Quantitative Endosurgery Process Analysis by Machine Learning Method

Bojan NOKOVIC

McMaster University Hamilton, Ontario, Canada

Andrew LAMBE

Able Data Science Consulting Mississauga, Ontario, Canada

ABSTRACT

This study demonstrates how endoscopic surgical data can be analyzed using a supervised machine learning (ML) classifier. Before the process begins, a computergenerated 3D image representing a safe zone is inserted into the endoscopic view. During surgery, the Laparo-Guard Augmented Reality System collects positional data. We perform two types of analysis on the collected data. First, we analyze how the surgeon handles laparoscopic surgical tools based on the angular velocity and angular acceleration of the tool. Next, we examine the risk associated with the entire surgical process in relation to the safe zone using all collected data, including the average linear and angular speeds of the surgical tool.

Keywords: Augmented Reality, Endosurgery, Data Processing, Machine Learning, Supervised Learning

1. INTRODUCTION

In this work, we extend our previous work [1], [2] and create an automated framework to assess minimally invasive surgical (MIS) procedures. The solution uses augmented reality (AR) technology to collect data and machine learning (ML) algorithms to classify risks observed in the collected data.

Endosurgery is a subfield of MIS that uses an endoscope to look inside the body [3]. The endoscope consists of a small camera and light inserted into the body, allowing the surgeon to look at what is wrong inside the body. A related subfield is laparoscopic surgery, also referred to as keyhole surgery. It is a surgical technique in which operations are performed through small incisions (usually 0.5-1.5 cm). The procedure is observed using an endoscope camera that projects a view of the body in real time to the surgical display. The main advantage of MIS surgery compared to open surgery is reduced pain and shorter recovery time for the patient. However, these advantages are achieved only if the procedure is performed without effective errors. Unfortunately, such errors are not uncommon in laparoscopic surgeries. Indeed, intraoperative and post-operative complications are prevalent with laparoscopic surgery procedures [4]. Because of this, there is a need to improve patient safety during laparoscopic surgery so that the benefits derived from such

procedures are achieved while the drawbacks are reduced or eliminated. One of the most profound drawbacks of laparoscopic surgery is the occurrence of unintentional or inadvertent injuries to tissue structures adjacent to or, occasionally, distant from the intended surgical site or field. Bleeding has been reported to occur with an incidence of up to nearly 10% in various series [5]. In the pelvic cavity, for example, bowels, ureters, large organs, and blood vessels can be injured either directly from the heat or sharpness of the laparoscopic instruments, or indirectly through the conduction of heat through nearby tissues. Typically, such injuries are not appreciated at the time of surgery because the specific injury sites are hidden by blood or other tissues. A further complication of such unintended ("iatrogenic") injuries is that the body's response to the injury is often a delayed one. This delayed response can be traumatic to the patient and can sometimes result in one or more further, previously unnecessary surgeries [6].

Augmented reality (AR) is a real-time view of the physical world combined with computer-generated images. The amount of information in AR is always greater than in reality itself. AR enhances the user's perception of and interaction with the real world. Using the latest AR techniques and technologies, the information about the surrounding real world becomes interactive and usable [7]. Areas of current AR implementation are advertising and commercial, entertainment and education, medical, and mobile applications for smartphones [7]. In this paper, we are focused on AR in medical, and surgical procedures. AR indicates the safe or unsafe zone and warns the surgeon in real time if the surgical tool is approaching the safe zone boundary. The computergenerated image of the safe zone is combined with the endoscope camera image in real time.

Machine learning (ML) is a branch of artificial intelligence (AI) that enables a computer to learn to perform tasks by analyzing a large dataset without being explicitly programmed [8]. AI is based on the assumption that the process of human thought can be mechanized. While many famous scientists, from Leibniz to Boole, Tesla, and Turing, were studying the theory of learning, the term *artificial intelligence* was coined in 1955 by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon. ML-based AI has become popular due to its ability to find complex patterns in large sets of data, whether that data comes from tables, plain text, or images.

There are three broad categories of ML algorithms: supervised, unsupervised, and reinforcement. In this paper, we use a supervised learning algorithm to classify surgical risks.

2. RELATED WORKS

ML applications in surgery have been rapidly expanding in recent years [9], [10]. ML has been used to evaluate surgeon performance [11] based on video from an endoscopic camera. ML can help improve robotic surgical training by tracking surgeon eye movements, physical motion, and cognitive function [12].

Significant research has been done to incorporate AR with medical imaging and instruments to enhance the physician's intuitive abilities. An example is a video image of the operating site inside the body recorded by an endoscopic camera device and presented on the surgical monitor. However, the image and positions of the surgical tools are viewed in 2D. That limitation can be partially eliminated by tool navigation techniques that augment the physician's view [13].

The Italian research group ICAR CNR has worked on several projects related to the visual presentation of noninvasive imaging on an advanced multifunction display. One of their projects aims to develop software technologies and novel methodologies for morphological and functional medical imaging applications [14].

AR can be used to manage a patient's medical history, post-stroke rehabilitation, treatment of psychological disorders, and improve navigation for the visually impaired, among other applications [7]. For instance, doctors can check a patient's medical record by putting on a head-mounted display (HMD) and looking over the patient to see virtual labels showing the patient's past injuries and illnesses [7]. So far, HMDs have been used for 3D visualization. The motion of the HMD is tracked to compute an appropriate perspective for the physician and tested to increase clinical acceptance. However, HMD systems still have a problem of lag for motion parallax and cannot provide a natural view for multiple observers [15].

Visual tracking in MIS that uses image processing of surgical tools with special markers is presented in [16]. The main idea is that surgeons can use a gesture as a command. For instance, opening and closing the laparoscopic instrument, e.g., a grasper, is recognized by advanced image processing systems as the command to overlay previously taken computed tomography (CT) or magnetic resonance imaging (MRI) images onto the endoscopic camera view.

Ultrasound augmented with virtual reality for intracardiac surgery is presented in [17]. Tracking is provided by an Aurora magnetic tracking system that determines spatial position (sub-millimetre) and orientation (subdegree) tracking. They clearly showed that the limitation of ultrasound tracking (in the absence of direct vision) can be corrected by applying AR to tracked 2D or 3D ultrasound images.

In addition to technical challenges, AR deployment in endoscopic surgery faces other issues related to the education and retraining of the medical staff. To the best of our knowledge, this is the first work that describes a framework of how to automatically assess the safety of endoscopic surgical procedures by ML classifier using data from an optical measurement system [18].

3. RISK ESTIMATION

An infrared (IR) camera [18] collects the position and orientation of the rigid tool body at a rate of up to 60 frames per second. We collect data for all six degrees (6D) of freedom (x, y, z, roll, pitch, yaw) of the surgical tool tip, as shown in Figure 1.



Fig. 1: Moving Directions in 6D

Based on the input data, we perform two types of analysis:

- A) First, we analyze the input data without considering a virtual safe zone. We consider only the recorded angular velocity and angular acceleration of the tool.
- B) Second, we analyze the data with a predefined safe zone. We examine the risk of the entire surgical process considering four features of the preprocessed data: angular velocity, angular acceleration, average linear speed, and average angular speed.

A. Data Analysis without Considering a Safe Zone

This kind of analysis focuses on estimating the surgeon's skills. For example, if the hands of the surgeon start to shake, that will be noticed immediately in the data. Furthermore, we can identify the time when this event, or some other irregular behavior, happens. We quantify how the surgeon *drives* the surgical tool and automatically analyze surgical skills.



Fig. 2: Average Linear Speed (ALS)

From Figure 2 we can see that during the first two seconds, the average speed in the x direction is higher than the average speed in the y and z directions. In three seconds, the average speed in x decreases from 200 mm/sec to around 40 mm/sec. That is an indication that the instrument is inserted vertically and in approximately three seconds reaches the target of the surgical procedure. From that moment, the average speed in the y and z directions and it stays in the range between 20 mm/sec and 40 mm/sec. This is an example of a normal start of the surgical procedure.



Fig. 3: Average Angular Speed (AAS)

Average angular speed is shown in Fig 3. The angular speed around the z axis (yaw) decreases from 0.45 rad/s to about 0.10 rad/s in three seconds. That trend follows that of the average x speed in the first three seconds, but the variation between the angular speeds is less than the variation in the linear speeds because the instrument has to be moved to the target at the beginning of the procedure.



Fig. 4: Angular Velocity (AV)



Fig. 5: Angular Acceleration (AA)

The angular velocity components *yaw*, *pitch*, and *roll* are shown in Fig 4. This figure shows the small variations and changes of direction not visible in the average speed shown in Fig 3. Finally, we present the angular acceleration in Fig 5, which is the change in angular velocity. We can see that the acceleration is strongest in the *roll* dimension, varying from 6 rad/s^2 to -8 rad/s^2 . This wide range may be the result of the type or nature of the instrument used. For example, the *Scissors* tool may be easier to manipulate than the *Graspers* tool, resulting in different angular velocity.

To illustrate the process of risk assessment, we choose several boundary values in Tables I to IV. In a real application, those boundaries should be configurable and chosen based on the situation. For example, we estimate the risk level due to average linear speed (ALS) according to Table I. If the ALS is less than 30 mm/s the risk is low, if it is in the range 30 to 50 mm/s the risk is medium, and if it is greater than 50 mm/s the risk is high.

TABLE I: The risk level with respect to Average Linear Speed

ALS $[mm/s]$	Risk Level
< 30	Low
[30, 50]	Medium
> 50	High

Tables II, III, IV list the risk levels for other measured components. Note that the units for the features are different $(mm/s, rad/s, rad/s, rad/s^2)$ but they are ignored in the supervised learning training process.

TABLE II: The risk level concerning Average Angular Speed

AAS $[rad/s]$	Risk Level
< 0.1	Low
[0.1, 0.2]	Medium
> 0.2	High

TABLE III: The risk level concerning Angular Velocity

AV $[rad/s]$	Risk Level
< 0.15	Low
[0.15 , 0.17]	Medium
> 0.17	High

TABLE IV: The risk level with respect to Angular Acceleration

AA $[rad/s^2]$	Risk Level	
< 2	Low	
[2 , 6]	Medium	
> 6	High	

Combining risk values of *yaw*, *pitch* and *roll* for AV and AA from tables III, IV we create a training data set for the supervised ML model as shown in Listing 1.

In addition to the six input data columns, we add a seventh (output) column for the risk class. The risk class can be *low*, *medium*, *high*, or *very-high*. The class is determined by assigning numerical values to each risk attribute, 0 for *low*, 1 for *medium*, and 2 for *high*, and then summing up these values. Since we track six attributes, the maximum possible score is 12. Table V shows the risk classes assigned to each possible sum of attribute risks.

TABLE V: The risk level assigned to the sum of attribute risks

Sum of risk values	Risk Level
≤ 3	Low
[4, 5]	Medium
[6, 7]	High
≥ 8	Very High

The file snippet in Listing 1 shows a few examples of attribute ranges and risk values. This file is formatted as an ARFF (Attribute-Relation File Format) file, used by the Weka machine learning software [19]. Because the file has six input columns, each with three possible risk levels, the file has $3^6 = 729$ rows to cover all possible risk scenarios.

For example, the first line below @DATA in the Listing 1 states if the angular velocity components AV_yaw , AV_pitch , and AV_roll are all less than 0.15 rad/s, and angular acceleration components AA_yaw , AA_pitch , AA_roll are less than $\pm 2 rad/s^2$, the overall risk is *low*.

B. Data Analysis Concerning the Safe Zone

In this process, we do a similar analysis as described in Section 3-A, but now we include the safe zone. We detect movements that can be potential violations of the safe zone. Our safe zone is a virtual 3D cone connecting the entry point of the surgery with the operation target. Figure 6 sketches this zone. To illustrate the position of the tip of the surgical tool, we introduced the condition that the target site and the surgical entry point are connected by concentric circles of decreasing radius. Listing 1: ARFF File Snippet

```
@DATA
lt -1/15, lt -1/15, lt -1/15,
1t - 2, 1t - 2, 1t - 2, 1ow
lt -1/15, lt -1/15, lt -1/15,
1t - 2, 1t - 2, 2 - 6, low
lt -1/15, lt -1/15, lt -1/15,
gt-6, 2-6, 1t-2, medium
lt -1/15, lt -1/15, lt -1/15,
gt-6, 2-6, 2-6, medium
lt -1/15, lt -1/15, 1/15 - 2/17, gt -6,
gt-6, lt-2, high
lt -1/15, lt -1/15, 1/15 - 2/17, gt -6,
gt-6, 2-6, high
lt -1/15, lt -1/15, 1/15 - 2/17, gt -6,
gt-6, gt-6, very-high
. . .
```



Fig. 6: Conical Virtual 3D Safe Zone

The process of inserting virtual objects into a real video stream is described in [20]. The position of the endoscopic camera and the orientation of the target object are determined by the position of the attached fiducial markers.



Fig. 7: Virtual 3D zone with guidance data and distance to safe zone boundary indication

Figure 7 shows the safe zone with additional guidance data to indicate the distance Δ from the tip of the surgical tool (x_t, y_t, z_t) to the closest point on the surface of the safe zone (x_s, y_s, z_s) .

$$\sqrt{(x_t - x_s)^2 + (y_t - y_s)^2 + (z_t - z_s)^2} = \Delta \quad (1)$$

Similar to before, the risk level is a combination of ALS, AAS, AV, and AA data, but now we have one more component: the shortest distance between the surgical tool and the safe zone Δ

$$(ALS, AAS, AV, AA, \Delta) \tag{2}$$

Using these five features and the total risk level computed from Table V, we can construct a training dataset similar to that found in Listing 1 for this case.

Table VI shows the risk level with respect to different values of Δ . As long as the distance between the tip of the surgical tool and the safe zone is greater than 4mm, the procedure is deemed *safe* or low risk. If the distance is between 3mm and 4mm, the risk is *medium*. If the surgical tool is even closer to the surface of the safe zone, the risk level is higher.

TABLE VI: Risk level with respect to Δ .

$\Delta \ [mm]$	Risk Level	
> 4	Low	
[3, 4]	Medium	
[1.5, 3]	High	
[1.1, 1.5]	Severe	
≤ 1.1	Critical	

4. DATA ANALYSIS EXAMPLE

A. ML Models

We now create supervised classifiers, machine learning models to classify risk using the Waikato Environment for Knowledge Analysis (Weka) machine learning and data analysis experimental software developed at the University of Waikato, New Zealand [19].

We created two models, one without considering the safe zone and the other considering it. The model that does not consider the safe zone uses the ARFF file described in 3-A as training data. As stated above, this file has 729 rows $((3 \cdot 3 \cdot 3)^2 = 729)$ to account for all of the risk levels in angular velocity and acceleration.

We use a second ARFF file for the model that considers a safe zone. In this model, we assess risk using the highest risk component (in the x, y, or z axes) of angular velocity, angular acceleration, average linear speed, and average angular speed, in addition to the risk level given by Δ . In this file, there are 405 rows ($3 \cdot 3 \cdot 3 \cdot 5 = 405$) to account for all possible combinations. While we could look at every component of the input data, the resulting data file would have $5 \cdot 3^{12}$ or about 2.66 million rows. Such a large dataset would take significantly longer to train.

Both models are relatively small. The first is created on risk assumptions from tables III and IV, while the second is created on risk assumptions from tables I- IV and VI. Note that the values for each of the measured components are hard coded just to demonstrate the process. In the full implementation, these values should be configurable based on experience and experiment.



Fig. 8: Risk Estimation Process

B. Data Prepossessing

The risk estimation process is shown in Figure 8. Before using the ML Model on real data, we need to preprocess it using the risk level tables we defined earlier. For the first model, we only use Tables III and IV with angular velocity and acceleration data.

For the second model, we need additional processing. This means calculating in real-time the average linear and angular speeds since the start of the procedure. Then we take the maximum of the values in the *x*, *y*, and *z* directions and classify these results according to Tables I and II. Finally, we compute Δ , the distance from the safe zone edge, and use Table VI to classify the result. By selecting only the worst velocity and acceleration components for this model, we estimate the *worst case* risk.

C. Risk Estimation Contingency

We use the PART decision list classifier algorithm to create our ML classifiers[19], [21]. This algorithm builds a partial C4.5 decision tree in each iteration and makes the *best* leaf into a rule. The algorithm option attributes are set to default values.

After training, our first ML model, which estimates risk without taking into account the safe zone, can correctly classify 559 out of 729 situations or 76.7%. The confusion matrix generated by the model is shown in Table VII.

TABLE VII: Confusion matrix - the first ML model

Low	Medium	High	Very-high	\leftarrow Classified as
13	15	0	0	Low
4	105	30	1	Medium
0	34	176	57	High
0	0	29	265	Very High

The number of correctly classified instances is the sum of diagonal values; all other instances are incorrectly classified. The *precision* is defined per class as the proportion of the examples that were correctly predicted as a particular class among all those that were classified as that class. When looking at the confusion matrix, this value is computed by dividing the value on the diagonal by the sum of values in the same column. For example, for the *very-high* risk class, the precision is 265/(265+57+1) or 82.0%. For the *high* risk class, the precision is 176/(176+30+29) or 74.9%.

Similarly, *recall* is defined per class as the proportion of examples in that class that were correctly classified. In the confusion matrix, this value is computed by dividing the value on the diagonal by the sum of values in the same row. For example, for the *very-high* risk class, the recall is 265/(265+29) or 90.1%. For the *high* risk class, the precision is 176/(176+57+34) or 73.3%. High values of recall and precision increase our confidence that the model can correctly identify the most high-risk scenarios.

The confusion matrix for the second ML model, which does incorporate information about the safe zone, is shown in Table VIII. This model correctly classified 385 out of 410 instances or 95.1%. The precision of the very high- and high-risk cases are 100% and 93.3% respectively, while the recall of these cases is 85.7% and 93.3% respectively. We also see zero cases in which the risk classification is two or more levels higher or lower than the true value. (e.g., there are no instances in which a very high-risk situation is classified as medium- or low-risk.)

TABLE VIII: Confusion matrix - the second ML model

Low	Medium	High	Very-high	\leftarrow Classified as
105	0	0	0	Low
5	80	0	0	Medium
0	10	140	0	High
0	0	10	60	Very High

5. CONCLUSION AND FUTURE WORK

In this work, we describe an automated process to assess risk by analyzing data collected during laparoscopic procedures. The input data represent the position of the tip of the laparoscopic surgical tool during the procedure. Based on the movement of the tool, we estimate the quality of the surgical process. We train a supervised ML model to classify the risk of the process and apply this method to one surgical procedure that does not consider a safe zone and one that does. Quantifying the risk enables us to estimate the surgeon's skill level and the overall quality of the endoscopic procedure.

There are three areas for improvement in our work. The first area concerns performing analysis in real time. Currently, we analyze previously recorded data. In the future, we intend to implement real-time analysis to provide immediate warnings to surgeons if a risky maneuver is detected. When the safe zone is not considered, implementing risk estimation using our first ML model is a straightforward process. For the second model, we anticipate a slight delay in generating warnings due to the need to calculate the ALS, AAS, and Δ values. At a frame rate of approximately 30 to 60 frames per second,

averaging values over 10 frames results in a delay of approximately 0.165 to 0.33 seconds.

The second area for improvement involves customizing the ML model to enhance estimation accuracy. One approach is to develop a customized PART classifier. Alternatively, selecting other model types—such as the Multi-layer Perceptron (MLP) classifier—may further improve accuracy.

The third area for improvement involves implementing an early warning feature. Real-time analysis enables the prediction of future movements, with some probabilistic confidence, based on past motion data. This feature allows for advance warnings to surgeons when risky movements are detected, enabling them to adjust their actions proactively, such as by reducing their speed.

6. ACKNOWLEDGMENT

We would like to thank Mitch Wilson, the President of Mariner Endosurgery, for providing us with the data from the Optical Tracker.

REFERENCES

- B. Nokovic and T. Zhang, "Future Augmented Reality in Endosurgery," 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:115135771
- [2] B. Nokovic, N. Djosic, and W. Li, "API Security Risk Assessment Based on Dynamic ML Models," 11 2020, pp. 247–252.
- [3] T. Smith and B. M. Association, The British Medical Association Complete Family Health Encyclopedia. Dorling Kindersley, 1995. [Online]. Available: https://books.google.ca/ books?id=6Y8QAAAACAAJ
- [4] L. Draghici, I. Draghici, A. Ungureanu, C. Copaescu, M. Popescu, and C. Dragomirescu, "Laparoscopic surgery complications: Postoperative peritonitis," Journal of Medicine and Life, vol. 5, pp. 288–96, 09 2012.
- [5] R. Kaushik, "Bleeding complications in laparoscopic cholecystectomy: Incidence, mechanisms, prevention and management," Journal of Minimal Access Surgery, vol. 6, pp. 59–65, 07 2010.
- [6] B. Nokovic, M. Wilson, and N. Nedialkov, "Laparoscopic Surgery System Calibrator and Method for Using the Same," 2016, Provisional patent US 62/439,845.
- [7] B. Furht, **Handbook of Augmented Reality**. Springer Publishing Company, Incorporated, 2011.
- [8] Merriam-Webster, 2024. [Online]. Available: https://www. merriam-webster.com/
- [9] P. Satapathy, K. Pradhan, S. Rustagi, V. Suresh, Z. Al-qaim, B. Padhi, and R. Sah, "Application of Machine Learning in Surgery Research: Current Uses and Future Directions," International Journal of Surgery (London, England), vol. Publish Ahead of Print, 04 2023.
- [10] M. Morris, A. Rajesh, M. Asaad, A. Hassan, R. Saadoun, and C. Butler, "Deep Learning Applications in Surgery: Current Uses and Future Directions," **The American Surgeon**, vol. 89, pp. 36–42, 01 2023.
- [11] A. Hung, J. Chen, and I. Gill, "Automated Performance Metrics and Machine Learning Algorithms to Measure Surgeon Performance and Anticipate Clinical Outcomes in Robotic Surgery," JAMA Surgery, vol. 153, 06 2018.
- [12] M. Egert, J. E. Steward, and C. P. Sundaram, "Machine Learning and Artificial Intelligence in Surgical Fields," Indian Journal of Surgical Oncology, vol. 11, no. 4, pp. 573–577, 2020.
- [13] C. Bichlmeier, F. Wimmer, S. M. Heining, and N. Navab, "Contextual Anatomic Mimesis Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality," in 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, Nov 2007, pp. 129–138.

- [14] L. di Calcolo e Reti ad Alte Prestazioni del Consiglio Nazionale delle Ricerche, "Non Invasive Imaging For Advanced M.F.D." 2017. [Online]. Available: http://www.icar.cnr.it/en
- [15] H. Liao, **3D Medical Imaging and Augmented Reality** for Image-Guided Surgery. New York, NY: Springer New York, 2011, pp. 589–602. [Online]. Available: http: //dx.doi.org/10.1007/978-1-4614-0064-6_27
- [16] S. Payandeh, Visual Tracking in Conventional Minimally Invasive Surgery. Chapman and Hall/CRC, 2016.
- [17] D. Bainbridge, D. L. Jones, G. M. Guiraudon, and T. M. Peters, "Ultrasound Image and Augmented Reality Guidance for Off-pump, Closed, Beating, Intracardiac Surgery," Artificial Organs, vol. 32, no. 11, pp. 840–845, 2008. [Online]. Available: http://dx.doi.org/10.1111/j.1525-1594.2008.00639.x
- [18] NDI, "3D Measurement Tracking Technology," 2024. [Online]. Available: http://www.ndigital.com/
- [19] M. A. H. Eibe Frank and I. H. Witten, Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 2016.
- [20] J.-S. Park, M. Y. Sung, and S.-R. Noh, "Virtual Object Placement in Video for Augmented Reality". Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 13–24.
- [21] E. Frank and I. Witten, "Generating Accurate Rule Sets Without Global Optimization," Machine Learning: Proceedings of the Fifteenth International Conference, 06 1998.