Suturing Support by Human Cooperative Robot Control Using Deep Learning

TAKUTO MIKADA[1], **TAKAHIRO KANNO**[2], **(Member, IEEE),**
TOSHIHIRO KAWASE[3,4], **(Member, IEEE), TETSURO MIYAZAKI**[5], **(Member, IEEE),**
AND KENJI KAWASHIMA[5], **(Member, IEEE)**

[1]Department of Biomechanics, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, Chiyoda 101-0062, Japan
[2]Riverfield Inc., Shinjuku 160-0017, Japan
[3]Department of Biomedical Information, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, Chiyoda 101-0062, Japan
[4]Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8503, Japan
[5]Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo 113-8656, Japan

Corresponding author: Takuto Mikada (tmdu.mikada@gmail.com)

This work was supported in part by The University of Tokyo.

ABSTRACT Considering the widespread use of surgical robots in recent years, computer-assisted surgery is becoming significantly popular. Because the automation of surgical tasks without human intervention remains complex owing to individual patient differences, a human cooperative control is proposed in this study. A system in which an operator manipulates one surgical instrument to insert a suture needle was developed, along with another surgical instrument that automatically pulls out the needle from the operated instrument. In the proposed method, YOLOv3 and a standard convolutional neural network (CNN) are used to estimate the penetration and pull state of the needle. An image-based state estimator classifies the state regardless of the stiffness of the object to which the suture needle is inserted. Furthermore, after the pull state is detected, despite a failure of the needle pulling, the position can be corrected and the automated surgical instrument can approach the needle again. Experiments on human cooperative control demonstrated the effectiveness of state estimation using the proposed method. In addition, the failure of grasping observed in a previous study caused by the needle angle error was reduced.

INDEX TERMS Computer-assisted surgery, deep learning, state estimation, human cooperative control.

I. INTRODUCTION

Endoscopic surgery, wherein only a few small incisions are required to insert surgical instruments into the body, has recently become popular. Postoperative esthetics and rapid recovery are of significant benefit to patients, however, the minimally invasive operation becomes more complex than the open surgery. For example, endoscopic surgery involves pivoting the surgical instrument around its insertion point in a narrow working space within the body. The longer instruments compared with those of open surgery also require training of the physician for precision. Surgical robots have been developed to solve the problems of endoscopic surgery [1], [2]. Surgical robots have eliminated complex operations around the insertion point and introduced hand tremor suppression and motion scale features to ensure accurate surgery. Intuitive Surgical's da Vinci is the most popular surgical support robot and is used for abdominal, pelvic, and thoracic surgeries [3], [4]. Most surgical robots are master—slave robots, and the surgeon continuously operates the position of the surgical instruments through endoscopic images. Therefore, the automation of complicated tasks during surgery using precise positioning control of the robot would lead to efficient surgery.

Murali *et al.* automated surgical cutting and debridement subtasks through "Learning By Observation" with a surgical support robot [5]. However, the color of the suture needle and the shape of the forceps were different from those of real surgical tools. Khoorjestan *et al.* proposed an automatic suture device that has greater strength than a hand-sutured specimen [6]. However, it is difficult to use this open-space device in minimally invasive surgery. Several systems have been developed to automate suture tasks in minimally invasive surgery [7], [8]. Shademan *et al.* verified the effectiveness of suturing task automation by experiments using pigs [9]. However, automation of surgical tasks without human intervention is still difficult due to individual differences in human bodies.

The associate editor coordinating the review of this manuscript and approving it for publication was Xiwang Dong.

Watanabe *et al.* proposed a single-master dual-slave system in which manually manipulated forceps insert a suture needle and autonomously operated forceps pull it out [10]. This method, which was previously proposed, shortens the time of the entire task and demonstrates that the semi-autonomous assistance to the operator is useful. However, there are a few remaining issues. This method uses the product of the external force estimated at the tip of the forceps and the angle of the roll axis of the holder robot to detect needle penetration. Penetration is not detected when the needle is inserted into an organ softer than the one used to define the threshold. For penetration into a rigid organ, the threshold is exceeded prior to penetration. Furthermore, because it cannot detect whether the suture needle was pulled, the automated slave conducts a hand-off motion without the needle when it fails to pull out the needle.

Image processing is effective for autonomous surgery and several automation methods are based on it. Stefan *et al.* verified the phase estimation of surgical videos from images using machine-learning technology and demonstrated its usefulness [11]. Compared to previous our methods that only use information regarding kinematics and force, image-processing-based methods can detect needle penetration despite the varying stiffness of organs. Furthermore, image processing can be used to detect failure to pull the needle and retry pull-out operation. For efficient and accurate needle state estimation, the following two-step image processing is effective: tracking the needle and cropping the image around the needle first, followed by estimating the state using the cropped image. Speidel *et al.* detected suture needles by color-based and geometry-based segmentation, and Chen *et al.* detected suture needles using a random forest model connected with a particle filter [12], [13]. The previous needle detection methods include sequential processing

and template matching, and are computationally extensive. Considering the burden on the patient, it is desirable to perform image processing quickly for use of the needle state in a feedback control system.

In this study, a two-step image processing method is proposed using two different deep neural networks. A laparoscopic image around the suture needle is cropped using a convolutional neural network (CNN), and then another CNN estimates the state of the needle. The proposed method can be applied to organs of varied stiffness because the penetration and pull states are determined only by the endoscope image, in contrast to the previous method [10], which was insufficient for managing stiffness variation. In this study, YOLOv3 was used to detect a suture needle, and a standard CNN was used to estimate the state from the image area of the suture needle. YOLOv3 is an object detection algorithm that can perform smoothing processing using deep learning [14].

The remaining of this study includes the following sections: Section II presents the devices used in this study. Section III presents the details of the state estimator using deep learning and human cooperative control. Section IV presents the experimental results regarding the learning of state estimators and the automation of suturing tasks using human cooperative control. Section V presents the improved functions compared with the previous study. Finally, the conclusions are presented in Section VI.

II. EXPERIMENTAL DEVICE
A. SLAVE MANIPULATOR
The slave manipulator used in this study is shown in Figure 1(a), which consists of a holder robot developed by Tadano *et al.* [15] and forceps developed by Haraguchi *et al.* [16]. The holder robot has a joint of yaw q_1,
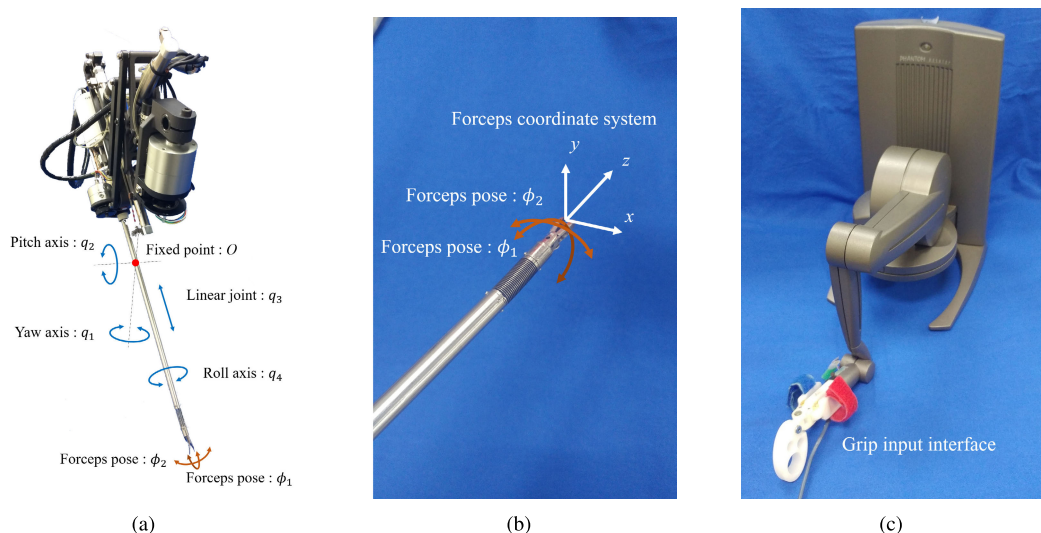
FIGURE 1. Slave Manipulator with six degrees of freedom and master device for manipulating slave. (a) Pneumatically driven surgical support robot with force feedback (b) Flexion motion and coordinate system of the forceps tip. (c) Master device with gripper opening/closing interface on 3D Systems' PHANTOM Desktop.
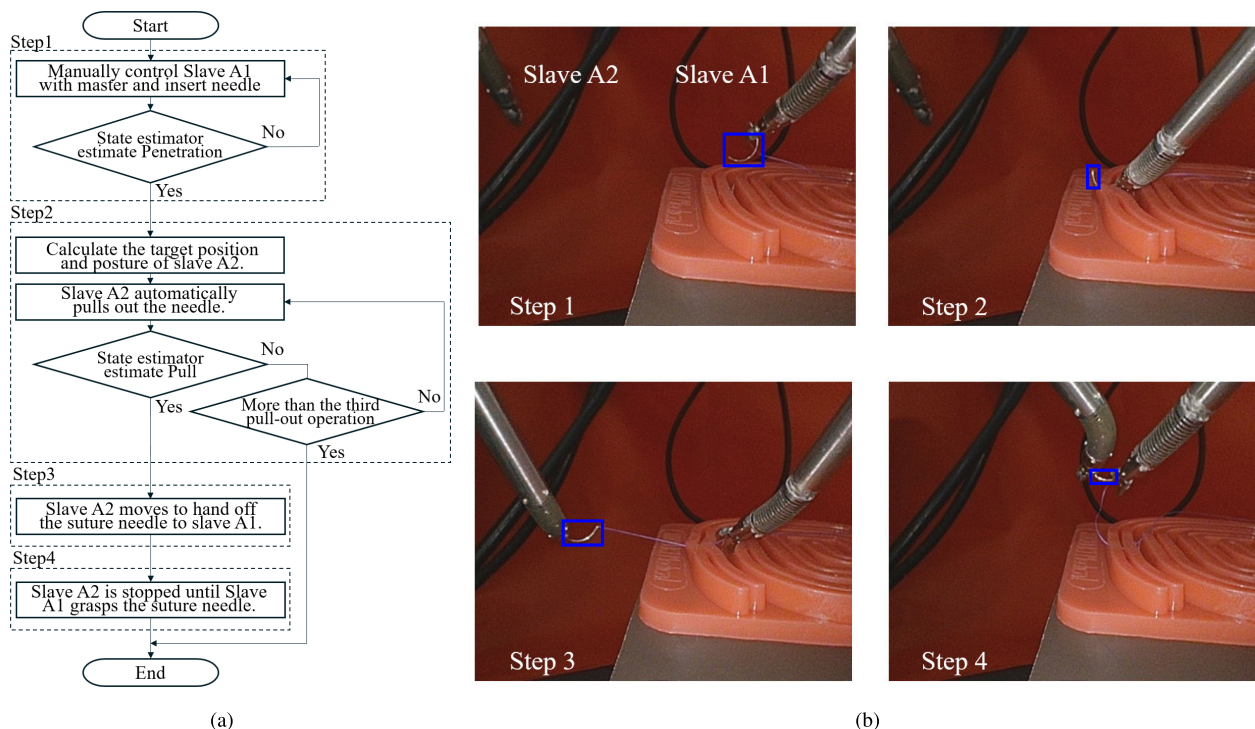
FIGURE 2. Human cooperative control based on the state of the suturing task. (a) Human cooperative control flowchart. (b) Endoscopic image corresponding to each step of the flowchart. The blue box in the endoscopic image indicates the detected area of the suture needle.

pitch q_2, linear motion joints q_3, and roll q_4, and operates as shown in Figure1(a). The holder robot was designed such that the intersection of the yaw and pitch axes of the holder and the axis of the forceps becomes a fixed point O, and the rotation around the fixed point. The tip of the forceps mounted on the holder robot is shown in Figure 1(b). The forceps provide a mechanism for bending the flexible joint by the push-pull of the two opposing nickel titanium wires. The bending postures ϕ_1 and ϕ_2 of the tip were controlled using four wires arranged at equal intervals. The opening and closing movements of the gripper attached to the tip grasps the organs and suture needles.

B. MASTER DEVICE

The 3D Systems' PHANTOM Desktop used as a master device in this study is shown in Figure 1(c). The master device is an input interface for the position and posture necessary for the control of the slave manipulator. An input interface attached to the distal end of the master device controls the opening and closing of the gripper at the tip of the forceps. The operator manipulates the input interface of the gripper through one's finger. The master device also has a triaxial force presentation function to feedback the force estimated by the forceps to the operator.

C. ENDOSCOPE

A stereoscopic endoscopy (Olympus ENDOEYE FLEX3D) was used in this study. The endoscope outputs 680 px × 540 px images at 30 Hz. This endoscopy can select two modes

of a three-dimensional image mode and a two-dimensional image mode. Because the scope of this study is to confirm the effectiveness of the proposed method, the two-dimensional image mode was selected from a simple implementation point of view to estimate the state of the suturing task.

III. METHOD

A. HUMAN COOPERATIVE CONTROL

A flowchart of the human cooperative control and the corresponding endoscopic image are shown in Figure 2. The system consists of an endoscope, Slave A1 operated by the master device, and Slave A2 for autonomous assistance. The state estimator with image processing is used as a switch to control the [Auto, Pause] mode of Slave A2. To assist the suturing task, three states were defined: non-penetration, penetration, and pull. When Slave A1 grasps the needle, whether the needle penetrates is represented by non-penetration or penetration; when Slave A2 grasps the needle, the state is set to pull. The suturing task is divided into four steps, and processing is performed based on the output of the state estimator. The processing at each step is presented as follows:
Step 1: The operator manually operates Slave A1 and inserts a needle into the training kit. For simplicity, Slave A1 already holds the needle from the beginning of the task. At this time, the state is defined as "non-penetration". While the state estimator continues to output non-penetration, Slave A2 is in pause mode. The position of Slave A2 is the initial position.

Step 2: The target position and posture of Slave A2 are calculated to pull out the suture needle, and the pulling

operation of the suture needle is controlled. Slave A1 and Slave A2 are mechanically fixed and their positional relationship is known; the orientations of their bases are the same. The target position of Slave A2 is first given in the Slave A1 coordinate considering the needle shape, and then transformed to the Slave A2 coordinate by using an offset vector. However, the needle may not be held in an ideal manner because it rotates freely around the grasping position of the gripper of Slave A1. To solve this problem, an algorithm in which Slave A2 corrects the target position was added, which considers the rotation of the needle when the pull out of the suture needle fails.
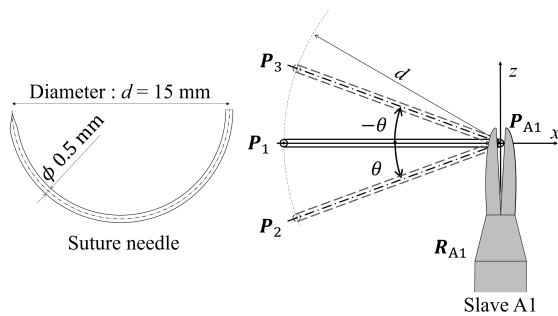
FIGURE 3. Geometry of the suture needle and anticipated rotation of the suture needle. P_1, P_2, and P_3 are needle tip positions corresponding to the angle of the suture needle rotation.

Figure 3 presents the geometry of the needle and assumed needle rotation. Slave A2 first moves to position P_1, assuming that the needle is ideally grasped. Then, if the first attempt fails, the position is corrected in the order of P_2, P_3. The target positions P_1, P_2, and P_3 of Slave A2 are calculated as follows on the coordinate system of the tip of Slave A1:

$$P_1 = P_{A1} - \begin{bmatrix} d \\ 0 \\ 0 \end{bmatrix}$$

$$P_2 = P_{A1} - \begin{bmatrix} d\cos(\theta) \\ 0 \\ d\sin(\theta). \end{bmatrix}$$

$$P_3 = P_{A1} - \begin{bmatrix} d\cos(-\theta) \\ 0 \\ d\sin(-\theta). \end{bmatrix} \quad (1)$$

where P_{A1} is the tip position of Slave A1, and d is the diameter of the semicircular arcuate needle shown in Figure 3. θ is the rotation angle of the suture needle assumed in this study. The target posture of Slave A2 is calculated as follows:

$$R_{A2} = R_z(30) \cdot R_{A1} \quad (2)$$

where $R_z(30)$ is a matrix rolling the z-axis shown in Figure 3 by 30 °, and R_{A1} is the posture of Slave A1. In this study, Slave A2 is rotated by 30 ° around the z-axis to make the posture suitable for grasping the suture needle. The modified target positions P_2 and P_3 depend on the needle rotation angle θ, and was set to 20 ° in this study. Under this setting,

if the rotation is within 20 °, then the needle will be within the graspable range of the gripper in less than three approaches.

When the target position and posture of Slave A2 are calculated, Slave A2 shifts to the automatic control and pulls out the needle. Slave A2 conducts the pulling operation with the gripper closed, and the state estimator confirms the success or failure of the needle pull-out. When the state estimator does not detect a "pull" after the pulling operation, Slave A2 performs the pulling operation by correcting the target position, as shown in Figure 3. While the state estimator continues to output the penetration, Slave A2 is in Auto mode.

Step 3: The pull of the needle is detected, and Slave A2 is controlled to hand off the suture needle to Slave A1. The position where slave A2 hands off the needle is the position that moved 20 mm in the *y* direction from the position where the state estimator detected penetration. While the state estimator estimates the pull, Slave A2 is set to the auto mode and moves to the needle transfer position.

Step 4: The operator manipulates Slave A1 to grasp the needle held by Slave A2. After Slave A2 moves to the needle transfer position, Slave A2 shifts to the pause mode until Slave A1 grasps the needle. When the operator manually operates Slave A1 and grasps the needle, Slave A2 releases the needle and moves to the initial position. After the operation is completed, return to Step 1.

The correspondence between the steps of automation, the mode of Slave A2, and the output of the state estimator is shown in Table 1.

TABLE 1. Correspondence between automation step, Slave A2 mode and output of state estimator.

	Pause	Auto
Non-penetration	Step 1	-
Penetration	-	Step 2
Pull	Step 4	Step 3

B. STATE ESTIMATION

State estimation is one of the most important factors to achieve task automation because it is used as the switch of automated control. In this study, a method to estimate a task state is proposed, as shown in Figure 4. The suturing task was performed using the suture training kit by KOTOBUKI Medical Inc., as shown in Figure 4. In this study, the following two deep neural networks were combined: one for needle detection and the other for task state estimation. YOLOv3 and a standard CNN were used to estimate the three states (non-penetration, penetration, and pull) required for human cooperative control. YOLOv3 is a method based on deep learning that detects a target object from an image and outputs a bounding box that operates at a high speed [14]. The proposed method uses YOLOv3 to obtain the bounding box of the needle from the 680 px × 540 px image and resizes it to 256 px × 256 px for input to the CNN. The resized image is annotated with [non-penetration, penetration, and pull] labels, and the CNN estimates the state of the three classes.
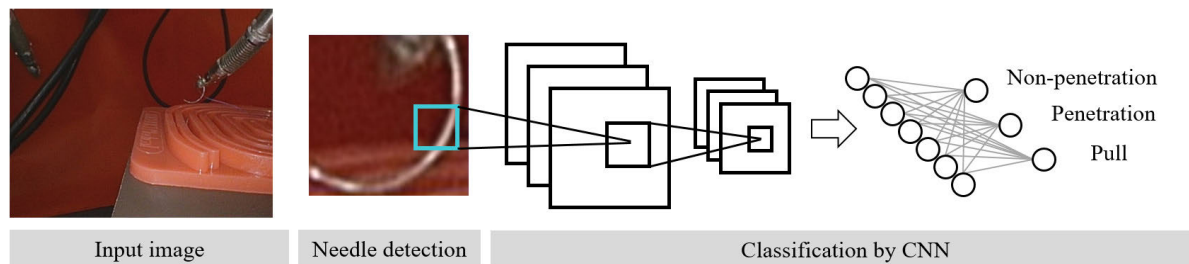
FIGURE 4. A method for estimating the state of the suturing task using YOLOv3 and standard CNN. YOLOv3 detects a suture needle from an endoscopic image of 680 px × 540 px. For input to the CNN, the detected region is cropped and resized to 256 px × 256 px. CNN classifies suture needle images into three states of the suturing task.

Image data was used when a suturing task was executed by the master—-slave device in Figure 1 for the learning of state estimation. The total of 15,505 images were created for learning, and the ratio of the number of each state was approximately [non-penetration : penetration : pull] = [1 : 1 : 1]. The learning data was manually annotated with 1,471 images that were resampled every 10 samples, excluding images where the needle was not visible. YOLOv3 learning is a single classification of suture needles. Object detection was conducted for all the images using learned YOLOv3. Images in which the suture needle could not be detected were qualitatively removed. The image in which the suture needle was finally detected was 13,386, and the learning of state estimation was performed using these images.

The structure of the constructed CNN is shown in Table 2. The CNN for state estimation receives the detected image of the suture needle resized to 256 px × 256 px and classifies it as a [non-penetration, penetration, or pull] state. The weights and biases of the CNN were initialized to the initial value of He *et al.* [17]. All convolution layers were activated by the Relu function and optimized using Adam [18], [19]. The extracted image vector is smoothed after convolution layers and classified into three states using the softmax function. In this study, a standard CNN was designed to focus on human collaboration and high processing speed, not Alexnet or Resnet, which scored significantly well in the image categories [20], [21]. In this study, the images of the suture needle detected by YOLOv3 were prepared for learning the CNN. The training data were 13,386 images in which the suture needle was detected by YOLOv3.

TABLE 2. Structure of CNN for state estimation.

Layer name	Kernel	Stride	Shape (H×W×C)
Convolution 1	3×3	2	128×128×32
Convolution 2	3×3	2	64×64×32
Convolution 3	3×3	2	32×32×64
Convolution 4	3×3	2	16×16×64
Convolution 5	3×3	2	8×8×128
Convolution 6	3×3	2	4×4×128
Convolution 7	3×3	2	2×2×256
Flutten	-	-	1024
Output	-	-	3

To avoid control errors due to misdetection of the state estimator, the state transition was programmed to occur when the penetration or pull was estimated 20 times continuously. In addition, the pull was processed to transition only from the penetration.

IV. RESULTS
A. HUMAN COOPERATIVE EXPERIMENTS
Verification experiments of human cooperative control in the suturing task were performed using the state estimator. The operator manipulates a single Slave A1 through a single master. A trial that achieved Step 4 is considered a success. On the other hand, the trial Slave A2 missed gripping the needle and did not reach Step 4, which was considered a failure. The initial condition of the task was that the system state is in Step 1 and Slave A1 is grasping the needle. The termination condition is that the system is in Step 4 and Slave A1 grasps the needle held by Slave A2. We conducted 20 trials. The proposed method estimates the state of the suturing task at a sampling rate of approximately 10 Hz. The experimental results are presented in Table 3. In the 20 trials, 12 suturing tasks succeeded in the first approach to the suture needle. The number of trials that succeeded after approaching the needle two or three times were 4 and 2, respectively. There were two trials in which the pull of the suture needle was not confirmed after three approaches.

TABLE 3. Results of automated suturing task.

Successful grasping operations	number of trials
1st time	12
2nd time	4
3rd time	2
Failure	2

B. STATE ESTIMATION
The accuracy of needle detection by YOLOv3 and the accuracy of state estimation by CNN are shown in Table 4. The accuracy of state estimation was verified only for images in which the suture needle was detected by YOLOv3. Table 4 presents the needle detection and state estimation accuracy

TABLE 4. Accuracy of needle detection and state estimation.

Range	Needle detection [%]	State estimation [%]
Step 1	81.4	99.1
Step 2	67.0	94.7
Step 3	83.8	97.5
Step 4	60.7	39.7
Total	72.4	93.2

of each step. Because Steps 3 and 4 were not performed in the 2 tasks in which the suturing task failed, the accuracy of Step 3 and Step 4 indicates the average of the 18 tasks.

V. DISCUSSION

A. HUMAN COOPERATIVE EXPERIMENTS

Results of the human cooperative experiments for the suturing task demonstrated that the proposed method detected the penetration and pull of the needle, and automated part of the suturing task. State estimation using an image is useful in that the state is estimated regardless of the stiffness of an object to be sutured because the penetration or pull of a needle is detected from the image feature around the needle. When the needle posture has an error of less than 20 ° from the ideal posture, the proposed method succeeds by the second or third grasping motions. Note, the indicated suturing task would have failed using the previously proposed method [10]. This is because the previous method assumed an ideal needle posture and did not have a system to detect pulling failure. An image-based state estimator has a significant advantage of the proposed method because it has the potential for successful suturing tasks despite the needle tip not being in an ideal position.

However, the experimental results recorded 2 failures in 20 trials. The reason for these failures was that the posture of the needle was rotated more than the expected angle of 20 °. In addition, there was a scene in which Slave A2 grasped the simulated organ by mistake; however, it is desirable for the needle to be accurately pulled out by one grasping operation considering the damage to a patient' s organ. These concerns may be resolved by measuring the tip position of the suture needle. Image-based estimation of needle posture is possible, as Kurose et al. demonstrated, using a 3D model of a suture needle and template matching [22]. Furthermore, the use of external force information acting on forceps will lead to safe, human cooperative robot control.

In this study, when automatically controlling Slave A2, a control signal was given for a linear movement to a target position. However, in real-time surgery, such a simple linear motion might cause an undesirable contact between the organ and Slave A2. Therefore, it is desirable to develop a method for generating trajectories that avoid organs for clinical application. The development of a trajectory generation system that uses YOLOv3 to detect organs and avoid undesirable contact during movement is considered for future studies.

B. STATE ESTIMATION

The false detection of the state estimation in Steps 1 and 2 leads to the failure of the suturing task. This is because human cooperative robot control is performed based on the results of state estimation in Steps 1 and 2. As Step 3 and Step 4 execute automated control regardless of the estimated state, there is no failure of the suturing task due to erroneous detection of the state. Table 4 presents that the state estimation accuracy of Steps 1 and 2 is high. The effect of false detection in the same step was removed by the processing, which shifted the state when the same state was estimated continuously for 20 frames.

Examples of state estimation errors are shown in Figure 5. Although Figure 5(a) is an example of the pull state, the state estimator indicated non-penetration in Step 4. The error is considered to have occurred because the learning data was not created for the two slaves located near the needle. However, the state estimation error that occurred in Step 4 does not affect human cooperative control because Slave A2 is in pause mode. An example of the state estimation error that occurred in Step 1 is shown in Figure 5(b). Figure 5(c) presents non-penetration, whereas the state estimator indicates penetration. Because the state estimation error that occurred in Step 1 can possibly transition to Step 2 before the actual penetration, an improvement in accuracy is required. Therefore, expanding the learning data would be effective in improving the estimation accuracy. By automatically annotating data using previous needle detection methods, it is possible to expand the learning data without manual annotation [12], [13], [22].
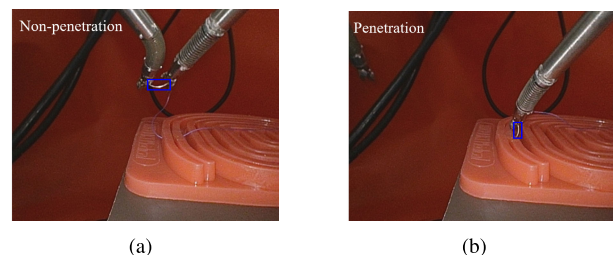
FIGURE 5. Example images that the state estimation error occurred. (a) Misdetection of non-penetration. (b) Misdetection of penetration.

In this study, the proposed state estimation method was validated using a suturing training kit. However, in real-time surgery, the background has complicated shapes and colors, such as organs, thus the proposed method is planned to be validated in vivo in the future. In addition, tissue mechanics is also an essential problem to achieve the suturing task in a real surgical situation [23]. Therefore, kinematic conditions for robot motion generation along with conditions for tissue mechanics will be added as the next step.

VI. CONCLUSION

In this study, a human cooperative suturing task control system was developed using a state estimator that combines YOLOv3 and a standard CNN. The proposed method

estimates the penetration state regardless of the stiffness of the sutured object by estimating the penetration state from only the image. The effectiveness of the proposed method was confirmed when there was an error in the position of the suture needle, in which the robot using the method in a previous study [10] failed to grasp the needle. The following will be considered for future studies: By adding an algorithm to the proposed method to estimate the three-dimensional posture of the suture needle, the needle will be pulled out at once regardless of the posture error. Because the accuracy of suture needle detection and state estimation for suturing tasks performed in an unlearned environment is insufficient, the expansion of learning data using existing suture needle detection methods is planned to improve the estimation accuracy. For clinical application, a trajectory generation method that detects organs and avoids undesirable contact during movement will be developed. In addition, tissue mechanics conditions will be added to the robot motion generation as the next step.

REFERENCES

[1] H. M. Le, T. N. Do, and S. J. Phee, "A survey on actuators-driven surgical robots," *Sens. Actuators A, Phys.*, vol. 247, pp. 323–354, Aug. 2016, doi: 10.1016/j.sna.2016.06.010.

[2] J. J. Rassweiler, R. Autorino, J. Klein, A. Mottrie, A. S. Goezen and J.-U. Stolzenburg, "Future of robotic surgery in urology," *BJU Int.*, vol. 120, no. 6, pp. 822–841, 2017, doi: 10.1111/bju.13851.

[3] A. Darzi and Y. Munz, "The impact of minimally invasive surgical techniques," *Annu. Rev. Med.*, vol. 55, no. 1, pp. 223–237, Feb. 2004, doi: 10.1146/annurev.med.55.091902.105248.

[4] G. Guthart and J. K. Salisbury, "The *textIntuitive*^{TM} telesurgery system: Overview and application," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 1, 2000, pp. 618–621, doi: 10.1109/ROBOT.2000.844121.

[5] A. Murali, S. Sen, B. Kehoe, A. Garg, S. McFarland, S. Patil, W. D. Boyd, S. Lim, P. Abbeel, and K. Goldberg, "Learning by observation for surgical subtasks: Multilateral cutting of 3D viscoelastic and 2D orthotropic tissue phantoms," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 1202–1209, doi: 10.1109/ICRA.2015.7139344.

[6] S. Mosafer Khoorjestan and G. Rouhi, "An automatic suturing machine for intestinal anastomosis: Advantages compared with hand-suturing technique," *Surgical Innov.*, vol. 26, no. 2, pp. 209–218, Apr. 2019, doi: 10.1177/1553350618808007.

[7] 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 *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 4111–4117, doi: 10.1109/IROS.2013.6696945.

[8] 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 *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 4178–4185, doi: 10.1109/ICRA.2016.7487611.

[9] A. Shademan, R. S. Decker, J. D. Opfermann, S. Leonard, A. Krieger, and P. C. W. Kim, "Supervised autonomous robotic soft tissue surgery," *Sci. Transl. Med.*, vol. 8, no. 337, pp. 337–364, May 2016, doi: 10.1126/scitranslmed.aad9398.

[10] K. Watanabe, T. Kanno, K. Ito, and K. Kawashima, "Single-master dual-slave surgical robot with automated relay of suture needle," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6335–6343, Apr. 2018, doi: 10.1109/TIE.2017.2786206.

[11] S. Petscharnig and K. Schöffmann, "Learning laparoscopic video shot classification for gynecological surgery," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8061–8079, Apr. 2018, doi: 10.1007/s11042-017-4699-5.

[12] S. Speidel, A. Kroehnert, S. Bodenstedt, H. Kenngott, B. Muller-Stich and R. Dillmann, "Image-based tracking of the suturing needle during Laparoscopic interventions," *Proc. SPIE Med. Imag.*, vol. 9415, Apr. 2015, Art. no. 94150B, doi: 10.1117/12.2081920.

[13] Y. Chen, M. M. Marinho, Y. Kurose, A. Nakazawa, K. Deie, K. Harada, and M. Mitsuishi, "Towards robust needle segmentation and tracking in pediatric endoscopic surgery," *Proc. SPIE Med. Imag.*, vol. 10576, pp. 698–705, Apr. 2018, doi: 10.1117/12.2292418.

[14] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*. [Online]. Available: http://arxiv.org/abs/1804.02767

[15] K. Tadano, K. Kawashima, K. Kojima, and N. A. Tanaka "Development of pneumatic surgical manipulator IBIS IV," *J. Robot. Mechatronics*, vol. 22, no. 2, pp. 179–188, 2010, doi: 10.20965/jrm.2010.p0179.

[16] D. Haraguchi, T. Kanno, K. Tadano, and K. Kawashima, "A pneumatically driven surgical manipulator with a flexible distal joint capable of force sensing," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 6, pp. 2950–2961, Dec. 2015, doi: 10.1109/TMECH.2015.2415838.

[17] K. He, X. Zhang, S. Ren, and J. Sun, and "Delving deep into rectifiers: Surpassing humanlevel performance on ImageNet Classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Apr. 2015, pp. 1026–1034, doi: 10.1109/ICCV.2015.123.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–4.

[20] K. Alex, S. Ilya, and H. E. Geoffrey, *ImageNet Classification With Deep Convolutional Neural Networks*. Red Hook, NY, USA: Curran Associates, 2012, pp. 1097–1105.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[22] Y. Kurose, Y. M. Baek, Y. Kamei, S. Tanaka, K. Harada, S. Sora, A. Morita, N. Sugita, and M. Mitsuishi, "Preliminary study of needle tracking in a microsurgical robotic system for automated operations," in *Proc. 13th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2013, pp. 627–630, doi: 10.1109/ICCAS.2013.6703943.

[23] S. M. Khoorjestan, G. Rouhi, and K. Toolabi, "An investigation of the effects of suture patterns on mechanical strength of intestinal anastomosis: An experimental study," *Biomed. Eng. / Biomedizinische Technik*, vol. 62, no. 4, pp. 429–437, Aug. 2017, doi: 10.1515/bmt-2016-0103.

TAKUTO MIKADA received the B.S. degree in mechanical engineering from the Department of Systems Design Engineering, Faculty of Engineering Science, Akita University, Akita, Japan, in 2018. He is currently pursuing the master's degree in health sciences and biomedical engineering track with Tokyo Medical and Dental University, Tokyo, Japan.

His research interests include medical robotics, machine learning, and control engineering.

TAKAHIRO KANNO (Member, IEEE) received the Ph.D. degree in engineering from the Department of Mechanical Engineering and Science, Kyoto University, in 2013.

In 2013, he worked as a Postdoctoral Researcher with the Precision and Intelligence Laboratory, Tokyo Institute of Technology. From 2013 to 2019, he worked as an Assistant Professor with Tokyo Medical and Dental University, where he has been an Associate Professor, since June 2019. Since 2020, he has been the Chief Technology Officer (CTO) at Riverfield Inc. His research interests include medical robotics, control engineering, and teleoperation.

TOSHIHIRO KAWASE (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the Tokyo Institute of Technology, Tokyo, Japan, in 2007, 2009, and 2012, respectively.

He was a Research Fellow with the Research Institute of National Rehabilitation Center for Persons with Disabilities, from 2012 to 2015, and worked as a Postdoctoral Fellow and a Specially Appointed Assistant Professor with the Tokyo Institute of Technology, from 2015 to 2017. He is currently an Assistant Professor with Tokyo Medical and Dental University and the Tokyo Institute of Technology. His research interests include medical robots, rehabilitation robotics, and biological signal processing.

KENJI KAWASHIMA (Member, IEEE) received the Ph.D. degree in engineering from the Department of Control Engineering, Tokyo Institute of Technology, in 1997.

From 1997 to 2000, he worked as a Research Assistant with the Tokyo Metropolitan College of Industrial Technology. He then worked as an Associate Professor with the Precision and Intelligence Laboratory, Tokyo Institute of Technology. From 2013 to 2020, he worked as a Professor with Tokyo Medical and Dental University. Since April 2020, he has been a Professor with The University of Tokyo. His research interests include medical robotics, control engineering, and fluid measurement and control.

• • •

TETSURO MIYAZAKI (Member, IEEE) received the Ph.D. degree in engineering from the Department of Mechanical Sciences and Engineering, Tokyo Institute of Technology, in 2014.

From 2014 to 2017, he worked as a Research Assistant (from 2014 to 2015) and an Assistant Professor (from 2015 to 2017) with Yokohama National University. From 2017 to 2020, he worked as an Assistant Professor with Tokyo Medical and Dental University. Since April 2020, he has been an Assistant Professor with the Graduate School of Information Science and Technology, The University of Tokyo. His research interests include mechanical engineering, control engineering, power assistive device, and medical welfare robotics.