Standard Operating Procedures: Open Surgery Skill Assessment

Contents

Introduction

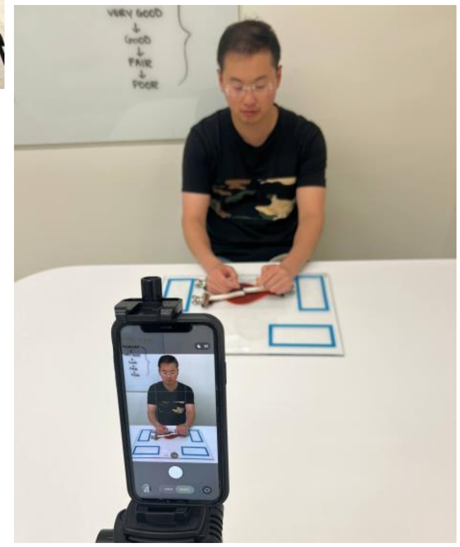
Purpose: To objectively evaluate fundamental surgical technical skills (economy of motion, spatial organization, and flow of motion) using an AI-based markerless computer vision framework.
What it does: The software utilizes a deep learning algorithm to track 21 hand and wrist joints from a standard 2D smartphone video. It translates hand kinematics into interpretable, cohort-normalized domain-level performance scores.
Why we have it: Traditional observational assessment is subjective and time-intensive. While wearable sensors provide objective data, they require complex hardware. This AI framework provides an accessible, unbiased assessment using standard cameras.

Reference: Zulbaran-Rojas, A., Rouzi, M. D., et al. (2026). AI-Based Markerless Computer Vision Framework for Open Surgery Skill Assessment: A Prototype Assessment Framework. Surgical Innovation.

Setup Instructions

Materials

- Smartphone
- Adjustable Tripod
- Surgical Board
- Pre-tied 2.5-foot #2-0 silk suture
- Table/Workstation for the surgical task
- Optional: Standard surgical non-blue gloves



Setup

1. Mount the pre-tied 2.5-foot silk suture onto the Knot Tying Board, ensuring the ends are symmetrical.
2. Place the Knot Tying Board on a flat, stable table at a comfortable height for the participant.
3. Set up the tripod in front of the table, facing where the participant will sit or stand.
4. Mount the smartphone on the tripod.
5. Adjust the camera. No predefined camera angle or specific distance is strictly required, provided both of the participant's hands remain completely visible within the frame during the entire task.

⚠️ Important Notes:

1. The participant's HEAD and BOTH HANDS (wrists to fingertips) MUST remain completely visible in the frame at all times! Ensure no obstructions block the camera. Only the participant should be in the video!
2. Ensure no objects block the line of sight between the smartphone and the participant's hands
3. Make sure only the participant is captured in the video. No other people should be in the video.

How to record a small video size suitable for faster AI-analysis?

We recommend applying these settings in the phone to capture smaller video sizes so analyzing them would not be time-consuming.

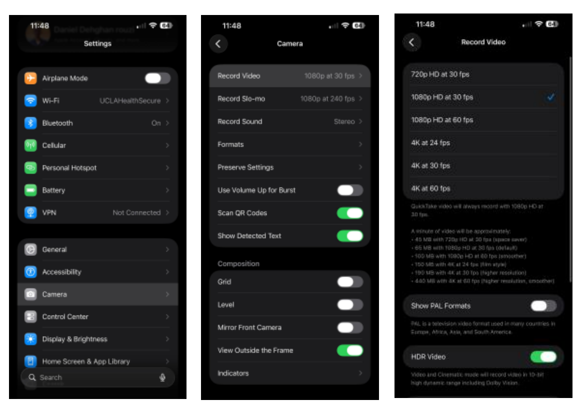
IPhone:

Go to settings > Camera > Record Video > Choose the option shown here:

Then go to Formats > Choose "High Efficiency" like below:

Android:

Go to Camera app > Settings > Video Resolution > Select "FHD / 1080p at 30 fps"



Analysis

According to the setup, record a video of the participant.

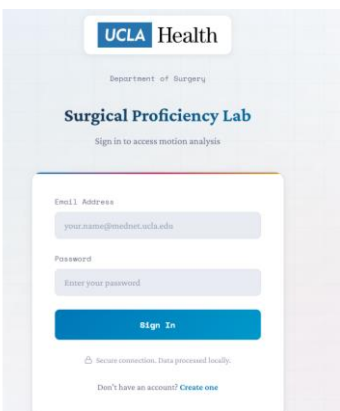
1.

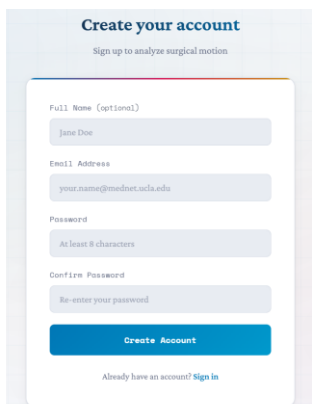
> ⚠️ Important Note:
>
> 1. In our paper the task was for one-handed knot tying. You may record any other surgery task and analyze your data using our AI model.
> 2. The scoring system is designed for 10 knots that we defined in our paper. If you want to do more than 10 knots, please ignore the scorings.

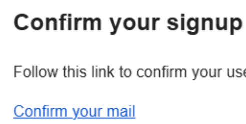
After recording the video, transfer the recorded video file (.mp4 or .mov) from the smartphone to the secure lab computer or designated encrypted drive.
2. Rename the file using the following convention: [ParticipantID]_KnotTying.mp4 (e.g., 001_KnotTying.mp4).
3. Store the video in the designated "Surgical_Proficiency_Tasks" folder.

AI-web Application

Scan this QR code or open this link: https://surgical-prof.blog/login
Account creation: If it's your first time visiting this website, you need to create an account.
Click on "Create one" > Enter your Name, Email, and Password > Confirm the link sent to your email. See below:

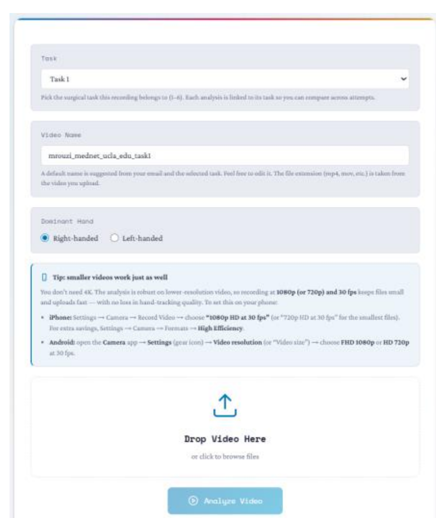




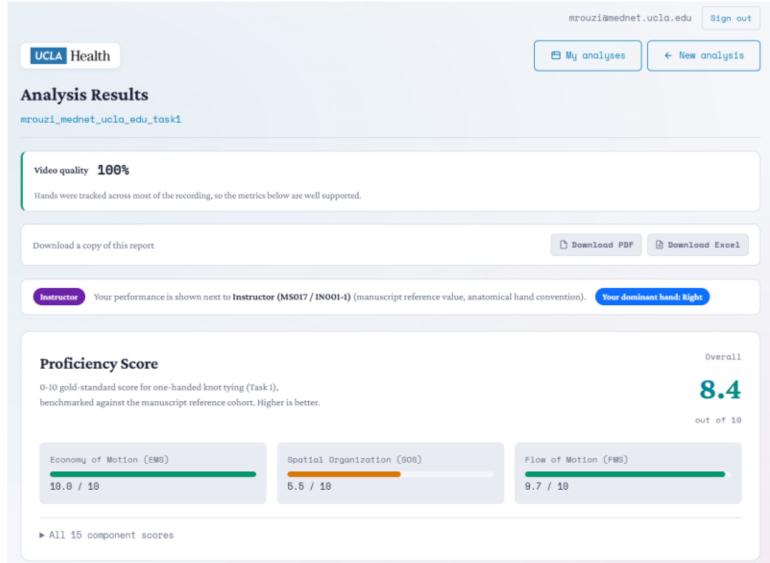


After login, you will see Analyze a new video and Previous analysis options.
Analyzing a new video: Now click on "Analyze a new video" and you will see this screen:

- Select the Task. (For now, "Task 1" is the knot hand tying task.)
- Then if the video name looks good to you keep it as it is. You can edit the name if you want to.
- Then select the participant's dominant hand.
- Then you can upload the video or Drag and Drop it in the box and click on "Analyze video"
- In this step, depending on the size of the video, you might wait anytime between few seconds to few minutes maximum. (5 seconds to 2 minute most of the time)



Results and Interpretations



Video quality

The video quality indicates how well you followed the protocols of the SOP. This number should be above 95%.

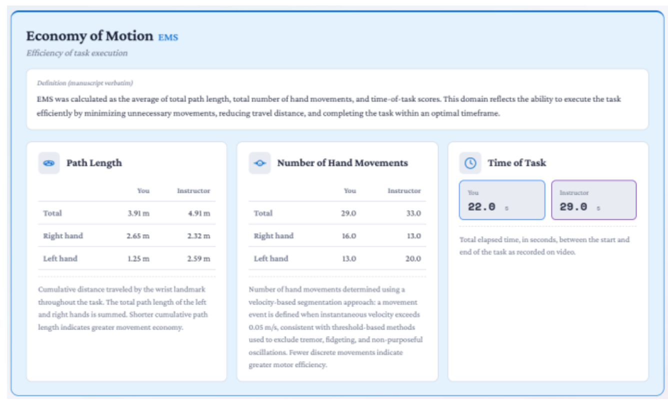
Download results:

Two download buttons below show that you can download results in EXCEL and PDF format. You can also download the GIF files and Hand Working Area charts by clicking on "Download" button.

Proficiency Score

Important note: This score is designed in our paper based on 10 one handed knot tying task. If your task is any other surgery task, please ignore the score values and you can interpret all other analysis in your results other than 0-10 scores.

In a case that you follow our 10 knots protocol in our paper, the scores are for the participant's performance for "Economy of Motion" (EMS), "Spatial Organization" (SOS), and "Flow of Motion" (FOM). On the side the average of all parameters score which is 8.4 in this example is shown. Scores are between 0 to 10. The higher score shows higher proficiency.



Economy of Motion Results:

EMS was calculated as the average of total path length, total number of hand movements, and time-of-task scores. This domain reflects the ability to execute the task efficiently by minimizing unnecessary movements, reducing travel distance, and completing the task within an optimal timeframe.

Path length: Cumulative distance traveled by the wrist landmark throughout the task. The total path length of the left and right hands is summed. Shorter cumulative path length indicates greater movement economy.

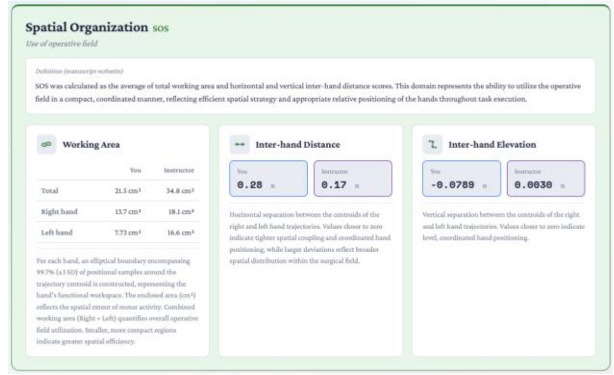
Number of Hand Movements: It's determined using a velocity-based segmentation approach: a movement event is defined when instantaneous velocity exceeds 0.05 m/s, consistent with threshold-based methods used to exclude tremor, fidgeting, and non-purposeful oscillations. Fewer discrete movements indicate greater motor efficiency.

Time of Task: Total elapsed time, in seconds, between the start and end of the task as recorded on video. Lower time, shows higher proficiency. In screen you can compare participant's performance compared to instructor.

Spatial Organization Results:

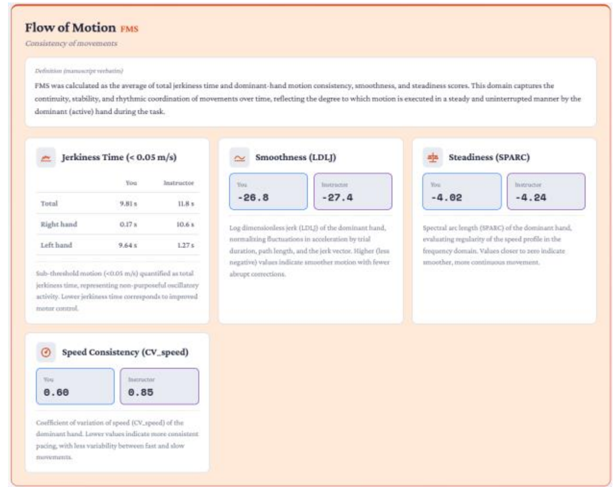
SOS was calculated as the average of total working area and horizontal and vertical inter-hand distance scores. This domain represents the ability to utilize the operative field in a compact, coordinated manner, reflecting efficient spatial strategy and appropriate relative positioning of the hands throughout task execution.

Working Area: For each hand, an elliptical boundary encompassing 99.7% (±3 SD) of positional samples around the trajectory centroid is constructed, representing the hand's functional workspace. The enclosed area (cm^2)



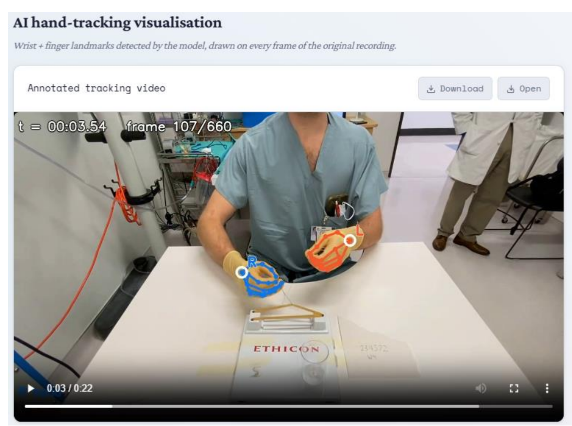
reflects the spatial extent of motor activity. Combined working area (Right + Left) quantifies overall operative field utilization. Smaller, more compact regions indicate greater spatial efficiency.
Inter-hand Distance: Horizontal separation between the centroids of the right and left hand trajectories. Values closer to zero indicate tighter spatial coupling and coordinated hand positioning, while larger deviations reflect broader spatial distribution within the surgical field. Lower distance shows higher proficiency.
Inter-hand Elevation: Vertical separation between the centroids of the right and left hand trajectories. Values closer to zero indicate level, coordinated hand positioning. In screen you can compare participant's performance compared to instructor.

Flow of Motion Results:



FMS was calculated as the average of total jerkiness time and dominant-hand motion consistency, smoothness, and steadiness scores. This domain captures the continuity, stability, and rhythmic coordination of movements over time, reflecting the degree to which motion is executed in a steady and uninterrupted manner by the dominant (active) hand during the task.
Jerkiness Time (< 0.05 m/s): Sub-threshold motion (<0.05 m/s) quantified as total jerkiness time, representing non-purposeful oscillatory activity. Lower jerkiness time corresponds to improved motor control.
Smoothness (LDLJ): Log dimensionless jerk (LDLJ) of the dominant hand, normalizing fluctuations in acceleration by trial duration, path length, and the jerk vector. Higher (less negative) values indicate smoother motion with fewer abrupt corrections.
Steadiness (SPARC): Spectral arc length (SPARC) of the dominant hand, evaluating regularity of the speed profile in the frequency domain. Values closer to zero indicate smoother, more continuous movement.
Speed Consistency (CV_speed): Coefficient of variation of speed (CV_speed) of the dominant hand. Lower values indicate more consistent pacing, with less variability between fast and slow movements.
In screen you can compare participant's performance compared to instructor.

AI tracking visualization:

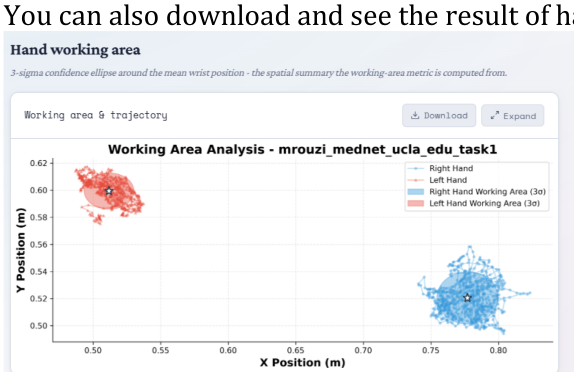


GIF file

You can download the GIF file of the AI tracking the hands in this section:

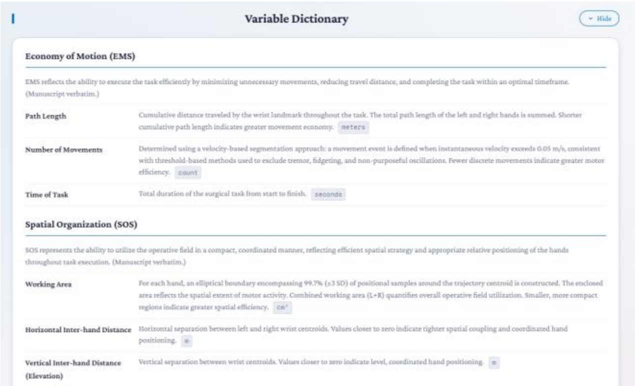
Hand Working Area

You can also download and see the result of hand working area:



Variable dictionary

In this section all variables definition and metrics has been mentioned:



Variable Dictionary

Hide

Economy of Motion (EMS)

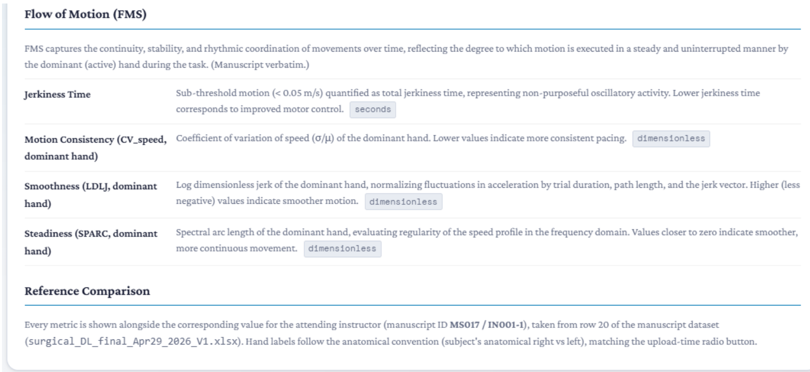
EMS reflects the ability to execute the task efficiently by minimizing unnecessary movements, reducing travel distance, and completing the task within an optimal timeframe. (Manuscript verbatim.)

Path Length	Cumulative distance traveled by the wrist landmark throughout the task. The total path length of the left and right hands is summed. Shorter cumulative path length indicates greater movement economy. `meters`
Number of Movements	Determined using a velocity-based segmentation approach: a movement event is defined when instantaneous velocity exceeds 0.05 m/s, consistent with threshold-based methods used to exclude tremor, fidgeting, and non-purposeful oscillations. Fewer discrete movements indicate greater motor efficiency. `count`
Time of Task	Total duration of the surgical task from start to finish. `seconds`

Spatial Organization (SOS)

SOS represents the ability to utilize the operative field in a compact, coordinated manner, reflecting efficient spatial strategy and appropriate relative positioning of the hands throughout task execution. (Manuscript verbatim.)

Working Area	For each hand, an elliptical boundary encompassing 99.7% (±3 SD) of positional samples around the trajectory centroid is constructed. The enclosed area reflects the spatial extent of motor activity. Combined working area (L+R) quantifies overall operative field utilization. Smaller, more compact regions indicate greater spatial efficiency. `cm²`
Horizontal Inter-hand Distance	Horizontal separation between left and right wrist centroids. Values closer to zero indicate tighter spatial coupling and coordinated hand positioning. `m`
Vertical Inter-hand Distance (Elevation)	Vertical separation between wrist centroids. Values closer to zero indicate level, coordinated hand positioning. `m`



Flow of Motion (FMS)

FMS captures the continuity, stability, and rhythmic coordination of movements over time, reflecting the degree to which motion is executed in a steady and uninterrupted manner by the dominant (active) hand during the task. (Manuscript verbatim.)

Jerkiness Time	Sub-threshold motion (< 0.05 m/s) quantified as total jerkiness time, representing non-purposeful oscillatory activity. Lower jerkiness time corresponds to improved motor control. `seconds`
Motion Consistency (CV_speed, dominant hand)	Coefficient of variation of speed (σ/μ) of the dominant hand. Lower values indicate more consistent pacing. `dimensionless`
Smoothness (LDLJ, dominant hand)	Log dimensionless jerk of the dominant hand, normalizing fluctuations in acceleration by trial duration, path length, and the jerk vector. Higher (less negative) values indicate smoother motion. `dimensionless`
Steadiness (SPARC, dominant hand)	Spectral arc length of the dominant hand, evaluating regularity of the speed profile in the frequency domain. Values closer to zero indicate smoother, more continuous movement. `dimensionless`

Reference Comparison

Every metric is shown alongside the corresponding value for the attending instructor (manuscript ID **MS017 / IN001-1**), taken from row 20 of the manuscript dataset (`surgical_DL_final_Apr29_2026_V1.xlsx`). Hand labels follow the anatomical convention (subject's anatomical right vs left), matching the upload-time radio button.

Reference for variables

Zulbaran-Rojas, A., Rouzi, M. D., Hansraj, N., Erstad, D., Bargas-Ochoa, M., D'Silva, E., ... & Najafi, B. (2026). AI-Based Markerless Computer Vision Framework for Open Surgery Skill Assessment: A Prototype Assessment Framework. Surgical Innovation, 15533506261451990.
Link: https://journals.sagepub.com/doi/full/10.1177/15533506261451990

Troubleshooting

- Hands leave the camera frame: If the participant moves their hands completely out of the camera's view during the task, the deep learning algorithm will fail to track the 21 joint landmarks. Say "Stop," halt the recording, ask the participant to reset, adjust the camera distance if necessary to widen the field of view, and restart the trial.
- Environmental noise/Lighting: While the algorithm is robust to lighting and background changes, attempt to perform the task in a well-lit room to minimize blur during fast hand movements.
- No one except the participant should be in the video. If there are other people in the video, this can impact the video quality percentage after analysis.
- If the quality of the video after analysis is below 95%, please review the SOP again and follow instructions to capture a high-quality video for the analysis.
- If your video is big size and you want to reduce its size, our AI model automatically does that for you when you upload the video. However, if you want to reduce the video size on your own, make sure the dimensions of the video are not distorted. This can be one of the reasons that the quality of your video might be low after analysis.