Using AI to assess surgical performance

March 12 2021

Credit: Pixabay/CC0 Public Domain

A research team at Inselspital, Bern University Hospital and Caresyntax has succeeded in proving that artificial intelligence can reliably assess surgeons' skills. A method involving a three-stage procedure has been presented that correctly designates good and mediocre performance with

a high accuracy rate. This paves the way for further steps towards AI-supported expert systems.

More than one million operations are performed in Switzerland every year. A surgeon's skill has a direct impact on the outcome of the operation. Training and experience, as well as momentary fatigue and other influencing factors all play a role. At present, skill is tested by experts, either directly during an operation or by evaluating video footage. This approach is very costly and only a limited number of experts are available. Moreover, the assessment may vary and is not always fully reproducible. For some time, attempts have been made to automate and objectify the assessment of surgeons' skills.

Proof of feasibility

The key result of the study is the proof of the fundamental feasibility of an artificial intelligence (AI)-based assessment of a surgeon's skill in the context of a surgical procedure. The AI used in the study identified good or moderate surgical skill with 87 percent accuracy. This can be considered a very good finding. Lead author Joël Lavanchy explains: "What was surprising was the high degree of algorithms' accuracy with the selected method. Our method of assessing surgical skills is based on the analysis of instrument movement. Surgical instruments were identified using computer algorithms and their movement was analyzed during the time period."

Innovative, three-stage approach with AI

The research team used a newly developed, three-stage approach. The study was based on 242 videos of laparoscopic gallbladder removal procedures. The first step was to identify the instruments used. For this purpose, a convolutional neural network (CNN) was trained to recognize

the instruments. In the second step, the movements were analyzed, and their patterns were extracted. In the third step, the extracted movement patterns correlated with rating results by experts using linear regression.

Broader database and in-depth training of algorithms is needed

The present study is an important first step towards assessing surgical performance. More in-depth steps are needed before the technology can be used in clinical practice. For one thing, the AI algorithms need to be trained on a broader database to further improve instrument recognition. For another thing, additional surgeries need to be investigated and, in the medium term, videos of open surgeries as well as procedures apart from the abdominal area can be addressed.

Dr. Enes Hosgor, a co-author of the study who leads the AI division at caresyntax, a medical technology company headquartered in Berlin and Boston, classifies the results as follows: "AI has mainly been used thus far to identify instruments or specific surgical phases. In our study, we now assess surgical skill based on surgical videos. In the future, the use of AI can solve problems at multiple levels: it is available on-demand peri-operatively (not dependent on a few hard-to-find experts); it is objective using algorithm-driven standards; it is comparable at a transregional level as well as surgeon level and could thus provide important support for decision-making processes at certification institutes."

AI at medical location in Bern: CAIM as an opportunity

The project provides an important indication of the future development of the use of AI in medicine. In the future, it will shift from the

erstwhile evaluation of image material to the provision of expert systems. Prof. Guido Beldi, head of the study, clarifies: "The study is a first step. Now that we have demonstrated the fundamental feasibility, we can start planning assistance systems that will support surgeons during operations. For example, they will be alerted when fatigue is detected, thereby helping to prevent complications."

More information: Joël L. Lavanchy et al, Automation of surgical skill assessment using a three-stage machine learning algorithm, *Scientific Reports* (2021). DOI: 10.1038/s41598-021-84295-6

Provided by Inselspital, Bern University Hospital