Automated Extraction of Surgical Needles from Tissue Phantoms

Priya Sundaresan[1], Brijen Thananjeyan[1], Johnathan Chiu[1], Danyal Fer[2], Ken Goldberg[1]

***Abstract*— We consider the surgical subtask of automated extraction of embedded suturing needles from silicone phantoms and propose a four-step algorithm consisting of calibration, needle segmentation, grasp planning, and path planning. We implement autonomous extraction of needles using the da Vinci Research Kit (dVRK). The proposed calibration method yields an average of 1.3mm transformation error between the dVRK end-effector and its overhead endoscopic stereo camera compared to 2.0mm transformation error using a standard rigid body transformation. In 143/160 images where a needle was detected, the needle segmentation algorithm planned appropriate grasp points with an accuracy of 97.20% and planned an appropriate pull trajectory to achieve extraction in 85.31% of images. For images segmented with >50% confidence, no errors in grasp or pull prediction occurred. In images segmented with 25-50% confidence, no erroneous grasps were planned, but a misdirected pull was planned in 6.45% of cases. In 100 physical trials, the dVRK successfully grasped needles in 75% of cases, and fully extracted needles in 70.7% of cases where a grasp was secured.**

I. INTRODUCTION

Robotic minimally invasive surgery (RMIS) is an established method for surgical operation that involves a surgeon performing an operation via master-slave teleoperation of a Robot Surgical Assistant (RSA), such as Intuitive Surgical's da Vinci [12] and the Raven-II surgical robot [8]. The surgeon exercises full control and oversight of the robot's motions, so tedious tasks such as suturing, which require a high degree of precision and delicacy, can lead to surgeon exhaustion. Recent work investigates automation of robot-assisted suturing subtasks including tissue localization and debridement [7, 20, 30], needle manipulation [5, 10, 31], and knot-tying [2, 3]. Automation of suturing subtasks can potentially reduce exertion on the part of the surgeon, improve workflow by reducing procedure times, avoid tool hand-offs, and prevent the infrequent yet dangerous event of needle loss and retention in a patient [9, 11].

Loss and subsequent recovery of surgical needles has been found to account for up to 76% of all 'near miss' events (situation where patient is exposed to a hazardous situation but does not experience harm due to early detection or luck) [4], leading to increased operative times, cost and risk to the patient. Improved extraction of the surgical needles can reduce ischemia through reducing the number of repeated needle throws and number of lost needles.

The focus of this paper is automating the subtask of locating an embedded surgical needle in a phantom in the RSA's camera view and extracting it from the phantom. Detecting a suturing needle on a surgical phantom poses several challenges in image processing. Suturing needles typically range from 0.02 to 0.8mm in thickness and 32 to 55mm in length. One issue is that needles are semicircular in shape but can appear curved to linear once embedded depending on the angle of insertion and the angle of view in an RSA's camera frame. Needles can also have a highly reflective metallic surface which can cause needles to be overexposed or discolored in different configurations.

Aside from perception challenges, the task of embedded needle extraction has control-based obstacles as well. Cable-driven RSA's including the da Vinci can be kinematically inaccurate due to inconsistent cable tensioning across a workspace, cable self-collisions, and cable friction [13, 15,

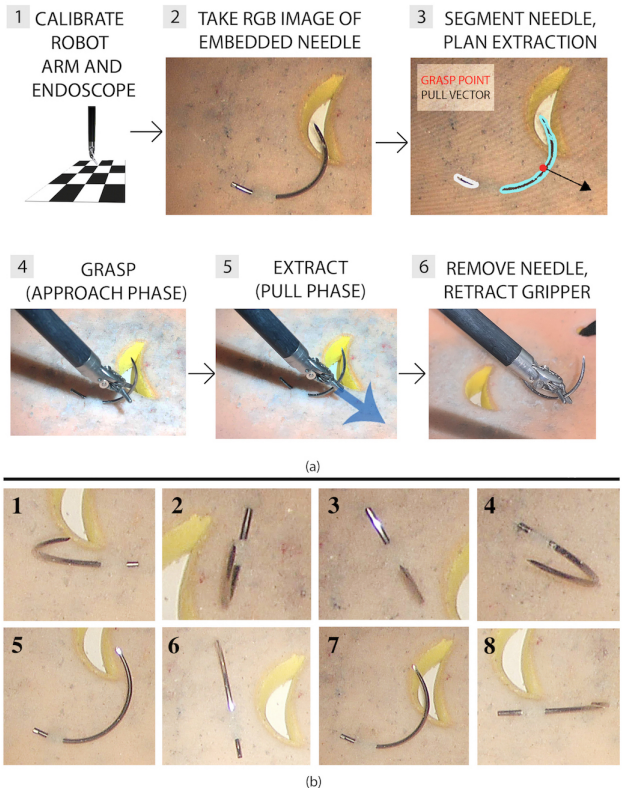

Fig. 1. (a) Needle extraction pipeline. Given an image of a suturing needle embedded in a tissue phantom, we segment the needle to identify a grasp point and a direction in which to pull the protruding needle and execute the extraction using a learned correspondence between the robot end-effector and camera coordinate space. (b) Challenges of vision-based needle segmentation. Sub-figures 1-8 present a sample of the many varying poses embedded suturing needles can take on, which complicates segmentation. Examples 1, 2, and 4 exhibit both the stark highlights and shadows that can appear on an embedded needle and examples 3 and 6 exhibit specularity, both of which interfere with detecting a needle contour in full. Examples 2, 3, 6, and 8 show linearity while examples 1, 4, 5, and 7 exhibit curvature.

Authors are affiliated with:
[1]AUTOLAB at UC Berkeley; @berkeley.edu
[2]UC San Francisco East Bay; @ucsf.edu
{priya.sundaresan, bthananjeyan, johnathanchiu, danyal.fer, goldberg}

21]. In particular, the da Vinci Research Kit (dVRK), assembled from parts of the da Vinci Surgical System, has two cable-driven arms called patient side manipulators (PSMs) [12]. 4 DoF tools can be secured to these PSMs via a cable system, which yields inherent control error when the tool's end-effector is commanded to a particular orientation and position in the PSM coordinate frame. Every calibration method has some error associated with the transformation between the dVRK's endoscope and PSM frames [22, 27].

In this paper, we present an algorithm for needle detection with an endoscopic stereo camera and extraction of embedded suturing needles with the dVRK. In all experiments, we use one PSM, and one suturing needle. We focus on cases where the majority of the needle is protruding from the phantom in order to be able to secure a solid grasp and to avoid trauma to the tissue-simulating phantom. To improve the calibration of the rigid body transformation between the endoscopic and PSM frames, we find corresponding points in each frame and project those points to their respective best-fit planes and then apply a rigid body transformation.

Summary of Contributions

The approach in this paper draws from several prior works on suturing needle detection and pickup/extraction and makes three contributions:

1) A calibration algorithm that builds upon [18] which results in an end-effector to endoscope transformation error of 1.3mm, compared to a transformation error of 2.0mm associated with a standard rigid body transformation.
2) An algorithm for robust needle detection using ellipse-fitting and contour-based segmentation that does not require marking needles with colored indicators as in [5, 24, 31, 33] or providing a high contrast background to isolate needles as in [14].
3) An algorithm for finding a grasp point for an embedded needle and the appropriate direction in which to pull the needle to extract it from a surgical phantom with an implementation for performing the extraction with the dVRK. The procedure identified a needle from 143/160 workspace images, and of those, planned an appropriate grasp in 97.20% of cases and an appropriate pull in 85.31% of cases. Given an appropriate segmentation in 100 trials, the dVRK successfully grasped an embedded needle in 75.0% of trials and of those, fully extracted the needle in 70.7% of cases.

II. RELATED WORK

A. Calibration

Calibration, particularly fine-tuning the correspondence between an RSA's endoscope viewing frame and the frame of its arm, is crucial for surgical robots like the dVRK to be able to perform tasks that require a high degree of precision and dexterity, especially in unstructured environments. Prior work has relied on using multiple sensors such as angled lasers, stereo cameras, monocular cameras as in [28], learning non-linear correction factors for cable-driven RSAs as in [19, 30], learning hand-eye coordination through visual-servoing as in [16], or scanning the arm's tooltip across known points in three planes in the workspace and using plane interpolation to register a point cloud as in [5]. Mahler *et al.* [19] used Gaussian Process Regression and data-cleaning to achieve a positional error of 2mm with the Raven II Surgical Robot. Seita *et al.* [30] achieved a transformation error of 1.08mm by allowing a dVRK arm to move through open-loop trajectories at random and training a DNN and Random Forest (RF) on the camera pixels and internal pose of the end-effector at various points in those trajectories to learn and correct the positional bias of the robot. Using the three plane scanning method, D'Ettorre *et al.* [5] constructed a point cloud with an error of 0.88 mm with the dVRK. In this paper, we explore the practicality of achieving transformational error minimal enough for implementing grasping and needle extraction. We explore calibration methods that are comparatively not nearly as data-intensive as the aforementioned techniques, which require additional sensors as in [28] or extensive raw data collection as in [5, 19, 30].

B. Needle Tracking

Prior work on needle tracking includes using coloration to mark a suturing needle and track its position and orientation. D'Ettorre *et al.* [5] labeled suturing needles with a green marker on either endpoint of the needle and at the center of the needle in order to streamline pose registration and command the dVRK to grasp at the central green marker and Speidel *et al.* [33], Sen *et al.* [31], and Nageotte *et al.* [24] used a similar color-based segmentation technique for pose registration. Kil *et al.* [14] analyzed suturing needle position and orientation for the sake of assessing suturing skill and technique in surgeons using a high-contrast, LED backlit workspace for ease of needle segmentation.

These works highlight a common obstacle across a variety of automated suturing subtasks of accurately segmenting a suturing needle efficiently without the use of visual markers. In contrast, we study needle segmentation in situations including workspace lighting variation, a tissue phantom with visible defects, and no color-based markers for tracking suturing needles.

C. Suturing Subtask Automation

Learning from both physical and artificial expert demonstrations has been used in prior work to automate various suturing subtasks [17, 25, 35, 36]. Schulman *et al.* [29] explored performing different suturing trajectories including insertion and extraction by learning from demonstrations and using a non-rigid transformation to adapt the task to a warped environment subject to rotations, stretching, and scaling compared to the demonstration environment. This approach uses manual annotations to map a coordinate frame in a demonstration to a coordinate space in a newly warped scenario. The tasks were scaled up several times larger than suturing scenarios to provide an error margin on the order of centimeters. In this paper, we explore needle extraction

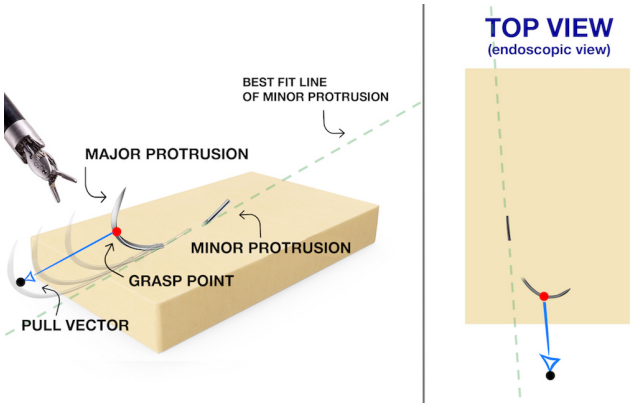

Fig. 2. This diagram defines terms that are used throughout the paper, namely *major protrusion*, *minor protrusion*, *grasp point*, and *pull vector*. The side view is a generic angled view of the workspace while the top view visualizes the embedded needle seen from the overhead endoscopic frame.

without demonstrations with error tolerance within a margin on the order of millimeters.

Sen *et al.* [31] approached the task of suturing as a convex optimization problem with the entry and exit positions for a given trajectory on a surgical phantom and demonstrated execution of multilateral throw suturing. Drawing from this work, we focus on segmentation, planning, and extraction of needles from uncertain configurations.

III. PROBLEM STATEMENT

The algorithm is provided an endoscopic image of the workspace containing a partially embedded suturing needle. The visible components of the needle are segmented to determine a point on the needle to grasp it and a point in the workspace to pull it to.

A. Definitions and Notation

In this paper, we will use the following notation:

- We denote a vector $\vec{x} \in \mathbb{R}^3$ in the *camera* frame using a superscript c as in $\vec{x}^c = [c_x, c_y, c_z]^T$. The three-dimensional camera coordinate frame is constructed by finding corresponding chessboard corner points in both the left and right stereo cameras and using their pixel-wise disparity to infer depth.
- We denote a vector $\vec{x} \in \mathbb{R}^3$ in the *base* frame of the PSM of the dVRK using a superscript b as in $\vec{x}^b = [b_x, b_y, b_z]^T$.
- The orientation of the PSM $\in \mathbb{R}^3$ in the base frame of the PSM is given by $\vec{\phi} = [\phi_y, \phi_p, \phi_r]^T$, where the elements are the yaw, pitch, and roll, respectively, which define the orientation of the tooltip.
- Scalars are represented using plain letters (e.g. α).
- A transformation $\in \mathbb{R}^{3\times3}$ from the camera frame to base frame or base frame to camera frame is represented as ${}^c\mathbf{T}_b$ or ${}^b\mathbf{T}_c$, respectively.
- A vector $\in \mathbb{R}^3$ transformed from the camera frame to the base frame is given by $\vec{x}^b = {}^c\mathbf{T}_b \vec{x}^c$ and vice versa.
- We define the *minor protrusion* and the *major protrusion* to be the smaller and larger segments of a needle protruding from a phantom once embedded (See Figure 2).
- We define the *grasp point* to be the point on the needle where the PSM gripper encloses the needle and the *pull vector* to be a point in space where the gripper and its contained needle can extract the needle (see Figure 2).
- A needle is considered fully extracted if the minor protrusion is completely freed from the phantom after a pull, and a partial extraction occurs when the needle remains partially embedded.

B. Assumptions

1) Needle Pose: We assume that the majority of the needle is protruding from the exit point on the phantom, and there is a minor protrusion of the needle at the point of the insertion that must be guided through the exit point to achieve extraction. We assume that both protrusions are unobscured in the endoscope camera. Furthermore, we assume that the minor protrusion can be approximated as a linear segment due to its minimal size despite the natural curvature of the suturing needle, and that the line of best-fit of the minor protrusion is a permissible heuristic for the direction of pull required to achieve extraction (Figure 2).

2) Equipment: We assume access to an fixed overhead 1920×1080px endoscopic stereo camera which is standard across da Vinci systems. The endoscope is situated above the tissue phantom and the workspace is illuminated by a standard surgical lamp with adjustable intensity. We assume the tissue phantom is a deformable, flat 7.5×7.5cm silicone-based platform and that the phantom may have physical defects or discoloration in areas. We use one PSM of the dVRK equipped with a large needle driver, or gripper-based end effector (See Figure 3). With fully open jaws, the gripper width is 10mm across with a range of 75°.

3) Experimental Setup: We assume that there is initially one suturing needle embedded into the tissue phantom that is graspable by the large needle driver with fully open jaws. We do not assume access to force sensing or haptic feedback as in [23]. Lastly, we assume needles can be successfully extracted by pulling a constant distance from the grasp point, a simple, experimentally reliable heuristic that tends to overestimate the actual distance needed to achieve extraction. This constant, the magnitude of the pull vector, jointly characterizes the needle extraction action with the pull direction and is hand-tuned to be 350px in a segmented image.

C. Input

The input consists of one image of the embedded needle taken from the endoscope equipped by the dVRK.

D. Output

The output is a grasp point and a pull vector. The grasp point is a point on the needle to grasp and the pull vector defines the direction to pull the needle. Both points are in the PSM frame. Extraction is considered successful if the PSM pulls the needle completely out of the phantom.

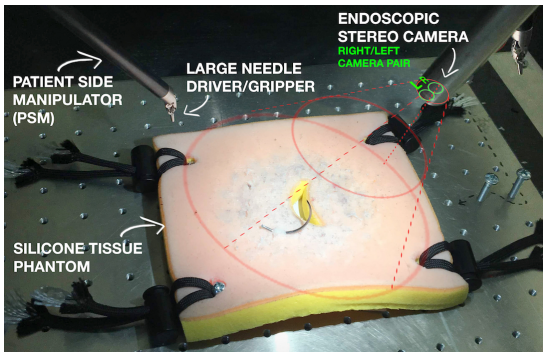

Fig. 3. Experimental setup.

IV. ALGORITHMS

A. Physical Calibration Algorithm

We learn a mapping between the base frame (PSM frame) and the endoscope frame using a corrective rigid body transformation on a set of corresponding points in the base and endoscope frames [18]. We evaluate this approach relative to standard approaches in Section V.

We locate a 5×5 chessboard grid in both the endoscope frame and the PSM frame, giving $\{\vec{x}_1^c, \vec{x}_2^c, \ldots, \vec{x}_{25}^c\} \in \mathbb{R}^3$ and $\{\vec{x}_1^b, \vec{x}_2^b, \ldots, \vec{x}_{25}^b\} \in \mathbb{R}^3$, respectively.

The close proximity of the left and right cameras, which yield a 3D coordinate frame using ray triangulation and pixel-wise disparity, can lead to scene geometry distortion and a lack of accurate depth perception. To correct for this, we project each $\vec{x}_i^c$ to the best-fit plane of $\{\vec{x}_1^c, \vec{x}_2^c, \ldots, \vec{x}_{25}^c\}$ prior to doing a standard rigid body transformation.

Define $\mathbf{A} \in \mathbb{R}^{25\times 3}$ as a one-padded matrix of the x and y coordinates of the scene points:

$$\mathbf{A} = \begin{bmatrix} \vec{x}_1^c[0] & \vec{x}_1^c[1] & 1 \\ \vec{x}_2^c[0] & \vec{x}_2^c[1] & 1 \\ \vdots & \vdots & \vdots \\ \vec{x}_{25}^c[0] & \vec{x}_{25}^c[1] & 1 \end{bmatrix}$$

Define $\vec{z} = [\ \vec{x}_1^c[2]\ \ \vec{x}_2^c[2]\ \ \ldots\ \ \vec{x}_{25}^c[2]\ ]^T \in \mathbb{R}^3$ to be the vector of the z coordinate of the scene points.

The vector of planar coefficients $\vec{v}$, is the least squares solution to $\mathbf{A}\vec{v} = \vec{z}$. The corrected z coordinate vector of the scene points, $\vec{z'}$, is given by:

$$\vec{z'} = a \begin{bmatrix} \vec{x}_1^c[0] \\ \vec{x}_2^c[0] \\ \vdots \\ \vec{x}_{25}^c[0] \end{bmatrix} + b \begin{bmatrix} \vec{x}_1^c[1] \\ \vec{x}_2^c[1] \\ \vdots \\ \vec{x}_{25}^c[1] \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{1}$$

where $a, b, c = \vec{v}[0], \vec{v}[1], \vec{v}[2]$ are the coefficients in the standard normal form for a plane.

Let $\vec{x}_i'^c = [\ \vec{x}_i^c[0]\ \ \vec{x}_i^c[1]\ \ \vec{z'}[i]\ ]^T$, or the original $\vec{x}_i^c$ with a corrected z-coordinate.

Then, we learn ${}^c\mathbf{T}_b$ as the rigid body transformation between $\{\vec{x}_1'^c, \vec{x}_2'^c, \ldots, \vec{x}_{25}'^c\}$ and $\{\vec{x}_1^b, \vec{x}_2^b, \ldots, \vec{x}_{25}^b\}$ as in [32].

B. Camera Calibration Algorithm

In all segmentation methods used, we use only the left camera contained in the stereo endoscope to detect a set of points of interest (i.e. contours, centroids, grasp points, pull points, etc.) $\{\vec{x}_1^{c_L}, \vec{x}_2^{c_L}, \ldots, \vec{x}_n^{c_L}\}$ and locate them in the right camera by learning the rigid body transformation ${}^{c_L}\mathbf{T}_{c_R}$ between the left and right cameras on a known set of corresponding chessboard points in each respective 2D frame. This transform has an associated error of 2.55px, and results in a set of points $\{\vec{x}_1^c, \vec{x}_2^c, \ldots, \vec{x}_n^c\}$ in the 3D endoscopic frame. This reduces the complexity of the segmentation algorithm by preventing the need to exhaustively find contours in two separate images and also allows us to take advantage of the favorable camera quality of the left camera over the right, which is specific to our endoscope.

C. Needle Segmentation Algorithm

The input to the needle detection system is a raw RGB endoscopic image of the workspace. This image must be preprocessed to remove noise and then filtered by contour size and shape to segment needles (see Figure 4).

1) Preprocessing: We convert an input workspace image to grayscale and threshold it using Otsu's method, which exhaustively finds the binarization that minimizes intra-class variance amongst pixels categorized as dark or light [26]. This step makes the thresholded image robust to varying lighting conditions, helps to normalize shadows, glare, and highlights on embedded needles, and suppresses visual distractions such as defects on the tissue phantom. We test robustness to these conditions in Section V.

2) Contour Segmentation: We detect and iterate through all contours in the preprocessed image using the approach in Suzuki *et al.* [34] to find the major protrusion of the embedded needle. A contour is a curve joining a set of continuous points $\{\vec{x}_1^c, \vec{x}_2^c, \ldots, \vec{x}_n^c\}$ with shared coloration or intensity [1], and a contour is considered a valid major protrusion candidate if the raw contour area is within a fixed margin and the area of the contour fit to an elliptical shape (denoted $\mathcal{E}$ and found using the Fitzgibbon least squares method) is also within a fixed margin [6] (see Figure 4).

The purpose of the raw area filter is to disregard contours that are too large or small to be considered part of a needle. The filter for the ellipse-fit contour draws from the fact that an embedded needle is capable of appearing anywhere from linear at a given insertion angle to nearly semicircular. In the former case, the $\mathcal{E}$ is expected to be very similar in area to the raw area lower bound, since the natural contour and $\mathcal{E}$ are expected to look like roughly linear segments. In the latter, $\mathcal{E}$ is expected to be similar in area to the natural circular area traced out by the ellipse-fit of a non-embedded suturing needle contour resting on its side.

A contour that satisfies these bounding conditions is considered a valid major protrusion $\mathcal{P}$ and is defined as:

$$\left\{\{\vec{x}_1^c, \ldots, \vec{x}_n^c\} \middle| \alpha_L < \iint_{D_{\mathcal{R}}} dA < \alpha_H, \epsilon_L < \iint_{D_{\mathcal{E}}} dA < \epsilon_H \right\}$$

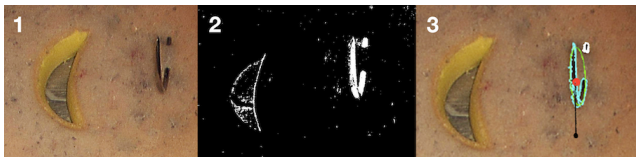

Fig. 4. Segmentation Pipeline. 1. Raw input RGB Image. 2. Otsu thresholded image. 3. Exhaustive contour search yields the major protrusion (blue), verified by fitting an ellipse (green) and ensuring the area of the raw contour and the ellipse-fitted contour is properly bounded. The minor protrusion (white) is also found iteratively and the grasp/pull is planned.

where α_L and α_H represent the raw contour area constraint for $D_{\mathcal{R}}$, the region enclosed by the raw contour, and ϵ_L and ϵ_H represent the raw contour area constraint for $D_{\mathcal{E}}$, the region enclosed by the contour fit to an ellipse. $\alpha_L, \alpha_H, \epsilon_L$, and ϵ_H are hand-tuned to be 1800px^2, 20000px^2, 1300px^2, and 210000px^2, respectively.

Once $\mathcal{P}$ has been identified, we compute a central grasp point on the contour by first determining the center of mass $\vec{x}^c_{\mathcal{P}_{COM}}$ of the contour from its moments as outlined in [1].

Since suturing needles can appear semicircular, the center of mass may be concentrated away from the contour itself but centered within the hollow of the curve. We consider a valid grasp point $\vec{g}^c$ to be on the major protrusion itself however, so we take the grasp point to be the closest point on the contour to the center of mass:

$$\vec{g}^c = \arg\min_{\vec{x}^c_i \in \mathcal{P}} || \vec{x}^c_{\mathcal{P}_{COM}} - \vec{x}^c_i || \tag{2}$$

Lastly, we find the major protrusion's associated minor protrusion, $\mathcal{R}$. We iterate through all image contours within an experimentally-determined area range (250 to 2000px^2) specified for minor protrusions and within a fixed radius (120 to 400px^2) of the major protrusion. The purpose of this second condition is to ensure that in cases where a major protrusion is segmented partially instead of fully (see Figure 5: trials 3 and 4), the predicted minor protrusion is not actually part of the major protrusion. Bounding the candidates to be a fixed distance away from the major protrusion ensures that a split major protrusion does not yield a predicted major/minor protrusion pair, and setting an outer bounding radius ensures that the minor protrusion selected is a reasonable choice rather than a stray mark far away from the major protrusion. Given this filtered set of candidates, the contour with a center of mass that is is closest to $\vec{g}^c$ in terms of Euclidean distance is selected as the corresponding minor protrusion.

D. Path Planning and Physical Extraction

With a major protrusion $\mathcal{P}$ with a grasp point $\vec{g}^c$ and its associated minor protrusion $\mathcal{R}$ identified, we determine a pull vector $\vec{p}^c$, or the offset the grasp point must be displaced to extract a needle, using analytical methods.

We assumed that the set $\mathcal{R}$ could be linearly approximated due to its small size and we further assumed that the direction of insertion is a reasonable heuristic for direction of extraction, so it follows that the best-fit line of points in $\mathcal{R}$ provides a direction for pull for $\mathcal{P}$. We take the direction of the pull to be the slope of the least-squares best-fit line of points in $\mathcal{R}$, denoted $\vec{m}^c$. Given $\vec{m}^c$ and $\vec{g}^c$, we compute $\vec{p}^c$ with direction $\vec{m}^c$ and $||\vec{p}^c||$ = 350px, the standard hand-tuned offset between grasp and pull points across all scenarios.

Using the corrected transformation ${}^c\mathbf{T}_b$ computed in the calibration phase, we locate the grasp point $\vec{g}^b$ and pull vector $\vec{p}^b$ in the base frame of the PSM as $\vec{g}^b = {}^c\mathbf{T}_b \vec{g}^c$ and $\vec{p}^b = {}^c\mathbf{T}_b \vec{p}^c$

From any fixed orientation $\vec{\phi}$, the PSM first moves to $\vec{g}^b$ with opened jaws, grasps the suturing needle, moves to $\vec{p}^b$, and releases the needle to complete the extraction task.

V. SEGMENTATION EXPERIMENTS

In this experiment, the system and baselines are evaluated for needle segmentation accuracy and robustness on a dataset of images collected from the experimental setup.

A. Experimental Setup

We follow the experimental setup described in Section III, with the overhead workspace lamp intensity set to 50% brightness, which we found produces less glare from the needle while still illuminating the workspace. The algorithm is evaluated for robustness to lighting variation.

B. Segmentation Confidence

We introduce a segmentation confidence metric δ given $(\mathcal{P}, \mathcal{E}, \mathcal{R})$ for images as follows:

$$\delta = \frac{\mathcal{E}_{minor}}{\mathcal{E}_{major}} \cdot \mathbb{1}_{\{\alpha_{min} < \iint_{D_{\mathcal{R}}} dA\}} \cdot \mathbb{1}_{\{\alpha_{maj} < \iint_{D_{\mathcal{P}}} dA\}} \tag{3}$$

δ is the aspect ratio of the ellipse fit to $\mathcal{P}$ — the ratio of its minor to major axis — if both a minor/major protrusion are identified and the area of the minor and major protrusions are above hand-tuned area thresholds ($\alpha_{maj} = 1900\text{px}^2, \alpha_{min} = 300\text{px}^2$), and 0 otherwise. It is maximized at 1 when $\mathcal{E}$ is semicircular and decreases as $\mathcal{E}$ is less eccentric. The rationale for this is that we are more confident that a detected curved contour belongs to a suturing needle than a linear contour, due to the inherent curvature of a needle. Additionally, we have no confidence in a segmentation without a minor/major protrusion pair detected. In categorizing the failure modes of Section V-C based on the confidence of each of the segmentations, we find that a lack of confidence in segmentation tends to introduce a greater spread of failure modes and makes segmentation of the major protrusion more challenging (Type C error more prevalent, see Figure 5).

C. Segmentation Accuracy: Setup and Failure Modes

We manually inserted one suturing needle in varying positions and orientations into the tissue phantom and recorded an endoscopic image of the workspace 160 times. For each image, we ran the segmentation and path-planning algorithm and overlaid predicted grasps and pulls on each image, and a human judge noted the appropriateness of both the planned grasp and the planned pull direction and the failure mode if applicable. For each trial, the outcome(s) were specified

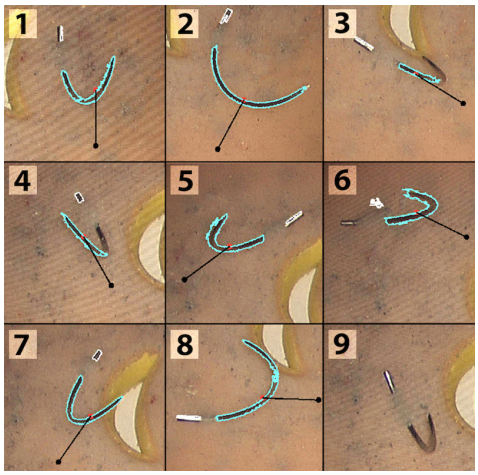

Fig. 5. Results for 9 out of the 160 trials of embedded needle detection and extraction planning with the setup in Section V-A. Blue and white outlines are the detected contours of the major protrusion and minor protrusion, respectively. The black line segment represents the predicted pull vector. The grasp and pull was successfully planned in 7/9 images, except for 6 where the incorrect minor protrusion was identified and in 9 where no major or minor protrusions were detected due to the effect of uneven shadows/highlights and specularities.

as follows, noting that images could be categorized under several failure modes (e.g. wrong minor protrusion, no major protrusion would be labeled as A, D):

- **A**: Contour misclassified as a minor protrusion (improper pull direction)
- **B**: No minor protrusion detected (no pull direction)
- **C**: Contour misclassified as a major protrusion (improper grasp)
- **D**: No big protrusion detected (no grasp)

D. Segmentation Accuracy: Results

We ran the segmentation algorithm across all 160 static images, of which a small subset of the results is visualized in Figure 5. The results suggest the reliability of our segmentation and path-planning algorithm.

Of 160 needle trials, in 17 trials (10.63%) no needle was detected, typically when the effects of specularity were especially noticeable or when the needle was positioned at the boundary of the phantom. Of the remaining 143 trials (89.38%), the major protrusion was correctly identified in 139 (97.20%) trials, the minor protrusion was correctly identified in 123 (86.01%) trials, and an appropriate extraction (correct major and minor protrusions identified) was planned in 122 (85.31%) trials.

Type A error occurred in 9/160 images (5.63%), Type B error occurred in 21/160 images (13.13%), Type C error occurred in 4/160 images (2.50%), and Type D error occurred in 17/160 images (10.63%).

We observe that the occurrence of inappropriate grasp was 4.25 times less likely than the occurrence of no grasp detected ($\frac{\%\text{Type D}}{\%\text{Type C}}$), and that the occurrence of a misdirected pull was 2.33 times less likely than the occurrence of no pull detected ($\frac{\%\text{Type B}}{\%\text{Type A}}$) across the test set of images, indicating that the proposed algorithm is relatively conservative in the context of planning needle extraction, which is especially relevant in a medical context.

The majority of failure modes were accounted for by Type B and Type D error, and these cases occurred most often when the inserted needle was placed on the perimeter of the workspace where the overhead workspace lighting was most dim and the needle contour was indistinguishable from its darkened containing region. Type A error was most frequently attributed to cases in which a dark defect on the phantom was recognized as a minor protrusion, and Type C error was attributed to cases in which the needle appeared close in size in both protruding ends, and its minor protrusion was taken to be the major protrusion.

E. Thresholding Baseline Comparison

We carry out segmentation on the same test set and experimental setup as in Section V-D but use a global threshold instead of Otsu's thresholding to compare segmentation accuracy. Using a global threshold level of 100, fine-tuned based on the workspace lighting, an appropriate grasp was planned in 84/160 images (52.50%), an appropriate pull given a planned grasp was planned in 43/84 images (51.19%), an an appropriate pull was planned overall in 43/160 images (26.88%). The occurrence of Types A, B, C, and D error were 60.0%, 13.13%, 38.75%, and 8.75%, respectively. The comparison of both segmentation accuracy and failure mode occurrence suggests the effectiveness of Otsu's method over a global threshold.

F. Robustness

In these experiments, we evaluate the system's robustness to various sources of noise. Needle segmentation accuracy is evaluated for images that are artificially lightened and darkened. The segmentation accuracy is also tested under the presence of textures and other objects in the images.

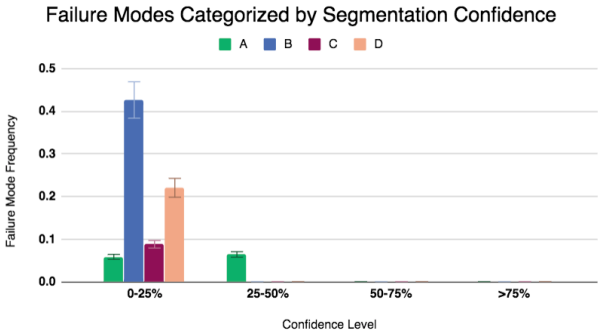

Fig. 6. Failure Modes vs. Segmentation Confidence. We show the occurrence of failure modes in images grouped by segmentation confidence. A and C correspond to wrong minor/major protrusions detected, respectively, and B and D correspond to no minor/major protrusions detected, respectively. We note the slight recurrence of Type A error as confidence increases. We hypothesize that this is due to the fact that needles segmented with higher confidence are likely to have a larger, more circular major protrusion and smaller minor protrusion, making the minor protrusion more difficult to identify.

1) Lighting Robustness: The above experiment was repeated on the same test set but with artifically brightening all 160 images using a gamma correction of 0.5. We also replicated the experiment with artificially darkening all images with a gamma correction of 1.5, and the results for grasp and pull detection were unchanged. This result suggests the robustness of the segmentation algorithm to lighting fluctuations and is likely attributed to Otsu's thresholding.

2) Robustness to Workspace Occlusions: The segmentation algorithm was also tested on a set of 41 images with an experimental setup as in Section V-A, but with strands of black and white suturing thread arbitrarily placed on the phantom to test the system's robustness to occlusions that are realistic in practice.

In this scenario, an appropriate grasp was planned in 30/41 images (73.1%), meaning the major protrusion was appropriately identified rather than a segment of thread or other contour, and of those cases, an appropriate pull was planned in 28/30 of cases (93.3%). An appropriate pull was planned overall in 28/41 (68.3%) of the images. The occurrence of Types A, B, C, and D error were 9.76%, 21.95%, 24.39%,and 2.44%, respectively, with Types A and C error being more prevalent in this experiment compared to a non-occluded environment due to suturing thread misidentified as a needle.

VI. PHYSICAL EXPERIMENTS

The full system is evaluated on a hernia mesh phantom with randomly embedded needles with a baseline calibration method and the proposed corrective calibration method.

A. Autonomous Extraction Accuracy: Setup/Failure Modes

We conducted 100 physical trials of extraction on arbitrarily inserted needles given an appropriate segmentation and grasp/pull plan, verified by a human judge, using the corrected calibration. A trial is considered successful if the robot is able to autonomously extract the needle from the phantom fully. The failure modes are categorized as follows:

- **E**: Missed grasp
- **F**: Gripper strikes and displaces needle during approach
- **G**: Partial extraction from phantom

These failure modes are not mutually exclusive, as a given trial could be categorized as E, F if the dVRK hits a needle during descent, causing it to miss a grasp.

B. Needle Extraction: Corrected Rigid Body Transform

With the corrected rigid body transformation, the dVRK successfully grasped the embedded needle in 75%, or 75/100 trials and fully extracted the needle in 70.67%, or 53/75 of these cases where a grasp was secured (and 53% or 53/100 cases overall). Given that a grasp was attained in 75% of trials, we hypothesize that the lower overall extraction rate can be attributed to the fixed offset distance a needle is pulled upon grasp. A longer, more aggressive pull would likely yield a much higher overall extraction rate, or a policy that factors the insertion depth, haptic feedback, or the elasticity of the phantom. However, a more forceful extraction could also yield damage to the tissue phantom or test the dVRK's joint limitations. Overall, the needle extraction results suggests the effectiveness of the fine-grained calibration in perceiving depth and accounting for scene geometry distortion. The occurrence of Types E, F, and G error were 24.0%, 10.0%, and 23.0%, respectively, noting that a given trial could be categorized under more than one failure mode.

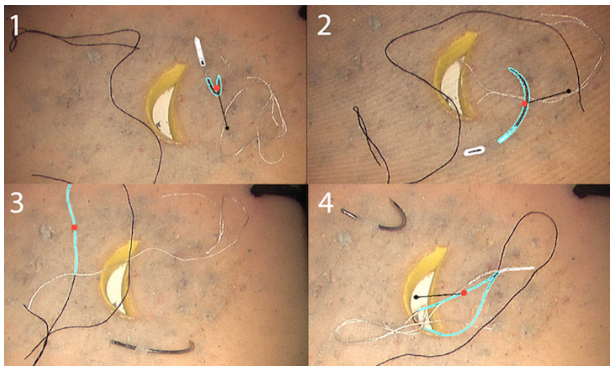

Fig. 7. Trials of *Robustness to Workspace Occlusions* segmentation experiment. Top (1, 2): Two appropriately segmented images isolating suturing needles from suturing thread with appropriately planned pull vectors. Bottom (3, 4): Suturing thread mistakenly segmented as needles.

VII. CONCLUSION AND FUTURE WORK

This paper proposes automating suturing needle extraction via a multi-step process of calibration, needle detection, path planning, and physical extraction. Experiments suggest that this task can be performed robustly and in relatively unstructured scenarios.

To improve segmentation, we will explore semantic segmentation on embedded needle amongst other surgical workspace objects. We also propose point-source lighting to identify needles based on specularity. We will explore the improvement of physical extraction, particularly through a non-linear heuristic for pull direction, greedy extraction policies for reattempting on failed grasps or pulls, and visual servoing. In addition, we will look at ways to use domain randomization to transfer needle extraction in simulation to physical execution with the dVRK.

ACKNOWLEDGMENTS

This research was performed at the AUTOLAB at UC Berkeley in affiliation with BAIR and the CITRIS "People and Robots" (CPAR) Initiative: http://robotics.citris-uc.org in affiliation with UC Berkeley's Center for Automation and Learning for Medical Robotics (Cal-MR). The authors were supported in part by donations from Intuitive Surgical, Siemens, Google, Toyota Research Institute, Autodesk, Honda, Intel, Comcast, Hewlett-Packard and by equipment grants from PhotoNeo, and NVidia. The da Vinci Research Kit is supported by the National Science Foundation, via the National Robotics Initiative (NRI), as part of the collaborative research project "Software Framework for Research in Semi-Autonomous Teleoperation" between The Johns Hopkins University (IIS 1637789), Worcester Polytechnic Institute (IIS 1637759), and the University of Washington (IIS 1637444).) We thank Daniel Seita and Jessica Ji for their extensive feedback on this manuscript.

REFERENCES

[1] G. Bradski and A. Kaehler, *Learning opencv: Computer vision with the opencv library*. " O'Reilly Media, Inc.", 2008.

[2] D.-L. Chow, R. C. Jackson, M. C. Çavuşoğlu, and W. Newman, "A novel vision guided knot-tying method for autonomous robotic surgery," in *2014 IEEE International Conference on Automation Science and Engineering (CASE)*, IEEE, 2014, pp. 504–508.

[3] D.-L. Chow and W. Newman, "Improved knot-tying methods for autonomous robot surgery," in *2013 IEEE International Conference on Automation Science and Engineering (CASE)*, IEEE, 2013, pp. 461–465.

[4] R. R. Cima, A. Kollengode, J. Garnatz, A. Storsveen, C. Weisbrod, and C. Deschamps, "Incidence and characteristics of potential and actual retained foreign object events in surgical patients," *Journal of the American College of Surgeons*, vol. 207, no. 1, pp. 80–87, 2008.

[5] C. D'Ettorre, G. Dwyer, X. Du, F. Chadebecq, F. Vasconcelos, E. De Momi, and D. Stoyanov, "Automated pick-up of suturing needles for robotic surgical assistance," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2018, pp. 1370–1377.

[6] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 21, no. 5, pp. 476–480, 1999.

[7] A. Garg, S. Sen, R. Kapadia, Y. Jen, S. McKinley, L. Miller, and K. Goldberg, "Tumor localization using automated palpation with gaussian process adaptive sampling," in *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, IEEE, 2016, pp. 194–200.

[8] B. Hannaford, J. Rosen, D. W. Friedman, H. King, P. Roan, L. Cheng, D. Glozman, J. Ma, S. N. Kosari, and L. White, "Raven-ii: An open platform for surgical robotics research," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 954–959, 2013.

[9] R. C. Jackson and M. C. Çavuşoğlu, "Modeling of needle-tissue interaction forces during surgical suturing," in *2012 IEEE International Conference on Robotics and Automation*, IEEE, 2012, pp. 4675–4680.

[10] R. C. Jackson and M. C. Çavuşoğlu, "Needle path planning for autonomous robotic surgical suturing," in *2013 IEEE International Conference on Robotics and Automation*, IEEE, 2013, pp. 1669–1675.

[11] R. Jayadevan, K. Stensland, A. Small, S. Hall, and M. Palese, "A protocol to recover needles lost during minimally invasive surgery," *JSLS : Journal of the Society of Laparoendoscopic Surgeons*, vol. 18, no. 4, 2014.

[12] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da vinci surgical system," in *IEEE Intl. Conf. on Robotics and Auto. (ICRA)*, Hong Kong, China, Jun. 1, 2014, pp. 6434–6439.

[13] B. Kehoe, G. Kahn, J. Mahler, J. Kim, A. Lee, A. Lee, K. Nakagawa, S. Patil, W. D. Boyd, P. Abbeel, *et al.*, "Autonomous multilateral debridement with the raven surgical robot," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2014, pp. 1432–1439.

[14] I. Kil, A. Jagannathan, R. B. Singapogu, and R. E. Groff, "Development of computer vision algorithm towards assessment of suturing skill," in *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, IEEE, 2017, pp. 29–32.

[15] K. Kozak, Q. Zhou, and J. Wang, "Static analysis of cable-driven manipulators with non-negligible cable mass," *IEEE Transactions on Robotics*, vol. 22, no. 3, pp. 425–433, 2006.

[16] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *The International Journal of Robotics Research*, vol. 37, no. 4-5, pp. 421–436, 2018.

[17] J. Liang, J. Mahler, M. Laskey, P. Li, and K. Goldberg, "Using dvrk teleoperation to facilitate deep learning of automation tasks for an industrial robot," in *Automation Science and Engineering (CASE), 2017 IEEE International Conference on*, IEEE, 2017.

[18] K. M. Lynch and F. C. Park, *Modern Robotics: Mechanics, Planning, and Control*. Cambridge University Press, 2017.

[19] J. Mahler, S. Krishnan, M. Laskey, S. Sen, A. Murali, B. Kehoe, S. Patil, J. Wang, M. Franklin, P. Abbeel, *et al.*, "Learning accurate kinematic control of cable-driven surgical robots using data cleaning and gaussian process regression," in *2014 IEEE International Conference on Automation Science and Engineering (CASE)*, IEEE, 2014, pp. 532–539.

[20] S. McKinley, A. Garg, S. Sen, D. V. Gealy, J. P. McKinley, Y. Jen, M. Guo, D. Boyd, and K. Goldberg, "An interchangeable surgical instrument system with application to supervised automation of multilateral tumor resection," in *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, IEEE, 2016, pp. 821–826.

[21] M. Morris and M. Shoham, "Applications and theoretical issues of cable-driven robots," May 2009.

[22] F. Mourgues, È. Coste-Manière, C. Team, *et al.*, "Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2002, pp. 25–34.

[23] A. Munawar and G. Fischer, "Towards a haptic feedback framework for multi-dof robotic laparoscopic surgery platforms," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2016, pp. 1113–1118.

[24] F. Nageotte, C. Doignon, M. de Mathelin, P. Zanne, and L. Soler, "Circular needle and needle-holder localization for computer-aided suturing in laparoscopic surgery," in *Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display*, International Society for Optics and Photonics, vol. 5744, 2005, pp. 87–99.

[25] T. Osa, N. Sugita, and M. Mitsuishi, "Online trajectory planning in dynamic environments for surgical task automation.," in *Robotics: Science and Systems*, 2014, pp. 1–9.

[26] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[27] K. Pachtrachai, M. Allan, V. Pawar, S. Hailes, and D. Stoyanov, "Hand-eye calibration for robotic assisted minimally invasive surgery without a calibration object," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2016, pp. 2485–2491.

[28] V. Pradeep, K. Konolige, and E. Berger, "Calibrating a multi-arm multi-sensor robot: A bundle adjustment approach," in *Experimental robotics*, Springer, 2014, pp. 211–225.

[29] J. Schulman, A. Gupta, S. Venkatesan, M. Tayson-Frederick, and P. Abbeel, "A case study of trajectory transfer through non-rigid registration for a simplified suturing scenario," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2013, pp. 4111–4117.

[30] D. Seita, S. Krishnan, R. Fox, S. McKinley, J. Canny, and K. Goldberg, "Fast and reliable autonomous surgical debridement with cable-driven robots using a two-phase calibration procedure," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2018, pp. 6651–6658.

[31] S. Sen, A. Garg, D. V. Gealy, S. McKinley, Y. Jen, and K. Goldberg, "Automating multi-throw multilateral surgical suturing with a mechanical needle guide and sequential convex optimization," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2016, pp. 4178–4185.

[32] O. Sorkine-Hornung and M. Rabinovich, "Least-squares rigid motion using svd," *Computing*, vol. 1, no. 1, 2017.

[33] S. Speidel, A. Kroehnert, S. Bodenstedt, H. Kenngott, B. Mueller-Stich, and R. Dillmann, "Image-based tracking of the suturing needle during laparoscopic interventions," in *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*, International Society for Optics and Photonics, vol. 9415, 2015, 94150B.

[34] S. Suzuki *et al.*, "Topological structural analysis of digitized binary images by border following," *Computer vision, graphics, and image processing*, vol. 30, no. 1, pp. 32–46, 1985.

[35] B. Thananjeyan, A. Tanwani, J. Ji, D. Fer, V. Patel, S. Krishnan, and K. Goldberg, "Optimizing robot-assisted surgery suture plans to avoid joint limits and singularities," in *2019 IEEE International Symposium on Medical Robotics*, IEEE, 2019.

[36] J. Van Den Berg, S. Miller, D. Duckworth, H. Hu, A. Wan, X.-Y. Fu, K. Goldberg, and P. Abbeel, "Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations," in *2010 IEEE International Conference on Robotics and Automation*, IEEE, 2010, pp. 2074–2081.