This full text paper was peer-reviewed at the direction of IEEE Instrumentation and Measurement Society prior to the acceptance and publication.

Tracking Instrument Handling During Laparoscopic Surgical Training Using Hand Landmarks

Kade MacWilliams[1], James R. Green[1], Ahmed Nasr[2], Georges Azzie[3], and Carlos Rossa[1]
[1]*Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada*
[2]*Division of Paediatric Surgery, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada*
[3]*Department of Surgery, Hospital for Sick Children, Toronto, ON, Canada*

***Abstract*—While laparoscopy has revolutionised surgical procedures by enabling faster recovery and reduced trauma compared to open surgery, it has a steep learning curve, requiring much practice to master the stroke mechanics required for complicated procedures. Instrumented laparoscopic surgical trainers provide a safe environment for surgeons to hone their motor skills, enabling them to perform procedures with confidence, while capturing kinematic and video data from the instruments for performance assessment. Although hand grip is known to affect force control, precision, and most efficiency, most simulators lack sensors outside of the box and cannot capture subtle variations in instrument grip. Accurately detecting such variations would allow teachers to provide more informative guidance to trainees.**

Leveraging MediaPipe Hand Landmarker, we propose an unobtrusive vision-based hand pose sensing framework for detecting different laparoscopic instrument grip and handling techniques during training. By clustering finger-tracking landmarks from feature vectors detected during selected instrument movements, we demonstrate that we can accurately detect two common instrument grip positions as well as different modes of axial instrument rotation that kinematics sensors cannot distinguish. Experimental results obtained using a laparoscopic simulator demonstrate that the algorithm can differentiate tool hand grip changes and two different rotation strategies with 99% accuracy, across 5 subjects. Machine learning models can leverage this new capability to provide more precise guidance to trainees on instrument handling, something current training systems cannot achieve.

I. Introduction

Laparoscopy has revolutionized modern surgical practices with less trauma, faster recovery, and shorter hospital stays compared to open surgery [1], [2]. Laparoscopic procedures are a form of minimally invasive surgery (MIS) where thin surgical instruments and a laparoscope (a tube-like camera) are passed through small incisions in the patient's chest or abdomen [1]. The surgeon must manoeuvre instruments through these access points based solely on a 2D video feed with a limited field of view and constrained motion patterns. Consequently, laparoscopic procedures have a steep learning curve for developing the required cognitive, technical, and motor skills [3], [4]. Current training methods (including intraoperative observation, cadaveric models, virtual reality simulators, and box trainers) provide environments in which trainees can develop and hone their laparoscopic skills [5], [6].

Box trainers offer a safe environment for mastering the stroke mechanics required in laparoscopic procedures. The standard trainer is a box with an internal camera and ports for inserting instruments, along with internal modules for training tasks. These training drills, such as a bean drop and needle passing, have been shown to improve performance in the operating room after 30-35 repetitions and over six months of practice [7], [8]. Box simulators are the most widely used training tool outside of the operating room due to their affordability and ease of deployment. However, a significant limitation of box trainers is the lack of automated assessment and feedback provided to the trainee. An expert surgeon is often required to oversee the training regime, observing trainee drills and providing actionable guidance. Recent advances in automated surgical assessment and feedback evaluate a surgeon's performance in the simulator either through a scoring system, such as the global rating scale, or by estimating their expertise level as novice, intermediate, or expert [9], [10].

To provide automated guidance and assessment, a box trainer is often equipped with sensors and/or cameras to record the position of surgical tools over time [9]. The most basic box trainer uses an internal camera and requires expert assessment of the trainee's performance based on the recorded video feed. Other simulators have attached sensors to the surgical tools or to the trocar site (entry points to the body cavity), such as an inertial measurement unit to capture the orientation of the tool, and optical sensors or time-of-flight sensors to measure surge and roll of the tool [6]. With this information, the location of the tool tip inside the box can be calculated and recorded. More advanced trainers also measure gripping force and tool/tissue contact force. Table I summarizes various laparoscopic trainers and the sensors they use.

Beyond grip strength, current instrumented laparoscopic surgical simulators do not explicitly track the position or movement of the hands, only the resultant movements of the tool tip. This results in ambiguity when the same tool motion can be accomplished with different hand movements or using

Email: kademacwilliams@cmail.carleton.ca (K. MacWilliams); jrgreen@sce.carleton.ca (J.R. Green); anasr@cheo.on.ca (A. Nasr), georges.azzie@sickkids.ca (G. Azzie), carlosrossa@cunet.carleton.ca (C. Rossa).

This research is supported by the Natural Sciences and Engineering Research Council - Canadian Graduate Scholarship (Master's), and the Children's Hospital Academic Medical Organization (CHAMO) innovation fund.

Cette recherche est appuyée par le Conseil de recherches en sciences naturelles et en génie - Bourse d'études supérieures du Canada (Maîtrise) et par les fonds d'innovation CHAMO.

TABLE I: Different laparoscopic trainers and the sensors used to record trainee data. Adapted from [9].

Trainer		Sensors	Tracks
Box Trainer	[11]	Camera	N/A
Box Trainer	[12]	Accelerometer, gyroscope, magnetometer	Tool
Box Trainer	[13]	Force, IMU	Tool
LAPKaans	[14]	Orientation, depth, speed, grasping force	Tool
Virtual reality	[15]	Camera, position, orientation	Tool
Lapro Apex	[16]	Position, camera	Tool
BlueDRAGON	[17]	Position, force	Tool
CyberGlove Sim	[18]	Position, Orientation	Hand

a different grip.

Figure 1 shows a typical laparoscopic surgical tool. To rotate the end of the tool to the desired angle, the user can either rotate the instrument axially by rotating their hand (wrist rotation), or by using their index finger to actuate the rotation knob located around the instrument shaft, which rotates the shaft without rotating the grip. Current instrumented simulators cannot discriminate axial rotation generated by hand motion from rotation produced through the rotation knob. This ambiguity limits the utility of feedback to a trainee, since there may be strategic reasons for preferring one rotation motion over the other in a given situation. Furthermore, these rotations reflect different motor skills and should therefore be analyzed separately.

Hand grip is known to affect force control, precision, and motion efficiency [19]. There are also known associations between poor handling of instruments and misuse injuries among surgeons. Therefore, tracking the way a trainee holds the instrument, rather than only its position and orientation, can allow for a more in-depth analysis of their stroke mechanics, and also allow assessment and feedback models to teach them the correct way of manipulating the tool. Several approaches have been proposed to track a subject's hand position/orientation in space. One approach is to use an instrumented glove that captures the position and orientation of the hand [18]. However, this approach is cumbersome and may interfere with surgical performance and alter the tactile feedback perceived by the user.

In this paper, we propose to leverage advances in landmark detection in live videos to accurately track various points on the hand during laparoscopic training. Google's MediaPipe Hand Landmarker allows for the detection and tracking of 21 landmark points on a person's hand [20], as shown in Fig. 2. Using the algorithm with a depth camera allows for these 21 landmarks to be tracked in 3D space. However, the use of this algorithm to track a hand while handling a surgical tool has not been done before.

Using a box trainer instrumented with external depth cameras and a self-supervised clustering algorithm, we demonstrate that, by combining the tracker algorithm with k-means clustering, we can differentiate axial rotation resulting from wrist rotation from axial rotation generated through the rotation knob, across 5 different subjects. In addition, the proposed algorithm can also detect the two commonly used forms of grasping the surgical instrument shown in Fig. 1(a) and (b),

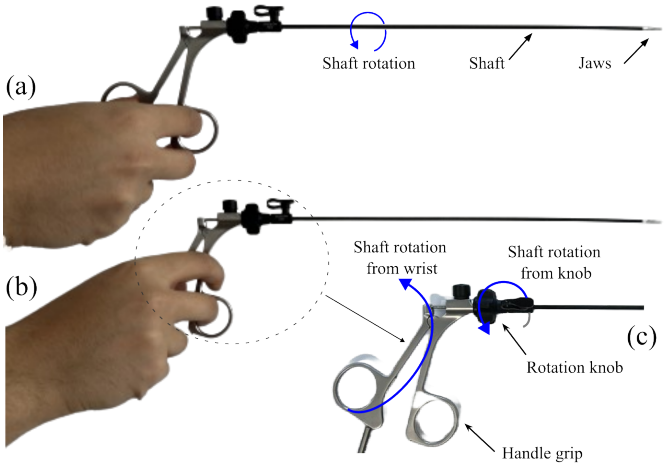

Fig. 1: Two variations of hand grip on the laparoscopic instrument. In (a) the instrument is grasped with the middle finger and thumb. In (b) the instrument is grasped with the index finger and thumb. Figure (c) shows the rotation knob and the handle grip.

i.e., using either the index finger or the middle finger to actuate the griper, also across 5 subjects. Together, this additional information provides a more detailed understanding of hand position and movement strategies than a traditional box trainer instrumented with only internal sensors. To the best of our knowledge, this is the first implementation of visual hand landmark tracking during laparoscopic surgery training.

The remainder of the paper is organized as follows. Section II describes the algorithm used for hand tracking and how it is adapted to laparoscopy training, followed by the unsupervised clustering algorithm for the detection of grip changes and rotation methods. The proposed method is then validated experimentally in Section III using an instrumented laparoscopic simulator equipped with an external depth camera to measure the hand movements. The paper concludes with a discussion of the obtained results, the limitations of the work, and directions for future research.

II. Visual Hand Tracking and Clustering Algorithm

Google's MediaPipe Hand Landmarker algorithm determines the 3-dimensional coordinates (x, y, z) of landmarks defined on the human hand from the RGB-Depth image acquired by a camera [21]. An example is shown in Fig. 2(a), where the spatial location information of 21 points on a trainee's hand are shown. The two dimensional position of each landmark (x, y) is calculated from each image using machine learning algorithms trained on real-world and synthetic datasets, while the z position of each landmark is inferred based on the estimated distance between the hand to the camera. The algorithm can also identify whether the left or right hand is being imaged.

For a video with K frames, the algorithm returns, at discrete time steps $t = 1, 2, ..., \tau$, the x, y, z position of all $d = 21$ tracked landmarks, where $1 \leq i \leq d$. In addition, the algorithm

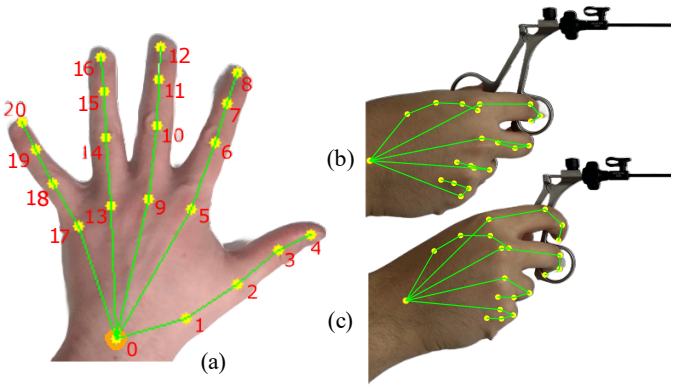

Fig. 2: Example of landmark locations detected by the MediaPipe algorithm. In (a), all 21 landmarks are visible. In (b) and (c), landmarks are shown on the user's hand during instrument manipulation.

returns a binary flag $f_i(t) = 1$ if landmark i is present in the image at time step k, and $f_i(t) = 0$ otherwise,

During tracking, we arrange the τ measured positions of landmark i into vector $\mathbf{v}_i(k)$:

$$\mathbf{v}_i(t) = [v_i(1), v_i(2), \ldots, v_i(\tau)]^T.$$

Since each $v_i(t)$ contains the x, y, z position of a landmark, they are 1×3 vectors. For simplicity, we henceforth denote them as scalars. However, the same process is applied to all dimensions of $v_i(t)$.

To ensure that the position of landmark i taken at time step t is valid, we check whether landmark i is present in the image at time t. If the landmark is not present, the invalid measurement is replaced with the median of landmark i position over the entire length of $\mathbf{v}_i$, that is:

$$v_i(t) \leftarrow \begin{cases} v_i(t), & \text{if} \quad f_i(t) = 1 \\ \text{median}\left(\mathbf{v}_i(t)\right), & \text{otherwise} \end{cases}$$

Invalid measurements were replaced with the median value of valid measurements to ensure consistent temporal length across all dimensions within each sample. The median is robust to outliers, avoids error propagation, and preserves the underlying data distribution.

A. Feature Vector Definition

For each dimension i the mean μ_i, standard deviation σ_i and range r_i are calculated over the trial length τ, as follows. The mean is

$$\mu_i = \frac{1}{\tau} \sum_{t=1}^{\tau} v_i(t). \tag{1}$$

The standard deviation of landmark i is given as

$$\sigma_i = \sqrt{\frac{1}{\tau - 1} \sum_{t=1}^{\tau} ||v_i(t) - \mu_i)||^2}, \tag{2}$$

and the range, i.e., the magnitude of change of a specific landmark over a trial is

$$r_i = \max(\mathbf{v_i}(t)) - \min(\mathbf{v_i}(t)). \tag{3}$$

From these statistics, a feature matrix $\mathbf{G} \in \text{Re}^{d \times 3}$ representing a trial can be constructed as:

$$\mathbf{G} = \begin{bmatrix} \mu_1 & \sigma_1 & r_1 \\ \mu_2 & \sigma_2 & r_2 \\ & \vdots & \\ \mu_d & \sigma_d & r_d \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_d \end{bmatrix}. \tag{4}$$

Assume a set of N trials, where each trial is given the index j, with $1 \leq j \leq N$. Each of the j feature vectors, $\mathbf{G}^j$, are standardized across all N trials by setting:

$$\hat{\mathbf{G}}^j = \begin{bmatrix} \frac{G^j_{1,1} - \hat{\mu}_{1,1}}{\hat{\sigma}_{1,1}} & \frac{G^j_{1,2} - \hat{\mu}_{1,2}}{\hat{\sigma}_{1,2}} & \frac{G^j_{1,3} - \hat{\mu}_{1,3}}{\hat{\sigma}_{1,3}} \\ \frac{G^j_{2,1} - \hat{\mu}_{2,1}}{\hat{\sigma}_{2,1}} & \frac{G^j_{2,2} - \hat{\mu}_{2,2}}{\hat{\sigma}_{2,2}} & \frac{G^j_{2,3} - \hat{\mu}_{2,3}}{\hat{\sigma}_{2,3}} \\ & \vdots & \\ \frac{G^j_{d,1} - \hat{\mu}_{d,1}}{\hat{\sigma}_{d,1}} & \frac{G^j_{d,2} - \hat{\mu}_{d,2}}{\hat{\sigma}_{d,2}} & \frac{G^j_{d,3} - \hat{\mu}_{d,3}}{\hat{\sigma}_{d,3}} \end{bmatrix} \tag{5}$$

with $G^j_{n,m}$ being, respectively, the n^{th} and m^{th} row and column of $\mathbf{G}^j$, and $\hat{\mu}_{n,m}$ and $\hat{\sigma}_{n,m}$ the average and standard deviation of element $G^j_{n,m}$ $\forall\ j \in [1, N]$ across all trials, i.e.,

$$\hat{\mu}_{n,m} = \frac{1}{N} \sum_{j=1}^{N} G^j_{n,m}, \tag{6}$$

and

$$\hat{\sigma}_{n,m} = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} (G^j_{n,m} - \hat{\mu}_{n,m})^2}. \tag{7}$$

B. k-Means Clustering

Once the feature matrices, henceforth referred to as samples, are defined, the objective is to assign each sample to one of the two possible clusters, based on feature similarity. The two clusters represent either the two forms of rotating the tool's shaft, i.e., knob rotation (cluster 1) or wrist rotation (cluster 2), or the two forms of grasping the tool, i.e., with middle finger (cluster 1) or index finger (cluster 2) as previously shown in Fig. 1.

To this end, we apply an unsupervised k-means clustering algorithm, where each sample is assumed to belong to one of k clusters or classes. The objective is to determine the cluster membership that minimizes the sum of the Euclidean distances between all samples assigned to a cluster and the centroid of that cluster, across all clusters.

The centroid of a cluster k_e, with $1 \leq e \leq K$, where K is the total number of clusters is

$$\boldsymbol{\eta}^e = \frac{1}{|k_e|} \sum_{\mathbf{G}^j \in k_e} \hat{\mathbf{G}}^j \tag{8}$$

where $|k_e|$ is the total number of samples assigned to cluster k_e. The clustering algorithm minimizes the sum of squared

Euclidean distances between the centroid of all clusters and each element in a cluster, that is:

$$\min \sum_{e=1}^{K} \sum_{\mathbf{G}^j \in k_e} \sum_{n,m} \left(\hat{G}^j_{n,m} - \eta^e_{n,m} \right)^2. \tag{9}$$

To evaluate the accuracy of the hand tracking method, we check whether all samples from a given class (e.g., hand rotation vs. knob rotation) fall within the same cluster.

C. Uniform Manifold Approximation and Projection

To visualize the samples, we reduce the dimensionality of the feature matrices $\hat{\mathbf{G}}^j$ to 2D, then plot the samples in a reduced two dimensional space. Dimensionality reduction is performed using Uniform Manifold Approximation and Projection (UMAP) [22].

Let ρ_j be the Euclidean distance between matrix $\hat{\mathbf{G}}^j$ and its closest neighbour $\hat{\mathbf{G}}^r$ in the original $3 \times d$ space, with $r \in [1, N]$, defined as:

$$\rho_j = D(\hat{\mathbf{G}}^j, \hat{\mathbf{G}}^r)$$

in which the operator $D(\mathbf{X}, \mathbf{Y})$ calculates the Euclidean distance between matrices $\mathbf{X}$ and $\mathbf{Y}$. Let σ be a normalization parameter for each feature matrix, chosen to satisfy:

$$\sum_{j,q \in [1,N] \ \forall j \neq q} e^{\left(\frac{\max\left[0, D\left(\hat{\mathbf{G}}^j, \hat{\mathbf{G}}^q\right)\right]}{\sigma} \right)} = \log_2(\psi) \tag{10}$$

where ψ is a hyperparameter defining the size of the local neighbourhood. A conditional edge weight, or similarity score, between matrices $\hat{\mathbf{G}}^j, \hat{\mathbf{G}}^q$ can be defined as:

$$w_{jq} = e^{\left(-\frac{\max\left[0, D\left(\hat{\mathbf{G}}^j, \hat{\mathbf{G}}^q\right) - \rho_j\right]}{\sigma} \right)}. \tag{11}$$

The asymmetric weights are then symmetrized to construct a fuzzy simplicial set representation of the local neighbourhood graph:

$$\tilde{w}_{jq} = w_{jq} + w_{qj} - w_{jq} w_{qj} = \tilde{w}_{qj}. \tag{12}$$

UMAP uses a heavy tailed, smooth kernel to define similarity between the higher dimensional space and lower dimensional space. It does so by minimizing a loss function that allows clusters in lower dimensional space to preserve their original structure in high dimensional space. The optimal dimension reduction is obtained by minimizing the cross-entropy between the high dimensional fuzzy graph $\tilde{w}_{rq}$ resulting in the following cost function:

$$\mathcal{C}_{\tilde{w}_{jq}, w_{jq}} = \sum_{j,q} \left[\tilde{w}_{jq} \ln \left(\frac{\tilde{w}_{jq}}{\phi_{jq}} \right) + (1 - \tilde{w}_{jq}) \ln \left(\frac{1 - \tilde{w}_{jq}}{1 - \phi_{jq}} \right) \right]. \tag{13}$$

where

$$\phi_{jq} = \left(1 + \alpha ||\mathbf{H}_j - \mathbf{H}_q||^{2\beta}\right)^{-1}, \tag{14}$$

defines the lower dimensional counter part of $\tilde{w}_{jq}$, in which the lower dimensional (2D) representation of $\hat{\mathbf{G}}^j$ is written as H_j. Here, α and β are fixed hyperparameters that control the shape of the lower dimensional distance decay.

The outer most points of the UMAP clusters were then connected by a dashed line using a convex hull method for easy visualization.

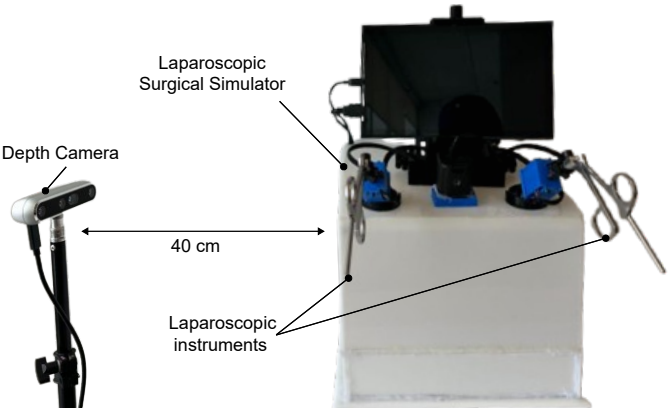

Fig. 3: Laparoscopic simulator and external RGB-Depth camera using during experimental validation.

III. Experimental Validation and Results

To validate the proposed algorithm, we use the instrumented laparoscopic box trainer shown in Fig. 3. The simulator consists of a box with two trocars through which two of the instruments shown in Fig. 1 are inserted. An Intel RealSense D435 colour-depth (RGB-D) camera is positioned on the side of the simulator, 40 cm from the side of the box trainer. The camera is mounted on a tripod so as to be level with the starting location of the laparoscopic tools and to capture a lateral view of the user's hand. Using the camera, we collect both RGB and depth images at a uniform sampling rate of 30 frames per second, so that each RGB image has a corresponding depth image taken at the same time. Both the RGB and depth images were captured with a resolution of 640×480 pixels. The camera captures a depth feed and RGB video feed of the surgeon's hands, and through the MediaPipe Hand Landmarker algorithm, it can track either the left or right hand.

The hand landmarker algorithm does not strictly require depth information, as the model has been trained to infer depth from an image. However, we use the depth information to ensure accurate depth tracking. The camera has onboard calibration software and calculates the distance from the depth camera to 21 landmarks on the hand in real time. The collected data is then passed into the UMAP dimensionality reduction and K-means clustering algorithms described in the previous section. Two experimental scenarios are evaluated. In each experimental setup, five subjects performed each of the four scenarios 10 times, resulting in 100 total trials for rotation methods and 100 trials for varied laparoscopic tool grips. The five subjects had varying hand sizes and skin tones, and data were collected throughout the workday under different lighting conditions.

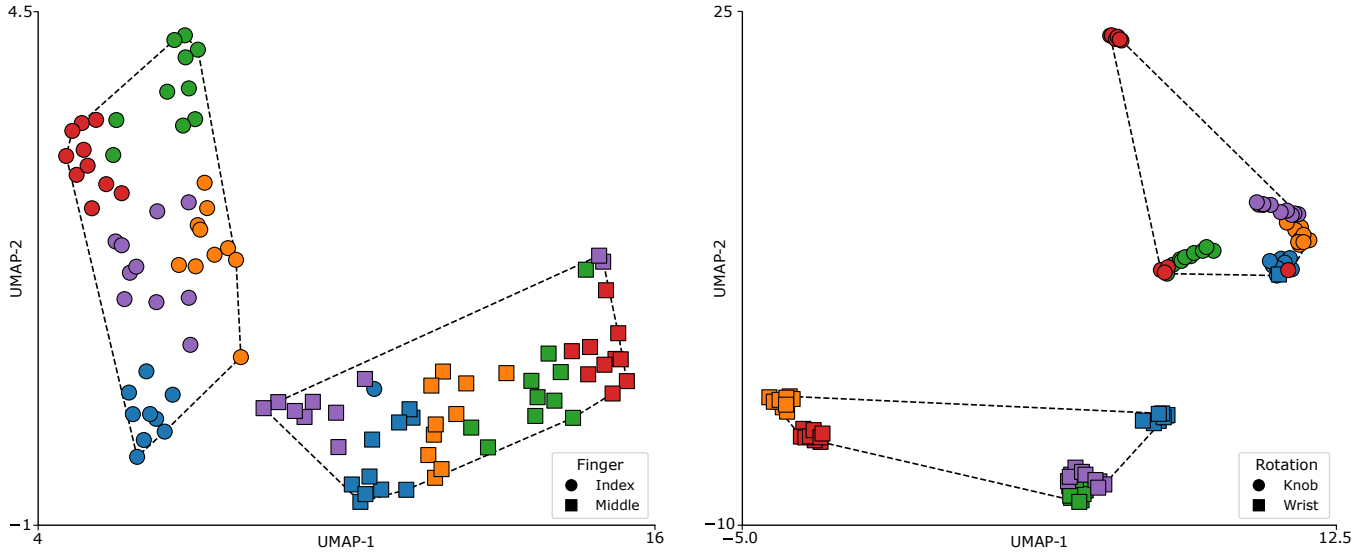

Fig. 4: UMAP visualization of hand landmark trajectories clustered using unsupervised k-means ($k = 2$). On the left, marker colour and shape denote the user and finger, respectively, used to manipulate the instrument handle. On the right, marker colour and shape denote the user and rotation methods, respectively, used to rotate the laparoscopic instrument by 90 degrees.

A. Scenario 1 - Tool Rotation Method Detection

In this scenario, each trainee was instructed to hold the laparoscopic instrument and rotate the tool shaft 90° clockwise and back to the start position three times, as follows:

- Ten trials of three repetitions each, where the tool shaft rotation is performed by rotating the gear on the laparoscopic tool, as shown in Fig. 1,
- Ten trials having three repetitions each, where the wrist is rotated rather than the knob, as shown in Fig. 1.

The unsupervised k-means clustering algorithm yields an accuracy of 99% (99 correct out of 100 total trials) in creating two clusters, across the five subjects, with a single trial from the wrist rotation set being placed in the knob rotation set. For visualization purposes, the dimensionality of all samples is reduced to 2D using UMAP, and the samples are coloured by the subject and the shape denotes movement type (ie gear vs wrist rotation, index vs middle find grip) cluster. Fig. 4 (right) shows the resulting UMAP clustering representation of the tool rotation trials. As can be seen, using the tool handle's gear to rotate the shaft is clearly distinguishable from using a wrist rotation.

B. Scenario 2 - Grip Detection

In the second scenario, hand landmark tracking is used to capture hand position data of each trainee pulling the tool back 4 cm along the surge axis, rotating clockwise by 90°, rotating back 90° counter-clockwise, and surging forward 4 cm back to the starting location. This movement is effected using two different variations of grip on the laparoscopic tool handle:

- Grasping the instrument finger holes on the handle with the thumb and **middle** finger;
- Grasping the instrument finger holes on the handle with the thumb and **index** finger.

Each movement variation is repeated a total of 10 times and lasts 20 seconds each. The collected data are then processed in the same way as in Scenario 1.

The results of the k-means clustering algorithm yield the same results as Scenario 1, with 99% accuracy (99 correct out of 100 total trials) in clustering the two variations of laparoscopic tool grips. Fig. 4 (right) illustrates the data after UMAP is used to reduce the dimensionality to 2D and samples are coloured according to their assigned clusters, which represent the two tool handle grip variations. As can be seen, grasping the tool with the thumb and middle finger is clearly distinguishable from grasping the tool with the thumb and index finger, even when performing identical movements.

Tracking the position of landmarks in space in real time using an external camera can provide additional feedback to the trainee on how to handle the instrument and position their hands for optimal dexterity and control. To verify the accuracy of these measurements, we compared the 3-DOF positions measured by the camera with those obtained from an electromagnetic tracking system. Using the Aurora tracking system from NDI, a sensor was placed on the user's hand at the same location as landmark zero (see Fig. 2a). The user was instructed to move their hand arbitrarily while holding the surgical instrument. The results shown in Fig. 5 demonstrate good agreement between the two measurement systems. The tracking error for each of the 3 tracked axes (x,y,z) is under a 3% symmetric mean absolute percentage error and a root mean square error of under 0.02m.

IV. Conclusion

To enhance traditional laparoscopic surgical training, cotemporary surgical trainers are often equipped with sensors that measure relevant data during task execution. The collected data

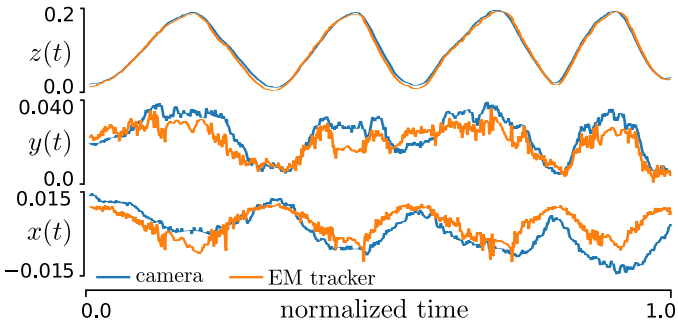

Fig. 5: Comparison between position tracking of landmark 0 in 3-DOF (x, y, z) using an electromagnetic tracking and MediaPipe during instrument manipulation. Displacement is in metres.

can then feed models that provide feedback and assessment without an expert surgeon having to be continuously present. Currently, instrumented box trainers capture tool movements from sensors or cameras located within the box, but are unable to assess how a trainee handles the surgical instrument. Yet, appropriate tool handling strongly affects dexterity, range of motion, and accuracy. To better assess how a trainee handles the surgical instrument, we demonstrate that an external depth camera and tracking algorithm, combined with k-means clustering, can accurately track the trainee's hand movements and differentiate between two common ways of grasping the tool and rotating the tool shaft.

The experimental evaluation of the proposed algorithm was carried out with five trainee's performing a series of pre-defined hand movements using a box simulator. The results showed that the clustering algorithm can differentiate/generalize to multiple subjects and trials accurately. To extend experimental validation to more realistic training conditions, future work will consider the complex and natural movement of the hand used during common training drills, such as object manipulation and tissue suturing [5], [23]. Additional clusters will have to be defined for additional conditions not currently captured in this study, such as idling and combined wrist and wheel rotation, for example.

Extending the proposed algorithm to work in real-time can enable trainees to receive real-time formative feedback on how to properly handle the surgical instrument, and learn proper techniques based on data collected from expert surgeons performing the same surgical tasks. While traditional training only considers tool tip trajectory, speed, and deviation for assessment, they neglect tool handling and posture, which are often attributed to errors in task execution. Correcting these issues can not only lead to better motor skill development, preventing the consolidation of suboptimal motion patterns, and improving dexterity during complex surgical tasks, but allow the thoughtful handling of instruments, which has been linked to decreased misuse injuries among surgeons.

References

[1] A. Johnson, "Laparoscopic surgery," *The Lancet*, vol. 349, no. 9052, pp. 631–635, 1997.

[2] B. Jönsson and N. Zethraeus, "Costs and benefits of laparoscopic surgery—a review of the literature," *European Journal of Surgery*, vol. 166, no. S12, pp. 48–56, 2000.

[3] Y. Tanagho *et al.*, "2D versus 3D visualization: Impact on laparoscopic proficiency using the fundamentals of laparoscopic surgery skill set," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 22, no. 9, pp. 865–870, 2012.

[4] S. Dreier, S. Mona, L. Konge, and F. Bjerrum, "Three-dimensional versus two-dimensional vision in laparoscopy: a systematic review," *Surgical Endoscopy*, vol. 30, no. 1, pp. 11–23, 2016.

[5] L. Villegas *et al.*, "Laparoscopic skills training," *Surgical Endoscopy And Interventional Techniques*, vol. 17, no. 12, pp. 1879–1888, 2003.

[6] G. Azzie *et al.*, "Development and validation of a pediatric laparoscopic surgery simulator," *Journal of Pediatric Surgery*, vol. 46, no. 5, pp. 897–903, 2011.

[7] J. Emken, E. Mcdougall, and R. Clayman, "Training and assessment of laparoscopic skills," *JSLS : Journal of the Society of Laparoendoscopic Surgeons*, vol. 8, pp. 195–9, 04 2004.

[8] M. Caban *et al.*, "Use of collapsible box trainer as a module for resident education," *JSLS*, vol. 17, no. 3, pp. 440–444, 2013.

[9] K. MacWilliams, J. Green, A. Nasr, G. Azzie, and C. Rossa, "Optimizing sensor selection in laparoscopic simulators: Lessons learned in a robotic platform," in *2025 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2025, pp. 1–6.

[10] S. Albasri *et al.*, "A novel distance for automated surgical skill evaluation," in *2019 7th E-Health and Bioengineering Conference*, 2019.

[11] J. Wong *et al.*, "Construction and validation of a low-cost laparoscopic simulator for surgical education," *Journal of Surgical Education*, vol. 70, no. 4, pp. 443–450, 2013.

[12] C. Heiliger *et al.*, "Tracking and evaluating motion skills in laparoscopy with inertial sensors," *Surgical Endoscopy*, vol. 37, no. 7, pp. 5274–5284, 2023.

[13] B. Gavrilovic *et al.*, "Development of an open-source laparoscopic simulator capable of motion and force assessment: High tech at low cost," *Journal of Laparoendoscopic & Adv. Surgical Techniques*, 2018.

[14] L. H. Olivas-Alanis *et al.*, "LAPKaans: Tool-motion tracking and gripping force-sensing modular smart laparoscopic training system," *Sensors*, vol. 20, 2020.

[15] V. Lahanas *et al.*, "Virtual reality-based assessment of basic laparoscopic skills using the leap motion controller," *Surgical Endoscopy*, vol. 31, no. 12, pp. 5012–5023, 2017.

[16] Laparosimulators, "Apexpro laparoscopic simulator," accessed: 2024-10-23. [Online]. Available: https://laparosimulators.com/apexpro/

[17] J. Rosen *et al.*, "The bluedragon - a system for measuring the kinematics and dynamics of minimally invasive surgical tools in-vivo," in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, vol. 2, 2002, pp. 1876–1881 vol.2.

[18] F. M. Sánchez-Margallo, J. A. Sánchez-Margallo, J. B. Pagador, J. L. Moyano, J. Moreno, and J. Usón, "Ergonomic assessment of hand movements in laparoscopic surgery using the cyberglove®," in *Computational biomechanics for medicine*. Springer, 2010, pp. 121–128.

[19] K. D. Tung, R. M. Shorti, E. C. Downey, D. S. Bloswick, and A. S. Merryweather, "The effect of ergonomic laparoscopic tool handle design on performance and efficiency," *Surgical endoscopy*, vol. 29, no. 9, pp. 2500–2505, 2015.

[20] F. Zhang *et al.*, "Mediapipe hands: On-device real-time hand tracking," *arXiv preprint arXiv:2006.10214*, 2020.

[21] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.-L. Chang, and M. Grundmann, "Mediapipe hands: On-device real-time hand tracking," 2020. [Online]. Available: https://arxiv.org/abs/2006.10214

[22] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," 2020. [Online]. Available: https://arxiv.org/abs/1802.03426

[23] D. M. Rooney, B. F. Santos, and E. S. Hungness, "Fundamentals of laparoscopic surgery (fls) manual skills assessment: Surgeon vs nonsurgeon raters," *Journal of Surgical Education*, vol. 69, no. 5, pp. 588–592, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1931720412001547