Vessel and tissue recognition during third-space endoscopy using a deep learning algorithm

Alanna Ebigbo, Robert Mendel, Markus W Scheppach, Andreas Probst, Neal Shahidi, Friederike Prinz, Carola Fleischmann, Christoph Römmele, Stefan Karl Goeldner, Georg Braun, David Rauber, Tobias Rueckert, Luis A de Souza Jr, Joao Papa, Michael Byrne, Christoph Palm, Helmut Messmann

MESSAGE

In this study, we aimed to develop an artificial intelligence clinical decision support solution to mitigate operator-dependent limitations during complex endoscopic procedures such as endoscopic submucosal dissection and peroral endoscopic myotomy, for example, bleeding and perforation. A DeepLabv3-based model was trained to delineate vessels, tissue structures and instruments on endoscopic still images from such procedures. The mean cross-validated Intersection over Union and Dice Score were 63% and 76%, respectively. Applied to standardised video clips from third-space endoscopic procedures, the algorithm showed a mean vessel detection rate of 85% with a false-positive rate of 0.75/min. These performance statistics suggest a potential clinical benefit for procedure safety, time and also training.

IN MORE DETAIL

Endoscopic submucosal dissection (ESD) is an established organ-sparing curative endoscopic resection technique for premalignant and superficially invasive neoplasms of the GI tract. However, ESD and peroral endoscopic myotomy (POEM) are complex procedures with an elevated risk of operator-dependent adverse events, specifically intra-procedural bleeding and perforation. This is due to inadvertent transection through submucosal vessels or into the muscularis propria, as visualisation and cutting trajectory within the expanding resection defect is not always apparent. An effective mitigating strategy for intra-procedural adverse events has yet to be developed.

Artificial intelligence clinical decision support solution (AI-CDSS) has rapidly proliferated throughout diagnostic endoscopy. We therefore sought to develop a novel AI-CDSS for real-time intra-procedural detection and delineation of vessels, tissue structures and instruments during ESD and POEM.

Sixteen full-length videos of 12 ESD and 4 POEM procedures using Olympus EVIS X1 series endoscopes (Olympus, Tokyo, Japan) were extracted from the Augsburg University Hospital database. A total of 2012 still images from these videos were annotated by minimally invasive tissue resection experts (ESD experience ≥500 procedures) using the computer vision annotation tool for the categories electro-surgical knife, endoscopic instrument, submucosal layer, muscle layer and blood vessel. A DeepLabv3+ neural network architecture with 52 predefined vessels (online supplemental methods) was trained with these data. The performance of the algorithm was measured in an internal fivefold cross validation, as well as a test on 453 annotated images from 11 separate videos using the parameters Intersection over Union (IoU), Dice Score and pixel accuracy (online supplemental methods). The IoU and Dice Score measure the percentual overlap between the algorithm’s delineation and the gold standard. The pixel accuracy measures the percentage of true pixel predictions per image and over all classes. The validation metrics were calculated by accumulating the per-fold outputs. The cross validation was done without hyperparameter tuning.

Three further full-length videos (1× POEM, 1× rectal ESD and 1× oesophageal ESD) were used for an evaluation of the algorithm on video. Thirty-one clips with 52 predefined vessels (online supplemental methods) were evaluated frame by frame with artificial intelligence (AI) overlay for true and
false vessel detection, and a vessel detection rate (VDR) was determined.

The cross-validated mean IoU, mean Dice Score and pixel accuracy were 63%, 76% and 81%, respectively. On the test set, the AI-CDSS achieved scores of 68%, 80% and 87% for the same parameters. The individual per class values and 95% CIs are shown in figure 1. Examples of the original frames, expert annotations and AI segmentations are shown in figure 1. The mean VDR was 85%. The VDR for rectal ESD, oesophageal ESD and POEM were 70%, 95% and 92%, respectively. The mean false-positive rate was 0.75 /min. The algorithm spotted seven out of nine vessels, which caused intraprocedural bleeding. It also recognised the two vessels which required specific haemostatic forceps for major bleeding.

To demonstrate the performance of the AI-CDSS without computing quantitative performance measures, we show an example of an internal POEM procedure with AI overlay. For visualisation of the experiment, we show six video clips, which were used for the evaluation of VDR in the same video (2× POEM, 2× rectal ESD and 2× oesophageal ESD; online supplemental video 1). For a test in robustness, the algorithm was also applied to a randomly selected highly compressed YouTube video of a gastric per-oral endoscopic myotomy procedure (ENDOCLUNORD 2020, https://www.youtube.com/watch?v=VKFHWozYDGM; online supplemental video 2). The individual output is the result of an exponential moving average between the current and past predictions which smooths the predictions and is a simple way to include temporal information.

COMMENTS

This preliminary study aims at investigating the potential role of AI during therapeutic endoscopic procedures such as ESD or POEM. The algorithm delineated tissue structures, vessels and instruments in frames taken from endoscopic videos with a high overlap to the gold standard provided by expert endoscopists. Analogous technology has been demonstrated for application in laparoscopic cholecystectomy to differentiate between safe and dangerous zones of dissection with a mean IoU of 53% and 71%, respectively.

On video clips with standardised and predefined vessels, the algorithm showed a VDR of 85%. The lower performance of 70% in rectal ESD compared with excellent detection of over 90% in oesophageal ESD and POEM might be explainable by poorer visualisation of the structures and more intraprocedural bleeding, which is in agreement with clinical experience.

Numerous preclinical and clinical studies on AI in GI endoscopy have been published, but until now, the application of AI has been limited largely to diagnostic procedures such as the detection of polyps or the characterisation of unclear lesions. In abdominal surgery, AI has been applied with promising results for various tasks, including the detection of surgical instruments, image guidance, navigation and skill assessment (‘smart surgery’). The results of this study suggest that AI may have the potential to optimise complex endoscopic procedures such as ESD or POEM in analogy to the mentioned research (‘smart ESD’). By highlighting submucosal vessels and other tissue structures, such as the submucosal cutting plane, therapeutic procedures could become faster and burdened with fewer adverse events such as intraprocedural or postprocedural bleeding and

Table 1 Performance results of the AI-CDSS in the internal cross validation and the test data set: IoU and Dice Score for all categories as well as their means across all categories, pixel accuracy for complete frames and 95% CI in brackets

<table>
<thead>
<tr>
<th>Internal cross validation</th>
<th>Vessel detection</th>
<th>Tissue differentiation</th>
<th>Instrument detection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vessel</td>
<td>Submucosa</td>
<td>Muscularis</td>
</tr>
<tr>
<td>Vessel Dice Score</td>
<td>55.15</td>
<td>75.51</td>
<td>70.64</td>
</tr>
<tr>
<td></td>
<td>(54.10 to 56.18)</td>
<td>(74.88 to 76.12)</td>
<td>(69.32 to 71.88)</td>
</tr>
<tr>
<td>IoU</td>
<td>38.07</td>
<td>60.65</td>
<td>54.60</td>
</tr>
<tr>
<td></td>
<td>(37.08 to 39.07)</td>
<td>(59.85 to 61.44)</td>
<td>(53.05 to 56.10)</td>
</tr>
<tr>
<td>Pixel accuracy</td>
<td>86.89</td>
<td>80.71</td>
<td>74.88</td>
</tr>
<tr>
<td></td>
<td>(85.86 to 87.70)</td>
<td>(79.50 to 81.82)</td>
<td>(69.40 to 74.99)</td>
</tr>
</tbody>
</table>

AI-CDSS, artificial intelligence clinical decision support solution.
perforation. In the future, AI assistance may have the potential to accelerate the learning curve of trainees in endoscopy.

The major limitation of this study is the small number of videos used for training and validation; however, every video contained a complete therapeutic ESD procedure with a full range of procedural situations. The study is further limited by the fact that the algorithm was not yet tested in a real-life setting. However, the AI model was tested on externally generated video sequences and was able to recognise submucosal vessels and the cutting plane. Furthermore, surrogate parameters such as the detection of vessels, which bled later during the procedures, give rise to the conclusion that these complications might have been preventable by the application of the AI-CDSS. This is a first preclinical report on a novel technology; further research is needed to evaluate a potential clinical benefit of this AI-CDSS in detail.

Author affiliations
1 Department of Gastroenterology, Universitätsklinikum Augsburg, Augsburg, Germany
2 Regensburg Medical Image Computing (ReMIC), Ostbayerische Technische Hochschule Regensburg, Regensburg, Germany
3 Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada
4 Department of Gastroenterology, Oostal-Kliniikum Aalen, Aalen, Germany
5 Department of Computing, Federal University of São Carlos, São Carlos, Brazil
6 Department of Computing, São Paulo State University, Botucatu, Brazil
7 Vancouver General Hospital, The University of British Columbia, Vancouver, British Columbia, Canada

Twitter: Joao Papa @papa_joaopaulo and Christoph Palm @ReMIC_OTH

Contributors AE and MWS: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript and critical revision of the manuscript. RM: study concept and design, software implementation, analysis and interpretation of data, drafting of the manuscript and critical revision of the manuscript. AP: study concept and design, acquisition of data and critical revision of the manuscript. RS: analysis and interpretation of data, drafting of the manuscript and critical revision of the manuscript. NC: analysis and interpretation of data, critical revision of the manuscript. FJ: analysis and interpretation of data, critical revision of the manuscript. DR: software implementation and experimental evaluation and critical revision of the manuscript. LS: statistical analysis and critical revision of the manuscript. LS: statistical analysis and critical revision of the manuscript. JP: study concept and design, analysis and interpretation of data, drafting of the manuscript and critical revision of the manuscript. LM: study concept and design, analysis and interpretation of data, drafting of the manuscript and critical revision of the manuscript. LM: statistical analysis, critical revision of the manuscript, administrative and technical support and study supervision. HM: study concept and design, acquisition of data, critical revision of the manuscript, administrative and technical support, and study supervision. AE: guarantor. All authors: had access to the study data and reviewed and approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.


Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable. Ethics approval was obtained from the ethics committee of Ludwigs-Maximilians-Universität, Munich (project number 21–1216).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s).

References
7 Shahidi N, Bourke MJ. How to manage the large Nonpedunculated colorectal polyp. Gastroenterology 2021;160:2239–43.