Positioning a Stapler in 3D Space: Metric Generation for Robotic Surgery

Anand Srinivasan
anand.srinivasan@yale.edu

Advisor: Daniel Rakita
daniel.rakita@yale.edu

A Senior Thesis as a partial fulfillment of requirements for the Bachelor of Science in Computer Science

Department of Computer Science
Yale University
May 1, 2025

Acknowledgements

I want to express my gratitude to my advisor Dr. Daniel Rakita and my two PhD student mentors Xiatao Sun and TJ Vitchutripop for their support and guidance throughout this project. Additional thanks to Dr. Jay Bader and the RoboSurgery team.

Contents

Positioning a Stapler in 3D Space: Metric Generation for Robotic Surgery

Anand Srinivasan

Abstract

Minimally invasive surgery using robotic instruments has seen widespread expansion since its introduction in the 1980s. These procedures require precise positioning of surgical tools and careful manipulation of tissue. Laparoscopic procedures, however, lack quantitative tools to assess technique. Measuring angles that capture the three-dimensional (3D) orientations of relevant surgical structures could support the development of robust metrics. Such measurements may help identify optimal angle ranges during critical steps of laparoscopic procedures, facilitating training and standardization. To address this need, we present a semi-automated pipeline for estimating 3D angles from monocular surgical video. The method integrates image segmentation, zero-shot point cloud estimation, and principal component analysis to compute directional vectors for key features. Angle measurements are derived from these vectors relative to visual axes. To validate the pipeline, we generated an in vitro dataset by manipulating a surgical stapler in an open-air setting and using a stereo vision system to obtain ground truth depth angles. The pipeline achieved a mean absolute error of 7.67° across 670 frames, showing promising performance in inferring depth angles. We apply our tool to analyze eight surgical videos of Inferior Mesenteric Artery (IMA) division, a key maneuver in many colorectal procedures for the removal of diseased tissue. Specifically, we measured angles between the IMA, a surgical stapler device, and the depth axis. The pipeline revealed considerable variation in stapler and IMA angles, indicating a lack of standardization. This work lays the foundation for quantitative assessment of robotic IMA division and other laparoscopic surgeries. Future work should focus on improving zero-shot 3D point cloud estimation in laparoscopic settings to further enhance accuracy.

1 Introduction

In laparoscopic procedures, surgeons operate through small incisions using long, articulated instruments, often with the assistance of robotic platforms such as the Da Vinci system. These procedures require precise positioning of surgical tools and careful manipulation of tissue within a constrained visual field [1, 2]. Despite the high precision requirements, laparoscopic surgery currently lacks quantitative tools to assess and inform technique. In particular, no standard metrics exist to evaluate the relative spatial orientation of surgical instruments and anatomical regions of interest (ROIs) during critical steps. Such metrics could support training, improve standardization across institutions, and enhance intraoperative decision-making.

Inferior mesenteric artery (IMA) division is one such step that exemplifies the challenges and opportunities for quantitative assessment. Commonly performed in colorectal cancer resections to facilitate lymph node removal, IMA division involves the precise application of a surgical stapler to transect the artery near its origin from the aorta [3, 4, 5]. Surgeons typically perform this step laparoscopically using a Da Vinci stapler, a high-precision device equipped with real-time feedback and articulated control. To avoid complications such as tissue ischemia or postoperative anastomotic leakage, the stapler must be carefully positioned relative to the IMA and nearby vessels [6]. However, the absence of standard metrics for stapler orientation means that decisions are often guided by subjective judgment. Even with best practices, anastomotic leakage occurs in 6–10% of patients [7, 8], underscoring the need for robust surgical evaluation tools.

As an initial step toward developing such tools, the Applied Planning, Learning, and Optimization (APOLLO) Lab previously created a semi-automated pipeline for two-dimensional (2D) angle analysis from laparoscopic video. Early findings using this tool revealed consistent patterns in stapler positioning relative to the IMA and image axes, suggesting that 2D angular metrics may help define optimal instrument orientations. However, 2D analyses omit depth information, an essential component of true 3D pose estimation.

To address this limitation, we developed a semi-automated 3D angle estimation pipeline for measuring angles involving relevant surgical instruments and anatomical structures using monocular laparoscopic video. The pipeline consists of three main steps: segmentation of ROIs using the Segment Anything Model 2 (SAM 2) [9], zero-shot 3D projection using the Monocular Geometry Estimation (MoGe) model [10], and principal component analysis (PCA) to derive directional vectors from 3D projections. Angles are then calculated between these vectors or with respect to camera-based visual axes.

To validate the accuracy of this pipeline, a stereo vision camera setup was built and verified to produce reliable angle measurements. An in vitro dataset was subsequently created by manipulating a DaVinci stapler across a range of clinically relevant depth an-

gles in an open-air setting, with ground-truth angle labels obtained by the stereo setup. The 3D angle estimation pipeline was then tested using monocular input from one of the stereo cameras. Across 670 frames, the pipeline achieved a mean absolute error of 7.67°, demonstrating that accurate depth angle estimation is feasible with monocular video input.

We applied our system to eight real-world surgeries involving IMA division, measuring angles between the IMA, stapler, and depth axis at the moment of division. Our analysis revealed considerable variation in 3D stapler and IMA angles across cases, suggesting a lack of standardization in instrument positioning during this critical step. These findings highlight the potential for angle-based metrics to guide and standardize surgical technique. More broadly, this work lays the foundation for angular and positional 3D analysis, applicable to a range of minimally invasive procedures. Applied retrospectively to surgical video data at scale, this tool could help identify positional metrics across a range of laparoscopic surgeries, allowing for better informed surgical decision making and eventually leading to improved clinical outcomes.

2 Methodology

2.1 3D Angle Estimation Pipeline Overview

A 3D angle estimation pipeline is proposed. It consists of three major parts. First, ROIs are segmented. Next, segmented regions are projected into 3D space using a zero-shot 3D geometry estimation model. Finally, PCA is applied to the 3D projections to determine directions vectors for the ROIs, from which angle measurements can be determined. An overview of the pipeline applied to an IMA division surgery is shown in Figure 1. All code for the pipeline is available at https://github.com/SrinivasanAnand/angle_estimation_3D.

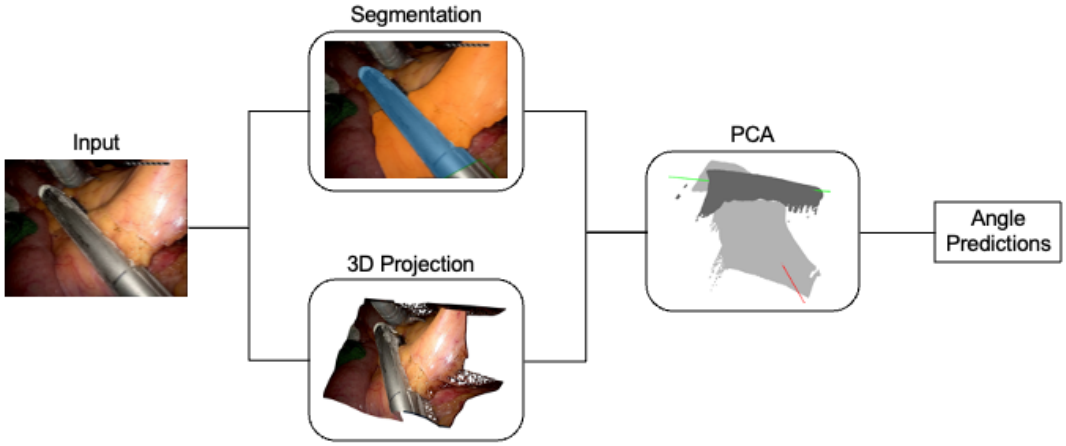

Fig. 1: Overview of 3D Angle Estimation Pipeline. The stapler and IMA are segmented from the input frame using the Segment Anything Model 2 (SAM 2) and projected into 3D space via a zero-shot prediction from the MoGe model. Principal component analysis (PCA) is then applied to the segmented, 3D-projected points to determine the primary axes of the stapler and IMA. These axes serve as directional vectors for subsequent 3D angle calculations.

2.1.1 Segmentation

ROIs were segmented using the Segment Anything Model 2 (SAM 2). SAM 2 is a segmentation model created by Meta designed to generalize zero-shot to unfamiliar settings. A graphical user interface (GUI) was designed using the PyQt6 Python library to allow a user to interactively select segmented regions corresponding to each ROI. Exclusion boxes can be drawn to omit regions of the segmentation that might hinder analysis. After

selecting segmentations for ROIs in the first frame of a surgical video, these segmentations were propagated automatically to all subsequent frames, resulting in an image mask corresponding to each ROI for each frame.

2.1.2 Point Cloud Estimation

Two depth estimation models and one monocular 3D geometry estimation model were tested in their ability to recover point clouds from monocular surgical video frames. Depth Anything is a general-purpose depth estimation model trained on over 63 million images [cite]. EndoDAC is a self-supervised depth estimation framework meant to adapt DepthAnything to endoscopic scenes. Finally, MoGe is a general-purpose 3D geometry estimation model which estimates depth, the camera intrinsic matrix, and 3D point clouds from a monocular image input. Since ground truth depth information is not available for the IMA division videos, all models are used as zero-shot prediction models.

Depth maps were used in conjunction with the MoGe estimation of the monocular camera intrinsic matrix to predict 3D point clouds within the camera's local reference frame. The 3D projection of a pixel (u, v) is calculated as follows:

$$p_x = \frac{d \cdot (u - c_x)}{f_x} \tag{2.1}$$

$$p_y = \frac{d \cdot (v - c_y)}{f_y} \tag{2.2}$$

$$p_z = d \tag{2.3}$$

where d is the depth value of pixel (u, v) determined by the depth map, (c_x, c_y) is the camera's principal point, (f_x, f_y) are the camera's horizontal and vertical focal lengths, and $p = (p_x, p_y, p_z)$ is the 3D projection of pixel (u, v). All three models were tested across multiple IMA division surgical frames and performance was assessed qualitatively. The best performing model (MoGe) was incorporated into the angle-estimation pipeline.

2.1.3 Principal Component Analysis

Principal component analysis (PCA) was applied to the segmented and 3D-projected ROIs in each video frame to determine their primary axes. PCA is a widely used linear dimensionality reduction technique that identifies the directions (principal components) along which the variance in the data is maximized. For the 3D segmentations, the first principal component corresponds to the directional vector of the region of interest, capturing its dominant orientation in space. The principal axes were computed by calculating the covariance matrix of the segmented and 3D-projected points:

For each ROI, the top two principal components were obtained from the eigenvectors of the covariance matrix. The first principal component, corresponding to the direction of greatest variance, was taken as the directional vector for each segmentation.

2.1.4 Angle Calculation

Depth angles were measured as the angle between a directional vector and the positive z-axis. An angle θ between two vectors $\mathbf{v1}$ and $\mathbf{v2}$ was calculated as follows:

$$\theta = \arccos\left(\frac{\mathbf{v1} \cdot \mathbf{v2}}{||\mathbf{v1}||\,||\mathbf{v2}||}\right) \tag{2.4}$$

2.2 Validation

A stereo vision setup using the ZED2i stereo camera was employed to evaluate the accuracy of the proposed pipeline. In a stereo vision setup, two spatially separated cameras capture synchronized images of the same scene. By identifying corresponding pixels in the left and right images, the depth of objects can be estimated based on the disparity between their positions in each image. Specifically, depth is computed as follows:

$$d = \frac{f \cdot B}{disp} \tag{2.5}$$

where f is the focal length of the cameras, B is the baseline distance between the cameras, $disp$ is the pixel disparity between corresponding pixels, and d is the calculated depth value.

To calculate depth angles using the stereo vision setup, two points were selected along the directional vector of interest in the 2D image space. These points were identified in both the left and right stereo images, and their depths were calculated using the disparity-based formula described earlier. Each point was then projected into 3D space following the procedure outlined in the point cloud estimation section. A linear 3D directional vector was defined by connecting the two projected points. Finally, the angle between this 3D directional vector and the positive z-axis was calculated according to formula equation 2.4.

Before evaluating the accuracy of the proposed angle estimation pipeline, the stereo vision setup was systematically tested using a newly constructed Cubes validation dataset (see Section 2.2.1) to verify its reliability in measuring depth angles. Once reliability was established, an in vitro stapler validation dataset was generated, and depth angle predictions from the proposed pipeline were compared to ground truth depth angles calculated using the stereo vision setup.

2.2.1 Cubes Validation Dataset

The Cubes dataset was constructed by positioning two cubes on flat graph paper. 40 stereo vision image pairs were collected with the cubes placed in varying relative positions. The objective was to calculate the depth angle of the vector connecting the two cubes. For stereo vision-based angle calculations, two points corresponding to the top-left corner of each cube were manually selected in each stereo image pair and used as input to the angle

calculation procedure described above. Ground truth depth measurements were obtained by manually measuring the distances between the cubes along the horizontal and depth axes of the graph paper, and the corresponding depth angles were computed using basic trigonometric relationships. Mean absolute error for the stereovision setup was calculated over the 40 angle measurements. Figure 2 shows an overview of the construction process of the Cubes validation dataset.

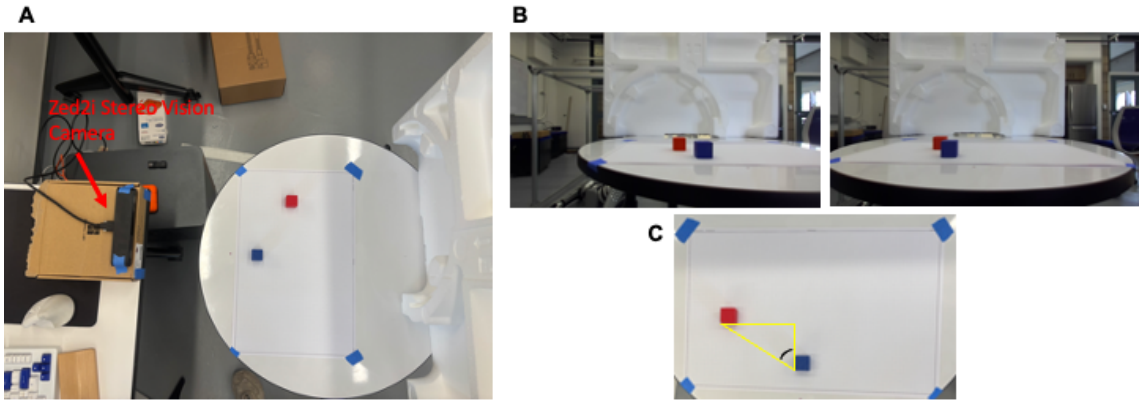

Fig. 2: Overview of Cubes Dataset Construction Process. A) Top-down view of two cube stereo vision setup. The Zed2i stereo vision camera can be seen on the left side of the image. **B)** An example stereo vision image pair from which a depth angle is measured. **C)** Top-down view of the two cubes with the depth angle annotated in yellow.

2.2.2 Stapler Validation Dataset

The stapler validation dataset was created by manipulating a surgical stapler through a range of depth angles in an open-air environment. The stapler was recorded using the ZED2i camera at 1080p HD resolution and 30 frames per second.

Video from the left camera was used an input into the 3D angle estimation pipeline to extract estimated stapler depth angles at each frame. Ground truth was measured using the stereo vision setup. Two circular markers were placed along a generator line of the cylindrical stapler body to facilitate stereo vision-based angle measurements. In each video frame, these markers were segmented using the same segmentation pipeline applied to the stapler and IMA. The segmented markers were used to extract pixel coordinates, which then served as input to the stereo-based angle calculation procedure. Mean absolute error for the angle estimation pipeline was calculated over the 670 input video frames from the two validation videos. An overview of the validation procedure is shown in Figure 3.

2.3 IMA Division Angle Analysis

The angle estimation pipeline was applied to analyze eight surgical videos involving an IMA division from procedures performed by two different colorectal surgeons at Yale

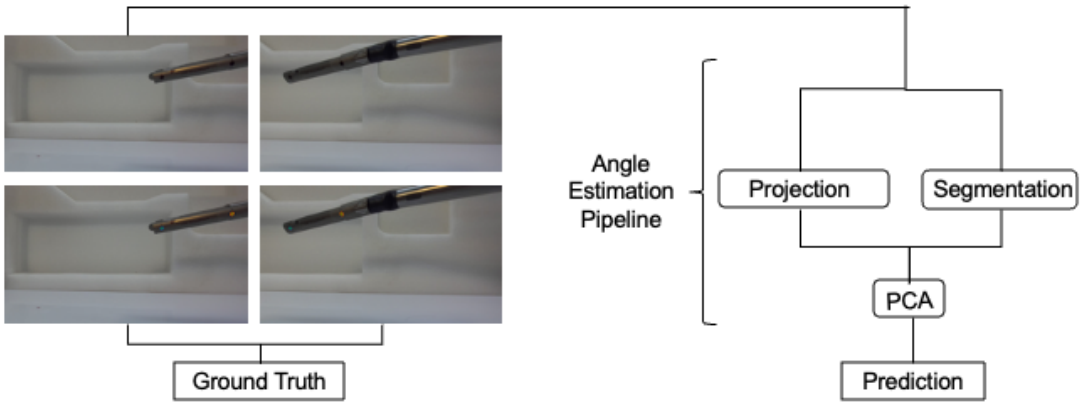

Fig. 3: Overview of Stapler Depth Angle Validation Process. The surgical stapler is manipulated through a diverse range of depth angles and recorded using the Zed2i stereo vision camera. Two circular markers are drawn onto the surface of the cylindrical stapler along a generator line to allow for quick stereo vision ground truth angle calculations. Video input from the left camera was used an input into the 3D angle estimation pipeline to generate a prediction of the stapler depth angle from monocular input. Mean absolute error was calculated for the 3D angle estimation pipeline predictions over all input frames.

New Haven Health hospitals. Each video was trimmed to a 3–15 second interval surrounding the moment of IMA division. These intervals were chosen to capture a period in which both the stapler and the IMA remained stable, allowing for reliable angle analysis during the incision.

Three angles were calculated for each video. These angles include the IMA-depth angle (the angle between the IMA and the positive z-axis), the stapler-depth angle (the angle between the stapler and the positive z-axis), and the IMA-stapler angle.

3 Results

3.1 Point Cloud Estimation

Zero-shot depth estimation results for all three models and their corresponding point cloud projections are shown in Figure 4. All projections use the camera intrinsic matrix estimated by the MoGe model for point cloud projection. All models estimate 3D geometry with reasonable success. Notably, the MoGe and Depth Anything models seem to better distinguish tool edge geometry in comparison to the EndoDAC model. Overall, the MoGe model seems to visually achieve the best point cloud prediction.

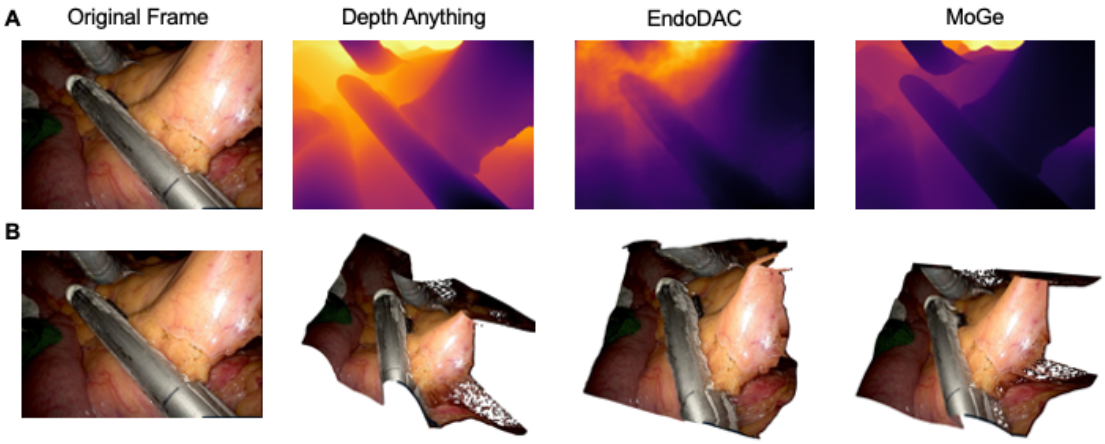

Fig. 4: Depth Map and Point Cloud Comparisons. A). Zero-shot depth map predictions for Depth Anything, EndoDAC, and MoGe models on a representative surgical video frame. **B)** Corresponding 3D projections of depth map predictions for all models using camera intrinsic matrix estimated by the MoGe model. Qualitatively, the MoGe model seems to achieve the best point cloud projection.

3.2 Validation

3.2.1 Stereo Vision Validation: Cubes Dataset

The stereo vision setup was evaluated using forty stereo image pairs to measure the depth angle of the vector connecting two cubes within the field of view. Across these measurements, the setup achieved a mean absolute error of 0.91°, demonstrating that it can reliably estimate depth angles with minimal error.

3.2.2 Angle Estimation Pipeline Validation: Stapler Dataset

The 3D angle estimation pipeline was evaluated across 650 image frames from two in vitro stapler validation videos to measure the depth angle of the surgical stapler. Ground truth depth angles were determined using the stereo vision setup. The pipeline achieved a mean absolute error of 7.67°, demonstrating that depth angles can be estimated with reasonable accuracy from monocular video input.

3.3 IMA Division Angle Measurement Results

3D angle measurements determined from the 3D angle estimation pipeline for all eight surgical videos are shown in Figure 5. Three angles are calculated for each video: the angle between the IMA and the positive z-axis, the angle between the stapler and the positive z-axis, and the angle between the stapler and the IMA. All angles show high variance across the eight surgeries that were analyzed, suggesting that depth angles are not currently standardized across IMA division procedures.

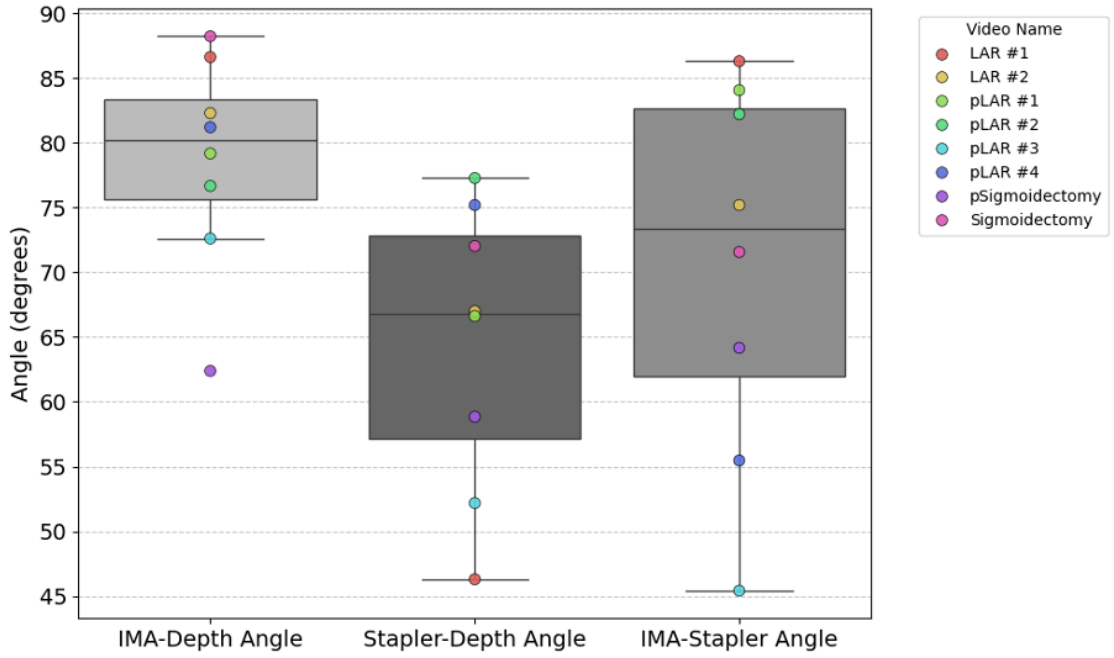

Fig. 5: 3D Angle Measurements for IMA Division Surgeries. Box and whisker plots of 3D angle measurement across all eight surgical videos analyzed. Three 3D angles were measured for each surgical video including the angle between the IMA and the positive z-axis (IMA-Depth Angle), the angle between the stapler and the positive z-axis (Stapler-Depth Angle), and the angle between the IMA and the stapler (IMA-Stapler Angle). All measured angles show a large range and variance across surgeries.

4 Related Work

Prior work has examined 2D angular metrics and demonstrated their relevance to surgical performance. In laparoscopic porcine enterotomy procedures, for example, 45° and 60° manipulation angles resulted in shorter execution times than 90° angles [11]. Similarly, a 60° angle was associated with both shorter task durations and higher performance scores in a standardized endoscopic suturing task when compared to 30° and 90° configurations [12]. These studies suggest that angular positioning may influence technical efficiency and overall performance.

In contrast, limited work has focused on measuring 3D angles or positional relationships in surgical settings or correlating these measurements with performance outcomes. Some prior efforts have leveraged specialized tools or camera systems to enable 3D tracking. For instance, one study developed a tool with three co-linear markers to enable estimations of position and ergonomic angles with millimeter-level precision [13]. While this approach yields accurate 3D pose estimates, it depends on a custom-marked instrument and a specialized high-performance camera not typically found in laparoscopic operating rooms.

Other systems employ stereoscopic laparoscopes, which add a second camera to the standard laparoscopic setup to enhance depth perception and enable 3D reconstruction [14, 15]. Stereoscopic visualization has been shown to improve task performance and potentially enhance surgical outcomes [16, 17]. Notably, stereo laparoscopic video could be readily incorporated into the proposed pipeline to improve 3D point cloud estimation to yield more precise angle measurements.

For surgical videos that lack stereoscopic data or instrument-mounted markers, monocular depth and point cloud estimation techniques are required. Estimating depth and, relatedly, 3D position from monocular images remains an open problem in computer vision. Recent advances in both generalized and surgical context depth estimation have enabled reasonably accurate predictions [18, 19]. Some of these models, including Depth Anything [20], EndoDAC [21], and MoGe [10] are explored here.

5 Discussion and Future Work

We developed and validated a 3D angle estimation pipeline to measure spatial relationships during IMA division. The pipeline combines segmentation to identify ROIs, point cloud estimation to project these regions into 3D space, and principal component analysis to determine their directional vectors. In validation experiments, our method achieved a mean absolute error of 7.67°. While not highly precise, this level of accuracy is sufficient to detect moderate positional differences, making the angle estimation system appropriate for analyzing coarse-grained variations in surgical technique.

Using this tool to analyze eight IMA division surgeries, we observed that key depth angles—specifically, the IMA-depth angle, stapler-depth angle, and IMA-stapler angle—exhibit substantial variability. This variation suggests that these angles are not standardized in current practice, reflecting a lack of consistent conventions. As depth angles directly reflect the spatial orientation of surgical instruments and anatomical targets, these findings highlight an opportunity to establish more formalized positional targets for IMA division. A standardized range of angles could help guide stapler placement, inform surgical decision-making, and support consistency in technique across surgeons and institutions.

Importantly, the angle estimation pipeline developed in this work provides a practical framework for such standardization. When deployed at scale across a larger number of surgical videos, this system could enable quantitative comparisons of technique, allowing for the identification of links between 3D positions and postoperative outcomes such as Ischemia or anastomotic leakage. If strong associations are identified, angle metrics could be incorporated to facilitate training and improve outcomes. Notably, this pipeline is highly generalizable, thereby providing a framework for 3D analysis that could lead to better clinical outcomes across a range of laparoscopic surgeries.

Nonetheless, several limitations remain. Most notably, the accuracy of the angle estimation method is constrained by the performance of the underlying 3D point cloud predictions. As our pipeline currently relies on zero-shot monocular geometry estimation, frames with poor lighting condition or low contrast may be particularly vulnerable to performance degradation. The development of better general-purpose or endoscopic monocular 3D point cloud reconstruction algorithms could enable more accurate angle estimations.

Additionally, validation was conducted using a rigid stapler in an open-air environment, which does not fully replicate the challenges of segmenting and estimating depth from soft, deformable, and low-contrast tissue. In particular, the IMA's appearance can vary significantly depending on lighting, manipulation, and the surrounding tissue context. Future work should include validation using structures that mimic tissue-like deformity and texture or annotated in vivo datasets to test performance more rigorously under surgical conditions.

Currently, the system operates in a semi-automated fashion, requiring manual selection of the stapler and IMA segmentations in the first video frame. Developing a classification model to assign SAM 2 outputs to anatomical categories would enable full automation of the workflow and improve scalability.

Bibliography

[1] Yeisson Rivero-Moreno et al. “Robotic Surgery: A Comprehensive Review of the Literature and Current Trends”. In: *Cureus* 15.7 (), e42370. ISSN: 2168-8184. DOI: 10.7759/cureus.42370. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10445506/ (visited on 05/02/2025).

[2] Ace St. John et al. “The Rise of Minimally Invasive Surgery: 16 Year Analysis of the Progressive Replacement of Open Surgery with Laparoscopy”. In: *JSLS : Journal of the Society of Laparoscopic & Robotic Surgeons* 24.4 (2020), e2020.00076. ISSN: 1086-8089. DOI: 10.4293/JSLS.2020.00076. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7810432/ (visited on 05/02/2025).

[3] Giulio Mari et al. “”High or low Inferior Mesenteric Artery ligation in Laparoscopic low Anterior Resection: study protocol for a randomized controlled trial” (HIGHLOW trial)”. In: *Trials* 16 (Jan. 27, 2015), p. 21. ISSN: 1745-6215. DOI: 10.1186/s13063-014-0537-5.

[4] Jun Yu et al. “High and low inferior mesenteric artery ligation in laparoscopic low anterior rectal resections: A retrospective study”. In: *Frontiers in Surgery* 9 (Jan. 13, 2023), p. 1027034. ISSN: 2296-875X. DOI: 10.3389/fsurg.2022.1027034. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9881683/ (visited on 05/02/2025).

[5] Antonio Brillantino et al. “Inferior Mesenteric Artery Ligation Level in Rectal Cancer Surgery beyond Conventions: A Review”. In: *Cancers* 16.1 (Jan. 2024). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 72. ISSN: 2072-6694. DOI: 10.3390/cancers16010072. URL: https://www.mdpi.com/2072-6694/16/1/72 (visited on 05/02/2025).

[6] A. Balcerzak et al. “Types of inferior mesenteric artery: a proposal for a new classification”. In: *Folia Morphologica* 80.4 (2021). Number: 4, pp. 827–838. ISSN: 1644-3284. DOI: 10.5603/FM.a2020.0115. URL: https://journals.viamedica.pl/folia_morphologica/article/view/FM.a2020.0115 (visited on 05/02/2025).

[7] Melissa N. N. Arron et al. “Mesenteric occlusive disease of the inferior mesenteric artery is associated with anastomotic leak after left-sided colon and rectal cancer surgery: a retrospective cohort study”. In: *International Journal of Colorectal Disease* 37.3 (Mar. 2022), pp. 631–638. ISSN: 1432-1262. DOI: 10.1007/s00384-021-04089-0.

[8] S. Fujii et al. “Randomized clinical trial of high versus low inferior mesenteric artery ligation during anterior resection for rectal cancer”. In: *BJS open* 2.4 (Aug. 2018), pp. 195–202. ISSN: 2474-9842. DOI: 10.1002/bjs5.71.

[9] Nikhila Ravi et al. “SAM 2: Segment Anything in Images and Videos”. In: The Thirteenth International Conference on Learning Representations. Oct. 4, 2024. URL: https://openreview.net/forum?id=Ha6RTeWMd0 (visited on 05/02/2025).

[10] Ruicheng Wang et al. *MoGe: Unlocking Accurate Monocular Geometry Estimation for Open-Domain Images with Optimal Training Supervision*. Apr. 15, 2025. DOI: 10.48550/arXiv.2410.19115. arXiv: 2410.19115[cs]. URL: http://arxiv.org/abs/2410.19115 (visited on 05/02/2025).

[11] S. Manasnayakorn, A. Cuschieri, and G. B. Hanna. “Ideal manipulation angle and instrument length in hand-assisted laparoscopic surgery”. In: *Surgical Endoscopy* 22.4 (Apr. 1, 2008), pp. 924–929. ISSN: 1432-2218. DOI: 10.1007/s00464-007-9520-5. URL: https://doi.org/10.1007/s00464-007-9520-5 (visited on 05/02/2025).

[12] G. B. Hanna, S. Shimi, and A. Cuschieri. “Influence of direction of view, target-to-endoscope distance and manipulation angle on endoscopic knot tying”. In: *BJS (British Journal of Surgery)* 84.10 (1997). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2168.1997.02835.x, pp. 1460–1464. ISSN: 1365-2168. DOI: 10.1111/j.1365-2168.1997.02835.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2168.1997.02835.x (visited on 05/02/2025).

[13] Sangkyun Shin et al. “A single camera tracking system for 3D position, grasper angle, and rolling angle of laparoscopic instruments”. In: *International Journal of Precision Engineering and Manufacturing* 15.10 (Oct. 1, 2014), pp. 2155–2160. ISSN: 2005-4602. DOI: 10.1007/s12541-014-0576-6. URL: https://doi.org/10.1007/s12541-014-0576-6 (visited on 05/02/2025).

[14] Katie Schwab et al. “Evolution of stereoscopic imaging in surgery and recent advances”. In: *World Journal of Gastrointestinal Endoscopy* 9.8 (Aug. 16, 2017), pp. 368–377. ISSN: 1948-5190. DOI: 10.4253/wjge.v9.i8.368. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565502/ (visited on 05/02/2025).

[15] L. Maier-Hein et al. “Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery”. In: *Medical Image Analysis* 17.8 (Dec. 1, 2013), pp. 974–996. ISSN: 1361-8415. DOI: 10.1016/j.media.2013.04.003. URL: https://www.sciencedirect.com/science/article/pii/S1361841513000480 (visited on 05/02/2025).

[16] Martin Schoenthaler et al. “Stereoscopic (3D) versus monoscopic (2D) laparoscopy: comparative study of performance using advanced HD optical systems in a surgical simulator model”. In: *World Journal of Urology* 34.4 (Apr. 1, 2016), pp. 471–477. ISSN: 1433-8726. DOI: 10.1007/s00345-015-1660-y. URL: https://doi.org/10.1007/s00345-015-1660-y (visited on 05/02/2025).

[17] Najib Isse Dirie, Qing Wang, and Shaogang Wang. “Two-Dimensional Versus Three-Dimensional Laparoscopic Systems in Urology: A Systematic Review and Meta-Analysis”. In: *Journal of Endourology* 32.9 (Sept. 1, 2018), pp. 781–790. ISSN: 0892-7790. DOI: 10.1089/end.2018.0411. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6156697/ (visited on 05/02/2025).

[18] Yue Ming et al. “Deep learning for monocular depth estimation: A review”. In: *Neurocomputing* 438 (May 28, 2021), pp. 14–33. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2020.12.089. URL: https://www.sciencedirect.com/science/article/pii/S0925231220320014 (visited on 05/02/2025).

[19] Yasuhide Hirohata et al. “Confidence-aware self-supervised learning for dense monocular depth estimation in dynamic laparoscopic scene”. In: *Scientific Reports* 13.1 (Sept. 16, 2023). Publisher: Nature Publishing Group, p. 15380. ISSN: 2045-2322. DOI: 10.1038/s41598-023-42713-x. URL: https://www.nature.com/articles/s41598-023-42713-x (visited on 05/02/2025).

[20] Lihe Yang et al. *Depth Anything V2*. Oct. 20, 2024. DOI: 10.48550/arXiv.2406.09414. arXiv: 2406.09414[cs]. URL: http://arxiv.org/abs/2406.09414 (visited on 05/02/2025).

[21] Beilei Cui et al. “EndoDAC: Efficient Adapting Foundation Model for Self-Supervised Depth Estimation from Any Endoscopic Camera”. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*. Ed. by Marius George Linguraru et al. Cham: Springer Nature Switzerland, 2024, pp. 208–218. ISBN: 978-3-031-72089-5. DOI: 10.1007/978-3-031-72089-5_20.