2 February 2023

جامعـــة الأمـــيـــر سطام بن عبدالعزيز
PRINCE SATTAM BIN ABDULAZIZ UNIVERSITY

Prince Sattam bin Abdulaziz University

College of Computer Engineering and Sciences

Information Systems department

IS2121 | Foundations of Information Systems

Surgical Data Science

Shrouq Mohammed
442052015

Sama Mohammed
442960460

Supervisor: Dr. Fatima Al-Qahtani

Abstract

Surgical data science aims to improve. The quality of interventional healthcare and its value through the capture, organization, analysis, and modelling of data. Structures with a diagnostic, prognostic, or therapeutic goal, such as surgery, Data may pertain to any part of the patientcare may concern. the patient, caregivers and/or technology used to deliver care Validation, evaluation, and application of SDS in technical infrastructure, facilitate sharing of data across organizations and the creation of extensively annotated, Validation centered on understanding surgical procedures in explaining and sharing data, refers to systems or devices that simulate human intelligence ،machine learning algorithms to analyze and aggregate unlabeled data sets. These algorithms self-discover hidden patterns in the data without human intervention. Learning is an ideal solution as it relies on learning representations from unclassified data in Self-supervised learning in surgical data science.

Table of Content

List of figures

Introduction

Since health topics are among the most common topics currently, due to their importance in achieving the goal of technology in serving people and harnessing them for their convenience, and when we searched in data science and health, we found a topic out of sight, which is surgical data science, until we thought to expand our knowledge in this field, which turns out to be somewhat modern. But it is of great value. [1]

And with the connection of building this solution with data science and artificial intelligence, working with them behind the scenes, machine learning, which we did not know the role and importance of, which is considered as the engine of these two, which will have a role here in development and learning to create an ideal model that responds to the continuous variables of things that we discovered that there are two types From training to the model, which is supervised training and unsupervised training. Therefore, we will present in detail here what we have achieved from this research."()" [2]

Surgical data science

It aims to improve the quality of interventional healthcare by capturing, organizing, analyzing, and modeling data while an increasing number of data-driven approaches and clinical applications have been studied in the fields of radiological and clinical data science. [1]

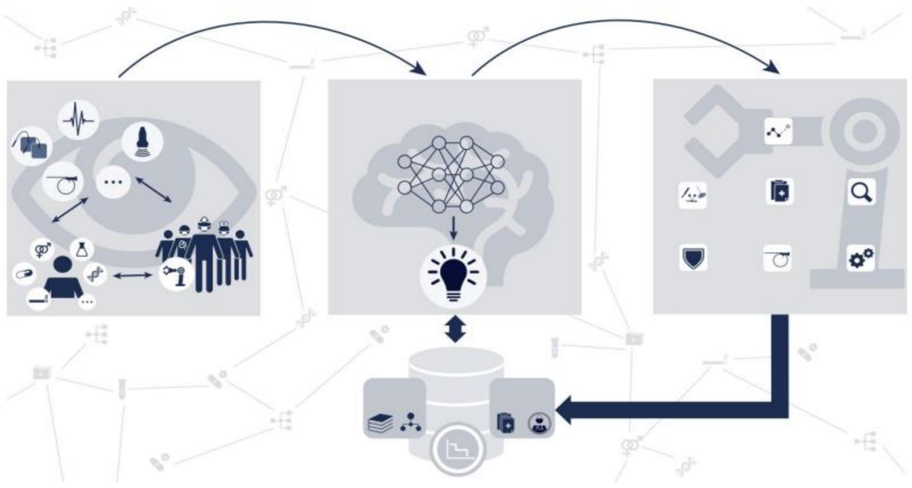


Figure 1-1 Building blocks of a surgical data science (SDS) system.

Relevant data is perceived by the system, effectors include humans and devices that manipulate the patient including surgeons, operating room (OR) team, anesthesia team, nurses, and robots. Sensors are devices for perceiving patient- and procedure-related data such as images, vital signals, and motion data from effectors. Data about the patient includes preoperative images and laboratory data. Applications of SDS are manifold, ranging from surgical education to various clinical tasks, such as early detection, diagnosis, and therapy assistance. [1]

Some of the missions for developing surgical data science

THE FIRST TASK: TECHNICAL INFRASTRUCTURE

Validation, evaluation, and application of SDS algorithms can be accessed and shared among researchers, healthcare professionals, and other stakeholders, either directly or retrospectively.
like:
1- Develop standards for data storage with respect to key aspects including data structure, format, and longevity
2- Develop new intraoperative imaging methods to obtain relevant information on tissue function, morphology, and pathology [1]

THE SECOND TASK: EXPLAINING AND SHARING DATA

facilitate sharing of data across organizations and the creation of extensively annotated, representative, and quality-checked databases.

like:

1- Develop best practices for evaluating and ensuring the quality of annotations

2- Establishing the Unified Ontology of Surgical Data Science

Data analytics:

Aligning SDS research methods with clinical goals and priorities

Develop validation methods and concepts focused on robustness, including generalizability.

Validation centered on understanding surgical procedures, including variance resulting from technical factors, staff, patient, and specific environment.

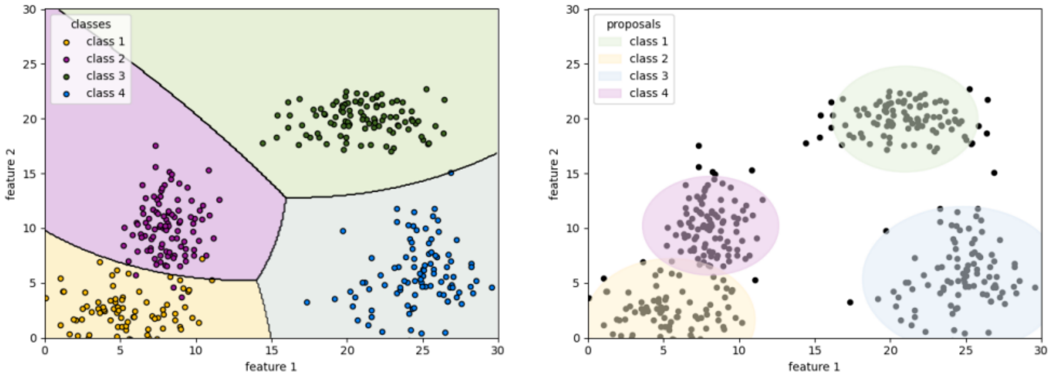
So, to achieve all these things technically, we need artificial intelligence techniques, machine learning, and most importantly data science. [1]

Machine learning

Machine learning is a subset of artificial intelligence (AI) that focuses on creating systems that learn - or improve performance - based on the data they consume. Artificial intelligence is a broad term that refers to systems or devices that simulate human intelligence. Machine learning and artificial intelligence are often discussed together, and the terms are sometimes used interchangeably, but they don't mean the same thing. It is important here to mention that although all machine learning techniques are AI, not all AI is machine learning.

There are two types of machine learning: supervised machine learning and self-learning (unsupervised machine learning). [3]

The first type These datasets are designed to train or "supervise" algorithms to accurately classify data or predict results. With tagged inputs and outputs, the model can measure its accuracy and learn over time. I cited here a personal experience in one of the trainings provided to us by the university in cooperation with SCAI to train an artificial intelligence model to assess the extent to which restaurants and others in the Kingdom adhere to precautionary and hygiene measures, etc., by working on classifying data, which were images according to specific criteria, to train the model.
Unsupervised learning uses machine learning algorithms to analyze and aggregate unlabeled data sets. These algorithms self-discover hidden patterns in the data without human intervention.[4]



(a) Supervised learning

(b) Unsupervised learning

Figure 2-1 Explains the difference between the two types, as in a, you are the one who defines the classifications for the model, and he puts them in the correct classification. In b, the model is the one that describes the structure according to the data given to it.[5]

A class of machine learning algorithms called Fundamentals of Neural Networks is inspired by the work of the brain by simulating complex interactions between neurons and their connections.
It was inspired by the power of the human brain that outperforms many information processing systems, as it can perform highly complex, non-linear, and parallel processing by

Organizing its structural components (neurons) to perform tasks such as accurate predictions, pattern recognition, perception, motor control, etc. In addition to his speed.[5]

SELF- SUPERVISED LEARNING IN LEARNING IN SURGICAL DATA SCIENCE

Always! The biggest problem we face here is the lack of classified data and here the data is not just a little! It's even rarer, and it's a problem that affects both communities. The computer vision community and the surgical data science community. In the medical field, it's even considered one of the major limitations of developing and translating deep learning-based models into clinical routines. When dealing with small data sets, most models are semi-supervised learning approaches.[6] Supervised aims to learn domain representation while cutting out manually generated annotations. dataset and subsequent refinement of a small set of training data from the target task. An alternative approach that does not require a labeled dataset is self-supervised learning. Rather than having a massive number of labels, the raw data is used in the additional task as its source of oversight.[7]

Given the high cost of providing classified data, especially in the clinical field, self-supervised learning is an ideal solution as it relies on learning representations from unclassified data. However, it is considered an unknown path until now because the process is complex, but there are solutions that the world is currently seeking to develop to serve the surgical-clinical field.[8]

AVAILABLE DATASETS FOR SURGICAL DATASETS FOR SURGICAL DATA SCIENCE:

We can say that the data available here is very little, and even with its availability, the confidentiality and privacy of the data may stand in your way, the concerns related to privacy and confidentiality of both patients and caregivers pose important legal and ethical issues that must be addressed for data science to be possible.[2]

So, the validity of accessing and using it is difficult, so training a good model with this small number of data will not be useful and will not give us the required results, so we need another method commensurate with the size of the data available and limited.[9]

Results

Based on these matters, the most important results they reached and start working on:

1- Align with the Objection: Incorporating Unlabeled Data into Training Deep Learning Models by providing an approach for how unlabeled data can be used to improve the training of deep learning models when only a small amount of training data is available. The basic idea is to train a network that can recolor endoscopic images. The assumption is that through this task, the network learns important context information about the images, which is useful for medical instrument segmentation.[10]

2- Aligned with the objective: to establish a quality control dataset for a multi-case surgical instrument segmentation task. One lesson that can be learned from the previous section is that the number of data categorized can lead to performance boosts. Thus, this section is about creating a representative multi-instance and quality-control segmentation dataset for surgical instrument segmentation. [10]

3- Aligned with the objective: Systematic evaluation of the latest performance of multi-instance surgical instrument segmentation by offering community-assisted state-of-the-art identification of binary and multi-instance segmentation.

The goal of this section was to create a fair benchmark between the different methods and see if it increases performance with more data. [10]

4- Align with the objective: Systematic problem-solving analysis of the latest approaches to segmentation of the multi-instance surgical instrument segmentation tool by providing an approach to how image properties (e.g., presence of blood, smoke), etc. influence the algorithm to perform the latest algorithms from the previous section. For this purpose, firstly, the properties of the image are explained, and secondly, their effect on the performance of the algorithm is analyzed statistically. The purpose was to identify characteristics that are detrimental to performance. [10]

5- Aligned with the objective: a problem-driven multi-instance surgical instrument segmentation algorithm, where the results from the third outcome are used to design an explicit algorithm and address some of the characteristics identified in the fourth, which degrade the performance of hardware segmentation algorithms. [10]

The world of machine learning is interesting and deep, and its importance for building this model is indescribable because it will be one of the reasons for making this model more useful with the different and continuous operating room variables that speak quickly, which requires building a fast learning and development model to keep pace with it.

Conclusion

In this simple research, we discussed surgical data science, to discuss what it is and the main challenges it faces, which is the problem of limited data availability and how to deal with data scarcity.
And the use of the second type of machine learning methods, which is self-learning (unsupervised machine learning).
As it is expected that the problem of data availability will not be easy to solve very soon
The future and methods that deal with less data are still subject to research, this dataset has the potential to enable research to train and scale algorithms on a larger amount of data.
This is the appropriate way to develop algorithms as targeted as possible.
In conclusion, one can confidently assume that the combination of data generation, problem-driven algorithm development and design has the potential to bridge the gap between research and clinical transformation. We also saw rapid and significant progress in the field, as this heralds that the future is bright here, and we will be able to develop algorithms to be implemented in a more targeted manner.

Recommendations

1- Expanding research into the applicability of unlabeled data capabilities, the exciting possibility of training deep learning algorithms when only a limited number of labels is available. Although the current performance is not high enough to be applied in a real clinical environment, this method is an important step towards overcoming the challenge of data availability through Sparse data management.

2- It may be good to use the self-learning method here according to the appropriate data, but we also recommend finding other data sources that support the application of the idea and do not violate data ethics and privacy because here we need an ideal model that deals with different variables every second collaborate between hospital and data scientist maybe help, so it will be very valuable to have enough data.

3- In conclusion, we recommend increasing the research and completing the development process in this field, because it will be a distinguished shift for the science of surgery. Because of its importance, we see that shedding light on this topic and the challenges it faces may help us to find creative ideas that facilitate and support the development process.

References

[1] L. Maier-Hein *et al.*, “Surgical data science–from concepts toward clinical translation,” *Med Image Anal*, vol. 76, p. 102306, 2022.

[2] L. Maier-Hein *et al.*, “Surgical data science for next-generation interventions,” *Nat Biomed Eng*, vol. 1, no. 9, pp. 691–696, 2017.

[3] Oracle, “What is Machine Learning?” https://www.oracle.com/sa/artificial-intelligence/machine-learning/what-is-machine-learning/ (accessed Feb. 02, 2023).

[4] Julianna Delua, “Supervised vs. Unsupervised Learning: What’s the Difference?” https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning (accessed Feb. 02, 2023).

[5] T. Roß, “Surgical data science in endoscopic surgery.” pp. 16–17, 2022. Accessed: Feb. 02, 2023. [Online]. Available: https://www.ub.uni-heidelberg.de/archiv/30928

[6] T. Roß, “Surgical data science in endoscopic surgery.” pp. 40–41, 2022. Accessed: Feb. 03, 2023. [Online]. Available: https://archiv.ub.uni-heidelberg.de/volltextserver/30928/

[7] T. Roß, “Surgical data science in endoscopic surgery.” pp. 36–37, 2022. Accessed: Feb. 02, 2023. [Online]. Available: https://archiv.ub.uni-heidelberg.de/volltextserver/30928/

[8] S. Ramesh *et al.*, “Dissecting Self-Supervised Learning Methods for Surgical Computer Vision,” *arXiv preprint arXiv:2207.00449*, 2022.

[9] T. Roß, “Surgical data science in endoscopic surgery.” pp. 37–38, 2022. Accessed: Feb. 02, 2023. [Online]. Available: https://archiv.ub.uni-heidelberg.de/volltextserver/30928/

[10] T. Roß, “Surgical data science in endoscopic surgery.” pp. 45–46, 2022. Accessed: Feb. 02, 2023. [Online]. Available: https://archiv.ub.uni-heidelberg.de/volltextserver/30928/