SURGICAL TEAM PERFORMANCE ANALYSIS USING COMPUTER VISION: A METHODOLOGY and USE CASE STUDY

Mahdi Ebnali, Harvard Medical School, Department of Emergency Medicine,
Mass General Brigham (mebnali-heidari@bwh.harvard.edu)
Marco A. Zenati, Harvard Medical School, VA Boston Healthcare System
Matthew C. Gombolay, Georgia Institute of Technology
Roger D. Dias, Harvard Medical School, Department of Emergency Medicine, Mass General Brigham

Current methods for assessing the performance of surgical teams, such as observational rating scales and team self-assessments, have significant limitations. Existing methods primarily focus on subjective evaluations of surgical team performance or use an unstructured and descriptive method, resulting in inaccurate, biased evaluations. A computer vision (CV) -based method has been proposed in this study to overcome these challenges. This technique is based on estimating the position and movement of OR staff during procedures using video data captured by standard cameras to assess the performance of Operation Room (OR) staff. We examined the feasibility of this approach by extracting OR team motion metrics from 30 videos from real-life cardiac surgery operations. Further studies are needed to provide more validation of the efficiency of this method to quantify the performance of surgical teams.

INTRODUCTION:

Surgical procedures take place in high-risk and complex environments, where multiple specialized professionals work together to deliver effective care to patients (Roger D. Dias et al. 2021). Therefore, teamwork is critical for the successful performance of surgical teams during open, minimally invasive, and robotic procedures. Team performance during a surgical procedure varies depending on the team's work experience and cognitive demand on individual operating room (OR) personnel (Avrunin et al. 2018) Understanding team performance, therefore, is essential for establishing a more systematic mechanism for improving patient outcomes (Molina and Heng 2009; Dearani et al. 2019; Roger D. Dias, Shah, and Zenati 2020; R. D. Dias et al. 2022; Ebnali et al. 2022).

Assessment of surgical competency at individual and team levels has become a priority for both surgical educators and licensing boards (Liu et al. 2021). Methods currently used to assess team performance in training laboratories and operating rooms, such as observational rating scales and team self-assessment, suffer from significant shortcomings. Most existing methods focus on subjectively evaluating performance or utilizing an unstructured and descriptive approach, leading to potentially biased and inaccurate results.

Several promising quantitative methods have also been examined in recent studies to evaluate surgical team performance, including gesture detection using special cameras (e.g., the Kinect)(Kim et al. 2014), body markers (Uslu et al. 2021), wearable device sensors(Zulbaran-Rojas et al. 2021), and instrument tracking(Bouarfa et al. 2012). These technologies are, however, difficult to implement in OR environments because of privacy and patient safety concerns (Varadarajan et al. 2009).

To address these challenges, our study proposes a computer vision (CV) based method for analyzing the performance of OR staff during procedures by estimating the position and movement of their bodies as captured in video data without expert monitoring or using specialized technologies. The effectiveness of CV techniques in evaluating human performance has been demonstrated through a number of applications, both within and outside healthcare. More specifically, CV-based techniques have been investigated in the assessment of technical (Wang and Majewicz Fey 2018; Jin et al. 2018; Louis et al. 2022) and non-technical (Kennedy-Metz et al. 2021; Roger D. Dias, Shah, and Zenati 2020; Roger D. Dias et al. 2022) surgical skills. Previous studies also used these methods to inform surgical teams with quantitative and actionable feedback, and enhance machines' capabilities by acquiring transportable teamwork competencies (Mumtaz et al. 2022) to create better coordination with the human and machine counterparts in a team and allowing other team members to determine when and how best to communicate with teammates, further enhancing ability and trust in human-machines (Carroll et al. 2019).

METHOD:

The proposed method is based on the OpenPose library (*Openpose: OpenPose: Real-Time Multi-Person Keypoint Detection Library for Body, Face, Hands, and Foot Estimation* n.d.) and video data captured by standard cameras that are available in OR environments without needing the attachment of makers or sensors to the OR staff. OpenPose is a posture-tracking algorithm employed to estimate the posture (keypoints) of multiple people from the images captured by a monocular camera using a multi-stage convolutional neural network (CNN), to predict confidence 2D maps of body part locations, and estimate Part Affinity Fields (PAF) which encode the degree of association between body parts.

Training and validation of the OpenPose algorithms were evaluated on two benchmarks for multi-person pose estimation: the MPII human multi-person dataset and the COCO 2022 keypoints challenge dataset. Both datasets had images collected from diverse real-life scenarios, such as crowding, scale variation, occlusion, and contact.

The following is a summary of the steps involved in using this method:

Step 1: Position keypoints generation from recorded OR videos:

As shown in Figure 2, after recording the video during the OR procedure, the OpenPose library can be used to identify the position of keypoints (17 keypoints, Figure 1) in the OR staff's body. A human pose skeleton denotes the orientation of an individual in a particular format. Fundamentally, it is a set of data points that can be connected to describe an individual's pose. Each data point in the skeleton can also be called a part or coordinate, or keypoint. A relevant connection between two coordinates is known as a limb or pair.

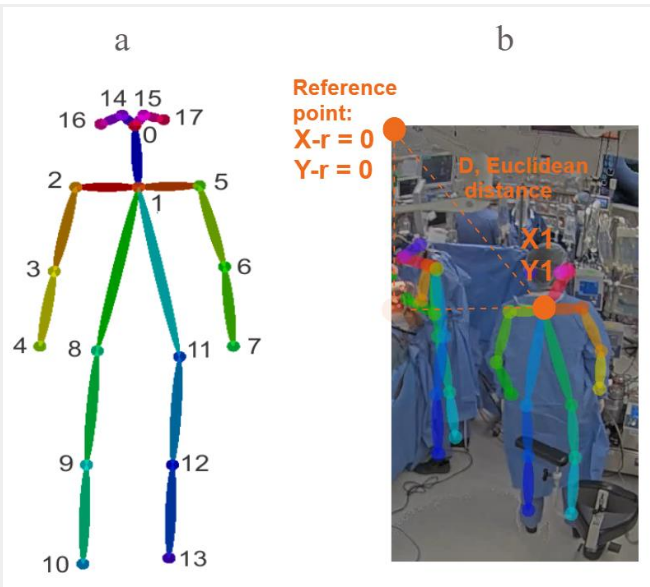

Figure 1. (a) Keypoints that can be detected by the OpenPose algorithm on OR staffs, (b) example of how Euclidean distance is used to measure the displacement of each individual of OR staff.

Step 2: Position keypoint extraction:

OpenPose creates JSON files for each frame containing all keypoint positions, which can be merged to produce a single file for all frames and persons. Aggregating data from all frames creates a larger file containing the x and y locations of keypoints of each OR staff in each time frame.

Step 3: Pre-processing position data:

According to our previous studies (Roger D. Dias et al. 2022), we suggest using only the neck keypoint (keypoint 1) during the pre-processing stage since changing the coordination of this keypoint resulted in higher accuracy in capturing the whole body movements of a person. For more specific aspects of team movements, other keypoints can also be used individually or collectively. After dealing with outliers and missing data, the x and y coordinates of the neck keypoint are suggested to calculate the Euclidean distance between the neck and a reference point ($x = 0$, $y = 0$) (Figure 1). The average displacement per frame in pixels (px) then can be calculated across all team members to capture the entire team motion.

Step 4: Team displacement data analysis:

Team displacement data measured in Step 3 can be analyzed using time-domain methods such as mean, median, standard deviation (SD), max, and min of team displacement data. This data can also be transferred to the frequency domain for further analysis using power spectral density to identify peak and frequency bands' statistics. Our previous studies have also shown that nonlinear methods such as Shannon entropy can be useful in evaluating team performance based on motion and physiological data(Roger D. Dias et al. 2021).

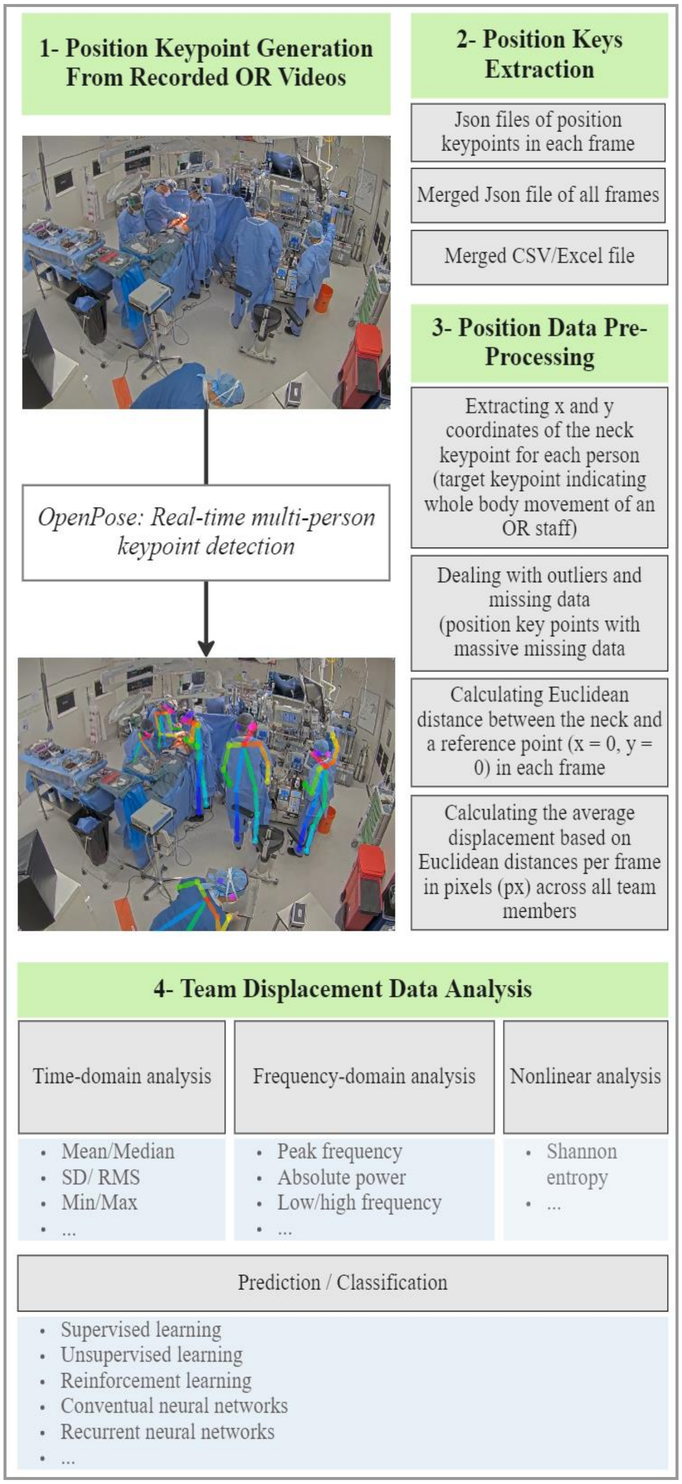

Figure 2. A computer vision (CV) based method for analyzing the performance of OR staff during procedures

USE CASE STUDY:

We examined the feasibility of this approach by extracting OR team motion metrics from 30 videos from real-life cardiac surgery operations. Afterward, we explored the relationship between these metrics and non-technical skills (NTS). An NTS assessment tool was used by 3 trained raters to assess these 30 videos. Raters used a Likert scale to rate 16 behavioral skills in the following categories: situational awareness, leadership, communication and teamwork, and decision-making. All ratings were averaged to derive an overall NTS score. Teams with ratings within the first quartile were classified as "Poor NTS", and those teams within the fourth quartile were classified as "Good NTS".

The metrics related to power spectral density and entropy were calculated to estimate the level of team organization and identify variations that may arise from a high cognitive load, medical errors, emergencies, or flow disruptions. We found that teams with non-technical scores exhibited significantly lower entropy in motion patterns, reflecting more consistent body motion during a critical cardiac surgery phase. Our study demonstrated that this methodology can be used as a feasible approach to understanding team performance based on metrics extracted from surgical video data using position estimation of OR personals with the OpenPose library.

CONCLUSION:

This study proposed a novel approach to analyze dynamic changes in OR teams by extracting and analyzing the position of OR staff's body in video data recorded by standard cameras during OR procedures. We also evaluated the feasibility of this method in real cardiac surgical data in 30 videos from real-life cardiac surgery operations. The findings suggested the proposed method could be a useful approach to characterize dynamic changes in OR teams, helping to better understand team performance with an automated approach. Further studies are needed to provide more validation of the efficiency of this method to quantify the performance of surgical teams.

ACKNOWLEDGMENT:

This work was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health [R01HL126896, R01HL157457]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank the clinical staff from the Veterans Affairs (VA) Boston Healthcare System for their valuable insights and collaboration in this study.

REFERENCES:

Avrunin, George S., Lori A. Clarke, Heather M. Conboy, Leon J. Osterweil, Roger D. Dias, Steven J. Yule, Julian M. Goldman, and Marco A. Zenati. 2018. "Toward Improving Surgical Outcomes by Incorporating Cognitive Load Measurement into Process-Driven Guidance." *Software Engineering in Healthcare Systems (SEHS), IEEE/ACM International Workshop on* 2018 (May): 2–9.

Bouarfa, Loubna, Oytun Akman, Armin Schneider, Pieter P. Jonker, and Jenny Dankelman. 2012. "In-Vivo Real-Time Tracking of Surgical Instruments in Endoscopic Video." *Minimally Invasive Therapy & Allied Technologies: MITAT: Official Journal of the*

Society for Minimally Invasive Therapy 21 (3): 129–34.
Carroll, Micah, Rohin Shah, Mark K. Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. 2019. “On the Utility of Learning about Humans for Human-Ai Coordination.” *Advances in Neural Information Processing Systems* 32. https://proceedings.neurips.cc/paper/2019/hash/f5b1b89d98b7286673128a5fb112cb9a-Abstract.html.
Dearani, Joseph A., Todd K. Rosengart, M. Blair Marshall, Michael J. Mack, David R. Jones, Richard L. Prager, and Robert J. Cerfolio. 2019. “Incorporating Innovation and New Technology Into Cardiothoracic Surgery.” *The Annals of Thoracic Surgery* 107 (4): 1267–74.
Dias, R. D., M. A. Zenati, G. Rance, Rithy Srey, D. Arney, L. Chen, R. Paleja, L. R. Kennedy-Metz, and M. Gombolay. 2022. “Using Machine Learning to Predict Perfusionists’ Critical Decision-Making during Cardiac Surgery.” *Computer Methods in Biomechanics and Biomedical Engineering. Imaging & Visualization* 10 (3): 308–12.
Dias, Roger D., Lauren R. Kennedy-Metz, Steven J. Yule, Matthew Gombolay, and Marco A. Zenati. 2022. “Assessing Team Situational Awareness in the Operating Room via Computer Vision.” *IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)* 2022 (June): 94–96.
Dias, Roger D., Julie A. Shah, and Marco A. Zenati. 2020. “Artificial Intelligence in Cardiothoracic Surgery.” *Minerva Cardioangiologica* 68 (5): 532–38.
Dias, Roger D., Steven J. Yule, Lauren Kennedy-Metz, and Marco A. Zenati. n.d. “Psychophysiological Data and Computer Vision to Assess Cognitive Load and Team Dynamics in Cardiac Surgery.” Accessed January 19, 2023. https://ipcai2019.github.io/pdf/Long_Abstracts/Diaz_etal_IPCAI_2019_Long_Abstract.pdf.
Dias, Roger D., Marco A. Zenati, Heather M. Conboy, Lori A. Clarke, Leon J. Osterweil, George S. Avrunin, and Steven J. Yule. 2021. “Dissecting Cardiac Surgery: A Video-Based Recall Protocol to Elucidate Team Cognitive Processes in the Operating Room.” *Annals of Surgery* 274 (2): e181–86.
Ebnali, Mahdi, Lauren R. Kennedy-Metz, Heather M. Conboy, Lori A. Clarke, Leon J. Osterweil, George Avrunin, Christian Miccile, et al. 2022. “A Coding Framework for Usability Evaluation of Digital Health Technologies.” In *Human-Computer Interaction. Theoretical Approaches and Design Methods*, 185–96. Springer International Publishing.
Jin, Amy, Serena Yeung, Jeffrey Jopling, Jonathan Krause, Dan Azagury, Arnold Milstein, and Li Fei-Fei. 2018. “Tool Detection and Operative Skill Assessment in Surgical Videos Using Region-Based Convolutional Neural Networks.” *arXiv [cs.CV]*. arXiv. http://arxiv.org/abs/1802.08774.
Kennedy-Metz, Lauren R., Pietro Mascagni, Antonio Torralba, Roger D. Dias, Pietro Perona, Julie A. Shah, Nicolas Padoy, and Marco A. Zenati. 2021. “Computer Vision in the Operating Room: Opportunities and Caveats.” *IEEE Transactions on Medical Robotics and Bionics* 3 (1): 2–10.
Kim, Yonjae, Simon Leonard, Azad Shademan, Axel Krieger, and Peter C. W. Kim. 2014. “Kinect Technology for Hand Tracking Control of Surgical Robots: Technical and Surgical Skill Comparison to Current Robotic Masters.” *Surgical Endoscopy* 28 (6): 1993–2000.
Liu, Daochang, Qiyue Li, Tingting Jiang, Yizhou Wang, Rulin Miao, Fei Shan, and Ziyu Li. 2021. “Towards Unified Surgical Skill Assessment.” In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9522–31. IEEE.
Louis, Nathan, Luowei Zhou, Steven J. Yule, Roger D. Dias, Milisa Manojlovich, Francis D. Pagani, Donald S. Likosky, and Jason J. Corso. 2022. “Temporally Guided Articulated Hand Pose Tracking in Surgical Videos.” *International Journal of Computer Assisted Radiology and Surgery*, October. https://doi.org/10.1007/s11548-022-02761-6.
Molina, J. A., and Bee Hoon Heng. 2009. “Global Trends in Cardiology and Cardiothoracic Surgery-an Opportunity or a Threat.” *Annals of the Academy of Medicine, Singapore* 38 (6): 541–45.
Mumtaz, Hassan, Muhammad Saqib, Farrukh Ansar, Durafshan Zargar, Madiha Hameed, Mohammad Hasan, and Pakiza Muskan. 2022. “The Future of Cardiothoracic Surgery in Artificial Intelligence.” *Annals of Medicine and Surgery (2012)* 80 (August): 104251.
Openpose: OpenPose: Real-Time Multi-Person Keypoint Detection Library for Body, Face, Hands, and Foot Estimation. n.d. Github. Accessed January 8, 2023. https://github.com/CMU-Perceptual-Computing-Lab/openpose.
Uslu, Tuğrul, Erkin Gezgin, Seda Özbek, Didem Güzin, Fatih Cemal Can, and Levent Çetin. 2021. “Utilization of Low Cost Motion Capture Cameras for Virtual Navigation Procedures: Performance Evaluation for Surgical Navigation.” *Measurement* 181 (August): 109624.
Varadarajan, Balakrishnan, Carol Reiley, Henry Lin, Sanjeev Khudanpur, and Gregory Hager. 2009. “Data-Derived Models for Segmentation with Application to Surgical Assessment and Training.” *Medical Image Computing and Computer-Assisted Intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 12 (Pt 1): 426–34.
Wang, Ziheng, and Ann Majewicz Fey. 2018. “Deep Learning with Convolutional Neural Network for Objective Skill Evaluation in Robot-Assisted Surgery.” *International Journal of Computer Assisted Radiology and Surgery* 13 (12): 1959–70.
Zulbaran-Rojas, Alejandro, Bijan Najafi, Nestor Arita, Hadi Rahemi, Javad Razjouyan, and Ramyar Gilani. 2021. “Utilization of Flexible-Wearable Sensors to Describe the Kinematics of Surgical Proficiency.” *The Journal of Surgical Research* 262 (June): 149–58.