Artificial intelligence in gastrointestinal endoscopy

Rodrigo Jover
Hospital General Universitario de Alicante
Alicante, SPAIN
Vienna, 19-10-2018
Deep Learning
The subset of machine learning composed of algorithms that permit software to train itself to perform tasks, like speech and image recognition, by exposing multilayered neural networks to vast amounts of data.

Machine Learning
A subset of AI that includes abstruse statistical techniques that enable machines to improve at tasks with experience. The category includes deep learning.

Artificial Intelligence
Any technique that enables computers to mimic human intelligence, using logic, if-then rules, decision trees, and machine learning (including deep learning).
Early efforts: AI with subhuman performance is occasionally used in commercial expert systems with varying degrees of utility.

Current state: Narrow task-specific AI has started to match and, in some instances, exceed human performance in tasks including conversational speech recognition, driving vehicles, playing Go and classifying skin cancer.

Future outlook: General AI exceeds human performance and reasoning in complex tasks, including writing best-selling novels and performing surgery. Human intelligence improves as we learn from AI.
1st layer: combination of edges form motifs
2nd layer: motifs assemble into parts
3rd layer: parts form objectes

Mathematical & Statistical rules

Layers are not designed by human engineers, they are learned from large data-sets

Rodrigo Jover, Alicante, SPAIN
# AI evolution

<table>
<thead>
<tr>
<th>Topic</th>
<th>DDW 2017</th>
<th>DDW 2018</th>
<th>UEGW 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colon detection</td>
<td>2</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Colon diagnosis</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>CRC T1</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Complications polypectomy</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Follow-up UC</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Barrett’s esophagus</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Esophageal cancer</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Gastric cancer</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Small bowel</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Pancreas cancer</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>3</strong></td>
<td><strong>16</strong></td>
<td><strong>26</strong></td>
</tr>
</tbody>
</table>
QUALITY. Cecal intubation

Chiu, DDW 2018

Rodrigo Jover, Alicante, SPAIN
QUALITY. Bowel cleansing

Confusion Matrix - Bowel Prepscore Classifier (MobileNet <1ms)

Examples of “False Negatives” (Scored 0-1, Predicted 2-3)

Examples of “False Positives” (Scored 2-3, Predicted 0-1)

<table>
<thead>
<tr>
<th></th>
<th>Proportion Estimate</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Upper</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.9478</td>
<td>0.9396</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9560</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.9786</td>
<td>0.9734</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9839</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>0.9769</td>
<td>0.9713</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9826</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.9515</td>
<td>0.9439</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9592</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.9636</td>
<td>0.9587</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9684</td>
</tr>
</tbody>
</table>

Karnes, DDW 2018

Rodrigo Jover, Alicante, SPAIN
POLYP DETECTION

Jover, Endoscopy 2016
Detection

Urban, Gastroenterology 2018

Real time polyp localization with Deep Learning with 96% accuracy

*Trained on 8000 images (4000 unique polyps)

>20% additional polyps found

Gastroenterology

Rodrigo Jover, Alicante, SPAIN
1. Training: 8641 colonoscopy images from 2000 patients
2. 2nd training: 1330 colonoscopy images
3. Feasibility study: 9 colonoscopy videos
4. 44947 image frames from the 9 videos
5. Validation: 11 challenging colonoscopy videos

Urban, Gastroenterology 2018
Polyps found with AI assistance

Urban, Gastroenterology 2018
Optical diagnosis

Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study

Rees, Gut 2017
Optical diagnosis in clinical practice

More problems

- Need of training
- Time-consuming
- Lack of motivation
- Legal problems
- What for me?
- Lack of confidence

Better call Pathologist
Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model

Gut, 2017

Michael F Byrne, Nicolas Chapados, Florian Soudan, Clemens Oertel, Milagros Linares Pérez, Raymond Kelly, Nadeem Iqbal, Florent Chandelier, Douglas K Rex

Rodrigo Jover, Alicante, SPAIN
Low-confidence in 19/125 polyps

Byrne, Gut 2017
Characterization of Colorectal Lesions Using a Computer-Aided Diagnostic System for Narrow-Band Imaging Endocytoscopy

Misawa, Gastroenterology 2018

Computer-Aided diagnosis (CAD) system for endocytoscopy with NBI

1079 images (431 non-neoplastic and 648 neoplastic) from 85 polyps

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>84.5 (72.6–92.7)</td>
<td>97.6 (87.4–99.9)</td>
<td>90.0 (82.4–95.1)</td>
<td>98.0 (89.4–99.9)</td>
<td>82.0 (68.6–91.4)</td>
</tr>
<tr>
<td>High confidence</td>
<td>97.6 (87.1–99.9)</td>
<td>95.8 (78.9–99.9)</td>
<td>96.9 (89.3–99.6)</td>
<td>97.6 (87.1–99.9)</td>
<td>95.8 (78.9–99.8)</td>
</tr>
</tbody>
</table>
Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy
A Prospective Study

Mori, Ann Intern Med 2018

- 466 ≤ 5 mm polyps
- 325 patients

No training, no confidence problems, highly reproducible

ASGE recommended threshold

OVERALL  RECTOSIGMOID
The future?
Strenghts

• Able to solve problems related to quality, improve detection and characterization of colonic lesions
• Space for improvement: the more you train, the better performance you get
• Improvement of computer systems
• But...
Limitations

- Need of multicenter and real life validation
- Need of high quality procedures: appropriate withdrawal time and withdrawal technique
- Possible effect on inspection behavior
- Deep learning performance may vary by indication
- Need of cost-effectiveness validation: colonoscopy time, pathology costs, irrelevant findings,…
- Legal and regulatory doubts, reimbursement
What gastroenterologists need to know?

• Be familiar with AI terminology. Standardize definition for AI terms
• What should we require for AI studies?
• How should we use AI?
• Which improvement does worth?
• Study design needed to evaluate AI usefulness

Rodrigo Jover, Alicante, SPAIN
ESGE research position statement on AI

• AIMS
  – Define AI terminology for endoscopic community
  – Establish standards: what this technology must be able to provide
  – Recommend adequate study designs for evaluation of AI in clinical practice
  – Recommend adequate performance measure levels: sensitivity, specificity, PPV, NPV