VeniBot: Towards Autonomous Venipuncture with Semi-supervised Vein Segmentation from Ultrasound Images

Yu Chen¹, Yuxuan Wang¹, Bolin Lai², Zijie Chen¹, Xu Cao¹, Nanyang Ye³, Zhongyuan Ren⁴, Junbo Zhao⁵, Xiao-Yun Zhou⁶, Peng Qi∗

Abstract—In the modern medical care, venipuncture is an indispensable procedure for both diagnosis and treatment. In this paper, unlike existing solutions that fully or partially rely on professional assistance, we propose VeniBot — a compact robotic system solution integrating both novel hardware and software developments. For the hardware, we design a set of units to facilitate the supporting, positioning, puncturing and imaging functionalities. For the software, to move towards a full automation, we propose a novel deep learning framework — semi-ResNeXt-Unet for semi-supervised vein segmentation from ultrasound images. From which, the depth information of vein is calculated and used to enable automated navigation for the puncturing unit. VeniBot is validated on 40 volunteers, where ultrasound images can be collected successfully. For the vein segmentation validation, the proposed semi-ResNeXt-Unet improves the dice similarity coefficient (DSC) by 5.36%, decreases the centroid error by 1.38 pixels and decreases the failure rate by 5.60%, compared to fully-supervised ResNeXt-Unet.

I. INTRODUCTION

As a popular invasive procedure for venous access, the venipuncture has been widely used and acknowledged for clinical evaluation and treatment. This procedure opens up broad clinical practices including but not limited to biochemical analysis of the blood sample, infusion of fluid and cannulation for catheterization. For most adults, venipuncture is conventionally performed by medical practitioners under direct visual monitorization. Despite its simplicity, the success rate of venipuncture is potentially influenced by a variety of factors. In some adverse cases, the rate can be below 50% under improper settings [1]. On one hand, vein puncture is an experience-based procedure requiring extensive practices, which is hard for novices. On the other hand, in specific populations such as the elderly or infants, people with dark complexion, or patients in shock state, veins are not easily accessible. In addition, medical practitioners are at a high risk of needle injury and infection of blood-borne diseases like hepatitis and human immunodeficient virus (HIV) [2] or airborne diseases like corona virus disease [3]. Hence, novel solutions are necessitated to provide better and safe healthcare service.

Robotics has been a popular paradigm to automate many traditional manual tasks, for example, autonomous driving [4], [5], [6] and autonomous surgical operation [7], [8], [9]. Some robots have been developed and widely used for vein puncture. For example, Veebot, which was founded in 2010, proposed a venipuncture robot that combines a robotic arm and a compact puncturing unit [10]. It identifies a coarse vein with the help of infrared light and examines whether the vein is suitable or not for puncture with ultrasound. Venouspro was proposed by Vasculogic. It contains a six degree of freedom (DOF) positioning unit and a three DOF distal manipulator [11]. It is more compact and portable, compared to the robotic arm proposed by Veebot. Magic-Nurse, which is an automatic blood drawing robot, draws the blood and exchanges the puncture needle automatically. It detects the forearm vein with NIR images and estimates the depth of the needle with force sensors. The device has more than 20 DOF and is as big as a refrigerator. KUKA AG invented a robot arm holding a wireless ultrasound probe as an automatically robot-assisted tracking system to scan the vessel. It uses active contours to segment the vessel and it works on phantoms [12].

In this paper, a compact venipuncture device - VeniBot is designed and built for autonomous venipuncture. The solidworks (CAD) model of VeniBot is shown in Fig. 1. It is consisted of four units, including the supporting unit used to support the whole robot, the positioning unit mainly used to move VeniBot to the puncture target, the puncturing unit mainly used to puncture, and the imaging unit used to mount the near-infrared (NIR) camera and ultrasound device. The positioning unit contains three single axis robots driven with direct current (DC) motors and one motor that drives directly, and is navigated by a NIR camera mounted on the imaging unit, while the puncturing unit contains two single axis robots, and is navigated by an ultrasound device mounted on the imaging unit.

Except the hardware development, software for localization and navigation is also important and essential for robotic automation, typical examples are in medical robotics [13], [14] and general robotics [15]. In the developed VeniBot, there are two navigation problems: (1) the puncture area determination from the NIR images; (2) the vein segmentation from the ultrasound images. In this paper, we mainly focus on the second challenge while the first challenge is focused by another IROS submission. Deep learning has been
Fig. 1: An illustration of the solidworks (CAD) model of the proposed VeniBot, with the white color indicating the supporting unit, the blue color indicating the positioning unit, the red color indicating the puncturing unit, and the purple color indicating the imaging unit.

a popular method for image segmentation, however, it relies heavily on the training data with carefully manual labels. A more efficient way is to fully use the unlabeled data and train with semi-supervised learning. Popular semi-supervised methods include Π-model [16], temporal ensemble [16], mix match [17] and mean teacher [18]. In this paper, inspired by mean teacher [18] and Unet [19], we propose semi-ResNeXt-Unet for vein segmentation from limited labeled training data and large amount of unlabeled training data.

The two main novelties and contributions of this paper are:

- VeniBot is proposed and built with four units (supporting unit, positioning unit, puncturing unit and imaging unit) and six motors (four motors for the positioning unit and two motors for the puncturing unit).
- Semi-ResNeXt-Unet is proposed to segment the vein from ultrasound images with semi-supervised learning,
B. Automatic Vein Segmentation and Center Position Determination

To achieve autonomous venipuncture with the proposed VeniBot, it’s of great importance to develop an automatic method to localize centers of veins from ultrasound images, as it indicates the depth of the movement of motor 5 and 6. In this paper, centers of veins are calculated from their segmentation masks. However, vein segmentation from ultrasound images is very challenging, as the image quality is susceptible to the experience of operators and low quality usually exists even scanned by well-trained operators. Another challenge is the deformation caused by transducer. Since ultrasound wave cannot transit through the air, the transducer has to touch the patient’s forearm skin with suitable pressure. Large pressure makes the vein squashed, while small pressure remains some air between the skin and the transducer, resulting in all-black ultrasound images. Even for well-obtained ultrasound images, the vein bound can be obscure and its shape can differ dramatically, indicating the difficulty of labeling large amounts of ultrasound images manually. See Fig. 2 for some examples.

To address the time-consuming and labor-intensive annotations, we propose a semi-supervised segmentation method — semi-ResNeXt-Unet which first trains a fully-supervised model on limited labeled data and then improves its performance on more unlabeled data. Semi-ResNeXt-Unet contains 10 convolutional layers and 4 deconvolutional layers as demonstrated in Fig. 3 and Table I.

The semi-supervised method is based on the state-of-the-art mean teacher framework [18]. Assume $M$ labeled images and $N$ unlabeled images are available. Thus we have $S = \{x_i, y_i\}_{i=1}^M$ and $U = \{x_i\}_{i=1}^N$ for labeled data and unlabeled data, respectively. $x_i$ and $y_i$ denote input images and pixel-wise annotations. As illustrated in Fig. 4 the proposed semi-ResNeXt-Unet works in the following steps:

1. Train a fully-supervised network based on $S$. The

![Fig. 3: An illustration of the proposed ResNeXt-Unet network structure used for vein segmentation.](image)

![Fig. 4: An illustration of the semi-ResNeXt-Unet framework for semi-supervised learning.](image)

### Table I: The layer details of the proposed ResNet-Unet network structure used for vein segmentation.

| stage | output | ResNet-Unet |
|-------|--------|-------------|
| conv1 | $32 \times 32$ | $3 \times 3$, $64$, stride 2 |
| conv2 | $16 \times 16$ | $3 \times 3$ maxpool, stride 2 |
| conv3 | $8 \times 8$ | $1 \times 1$, $128$ |
| deconv4 | $16 \times 16$ | $3 \times 3$, $256$, stride 2 |
| conv5 | $16 \times 16$ | $3 \times 3$, $256$, stride 1 |
| conv6 | $16 \times 16$ | $3 \times 3$, $256$, stride 1 |
| deconv7 | $32 \times 32$ | $3 \times 3$, $128$, stride 2 |
| conv8 | $32 \times 32$ | $3 \times 3$, $128$, stride 1 |
| conv9 | $32 \times 32$ | $3 \times 3$, $128$, stride 1 |
| deconv10 | $64 \times 64$ | $3 \times 3$, $64$, stride 2 |
| conv11 | $64 \times 64$ | $3 \times 3$, $64$, stride 1 |
| conv12 | $64 \times 64$ | $3 \times 3$, $64$, stride 1 |
| conv13 | $64 \times 64$ | $3 \times 3$, $2$, stride 1 |

1Following [20], segmentation masks are used to localize centers of targets rather than detection bounding boxes because pixel-wise masks contain more information of shape and hence can further supply more accurate center location.
binary cross-entropy (BCE) loss is used at first. When reasonable prediction masks are obtained from the network. Focal loss [21] is used in place of BCE loss to address the extreme imbalance of foreground and background in ultrasound images. After the convergence, the parameters are used as initialization for the teacher and student models.

(2) Input labeled images $S$ into the student model. Focal loss is still used to train the network with other training protocols the same as that used for the fully-supervised training.

(3) For unlabeled images $U$, raw images are augmented by spatial transformation and input into the teacher model to get reliable pseudo probability maps. Then they are further disturbed by extra intensity augmentation and input into the student model. We use mean square error (MSE) as the consistency loss to penalize any difference between teacher and student model’s outputs. Then the total loss can be expressed as a weighted summation of supervised and semi-supervised losses.

$$L_{total} = \frac{\lambda_1 M L_{sup} + \lambda_2 N L_{semi sup}}{M + N},$$

(1) where the $\lambda_1$ and $\lambda_2$ are hyper-parameters used to balance the two losses.

(4) In the $k$ iteration, parameters of the teacher model ($\theta_T$) are frozen at first and the student model’s parameters ($\theta_S$) are updated by back-propagation of the focal loss and MSE loss. Then teacher model’s parameters are updated by moving average from the student model according to the following equation

$$\theta^k_T = \alpha \theta^{k-1}_T + (1 - \alpha) \theta^k_S,$$

(2) where $\alpha$ is a hyper-parameter to control the pace of update. After training, either the teacher model or the student model, which shows better performance on the validation set, is used for inference.

After segmenting the area of vein, to minimize the center detection error of the network and make sure the target vein is safe to access, a group of operation and evaluation is necessary. The image after binarization will go through opening operation to remove outliers and closing operation to link broken regions of the vein. We remain the biggest connected component and filter out other foreground areas. The target for venipuncture is the centroid of the segmented vein mask, which can be calculated by averaging the $x$ coordinates and $y$ coordinates of all pixels in the segmentation mask, respectively. It can be expressed as

$$(x_c, y_c) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n}\right),$$

(3) where $(x_c, y_c)$ and $(x_i, y_i)$ are coordinates of the centroid of the vein and $i$-th pixel in the mask. $n$ is the total number of pixels. $(x_c, y_c)$ further determines the final position of axis 5 and axis 6.

### III. VALIDATION

#### A. Experiments

The real VeniBot and the experimental setup are demonstrated in Fig. 5. The supporting unit was assembled with aluminium pieces. The skeleton of positioning unit and puncturing unit were built with machine work, while the base of ultrasound probe and NIR camera was processed with 3D metal printing. All of the translation pairs were driven by the combination of DC motors and single axis robots. Axis 4, a rotation pair, was directly driven by the DC motor. In total, four HIWIN KK40 single axis robots, one HIWIN KK50 single axis robot, three EC-max16 DC motors and three EC-max22 DC motors from Maxon Motor Inc. are used in VeniBot. Note that all ultrasound images used in this paper were collected by our machine, validating the robustness and stability of the proposed VeniBot.

Painting liquid or solid ultrasonic coupling agent on the skin is necessary to drive away the air and keep the ultrasound wave propagating in solid or liquid media. From the aspect of computer graphic method, the difference between the vein and the background and the characteristics of vein section are obvious enough to distinguish the vein from background. The intensity of vein is low because blood, in the state of liquid, does not reflect the ultrasound wave. Moreover, the vein section is an approximate circle or ellipse.
TABLE II: The range of the spatial and intensity augmentation added on the training of the semi-ResNeXt-Unet model.

| augmentation | size, range or probability |
|--------------|-----------------------------|
| resizing     | to $64 \times 64$            |
| horizontal flipping | 50% probability |
| rotation     | $-15^\circ \sim 15^\circ$   |
| shearing     | $-15^\circ \sim 15^\circ$   |
| changing aspect ratio | $0.01 \sim 0.01$ |
| adding       | $0 \sim 15$                 |
| multiplying  | $0.8 \sim 1.2$              |
| dropout      | $0 \sim 0.05$               |
| contrast normalization | $0.8 \sim 1.2$ |

Most of the veins are at the size of $1\sim2$mm in diameter. These characteristics above can be manually measured with intensity, circularity, size, convexity, concavity and aspect ratio, indicating the availability of manual labeling of ground truth (GT). We use VeniBot to collect ultrasound images from 40 volunteers (27 males and 13 females) at the age of 18~70 to validate our methods. Each of them provided 30 ultrasound images of forearm vein, i.e. 1200 images in total, with the same size of $70 \times 70$. By adjusting the parameters of multiple filters to segment the vein area one by one, we segmented 300 images as labeled training data and used images from the rest 30 volunteers as unlabeled data.

Two kinds of augmentations are implemented in fully- and semi-supervised training: spatial augmentation and intensity augmentation. Spatial augmentation includes resizing, flipping, rotation, shearing and changing the aspect ratio. Intensity augmentation includes randomly adding or multiplying a certain value, contrast normalization and dropout. The range of the spatial and intensity augmentation is listed in Table II.

The ResNeXt-Unet model is trained with 300 labeled images first. The model is first trained using BCE loss and with the learning rate of $1 \times 10^{-3}$. Then it is further trained using focal loss with a learning rate of $5 \times 10^{-4}$ until convergence. The batch size is set as 8 and both intensity and spatial augmentations are used. In the semi-supervised training, labeled images and unlabeled images are randomly selected in each batch. The same spatial augmentations are used on student and teacher model’s inputs, while intensity augmentations are only implemented on student model’s inputs.

B. Results

To validate the kinematic and assembly accuracy of VeniBot, a $10 \times 10$ grid paper with a grid size of $2$mm $\times 2$mm is used. We assembled a pencil at the end of axis 6 and controlled Venibot to move $2$mm along axis x or y once at a time. After each movement, the pencil moved along axis 6 to draw a short line in the corresponding grid. As shown in Fig. 7, VeniBot successfully controlled a pencil drawing short lines inside each grid, which shows the crucial and fundamental ability of VeniBot to achieve safe and viable venipuncture.

In order to show the advantage of the proposed semi-ResNeXt-Unet, three strong semi-supervised baselines are used: (1) pseudo label training [22] which generates pseudo labels for unlabeled data by training a model on labeled data and using it to predict labels for unlabeled data; (2) teacher-student learning, where a teacher network is trained on labeled data and a student network is trained on both labeled and unlabeled data; (3) temporal ensemble, where multiple models are trained on labeled data and their predictions are averaged to produce the final output.
TABLE III: The mean and std DSC of training the fully-supervised model, pseudo label training model, Π-model, temporal ensemble model and the proposed semi-ResNeXt-Unet model on the five cross validations.

| experiment         | model 1       | model 2       | model 3       | model 4       | model 5       | average       |
|--------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| fully-supervised   | 0.691±0.3565  | 0.650±0.3760  | 0.733±0.2949  | 0.743±0.2817  | 0.708±0.3135  | 0.697±0.3245  |
| pseudo label [22]  | 0.674±0.3697  | 0.629±0.3930  | 0.712±0.3197  | 0.715±0.3115  | 0.674±0.3446  | 0.672±0.3477  |
| Π-model [16]       | 0.720±0.3150  | 0.663±0.3636  | 0.768±0.2671  | 0.781±0.2547  | 0.737±0.2831  | 0.732±0.2972  |
| temporal ensemble  | 0.669±0.3576  | 0.613±0.4014  | 0.705±0.3235  | 0.728±0.3047  | 0.684±0.3630  | 0.661±0.3501  |
| semi-supervised    |               |               |               |               |               |               |
| student            | 0.729±0.3092  | 0.679±0.3441  | 0.749±0.2517  | 0.831±0.2149  | 0.758±0.2584  | 0.751±0.2557  |
| teacher            | 0.704±0.3265  | 0.679±0.3441  | 0.759±0.2682  | 0.771±0.2625  | 0.744±0.2706  | 0.730±0.2947  |

TABLE IV: The mean and std centroid error of training the fully-supervised model, pseudo label training model, Π-model, temporal ensemble model and the proposed semi-ResNeXt-Unet model on the five cross validations.

| experiment         | model 1       | model 2       | model 3       | model 4       | model 5       | average       |
|--------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| fully-supervised   | 14.19±8.48   | 12.40±14.80   | 8.39±12.50    | 5.57±8.26     | 7.50±9.96     | 9.61±12.80   |
| pseudo label [22]  | 14.81±16.77  | 15.52±16.56   | 9.21±12.87    | 6.60±9.06     | 8.50±11.54    | 10.94±13.36  |
| Π-model [16]       | 10.26±13.40  | 14.40±14.01   | 7.27±10.56    | 5.21±7.79     | 7.47±9.53     | 8.32±11.06   |
| temporal ensemble  | 10.59±15.30  | 14.93±16.16   | 7.90±10.08    | 5.85±7.96     | 12.75±12.51   | 10.48±12.40  |
| semi-supervised    |               |               |               |               |               |               |
| student            | 8.78±14.09   | 13.20±12.77   | 7.07±10.58    | 4.30±7.65     | 7.80±10.34    | 8.23±11.09   |
| teacher            | 8.08±12.28   | 13.05±12.12   | 7.80±10.50    | 4.86±7.66     | 9.88±10.78    | 8.73±10.67   |

TABLE V: The mean failure rate of training the fully-supervised model, pseudo label training model, Π-model, temporal ensemble model and the proposed semi-ResNeXt-Unet model on the five cross validations.

| experiment         | model 1       | model 2       | model 3       | model 4       | model 5       | average       |
|--------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| fully-supervised   | 21.67%        | 6.67%         | 5.00%         | 3.00%         | 1.67%         | 7.60%         |
| pseudo label [22]  | 25.00%        | 21.67%        | 8.33%         | 3.33%         | 8.33%         | 13.33%        |
| Π-model [16]       | 10.00%        | 6.67%         | 6.67%         | 6.67%         | 0.00%         | 4.67%         |
| temporal ensemble  | 1.67%         | 11.67%        | 5.00%         | 1.67%         | 1.67%         | 4.34%         |
| semi-supervised    |               |               |               |               |               |               |
| student            | 6.67%         | 3.33%         | 0.00%         | 0.00%         | 0.00%         | 2.00%         |
| teacher            | 5.00%         | 0.00%         | 0.00%         | 0.00%         | 0.00%         | 1.00%         |

GT for the unlabeled data with the converged model in the fully-supervised step, (2) Π-model [16] which only uses one model with the consistency loss, temporal ensemble [16] which updates the pseudo GT after every training epoch using moving average method. The DSC of our method and baselines with 5-fold cross validation is illustrated in Tab.III. It can be observed that temporal ensemble and pseudo label training achieve the lowest performance, which is even lower than the fully-supervised model. One possible reason is that the challenging nature of vein segmentation causes lots of error in the pseudo GT from the fully-supervised model, which misleads the model in semi-supervised training. In contrast, based on the consistency loss, Π-model and our method boost the performance by a significant margin. The proposed semi-ResNeXt-Unet still shows its superiority and outperforms the Π – model by 1.91%. Compared with the fully-supervised counter-part, the mean DSC is finally improved by 5.36%.

We further show the centroid MSE error which measures the distance between the centroid of prediction and GT. If there is no vein predicted, we count this situation as failure cases and see its prediction at the top left corner. The centroid MSE error and the failure rate are shown in Tab.IV and V respectively. For the centroid error, the proposed semi-ResNeXt-Unet outperforms the fully-supervised method, pseudo label training, and temporal ensemble with notable margins, while outperforms Π-model with a small margin. For the failure rate, the proposed semi-ResNeXt-Unet achieves the lowest value for four of the five cross validations and the overall average failure rate, indicating the robustness of the proposed semi-ResNeXt-Unet to hard cases, i.e., images with small veins.

Finally, seven vein segmentation results, segmented by the four methods, were randomly selected and shown in Fig. 8. The results of baselines lose some significant vein areas or even totally fail to generate reasonable masks. The proposed semi-ResNeXt-Unet precisely catches the vein and outputs more reliable masks, which is an important perceptual evaluation of the superiority of our method.

IV. CONCLUSION

In the paper, we proposed a compact venipuncture robot – VeniBot. For its automation, we proposed ResNeXt-Unet to automatically segment the vein from ultrasound images and calculate the depth from skin to the vein centroid. Given the expense of collecting a large labeled dataset, we also proposed a semi-supervised segmentation network – semi-ResNeXt-Unet to obtain reasonable vein masks based on a limited set of labeled training data as well as more unlabeled images. It prominently improves the performance by 5.36% DSC compared with the fully-supervised counter-part and surpasses other semi-supervised baselines by at least 1.91% DSC. It’s a promising step towards a wide and automatic application in the future.
REFERENCES

[1] D. Mbamalu and A. Banerjee, “Methods of obtaining peripheral venous access in difficult situations,” *Postgraduate medical journal*, vol. 75, no. 886, pp. 459–462, 1999.

[2] D. A. Mengistu, S. T. Tolera, and Y. M. Demmu, “Worldwide prevalence of occupational exposure to needle stick injury among healthcare workers: A systematic review and meta-analysis,” *Canadian Journal of Infectious Diseases and Medical Microbiology*, vol. 2021, 2021.

[3] G. Scoppettuolo, D. G. Biasucci, and M. Pittiriti, “Vascular access in covid-19 patients: smart decisions for maximal safety,” 2020.

[4] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, et al., “Towards fully autonomous driving: Systems and algorithms,” in *2011 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2011, pp. 163–168.

[5] S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu, “A survey of deep learning techniques for autonomous driving,” *Journal of Field Robotics*, vol. 37, no. 3, pp. 362–386, 2020.

[6] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogumani, and P. Pérez, “Deep reinforcement learning for autonomous driving: A survey,” *arXiv preprint arXiv:2002.00444*, 2020.

[7] T. S. Perry, “Profile: Veebot drawing blood faster and more safely than a human can,” 2013.

[8] A. I. Chen, M. L. Balter, T. J. Maguire, and M. L. Yarmush, “Deep learning robotic guidance for autonomous vascular access,” *Nature Machine Intelligence*, vol. 2, no. 2, pp. 104–115, 2020.

[9] M. Unger, J. Berger, B. Gerold, and A. Melzer, “Robot-assisted ultrasound-guided tracking of anatomical structures for the application of focused ultrasound,” *Current Directions in Biomedical Engineering*, vol. 6, no. 3, pp. 123–126, 2020.

[10] Q. Ma, E. Kobayashi, H. Suenaga, K. Hara, J. Wang, K. Nakagawa, I. Sakuma, and K. Masamune, “Autonomous surgical robot with camera-based markerless navigation for oral and maxillofacial surgery,” *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 2, pp. 1084–1094, 2020.

[11] G. Scoppettuolo, D. G. Biasucci, and M. Pittiriti, “Vascular access in covid-19 patients: smart decisions for maximal safety,” 2020.

[12] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogumani, and P. Pérez, “Deep reinforcement learning for autonomous driving: A survey,” *arXiv preprint arXiv:2002.00444*, 2020.

[13] M. Unger, J. Berger, B. Gerold, and A. Melzer, “Robot-assisted ultrasound-guided tracking of anatomical structures for the application of focused ultrasound,” *Current Directions in Biomedical Engineering*, vol. 6, no. 3, pp. 123–126, 2020.

[14] Q. Zhang, X.-G. Han, Y.-F. Xu, M.-X. Fan, J.-W. Zhao, Y.-J. Liu, D. He, and W. Tian, “Robotic navigation during spine surgery,” *Expert review of medical devices*, vol. 17, no. 1, pp. 27–32, 2020.

[15] X.-Y. Zhou, J. Lin, C. Riga, G.-Z. Yang, and S.-L. Lee, “Real-time 3-d shape instantiation from single fluoroscopy projection for fenestrated stent graft deployment,” *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1314–1321, 2018.

[16] C. G. Pachón-Suescún, C. J. Enciso-Aragón, and R. Jiménez-Moreno, “Robotic navigation algorithm with machine vision,” *International Journal of Electrical & Computer Engineering* (2088-8708), vol. 10, no. 2, 2020.

[17] S. Laine and T. Aila, “Temporal ensembling for semi-supervised learning,” *arXiv preprint arXiv:1610.02242*, 2016.

[18] D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, and C. Raffel, “Mixmatch: A holistic approach to semi-supervised learning,” *arXiv preprint arXiv:1905.02249*, 2019.

[19] A. Tarvainen and H. Valpola, “Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results,” in *NeurIPS*, 2017, pp. 1195–1204.

[20] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[21] C. G. Pachón-Suescún, C. J. Enciso-Aragón, and R. Jiménez-Moreno, “Robotic navigation algorithm with machine vision,” *International Journal of Electrical & Computer Engineering* (2088-8708), vol. 10, no. 2, 2020.

[22] S. Laine and T. Aila, “Temporal ensembling for semi-supervised learning,” *arXiv preprint arXiv:1610.02242*, 2016.

[23] D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, and C. Raffel, “Mixmatch: A holistic approach to semi-supervised learning,” *arXiv preprint arXiv:1905.02249*, 2019.

[24] A. Tarvainen and H. Valpola, “Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results,” in *NeurIPS*, 2017, pp. 1195–1204.

[25] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[26] C. G. Pachón-Suescún, C. J. Enciso-Aragón, and R. Jiménez-Moreno, “Robotic navigation algorithm with machine vision,” *International Journal of Electrical & Computer Engineering* (2088-8708), vol. 10, no. 2, 2020.

[27] S. Laine and T. Aila, “Temporal ensembling for semi-supervised learning,” *arXiv preprint arXiv:1610.02242*, 2016.