Moritz Spiller*, Nazila Esmaeili, Thomas Sühn, Axel Boese, Michael Friebe, Salmai Turial, and Alfredo Illanes

Towards AI-driven minimally invasive needle interventions

https://doi.org/10.1515/cdbme-2023-1140

Abstract: The overall complication rate during laparoscopic access is estimated to be as high as 14 %. Surgeons have to rely heavily on their experience and haptic perception while inserting the Veress needle or a trocar into the peritoneal cavity. Surgical Audio Guidance (SURAG) is a promising alternative to current techniques. It acquires instrument-born vibroacoustic (VA) waves to track the insertion of the instrument and provide real-time feedback to surgeons. This article presents an initial evaluation of the SURAG technology through two sets of experiments to classify Veress needle events using different AI-models. The results demonstrate the feasibility of using AI for classifying Veress needle events and the potential of the SURAG technology to support surgeons during laparoscopic access and minimally invasive needle interventions in general.

Keywords: Artificial Intelligence, Needle Interventions, Audio Sensing, Surgical Support Systems

1 Introduction

Establishing laparoscopic access is still a risky step in laparoscopic and robot-assisted surgical procedures. Between 30% and 50% of all complications in laparoscopy occur during the creation of laparoscopic access [4]. The incidence of serious intraoperative complications, such as the injury of intraabdominal organs, is reported to be as high as 4%, while the overall complication rate is up to 14 % [4].

***Corresponding author: Moritz Spiller, Nazila Esmaeili, Thomas Sühn, Axel Boese, Michael Friebe, Alfredo Illanes,** SURAG Medical GmbH, Brenneckestraße 20, 39120 Magdeburg, Germany, moritz@surag-medical.com
Axel Boese, Michael Friebe, INKA Innovation Laboratory for Image Guided Therapy, Otto-von-Guericke University Magdeburg, 39120 Magdeburg, Germany
Salmai Turial, Faculty of Pediatric Surgery, Department of General, Visceral, Vascular and Transplantation Surgery, Otto-von-Guericke University Magdeburg, 39120 Magdeburg, Germany
Michael Friebe, Department of Biocybernetics and Biomedical Engineering, AGH University of Science and Technology, 30-059 Kraków, Poland
Nazila Esmaeili, Department of Otorhinolaryngology, Head and Neck Surgery, Faculty of Medicine, Justus Liebig University of Giessen, Giessen, Germany

The goal of laparoscopic access is the insufflation of carbon dioxide into the peritoneal cavity to ensure sufficient operative space and field of vision inside the abdomen during the laparoscopic procedure. For that, the surgeon first needs to place the Veress needle's tip into the peritoneal cavity. Depending on the insertion location, the needle has to puncture one (Umbilical point) or two fascias (Palmer's point) and the peritoneum before arriving to the cavity. During the insertion of the Veress needle, surgeons can feel a decreasing resistance after passing through a fascia or the peritoneum. Thereby, and by taking the spring mechanism of the Veress needle into account, surgeons can orientate themselves. According to a recent study, 36 % of surgeons reported feeling unsafe and perceiving a risk of injuring the patient during laparoscopic access. Moreover, 55 % of those surgeons stated that the spring mechanism of the Veress needle does not clearly indicate when it has reached the peritoneal cavity. [8] A direct relationship between the surgeons' experience and their perception of the feedback was reported [5].

Due to the complications that are associated with standard Veress needles and trocars, needles and trocars with added guidance characteristics have been developed and are commercially available today.

Single- and multi-use optical trocars allow using a laparoscope that transmits real-time monitor images while transecting abdominal wall tissue layers. However, those instruments have not proven to provide significantly more safety than the Veress needle [1].

Additionally, the scientific community has suggested multiple force sensor based solutions for laparoscopic access [2], but they were neither tested in realistic scenarios, nor pursued further.

Surgical Audio Guidance (SURAG) is a novel guidance technology for obtaining complementary information from surgical instruments that has shown high potential for guiding needles during minimally invasive interventions [3, 6]. Vibroacoustic (VA) waves resulting from any interaction between the needle's tip and the tissue naturally propagate through the needle's shaft. They can be acquired and detected using a sensing module located at the proximal end of the instrument. Subsequently, the signals can be processed to extract meaningful guidance information that can be further translated into interpretable feedback to surgeons. Events such as tissue-tissue passages and punctures of layers like fascia or peritoneum can

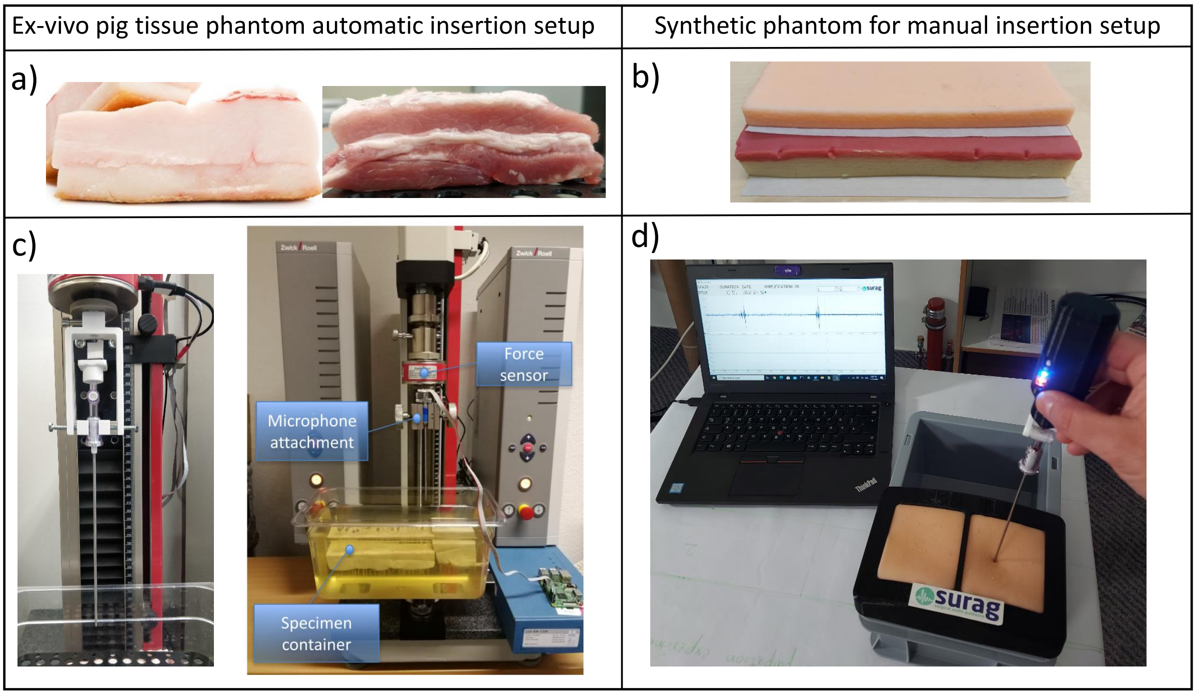

Fig. 1: Experimental setup with the Zwick force testing machine.

be automatically identified, magnified and translated to visual and/or acoustic feedback.

This article summarizes two sets of experiments that evaluated the feasibility of using Artificial Intelligence (AI) for classifying Veress needle events to support surgeons during laparoscopic access. At first, a study was performed on an artificial phantom to evaluate if the time and frequency domain features suggested in [7] can be used to distinguish different Veress needle events such as tissue-tissue passage and arrival to the peritoneal cavity. The second set of experiments was performed to study if AI-models that are able to classify Veress needle events can also run on an embedded system.

The results indicate that the classification of punctures to the fascia and the peritoneum during laparoscopic access can be classified with high accuracy using simple, and thus real-time capable, machine learning classifiers such as Random Forest. Even though AI-models embedded in a microcontroller do not yield good performance yet, this could be a significant step towards AI-driven interventions in minimally invasive surgery.

2 Materials & Methods

Both experiments were conducted using an audio-based sensing device (SURAG Medical GmbH) attached to the proximal end of a Veress needle. The device recorded the instrument-born vibrations at 16,000 Hz sampling frequency with a resolution of 32-bit.

In the first experiment, a synthetic phantom with two layers of artificial TPE material was used to simulate the skin, subcutaneous fat and pre-peritoneal fat. As it is shown in Figure 1 b and d, these layers were separated by a non-woven fabric to mimic the fascia and peritoneum layers. A total of 400 vibroacoustic signals were recorded while manually inserting a Veress needle into the phantom, with two main tissue-tissue-passage events manually segmented and labeled as fascia and peritoneal cavity events. Data analysis included signal pre-processing, feature extraction, and ML-based classification steps. 17 features based on the time (T) and time-frequency (TF) domain characteristics of the signal were computed [7]. For classification, two supervised classifiers, Support Vector Machine (SVM) and Random Forest (RF), were used. The data set was randomly split into 80 % training and 20 % testing samples. In every classification scenario, the training set was used to perform the hyperparameter tuning using the grid search method.

Figure 1 a and c presents the experimental setup of the second experiment, where a phantom using the abdominal wall of a pig was created to evaluate the performance of the SURAG technology in combination with ML-based classification techniques in a more realistic condition. During the initial trials for validating the experimental setup, annotating needle-tissue-passage events turned out to be challenging due to the

				Training Set			Test Set		
Model no.	**Model**	**Dataset**	**Class**	**Accuracy**	**Precision**	**Recall**	**Accuracy**	**Precision**	**Recall**
1	SVM	T+F	Fascia	82.2 %	0.70	0.90	82.2 %	0.70	0.81
			Peritoneal Cavity		0.99	0.96		0.97	0.95
2	SVM	T	Fascia	82.7 %	0.77	0.77	79.9 %	0.80	0.70
			Peritoneal Cavity		0.99	0.96		0.95	0.93
3	SVM	F	Fascia	70.4 %	0.64	0.75	67.8 %	0.60	0.71
			Peritoneal Cavity		0.75	0.83		0.77	0.86
4	RF	T+F	Fascia	92.1 %	0.87	0.94	87.9 %	0.80	0.90
			Peritoneal Cavity		0.98	0.97		0.97	0.96
5	RF	T	Fascia	90.1 %	0.82	0.95	83.4 %	0.74	0.88
			Peritoneal Cavity		0.99	0.96		0.99	0.95
6	RF	F	Fascia	82.6 %	0.75	0.87	65.8 %	0.57	0.63
			Peritoneal Cavity		0.87	0.90		0.80	0.87

Tab. 1: Results of the Support Vector Machine (SVM) and the Random Forest (RF) classifier on the three different datasets consisting of time domain features (T), frequency domain features (F), and, both, time and frequency domain based features (T+F).

porcine tissue's complexity. Therefore, the VA data acquisition was conducted under a controlled condition where the Veress needle with the attached SURAG system was connected to a force testing machine (Zwick GmbH) that automatically inserted the needle into the phantom at a controlled velocity and force. In this experiment. a total of 1415 VA signals were acquired and stored by the SURAG system. The force measurements acquired by the force testing machine were used as the ground truth to automatically segment the two main needle-tissue-passage events in each VA signal and annotate them as fascia and peritoneal cavity events to create the data set. Finally, the data set included 756 and 285 samples of the fascia and peritoneal cavity events, respectively. The data analysis step included developing and testing two Convolutional Neural Networks (CNN) to evaluate the performance of these models in classifying needle-tissue-passage events of the VA signals and their ability to run on an embedded system. The VA signal pre-processing was conducted with DC bias removal, normalization, audio padding, and data augmentation. Then, the classification scenarios were conducted using a one-dimensional Residual Neural Network (ResNet) with five residual blocks and a depthwise separable CNN (DS-CNN). The pre-processed time domain VA signals were used as input for the ResNet architecture, while 2-dimensional (2D) spectrograms generated from VA signals were applied as input for DS-CNN. To train, validate and test the models, the generated data set was randomly divided into training and testing sets, including 80% and 20% of the total samples, respectively. The training set was again split into subsets to perform the validation step. An Adam optimizer with a learning rate of $5e^{-6}$ was used for models' training with up to 200 epochs. In order to avoid overfitting, the early stoppage was set with a patience of 5 epochs. The classification step was conducted on a computer with the NVidia P6000 24GB GPU. The classification performance of the ResNet and DS-CNN model was evaluated using accuracy, F1-score, precision, and recall metrics. Moreover, the size and number of parameters of every architecture were studied to assess the ability of the models to run on a microcontroller.

3 Results & Discussion

In the first experiment, six different classification scenarios were conducted to evaluate the performance of the extracted features on classifying VA events into the fascia and peritoneal cavity classes. Therefore, the data set was split into three subsets (time domain feature set, time-frequency domain feature set, combined time and time-frequency domain feature set), where every ML-based classifier was individually trained and tested on.

Table 1 represents the results of the classification scenarios in the first experiment. The best performance by both classifiers is achieved in scenarios 1 and 4, where the combination of two time and time-frequency domain feature sets are used for VA event classification. It is important to mention that the stand-alone application of the time-frequency domain feature set contributed to higher classification results compared to the classification scenarios that were conducted only with the time domain feature set. Moreover, both classifiers showed high accuracy in the classification of the peritoneal cavity event indicating the high performance of the features in detecting the arrival of the Veress needle into the cavity. Recent studies showed that this information is the most important input that surgeons need to avoid complications during laparoscopic access and other minimally invasive needle interventions [8].

Model	Validation Acc.	Test Acc.	Size	Params
ResNet	72.7 %	62.0 %	3.9 MB	303,683
DS-CNN	79.18 %	60.0 %	2 MB	0.19M

Tab. 2: Validation Accuracy, Test Accuracy, the model's size and the number of parameters for the trained Deep Learning models.

The classification results of the ResNet and DS-CNN in the second experiment are shown in Table 2. In general, the ResNet model with 62% testing accuracy showed higher performance than the DS-CNN in the classification of needle-tissue-passage events of the VA signals. Moreover, the ResNet has a smaller number of parameters than the DS-CNN, which could make it more suitable for real-time applications on a microcontroller, despite it being slightly larger in size. It is important to highlight that both architectures classified the fascia event better than the peritoneal cavity. However, with respect to precision, recall, and F1-score, the ResNet (0.81, 0.92, 0.86) outperformed the DS-CNN (0.69, 0.55, 0.61). This performance could be caused by the larger number of training samples for fascia. Moreover, ResNet and DS-CNN were trained based on different types of input data (time domain VA signal versus 2D spectrogram). Therefore, each model was built considering different characteristics of the fascia and peritoneal cavity events in the VA signal. This finding requires more research, but it could indicate that the time domain VA signal contains more information than the 2D spectrogram to differentiate between fascia and peritoneal cavity events.

Comparing the results of both experiments showed the importance of a detailed analysis of the acquired VA signals to extract features that represent the main attributes in relation to the target events. Although the second experiment was conducted in more controlled conditions, the data analysis applied in the first experiment showed better performance than the CNN-based techniques of the second experiment for peritoneal cavity event classification. Nevertheless, we need to consider that the synthetic phantom had more stable characteristics than the abdominal wall of a pig, which can affect the behavior of the acquired VA signal.

4 Conclusion

This article describes the evaluation of various AI algorithms for classifying vibroacoustic signals obtained using the novel guidance technology SURAG. The results indicate that the classification of navigation-relevant events during laparoscopic access can be classified with high accuracy using simple, and thus real-time capable, machine learning classifiers. The findings also confirm that SURAG could provide the information surgeons are currently missing when performing laparoscopic access. In the future, the evaluated classifiers should be trained and tested on real-world animal data and a real-time demonstrator should be developed. Additionally, Deep Learning models able to run on a microcontroller should be trained on bigger, more balanced datasets. Nevertheless, the SURAG technology could be a simple and cost-efficient means to introduce real-time AI algorithms into minimally invasive surgery to make it even more precise and safe.

Author Statement

Research funding: This research was partly funded by the German Federal Ministry for Economic Affairs and Energy, EXIST Transfer of Research 03EFOST032, and the European Union, European Social Fund for Germany. **Conflict of interest:** The author S.T. declares no conflict of interest. The authors T.S., N.E., M.S., A.B., M.F. and A.I. are shareholders of SURAG Medical GmbH, which provided the sensing device that was used in this study.

References

[1] S. Bhoyrul, M. A. Vierra, C. R. Nezhat, T. M. Krummel, and L. W. Way. Trocar injuries in laparoscopic surgery. *Journal of the American College of Surgeons*, 192(6):677–683, 2001.

[2] G. A. Fontanelli, L. R. Buonocore, F. Ficuciello, L. Villani, and B. Siciliano. A novel force sensing integrated into the trocar for minimally invasive robotic surgery, 11.04.2023.

[3] A. Illanes, A. Boese, I. Maldonado, A. Pashazadeh, A. Schaufler, N. Navab, and M. Friebe. Novel clinical device tracking and tissue event characterization using proximally placed audio signal acquisition and processing. *Scientific reports*, 8(1):12070, 2018.

[4] Javier F. Magrina. Complications of laparoscopic surgery : Clinical obstetrics and gynecology. *Clinical Obstetrics and Gynecology*, (45):469–480, June 2002.

[5] X.-L. Li, D.-F. Du, and H. Jiang. The learning curves of robotic and three-dimensional laparoscopic surgery in cervical cancer. *Journal of Cancer*, 7(15):2304–2308, 2016.

[6] M. Sabieleish, K. Heryan, A. Boese, C. Hansen, M. Friebe, and A. Illanes. Study of needle punctures into soft tissue through audio and force sensing: can audio be a simple alternative for needle guidance? *International Journal of Computer Assisted Radiology and Surgery*, 16(10):1683–1697, 2021.

[7] A. Schaufler, T. Sühn, N. Esmaeili, A. Boese, C. Wex, R. Croner, M. Friebe, and A. Illanes. Automatic differentiation between veress needle events in laparoscopic access using proximally attached audio signal characterization. *Current Directions in Biomedical Engineering*, 5(1):369–371, 2019.

[8] M. Spiller, M. Bruennel, V. Grosse, T. Sühn, N. Esmaeili, J. Stockheim, S. Turial, R. Croner, A. Boese, M. Friebe, and A. Illanes. Surgeons' requirements for a surgical support system to improve laparoscopic access. *BMC surgery*, 22(1):279, 2022.