were represented with their means and 95% confidence intervals. The precision-recall (PR) curve and area under the precision-recall curve were used to measure the performance of the model.

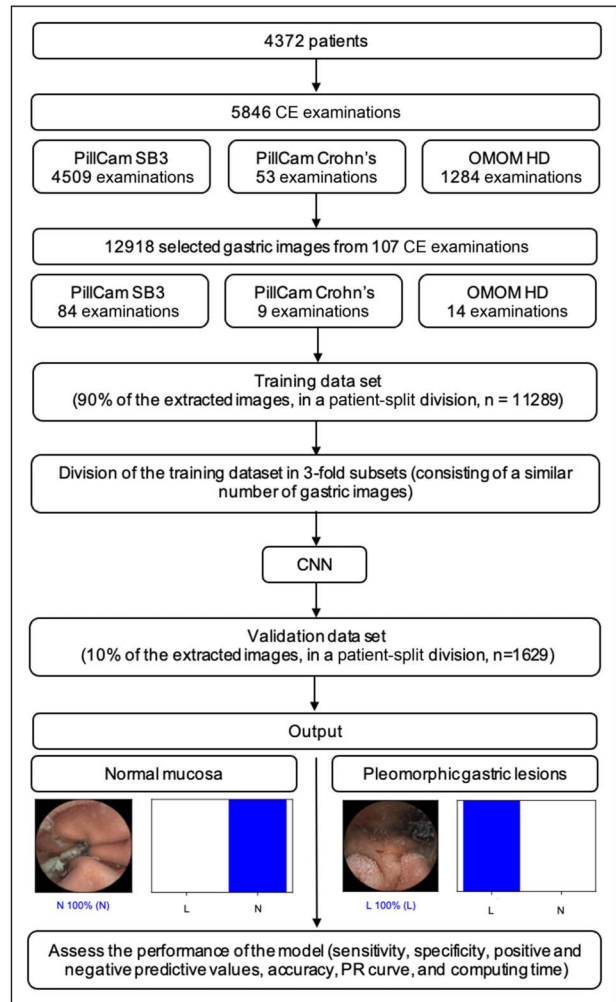
The performance results from the 3-fold subsets were used to identify the better parameters of the CNN. These parameters were then applied in the validation data set, which was evaluated afterward.

Statistical analysis was performed using Sci-Kit learn version 0.22.2 (20).

**RESULTS**

**Construction of the network**

From a total of 5,846 CE examinations, our group developed a CNN-based model with 12,918 gastric images. Each image was evaluated by the CNN, which predicted a classification,



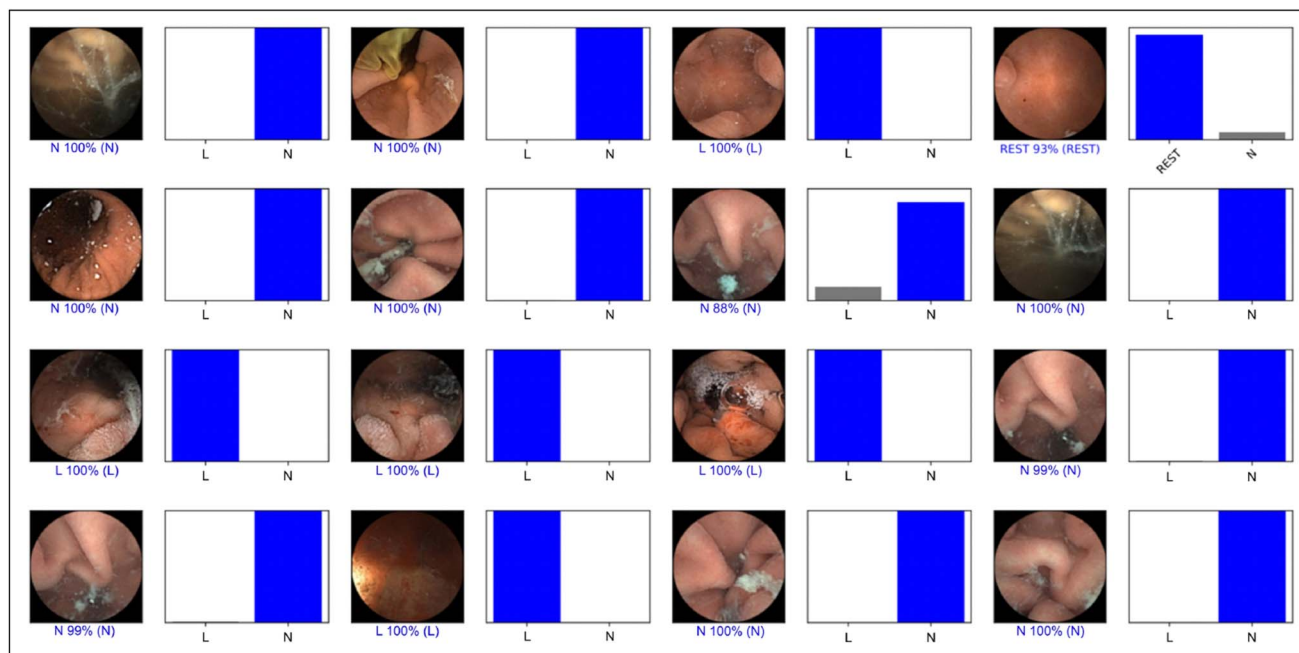
**Figure 2.** Study flowchart for the training and validation phases. CE, capsule endoscopy; CNN, convolutional neural network; L, pleomorphic gastric lesions; N, normal; PR, precision-recall.

subsequently compared with the experts’ labelling. The repeated data inputs in the subset folds of the training data set allowed the adjustment of the CNN parameters. Subsequently, the fine-tuned CNN performance was evaluated in the validation data set.

**Global performance of the network**

For the first analysis, a 3-fold cross-validation of the training data set was performed. The performance results of the 3-folds of the training model are presented in Table 1. Overall, the training data set had a mean sensitivity of 87.8%, specificity of 92.3%, PPV of 91.4%, and NPV of 89.2%. The mean accuracy of the model was 90.2%.

In the second analysis, the remaining 10% of the retrieved images were used as a validation data set for evaluation of the CNN’s performance. The confusion matrix between the trained CNN and experts’ classification is presented in Table 2. The CNN model identified pleomorphic gastric lesions with a sensitivity of 97.4%, specificity of 95.9%, PPV of 95.0%, NPV of 97.8%, and accuracy of 96.6% (Table 1). The model had an area under the precision-recall curve of 1.00.



**Figure 3.** Output obtained from the application of the CNN for pleomorphic gastric lesions. The bars represent the estimated probability by the CNN model. The finding with the highest probability was outputted as the predicted classification. The blue bars represent a correct prediction, whereas the red bars represent an incorrect prediction. CNN, convolutional neural network; L, pleomorphic gastric lesions; N, normal.

### Computational performance of the CNN

The CNN's image processing time of the validation data set was 115 images per second (8.7 milliseconds per frame), with a total viewing time of 14.2 seconds.

### DISCUSSION

In this proof-of-concept study, our group developed a WCE-based CNN capable of detecting multiple gastric lesions in both SBCE and CCE devices. This model demonstrated high performance levels in all the evaluated parameters, with 97.4% sensitivity, 95.9% specificity, and 96.6% global accuracy for pleomorphic gastric lesions. These results were achieved with an image processing time of 115 images per second. Furthermore, the diagnostic yield of the model was verified not only in SBCE devices but also in CCE, with 3 different types of CE devices. Thereby, our group reckons that the development of AI-powered WCE might change the landscape regarding the classification of pleomorphic gastric lesions (protuberant, vascular, ulcers, or erosions).

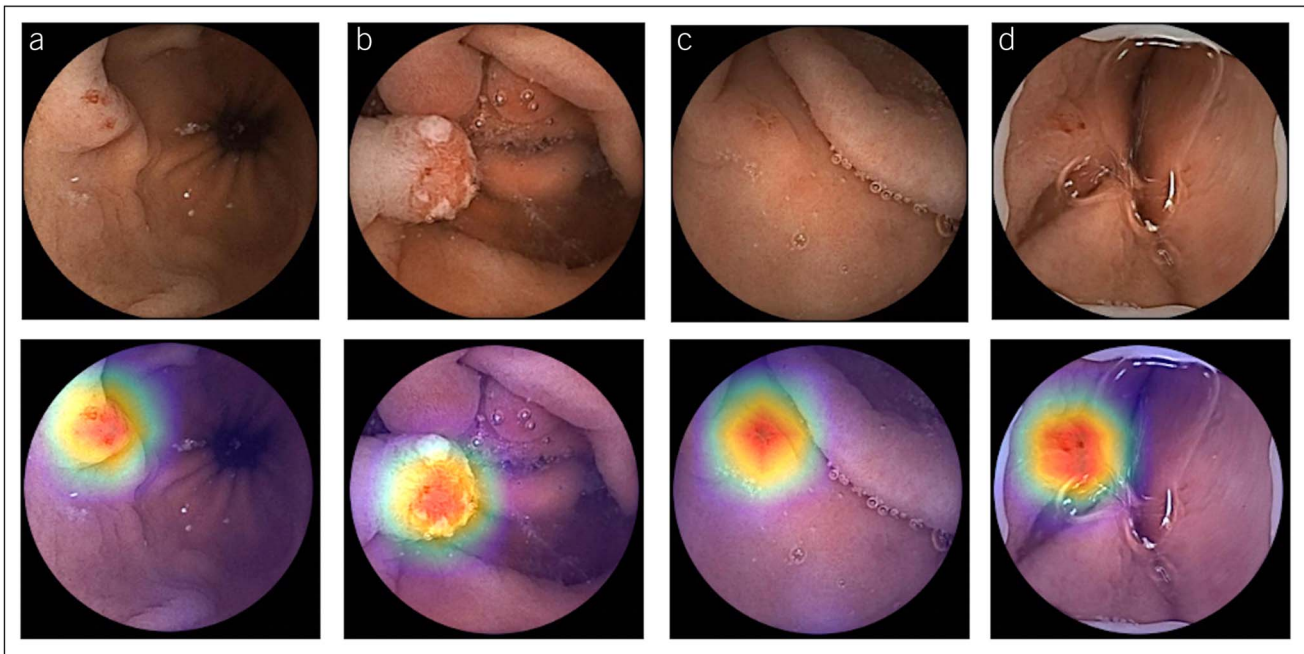
Furthermore, it is important to consider some methodology points about this study. The division between training and validation data sets was based on a patient-split design. Thus, all the given images from a single patient were included in the same data set. In the development of a CNN model, the inclusion of similar images in both data sets could imply an overfitting of the model (because the model would recognize a very similar image present in the training data set). Therefore, we believe that the risk of the model's overfitting was reduced. Moreover, a 3-fold cross-validation was performed. The data set was divided into 3 equal-sized, nonoverlapping folds, with images from a single examination allocated exclusively to one-fold. The average of these metrics across all iterations was used as an aggregate measure of the model's performance. This aggregated performance measure

provides a more reliable assessment of the model's ability to classify new, unseen images, increasing confidence in the model's generalizability. In addition, our group used PR curves instead of receiver-operating characteristic (ROC) curves for the model evaluation. In cases of data imbalance in a certain variable, ROC curves are known for being excessively optimistic in the evaluation of a model/biomarker performance (21), with PR curves being a proved alternative (22). In our study, this was defined by a larger representation of normal gastric mucosa than gastric lesions. Given our focus on determining all the lesion images, instead of identifying the commoner true negatives, which are implied in the ROC curve construction, the PR curve was then preferred.

The interoperability challenge is one of the main points of interest in the implementation of AI-based technologies in medicine (23), with the need for generalization of technology across multiple platforms and devices. The interoperability challenge has been a point of focus in the recent years in the electronic health records system (24), but a recent article by Tang et al stated interoperability between different systems to be a determinant factor for the application of AI tools in radiology (25). Therefore, the results of our work in 3 different CE devices, either in SBCE or CCE, are a proof of the interoperability of the CNN in different outlets, fundamental for its implementation in the clinical practice. This is, to our knowledge, the first CNN model capable of diagnosing pleomorphic gastric lesions in 3 different CE devices, comprising both SBCE and CCE.

Recently, the concept of WCE-based panendoscopy has been the focus of numerous studies (26). However, CE (particularly CCE) is a time and resource-consuming examination with a large number of image frames produced, which is a disadvantage against conventional endoscopy evaluation. Furthermore, despite numerous studies about CNN models in





**Figure 4.** Heatmaps obtained from the application of the CNN showing pleomorphic gastric lesions as identified by the CNN. CNN, convolutional neural network; L, pleomorphic gastric lesions; N, normal.

small bowel and colon evaluation by WCE (9,15), there are no published articles about CNN models for gastric evaluation in WCE. Therefore, implementation of a specific gastric CNN, associated with small bowel and colon CNN, is important for increasing the diagnostic yield and cost-effectiveness of WCE-based panendoscopy, giving it an advantage over conventional endoscopy by providing simultaneous evaluation of the full length of the GI tract in a minimally invasive single examination. Our specific gastric CNN, with proven results in both SBCE and CCE devices, may contribute to reduce WCE reading times while increasing its diagnostic yield, reducing the subjective bias in image evaluation by experts. The integration and application in the clinical practice of such CNN would be pivotal for the implementation of a minimally invasive WCE-based panendoscopy.

Nonetheless, there are intrinsic limitations of the WCE to be considered. WCE is performed without air insufflation and stomach cavity distention, and the devices' movement depends only on abdominal peristalsis. This results in a significant reduction in stomach surface visualization, mainly the more

proximal part and the fundus. The evaluation of gastric lesions by CE has regained a renewed interest with the development of magnetic controlled CE (MCE), a subtype of CE with good performance in evaluating gastric lesions, even in asymptomatic individuals (27,28). Xia et al (29) were the first to develop a CNN-based model for the detection of multiple gastric lesions and applied it to MCE. Their model focused on identifying 7 categories of gastric images (erosions, polyps, ulcers, submucosal tumors, xanthomas, normal mucosa, and invalid images). The CNN-based model had 96.2% sensitivity, 76.2% specificity, 16.0% PPV, 99.7% NPV, and 77.1% accuracy, with an area under the curve of 0.84. Although this study represents a significant breakthrough in the application of AI to minimally invasive techniques of gastric inspection, MCE is limited to very few research centers, contrary to widely available conventional WCE. In addition, MCE implies a learning curve for procedure performance (30). Besides, there is, to our knowledge, a lack of comparative studies between WCE and MCE diagnostic yield in the evaluation of gastric lesions. Whereas MCE may overcome the peristalsis dependence, it is still affected by the absence of

**Table 1.** Performance measures of the 3-fold cross-validation of the training data set and validation data set for detection of pleomorphic gastric lesions

→	Sensitivity, % (CI)	Specificity, % (CI)	PPV, % (CI)	NPV, % (CI)	Accuracy, % (CI)	AUPRC
Fold 1 (n = 3,754)	81.8 (79.9–83.5)	97.8 (97.1–98.4)	97.2 (96.3–97.9)	85.3 (84.0–86.5)	90.1 (89.1–91.0)	0.93
Fold 2 (n = 3,755)	90.8 (89.4–92.1)	83.8 (82.1–85.4)	83.8 (82.4–85.2)	90.8 (89.5–91.9)	87.2 (86.0–88.2)	0.96
Fold 3 (n = 3,780)	91.0 (89.5–92.2)	95.4 (94.3–96.2)	94.7 (93.6–95.6)	92.0 (90.9–93.0)	93.2 (92.4–94.0)	0.98
Training data set mean (n = 11,289)	87.8 (86.9–88.7)	92.3 (91.6–93.0)	91.4 (90.6–92.0)	89.2 (88.5–89.9)	90.2 (89.6–90.7)	
Validation data set (n = 1,629)	97.4 (96.0–98.4)	95.9 (94.4–97.1)	95.0 (93.3–96.3)	97.8 (96.7–98.6)	96.6 (95.6–97.4)	1.00

AUPRC, area under the precision-recall curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

**Table 2. Confusion matrix of the automatic detection vs the experts' classification in the validation data set of the CNN model**

		Experts' classification	
		N	L
CNN classification	N	865 (0.96)	19 (0.03)
	L	37 (0.04)	708 (0.97)

Number of cases (relative frequency).  
CNN, convolutional neural network; L, pleomorphic gastric lesions; N, normal gastric images.

stomach distention, with reduction in the proximal stomach visibility. In fact, a small study of MCE in healthy volunteers by Liao et al showed that, despite visualization of the more distal stomach regions by WCE in 100% of the patients, visualization of the gastric cardia and the fundus was only achieved in 82.4% and 85.3% of the patients, respectively (31). Moreover, our model showed better performance marks in all the evaluated parameters, including image processing. Thus, our group believes that the better performance of our model, combined with the worldwide availability of WCE devices, will contribute to make the generalization of this model more feasible. In addition, MCE primary focus is not a panendoscopic evaluation of the GI tract but a minimally invasive alternative for conventional EGD.

Our group developed the first CNN model for gastric lesion evaluation by SBCE and CCE with high sensitivity, specificity, accuracy, and image processing capacity. However, despite the promising results of our model, this is only a proof-of-concept study and the first step to the application of this model to a real clinical scenario.

Nonetheless, our study has several limitations. First, this study was conducted in a retrospective manner. Therefore, in the future, larger prospective multicentric studies are needed for assessment of the clinical utility of this tool. Furthermore, our results were based on still images, requiring studies with real-time WCE videos in the future.

In conclusion, the use of CNN models in clinical practice can become the standard of care in only a few years. In the gastroenterology field, optimization of WCE examinations with CNN-based technologies has been evolving recently, but the role of these systems for detection of gastric lesions in conventional WCE has not been explored yet.

Our CNN system was the first, to our knowledge, to detect gastric lesions with high accuracy and sensitivity, with excellent imaging processing times in SBCE and CCE devices. The application of these systems in clinical practice will favor the cost-effectiveness of WCE in a panendoscopy evaluation, with an associated standardization of the classification and reduction of the reading time of the examination. Larger multicentric prospective real-time studies are needed to confirm this proof-of-concept study.

#### CONFLICTS OF INTEREST

**Guarantor of the article:** Miguel Mascarenhas Saraiva, MD, PhD.  
**Specific author contributions:** M.M. and F.M.: equal contribution in study design, image extraction, drafting of the manuscript, and critical revision of the manuscript. T.R. and J.A.: bibliographic review, image extraction, and critical revision of the manuscript. P.C.

and M.M.: bibliographic review, drafting of the manuscript, and critical revision of the manuscript. J.F.: construction and development of the CNN, statistical analysis, and critical revision of the manuscript. P.A., H.C., M.M.S., and G.M.: study design and critical revision of the manuscript. All authors approved the final version of the manuscript.

**Financial support:** The authors recognize NVIDIA support for the graphic unit acquisition.

**Potential competing interests:** None to report.

## Study Highlights

### WHAT IS KNOWN

- ✓ Capsule endoscopy (CE) is a minimally invasive examination for evaluating the entire gastrointestinal tract.
- ✓ The diagnostic yield of CE for gastric lesions is suboptimal.

### WHAT IS NEW HERE

- ✓ Our group developed a convolutional neural network-based model for the classification of pleomorphic gastric lesions.
- ✓ The model's diagnostic yield was verified in both small bowel and colon CE devices.
- ✓ Artificial intelligence algorithms could increase the diagnostic yield of capsule panendoscopy.

## REFERENCES

- Levy I, Gralnek IM. Complications of diagnostic colonoscopy, upper endoscopy, and enteroscopy. *Best Pract Res Clin Gastroenterol* 2016; 30(5):705–18.
- Helmert RA, Dilling JA, Chaffee CR, et al. Overall cost comparison of gastrointestinal endoscopic procedures with endoscopist- or anesthesia-supported sedation by activity-based costing techniques. *Mayo Clin Proc Innov Qual Outcomes* 2017;1(3):234–41.
- Iddan G, Meron G, Glukhovskiy A, et al. Wireless capsule endoscopy. *Nature* 2000;405(6785):417.
- Pennazio M, Spada C, Eliakim R, et al. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy* 2015;47(4):352–86.
- Le Berre C, Trang-Poisson C, Bourreille A. Small bowel capsule endoscopy and treat-to-target in Crohn's disease: A systematic review. *World J Gastroenterol* 2019;25(31):4534–54.
- Rondonotti E, Pennazio M. Colon capsule for panendoscopy: A narrow window of opportunity. *Endosc Int Open* 2021;9(12):E1860–2.
- Chetcuti Zammit S, Sidhu R. Capsule endoscopy: Recent developments and future directions. *Expert Rev Gastroenterol Hepatol* 2021;15(2): 127–37.
- Kim JH, Nam SJ. Capsule endoscopy for gastric evaluation. *Diagnostics (Basel)* 2021;11(10):1792.
- Picciorelli S, Mussetto A, Bellumat A, et al. New generation express view: An artificial intelligence software effectively reduces capsule endoscopy reading times. *Diagnostics (Basel)* 2022;12(8):1783.
- Richards BA, Lillicrap TP, Beaudoin P, et al. A deep learning framework for neuroscience. *Nat Neurosci* 2019;22(11):1761–70.
- Hughes JW, Olgin JE, Avram R, et al. Performance of a convolutional neural network and explainability technique for 12-lead electrocardiogram interpretation. *JAMA Cardiol* 2021;6(11):1285–95.
- Kim GR, Lee E, Kim HR, et al. Convolutional neural network to stratify the malignancy risk of thyroid nodules: Diagnostic performance compared with the American College of Radiology Thyroid Imaging Reporting and Data System Implemented by Experienced Radiologists. *AJNR Am J Neuroradiol* 2021;42(8):1513–9.
- Agarwala S, Mata DA, Hafeez F. Accuracy of a convolutional neural network for dermatological diagnosis of tumours and skin lesions in a clinical setting. *Clin Exp Dermatol* 2021;46(7):1310–1.

14. Mascarenhas M, Afonso J, Andrade P, et al. Artificial intelligence and capsule endoscopy: Unravelling the future. *Ann Gastroenterol* 2021; 34(3):300–9.
15. Soffer S, Klang E, Shimon O, et al. Deep learning for wireless capsule endoscopy: A systematic review and meta-analysis. *Gastrointest Endosc* 2020;92(4):831–9.e8.
16. Rondonotti E, Spada C, Adler S, et al. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) technical review. *Endoscopy* 2018;50(4):423–46.
17. Rosa B, Margalit-Yehuda R, Gatt K, et al. Scoring systems in clinical small-bowel capsule endoscopy: All you need to know! *Endosc Int Open* 2021; 9(6):E802–23.
18. Leenhardt R, Li C, Koulaouzidis A, et al. Nomenclature and semantic description of vascular lesions in small bowel capsule endoscopy: An international Delphi consensus statement. *Endosc Int Open* 2019;7(3): E372–9.
19. Korman LY, Delvaux M, Gay G, et al. Capsule endoscopy structured terminology (CEST): Proposal of a standardized and structured terminology for reporting capsule endoscopy procedures. *Endoscopy* 2005;37(10):951–9.
20. Pedregosa FVG, Gramfort A, Michel V, et al. Scikit-learn: Machine learning in python. *J Mach Learn Res* 2011;12:2825.
21. Movahedi F, Padman R, Antaki JF. Limitations of receiver operating characteristic curve on imbalanced data: Assist device mortality risk scores. *J Thorac Cardiovasc Surg* 2023;165(4):1433–42.e2.
22. Fu GH, Yi LZ, Pan J. Tuning model parameters in class-imbalanced learning with precision-recall curve. *Biom J* 2019;61(3):652–64.
23. He J, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019;25(1):30–6.
24. Janett RS, Yeracaris PP. Electronic medical records in the American Health System: Challenges and lessons learned. *Cien Saude Colet* 2020; 25(4):1293–304.
25. Tang A, Tam R, Cadrin-Chenevert A, et al. Canadian association of radiologists white paper on artificial intelligence in radiology. *Can Assoc Radiol J* 2018;69(2):120–35.
26. Cortegoso Valdivia P, Elosua A, Houdeville C, et al. Clinical feasibility of panintestinal (or panenteric) capsule endoscopy: A systematic review. *Eur J Gastroenterol Hepatol* 2021;33(7):949–55.
27. Zhang Y, Zhang Y, Huang X. Development and application of magnetically controlled capsule endoscopy in detecting gastric lesions. *Gastroenterol Res Pract* 2021;2021:2716559.
28. Li J, Ren M, Yang J, et al. Screening value for gastrointestinal lesions of magnetic-controlled capsule endoscopy in asymptomatic individuals. *J Gastroenterol Hepatol* 2021;36(5):1267–75.
29. Xia J, Xia T, Pan J, et al. Use of artificial intelligence for detection of gastric lesions by magnetically controlled capsule endoscopy. *Gastrointest Endosc* 2021;93(1):133–9.e4.
30. Hu J, Yunpeng H, Wang S, et al. The operation training for magnetically controlled capsule endoscopy. *Eur J Gastroenterol Hepatol* 2021;33(1): 46–9.
31. Liao Z, Duan XD, Xin L, et al. Feasibility and safety of magnetic-controlled capsule endoscopy system in examination of human stomach: A pilot study in healthy volunteers. *J Interv Gastroenterol* 2012;2(4):155–60.

---

**Open Access** This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.