11-08-2023

From Bit to Bedside

Artificial Intelligence and its Potential in Surgery: These days a lot of scientific breakthroughs are enabled by artificial intelligence (AI).

Dr. med. Joël L. Lavanchy

As most surgeons are not familiar with computer science, this review article will explain basic concepts of AI and deep learning, how they are applied to surgery and what potential can be leveraged through AI in surgery. The term AI was first used in 1955 in a grant proposal by John Mc-Carthy a mathematician at Dartmouth College in Hannover, New Hampshire. McCarthy proposed "to make machines use language, [...] solve kinds of problems now reserved for humans, and improve themselves".[1] Even though the term AI was introduced 70 years ago, the idea of intelligent machines is over 2700 years old. It dates to Greek mythology where the saga of Talos describes a giant automaton (machine) that protects Europe from pirates and invaders. Thus, AI is more an idea or a concept rather than a precise method. It is often used as umbrella term to describe computational methods that are not explicitly programmed.

Machine learning in contrast is a computational method, that trains an algorithm to process data in an iterative way. This may be fully supervised, semi-supervised or un-supervised. To illustrate the different modes of supervision the task of image classification will be used as an example. Image classification is a classical computer vision task, that aims to classify the content of an image into predefined classes. Computer vision is a field within computer science that extracts information from images and videos using computational methods. As computer vision is older than recent developments in AI this traditionally did not include AI methods. Though, today most computer vision models use AI.

Full supervision means that the algorithm is fed with labelled training samples. Given we have a set of images of cats and dogs, a human annotator labels every image as belonging to the class of cat or dog (ground truth labels). In a semi-supervised setting, the algorithm has been trained on a subset of human labeled images. This prior knowledge will be used to automatically label the whole dataset with so-called pseudo-labels or weak labels, which will serve for model training as in the fully supervised method. An unsupervised machine learning algorithm does not learn from labelled training data but autonomously explores the differences between classes, in this example the visual appearance of cats and dogs.

How does an AI model learn?

Staying with the image classification example, the model is trained with labelled cat and dog images of the training dataset (Figure1). Another part of the labeled data is set apart for validation and testing of the model. The

building block of most computer vision models are neural networks (NN). They process the training images applying consecutive mathematical operations to predict an output. To understand this process, an image needs to be considered as an array of numbers. A color image of 125x125 pixels containing a red, a green and a blue channel (RGB image) is represented by an array of 3x125x125 numbers. Every mathematical operation performed on the input array of numbers is considered a layer of the NN. As the different layers learn from and update each other like synapses in the brain those networks are referred to as NN. In the beginning of NN, they often had only few input, intermediate (so called hidden) and output layers. But as methods evolved, NN became increasingly deep with hundreds and thousands of layers. Therefore, those models are also referred to as deep learning (DL). Once the NN has processed the whole training dataset (one training epoch), the predictions of the model (whether a given image is a cat or a dog) are compared to the ground truth labels. A loss function defines how the deviation of predicted labels to the ground truth labels updates the model weights i.e. the mathematical operations in every layer. Thus, the model learns to refine its predictions with every training epoch. To avoid overfitting, which means to ensure that the model can be generalized to the prior unseen images of dogs and cats in the test dataset, the model performance is constantly measured during training on the validation dataset.
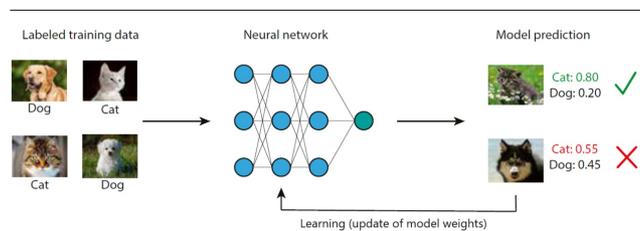

Figure 1: Illustration of a machine learning algorithm for image classification.
A neural network is trained with labeled training samples of cat and dog images. The trained model predicts the likelihood of images not used during training (test set) to belong to the dog or the cat class. In the top row the model prediction is confident and correct. In the bottom row model prediction is close to the classifier threshold of 0.5 i.e. the model is unconfident and a dog image is misclassified as a cat. Given that the test set is labeled as the training set, the model identifies the misclassification and updates the neural network to improve its predictions. This is how the model learns to classify dog and cat images.

Why do we need AI in surgery?

Surgical education undergoes a paradigm shift from quantitybased to competency-based assessment of surgical skills.[2] As direct observation of trainees for surgical skill assessment is timeconsuming and often lacks objectivity, video-based assessment has been introduced.[3] This allows for a retrospective review by multiple assessors, which increases availability and objectivity of video-based assessment. Surgical skills as assessed by peers in surgical videos have been shown to correlate with clinical outcomes. [4,5] Therefore, video-based assessment has become increasingly popular in education but also in quality improvement programs and research.

Even though video-based assessment has proven benefits it is not scalable to daily surgical practice, as a lot of manual work is involved. This is where automated video analysis using DL comes into play.[6]

How does AI learn surgery?

The most prevalent data source to teach a computer algorithm the understanding of surgery is endoscopic videos. During an endoscopic

procedure the surgeon performs the surgery almost exclusively by relying on what he sees on the screen. In retrospective video-based assessment no other source of data but the surgical video is available. Computer vision models have been trained to recognize phases and steps of interventions in endoscopic videos. [7,8] Another source of information are the tools being used during surgery as they inform about the surgical actions being performed. Recent AI models have been trained to recognize tools in endoscopic videos in a semi-supervised[9] or unsupervised way.[10] Once the tools are recognized by an AI model, they can be tracked over time to measure the movement trajectories of the instruments (Figure 2). To date, the most fine-grained representations of surgical activities recognized by AI algorithms are action triplets, defined as ‹object, verb, target› as in ‹grasper, grasps, gallbladder›.[11]
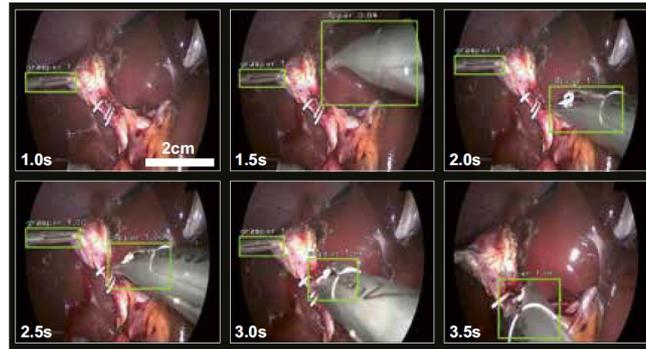

Figure 2: Tool tracking model of laparoscopic instruments.
A neural network was trained to recognize and track surgical instruments within a video of a laparoscopic cholecystectomy. The instruments are marked with bounding boxes (light green) and a class label (clipper or grasper) and a confidence score (between 0 and 1.0) are attributed. The figure shows consecutive images every 0.5 seconds from a video sequence (complete video https://doi.org/10.1038/s41598-021-84295-6).

What is the potential of AI in surgery?

Recognition of surgical phases by an AI model has been used to predict the remaining duration of the surgical procedure.[12] This facilitates the scheduling of subsequent operations and helps to organize the scarce and costly operation room resources. Based on tool tracking a three staged machine learning algorithm has been trained to assess surgical skills.[13] This study showed that good surgical skills are correlated with narrow and precise movement patterns and that an AI algorithm can be trained to assess surgical skills. As there is a large body of evidence showing that surgical skills are associated with postoperative outcomes,[4,5] further developments in this direction might help to predict postoperative outcomes from intraoperative analysis of endoscopic videos.

Other works used AI models to assess safety relevant tasks during laparoscopic cholecystectomy. The DL model ClipAssistNet was trained to assess whether a clip application is safe.[14] Once the clip fully encircles the cystic duct or artery (i.e. both tips of the clip are visible), a green light occurs. ClipAssistNet reassures surgeons that they do not unintentionally damage neighboring structures. Two other DL models increasing safety during laparoscopic cholecystectomy are using semantic segmentation. Semantic segmentation assigns anatomical labels to every pixel of an image allowing the DL model to learn anatomical structures. DeepCVS was trained to automatically assess the three critical view of safety (CVS) criteria.[15] Another model called GoNoGoNet was trained to predict safe (Go) and unsafe (No-Go) dissection zones during hepatocystic triangle dissection.[16] All three DL models (ClipAsisst-Net, DeepCVS, and GoNoGoNet) are

measures to guide surgeons intraoperatively, to enhance their decision making and to increase safety during laparoscopic cholecystectomy. Especially surgeons in training might benefit from those AI models.

Another area of great potential for AI in surgery is to reduce the burden of documentation. Based on phase and tool recognition the DL model EndoDigest automatically identifies and extracts a 2.5 minute video clip of the cystic duct division out of full-length videos of laparoscopic cholecystectomies.[17] Within 91% of the extracted video clips the three CVS criteria were assessable. This allows to use EndoDigest for retrospective video-based assessment or for the documentation of CVS achievement. A recent study explored the potential of surgical phase recognition to create automated operating room (OR) reports.[18] The more reliable AI models recognize surgical phases, tools and actions the better they become in understanding and documenting what happens in the OR.

While a lot of AI models in surgery have been trained on monocentric datasets of laparoscopic cholecystectomy, as this is a highly frequent and relatively standardized procedure, recent works translate those models to more complex procedures and multicentric datasets to prove generalizability.[8,19]

In summary, AI has been applied to automatically analyze workflows, assess surgical skills and safety checkpoint achievement. Thus, AI applications have tremendous potential to improve intraoperative guidance and enhance decision-making during surgery. Most AI models in surgery are still in the development stage and not in routine clinical use yet. Though, translation from "bit to bedside" is not far and will follow soon.

AUTOREN

Dr. med. Joël L. Lavanchy
joel.lavanchy@ihu-strasbourg.eu

REFERENZEN

References:

1. McCarthy J, Minsky ML, Rochester N, Shannon CE. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. AI Magazine. 2006;27:12

2. Sonnadara RR, Mui C, McQueen S, Mironova P, Nousiainen M, Safir O, Kraemer W, Ferguson P, Alman B, Reznick R. Reflections on Competency-Based Education and Training for Surgical Residents. Journal of Surgical Education. 2014;71:151–158

3. Pugh CM, Hashimoto DA, Korndorffer JR. The what? How? And Who? Of video based assessment. The American Journal of Surgery. 2021;221:13–18

4. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR, Dimick J, Banerjee M, Birkmeyer NJO. Surgical Skill and Complication Rates after Bariatric Surgery. New England Journal of Medicine. 2013;369:1434–1442

5. Fecso AB, Bhatti JA, Stotland PK, Quereshy FA, Grantcharov TP. Technical Performance as a Predictor of Clinical Outcomes in Laparoscopic Gastric Cancer Surgery. Annals of Surgery. 2019;270:115–120

6. Pedrett R, Mascagni P, Beldi G, Padoy N, Lavanchy JL. Technical Skill Assessment in Minimally Invasive Surgery Using Artificial Intelligence: A Systematic Review . Preprint medRxiv November 8, 2022. DOI: 10.1101/2022.11.08.22282058

7. Twinanda AP, Shehata S, Mutter D, Marescaux J, De Mathelin M, Padoy N. EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos. IEEE Transactions on Medical Imaging. 2017;36:86–97

8. Ramesh S, Dall'Alba D, Gonzalez C, Yu T, Mascagni P, Mutter D, Marescaux J, Fiorini P, Padoy N. Multitask temporal convolutional networks for joint recognition of surgical phases and steps in gastric bypass procedures. International Journal of Computer Assisted Radiology and Surgery. 2021;16:1111–1119

9. Vardazaryan A, Mutter D, Marescaux J, Padoy N. Weakly-Supervised Learning for Tool Localization in Laparoscopic Videos. In: MICCAI LABELS 2018. Lecture Notes in Computer Science. Cham: Springer. July 18, 2018. DOI: 10.1007/978-3-030-01364-6_19

10. Sestini L, Rosa B, De Momi E, Ferrigno G, Padoy N. FUN-SIS: A Fully UNsupervised approach for Surgical Instrument Segmentation. Medical Image Analysis. 2023;85:102751.

11. Nwoye CI, Gonzalez C, Yu T, Mascagni P, Mutter D, Marescaux J, Padoy N. Recognition of Instrument-Tissue Interactions in Endoscopic Videos via Action Triplets. In: MICCAI 2020. Lecture Notes in Computer Science. Cham: Springer. July 10, 2020. DOI: 10.1007/978-3-030-59716-0_35.

12. Aksamentov I, Twinanda AP, Mutter D, Marescaux J, Padoy N. Deep Neural Networks Predict Remaining Surgery Duration from Cholecystectomy Videos. In: MICCAI 2017. Lecture Notes in Computer Science. Cham: Springer International Publishing:586–593.

13. Lavanchy JL, Zindel J, Kirtac K, Twick I, Hosgor E, Candinas D, Beldi G. Automation of surgical skill assessment using a three-stage machine learning algorithm. Scientific Reports. 2021;11:5197

14. Aspart F, Bolmgren JL, Lavanchy JL, Beldi G, Woods MS, Padoy N, Hosgor E. ClipAssistNet: bringing real-time safety feedback to operating rooms. International Journal of Computer Assisted Radiology and Surgery. 2022;17:5–13

15. Mascagni P, Vardazaryan A, Alapatt D, Urade T, Emre T, Fiorillo C, Pessaux P, Mutter D, Marescaux J, Costamagna G, Dallemagne B, Padoy N. Artificial Intelligence for Surgical Safety. Annals of Surgery. 2022;275:955–961

16. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, Navarrete-Welton A, Sankaranarayanan G, Brunt LM, Okrainec A, Alseidi A. Artificial Intelligence for Intraoperative Guidance: Using Semantic Segmentation to Identify Surgical Anatomy During Laparoscopic Cholecystectomy. Annals of Surgery. 2022;276:363–369

17. Mascagni P, Alapatt D, Urade T, Vardazaryan A, Mutter D, Marescaux J, Costamagna G, Dallemagne B, Padoy N. A Computer Vision Platform to Automatically Locate Critical Events in Surgical Videos: Documenting Safety in Laparoscopic Cholecystectomy. Annals of surgery. 2021;274:e93–e95

18. Berlet M. Surgical reporting for laparoscopic cholecystectomy based on phase annotation by a convolutional neural network (CNN) and the phenomenon of phase flickering: a proof of concept. International Journal of Computer Assisted Radiology and Surgery. 2022;17:1991–1999

19. Lavanchy JL, Gonzalez C, Kassem H, Nett PC, Mutter D, Padoy N. Proposal and multicentric validation of a laparoscopic Roux-en-Y gastric bypass surgery ontology. Surg Endosc. . Epub ahead of print October 26, 2022. DOI: 10.1007/s00464-022-09745-2