

Spatiotemporal analysis of small bowel capsule endoscopy videos for outcomes prediction in Crohn's disease

Raizy Kellerman, Amit Bleiweiss, Shimrit Samuel, Reuma Margalit-Yehuda, Estelle Aflalo, Oranit Barzilay, Shomron Ben-Horin, Rami Eliakim, Eyal Zimlichman, Shelly Soffer^{id}, Eyal Klang and Uri Kopylov^{id}

Ther Adv Gastroenterol

2023, Vol. 16: 1–11

DOI: 10.1177/
17562848231172556

© The Author(s), 2023.
Article reuse guidelines:
sagepub.com/journals-
permissions

Abstract

Background: Deep learning techniques can accurately detect and grade inflammatory findings on images from capsule endoscopy (CE) in Crohn's disease (CD). However, the predictive utility of deep learning of CE in CD for disease outcomes has not been examined.

Objectives: We aimed to develop a deep learning model that can predict the need for biological therapy based on complete CE videos of newly-diagnosed CD patients.

Design: This was a retrospective cohort study. The study cohort included treatment-naïve CD patients that have performed CE (SB3, Medtronic) within 6 months of diagnosis. Complete small bowel videos were extracted using the RAPID Reader software.

Methods: CE videos were scored using the Lewis score (LS). Clinical, endoscopic, and laboratory data were extracted from electronic medical records. Machine learning analysis was performed using the TimeSformer computer vision algorithm developed to capture spatiotemporal characteristics for video analysis.

Results: The patient cohort included 101 patients. The median duration of follow-up was 902 (354–1626) days. Biological therapy was initiated by 37 (36.6%) out of 101 patients. TimeSformer algorithm achieved training and testing accuracy of 82% and 81%, respectively, with an Area under the ROC Curve (AUC) of 0.86 to predict the need for biological therapy. In comparison, the AUC for LS was 0.70 and for fecal calprotectin 0.74.

Conclusion: Spatiotemporal analysis of complete CE videos of newly-diagnosed CD patients achieved accurate prediction of the need for biological therapy. The accuracy was superior to that of the human reader index or fecal calprotectin. Following future validation studies, this approach will allow for fast and accurate personalization of treatment decisions in CD.

Keywords: artificial intelligence, capsule endoscopy, Crohn's disease

Received: 14 December 2022; revised manuscript accepted: 12 April 2023.

Introduction

Capsule endoscopy (CE) is a prime modality for diagnosis and follow-up in Crohn's disease (CD). Complete evaluation of the entire small bowel using CE or cross-sectional modality is recommended in all newly-diagnosed CD patients, or patients in need of re-evaluation for non-explained symptoms.^{1,2} CE is an accurate predictor of pending relapse in CD patients in clinical remission,³ and can change disease classification

in over 50% of CD patients.⁴ Location of the lesions is of importance as well, as jejunal CD was demonstrated to be associated with worse long-term prognosis and higher risk of surgery.⁵

Artificial intelligence (AI) and deep learning techniques are being rapidly implemented in both research and clinical practice gastroenterology, mostly for real-time polyp detection during colonoscopy.^{6–8} Additional fields for evaluation of AI

Correspondence to:

Shelly Soffer
Department of Internal
Medicine B, Assuta
Medical Center, 7747629,
Ashdod, Israel

Ben-Gurion University of
the Negev, Be'er Sheva,
Israel
soffer.shelly@gmail.com

Raizy Kellerman
Amit Bleiweiss
Shimrit Samuel
Estelle Aflalo
Intel, Petach Tikva, Israel

Reuma Margalit-Yehuda
Shomron Ben-Horin
Rami Eliakim
Uri Kopylov
Department of
Gastroenterology, Sheba
Medical Center, Tel
Hashomer, Israel and Tel
Aviv University, Tel Aviv,
Israel

Oranit Barzilay
Department of Internal
Medicine F, Sheba Medical
Center, Tel Hashomer,
Israel and Tel Aviv
University, Tel Aviv, Israel

Eyal Zimlichman
Sheba ARC and Hospital
Management, Sheba
Medical Center, Tel
Hashomer, Israel and Tel
Aviv University, Tel Aviv,
Israel

Eyal Klang
Sheba ARC, Sheba Medical
Center, Tel Hashomer,
Israel and Tel Aviv
University, Tel Aviv, Israel

in Gastrointestinal tract (GI) assist in early detection of gastric neoplasia,^{9,10} Barrett's esophagus,¹¹ endoscopic ultrasound,^{12,13} and grading of mucosal inflammation in ulcerative colitis.^{14–18}

An additional field with fast development of AI research is CE, with several publications evaluating deep learning for automated detection of inflammatory lesions,^{19–26} vascular lesions,^{27,28} protruding lesions and masses,²⁹ and scoring of bowel cleanliness.³⁰ The accuracy for most of these tasks was reported to be at least 90% and frequently much higher.

However, to date the AI research in GI was mostly focused on identification of individual lesions, and this goal seems to be quite successfully achieved in luminal endoscopy. However, the bigger and perhaps more important challenge is the potential utilization of AI for prognostication and personalization of treatment approach. Such evidence is still extremely scarce.

In CD, predictors of disease prognosis and response to treatment are still severely lacking.^{31–33} The impact of mucosal healing and the lesion size on disease outcomes in CD has been well described.^{34,35} It is likely that the vast abundance of visual data contained in the complete videos of endoscopic procedures such as CE (containing up to 12,000 still images per film) may contain factors and subtle findings that escape the human eye, however, could be picked by AI algorithms. We hypothesized that AI analysis of CE videos performed at diagnosis could be an accurate predictor of treatment outcomes.

The main aim of our study was to evaluate the utility of AI analysis of complete CE videos of CD patients at inception for prediction of the need for biological therapy.

Methodology

Cohort definition

This was a retrospective cohort study. The reporting of this study conforms to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. In this study, we included CD patients followed by the gastroenterology department of Sheba Medical Center between January 2011 and March 2021. Only patients who performed CE within 6 months

of diagnosis and before initiation of immunomodulator or biological therapy were included. Patients with less than 6 months of follow-up after CE were excluded. Demographic, clinical, and endoscopic data were extracted from patients' electronic medical records.

Capsule endoscopy

All CE procedures were performed using PillCam SB2 or SB3 small bowel capsules (Medtronic, Yokneam, Israel) and reviewed using the Reader software version 8 or 9 (Medtronic, Yokneam, Israel). The CE videos were graded using the Lewis score (LS).³⁶ We extracted complete small bowel videos starting with the first duodenal image and ending with the first cecal image using the RAPID Reader software v9.0 (Medtronic, Yokneam, Israel). The videos were saved in MPEG format.

Study outcome

The study outcome was the initiation of biological therapy within the duration of follow-up.

Machine learning algorithm

Machine learning analysis was performed by Intel Israel. For analyzing videos, we used the TimeSformer, Facebook Research, Menlo Park, California architecture.³⁷ An overview of the TimeSformer architecture is presented in Figure 1. This is a state-of-the-art algorithm introduced by Facebook. It is based on the Transformer paradigm from the natural language processing domain^{37,38} (Supplemental Section A).

We used a fivefold cross-validation train-test random split: in each split four folds were used as the train set and the fifth fold was used as the validation set. The raw training videos were augmented (Supplemental Table 1) after the train-test split.

Detailed hyperparameters of the experiments are presented in the supplements. Three TimeSformer models were experimented upon: FT K400, $8 \times 32 \times 224$; FT K600, $8 \times 32 \times 224$; and FT K400, $16 \times 16 \times 448$ (Supplemental Section B).

Model interpretability

To understand the influence of spatial and temporal features, respectively, we used Intel

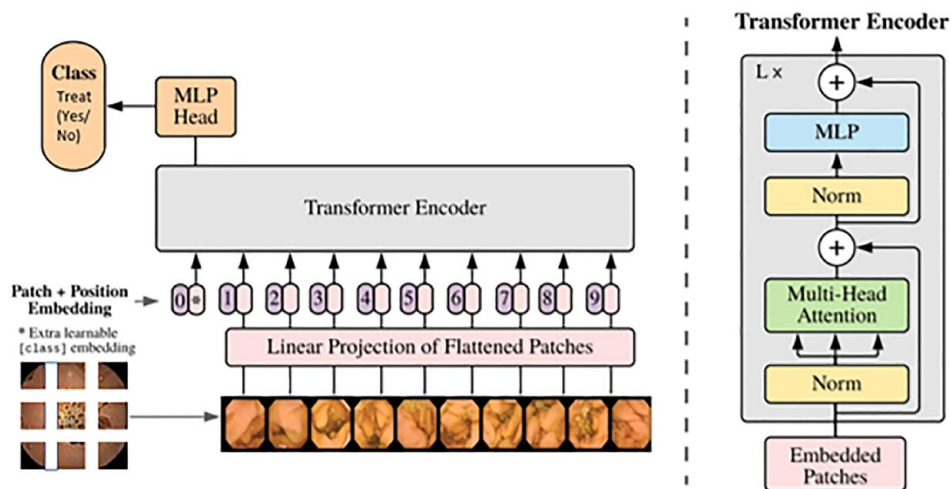


Figure 1. An overview of the TimeSformer architecture. On the right side: the video self-attention block, applied on the embedded patches. On the left side: the end-to-end architecture. A multilayer perceptron (MLP) is applied both at the end of each block and to the projected and concatenated vectors from all heads.

Labs' *InterpreT* toolbox. This toolbox allows us to investigate the model behavior and understand its internal structure (Supplemental Section C).

Statistical methods for clinical data

Descriptive statistics were presented as means \pm standard deviations/medians with interquartile ratios (IQRs) for continuous variables and percentages for categorical variables. Continuous variables were analyzed by *t*-test/Mann–Whitney test and categorical variables by chi-square/Fisher's exact test. A receiver operator curve (ROC) analysis was performed for the models, and LS and fecal calprotectin predictions of the study outcome were analyzed. A two-tailed *p*-value < 0.05 was considered statistically significant. The analysis was performed using IBM SPSS statistics (version 27.0) (Armonk, NY, USA).

Results

Cohort description

One hundred and one patients were recruited. Clinical and demographic characteristics of the patients appear in Table 1. The median LS was 450 (IQR 225–1012).

Table 1. Clinical and demographic characteristics of the included patients.

Characteristics	Median	N (%)
Age, years (median, IQR)	27 [23–42]	
Age at diagnosis, years (median, IQR)	25 [20–40]	
Gender, women		54 (53.5)
CD location		
Ileal		88 (87.1)
Colonic		1 (1)
Ileocolonic		12 (11.9)
CD phenotype		
Non-stricturing non-penetrating		92 (91.1)
Penetrating		6 (5.9)
Stricturing		3 (4)
Smoking		
Never		66 (73)
Active		17 (19)
Previous		8 (9)
Lewis score (median, IQR)	450 [225–1012]	
Fecal calprotectin, mg/g stool (median, IQR)	233 (99–656)	
CD, Crohn's disease; IQR, interquartile ratio.		

Table 2. Comparison of clinical characteristics between patients that did or did not initiate biological therapy within the duration of follow-up.

Characteristics	Received biological therapy		<i>p</i> (univariate)	<i>p</i> (multivariate)
	No (<i>N</i> =37)	Yes (<i>N</i> =64)		
	Median, IQR (<i>N</i> %)	Median, IQR (<i>N</i> %)		
Age, years (median, IQR)	27 (23–41)	27 (22–44)	0.78	
Age at diagnosis, years (median, IQR)	26 (21–38.5)	24 (20–41)	0.64	
CD location				
Ileal	59 (90.6)	30 (81.1)	0.19	
Colonic	1 (1.5)	1 (2.6)		
Ileocolonic	5 (7.7)	7 (18.9)		
CD phenotype				
Non-stricturing non-penetrating	62 (96.9)	30 (81.1)	0.48	
Penetrating	1 (1.5)	5 (13.5)		
Stricturing	1 (1.5)	2 (5.4)		
Smoking				
Never	44 (80%)	21 (60%)	0.11	
Active	7 (12.7%)	10 (28.6%)		
Previous	4 (7.3%)	4 (11.4%)		
Lewis score (median, IQR)	393 (225–900)	618 (450–1350)	0.01	0.46
Fecal calprotectin, mg/g stool (median, IQR)	130 (93–350)	360 (203–1000)	0.02	0.67
CD, Crohn's disease; IQR, interquartile ratio.				

Thirty-seven patients (36.8%) initiated biological therapy within the duration of follow-up with a median time of 17 months (4.7–38.5) for therapy. The initiated biological treatments included adalimumab (18, 48.6%), infliximab (13, 35.1%), or vedolizumab (6, 16.2%).

On univariate analysis, CD phenotype, LS, and fecal calprotectin were associated with initiation of biological therapy; however, on multivariate analysis, none retained significance (Table 2). For LS, the AUC for predicting biological treatment was 0.70 and for calprotectin 0.74.

Machine learning results

The optimal model achieved AUC 0.86 with accuracy 0.81. Below is a summarization of the TimeSformer experiments with respect to

the different model configurations (Table 3, Figure 2).

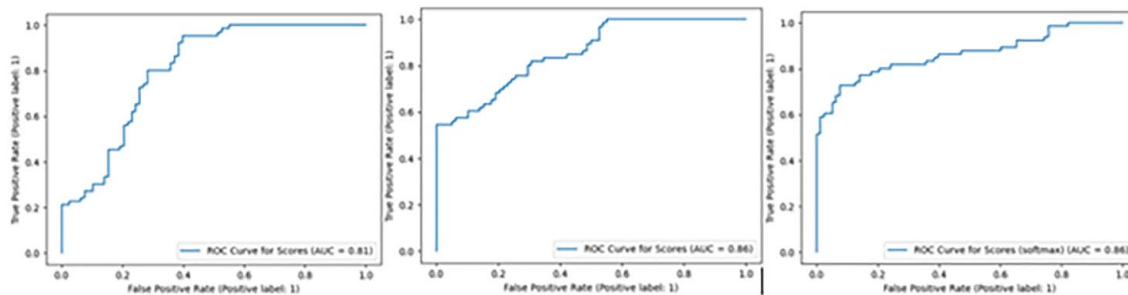
Model interpretability

Hidden state representations in a two-dimensional space. Supplemental Figure 1 shows the t-distributed stochastic neighbor embedding (t-SNE) representations of the CLS token of each video. This figure shows how positive and negative videos are clustered differently throughout the network's layers. Especially, two distinct clusters are seen at the last layer.

Then, we focused on the t-SNE representation of the regions of each frame. In Figure 3, pixel id represents the same region in each image, and example ID represents the video level.

Table 3. Machine learning algorithms performance for prediction of biological therapy in CD patients.

Machine learning model	AUC	Precision/positive predictive value	Sensitivity/true positive rate/recall	Specificity/true negative rate	Fall-out/false positive rate	Miss rate/false negative rate
FT K400, $8 \times 32 \times 224$	0.80	0.71	0.78	0.87	0.12	0.29
FT K600, $8 \times 32 \times 224$	0.85	0.82	0.71	0.92	0.07	0.28
FT K400, $16 \times 32 \times 448$	0.86	0.81	0.75	0.84	0.15	0.25

**Figure 2.** The ROC AUC plots for models FT K400 $8 \times 32 \times 224$, FT K600 $8 \times 32 \times 224$, and FT K400 $16 \times 32 \times 448$, respectively. AUC, Area under the ROC Curve.

Attention heatmap for ulcer detection. When comparing our results to ResNet101 Convolutional Neural Network classifier trained on a set of colonoscopy images, the TimeSformer attention mechanism was better focused on the ulcers and their localization (Figure 4), as seen in the Grad-CAM heatmaps.

Discussion

Our study demonstrates that machine learning analysis of complete CE videos of CD patients performed at initial diagnosis can accurately predict the need for biological therapy. The performance of TimeSformer algorithm was superior to that of the traditional LS or of fecal calprotectin.

In the recent years, AI algorithms achieved excellent accuracy in identification of findings on still images in multiple fields in gastroenterology, including CE.^{19,20,22,26,39,40} Moreover, real-time diagnosis of colonic polyps on colonoscopy is currently possible and was demonstrated to be associated with an improved polyp detection rate^{6–8}; AI technology for colonoscopy has already matured into clinically available and FDA-approved tools such as GI Genius (Medtronic, Dublin, Ireland).

However, it is quite plausible that the full-fledged potential of the use of AI in Inflammatory bowel disease (IBD) is yet to be realized. Prediction of disease course and outcomes is among the crucial challenges in CD; despite the massive abundance of research targeting personalization of treatment approach, the practical implication for the daily practice is still scarce.^{31,41–43} One of the cornerstones of the modern treatment approach in IBD is the achievement of mucosal healing^{1,2,44}; striving for complete mucosal healing in IBD is well justified as patients in mucosal healing have consistently better long- and short-term disease outcomes.⁴⁵ CE is a valuable diagnostic and monitoring tool in CD^{1,2} as it allows for accurate panenteric assessment of bowel inflammation. In established CD patients, proximal disease location is positively associated with a risk of intestinal resection.⁵ CE identifies mild to moderate mucosal inflammation in over 80% of CD patients in clinical remission, including the majority of patients in both clinical and biomarker remission⁴⁶; such smoldering inflammation was an accurate predictor of relapse within 24 months.³

An average complete CE video may include up to 10,000 images. The existing indices for grading inflammation on CE^{36,47–52} are based on a small number of semiquantitative parameters such as

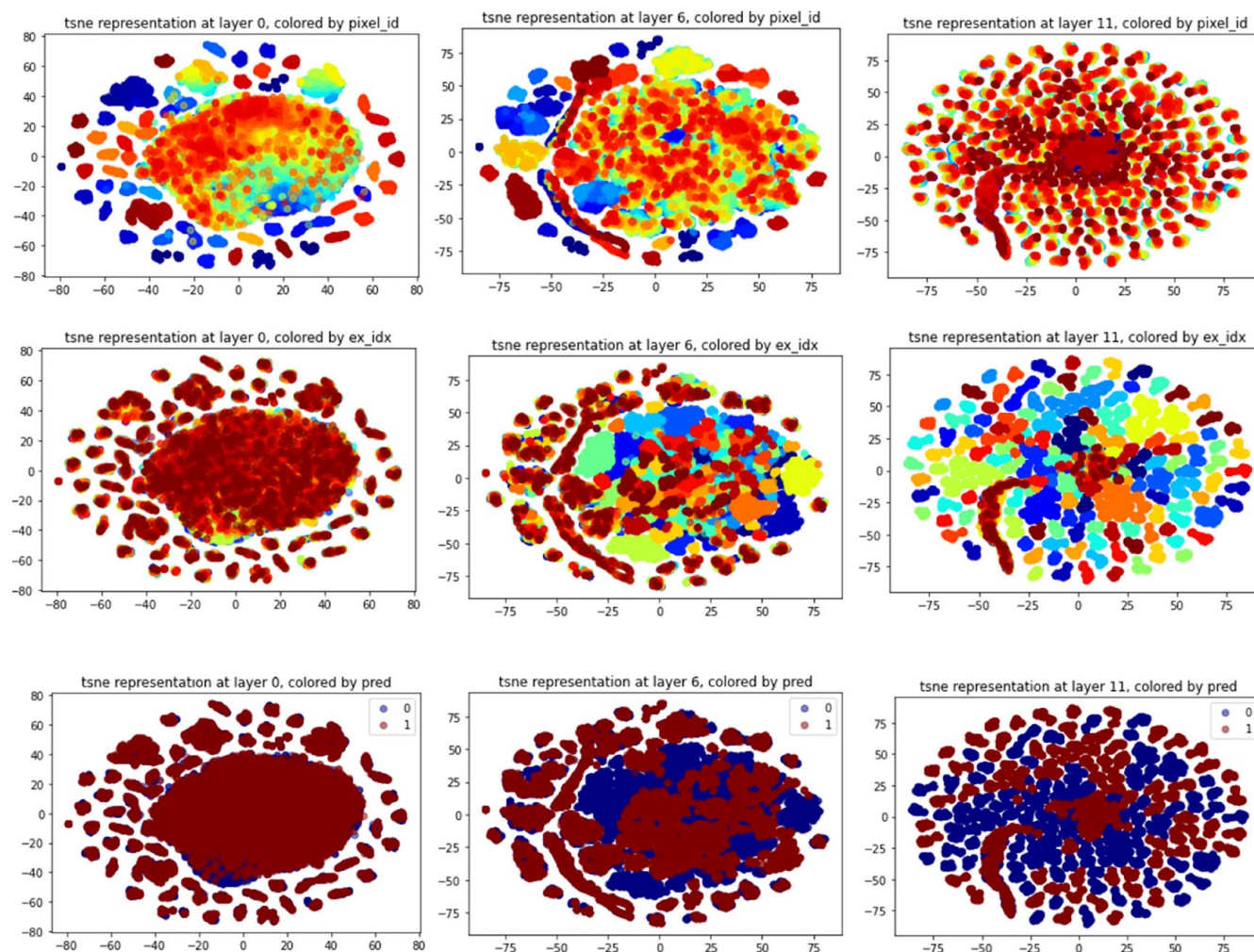


Figure 3. The evolved clustering of the frame regions after t-SNE embedding through the layers. The first row of the figure shows the evolution of the regions' representation colored by the pixel id while the second row shows this evolution colored by the example ID. These figures show that in the first layers, the clusters are location-specific: pixels of the same part of any different frames will be closer. However, at the end of the network, the clusters are example related. This implies that location embedding is represented in the first layers. In contrast, prediction is at the example level and not at region one, and it is represented in the last layers of the network. The third row shows that there is a separation between positive and negative videos in the last layer, while at the first layers, they were completely mixed and overlapping.

the number of ulcers (none/few/multiple), disease of ulceration (based on eye-balling of the extent), and presence of structures. Clearly, only a small amount of the available optical information is incorporated and processed in any of these scores. Thus, AI analysis of the entire available visual data may hold additional promise for the complete assessment of mucosal inflammatory burden. However, we assumed that not only the complete extent and morphology of the ulcerations and strictures may be of prognostic importance but the anatomical distribution of the

pathologies, which can only be gauged by the temporal location in relation to the first and last images, should also be of importance. Thus, a mere segmentation of videos to individual still images may not have been sufficient for the purpose.

Our approach is unique in that it classifies videos using both space and time features. To this end, we capture the sequential context by learning spatiotemporal features from the video capsule data, giving balanced representations to the local pixel

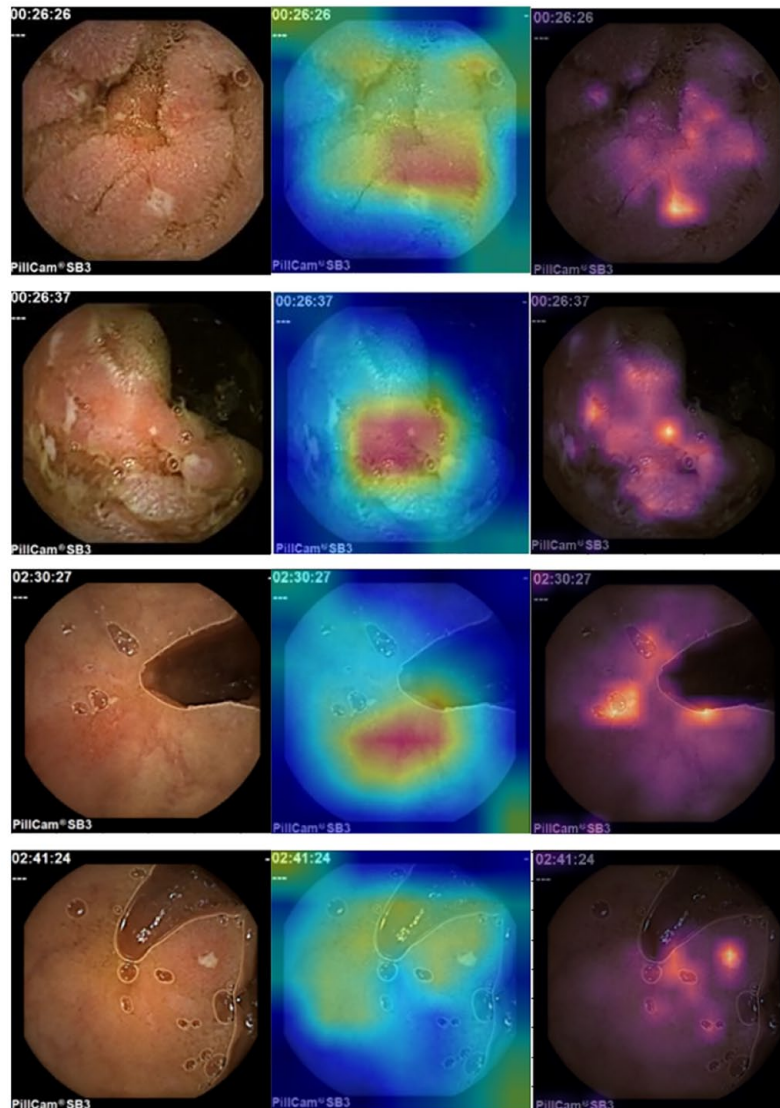


Figure 4. Ulcer detection – comparison of Grad-CAM visualizations to attention head visualizations to the CLS token in the spatial dimension. As can be seen, in the first two examples, the Grad-CAM method succeeds in detecting the lesions but with a diffused heatmap that might not be precise enough. The second two examples show how Grad-CAM also fails to detect ulcers correctly. Focusing on the attentions to the CLS token in the spatial dimension at the head (layer=3, head=8), we can see how the model indirectly learned to prioritize the ulcers. Compared to the Grad-CAM heatmaps, we can see how the detection is better focused on the ulcers, and we also see that even when Grad-CAM fails to detect the lesion, our model and this specific attention head succeed in localizing it.

structure as well as relationships across the video frames in time. TimeSformer algorithm was developed in order to capture and identify motion, such as sports activity, so we assumed that could be well suited to address the task of spatiotemporal video analysis.

Our study has several limitations. Primarily, a larger sample size could have resulted in

improved accuracy. For classification, we used a relatively ‘soft’ outcome of biological initiation. It is clear that the decision to initiate biological therapy is very heterogenous and could be influenced by a multitude of parameters such as patient and physician preferences, extraluminal disease characteristics (perianal disease, extraintestinal manifestations, etc.). It would have been preferable to select a less-biased

outcome such as need for surgery; however, such a cohort would be almost impossible to compile as patients with intestinal strictures are usually excluded from performing CE after evaluation of small bowel patency. In our cohort that included patients followed in a large tertiary center for 9 years, only one patient required surgery. In addition, our cohort included patients with relatively mild disease, as signified by LS and fecal calprotectin levels. In addition, our study focused on prediction of the outcome and did not suggest an algorithm for mucosal burden quantification. Such an algorithm would be of great practical utility and should be investigated in future studies.

Our current effort lays the groundwork for the creation of automated AI-based predictive systems for CE in CD. In fact, completely automated reading of AI can be easily feasible with the currently available tools for detection of pathologies, as they have already matured for excellent accuracy; nonetheless the ultimate goal of any diagnostic tool in early CD would be to predict the treatment outcome and the risk of complications, and for that purpose it appears that a spatiotemporal algorithmic approach that is able to incorporate disease distribution and location as well as the potential impact on GI motility is capable of catering to the task. The current algorithm requires further validation in extensive prospective cohorts, either as a standalone solution or as a component in a multiomic predictive model potentially incorporating multiple biological and visual inputs.

Declarations

Ethics approval and consent to participate

The Sheba Medical Center Institutional Review Board granted approval for this study (5039-18). Patient consent was waived due to a retrospective and anonymized format of the study.

Consent for publication

Not applicable.

Author contribution(s)

Raizy Kellerman: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing – original draft.

Amit Bleiweiss: Conceptualization; Formal analysis; Investigation; Writing – review & editing.

Shimrit Samuel: Conceptualization; Formal analysis; Writing – review & editing.

Reuma Margalit-Yehuda: Conceptualization; Data curation; Writing – review & editing.

Estelle Aflalo: Conceptualization; Formal analysis; Writing – review & editing.

Oranit Barzilay: Conceptualization; Data curation; Writing – review & editing.

Shomron Ben-Horin: Conceptualization; Data curation; Writing – review & editing.

Rami Eliakim: Conceptualization; Writing – review & editing.

Eyal Zimlichman: Conceptualization; Data curation; Writing – review & editing.

Shelly Soffer: Data curation; Writing – review & editing.

Eyal Klang: Conceptualization; Data curation; Investigation; Writing – review & editing.

Uri Kopylov: Conceptualization; Data curation; Visualization; Writing – original draft.

Acknowledgements

None.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by a grant from The Leona M. and Harry B. Helmsley Charitable Trust.

Competing interests

UK: research support – Medtronic, Janssen, Takeda. Speaker/advisory fees – AbbVie, BMS, Janssen, Pfizer, Takeda, MSD, Rafa. RE: Speaker for Takeda, Janssen, and Medtronic. SBH: received consulting and advisory board fees and/or research support from AbbVie, MSD, Janssen, Takeda, Pfizer, GSK, and CellTrion. EK, SheS, RMY, OB, and EZ: no competing interests. RK, AB, EA, and ShiS: employees of Intel Inc.

Availability of data and materials

The data underlying this article cannot be shared publicly due to the privacy of individuals that

participated in the study. The data will be shared on reasonable request to the corresponding author.

ORCID iDs

Shelly Soffer  <https://orcid.org/0000-0002-7853-2029>

Uri Kopylov  <https://orcid.org/0000-0002-7156-0588>

Supplemental material

Supplemental material for this article is available online.

References

1. Maaser C, Sturm A, Vavricka SR, *et al.* ECCO-ESGAR guideline for diagnostic assessment in IBD part 1: initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis* 2019; 13: 144–164.
2. Sturm A, Maaser C, Calabrese E, *et al.* ECCO-ESGAR guideline for diagnostic assessment in IBD part 2: IBD scores and general principles and technical aspects. *J Crohns Colitis* 2019; 13: 273–284.
3. Ben-Horin S, Lahat A, Amitai MM, *et al.* Assessment of small bowel mucosal healing by video capsule endoscopy for the prediction of short-term and long-term risk of Crohn's disease flare: a prospective cohort study. *Lancet Gastroenterol Hepatol* 2019; 4: 519–528.
4. Greener T, Klang E, Yablecovitch D, *et al.* The impact of magnetic resonance enterography and capsule endoscopy on the re-classification of disease in patients with known Crohn's disease: a prospective Israeli IBD research nucleus (IIRN) study. *J Crohns Colitis* 2016; 10: 525–531.
5. Flamant M, Trang C, Maillard O, *et al.* The prevalence and outcome of jejunal lesions visualized by small bowel capsule endoscopy in Crohn's disease. *Inflamm Bowel Dis* 2013; 19: 1390–1396.
6. Repici A, Badalamenti M, Maselli R, *et al.* Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. *Gastroenterology* 2020; 159: 512–520.e7.
7. Repici A, Spadaccini M, Antonelli G, *et al.* Artificial intelligence and colonoscopy experience: lessons from two randomised trials. *Gut* 2022; 71: 757–765.
8. 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* 2021; 93: 77–85.e6.
9. Wu L, Xu M, Jiang X, *et al.* Real-time artificial intelligence for detecting focal lesions and diagnosing neoplasms of the stomach by white-light endoscopy (with videos). *Gastrointest Endosc* 2022; 95: 269–280.e6.
10. Klang E, Barash Y, Levartovsky A, *et al.* Differentiation between malignant and benign endoscopic images of gastric ulcers using deep learning. *Clin Exp Gastroenterol* 2021; 14: 155–162.
11. Struyvenberg MR, de Groof AJ, van der Putten J, *et al.* A computer-assisted algorithm for narrow-band imaging-based tissue characterization in Barrett's esophagus. *Gastrointest Endosc* 2021; 93: 89–98.
12. Goyal H, Mann R, Gandhi Z, *et al.* Application of artificial intelligence in pancreaticobiliary diseases. *Ther Adv Gastrointest Endosc* 2021; 14: 2631774521993059.
13. Kuwahara T, Hara K, Mizuno N, *et al.* Usefulness of deep learning analysis for the diagnosis of malignancy in intraductal papillary mucinous neoplasms of the pancreas. *Clin Transl Gastroenterol* 2019; 10: 1–8.
14. Gottlieb K, Requa J, Karnes W, *et al.* Central reading of ulcerative colitis clinical trial videos using neural networks. *Gastroenterology* 2021; 160: 710–719.e2.
15. Ghoshal UC, Rai S, Kulkarni A, *et al.* Prediction of outcome of treatment of acute severe ulcerative colitis using principal component analysis and artificial intelligence. *JGH Open* 2020; 4: 889–897.
16. Takenaka K, Ohtsuka K, Fujii T, *et al.* Development and validation of a deep neural network for accurate evaluation of endoscopic images from patients with ulcerative colitis. *Gastroenterology* 2020; 158: 2150–2157.
17. Yao H, Najarian K, Gryak J, *et al.* Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2021; 93: 728–736.e1.
18. Stidham RW, Liu W, Bishu S, *et al.* Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis. *JAMA Netw Open* 2019; 2: e193963.
19. Klang E, Barash Y, Margalit RY, *et al.* Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule

- endoscopy. *Gastrointest Endosc* 2020; 91: 606–613.e2.
20. Klang E, Grinman A, Soffer S, *et al.* Automated detection of Crohn's disease intestinal strictures on capsule endoscopy images using deep neural networks. *J Crohns Colitis* 2021; 15: 749–756.
21. Klang E, Kopylov U, Mortensen B, *et al.* A convolutional neural network deep learning model trained on CD ulcers images accurately identifies NSAID ulcers. *Front Med* 2021; 8: 656493.
22. Barash Y, Azaria L, Soffer S, *et al.* Ulcer severity grading in video capsule images of patients with Crohn's disease: an ordinal neural network solution. *Gastrointest Endosc* 2021; 93: 187–192.
23. Aoki T, Yamada A and Koike K. The exceptional performance of deep learning for capsule endoscopy: will such quality be maintained in clinical scenarios? *Gastrointest Endosc* 2021; 93: 365–366.
24. Aoki T, Yamada A, Aoyama K, *et al.* 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.
25. Majtner T, Brodersen JB, Herp J, *et al.* A deep learning framework for autonomous detection and classification of Crohn's disease lesions in the small bowel and colon with capsule endoscopy. *Endosc Int Open* 2021; 9: E1361–E1370.
26. Mascarenhas Saraiva MJ, Afonso J, Ribeiro T, *et al.* Deep learning and capsule endoscopy: automatic identification and differentiation of small bowel lesions with distinct haemorrhagic potential using a convolutional neural network. *BMJ Open Gastroenterol* 2021; 8: e000753.
27. Leenhardt R, Vasseur P, Li C, *et al.* A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc* 2019; 89: 189–194.
28. Houdeville C, Souchaud M, Leenhardt R, *et al.* A multisystem-compatible deep learning-based algorithm for detection and characterization of angiectasias in small-bowel capsule endoscopy. A proof-of-concept study. *Dig Liver Dis* 2021; 53: 1627–1631.
29. Saito H, Aoki T, Aoyama K, *et al.* 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.
30. Leenhardt R, Souchaud M, Houist G, *et al.* A neural network-based algorithm for assessing the cleanliness of small bowel during capsule endoscopy. *Endoscopy* 2021; 53: 932–936.
31. Zilbauer M and Heuschkel R. Disease prognostic biomarkers in inflammatory bowel diseases – a reality check. *J Crohns Colitis* 2022; 16: 162–165.
32. Gisbert JP and Chaparro M. Clinical usefulness of proteomics in inflammatory bowel disease: a comprehensive review. *J Crohns Colitis* 2019; 13: 374–384.
33. Verstockt B, Verstockt S, Dehairs J, *et al.* Low TREM1 expression in whole blood predicts anti-TNF response in inflammatory bowel disease. *EBioMedicine* 2019; 40: 733–742.
34. Baert F, Moortgat L, Van Assche G, *et al.* Mucosal healing predicts sustained clinical remission in patients with early-stage Crohn's disease. *Gastroenterology* 2010; 138: 463–468; quiz e10.
35. Allez M, Lemann M, Bonnet J, *et al.* Long term outcome of patients with active Crohn's disease exhibiting extensive and deep ulcerations at colonoscopy. *Am J Gastroenterol* 2002; 97: 947–953.
36. Gralnek IM, Defranchis R, Seidman E, *et al.* Development of a capsule endoscopy scoring index for small bowel mucosal inflammatory change. *Aliment Pharmacol Ther* 2008; 27: 146–154.
37. Bertasius G, Wang H and Torresani L. Is space-time attention all you need for video understanding? *arXiv Preprint arXiv:2102.05095*, 2021.
38. Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. *arXiv Preprint arXiv:1706.03762*, 2017.
39. Ueyama H, Kato Y, Akazawa Y, *et al.* Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. *J Gastroenterol Hepatol* 2021; 36: 482–489.
40. Ghosh T and Chakareski J. Deep transfer learning for automated intestinal bleeding detection in capsule endoscopy imaging. *J Digit Imaging* 2021; 34: 404–417.
41. Kennedy NA, Heap GA, Green HD, *et al.* Predictors of anti-TNF treatment failure in anti-TNF-naïve patients with active luminal Crohn's disease: a prospective, multicentre, cohort study. *Lancet Gastroenterol Hepatol* 2019; 4: 341–353.
42. Atreya R, Neurath MF and Siegmund B. Personalizing treatment in IBD: hype or reality

- in 2020? Can we predict response to anti-TNF? *Front Med (Lausanne)* 2020; 7: 517.
43. Noor NM, Verstockt B, Parkes M, *et al.* Personalised medicine in Crohn's disease. *Lancet Gastroenterol Hepatol* 2020; 5: 80–92.
 44. Torres J, Bonovas S, Doherty G, *et al.* ECCO guidelines on therapeutics in Crohn's disease: medical treatment. *J Crohns Colitis* 2020; 14: 4–22.
 45. Shah SC, Colombel J-F, Sands BE, *et al.* Mucosal healing is associated with improved long-term outcomes of patients with ulcerative colitis: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2016; 14: 1245–1255.e8.
 46. Kopylov U, Yablecovitch D, Lahat A, *et al.* Detection of small bowel mucosal healing and deep remission in patients with known small bowel Crohn's disease using biomarkers, capsule endoscopy, and imaging. *Am J Gastroenterol* 2015; 110: 1316–1323.
 47. Monteiro S, Boal Carvalho P, Dias de Castro F, *et al.* Capsule endoscopy: diagnostic accuracy of lewis score in patients with suspected Crohn's disease. *Inflamm Bowel Dis* 2015; 21: 2241–2246.
 48. Yablecovitch D, Lahat A, Neuman S, *et al.* The lewis score or the capsule endoscopy Crohn's disease activity index: which one is better for the assessment of small bowel inflammation in established Crohn's disease? *Therap Adv Gastroenterol* 2018; 11: 1756283X17747780.
 49. Koulaouzidis A, Douglas S and Plevris JN. Lewis score correlates more closely with fecal calprotectin than Capsule Endoscopy Crohn's Disease Activity Index. *Dig Dis Sci* 2012; 57: 987–993.
 50. Eliakim R, Yablecovitch D, Lahat A, *et al.* A novel PillCam Crohn's capsule score (Eliakim score) for quantification of mucosal inflammation in Crohn's disease. *United European Gastroenterol J* 2020; 8: 544–551.
 51. Gal E, Geller A, Fraser G, *et al.* Assessment and validation of the new capsule endoscopy Crohn's disease activity index (CECDAI). *Dig Dis Sci* 2008; 53: 1933–1937.
 52. Niv Y, Ilani S, Levi Z, *et al.* Validation of the Capsule Endoscopy Crohn's Disease Activity Index (CECDAI or NIV score): a multicenter prospective study. *Endoscopy* 2012; 44: 21–26.

Visit SAGE journals online
journals.sagepub.com/
home/tag

 SAGE journals